31st SAINT PETERSBURG

INTERNATIONAL CONFERENCE

ON INTEGRATED

NAVIGATION SYSTEMS

27 – 29 May 2024

Saint Petersburg, Russia

CO-SPONSORED BY:

- INTERNATIONAL PUBLIC ASSOCIATION – ACADEMY OF NAVIGATION AND MOTION CONTROL (ANMC)
- NATIONAL RESEARCH UNIVERSITY ITMO
- CHINESE SOCIETY OF INERTIAL TECHNOLOGY (CSIT)
- JOURNAL GYROSCOPY AND NAVIGATION

This is a collection of plenary and poster papers of the 31st Saint Petersburg International Conference on Integrated Navigation Systems (27–29 May 2024).

Editor-in-Chief

Academician of RAS
Vladimir G. Peshekhonov

SESSION I – INTEGRATED NAVIGATION AND CONTROL SYSTEMS

PLENARY PAPER

POSTER PAPERS

SESSION II – INERTIAL SYSTEMS AND SENSORS

PLENARY PAPERS

POSTER PAPERS

SESSION III – RELEVANT ISSUES OF THEORY

Development of the RS ISS Local Positioning Systems

S.V. Bronnikov

S.P. Korolev Rocket and Space
Corporation "Energia"
Korolev, Moscow Region,
Russian Federation
sergey.bronnikov@rsce.ru

D.Yu. Karavaev

S.P. Korolev Rocket and Space
Corporation "Energia"
Korolev, Moscow Region,
Russian Federation
dmitry.karavaev@rsce.ru

A.S. Rozhkov

S.P. Korolev Rocket and Space
Corporation "Energia"
Korolev, Moscow Region,
Russian Federation
alexander.rozhkov@rsce.ru

D.N. Rulev

S.P. Korolev Rocket and Space
Corporation "Energia"
Korolev, Moscow Region,
Russian Federation
dmitry.rulev@rsce.ru

O.S. Rurin

LLC «ruCap»
Moscow, Russian Federation
mail@rucap.ru

A.K. Kalifatidi

LLC «ruCap»
Moscow, Russian Federation
kalifatidi@rucap.ru

Abstract— **The issues of increasing the efficiency of the space station crew activity by using local positioning systems to support such activities as photographing the Earth's surface and searching for objects stored on board are considered. The use of such tools makes it possible to increase the productivity of the crew, as well as to ensure the collection of data on the location of crew members, on the use of the useful volume of the station, which allows optimizing the design of compartments and the layout of on-board equipment. As part of the implementation of the Vizir space experiment on the Russian segment (RS) of the International Space Station (ISS), a study was carried out of the principles, methods of construction and technical characteristics of local positioning systems, on the basis of which devices to support the activities of spacecraft and station crews can be developed. Systems in three areas of target application have been studied: a goniometer system, an automated storage of objects, and a system for monitoring the location of the astronauts. The experimental use of the considered local positioning systems on the RS ISS made it possible to create new technologies for performing flight operations aimed to increase the efficiency of the crew and the efficiency of solving the target tasks of the station as a whole.**

Keywords—local positioning system, space station, crew, goniometric system, smart shelf

I.   INTRODUCTION

Astronauts' working time is one of the most expensive resource on the space station. Therefore, the issues of increasing the efficiency of the station crew are relevant.

The issues of increasing the efficiency of crew activities by automating them using technical means built on the basis of local positioning systems to support such activities as photographing the Earth's surface and searching for objects stored on board are considered.

The use of such tools makes it possible to increase crew productivity, as well as to ensure the collection of current and statistical information about the location of crew members, about the crew's use of the station's useful volume, which allows optimizing the design of compartments and the layout of on-board equipment.

II.   AUTOMATION OF THE PROCESS OF OBTAINING THE EARTH'S SURFACE IMAGES USING FREELY MOVABLE CAMERAS

During the flight, astronauts of the International Space Station (ISS) take a large amount of photographs - hundreds of thousands of images of the Earth's surface during a six-month flight. Images of the Earth's surface, obtained by the crews of orbiting manned spacecraft (SC) through windows with video and photo cameras, are used for scientific and practical purposes [1, 2] in many sectors of the economy.

However, until recently, the effectiveness of using these images has been low. This is due to the fact that in order to use an image of the Earth's surface, it is necessary to georeference it. The georeferencing (binding) is performed on Earth by special personnel. This is a complex and time-consuming process. Moreover, for many homogeneous images (forest, sea), it is impossible to georeference due to the lack of any unique landmarks.

There are two ways to survey the earth's surface from a manned spacecraft. The first one is to install the camera in a gimbal, which has sensors for the camera's angles of rotation relative to the spacecraft. The second is shooting with a freely movable camera, i.e. a camera that does not have a

rigid connection with the spacecraft body. Data on the orientation of the camera relative to the spacecraft in this case are not available.

When shooting using the first method, you can calculate the vector of the camera's optical axis and determine the coordinates of its intersection with the earth's surface - the longitude and latitude of the center of the image. In this case, it is relatively easy to link the image, since the intended identification area by the image specialist is a circle with a radius of about 15 km on the orthophotomap of the earth's surface around the calculated point of intersection of the camera's optical axis with the earth's surface.

Astronauts prefer the second shooting method. Weightlessness creates for a non-fixed camera the conditions of an ideal three-degree suspension, to which vibrations from the spacecraft structure are not transmitted, which allows the astronaut to point the camera at objects with high accuracy, track their passage, quickly move the camera to a new target, taking into account the current situation, quickly photograph phenomena and objects, appearing in the field of view.

However, when shooting with freely moving cameras, it is much more difficult to georeference the resulting images than with the first method. The position of the camera's optical axis relative to the SC associated coordinate system (ACS) at the time of shooting is unknown. Based on the time of photographing, we can calculate only the coordinates of the sub-satellite point. Therefore, the expected zone on the orthomosaic for shooting subject identification by an image specialist is a circle with a radius of about 800 km around the sub-satellite point.

Therefore, it is important to improve the means and methods for capturing the Earth's surface from a spacecraft by cosmonauts using freely movable cameras [3].

A. Creation of a goniometric system for freely movable cameras

To enhance imaging techniques for a freely moving camera, it was developed 4 options an goniometric system (GS) which, in the absence of a mechanical connection between the camera and the SC body, ensures the determination of the angles of the optical axis of the camera relative to the SC ACS.

Option 1.

The GS based on two rigidly attached photo cameras (PC) and external markers located on the SC's internal surface in the area of imaging equipment (GS-2PC).

The main disadvantage of this option is that the georeferencing of images of the earth's surface cannot be performed automatically and requires operator participation due to the variability of the interior of the manned spacecraft.

Option 2.

The GS is based on an angular velocity sensor TIUS-500 manufactured by Optolink, Russia (GS-TIUS) [4].

The error in calculating the coordinates of the image center based on GS-TIUS data is 20-30 m. It is acceptable. But from an ergonomic point of view, GS-TIUS turned out to be inconvenient, since it has a low limit to the linear measurement range (the equipment is sensitive to sharp turns and collisions), so 20% of the data was lost due to an

unintentional collision of the camera with the spacecraft structure during the shooting process.

Option 3.

The GS is based on an angular velocity meter MTi produced by Xsens Technologies B.V., the Netherlands (GS-MTi).

A significant disadvantage of the MTi device is that it has a large linear component of the systematic error in the values of rotation angles (up to 0.6 deg/s), which is unacceptable for a long shooting process.

Option 4.

The GS is based on ultrasonic sensors (GS-U), developed by the RUCEP company, Russia, according to the Technical Specifications of Rocket-Space Corporation Energia, Russia. The GS-U consists of the following components:

- platform attached to the camera, with a Wi-Fi emitter and ultrasonic emitters powered by the camera;

- platform (frame) placed around the porthole, with ultrasonic receivers,

- control unit (CU);

- Wi-Fi communication unit;

- set of cables located behind the internal panels of the SC;

The control unit and ultrasound receivers are powered from the crew computer. At the time of shooting, the distances between the emitters and receivers are measured. Next, the spatial orientation of the camera relative to the SC ACS is calculated.

The radio channel is used to synchronize receivers and transmitters.

B. Results of comparison of goniometric systems

An analysis of the GS options, which took into account the possibility of automatically calculating the angular position of the camera in the SC ACS, the possibility of pointing the camera according to target designations, the presence of restrictions on the work of the crew and other criteria, showed that option 4 is the most preferable.

This option has been selected for further use.

Photos of GS-U are shown in Figures 1,2.

To increase the efficiency of the process of surveying the earth's surface from a manned spacecraft using GS-U and processing the images, new technologies were developed:

- technology for georeferencing the Earth images,

- technology for supporting the astronaut's pointing of the camera at given objects,

- technology for planning observations of list of objects.

Fig. 1. Photocamera with a platform of emitters: 1 – platform attached to the camera; 2 – ultrasonic emitters.

Fig. 2. Receiver platform with control unit: 1 – platform located around the window, with ultrasonic receivers; 2 – porthole; 3 – cable connecting the control unit to the crew computer; 4 - control unit with Wi-Fi communication device; 5 - cable for connecting ultrasonic receivers to the control unit; 6 – ultrasonic receiver.

C. Technologies for image reference, guidance support and survey program planning

The following data is used as initial data for the Earth's images georeferencing technology:

- coordinates and orientation of the spacecraft at the moment of shooting;

- matrix of rotation of the receiver platform relative to a given porthole;

- matrix of porthole rotation relative to the associated spacecraft coordinate system;

- matrix of camera rotation relative to the emitter platform.

The main stages of georeferencing an image:

- calculation the rotation matrix of the emitter platform relative to the receiver platform based on CRS-U data;

- calculation the camera rotation matrix relative to the spacecraft coordinate system;

- calculation the coordinates of the center of the image as the point of intersection of the camera's optical axis with the earth's surface;

- identification by the operator of the shooting object on the orthomosaic of the earth's surface using the coordinates of the image center;

- converting an image file to orthographic projection and enters it into the database [5, 6].

Technology to support the astronaut's pointing of the camera at given objects, including calculation of the current position of the spacecraft, determination of the angular position of the camera lens axis relative to the spacecraft, calculation of the coordinates of the intersection of the lens axis with the earth's surface and generating visual and/or voice instructions to deflect the camera axis so that the center of the image aligns with a given object.

GS-U constantly calculates the direction of the camera axis. Based on these data, as well as the location of the spacecraft in orbit and its orientation, the coordinates of the intersection of the camera axis with the Earth's surface and the deviation of this point from a given object are calculated. In this case, the astronaut is given visual and/or voice instructions on where to deflect the camera axis so that the center of the image aligns with a given object. The camera shutter can be released manually by the astronaut, or automatically by a signal from the GS-U when the deviation of the center of the camera frame from a given object is acceptable.

Visual information to support guidance is provided using an additional small-sized display (Figure 3) located on the camera.

Voice information is provided through the on-board computer [7, 8].

Fig. 3. Emitter platform with a small-sized screen.

The technology for planning observations of ground-based objects includes predicting the SC orbit, calculating the position of objects along the flight path, estimating the time spent by the astronaut to point the camera at each

individual object and to reorient the camera between observations, and determining the required number of cosmonauts and cameras involved in the filming [3]. This technology is based on the experience of planning space experiment programs at Russian orbital stations, starting with the Salyut station [9 – 15].

The problem of planning a survey from an orbital spacecraft with hand-held cameras of a given list of objects is considered as a problem of choosing a route for a mobile observer, taking into account existing restrictions on the moments of surveying [16, 17]. It is required to construct a program for shooting a given list of objects in such a way as to ensure the specified duration of shooting of each object and minimize the total time of reorienting the camera from one object to another. If all objects cannot be photographed by one astronaut on one orbit, then the desired route should consist of segments, each of which is a sequence of objects photographed by a separate astronaut with a separate set of imaging equipment. Within the framework of graph theory, an algorithm for solving the problem is proposed as a problem of minimum dimension equal to the sum of the number of survey objects and the number of sets of observation equipment used. Taking into account the specifics of the problem under consideration, which consists in the fact that the possibility of surveying objects arises sequentially as the spacecraft passes over the survey area, an algorithm is proposed that provides optimized solutions without restrictions on the number of surveyed objects.

The technology for planning observations of ground-based objects includes estimating the average speed and accuracy of camera reorientation by an astronaut between objects, including both angular and linear movement of the camera near the window and between windows [18 - 21].

Based on this assessment, the possibility of including the object in the shooting program is determined.

When drawing up an observation program, the route of potential observations is calculated in the form of obvious simple solutions - ordered sequential objects in the order of time of their minimum distance from the ground path or in the order of time of their visibility zones from the SC (if such solutions are available).

In the absence of simple solutions, the exact shortest solutions are calculated, obtained by database search (if computational resources are available) and optimized solutions (by solving optimization problems).

Based on the results of the analysis of these decisions, the required number of astronauts and filming equipment is determined, and the best option is accepted for implementation.

Figure 4 shows examples of optimized solutions for a catalog of 40 objects under the following conditions: the SC orbital altitude is 400 km, the viewing angle of the observation equipment from nadir is 30°, the speed of reorientation of the sighting axis of the cosmonaut's observation equipment is 30°/s, the minimum visibility time catalog objects from the SC 20 sec.

In Figure 4, the graphs are shown in a reference system fixed along the ground path. The dotted line reflects the directions to the observed object from the SC at the beginning and end of the observation. The dashed line reflects the intervals corresponding to the projections on the ground path from the intervals when the SC was located when the object was available for observation.

(a)

(b)

Fig. 4. Optimized sequences of object observation with minimum 20 sec time of object visibility from the SC 400 km orbit: (a) - Δ=12 s, K=1; (b) - Δ=20 s, K=2 (Δ – minimum guaranteed shooting time for each object; K – number of parallel shooting processes).

The graphs show solutions with K=1, 2 sections of the route (each section corresponds to an individual set of observation equipment) and an indication of the duration of observation of each object Δ achieved in the solution.

If it is necessary to increase the volume of filming, the technology allows us to develop anoption with the involvement.

D. Conclusions

- A set of hardware and software tools and corresponding technologies have been developed to photograph the Earth's surface with freely movable cameras on board an orbital SC and to process the resulting images.

- The use of the presented goniometric system and camera pointing support tools was approved by the cosmonauts of the Russian segment of the ISS.

- The use of the presented tools makes it possible to increase the efficiency of the process of obtaining and processing images of the Earth's surface.

- The presented instruments can be used to capture and process images of various external objects of a manned SC.

III. AUTOMATION OF THE STORAGE PROCESS

On the ISS, an identification system based on bar codes is used to store and account for spare parts for crew life support equipment, spare parts, etc., hereinafter referred to as storage items (SITEs). The volume of SITEs on the RS ISS reaches 10,000. At the same time, the database stores a complete list of storage units, storage location and other additional information. If, when using a storage unit, each of SITE displacement is entered into the database, then the system allows you to find the required SITE in a short time. However, during the activities of astronauts, situations are possible when data is not entered into the database. This is due to the fact that errors may occur in the work of astronauts due to the need to switch attention when performing several flight operations simultaneously. The likelihood of errors increases when working under time pressure, the presence of disturbances, such as the receipt of messages from the control center, from other crew members, the receipt of data on the state of on-board systems, the activation of on-board automatic alarms, etc. As a result of errors, not all SITE replacements are recorded in the database. The amount of unreliable information is steadily increasing. This leads to the fact that a partial or complete inventory of storage facilities is planned on board, which can take several crew working days. Therefore, it is relevant to develop technology for storage and tracking of SITE, which reduces the likelihood of presence corrupted information in the storage database.

A. Development of a storage system

It was developed and investigated on RS ISS storage system for crew medicines. It is so called "smart shelf" (SS) based on radio tag technology (RFID technology) [22 - 24].

The SS includes the following units:

- a rack consisting of four compartments for storing medicines;

- antenna for reading RFID tags (in each compartment);

- control unit with Wi-Fi communication device;

- tablet PC;

- containers for medicines (CM).

Each CM is intended for one drug and is a plastic plate that has a mechanical interface with the SS, equipped with an RFID tag.

Figure 5 shows the preparation SITE for loading into SS on the RS ISS.

SS constantly monitors the availability of CM in the SS. Access to the SS equipment database is possible both using a tablet computer included in the SS, and through the on-board local network from the any on-board computer, as well as through a radio channel from the control center computer. This allows you to entrust control of the state of the contents of the SS to the MCC personnel, freeing the crew from this task.

Fig. 5. Cosmonaut Sergei Kud-Sverchkov holds a CM in his hand while testing the SS equipment tter platform with a small-sized screen.

In the event of a crew error, for example, failure to return the CM to the SS, the control center can quickly identify this event and give an appropriate recommendation to the crew. This allows preventing the accumulation of deviations in the SS database.

The probability of correctly identifying the contents of the SS for the selected design is 100%.

B. Conclusions

- During testing of the SS equipment, the performance, functionality and ergonomics of the selected SS structure were confirmed.

- The SS is integrated into the structure of on-board systems and the space complex, which allows real-time work with the SS from any on-board workstation connected to the on-board network, from a workplace in the MCC, if there is a MCC-on-board communication channel.

- It is proposed that promising space station modules be equipped with the required number of SS, taking into account the number and range of important SITEs, the presence of which must be monitored during the flight and to which quick access is required during the flight.

IV. AUTOMATION OF THE CREW MEMBERS' LOCATION TRACKING PROCESS

In the operation of manned orbital stations, there are a number of flight operations in which the mission control

team needs to know the location of crew members, as some instrument or piece of equipment that is critical. It is necessary to know the location of each crew member inside the station when making decisions in the process of eliminating emergency situations; when the crew performs dangerous operations (for example, EVA). In addition, to plan the work of astronauts, to improve flight operations and to design promising space stations, statistics on the location of crew members in various work areas of the station are needed. It should be taken into account that astronauts do not like to constantly be in the field of view of video cameras.

A. Development of a system for determining the coordinates of an astronaut's position in the compartments of the space station

To increase the efficiency of the process of determining the coordinates of an astronaut's location in the internal compartments of the space station on the RS ISS, a coordinate reference system using infrared sensors (CRS-I) [25 - 28] was developed and studied.

CRS-I includes the following elements:

- infrared autonomous receivers (IRR);
- autonomous infrared beacons (IRB);
- control unit (CU);
- on-board computer with software.

The diagram of the CRS-I is shown in Figure 6.

Fig. 6. Diagram of the CRS-I system in the interior of the space station: 1 – on-board laptop; 2 – control module; 3 – infrared beacon; 4 – infrared receivers.

Figure 7 shows a photograph of an astronaut with two CRS-I beacons attached to his clothes on his shoulders.

The CRS-I operates in a "pulse" mode. The MU launches each IRB located in the controlled space in turn via a radio channel. In response, each IRB emits an IR pulse. Each IRR that has received an IR pulse from the IRB transmits to the CU the two-dimensional coordinates of the center of the light spot formed as a result of the action of the IR pulse on the IRR sensor. Measurement data from the CU enters the computer, where the current three-dimensional coordinates of all IRBs currently located in the controlled space are calculated.

Fig. 7. Cosmonaut Oleg Novitsky with two IRB attached to his shoulders.

B. Methodology for calculating the coordinates of an astronaut's position

Determining the location of an object (cosmonaut) relative to the ISS coordinate system is carried out as follows.

After deployment of the system, calibration is carried out: the parameters of the relative position of the beacons installed in given fixed positions are determined, and the parameters of the relative position of several beacons fixed on the given moving parts of the astronaut are determined for several (at least two) given fixed positions of the moving parts of the astronaut.

During the monitoring process, the following operations are performed.

1) The control unit generates control actions on the beacons - it launches each beacon in turn via a radio channel.

2) The IR receivers measure and transmit to the control unit information about the center of the light spot formed as a result of the IR pulse hitting the receiver detector.

3) The control unit transmits measurement data to the on-board computer.

4) Based on the measurements received from the IR receivers and the specified parameters for the location of detectors and optical systems of the receivers, the current three-dimensional coordinates of all beacons currently located in the controlled space are calculated (note: to determine the coordinates, it is necessary to obtain measurements from at least four receivers).

5) Based on the current values of the beacon coordinates and the parameters of the relative position of the beacons, determined for given fixed positions of the astronaut's moving parts at the calibration stage, the parameters of the cosmonaut's current position relative to the RS ISS coordinate system are calculated.

C. SKP software

The initial data for the coordinate calculation algorithm are:

- coordinates of the light spot formed as a result of the IR pulse hitting the receiver sensor, in the receiver's coordinate system;

- coordinates of the infrared receiver in the coordinate system of the RS ISS;

- the value of the angles between the axes of the receiver coordinate system and the axes of the RS ISS coordinate system.

The following algorithm is used to calculate the three-dimensional coordinates of an object.

1) The translation the linear coordinates of the center of the light spot into angles that determine the direction to the beacon relative to the optical axis of the receiver. The resulting angles are converted into a three-dimensional vector in the coordinate system of the receiver. This vector is translated into the RS ISS coordinate system.

2) The beacon coordinates are calculated by finding the intersection of the point of closest convergence of two three-dimensional rays from receivers with the best signal amplitude (light spot brightness).

The accuracy of determining the astronaut's location is about 1 meter.

D. Conclusions

- During testing of the CRS-I equipment, the operability and functionality of the cosmonaut position control system on the space station was confirmed.

- The CRS-I equipment is integrated into the space complex, including on-board systems and ground control facilities, which makes it possible to determine the astronaut's location in real time from any on-board workstation connected to the on-board network, as well as from the MCC personnel workplace.

- CRS-I equipment can be used to monitor astronauts' location during certain flight operations (carrying out work in an emergency situation, during emergency situations training, to control the movement of robotic products moving independently around the station).

The disadvantages of CRS-I include the need to calibrate the equipment before each session, which does not allow its use in automatic mode.

REFERENCES

[1] M.Yu. Belyaev, L.V. Desinov, D.Yu. Karavaev, V.P. Legostaev, V.V. Ryazantsev, and O.A. Yurina, "Features of conducting and using the results of surveying the earth's surface performed by the crews of the Russian segment of the ISS," Kosm. Tekh. Tekhnol., 2015, No. 1(8), p. 17-30.

[2] V.I. Vasiliev, I.G. Sokhin, S.V. Bronnikov, N.V. Vasilyeva, and O.S. Gordienko, "Visual and instrumental observations from the International Space Station by crews of the Russian segment and the basic principles of preparation for their implementation," Manned flights into space, 2013, No. 2 (7), p. 23-29.

[3] S. Bronnikov, D. Karavaev, A. Rozhkov, and D. Rulev, "Enhanced method to perform crew earth observation onboard the ISS with use of relocatable cameras," 74th International Astronautical Congress (IAC), Baku, Azerbaijian, 2-6 October 2023. IAC-23-B3,4-B6.4,5,x75868. – 6p.

[4] S.V. Bronnikov, E.I. Malimenkov, A.S. Rozhkov, D.Yu. Karavaev, A.Yu. Kaleri, S.K. Krikalev, and P.V. Vinogradov, "Method for determining the geographic coordinates of images of objects on the surface of the planet when shooting from a manned spacecraft," Patent for invention of the Russian Federation No. RU 2353902. Bull. No. 12 dated 04/27/2009.

[5] S.V. Bronnikov, D.Yu. Karavaev, and A.S. Rozhkov, "Study of technology and means of georeferencing images of the Earth obtained on a manned spacecraft using freely moving cameras," Kosm. Tekh. Tekhnol., 2016, No. 2 (13), pp. 105-115.

[6] S.V. Bronnikov, A.S. Rozhkov, D.Yu. Karavaev, D.N. Rulev, O.S. Rurin, A.K. Kalifatidi, and I.G. Gorodetsky, "A method for determining the geographic coordinates of the observation area of observation equipment moving relative to a spacecraft, a system for its implementation and a device for placing emitters on the observation equipment," Patent for invention of the Russian Federation No. RU 2524045. Bull. No. 21 dated 07/27/2014.

[7] S.V. Bronnikov, A.S. Rozhkov, P.A. Borovikhin, and D.N. Rulev, "Method for orienting equipment moved on board a manned spacecraft," Patent for invention of the Russian Federation No. RU 2695739. Bull. No. 21 dated 07/25/2019.

[8] S.V. Bronnikov, P.A. Borovikhin, D.Yu. Karavaev, and D.N. Rulev, "Method for controlling surveillance equipment moved on board a manned spacecraft," Patent for invention of the Russian Federation No. RU 2780900. Bull. No. 28 dated 10/04/2022.

[9] M.Yu. Belyaev, "Operational planning of scientific experiments carried out using spacecraft," Cosmic Research, 1980, Vol. 18, pp. 235-243.

[10] M.Yu. Belyaev, Scientific experiments on spacecraft and orbital stations. Moscow: Mashinostroenie, 1984.

[11] M.Yu. Belyaev and D.N. Rulev, "Optimization of the experimental program for operational planning of research performed from spacecraft," Cosmic Research, 1987, Vol. 25, No. 1, pp. 30-36.

[12] M.Yu. Belyaev and D.N. Rulev, "Optimization of planning experiments performed at orbital stations," Cosmic Research, 2007, Vol. 45, No. 3, pp. 236-243.

[13] M.Yu. Belyaev, P.A. Borovikhin, A.M. Vetoshkin, D.Yu. Karavaev, and I.V. Rasskazov, "Guidance of the scientific equipment of the International space Station on the objects under study," Cosmic Research, 2022, Vol. 60, No. 1, pp. 1-10.

[14] M.Yu. Belyaev, P.A. Borovikhin, A.M. Esakov, D.Yu. Karavaev, and I.V. Rasskazov, "Optimization of pointing scientific equipment at observed objects in the Uragan experiment on the ISS," Gyroscopy and Navigation, 2023, Vol. 31, No. 3 (122), pp. 118-137.

[15] M.Yu. Belyaev, P.A. Borovikhin, D.Yu. Karavaev, and I.V. Rasskazov, "Optimizing algorithms for visual and instrumental observations taken by the crew of the Russian Segment of the International Space Station," 74th International Astronautical Congress (IAC), Baku, Azerbaijan, 2-6 October 2023. IAC-23-B3.5.8. -11p.

[16] D.N. Rulev, "Planning Optimization of Observations of Multiple Objects from a Flying Vehicle Following a Predetermined Trajectory," Kosmonavtika i raketostroenie, 2021, No. 5 (122), pp. 21-39.

[17] D.N. Rulev, "Optimizing the Planning of the Observation of a Catalog of Objects by a Mobile Observer, Taking the Implicated Limitations into Account," Journal of Computer and Systems Sciences International, 2022, Vol. 61, No. 2, pp. 195–212. © Pleiades Publishing, Ltd., 2022. Russian Text © The Author(s), 2022, published in Izvestiya Akademii Nauk, Teoriya i Sistemy Upravleniya, 2022, No. 2, pp. 62–79.

[18] S.V. Bronnikov, A.S. Rozhkov, P.A. Borovikhin, and D.N. Rulev, "Orientation system for equipment moved on board a manned spacecraft," Patent for invention of the Russian Federation No. RU 2695254. Bull. No. 21 dated 07/22/2019.

[19] S.V. Bronnikov, A.S. Rozhkov, P.A. Borovikhin, and D.N. Rulev, "Method for orienting equipment moved on board a manned spacecraft," Patent for invention of the Russian Federation No. RU 2695046. Bull. No. 20 dated 07/18/2019.

[20] S.V. Bronnikov, A.S. Rozhkov, D.Yu. Karavaev, and D.N. Rulev, "Method for orienting equipment moved on board a manned spacecraft," Patent for invention of the Russian Federation No. RU 2693634. Bull. No. 19 dated 07/03/2019.

[21] S.V. Bronnikov, A.S. Rozhkov, P.A. Borovikhin, and D.N. Rulev, "Method for orienting equipment moved on board a manned spacecraft," Patent for invention of the Russian Federation No. RU 2695041. Bull. No. 20 dated 07/18/2019.

[22] S.V. Bronnikov, A.S. Rozhkov, P.A. Borovikhin, and D.N. Rulev, "Orientation system for equipment moved on board a manned spacecraft," Patent for invention of the Russian Federation No. RU 2695960. Bull. No. 22 dated 07/29/2019.

[23] S.V. Bronnikov, D.Yu. Karavaev, and A.S. Rozhkov, "Study of positioning technologies in the space experiment Vizir," Materials of the 50th Scientific Readings in memory of K.E. Tsiolkovsky. Kaluga, September 15-17, 2015, pp. 113-114.

[24] S.V. Bronnikov, A.S. Rozhkov, P.A. Pozdnyakov, A.N. Leporsky, and D.A. Sokolov, "Application of radio tag technology on the international space station," Materials of the 51st Scientific Readings in memory of K.E. Tsiolkovsky. Kaluga, September 20-22, 2016, pp. 96-97.

[25] S.V. Bronnikov, A.S. Rozhkov, D.N. Rulev, P.A. Pozdnyakov, A.K. Kalifatidi, D.A. Volokhovsky, and Yu.A. Privalov, "A device for storing and identifying moving objects on a spacecraft," Patent for invention of the Russian Federation No. RU 2665914. Bull. No. 25, dated 09/04/2018.

[26] S.V. Bronnikov, A.S. Rozhkov, P.A. Pozdnyakov, D.N. Rulev, D.A. Volokhovsky, Yu.A. Privalov, and A.A. Nabok, "A method for determining the position of an object primarily relative to a spacecraft and a system for its implementation," Patent for invention of the Russian Federation No. RU 2600039, Bull. No. 29 dated 10/20/2016.

[27] S.V. Bronnikov, A.S. Rozhkov, and D.N. Rulev, "A method for controlling the movement of an astronaut to identifiable objects on the space station and a system for its implementation," Patent for invention of the Russian Federation No. RU 2669155. Bull. No. 28, dated 10/08/2018.

[28] S.V. Bronnikov, A.S. Rozhkov, and D.N. Rulev, "A method for controlling the movement of an astronaut relative to a spacecraft and a system for its implementation," Patent for invention of the Russian Federation No. RU 2604892. Bull. No. 35 dated 12/20/2016.

Regional Earth Survey Planning and Control of Mini-Satellites in Low-Orbit Constellations

Y. Somov
Dynamics and Motion Control
Samara Federal Research Scientific
Center, Russian Academy of Sci-
ences Samara, Russia
e_somov@mail.ru

S. Butyrin
Dynamics and Motion Control
Samara Federal Research Scientific
Center, Russian Academy of Sci-
ences Samara, Russia
butyrinsa@mail.ru

S. Somov
Dynamics and Motion Control
Samara Federal Research Scientific
Center, Russian Academy of Sci-
ences Samara, Russia
s_somov@mail.ru

Abstract — **The tasks of ballistic designing a constellation of mini-satellites for regional space survey with the maintenance of their sun-synchronous orbits and problems on motion control of mini-satellites in low-orbit constellations are considered.**

Keywords — low-orbit constellation design, earth observation mini-satellites, spatial motion control

I. INTRODUCTION

Current problems of Earth observation by small satellites from low orbits are presented in review articles [1-3], and approaches to their ballistic justification, – in well-known monographs [4-9]. With a coordinated survey, the orbit of each mini-satellite in the constellation must pass through a given ascending node. Sun-synchronous orbits (SSO) [10-17] have this property with a multiplicity of $n_g = T_d / T_p$, where T_d is the duration of sunny day and T_p is the nodal period of a spacecraft (SC).

For regional Earth survey using a constellation of mini-satellites, the most convenient orbits are with a daily multiplicity and a height of 570 km, when a period $T_d = 86400$ s

Fig. 1. The Earth survey SC

and multiplicity $n_g = 15$. Orbits with a higher multiplicity require an increase of the SC number in the constellation to ensure a given observation repeatability interval and with a decrease in the orbital height, fuel costs significantly increase to compensate for the influence of a resistance to the SC motion. In the article we study

(i) the tasks on ballistics of satellites constellation for regional Earth survey and maintaining the SSO with ensuring solar synchrony and stability of local mean solar time (LMST) [17] over long time intervals when correcting the SC motion by electric propulsion units (EPU) with small thrust;

(ii) the problems of satellite guidance and motion control using a gyroscopic moment cluster (GMC) based on gyrodines (GDs), when the measurement of the SC motion coordinates is performed by a strapdown inertial navigation system with correction by signals from GLONASS satellites and a cluster of star trackers.

The earth survey mini-satellite is equipped with a telescope with optoelectronic converters (OEC) in its focal plane, Fig. 1. When performing scanning routes of the earth's surface with a telescope, the OEC matrices operate in the time delay and integration (TDI) mode.

With a known orbit, the survey route is determined by the law of the SC angular guidance to ensure the required movement of the optical image on OEC matrices. This law is formed according to the analytical relationships between the movement of the image and the change in the coordinates of the SC spatial motion.

II. CONSTELLATION BALLISTIC DESIGN

To describe the SC motion, we use an inertial reference frame (IRF $O_\oplus X^1 Y^1 Z^1$, basis $\mathbf{I}_\oplus$), orbital (ORF, basis $\mathbf{O}$ with the unit vectors of radial $\mathbf{r}^\circ$, transversal $\boldsymbol{\tau}^\circ$ and binormal $\mathbf{n}^\circ$), the SC body (BRF $Oxyz$, Fig. 1, basis $\mathbf{B}$) reference frames with the origin at its mass center O, geodetic Greenwich reference frame (GRF), associated with the rotating Earth. The notation $\{\cdot\} = \mathrm{col}(\cdot)$, $[\cdot] = \mathrm{line}(\cdot)$, $(\cdot)^t$, $[\cdot\times]$, $\langle\cdot,\cdot\rangle$ and $\circ, \tilde{}$ are applied for vectors, matrices and quaternions, as well as $S_\alpha = \sin\alpha$, $C_\alpha = \cos\alpha$.

In the column $\mathbf{r} = \{\Omega, i, \omega_\pi, p, e, \nu\}$ of the SC location in orbit, we distinguish: (i) three osculating elements of its orientation – longitude of the ascending node (AN) Ω, inclination i, perigee argument ω_π and (ii) three elements that determine the size, shape of orbit, and the SC position – the focal parameter p, eccentricity e and the true anomaly ν, uniquely related to the argument of orbital latitude $u(t)$.

The design of constellation orbits consists in determining the initial data, when the required repeatability of the passage of their traces through given points of the region on the earth's surface is ensured. Here, the SSO of all satellites in the constellation have the same height and pass through a single AN of the orbit on the earth's surface, but at different times with the repeatability interval of T_r duration.

We will assume that the constellation consists of N local groups of mini-satellites and each local group has K satellites in the direction of their flight, for example, for $K = 3$: left, central and right, with similar values of their longitude of ascending node (LAN).

The initial data are the orbital parameters of the central satellite and duration of the repeatability interval. It is necessary to determine the number N of the central satellites (local groups) within the entire constellation and initial parameters of the satellites' orbits.

Fig. 2. Bands of regional Earth survey in the European part of the Earth Eastern hemisphere

With the same flight height of spacecraft, the parameters of their orbit differ only in the time moments when they pass their AN with a longitude Ω.

The number of central mini-satellites is determined taking into account the solar illumination of ground-based regional observation objects, which depends on their geodetic coordinates, date and time of observation. The best conditions for optoelectronic observation of the earth's surface occur in the range from 10 to 14 hours local time [17]. It is advisable to extend this range from $t_i = 9$ to $t_f = 15$ hours.

The location of satellites in each local group is determined by both type (route, object, areal, stereo, etc.), and the style of scanning optoelectronic survey – tracing, orthodromic, with equalization of longitudinal an image motion velocity (IMV) on OEC matrices. For example, in an areal survey [18], the orbital planes of 3 mini-satellites in a local group are spaced in the LAN by an angle $\Delta\Omega = 0.84$ deg.

If we set the observation repetition interval $T_r = 1$ h, then to observe the entire surface of the Earth, the central mini-satellites are located in 24 orbital planes in the IRF with a step of their relative rotation by an angle

$$\Delta\Omega_c = (361[\text{deg}] / [24\,\text{h}]) \cdot [1\,\text{h}] = 15.04 \text{ deg.}$$

A well-known system of equations for the spacecraft translational motion in osculating variables with standard notations is used. All satellites in local groups are numbered in the order of the orbital planes in the IRF in ascending order of their LAN.

Let the geodetic coordinates of a point on the earth's surface through which the path of the central satellite #1 passes at the moment t_o be given, for example, as $B_o = 5.53$ deg, $L_o = 10.01$ deg (Saint Petersburg), and on the calculated date 21.03.2023 the local time $t_o = 09:00:00$ is set.

The paths of other central satellites must pass through this ground object at an interval of $T_r = 1$ h during daylight $T_N = t_f - t_i + T_r = 15 - 9 + 1 = 7$ h, then the number of groups is $N = T_N / T_r = 7$. For example, for the central satellite #1, the LAN of its orbit in the GRF is calculated, in which at a given time t_o the satellite path passes through the observation object with the geodetic coordinates B_o and L_o.

The turn L_o of the SC orbit around the Earth rotation axis does not change the angular relationships for the orbit elements; therefore, numerical integration of the equations of motion of the central mini-satellite #1, until it reaches a given geodetic latitude B_o, allows determining the value of the orbit LAN $L_\Omega = L_o - L(t_o)$ in the GRF. Here the values were calculated:

Fig. 3.   Bands of regional Earth survey for the European part of the Russian Federation

$L_\Omega = 47.79$ deg; LMST $t_\Omega = 09:53:20$ when the satellite passes AN of the orbit (UTC $06:42:08$); in the IRF a longitude of the AN $\Omega = L_\Omega + S_o = 5.70$ deg.

Here S_o is the sidereal time of the Greenwich meridian at the time moment t_Ω, presented in angular measure.

The central satellites of other local groups move in orbital planes with a turn in the IRF at an angle of 15.04 deg.

At given number $n \in [1,2,...7]$ of the central satellite, the value of its LAN is calculated in degrees at the form $\Omega = (180/\pi)(L_\Omega + S_o) + 15.04(n-1)$.

If each local group has $K = 3$ mini-satellites, then the number of spacecraft in the entire constellation is $N_s = 21$.

The LAN values for the central spacecraft and the time parameters of the constellation are presented in Table I.

TABLE I

The LAN values and parameters of the central SC orbits

n, #SC	Ω, deg	Date, d : m : y	UTC, h : m : s	t_Ω, h : m : s	t_o, h : m : s
1	-33.13	21.03.2023	06:42:08	09:53:20	09:00:00
2	-18.09	21.03.2023	07:42:08	10:53:20	10:00:00
...	…	...	...	...	...
7	57.11	21.03.2023	12:42:08	15:53:20	15:00:00
1	-33.13	22.03.2023	06:42:08	09:53:20	09:00:00

When limiting $|\gamma| \le 20$ deg to the roll angle γ of each mini-satellite relative to the local vertical in the ORF, the band of available Earth survey (overview stripe) from geodetic latitude 32 deg to latitude 70 deg is shown in Fig. 2, gray central stripe (band). On the Earth surface, the average width of such an overview stripe is 445 km. Then the geodetic latitude changes along the route, the local time changes, here by 35 minutes, which has little effect on the solar lighting conditions of the corresponding part of the region.

Let us now consider the possibility of using an orbital constellation of mini-satellites to survey the Earth from neighboring orbits. Firstly, we determine the position of a possible overview stripe for the central satellite #1, when observation is carried out one orbit later. Then in Fig. 2, we get a left overview stripe with a route running along the Naples – Hamburg trajectory. It is not difficult to establish that in the survey from one orbit earlier, the right overview stripe in Fig. 2 is obtained – with a route running along the Ural ridge south to Kandahar.

In total, here 15 such regional overview strips will be located on the Earth surface, which have the variations in the UTC and LAN Ω values due to SSO precession. Local time and lighting conditions in all such strips are almost the same.

The problem of regional land survey of the European part of Russia attracts attention when the geodetic latitude of the observed objects is from 43 to 70 degrees, Fig. 3.

The 105 mini-satellite constellation used here consists of 5 groups with 21 satellites in each group, which is simultaneously represented by 7 local groups of 3 mini-satellites. Satellites in each local group move in close orbits with small differences in their LANs.

In the IRF, the orbital planes of the 7 central satellites of local groups are spaced apart by the LAN at an angle of 15.04 degrees, and mini-satellites with the same end-to-end numbers in the groups from the $1 \div 21$ range move in orbits that differ in the LAN values by 5.5 deg. Each of these groups provides one of the 5 overview bands in Fig. 3.

If required, then for each of these bands, it is possible to provide the same local survey time at the same geodetic latitude. For example, shown in Fig. 3, the calculation was performed under the condition that in all 5 groups the paths of the central satellites #2 (if they are numbered consecutively in the group) cross the geodetic latitude $B = 50$ deg at the same LMST value. The presented constellation of 105 mini-satellites provides imaging of objects with a given range of changes in geodetic latitude throughout the entire northern hemisphere of the Earth when using other orbits of mini-satellites in the constellation.

Let us briefly consider the problem of estimating fuel costs when shifting the earth overview bands to rearrange such an orbital constellation.

To do this, it is necessary to perform one of two options for changing the position of each satellite in the constellation: 1) rotating the orbital plane with changing its LAN; 2) change in the phase of the satellite's motion – change in orbital latitude. Both types require high fuel consumption.

First, let us consider a shifting overview band by changing the orbital LAN by an amount $\Delta\Omega = 2$ deg. Using the results of the authors on control of a space robot, an estimate of the EPU fuel consumption $\Delta m = 31.57$ kg was obtained for a maneuver duration of about 5 days.

The phase change consists of a temporary increase in the orbital altitude followed by the return of the satellite to its original orbit. Here, a decrease in the average angular velocity of the SC orbital motion is applied and the orbital latitude of its direction is changed to a sun-synchronous orbit.

As the satellite moves until it passes through the AN of its orbit, the Earth rotates and, therefore, the spacecraft's flight path shifts to the West, but then the effect of aerodynamic drag again shifts the satellite's path to the East. With such a maneuver, the spacecraft's flight altitude temporarily increases by 777 km, the duration of the maneuver is 1.03 days, and the fuel consumption is $\Delta m = 2.98$ kg.

III. KEEPING MINI-SATELLITES ON SSO

Methods for long-term keeping of earth survey satellites at given SSO using EPU have been studied in many scientific works [10-17]. Let's introduce a notation L_Ω^* for the required value for the geodetic longitude of the AN $L_\Omega = L(t_\Omega)$ in the GRF – the value of the longitude of the sub-satellite point at the time t_Ω, when the mini-satellite passes the ascending node of the corresponding orbit.

As is known, the correction of the LAN L_Ω is carried out by changing height $h(t)$ (i.e., the semi-major axis $a(t)$) of the orbit, but to maintain solar synchrony and stability of the LMST, correction of its inclination is necessary. The task is to estimate the parameters of the orbit corrections – their number, periodicity, duration of EPU operation intervals and fuel consumption. Deviation $\Delta L_\Omega^* = |L_\Omega - L_\Omega^*|$ of the LAN from required value, for example, a $\Delta L_\Omega^* = 0.1$ degree, corresponds to a length of 11 km on the earth's surface.

The relationship between the deviation ΔL_Ω^* and the change $\Delta a(t_c)$ in the semi-major axis of SSO is presented in [17] as $\Delta a(t_c) = k_a \sqrt{[8a(t_c) |\dot{h}| \Delta L_\Omega^* /(3\omega_e)]}$. Here, $a(t_c)$ is the value of the semi-major axis of the orbit at the time t_c for execution of the correction impulse on the SC velocity, $\dot{h}$ is the rate of change in the orbital height due to aerodynamic drag, ω_e is the angular velocity of the Earth's rotation, and coefficient k_a takes into account the influence of other disturbances.

In the developed calculation method, the values of the vectors of corrective impulses of the spacecraft velocity are first calculated on the basis of impulse theory, and then their implementation is calculated using low-thrust EPUs.

When predicting the motion of a satellite on the next orbit turnover, the time instants of the satellite passing perigee t_π and apogee t_α, are determined, the parameters of a two-pulse transition between circular orbits, the semi-major axes of which differ by Δa, and the values of the vectors of velocity pulses are calculated at perigee $\Delta\mathbf{v}_\pi$ and apogee $\Delta\mathbf{v}_\alpha$. These vectors are directed along the velocity unit vector $\boldsymbol{\tau}^\circ$ of the satellite translational motion.

Correction of orbital inclination is performed under the condition $|\Delta i| \geq \Delta i^*$, where Δi^* is a limit of the permissible difference in inclination from its nominal value of 97.67 deg.

In this case, impulses of the spacecraft velocity vector are formed at the nodes of orbit in the direction of its unit binormal vector.

Fig. 4. Deviations of the semi-major axis and inclination of the SSO from their initial values

Fig. 5. Deviations of the LAN and LMST from the nominal values when passing the ascending node

Let us briefly present the results of computer simulation of processes during correction of the SSO semi-major axis and inclination with a nominal height of 570 km for the satellite with a mass of 350 kg and the following data:

• the EPU contains two electric propulsion engines with a total thrust of 0.58 H, which allows any necessary orbit correction to be performed in 1 orbit turnover;

• the SC midsection area is 15 m^2, the aerodynamic drag coefficient $c_x = 2.5$;

• at the time moment $t_0 = 0$, orbital parameters $e = 10^{-4}$, $\omega_\pi = 0$, $L_\Omega^* = 41.8$ deg, $p = 6948.1359$ km and $i = 97.64$ deg are adopted, which differs from value $i = 97.67$ deg for minimizing $|\Omega(t_0) - \alpha_s(t_0)|$, where $\alpha_s(t)$ is the right ascension of the mean equatorial Sun.

The estimated date is 21.03.2023, the initial values on the first orbit turnover are as follows $a_1 = 6938.603$ km and $i_1 = 97.645$ deg, at the time moment t_0 satellite is in the ascending node at a longitude value $\Omega = L_\Omega^* + S_o$ in the IRF.

Figure 4 shows changes in the semi-major axis $\Delta a = a - a_1$ and inclination $\delta i = i - i_1$ of the SSO depending on the number n of its turnover, and in Fig. 5 – changes in LAN $\Delta L_\Omega(t_\Omega)$ and local mean solar time $\delta t_\Omega = t_\Omega - t_{\Omega 1}$ at the instants of passage of the ascending node on the n-th orbit turnover, see green dashed lines in the absence of inclination correction. The developed method makes it possible to keep the deviation of the SSO AN values within specified limits. Under the influence of gravitational disturbances, the LAN $\Delta L_\Omega(t)$ regularly shifts to the western limit of the permissible range, see Fig. 4.

TABLE II

The EPU operating characteristics and fuel consumption				
Corrected parameters	Period, day	Number per year	Duration per turnover, s	Consumption per year, kg
a	12,54	29	339.77	0.329
i	66.71	5	638.17	0.0538
Total	-	34	-	0.488

The inclination correction performed at the 1000th orbit leads to the restoration of solar synchrony of the orbit, stabilization of the "oscillation amplitude" both $\Delta L_\Omega(t)$ and δt_Ω. The created method ensures the maintenance of daily multiplicity, synchronism and stability of the LMST.

Numerical estimates of the frequency of corrections, the duration of operation of the EPU on an orbit, and fuel consumption when performing the above-described SSO corrections were obtained, which are presented in Table II.

IV. ANGULAR GUIDANCE OF SATELLITES

The orientation of the BRF in the IRF is determined by the quaternion Λ and the vector $\sigma = \{\sigma_i\}$ of modified Rodrigues parameters (MRP), and the orientation in the ORF – by the yaw ϕ_1, roll ϕ_2 and pitch ϕ_3. angles. Angular velocity ω and acceleration ε vectors are used.

The body of the satellite is considered a solid with mass m and the inertia tensor $\mathbf{J}$, with the position $\mathbf{r}$ and velocity $\mathbf{v}$ vectors the model of its motion in the IRF has the form

$$\mathbf{r}' + \boldsymbol{\omega} \times \mathbf{r} = \mathbf{v}; \quad m(\mathbf{v}' + \boldsymbol{\omega} \times \mathbf{v}) = \mathbf{P}^e + \mathbf{F}^d;$$

$$\dot{\boldsymbol{\Lambda}} = \boldsymbol{\Lambda} \circ \boldsymbol{\omega}/2; \quad \mathbf{J}\dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times \mathbf{G} = \mathbf{M}^g + \mathbf{M}^d.$$

Fig. 6.   Areal survey of cities in the Volga and North-Western regions of Russia in accessible overview strip with a geodetic latitude from 45.5 to 66 degrees

Here vector $\mathbf{P}^e$ reflects the EPU thrust, vector $\mathbf{G} = \mathbf{J}\boldsymbol{\omega} + \mathbf{H}$, $\mathbf{H}$ is the angular momentum (AM) vector of GMC, vector of its control torque is $\mathbf{M}^g = -\mathbf{H}'$, vectors $\mathbf{F}^d$ and $\mathbf{M}^d$ represent disturbing forces and torques, and $(\cdot)'$ is the symbol of the local derivative with respect to time.

The authors have created analytical methods for synthesizing the SC guidance laws during scanning surveys, which are based on the analysis of the IMV of ground objects on OEC matrices, angular guidance algorithms in the form of a set of smoothly conjugate vector splines MRP $\boldsymbol{\sigma}(t)$.

The problem of calculating the quaternion $\boldsymbol{\Lambda}$ of the basis $\mathbf{B}$ orientation in the inertial basis $\mathbf{I}_\oplus$, the vectors of angular velocity $\boldsymbol{\omega}$ and acceleration $\boldsymbol{\varepsilon}$ in the form of explicit functions at a given time interval, is solved on the basis of the vector addition of all elementary movements of the telescope in the GRF, taking into account the current observation perspective when specifying the initial coordinates of the ground object and the geodetic azimuth A of the scan.

For any point on the focal plane of the telescope, the longitudinal $\tilde{V}_y^i(\tilde{y}^i, \tilde{z}^i)$ and transverse $\tilde{V}_z^i(\tilde{y}^i, \tilde{z}^i)$ compo-

nents of the normalized IMV vector are calculated using the authors' analytical relation

$$\begin{bmatrix} \tilde{V}_y^i \\ \tilde{V}_z^i \end{bmatrix} = \begin{bmatrix} \tilde{y}^i & 1 & 0 \\ \tilde{z}^i & 0 & 1 \end{bmatrix} \begin{bmatrix} q^i \tilde{v}_{e1}^s - \tilde{y}^i \omega_{e3}^s + \tilde{z}^i \omega_{e2}^s \\ q^i \tilde{v}_{e2}^s - \omega_{e3}^s - \tilde{z}^i \omega_{e1}^s \\ q^i \tilde{v}_{e3}^s + \omega_{e2}^s + \tilde{y}^i \omega_{e1}^s \end{bmatrix}.$$

Here, vectors $\boldsymbol{\omega}_e^s$ and $\mathbf{v}_e^s$ represent the angular velocity and the velocity of the SC translational motion in the GRF, orthogonal matrix $\tilde{\mathbf{C}} = \|\tilde{c}_{ij}\|$ determines the telescope orientation in GRF; $\tilde{y}^i = y^i / f_e$ and $\tilde{z}^i = z^i / f_e$ are normalized coordinates when using the telescope focal length f_e; scalar function $q^i = 1 - (\tilde{c}_{21}\tilde{y}^i + \tilde{c}_{31}\tilde{z}^i)/\tilde{c}_{11}$ and for the observation range $D(t)$, the components of normalized velocity vector $\mathbf{v}_e^s$ are represented as $\tilde{v}_{ei}^s = v_{ei}^s(t)/D(t)$. This relation is used to calculate the components of quaternion $\boldsymbol{\Lambda}$, vectors $\boldsymbol{\omega}_e^s$ and MPR $\boldsymbol{\sigma}$.

Next, the guidance laws are approximated by vector splines and vectors $\boldsymbol{\omega}(t)$, $\boldsymbol{\varepsilon}(t)$, $\boldsymbol{\varepsilon}'(t)$ are analytically obtained for scanning survey [18,19].

When there are two adjacent scanning routes (SR) with given boundary conditions for quaternion $\mathbf{\Lambda}$, vectors $\mathbf{\omega}$, $\mathbf{\varepsilon}$ and $\mathbf{\varepsilon}'$ at the start time of the second route, the problem of the satellite rotating maneuver (RM) arises. Spatial RM on the time interval $t \in \mathrm{T_p} \equiv [t_i^p, t_f^p]$, $t_f^p \equiv t_i^p + T_p$, with the boundary conditions

$$\mathbf{\Lambda}(t_i^p) = \mathbf{\Lambda}_i; \ \mathbf{\omega}(t_i^p) = \mathbf{\omega}_i; \ \mathbf{\varepsilon}(t_i^p) = \mathbf{\varepsilon}_i;$$

$$\mathbf{\Lambda}(t_f^p) = \mathbf{\Lambda}_f; \ \mathbf{\omega}(t_f^p) = \mathbf{\omega}_f; \ \mathbf{\varepsilon}(t_f^p) = \mathbf{\varepsilon}_f; \ \mathbf{\varepsilon}'(t_f^p) = \mathbf{\varepsilon}_f'$$

and restrictions are not the only ones. The method developed by the authors for synthesizing the SC guidance law under RM with the specified boundary conditions is based on the necessary and sufficient condition for the solvability of Darboux problem. Here, quaternion $\mathbf{\Lambda}(t)$ is presented as the result of the addition of three simultaneous rotations of "nested" bases $\mathbf{E}_k$ around unit vectors $\mathbf{e}_k$ of Euler axes as $\mathbf{\Lambda}(t) = \mathbf{\Lambda}_i \circ \mathbf{\Lambda}_1(t) \circ \mathbf{\Lambda}_2(t) \circ \mathbf{\Lambda}_3(t)$.

Figure 6 presents the results of planning an areal survey of the cities in the Volga region (Volgograd and Nizhny Novgorod) and North-Western region (Vologda and Arkhangelsk) in accessible overview band with a geodetic latitude from 45.5 to 66 degrees. In this band, surveying is carried out by one local group of three mini-satellites: left SC #1, orange scans; central SC #2, green scans and right SC #3, blue scans. Here, we demonstrate the free assignment of targets for each spacecraft, for example, imaging Volgograd with three satellites, and imaging the remaining three cities with separate spacecraft.

Fig. 7.   The GMC scheme and the region of variation of its AM

V. Digital Control Algorithms

The attitude and orbit control system (AOCS) of the mini-satellite uses a GMC with four GDs according to the well-known multiple 2-*SPE* scheme, Fig. 7. The AM vector of the p-th GD $\mathbf{h}_p(\beta_p) = h_g \mathbf{h}_p(\beta_p)$, $p = 1 \div 4$, has an unit vector $\mathbf{h}_p$ and its own angular momentum h_g is the same for all GD. The normalized AM vector $\mathbf{h}(\beta)$ of the GD cluster is calculated as $\mathbf{h}(\beta) = \sum \mathbf{h}_p(\beta_p)$ with $\beta = \{\beta_p\}$, and vector of the GMC control torque is formed according to relationships $\mathbf{M}^g = -\mathbf{H}' = -h_g \mathbf{A}_h(\beta) \mathbf{u}^g$, $\dot{\beta} = \mathbf{u}_k^g$ with digital control $\mathbf{u}_k^g$ and matrix $\mathbf{A}_h(\beta) = \partial \mathbf{h}(\beta) / \partial \beta$.

The SC angular guidance law is represented by the program values of quaternion $\mathbf{\Lambda}^p(t)$, vectors of angular velocity $\mathbf{\omega}^p(t)$ and acceleration $\mathbf{\varepsilon}^p(t)$. An error quaternion $\mathbf{E} = (e_0, \mathbf{e}) = \tilde{\mathbf{\Lambda}}^p \circ \mathbf{\Lambda}$ with a vector $\mathbf{e} = \{e_i\}$ corresponds to a

matrix $\mathbf{C}^e = \mathbf{I}_3 - 2[\mathbf{e}\times]\mathbf{Q}_e^t$ with matrix $\mathbf{Q}_e = \mathbf{I}_3 e_0 + [\mathbf{e}\times]$ and an orientation error vector $\delta\mathbf{\phi} = \{\delta\phi_i\} = 2e_0\mathbf{e}$. The column of the angular velocity error is calculated in the form $\delta\mathbf{\omega} = \{\delta\omega_i\} = \mathbf{\omega} - \mathbf{C}^e \mathbf{\omega}^p(t)$.

With discrete filtering of the mismatch vector $\mathbf{\epsilon} = -\delta\mathbf{\phi}$, the values of vector $\mathbf{\epsilon}_k^f$ are obtained at the time moments t_k, $k \in \mathrm{N}_0 = [0,1,2,..)$ with a period T_u, which are used in the recurrent discrete control law for the GD cluster

$$\mathbf{g}_{k+1} = \mathbf{B} \, \mathbf{g}_k + \mathbf{C} \, \mathbf{\epsilon}_k^f; \quad \tilde{\mathbf{m}}_k = \mathbf{K}(\, \mathbf{g}_k + \mathbf{P} \, \mathbf{\epsilon}_k^f\,);$$

$$\mathbf{M}_k^g = \mathbf{\omega}_k \times \mathbf{G}_k + \mathbf{J}(\mathbf{C}_k^e \mathbf{\varepsilon}_\kappa^p + [\mathbf{C}_k^e \mathbf{\omega}_k^p \times]\mathbf{\omega}_k + \tilde{\mathbf{m}}_k).$$

Here, $\mathbf{G}_k = \mathbf{J}\mathbf{\omega}_k + \mathbf{H}_k$ is the AM vector and in order to eliminate the problem of redundancy of the GD cluster, vector $\mathbf{M}_k^g$ of its control torque is formed using the explicit distribution of the GMC AM and is "recalculated" into a vector of digital commands for the GD angular velocities [20].

Fig. 8.   Scanning surveying of the Vologda surroundings

Fig. 9.   The SC angular guidance law when surveying Vologda

VI. Computer Simulation of the AOCS Operation

Simulation of the AOCS operation of a mini-satellite with $m = 350\,\mathrm{kg}$, an inertia tensor $\mathbf{J} = \mathrm{diag}(438, 240, 390)$ kgm^2 on a SSO with an height of 570 km was performed when surveying the surroundings of Vologda (Fig. 8) using gyrodines when their AM $h_g = 30$ Nms and digital control period $T_u = 1/8$ s.

The angular guidance law of mini-satellite #1 in the first local group (see Fig. 6) in such an areal scanning survey is presented in Fig. 9, and timing characteristics of scans in given directions are shown in Table III.

TABLE III

Temporal characteristics of scans when surveying Vologda			
Scan #	Start, s	End, s	Duration, s
1 ↑	916.6875	932.875	16.2
2 ↓	947.6250	988.563	40.9
3 ↑	973.4375	1004.750	31.3

Fig. 10. The AOCS errors and GD speeds when surveying Vologda

Fig. 11. The SR #2, the AOCS errors and GD speeds

The AOCS errors in angles and velocities, as well as the GD angular speeds, are shown in Fig. 10, where the time intervals of three scanning routes are highlighted in lilac, see Tab. III. Note that SR #2 is performed with reverse accumulation of the electronic image on the OEC matrices, when the module of the SC programmed angular velocity in pitch reaches a large constant value, see Fig. 9, red color. Details of this effect are presented in Fig. 11. When performing such routes, errors can be reduced by introducing astatism into the AOCS pitch channel.

VII. CONCLUSIONS

The problems of creating low-orbit constellations of mini-satellites for regional Earth observation with maintaining their sun-synchronous orbits over long time intervals using low-thrust electric propulsion units have been studied. Here, for the first time, available Earth overview bands are used with sets of local groups of mini-satellites

that can be reconfigured to monitor the development of regional emergencies (floods, fires) with a given interval of repeatability. Modifications of authors' control laws for mini-satellites when surveying specified territories have been developed.

The results of planning the survey of Russian cities and computer simulation of mini-satellite control processes are presented, which demonstrate the effectiveness of the developed digital algorithms.

REFERENCES

[1] S. Rodriguez-Donaire et al., "Earth observation technologies: low-end-market disruptive innovation," in Satellites Missions and Technologies for Geosciences. In-techOpen, 2020, ch. 7, pp. 1–15.

[2] V. Lappas, V. Kostopoulos, "A survey on small satellite technologies and space missions for geodetic applications," in Satellites Missions and Technologies for Geosciences. IntechOpen, 2020, ch. 8, pp. 1–22.

[3] N. Crisp, P. Roberts, F. Romano et al., "System modelling of very low earth orbit satellites for earth observation," Acta Astronaut., vol. 187, pp. 475–491, 2021.

[4] P.E. Elyasberg, "Introduction to theTheory of Flight of Artificial Satellites of the Earth," Moscow: Nauka, 1965. 540 p.

[5] G.V. Mozhaev, "Synthesis of Orbital Structures of Satellite Systems," Moscow: Mashinostroyeniye, 1989. 303 p.

[6] R. Battin, "An Introduction to the Mathematics and Methods of Astro-dynamics," New York: AIAA Education Series. 1999. 796 p.

[7] A.A. Baranov, "Maneuvering the Spacecraft in the Vicinity of a Circular Orbit," Moscow: Sputnik +. 2016. 512 p.

[8] D. Vallado, "Fundamentals of Astrodynamics and Applications," Hawthorne: Microcosm Press. 2020. 1135 p.

[9] H. Curtis, "Orbital Mechanics for Engineering Students," Butter-worth-Heinemann: Elsevier. 2020. 946 p.

[10] L. Wang and A.A. Baranov, "Optimal keeping of a spacecraft with low-thrust engines in a sun-synchronous orbit," Vestnik of Bauman MSTU, Mashinostroyeniye series, no. 2, pp. 68-83. 2015.

[11] S. Ulybyshev, "The use of sun-synchronous orbits for a real-time glo-bal coverage satellite," Cosmic Res, vol. 54, no. 6, pp. 445–451, 2016.

[12] Y. Razoumny, "Fundamentals of the route theory for satellite constel-lation design for earth discontinuous coverage," Acta Astronaut., vol. 128; 129, pp. 722–740, 741–758; 447–458, 459–465, 2016.

[13] O. Kichigina, "Optimization of conditions for the formation of low-thrust engines of a multiple sun-synchronous orbit with a stable altitude profile under control restrictions," Space Engineering and Technology," no. 2(13), pp. 80-89. 2016.

[14] D. Vinogradov and E. Davydov, "Methodology for forming the stable near-circular sun-synchronous orbits when long-term existence of spacecraft," Engineering Journal: Science and Innovation, no. 6 (66), pp. 1-22. 2017.

[15] K. Mikhailovsky and M. Gorodetsky, "Development of methodology for determining and correcting parameters of the working orbit of the Earth remote sensing satellite," RUDN Journal of Engineering Research, vol. 18, no. 3. pp. 361-372. 2017.

[16] S. Ulybyshev and A. Lysenko, "Design of satellite constellations for operational global monitoring with a daily repeat of flight track," Cosmic Res, vol. 57, no. 3, pp. 204–212, 2019.

[17] A. Chernov and G. Chernyavsky, "Orbits of Earth Remote Sensing Satellites," Moscow: Radio i Svyaz. 2004. 200 p.

[18] Ye. Somov, S. Butyrin, and T. Somova, "Autonomous guidance and control of mini-satellites in a low-orbit constellation during areal scanning survey," in: Proc. 2023 30th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS). 2023, pp. 1-4.

[19] Y. Somov, S. Butyrin, S. Somov, and T. Somova, "Coordinated gui-dance and control of mini-satellites in constellations for space Earth survey," Izvestiya of Samara Scientific Centre RAS, vol. 25, no. 2, pp. 88-96. 2023.

[20] V. Matrosov and Ye. Somov, "Nonlinear problems of spacecraft fault tolerant control systems," in Nonlinear Problems in Aviation and Aerospace, vol. 12: Advanced in Dynamics and Control, S. Sivasun-daram Ed., CRC Press / Taylor & Francis. 2004, pp. 309-331.

Testing the Methods for GNSS-Based Relative Navigation of Moving Objects

E.A. Kasulin

Moscow Institute of Physics and Technology, MIPT, Phystech

Kasulin.ea@phystech.edu

«RSC «Energia»»,

Korolev, Russia

Abstract— **The paper deals with a new approach to finding the relative vector between two objects with centimeter error using satellite navigation equipment. For this purpose, at large distances (more than 2 km) the algorithm on code measurements of pseudorange, smoothed by integral phase with the help of modified Hatch filter is used. At smaller distances, on the other hand, the phase method is used; the phase uncertainty disclosure (i.e., the search for the unknown vector of U0 constants) is performed using a dynamic filtering algorithm. New measurements are continuously added to the system, using simple recurrence formulae to refine the value of U0 as each measurement arrives. Over time, the accuracy of the knowledge of this vector reaches a necessary threshold, which allows for a centimeter positioning error. The algorithms were tested on real satellite data from NASA's GRACE mission, in which two satellites flying in the same orbit approached to a distance of 1500 meters.**

Keywords— *relative vector, centimeter accuracy, integral phase, Hatch filter, phase method, phase uncertainty, U0 vector, recurrence formulae, dynamic filtering, GRACE mission.*

I. INTRODUCTION

Creation of a centimeter-precise algorithm that can calculate the relative position vector of two objects relying only on satellite navigation equipment would significantly simplify the operation of both spacecraft and airplanes. For example, the currently used satellite navigation equipment algorithms have an error of about 2 m at 10-20 m distance [1], therefore, when docking with the ISS, a heavy (more than 100 kg) and expensive system "Kurs" is still used, calculating the relative vector by readings of complex radio-electronic equipment. In group flight missions (GRACE, TerraSAR, TandemX, etc.), where centimeter accuracy of spacecraft relative positioning is also required, laser range finders are used. However, the latter can only measure the distance between the vehicles along the line of sight. The full three-dimensional relative vector cannot be obtained with their help. Therefore, the subject of the research is the development and testing on real satellite data of a relative navigation algorithm that satisfies the following requirements:

1. Sub-meter accuracy (per 1 km range).

2. No need to use any navigation equipment, except for only two receivers (one for each device) and inter-board radio line between them (this, among other things, will allow to abandon the use of ILS and MLS in aviation and land aircraft automatically even at remote improvised airfields). It is only enough that the algorithm meets ICAO CAT III standards (60 cm vertically and 4 m horizontally) [2].

All algorithms have been tested both with the help of simulation modeling of the group flight of satellites and on real data from the GRACE space mission (NASA/DLR): it consisted of two satellites flying together, the telemetry of which is publicly available. It includes data of post-processing of antenna phase center coordinates with an accuracy of several mm and the so-called "raw measurements" of GPS receivers themselves: pseudorange and integral phases. This particular mission was chosen to test the proposed formulae, because on December 10, 2005 the satellites were approaching at a distance of 0.5 - 2 km, at which it is easy to check the accuracy of the algorithms (Fig. 6). This will be the first study of these formulae in the conditions of free space flight of two spacecraft (before this investigation, similar formulae were tested on the calculation of the distance between the antennas on the ISS).

The vectors of double differences of integral phases of the signals of all observed satellites were taken as measurement vectors. This allows us not to take into account in the solution the errors of navigation receivers' clocks and errors strongly correlating in place and time (errors of ephemeris and satellite clocks, ionospheric errors). Let us introduce the notations:

$$u_{ij} = [(Pr_{i1} - Pr_{i2}) - (Pr_{j1} - Pr_{j2})] \quad (1.1)$$

$$\vec{b}_{ij} = \frac{\vec{X}_i - \vec{X}}{D_{i1}} - \frac{\vec{X}_j - \vec{X}}{D_{j1}} \quad (1.2)$$

Therefore,

$$u_{ij} = \vec{b}_{ij}^T \Delta \vec{X} + \Delta nois_{ij} \quad (1.3)$$

This is a system of linear equations with respect to the unknown vector ΔX. The value of u_{ij} is measured, X is the coordinates of the receiver on the passive spacecraft, and the coordinates of all NSs are calculated from ephemeris data, so b_{ij} is a known value. $\Delta nois_{ij}$ is the unknown noise error. Assume that at a given second we have a common constellation of n satellites for the on-board and ground receiver. Let us compose the measurement vector U for the current second, the directional cosine matrix B and the error vector Δ:

$$\vec{U} = \begin{pmatrix} u_{12} \\ u_{23} \\ ... \\ u_{(n-1)n} \\ u_{n1} \end{pmatrix}, B = \begin{pmatrix} \vec{b}_{12}^T \\ \vec{b}_{23}^T \\ ... \\ \vec{b}_{(n-1)n}^T \\ \vec{b}_{n1}^T \end{pmatrix}, \vec{\Delta} = \begin{pmatrix} \Delta nois_{12} \\ \Delta nois_{23} \\ ... \\ \Delta nois_{(n-1)n} \\ \Delta nois_{n1} \end{pmatrix} \quad (1.4)$$

Then the initial equation (1.1) will take the form:

$$\vec{U} = B\Delta\vec{X} + \vec{\Delta} \qquad (1.5)$$

If the matrix B is nongenerated, the optimal solution in terms of the minimum of the error modulus has the form

$$\Delta\vec{X}^* = (B^T B)^{-1} B^T \vec{U} . \qquad (1.6)$$

In the initial equation (1.4) there is an unknown noise error, which can severely degrade the accuracy of the final answer. To smooth these noises, we apply the so-called Hatch filter:

$$Pr_{ik}^* = \frac{1}{T} Pr_{ik} + \frac{T-1}{T}[Pr_{i(k-1)}^* + (CP_{ik} - CP_{i(k-1)})] \quad (1.7)$$

Here: Pr_{ik} is the "raw" pseudorange value received by the receiver from the i-th satellite at the current time instant k; Pr_{ik*} is the filtered pseudorange value at the same time instant; $Pr_{i(k-1)*}$ is the same but at the previous time instant k-1; CP_{ik} is the integral phase from the i-th satellite at the current time instant k; $CP_{i(k-1)}$ is the same at the previous time instant; T is the filter constant equal to 300. At the initial instant of time, take $Pr_{i0*} = Pr_{i0}$.

If now in formula (1.1) we replace pseudoranges by integral phases, we can obtain a system of equations similar to (1.5), in which the vector of integer wavelengths $U0$ appears as a free unknown term (phase uncertainty is the key problem of all phase methods). Here we recall that the integral phase represents the measurement of the distance between the receiver and the satellite, expressed in clock units of the carrier frequency:

$$\overrightarrow{U_{CP}} - \vec{U}_0 = B\Delta\vec{X} \qquad (1.8)$$

By writing this equation for each second of the time range under study, we can obtain a system of such similar equations. For each j-th second from the range, we can write the following expressions (E_n is a unit matrix):

$$\overrightarrow{\Delta X}_j = (B_j^T B_j)^{-1} B_j^T (\overrightarrow{U_{CPj}} - \vec{U}_0) \qquad (1.9)$$

$$\overrightarrow{U_{CPj}} = B_j(B_j^T B_j)^{-1} B_j^T (\overrightarrow{U_{CPj}} - \vec{U}_0) \qquad (1.10)$$

$$[E_n - B_j(B_j^T B_j)^{-1} B_j^T]\overrightarrow{U_{CPj}} = [E_n - B_j(B_j^T B_j)^{-1} B_j^T]\overrightarrow{U}_0 \quad (1.11)$$

We denote the expression in square brackets in (1.11) by D_j. This matrix is idempotent ($D_j D_j = D_j$). Accordingly, we can introduce the notations:

$$\widehat{U}_{CP} = (D_1\overrightarrow{U_{CP1}}...D_m\overrightarrow{U_{CPm}})^T \qquad (1.12)$$

$$\widetilde{B} = (D_1...D_m)^T \qquad (1.13)$$

Then the original system of equations will take the form (1.14), and the optimal estimate of the unknown vector U0 will be expressed through (1.15):

$$\widehat{U}_{CP} = \widetilde{B}\vec{U}_0 \qquad (1.14)$$

$$\vec{U}_0^* = (\widetilde{B}^T \widetilde{B})^{-1} \widetilde{B}^T \widehat{U}_{CP} \qquad (1.15)$$

If we denote $W_m = \widetilde{B}^T \widetilde{B}$ and $U_m = \widetilde{B}^T \widehat{U}_{CP}$, we can obtain simple recurrence formulae (1.16-1.17) (at m=0 the initial matrices are zero):

$$W_{m+1} = W_m + D_{m+1} \qquad (1.16)$$

$$U_{m+1} = U_m + D_{m+1}\widehat{U}_{CP(m+1)} \qquad (1.17)$$

Having accumulated measurements for m seconds, we obtain the estimation

$$\vec{U}_0^* = W_m^{-1} U_m \qquad (1.18)$$

After that it remains to solve the system (1.8) similarly to the previous algorithm. In reality, the matrices Wm are also normalized at each step so that the computing machine retains the ability to quickly invert them. Either way, the recurrent addition of new dimensions to the system contributes to the gradual refinement of U0.

Over time, the accuracy of the knowledge of this vector reaches a necessary threshold, which is what allows a centimeter positioning error to be obtained. Since the composition of the observed constellation of navigation satellites will almost inevitably change during the observation process, we have additionally developed algorithms for modifying the observation matrix when satellites appear and disappear. The results of the algorithm are given below.

The proposed method of phase uncertainty disclosure differs favorably from the existing ones, such as the method of ambiguity functions, the method of primary and secondary satellites, direct search using geometrically free combinations. Unlike them, the algorithm presented in this paper does not require brute-force operations, has low computational complexity and simple computational formulae. Although it is sensitive to failures (phase shifts by a certain number of wavelengths, so-called cycle slip), they can be avoided by using a dual-frequency receiver. Also for this algorithm, methods of modifying the observation matrix when the composition of the constellation of observed satellites changes have been developed.

The conducted studies allow us to conclude: the 2nd algorithm (with phase measurements) demonstrated an error not more than 10 cm (transverse error not more than 5 cm) on real GRACE data, however, it requires significant time (at least 500 s) for the convergence of the solution. Therefore, in practice, the positioning process will be two-staged. For example, when calculating the approaching of two vehicles, first, while the distance between them still exceeds 2-3 km, code-based pseudorange measurements are used. But at the same time the phase algorithm is also started. As soon as the latter reaches the required convergence accuracy, the navigation equipment starts using it, leaving the first algorithm for accuracy control and protection against failures. Determination of convergence accuracy criteria is the subject of further research.

MODELING

1st algorithm (pseudoranges)
err, м

M = 0.726
D = 0.095

2nd algorithm (phases)
err, м

M = 0.083
D = 0.001

REAL DATA (GRACE)

1st algorithm (pseudoranges)
err, м

M = 0.369
D = 0.117

2nd algorithm (phases)
err, м

M = 0.049
D = 0.0001

REFERENCES

[1] Mikrin, E.A., Mikhailov, M.V., Rozhkov, S.N. et al. The results of a flight experiment aboard the International Space Station to study the effect of multipath on navigation, attitude control and rendezvous using measurements of satellite navigation equipment. Gyroscopy Navig. 3, 104–113 (2012). https://doi.org/10.1134/S2075108712020095.

[2] International Standards, Recommended Practices and Procedures for Air Navigation Services – Annex 10, ICAO, April 1985.

[3] GRACE Tellus: Gravity Recovery & Climate Experiment [Электронный ресурс]. URL: https://grace.jpl.nasa.gov/ (дата обращения: 27.02.2024)

[4] A.I. Zhodzishskiy, O.V. Nesterov, A.S. Bukin. Determination of the Relative Position of Objects by the First Phase Measurment Differences of One Epoch. Rocket-space device engineering and information systems Vol. 5, Iss. 1, 2018, pp. 2–11. DOI 10.30894/issn2409-0239.2018.5.1.3.12.

Attitude Coordination for Small Satellites Constellation

J. Fadeeva
Baltic State Technical University «VOENMEH»
Saint Petersburg, Russia
ORCID 0000-0002-2304-9806

I. Kostin
Baltic State Technical University «VOENMEH»
Saint Petersburg, Russia
ORCID 0000-0001-9459-1929

E.E. Vorobeva
Baltic State Technical University «VOENMEH»
Saint Petersburg, Russia

A.M. Popov
Baltic State Technical University «VOENMEH»
Saint Petersburg, Russia
ORCID 0000-0002-0732-9111

V.Yu. Emelyanov
Baltic State Technical University «VOENMEH»
Saint Petersburg, Russia
ORCID 0000-0001-5549-0906

Abstract — **This paper is devoted to the development of a sliding mode quaternionic controller that allows synchronous change of angular positions of satellites in formation. The proposed regulator does not contain a "pure relay", which guarantees the continuity of the control signal. The modeling for a group of satellites was carried out, which showed good efficiency and robustness of the developed algorithm both under interference conditions and when the moments of inertia of the satellites change.**

Keywords — satellites formation, attitude coordination spacecraft; multiagent control; quaternion feedback; consensus

I. INTRODUCTION

Today, satellites play an important role in many areas of human activity. Determining and controlling the orientation of a spacecraft is a major part of its mission accomplishment. To date, many spacecraft have been successfully launched and most of them have successfully accomplished their tasks. Many scientific papers have been published on the problems of orientation determination and control [1]. Of particular interest is the control of angular stabilization of satellites when they form a formation. In such a case communication, computational and energy resource constraints significantly complicate the task.

The aim of the paper is to develop and investigate a decentralized algorithm for controlling the orientation of a constellation of satellites based on quaternions and augmented with a consensus algorithm.

The angular stabilization of a satellite is based on the classical problem of solid body rotation in space. Currently, 4 approaches are used to solve the problem [1]. They use models based on

- Euler angles;
- modified Rodrigue parameters;
- quaternions;
- directional cosine matrices.

Control systems based on Euler angles have proven to be very efficient because linearized models with Euler angles are controllable, and all the standard design methods of linear control systems. The drawbacks associated with this approach include that the model may fail when the satellite position is far from the point at which linearization is performed, and for any sequence of rotations there is a singular point at which the model degenerates.

Control systems constructed by modified Rodrigue parameters cannot globally stabilize the system and also have a singular point. In turn, for models with quaternions and models based on matrices of directional cosines, controllers that allow for global stabilization of nonlinear spacecraft systems have been developed. These systems are independent of rotation sequences and do not contain a singular point.

II. ALGORITHM DEVELOPMENT

A. Problem statement.

The problem of angular stabilization of a group of satellites is posed as follows: there is a leader satellite, which moves in an elliptical orbit, but with a small eccentricity, the follower satellites form a certain formation together with the leader satellite [2-4]. It is necessary to realize synchronous change of angular positions of the satellites in the formation and further preservation of the given angular positions.

Initially, the inertial coordinate system XYZ, the origin of which is associated with the center of mass of the Earth, is used to describe the formation.

The radius vector $\vec{r}_c$ and true anomaly θ are used to specify the position of the leader satellite in orbit. The true anomaly θ is the angle between the direction to the pericenter of the orbit and the position vector of the leader satellite $\vec{r}_c$ with its vertex at the Earth's center of mass. Further, the vector $\vec{r}_c$ and angle θ are given by the absolute position of the origin of the movable non-inertial (relative) coordinate system *xyz*. The mobile coordinate system is

rigidly connected with the leader satellite. The x-axis is directed along the radius-vector $\vec{r}_c$, the y-axis lies in the orbit plane and is perpendicular to the radius-vector in the direction of motion, the z-axis completes the triple of vectors to the right. The absolute position of the follower satellite is given by the vector $\vec{r}$, and the position in the relative coordinate system is given by the vector $\vec{\rho}$.

The coordinate systems are shown in Figure 1.

Figure 1 – Coordinate system

Orientation changes occur in non-Euclidean space, causing difficulties in synthesizing the orientation control law, since the entire analysis must be performed on the corresponding set of the state space. One of the natural sets is the space formed by the special orthogonal group SO(3), whose elements are all orthogonal matrices with determinant 1. Several controllers on SO(3) have been proposed in the literature, notably [5-6], but as shown in [7], control on SO(3) can at best provide near global stability, i.e., only trajectories on the open set converge to equilibrium. In addition, the complexity of the operation in SO(3) can make it difficult to analyze efficiency and robustness.

The nonlinear equations of the spacecraft motion can be represented by a quaternion as follows [8]

$$\dot{q} = \frac{1}{2} q \otimes \begin{bmatrix} 0 \\ \omega \end{bmatrix},$$
$$J\dot{\omega} = -\omega \times J\omega + M + d,$$
(1)

where $q = (q_0, \vec{q})$ – singular quaternion, $\otimes$ denotes the product of quaternions, $\omega \in \mathbb{R}^3$ – satellite angular velocity, J – symmetric and positively defined inertia tensor, $M \in \mathbb{R}^3$ – control moment, $d \in \mathbb{R}^3$ – unknown disturbance.

B. Satellite control low

Let's introduce an error quaternion $q_e = q_d^* \otimes q$, where q_d the given quaternion, and q_d^* conjugate quaternion. The error dynamic equations will be in the form

$$\dot{q}_e = \frac{1}{2} q_e \otimes \begin{bmatrix} 0 \\ \omega_e \end{bmatrix}.$$
(2)

The control objective is to minimise the quaternion error signals and the deviation of the current angular velocity ω from the target velocity ω_d:

$$\vec{q}_e \to 0 \ \text{при} \ t \to \infty,$$
$$\omega_e \to 0 \ \text{при} \ t \to \infty,$$
(3)

Where $\omega_e = \omega - \omega_d \in \mathbb{R}^3$

Let's define the sliding variable s

$$s = \omega_e + \lambda \operatorname{sgn}_+(q_{e_0})\vec{q}_e,$$
(4)

where $\omega_e = \omega - \omega_d, \lambda > 0$, and

$$\operatorname{sgn}_+(\cdot) := \begin{cases} 1 & if \cdot \geq 0 \\ -1 & if \cdot < 0 \end{cases}$$
(5)

function (5) is necessary to achieve global stability of the system. Due to non-ideal operation of relay devices, runout occurs in sliding mode control systems. To reduce the impact of this factor the bandwidth is introduced Φ.

In [8], the following algorithm for controlling the angle of a satellite is proposed. For the system (1) with limited dynamics $|d| \leq D$, and the inertia tensor J, which is additive $J = \hat{J} + \tilde{J}$, where $\hat{J}$ nominal inertia tensor, and the uncertain part $|\tilde{J}_{i,j}| \leq G_{i,j}$ is bounded, the sliding variable s converges in finite time in the area $|s| \leq \Phi$, the control law will take the form:

$$M = \hat{J}\dot{\omega}_d + \omega \times \hat{J}\omega - \lambda \hat{J}\operatorname{sgn}_+(q_{e_0})\dot{\vec{q}}_e - K\operatorname{sat}\left(\frac{s}{\Phi}\right),$$
(6)

where **sat** – a saturation function, $\Phi \in \mathbb{R}^3 > 0$ – half of the bandwidth, which depends on the noise level. Robustness is achieved at coefficients of

$$k_i \geq |\omega \times \tilde{J}\omega|_i + (G|\dot{\omega}_d + \lambda \operatorname{sgn}(q_e)\dot{\vec{q}}_e|)_i + D + \eta_i,$$
(7)

where $\eta_i > 0$ and $(\cdot)_i$ denotes $i-th$ element of vector $(\cdot)$. In addition, the quaternion error converges exponentially to the region of a given quaternion. That is, the control objective is achievable.

C. Satellite formation control law based on consensus algorithm

The grouping of $n+1$ satellites is considered. The leader satellite moves freely along its trajectory around the Earth, and the follower satellites move along certain predetermined trajectories in relative coordinates. It is necessary to develop such a control law that the follower satellites align their angular position relative to the position set by the leader and converge to the same set angular velocity.

The connections between satellites can be described using graph theory [9]. To do this, an adjacency matrix is set where if there is a connection between i and j satellites, and 0 otherwise.

Figure 2 shows the various connection graphs for the 5 satellites in the formation. In the diagrams shown, node 5 is the satellite leader, and the rest are followers. Case a) is trivial, here each follower receives the required angular position directly from the leader and the problem is solved using algorithm (6) for a single satellite. Cases b) and c) are much more complicated. Some solutions for such problems were obtained in [9-11]. In this paper, a new algorithm for decentralized group control in sliding modes is proposed.

The algorithm is in the form

$$M_i = \omega_i \times \hat{J}_i\omega_i - K_i\operatorname{sat}\left(\frac{s_i}{\Phi_i}\right),$$
$$s_i = \frac{1}{N_i}\sum_{j=1}^{n+1} a_{ij}s_{ij},$$
(8)(7)

where $i = 1,...,n$ – satellite number, N_i – number of satellites that giving information to i satellite, s_{ij} – the equation of the sliding surface (4) for a pair of satellites ij.

Figure 2 - Switching graph for three cases (a, b, c)

III. SIMULATION

For examining the developed control systems dynamics, a series of simulations were carried out in the MATLAB/Simulink software framework. The system considered was (1) and the control law (8). Three situations were modeled corresponding to three different connection graphs shown in Figure 2. A variant of setting the connection graph with switches is presented in:

The initial angular positions of the satellites were set arbitrarily. The inertia tensors were different, interference in the input signals and external disturbances in the form of d_i

= [1   1 -1] were taken into account. Coefficients for the control law: $\lambda = 8, K = 150, \Phi = 0.01$. : The results are presented in Figures 3-5.

IV. CONCLUSION

A quaternionic sliding mode controller is derived in this paper, which provides exponentially convergent sliding surface dynamics for any given rotation. The proposed controller does not contain a "pure relay", which guarantees the continuity of the control signal. The application of the algorithm is considered for three cases in which the satellites converge to the same given angular velocity, levelling their position during the transition.

Group angular position control is implemented based on the consensus algorithm. It is shown that the introduction of the consensus algorithm improves the accuracy of angular stabilization of the satellite constellation.

Computational experiments have shown good efficiency and robustness of the developed algorithm both under interference conditions and when the moments of inertia of the satellites change.

The proposed method can be used for satellite formation control in cases where satellite orientation is important, for example, during antenna pointing, in synchronous turning tasks. This approach requires a minimum of computations during the turn, and also avoids turning along a suboptimal trajectory due to interference.

Figure 3 – Simulation result for the case a)

Figure 4 – Simulation result for the case b)

Figure 5 – Simulation result for the case c)

REFERENCES

[1] F. Landis Markley, John L. Crassidis. Fundamentals of Spacecraft Attitude Determination and Control. Springer New York, NY. 2014.

[2] K. Alfriend, .S. Vadali, P. Gurfil. Spacecraft Formation Flying. Butterworth-Heinemann: Oxford, 2010.

[3] Andrievsky, B.; Kuznetsov, N.; Popov, A. Algorithms for aerodynamic control of relative motion two satellites in a near circular orbit. Differencialnie Uravnenia i Protsesy Upravlenia 2020, pp. 28–58. (In Russian).

[4] A. M. Popov; I. Kostin; J. Fadeeva; B. Andrievsky. Development and Simulation of Motion Control System for Small Satellites Formation. Electronics 2021, 10, 3111. https://doi.org/10.3390/electronics10243111

[5] T. Lee, M. Leok, and N. H. McClamroch, "Geometric tracking control of a quadrotor uav on se (3)," in 49th IEEE conference on decision and control (CDC), pp. 5420–5425, IEEE, 2010.

[6] Y. Zou and Z. Meng, "Velocity-Free Leader–Follower Cooperative Attitude Tracking of Multiple Rigid Bodies on SO(3)," in IEEE Transactions on Cybernetics, vol. 49, no. 12, pp. 4078-4089, 2019.

[7] S. P. Bhat and D. S. Bernstein, A topological obstruction to continuous global stabilization of rotational motion and the unwinding phenomenon, Systems & Control Letters, vol. 39, no. 1, pp. 63–70, 2000.

[8] B. T. Lopez and J. -J. E. Slotine, "Sliding on Manifolds: Geometric Attitude Control with Quaternions," 2021 IEEE International Conference on Robotics and Automation (ICRA), Xi'an, China, 2021, pp. 11140-11146, doi: 10.1109/ICRA48506.2021.9561867.

[9] W. Ren; R.W. Beard. Distributed consensus in multi-vehicle cooperative control; Springer London, 2008.

[10] He Cai, Jie Huang. The leader-following attitude control of multiple rigid spacecraft systems. Automatica. Volume 50, Issue 4. 2014.

[11] Haichao Gui, Anton H.J. de Ruiter, Global finite-time attitude consensus of leader-following spacecraft systems based on distributed observers, Automatica, Volume 91, 2018.

[12] A. M. Popov, D. G. Kostrygin, P. V. Krashanin and A. A. Shevchik, "Development of Algorithm for Guiding the Swarm of Unmanned Aerial Vehicles," 2022 29th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), Saint Petersburg, Russian Federation, 2022, pp. 1-4, doi: 10.23919/ICINS51784.2022.9815370.

Dynamic Design of a Small-sized Spacecraft with a Passive Stabilization System

E.V. Barinova
Inter-University Department of Space Research
Samara National Research University
Samara, Russia
L5545@yandex.ru

N.A. Elisov
NIL-102
Samara National Research University
Samara, Russia
mr07th@gmail.com

I.V. Belokonov
Inter-University Department of Space Research
Samara National Research University
Samara, Russia
belokonov.iv@ssau.ru

I.A. Timbai
Inter-University Department of Space Research
Samara National Research University
Samara, Russia
timbai@mail.ru

Abstract—**This paper discusses the dynamic design of a small-sized CubeSat spacecraft with a passive stabilization system. We propose a comprehensive approach for the joint selection of design parameters, such as the center of pressure displacement relative to the center of mass, moments of inertia, and setting requirements for the angular motion control system to stabilize the uncontrollable motion of the spacecraft in the desired direction. This approach is applied in the development of nanosatellites at Samara University.**

Keywords—**small-sized CubeSat spacecraft, aerodynamic moment, gravitational moment, spatial angle of attack, angle of own rotation, passive stabilization system**

I. INTRODUCTION

Currently, small-sized CubeSat standard spacecraft are widely used not only for educational purposes but also for scientific purposes. As of January 1, 2024, according to the nanosats.eu website [1], 2323 such spacecraft have been launched. To conduct most scientific and technological experiments in space, a certain orientation of the spacecraft is required, for example, for the placement of Earth remote sensing equipment [2], Sun tracking equipment [3], etc.

In some CubeSats, reaction wheels and/or magnetic coils are used to maintain the necessary orientation [4, 5], which requires significant energy consumption and additional space for their placement. However, due to the very limited size and mass of CubeSats, it is advisable to consider the use of passive or combined (passive in combination with active [6, 7]) stabilization systems, as they do not require or require reduced consumption of working fluid and energy.

Since most spacecraft are launched into low Earth orbits of about 300-700 km, where gravitational and aerodynamic moments are predominant, it is advisable to use both moments to provide passive stabilization of motion relative to the center of mass. The choice of a specific type of passive stabilization depends on the orbit altitude, mission objectives, and spacecraft parameters.

A significant amount of research has been dedicated to studying the influence of aerodynamic and gravitational moments on the motion of a satellite relative to the center of mass, for example, [8, 9]. It is important to consider that for small-sized spacecraft, the angular acceleration caused by the aerodynamic moment is significantly higher than for larger satellites (with the same values of relative static stability margin and volumetric density) [10], so passive aerodynamic stabilization can be used over a wider range of altitudes. For example, in studies [11–13], the problem of providing aerodynamic stabilization for CubeSat nanosatellite is considered by orienting solar panels at a specific angle to its longitudinal axis after separation from the launch container. The initial angular velocity is then dampened using hysteresis rods or an active magnetic control system (using the B-dot control algorithm).

When designing a passive stabilization system, an important task is to determine the equilibrium positions of the spacecraft relative to the center of mass under the action of external forces and the nature of the spacecraft's motion in their vicinity. In this case, the parameters of the spacecraft should be chosen in such a way that the required orientation corresponds to a stable equilibrium position. Significant attention in the literature is paid to determining equilibrium positions of spacecraft relative to its center of mass. For example, in works [14–16], the dynamics of a satellite with a shape close to spherical is considered, with displacement of the pressure center relative to the center of mass in three coordinates and three unequal principal moments of inertia, when it is subjected to restoring aerodynamic and gravitational moments.

When designing CubeSat nanosatellite, it is necessary to take into account the fact that they have a rectangular parallelepiped shape. Therefore, unlike spacecraft with axisymmetric shapes, the aerodynamic moment depends not only on the spatial angle of attack but also on the angle of its proper rotation. It is important to consider the possibility of resonant modes of motion due to the spacecraft's shape and the presence of small inertial-mass asymmetry when using aerodynamic stabilization based on velocity vector. Such modes of motion have been studied by the authors in works [17, 18]. They manifest in a sharp change in oscillation amplitude with respect to the spatial angle of attack when a linear integer combination of the oscillation frequency of the spatial angle of attack and the average frequency of proper rotation

is close to zero, which can lead to unpredictable orientation. Recommendations for preventing the occurrence of resonances are provided in work [17].

In [19], based on a probabilistic formulation, the authors considered various systems for passive stabilization of small-sized spacecraft of the CubeSat standard using aerodynamic and gravitational moments. Analytical functions for the distribution of maximum angular deviations of the spacecraft axes from the required directions (orbital velocity vector and local vertical) were derived, along with formulas for selecting design parameters.

Dynamic design of the small-sized spacecraft with a passive stabilization system in this study refers to a comprehensive approach to jointly selecting the design parameters of the spacecraft (inertia tensor, center of pressure displacement relative to the center of mass) and the attitude control system, ensuring the transition of the spacecraft to an equilibrium position after separation from the launch vehicle, corresponding to the goals and objectives of the space mission.

This work considers the problem of selecting parameters for a small-sized spacecraft to ensure the desired equilibrium position and range of angles and angular velocities in the vicinity of the equilibrium position that should be provided by the preliminary active damping system. As an example, the SamSat-ION nanosatellite (NS), created at Samara University, is considered.

II. Stages of Dynamic Design of Small-Sized Spacecraft with Passive Stabilization System

The proposed comprehensive approach consists of the following stages.

1. An analysis of the spacecraft mission is conducted, including goals and objectives, characteristics of scientific instruments, parameters of the orbit, considering the spacecraft format, and the possible presence of deployable structures. Based on this analysis, a decision is made on the choice of the type of passive stabilization system and the need for using a preliminary active damping system, such as magnetic coils.

In work [19], the following classification of passive stabilization systems is proposed: single-axis aerodynamic stabilization system along the velocity vector (region where aerodynamic moment exceeds gravitational); three-axis aerodynamic-gravitational stabilization system (region where aerodynamic moment exceeds gravitational); single-axis and three-axis gravitational stabilization systems (region where gravitational moment exceeds aerodynamic); three-axis gravitational-aerodynamic stabilization system (regions with any ratio of aerodynamic and gravitational moments).

2. Following the chosen type of passive stabilization system of the spacecraft, in accordance with the mission objectives and characteristics of the scientific equipment, constraints on the maximum angles of deviation from the required directions that may occur during the spacecraft operation are determined. Using the analytical expressions obtained in [19], based on the chosen law of distribution of initial angular velocities in the vicinity of the desired equilibrium position, nomograms are constructed. These nomograms are used to determine preliminary values of design parameters for the spacecraft, such as geometric dimensions,

margin of static stability, and moments of inertia to ensure the chosen type of stabilization.

3. After this, the small-sized spacecraft is designed in a computer-aided design (CAD) system. During the design process, in accordance with the chosen orientation, a layout is developed. The aim of this layout is to ensure the mutual arrangement of the main components of the spacecraft in such a way as to meet the requirements of the design parameters. This includes considering both the technical requirements for the functionality of each element and the constraints associated with the physical parameters and volume of the spacecraft itself.

Afterwards, the CAD system is used to determine the spacecraft's mass centering and inertial characteristics (MCIC).

4. After completing the design of the small-sized spacecraft in CAD, detailed calculations are carried out using a complete model of the spacecraft's motion relative to its center of mass. In this case, a refined model of aerodynamic characteristics is used, which more fully considers the external geometry of the small spacecraft, including the presence of deployable structures. Calculations are performed under various initial conditions to ensure that the small-sized spacecraft will provide the required angular motion in space.

At this stage, adjustments may be made to the design as necessary. For instance, the spacecraft layout may be modified to meet the prescribed motion parameters relative to the center of mass, or provisions may be made for installing balancing weights.

Upon confirming the required conditions for angular motion through calculations, the final design of the spacecraft is agreed and manufacturing begins.

5. After the spacecraft is manufactured, a series of experiments is conducted to determine its MCIC on a specialized stand established at Samara University. This stand ensures the determination of the center of mass coordinates with an accuracy of no more than 0.5 mm and the determination of axial moments of inertia with an accuracy of no more than 1.5% [20, 21]. Since the existing stand allows for determining the MCIC of the spacecraft only for the transport configuration, i.e., with the deployable structures folded, a recalculation is then performed for the flight configuration using the methodology described in [22].

The obtained characteristics are compared with those calculated in the CAD system, and if necessary, a decision is made regarding the installation of additional balancing weights. In this case, the CAD system is used to select the mass and optimal placement of these weights, considering the existing constraints. After the weights are installed, the MCIC of the spacecraft is determined experimentally again, considering the introduced changes.

6. Following this, the motion of the spacecraft is simulated using the experimentally determined MCIC, taking into account the uncertainty in parameter determination under various initial motion conditions in the vicinity of the equilibrium position.

7. Subsequently, target objectives are formulated for the preliminary active attitude control system in terms of angles and angular velocities to transition the spacecraft to the designated equilibrium position.

III. DYNAMIC DESIGN OF SAMSAT-ION NS

The described approach is employed in the development of scientific and educational CubeSats at Samara University. As an example illustrating the proposed dynamic design technology, the SamSat-ION NS is considered, which was launched into orbit on June 27, 2023 [23]. Its objective is to study the parameters of the upper ionosphere, plasma state, and Earth's magnetic field along an orbit at an altitude of 558 km.

According to the proposed technology, taking into account the altitude of the flight and the scientific objectives of the mission, the spacecraft parameters were chosen to ensure a passive gravitational-aerodynamic stabilization system for orienting the sensitive element of the plasma parameter sensor perpendicular to the direction of the oncoming flow.

After choosing the gravitational-aerodynamic stabilization system, a series of nomograms were constructed based on which the design parameter regions of the SamSat-ION nanosatellite were determined.

When selecting the layout and designing SamSat-ION in the CAD system, the main challenge was to maximize the difference between the transverse axial moments of inertia, as this spacecraft has a 3U form factor, making it quite difficult to manufacture it dynamically asymmetric. Therefore, it was proposed to install a balancing weight in the form of a cylinder on one of the antennas, which allowed increasing the required difference in moments of inertia.

Since the SamSat-ION NS has deployable structures: four rotating telemetry antennas, a magnetometer located on a remote rod, and a plasma parameter sensor with a deployable mechanism (Fig. 1), its aerodynamic characteristics differ from those of a spacecraft without deployable structures.

Fig. 1.   SamSat-ION NS in the flight configuration

Using CAD software, dependencies were calculated for the projection area onto the plane perpendicular to the oncoming flow, as well as the coordinates of the center of pressure in the velocity coordinate system depending on the orientation angles. Figure 2 shows the dependence of the pro-

jection area of the SamSat-ION on the plane perpendicular to the velocity vector of the oncoming flow. This dependency is well approximated by the formula:

$$S(\alpha,\varphi)=0{,}017\left|\cos\alpha\right|+0{,}047\sin\alpha\left|\cos\varphi\right|+0{,}039\sin\alpha\left|\sin\varphi\right|,$$

where α is the spatial angle of attack, φ is the angle of proper rotation [19].

Fig. 2.   Projection area of the SamSat-ION NS onto a plane perpendicular to the oncoming flow

Based on these dependencies, a model of aerodynamic characteristics was developed, which was used in the motion simulation relative to the center of mass.

As a result of a series of calculations of motion relative to the center of mass using the MCIC calculated in the CAD and refined aerodynamic characteristics, the layout of the spacecraft was finally agreed.

After the SamSat-ION NS was manufactured, a series of experiments were carried out on a specialized stand to determine the MCIC of the NS in the transport configuration, and then they were recalculated for the flight configuration. The obtained characteristics, taking into account the uncertainty of their determination, satisfied the requirements for design parameters, so no additional balancing weights were required. Table 1 shows the experimentally determined MCIC of the flight model of the SamSat-ION NS.

TABLE 1. MCIC OF THE SAMSAT-ION NS IN FLIGHT CONFIGURATION

Characteristics	Values
Mass m, g	3602 ±0,4
Center of mass coordinate x_c, mm	173,8± 0,5
Center of mass coordinate y_c, mm	49,5 ± 0,5
Center of mass coordinate, z_c mm	54,6 ± 0,5
Axial moment of inertia J_x, kg·m²	0,01007± 0,00013
Axial moment of inertia J_y, kg·m²	0,05402 ± 0,00013
Axial moment of inertia J_z, kg·m²	0,05258 ± 0,00013
Products of inertia J_{xy}, kg·m²	-0,00015± 0,00020
Products of inertia J_{xz}, kg·m²	-0,00310 ± 0,00020
Products of inertia J_{yz}, kg·m²	0,00009 ± 0,00020

Then, calculations of the motion relative to the center of mass were carried out using experimentally determined MCIC and the developed aerodynamic moment model, taking into account the presence of deployable structures. The

simulation was carried out for various parameters of the NS within the error limits of the experimental determination of the MCIC. Multiple simulations of spatial motion for the formed equilibrium position regions in terms of angles and angular velocities, considering the action of aerodynamic and gravitational moments, showed that deviations of the spacecraft axes from the required directions, even in extreme cases of parameter distribution, ensure three-axis gravitational-aerodynamic stabilization when initial values of angles and angular velocities are consistent.

Thus, the dynamic design technology proposed in the study allows for the maximum consideration of the peculiarities of small-sized spacecraft behavior in low orbits and minimizes the required energy expenditure to maintain the spacecraft's operational orientation.

REFERENCES

[1] Nanosats Database. URL: https://www.nanosats.eu/ (accessed: 20.04.2024).

[2] Zharkikh, R.N., Purikov, A.V., Kuznetsov, V.V., Shubin, M.I., Svinareva, A.M. and Eletskaya, L.L., Construction of wide-purpose space complexes based on the SXC6 platform, in AIP Conference Proceedings, 2021, vol. 2318.

[3] Roscosmos News. Norby. URL: https://www.roscosmos.ru/39444/ (accessed: 20.04.2024).

[4] Hakima, H., Bazzocchi, M.C.F., and Emami, M.R., A deorbiter CubeSat for active orbital debris removal, Adv. Sp. Res., vol. 61, no. 9, pp. 2377–2392, 2018.

[5] Monkel, M., Montalvo, C., and Spencer E., Using only two magnetorquers to de-tumble a 2U CubeSAT, Adv. Sp. Res., vol. 62, no. 11, pp. 3086–3094, 2018.

[6] Chesi, S., Gong, Q., Romano, M., Satellite attitude control by center-of-mass shifting, Advances in the Astronautical Sciences, 2014, vol. 150, pp. 2575–2594.

[7] Chesi, S., Gong, Q., Romano, M., Aerodynamic Three-Axis Attitude Stabilization of a Spacecraft by Center-of-Mass Shifting, Journal of Guidance, Control, and Dynamics, 2017, vol. 40, no. 7, pp. 1613–1626.

[8] Beletskii, V.V., Motion of a satellite relative to the center of mass in a gravitational field, Ed. Moscow University. 1975.

[9] Sarychev, V.A., Ovchinnikov, M.Y., Dynamics of a satellite with a passive aerodynamic orientation system, Kosm. Issled, vol. 32, no. 6, pp. 561–575, 1994.

[10] Belokonov, I.V., Timbai, I.A., and Orazbaeva, U.M., Special features of low-altitude aerodynamically stabilized nanosatellite movement, Izv. vuzov. Priborostroenie, 2016, vol. 59, no. 6, pp. 507-512.

[11] Rawashdeh, S., Jones, D., Erb, D., Karam, A., Lumpp, J.E. Jr, Aerodynamic attitude stabilization for a ram-facing CubeSat, Advances in the Astronautical Sciences, 2009, vol. 133, pp. 583-595.

[12] Rawashdeh, S.A. and Lumpp, J.E., Aerodynamic stability for CubeSats at ISS orbit, Journal of Small Satellites, 2013, vol. 2, no. 1, pp. 85–104.

[13] Rawashdeh, S.A., Attitude analysis of small satellites using model-based simulation, International Journal of Aerospace Engineering, 2019, vol. 2019.

[14] Sarychev, V.A., Mirer, S.A., Degtyarev, A.A., Duarte, E., Investigation of equilibria of a satellite subjected to gravitational and aerodynamic torques, Celestial. Mech. Dyn. Astron., 2007, vol. 97, no. 4, pp. 267–287.

[15] Sarychev, V.A., Gutnik, S.A., Dynamics of an axisymmetric satellite under gravitational and aerodynamic torques, Cosmic Research, 2012, V. 50, no. 5. pp. 367–375.

[16] Sarychev, V. A., Gutnik, S.A., Satellite Dynamics under the Influence of Gravitational and Aerodynamic Torques. A Study of Stability of Equilibrium Positions, Cosmic Research, 2015, vol. 54, no. 5, pp. 388–398.

[17] Barinova, E.V., Belokonov, I.V., and Timbai, I.A., Preventing resonant motion modes for low-altitude CubeSat nanosatellites, Gyroscopy and Navigation, 2021, vol. 12, no. 4, pp. 350–362.

[18] Belokonov, I. V., Timbai, I. A., and Nikolaev, P. N., Analysis and Synthesis of Motion of Aerodynamically Stabilized Nanosatellites of the CubeSat Design, Gyroscopy and Navigation, 2018, Vol. 9, No. 4, pp. 287–300.

[19] Belokonov, I.V., Timbai, I.A., and Barinova, E.V., Design parameters selection for CubeSat nanosatellite with a passive stabilization system, Gyroscopy and Navigation, 2020, vol. 11, no. 2, pp. 149–161.

[20] Vasin, P.V., Barinova, E.V., Methodology for determining the error of a stand for measuring mass-centering and inertial characteristics of nanosatellites using reference objects, Vestnik of Samara University. Aerospace and Mechanical Engineering, vol. 21, no. 3, 2022.

[21] Belokonov, I.V., Barinova, E.V., Klyuchnik, V.N., Ivliev, A.V., Boltov, E.A., Technology and method for experimental determination of mass-centering and inertial characteristics of CUBESAT format nanosatellites, Space Engineering and Technology, vol. 34, no. 3, pp. 83–95, 2021.

[22] Vasin, P.V., Barinova, E.V. Algorithm for determining the mass-centering and inertial characteristics of a prototype nanosatellite SamSat-ION in flight configuration, Proceedings of the Federal State Unitary Enterprise "NPTsAP". Control systems and devices, no. 2, pp. 26–34, 2023.

[23] Roscosmos News. Universat. URL: http://www.roscosmos.ru/39352/ (accessed: 20.04.2024).

Attitude Determination and Control Algorithms to Support Optical Payload of an Earth Observation Nanosatellite

D.V. Pershin
BSTU VOENMEH,
Special Technology Center LLC
St. Petersburg, Russia
pershingg@bk.ru

V.I. Kulakova
Special Technology Center LLC
St. Petersburg, Russia
nika_kulakova@mail.ru

A.S. Lysenko
Special Technology Center LLC
St. Petersburg, Russia
alex.stc@list.ru

Abstract—**Attitude determination and control algorithms to support optical payload onboard nanosatellites are considered. Attitude information is obtained from fiber-optic gyroscopes and a star tracker, and the issue of accounting for time delays in the star tracker measurements is addressed through the use of the Kalman filter. To meet pointing requirements, the desired satellite attitude and vector of the required angular velocity for ground target tracking are implemented onboard. The stabilization law is synthesized using the Lyapunov function. Simulation results demonstrate the achievability of accuracy characteristics in terms of both attitude angles and angular velocity.**

Keywords—nanosatellite, CubeSat, Earth observation, fiber-optic gyroscope, star tracker, time delay, Lyapunov function, PID controller, reaction wheels.

I. INTRODUCTION

The problem of high-resolution Earth remote sensing solved by Earth observation satellites is relevant nowadays [1-4]. Since a satellite constellation is necessary for the real-time monitoring of any region of the planet, such factors as the time to develop and manufacture a satellite or the cost of its launching into orbit are significantly important. Such advantages are provided by nanosatellites of the CubeSat format. The Earth observation nanosatellite is based on the CubeSat 12U platform; it has dimensions of 30×20×20 cm and mass about 20 kg. The achievable resolution of the Earth images is 2.7 m and the estimation accuracy of the target position is 1000 m at an orbit altitude of 570 km.

To meet the performance requirements, the attitude determination and control subsystem (ADCS) on the satellite must provide an absolute pointing accuracy for the optical payload better than 0.1 ° and a relative pointing accuracy (pointing stability) better than 0.03 °/sec. The latter requirement is related to the limitation for the maximum image jitter during the exposure time (e.g. part of a second). In order to fulfill the pointing requirements imposed by the mission, a small-sized star tracker (STR) and a three-axis medium accuracy fiber-optic gyroscope (FOG) were chosen for attitude measurement. Note that it is reasonable to install the optical payload in a small spacecraft with rigid mounting. Thus, the rotation of the camera is carried out by the rotation of the entire satellite.

The paper is organized as follows. Section 2 presents the structure of the ADCS and discusses the problem of the STR placement with respect to the satellite axes. In section 3, an Extented Kalman filter (EKF) for the satellite attitude estimation is presented, with particular attention given to the compensation of a time delay in the STR measurements when calculating the combined measurements. The results of the attitude estimator performance during orbital motion and camera pointing mode are also presented. Section 4 describes the synthesis of a control law using the Lyapunov function and analyzes the simulation results of the control subsystem performance when tracking the desired angular position.

II. ADCS COMPONENTS

The basic parameters of the selected STR and FOG are given in Tables I and II.

TABLE I. STAR TRACKER PARAMETERS

Parameter	Value
Dimensions, mm	56×60×93
Mass, g	193
Noise (1σ), " in cross-boresight (σ_y ,σ_x) in boresight (σ_z)	10 70
Maximum angular velocity, °/s	3
Field of view, °	22

TABLE II. FOG PARAMETERS

Parameter	Value
Dimensions, mm	80×95×62,5
Mass, g	700
In-run bias stability over temperature -40°C to +60 °C (1σ), °/h	0,3
Scale factor error over temperature -40°C to +60°C (1σ), %	0,02
Random walk (Allan variance), °/√h	0,01

The satellite's location on orbit is determined using a GNSS receiver, which provides an accuracy of 30 m. A redundant cluster of four reaction wheels is implemented to control the satellite three-axis attitude [5-7]. The reaction wheels parameters are presented in Table III.

TABLE III. REACTION WHEEL PARAMETERS

Parameter	Value
Maximum torque, mNm	±3
Maximum momentum, mNms	±58
Maximum speed, rpm	10000

The developed satellite includes one STR, but it's installation directions (see Fig. 1) are optimized to avoid the

interference of the stray light from the Sun or Earth during the Earth observation mode. Furthermore, the interference to the STR is also reduced by optimizing the choice of the horizontal attitude angle for pointing mode. The following expression is used to transform the output quaternion of the STR $Q_s^i(t)$ to the satellite body frame:

$$Q_b^i(t) = Q_s^i(t) \circ \tilde{Q}_s^b,$$

where «~» denotes the quaternion conjugation, «∘» defines the quaternion multiplication, and Q_s^b is the quaternion which defines the orientation of the STR with respect to the satellite body frame.

Fig. 1. STR installation in the satellite

It should be noted that the ADCS on the satellite should support other modes of operation imposed by the mission, such as dumping the accumulated momentum of the reaction wheels, solar panels pointing to the Sun or pointing a high-speed communication channel antenna to a control ground station, initial detumbling (dumping the angular velocity) of the satellite, etc. To fulfill these requirements, the ADCS includes: MEMS 3-axis gyroscopes, 3-axis magnetometers, sun sensors, infrared nadir finding sensors and magnetic torquers [8-10]. However, the ADCS operation in these modes is not in the focus of this study.

III. ATTITUDE ESTIMATOR

The attitude is represented by the quaternion $L_b^i(t)$, which defines the orientation of the satellite body frame relative to an inertial frame, while the EKF estimates small attitude errors. Similar to works [11-16] the state vector of the EKF, which uses FOG and STR measurements, includes biases and scale factor errors of the gyroscopes. In addition, the proposed EKF takes into account the time delay in STR measurements, which can reach 0.2-0.3 s and cause severe degradation in the ADCS accuracy. Typically, the STR provides the measurement delay time τ_0 relative to the attitude quaternion formation time. The final delay time τ will additionally include a random delay $\Delta\tau$ required for data transmission and reception.

The methods to incorporate the delayed STR measurements in the attitude filters has been discussed in several recent studies [17-19]. In [17, 18], the STR attitude quaternion $Q_s^i(t)$ is propagated to the current filter time. In [19] the filter re-calculation method is implemented, when the last EKF states are buffered and after EKF update at the exact timing of the STR measurement the entire trajectory of states is re-calculated until the current step. However, high computational costs limit the real-time application of this method. In this study, we propose two computationally simple approaches to incorporate the delayed STR measurements while meeting the pointing accuracy requirements.

A. Accounting for a known measurement time delay

The time delay value τ_0 transmitted by the STR can be incorporated in the EKF during measurement formation. To do this, the quaternion $L_b^i(t)$ is taken with a fixed delay τ_0 when computing the correction quaternion $\delta L(t)$:

$$\delta L(t) = Q_b^i(t) \circ \tilde{L}_b^i(t - \tau_0).$$

Then the measurement equation is defined as:

$$Z = 2 \cdot sign(\delta L_0) \cdot [\delta L_1 \quad \delta L_2 \quad \delta L_3]^T,$$

where δL_k are the elements of the quaternion δL, $k = 0 - 3$.

B. Accounting for a random measurement time delay

The random delay can be estimated in the EKF by introducing a Wiener process $\Delta\dot{\tau} = w$ into the state vector, where w is a zero-mean Gaussian white-noise process with spectral density q_w^2. Then the state vector is defined as:

$$x = [\Lambda \quad \Delta\omega \quad \Delta M \quad \Delta\tau]^T,$$

where $\Lambda = [\Lambda_x \quad \Lambda_y \quad \Lambda_z]$ defines the attitude errors presented in inertial frame, $\Delta\omega = [\Delta\omega_x \quad \Delta\omega_y \quad \Delta\omega_z]$ and $\Delta M = [\Delta M_x \quad \Delta M_y \quad \Delta M_z]$ are the FOG biases and scale factor errors. The full estimate of the STR mesurement delay is calculated as $\hat{\tau} = \tau_0 + \Delta\hat{\tau}$.

Assuming that the STR measurement errors are uncorrelated Gaussian white-noise processes, with standard deviations specified in Table I for the boresight and cross-boresight axes, the diagonal elements of the measurement noise covariance matrix can be expressed as follows [17]:

$$r_{jj} = c_{j1}^2 \cdot \sigma_x^2 + c_{j2}^2 \cdot \sigma_y^2 + c_{j3}^2 \cdot \sigma_z^2,$$

where $j = 1, 2, 3$, and c_{j1}, c_{j2}, c_{j3} are the elements of the rotation matrix C_s^i from the STR frame to the inertial frame.

The observation matrix could be given as follows:

$$H = [I_{3\times3} \quad 0_{3\times6} \quad \overline{\omega}],$$

where $I_{3\times3}$ and $0_{3\times6}$ are the identity and zero matrices, $\overline{\omega}$ is the average value of the absolute angular velocity vector over the time interval $[t - \hat{\tau}, t]$.

It should be noted that if the STR and onboard computer are synchronized in time, and the STR provides the quaternion measurement time in a common timescale, the data delay could be directly computed and incorporated in EKF while calculating the correction quaternion $\delta L(t)$.

C. Simulation results

To verify the attitude estimator performance, a software program was developed to simulate the orbital motion of the satellite. The satellite position in its orbit was calculated using the SGP4 model, which utilizes a set of orbital elements stored in a TLE file. The orbit altitude was set at 540 km. During the simulation, the angular motion of the satellite was analyzed, taking into account the impact of external torque disturbances, including gravitational and aerodynamic forces. To implement the attitude control torque, a reaction wheel assembly was represented as a second-order system. The speed control loop of the reaction wheel was equipped with a PI controller to minimize overshoot.

The bias and scale factor errors of the gyroscopes were simulated as normal distributed random values with 0.3 °/hr and 0.02 % STD. To obtain the FOG bias instability and measurement noise, we recorded actual gyroscope readings at a frequency of 200 Hz under stationary conditions (as depicted in Fig. 2), while compensating for any constant offset. The satellite's attitude was then computed at 100 Hz.

Fig. 2.   FOG readings

The STR errors were represented as Gaussian white-noise processes with STD according to Table II. A fixed time delay of 0.1 s and a random measurement delay ranging from 0 to 0.1 s were introduced. The measurement update frequency was 5 Hz. The STR shading from the Earth was not simulated.

During the first 50 s of the simulation, the satellite was rotated to the desired spatial position with a maximum angular velocity of 2 °/s (refer to Fig. 3). Following this maneuver, the camera's optical axis was directed towards the target for tracking purposes. The satellite passed over the observed ground target at approximately 250 s, resulting in a peak angular velocity of 1 °/s. Figures 4 and 5 illustrate the attitude estimation errors for the simulated data, without and with consideration of the time delay in STR measurements, respectively. In the first scenario, the attitude errors significantly exceeded the required values and were proportional to the angular velocity of the satellite. However, in the second scenario, the accuracy requirements were met.

Figure 6 presents an illustrative example of the estimation error of the random measurement time delay $\delta\tau = \Delta\hat{\tau} - \Delta\tau$, depicted as a solid line. The dashed line represents the tripled value of the error STD derived from the covariance channel. The results indicate that the STR measurement delay observability is provided during satellite maneuvering, with an error of less than 2 ms (3 STD) in the steady state.

Fig. 3.   Rotational motion of the satellite during simulation

Fig. 4.   Attitude estimation errors without consideration of the time delay in STR measurements

Fig. 5.   Attitude estimation errors with consideration of the time delay in STR measurements

Fig. 6.   Estimation error of the random measurement time delay

IV.   SATELLITE ANGULAR MOTION CONTROL

The Earth observation mode consists of two steps: the satellite rotation to the desired spatial position and the fine pointing control mode for tracking the ground target. The satellite maneuver is carried out by a control law with trajectory planning when the trajectory is synthesized with a restriction on the maximum angular velocity of the satellite during rotation.

Note, that in order to point the optical payload in the required direction it is also necessary to know the satellite position in its orbit at each moment of calculating the torque control signal for the reaction wheel system. For this purpose, an algorithm for orbit propagation (motion prediction) has been implemented, which allows estimating the satellite position between the GNSS measurements or in the case of a short-term absence of GNSS data.

The Lyapunov direct method is employed to develop a control law that provides asymptotic stability of the closed-loop system [20,21]. This control technique enables stabilization of the satellite in the desired spatial orientation, even in the presence of unaccounted external disturbances and modeling errors. The control law is synthesized using a positive-definite function, known as the Lyapunov candidate function, which only equals zero when the satellite's motion aligns with the commanded trajectory.

The Euler's dynamic and Poisson's kinematic equations are used to describe the satellite motion:

$$J\dot{\omega} + \omega \times J\omega = M_u + M_e, \omega(t_0),$$
$$2\dot{L} = L \circ \omega, L(t_0),$$
(1)

where ω is the vector of absolute angular velocity, J is the moment of inertia tensor, L is the quaternion that determines the orientation of the satellite relative to the inertial frame, and M_u, M_e are the control and disturbance torques.

Since orientation and angular velocity of the satellite varies during pointing mode, a pointing reference frame is introduced for the synthesis of the control law. One axis of this reference frame is aligned with the line-of-sight to the ground object, and the other axis is fixed in space to optimize the STR position. The orientation of the pointing reference frame relative to the inertial frame is determined by the matrix C_r^i, while its angular velocity relative to the inertial frame is denoted as ω_r. To determine C_r^i and ω_r, information regarding the current velocity vector of the satellite and the satellite to target vector is utilized.

It can be shown that the relative pointing error, which determines the displacement of the image relative to the pixel during camera exposure time (several ms), is determined as follows [20]:

$$\Delta\omega = \omega - C_r^b \cdot \omega_r,$$
(2)

where C_r^b defines the orientation of the pointing frame relative to the body frame. Note that the optical payload frame coincides with the body frame. Thus, the stabilization system's objective is to ensure that $C_r^b = I_{3\times3}$, $|\Delta\omega| = 0$.

To synthesize the control law, we introduce the Lyapunov candidate function and its derivative as following:

$$\begin{cases} V = \dfrac{1}{2}(\Delta\omega, \ J \cdot \Delta\omega) + k_p \cdot (1 - q_0), \\ \dot{V} = (\Delta\omega, \ J \cdot \Delta\dot{\omega} + k_p \cdot q), \end{cases}$$

where $(\ ,\ )$ denotes the vectors scalar product, q_0, q are the scalar and vector parts of the quaternion corresponding to matrix C_r^b, k_p is a positive scalar constant.

The Lyapunov candidate function must satisfy the Barbashin-Krasovskii theorem [21], which requires its derivative to be nonpositive. This condition is satisfied with following relationship [20]:

$$J\Delta\dot{\omega} + k_p q = -k_d \ \Delta\omega, \ k_d > 0,$$
(3)

where k_d is a positive scalar constant.

Taking into account expressions (1-3), the control torque can be expressed as:

$$M_u = \omega \times J\omega - J[\Delta\omega \times C_r^b \omega_r] + JC_r^b \dot{\omega}_r - (M_e + k_d \Delta\omega + k_p q),$$

where the last two terms represent a PD-controller [22] with gains k_p, k_d to be properly selected.

Since the control of the satellite angular motion is implemented by a reaction wheel assembly, it is necessary to take into account the dynamic equation for the reaction wheels: $\dot{H} + \omega \times H = -M_u$, where H is a wheel angular momentum.

Figure 7 shows the attitude and angular velocities errors (difference between the commanded and obtained angles and velocities) with respect to the inertial frame when aligning the optical payload boresight axis with the satellite to target direction. The simulation parameters are the same as in section 3. The maximum attitude error is 0.03 °, while the relative pointing error is 0.001°/s. It should be noted that the relative pointing error reaches its maximum magnitude when the satellite passes over the observed ground target and the commanded angular velocity of the satellite is at its highest value (see Fig. 3).

Fig. 7. Attitude and angular velocities errors for pointing mode

CONCLUSIONS

This paper presents algorithms for high-precision attitude determination and control for an Earth observation nanosatellite, utilizing information from gyroscopes and a STR. The proposed approach addresses the challenge of accounting for and estimating time delays in the STR measurements. A method for synthesizing control of reaction wheels is applied, providing maximum stability in the presence of unaccounted external disturbances. The presented simulation results confirm that the chosen approach ensures the necessary accuracy of spatial orientation (<0.1°) and stabilization (<0.03 °/s) of the optical payload.

In future work, the influence of gyro axes non-orthogonality on attitude error will be analyzed. Additionally, a control law with motion prediction will be

considered to reduce the pointing error. Future plans also include evaluating of the impact of aerodynamics and elastic support elements (as solar panels) on angular motion dynamics.

REFERENCES

[1] A. Kosari, A. Sharifi, A. Ahmadi, M. Khoshsima, "Remote sensing satellite's attitude control system: rapid performance sizing for passive scan imaging mode". Aircraft Engineering and Aerospace Technology, 92(7), 1073-1083, June 2020. https://doi.org/10.1108/AEAT-02-2020-0030.

[2] A.P. Danilkin, "Small spacecraft-based space system for real-time earthsurface monitoring", Space technics and technologies, 32(1), pp. 42-55, 2021.

[3] A.N. Kirilin, "Main design characteristics of small scientific and applied-purpose spacecraft based on the Aist-2 unifed platform". Space technics and technologies, 31(4), pp. 5-20. 2020.

[4] V.S. Lobanov, N.V. Tarasenko, V.N. Zboroshenko, "The development directions of attitude control and stabilization systems for space vehicles of various purpose", 89(2), 18-29, 2015.

[5] B. Ibrahim, "Study and Analyze the Attitude Determination and Control Subsystem ADCS of a Small Remote Sensing Satellite", 2020. DOI:10.6084/m9.figshare.12389300

[6] A.P. Belenki, "Algorithm for controlling the system of four engines-flywheels of METEOR-M №2 series spacecraft. Questions of electromechanics". Proceedings of VNIIEM, 134(3), 9-14, 2013.

[7] B. Goeree, B. Shucker, "Geometric attitude control of a small satellite for ground tracking maneuvers". Aerospace and Mechanical Engineering Dept. University of Arizona, Tucson, 1999.

[8] T. Nguyen, K. Cahoy, A. Marinan, "Attitude determination for small satellites with infrared earth horizon sensors". Journal of Spacecraft and Rockets, 55(6), pp. 1466-1475, 2018.

[9] N. Ibrahim, "Attitude and orbit control of small satellites for autonomous terrestrial target tracking". Diss. University of Toronto, 2013. (Q).

[10] Wertz, James R, David F. Everett, and Jeffery J. Puschell, "Space Mission Engineering: The New Smad", Hawthorne, CA: Microcosm Press, 2011.

[11] J. L. Crassidis, F. L. Markley, A. M. Kyle, K. Blackman, "Attitude determination designs for the GOES spacecraft", In GOES-8 and Beyond (Vol. 2812, pp. 824-835), October 1996.

[12] P. Rozin, "Earth observation «AURIGA» small spacecraft attitude control and stabilization system", Proceedings of MAI works, 90, p. 11, 2016.

[13] G. I. Emelyantsev, A. P. Stepanov, "Integrated inertial-satellite systems of attitude and navigation", St. Petersburg CSRI Elektropribor, 2016.

[14] E. F.Solorzano, R. E. Zee, "Fiber Optic Gyro-Based Attitude Determination for High-Performance Target Tracking", 2016. https://www.researchgate.net/publication/313055511.

[15] W. H. Steyn, "An attitude control system for SumbandilaSat an Earth observation satellite", In ESA 4S Symposium, pp. 1-12, may 2008. https://www.researchgate.net/publication/258302605.

[16] F. Landis Markley, John L. Crassidis (auth.), "Fundamentals of Spacecraft Attitude Determination and Control", Springer-Verlag New York (2014).

[17] O. Kahraman, H. E. Soken, "Incorporating delayed star tracker measurements in Gokturk-2 satellite attitude filter", Advances in Space Research, 2023. http://dx.doi.org/10.1016/j.asr.2023.03.005.

[18] E.A. Hogan, B.A. Woo, "Treatment of measurement variance for star tracker-based attitude estimation", In: Astrodynamics Specialist Conference. Stevenson, USA, 2017.

[19] S. Gaisser et al., "Improvements in attitude determination and control of the small satellite flying laptop", In: Proceedings of the 33rd Annual AIAA / USU Conference on Small Satellites, Technical Session 3: Year in Review 1, SSC19-III-06, 2019, pp. 1–8. Available at: https://digitalcommons.usu.edu/smallsat/2019/all2019/275/.

[20] Y.V. Mashtakov, "Using the direct Lyapunov method in spacecraft orientation control tasks", Cand. Sci. Dissertation, Moscow 2019.

[21] G.N. Yakovenko, "A short course in analytical dynamics", Moscow, BINOM, 2004.

[22] B. R. Andrievsky, A. L. Fradkov, "Selected chapters of the theory of automatic control", St. Petersburg, Science, 2000.

Applying Procedures Used for Space Experiments Onboard Orbital Stations to Attitude Control of Geostationary Communication Satellites Yamal

D.N. Sevastyanov

JSC Gazprom Space Systems
Shchelkovo, Moscow region, Russia
info@gazprom-spacesystems.ru

Yu.R.Banit

JSC Gazprom Space Systems
Shchelkovo, Moscow region, Russia
info@gazprom-spacesystems.ru

M.Y. Belyaev

S.P. Korolev Rocket and Space
Corporation
Korolev, Moscow region, Russia
mikhail.belyaev@rsce.ru

Abstract — **November 24, 2003 saw insertion into orbit of geostationary communications satellites Yamal-201 and Yamal-202. One of them, Yamal-202, has been operating in orbit for more than 20 years now. Contributing to successful operation of the satellites in this series were results of scientific and applied research conducted during the missions of our country's orbital stations, as well as the results of certain technology tests carried out in the Russian Segment of the International Space Station (ISS), which celebrated its 25th anniversary in orbit in 2023.**

The paper presents the results of studies conducted onboard orbital stations, which were then applied to controlling the mission of the Yamal-series satellites and were conducive to their longevity in orbit.

Key words *- space experiments, orbital stations, International Space Station, dynamic properties, SC attitude, geostationary communications satellite*

I. Introduction

Interpreting the results of many space experiments (SE) requires knowledge of the instrumentation attitude position during scientific measurements. The attitude position of the scientific equipment (SE) with respect to the spacecraft (SC) body axes is usually known, therefore the problem reduces to determining the SC attitude in space. Used for determining SC attitude were measurements from magnetometers, sun and star trackers and other instruments. Such attitude sensors were also included in telemetry systems of orbital stations (OS) Salyut, orbital complex Mir, and the Russian Segment of the International Space Station.

SC angular motion is calculated using local methods for monitoring its attitude and integral statistical algorithms based on the math model of SC motion relative to its center of mass. Measurement data processing with the use of such algorithms usually not only recovers the SC angular motion, but also updates parameters included in the math model of motion used: SC tensor of inertia, parameters of the aerodynamic moment applied to the SC, etc.

Methods and practical algorithms developed for orbital stations made it possible to significantly enhance the capability of interpreting scientific measurements from research equipment during non-oriented segments of flight of orbital stations Salyut-4; Salyut-6, and Salyut-7. The developed procedure for updating math models (MM) of OS angular motion from telemetry data made it possible to build highly accurate MMs for controlling the attitude of the orbital complex Mir, which were

continuously used through the entire mission of that station. The procedure for updating tensor of inertia and MM of the SC angular motion from measurements was refined during the ISS mission within the framework of SEs Tenzor (Tensor) and Sreda-MKS (Environment-ISS). Methods and procedures developed for orbital stations turned out to be also useful for controlling the missions of geostationary communication satellites (GCS) Yamal and enabled successful operation of GCS Yamal-202 for more than 20 years in spite of the anomalies that arose in the satellite attitude control system.

To control attitude of satellites using their predicted angular motion, methods and software were quickly developed based on the experience from the experiments conducted onboard OS, which make it possible to update GCS tensor of inertia and parameters of applied disturbance models. This enabled predicting the angular motion of the satellite with an accuracy exceeding $0.3°$ over intervals of up to 4 hours of flight. In addition to this, a new method was developed to determine GCS attitude from measurements of onboard transponder system signals, transmitted from satellite communication ground stations of the users of frequency resource, which significantly improved the satellite control capability.

Owing to the quickly developed and implemented procedure, Yamal-202 SC has been successfully providing space communications services to users in Russia and abroad for more than 20 years, which is a great achievement in spaceflight.

II. Origins of Yamal communications satellites

November 24, 2023 marked the 20th anniversary of the day when a Proton launch vehicle launched two communications satellites Yamal-200 into geostationary orbit [1-3] (Fig. 1). The effort to develop Yamal satellites was put into motion when NPO Energia specialists took part in an international contest Columbus-500, which was announced in tribute to the upcoming 500th anniversary of America's discovery by C. Columbus. According to the conditions of the contest, the participants were supposed to develop a solar sail-driven space vehicle and have it win a race against other participants. The designers of the winning designs for such vehicles from each continent were to be granted a considerable amount of funds in order to implement their design.

Having learned about the contest, V.A. Koshlev, an employee of NPO Energia, shared this information with V.S.Syromyatnikov, the chief of his section, who had a second job as the chairman of the administering sub-department at the famous FEST department of MSFU (now the Space Department of the Mytishchi branch of the Bauman MSTU) which was originally established on the initiative of S.P. Korolev. V.S. Syromyatnikov called a meeting of the members of the administering sub-department staff, who were also employees of NPO Energia, where they made a decision to take part in that contest. Since only non-governmental organizations were allowed to participate in the competition, it was decided to file the application in the name of a youth center that had been set up in those Perestroika years under the auspices of NPO Energia. V.S. Syromyatnikov was appointed to be the technical manger of the project, and E.M. Belikov, V.N. Branets and V.A.Koshelev were appointed his deputies.

Fig. 1.   Launch of the Proton LV carrying the Yamal SC

Soon after that, in view of launching the work on the project, the participants established a small enterprise – Space Regatta consortium, to fill the position of general director of which, at V.N. Branets' suggestion, a young specialist of NPO Energia, a 1984 graduate of MIPT, N.N. Sevastianov was picked out. The consortium eventually produced a conceptual design of a spacecraft with a solar sail to be deployed using the SC rotation. In spite of the initial contest conditions, this project was not developed further, and the organizers did not even hold the promised 'races in space' under solar sails. Some of the ideas that formed the basis for the design of the SC with solar sails were later implemented in the space experiment Znamya-2. The solar reflector deployment assembly with 8 spools, n which web sectors were wound, as well as rotary actuators for deployment under the action centrifugal forces, was installed on the cargo transportation spacecraft Progress M-15 launched in October of 1992 towards Mir space station. In February of 1993, after the cargo spacecraft undocked from Mir, the experiment Znamya-2 was successfully carried out.

The work on the project made it possible to gain a unique experience in spacecraft design, which was subsequently applied to the development of Yamal satellites intended for providing communications with the northern gas production enterprises of our country. That proposal was supported by the management of JSC Gazprom and Gazprombank, as well as the management of NPO Energia, where the development work on the communications satellites was to be performed. To carry out the work to provide communications to enterprises in the Far North, in November 1992, the company

JSC Gazcom was established headed by general director N.N. Sevastianiv, who ran it till 2005. During the first phase of its activities, JSC Gazcom established a network of ground stations in the North, which covered the remote and hard-to-access regions of our country. Establishing such a network enabled Gazcom to accomplish some tasks aimed at modernizing JSC Gazprom production processes [1].

The next phase of the effort to provide communications to regions in the Far North was the development of the Yamal-100 satellite. On the initiative of N.N. Sevastiyanov JSC Gazprom issued to NPO Energia a statement of work to develop, build and launch two communications satellites Yamal-100. At the suggestion of Yu.P. Semenov, this work at NPO Energia was headed by N.N. Sevastianov as a deputy general designer. Proposed and implemented in the course of the satellite development were numerous innovative solutions, which allowed turning Yamal-100 into the baseline model for a whole series of subsequent satellites [2-4]. Some of the techniques used in the operation of the Yamal-series satellites were developed in the course of conducting experiments onboard orbital stations. In particular, the methods for determining and predicting the angular motion of orbital stations using telemetry measurements turned out to be handy for controlling the mission of Yamal-200 satellites whenever off-nominal situations occurred.

III. DETERMINING ANGULAR MOTION OF OS SALYUT AND MIR DURING EXPERIMENTS

Early orbital stations had no means of re-fuelling in orbit. Therefore, the ability to point scientific instrumentation toward targets was limited. In order to enhance the ability to observe the targets, additional passive spinning modes were proposed for the stations. The proposed station spinning was implemented about its long axis X, which was put into a proper spatial attitude. This provided the scanning of the celestial sphere with telescopes installed along the +Y axis of the station. The effect of gravitational and aerodynamic disturbing torques on station was degrading the spinning. In order to restore the spatial attitude of the telescope axes during celestial sphere scanning modes, special methods and associated software were promptly developed [5, 6]. In order to monitor the attitude, geomagnetic field intensity sensors and sun trackers were installed on stations of the Salyut series. When the station was flying in the daylight part of its orbit, its attitude was immediately computed at the rate of telemetry data reception from the sun tracker and magnetometer using a two-vector algorithm [5]. Moreover, the attitude computation results, beginning with the mission of Salyut-4, were immediately, at the rate of telemetry data reception, output to the displays at the Mission Control Center, which at the time was a significant achievement. Computing station attitude in the dark portion of the orbit where no data from sun tracker are available is a more challenging problem. Imposed on its solution were constraints that had to do with the need to solve all the problems in real time at the Mission Control Center located near Yevpatoriya (MCC-Y). At the time, MCC-Y was equipped with computer M-220. To solve the problem of calculating the station motion and attitude in real time using the available computer hardware at that time required developing new methods of describing disturbed motion of SC, integration method, methods and algorithms for attitude determination [5-8].

The integration method that was developed made it possible to speed up the calculations by an order of magnitude [8]. The analytical solution for the station disturbed motion that was found was more precise and at the same time more simple that known solutions [7]. To calculate station attitude in the dark portion of the orbit, algorithms were developed for a motion mode that was close to the regular precession, and for strongly disturbed OS motion. Those algorithms were quick and required a minimum amount of random access memory in the computer, since during their implementation the amount of computer memory did not depend on the amount of telemetry data that was being computed and coming at the rate of reception [5, 6]. All of this was for the first time was done and implemented at MCC-Y during the mission of Salyut-4.

A decision was made to hand over the mission control for stations Salyut-6 and Salyut-7 to MCC-M, which was originally established in the town of Korolev to support the Apollo-Soyuz program. Used in the development of procedures for conducting and operationally supporting experiments onboard the third-generation stations was the experience of running experiments onboard Salyut-4 at MCC-Y in Yevpatoriya [6].

In the course of running a number of experiments onboard OS Salyut-6 and Salyut-7, the mode of gravitational attitude control of the station was actively used [9]. Determination of the angular position of OS Salyut - 6÷7 in passive modes used a math model of station motion that took into account gravitational and aerodynamic disturbances [9 - 11]. Angular positions of Salyut stations were being updated using measurements of stellar photometers included in their scientific equipment [5, 12]. When conducting studies onboard orbital complex (OC) Mir, the determination of the OC angular position from sun tracker and magnetometer data involved the use of star tracker Astro-1. The Optical Star Tracker (OST) Astro-1 was jointly developed by Carl Zeiss Jena, Space Research Institute (IKI) and NPO Energia. It was installed onboard Kvant-2 module of OC Mir. The OST was used for determining angular position of the OC relative to stars in inertial or orbital attitude control modes, as well as in passive motion with small ($\leq$0.5 deg/s) angular rate [13].

The tracker has three identical optical units $A_v(v = 1, 2, 3)$ installed on the same support structure that provide viewing of three areas of celestial sphere. The size of one area is 5.3° x 8°. The image of the area is formed on a cooled CCD matrix that has 520 x 580 picture elements. Electronic equipment of the tracker isolates in each image up to seven brightest stars, determines in the optical unit-fixed coordinates positions of the centers of brightness for the isolated stars and transmits results to the OC telemetry system.

After the systems have received coordinates of stars viewed by the instrument and determined their relative magnitudes, those stars are matched to stars from the catalogue that had been compiled specifically for working with OST. The catalog contains coordinates of 8749 stars with magnitudes $\lesssim 7^m$ and has the following properties: 1) if, when the line of sight of the optical unit A_v is pointing in a certain direction, four or more stars with magnitude no higher than 7^m, fall within its field of view, then no less than four of them that are the brightest are included in the

catalogue; 2) two adjacent stars that the instrument sees as a single star correspond in the catalogue to one star.

The matching is performed in the following manner. Let at some moment in time the optical unit Av detects no less than three stars. They are enumerated and for each pair of stars with numbers i and $j (i < j)$ the cosine d_{ij} of the angular distance between them is calculated. The process of matching begins with identifying a triplet of stars for reference. The catalog is searched through to find pairs of stars where the cosine of the angle between them differs from d_{ij} by no more than $5 \cdot 10^{-5}$. Out of these pairs a set M_{ij} is compiled. Based on comparison of pairs in sets M_{12}, M_{13} and M_{23}, stars 1, 2 and 3 are matched. Identification of the k-th star with k $\geq$ 4 is done by comparing cosines of its angular distances from the stars that have already been matched.

Matching at a certain point in time several sighted stars makes it possible to find the OC attitude at that point in time. The associated algorithms for determining attitude are called local algorithms and are described in [5], [14-16] and others.

Let us assume that through the use of OST at certain point in time unit vectors (v_{1m}, v_{2m}, v_{3m}) $(m = 1, 2, ..., n)$ of directions toward n stars were found in the structural coordinate system. The stars are identified, and in the base coordinate system (such as Greenwich) they correspond to the unit vectors (V_{1m}, V_{2m}, V_{3m}), calculated using data from the catalogue. Let us denote the matrix of transfer from the structural coordinate system to the base coordinate system as $A = \left\| a_{ij} \right\|_{i,j=1}^{3}$. Defining this matrix within the framework of the least squares method results in minimizing the expression

$$\psi = \sum_{m=1}^{n} \sum_{i=1}^{3} \left(V_{im} - \sum_{j=1}^{3} a_{ij} v_{jm} \right)^2 ,$$

in elements a_{ij} provided that the sought for matrix is orthogonal and has positive determinant.

The matrix A is constructed as follows. Let us compose the matrix

$$D = \left\| v_{i1} v_{j1} + v_{i2} v_{j2} + ... + v_{in} v_{jn} \right\|_{i,j=1}^{3},$$

and consider its singular decomposition $D = UQV^T$. Here, U and V are orthogonal matrices of order 3, $Q = diag (q_1, q_2, q_3)$, $q_1 \geq q_2 \geq q_3 \geq 0$. Let us suppose that $q_3 > 0$, that is, the matrix D is non-degenerate and $n \geq 2$. Then

$$A = V \, diag \, (1, 1, \det U \cdot \det V) U^T.$$

The local method does not permit finding the OC attitude at those points in time, when observations of two or more stars are not available, and determining its angular velocity. To solve this problem one could use an integral statistical procedure, within the framework of which dozens of measurements performed over a length of time approximately equal to the orbital period are processed together using integration of equations for OC motion relative to the center of mass. Used for calculating the OC orbital motion are point mass equations of motion in the Greenwich coordinate system [17]. Evaluated as a result of the processing are the OC

initial motion conditions and inertia tensor, as well as parameters that characterize the restoring aerodynamic torque acting on the complex. The procedure is close to procedure [11] for determining rotational motion of orbital stations Salyut-6 and Salyut-7 based on readings of the solar and magnetic sensors.

In order to update the SC inertia tensor one can also use measurements of angular momentum of gyrodynes, taken when keeping a stationary attitude in the inertial coordinate system [18, 19]. As a rule, measurements processing fails to determine all the components of the inertia tensor. The most that can be done is to find in the structural coordinate system its off-diagonal components and differences of its diagonal components. However, the knowledge of the said values is sufficient for making a prediction, because inertia tensor components enter into math models only through these values.

Two methods were developed for processing the actual measured data. In the first method the measurement data belong to one interval of maintaining the attitude, the inertia tensor is considered diagonal, two differences of diagonal components of this tensor are evaluated. The second method includes into the processing the data from measurements taken during several (two or three) time intervals with different station attitudes, off-diagonal components of the inertia tensor and differences of its diagonal components are evaluated. Usually, the processed intervals are 20 to 60 minutes long and contain dozens of measurements of each component of the total angular momentum of gyrodynes in the structural coordinate system. The first method of processing permits, as a rule, to evaluate with sufficient degree of accuracy (with an error of 5 – 10% against the a priory calculated value) only one of the differences of diagonal components of the station inertia tensor. The second method, in most cases, produces differences of diagonal components with an error of no more than several percent of their calculated values, and the estimates of the off-diagonal components give only the accurate order of magnitude when compared against the a priori calculated values. However, since off-diagonal components of the station inertia tensor are very small, their calculated values may contain large relative errors.

Out of external torques applied to the station, the ones that were taken into account were gravitational and aerodynamic restoring torques.

The procedure for updating inertia tensor was used for controlling the mission of OC Mir [20, 21], was refined in the course of the mission of the International Space Station ISS [22] and was used for controlling the missions of geostationary communications satellites Yamal [23-25]. The math models of motion about the center of mass for the geostationary satellites Yamal take into account, instead of the aerodynamic disturbing torque, the torque produced by the solar radiation pressure.

IV. CONTROLLING GEOSTATIONARY COMMUNICATIONS SATELLITES YAMAL-200

The project to develop communications satellites Yamal-201 and Yamal-202 was carried out by JSC Gazcom (since December 1, 2008 known as JSC Gazprom Space Systems, hereinafter referred to as GSS), which developed the payload, service control channel, ground infrastructure of the telecommunications system. Development of the satellite bus, its integration with the payload, as well as spacecraft commissioning in orbit were performed by S.P.Korolev Rocket and Space Corporation (RSC) Energia. Both satellites were placed in their respective orbital slots: Yamal-201 at 90^0E, Yamal-202 at 49^0E. Operation of the satellite Yamal-201 continued for more than 10 years, while Yamal-202 has been in operation for more than 20 years, but now in the orbital slot at 163.5^0E, where it was transferred in 2019, after a new geostationary communications satellite Yamal-601 was placed in the 49^0E slot. Yamal-202 has 18 C-band transponders. The expected active life in orbit for Yamal-200 satellite was 12.5 years.

The guidance, navigation and control system of satellites Yamal-201 and Yamal-202 is built around a strap-down inertial navigation system (SINS), uses flywheels for final control, and its primary instrument for determining the angular rate was ARVGM. In 2005, ARVGM devices failed in both satellites. Star sensors available onboard the satellites were intended for periodic corrections of SINS, and therefore were only activated twice a day (operational life of the star sensor was 10000 hours). To resolve the situation, a control mode was developed that used predictions of the flywheel angular momentum (Prediction Mode).

The principle of the attitude control method that was proposed for Yamal-201 and Yamal-202 SC can be described as follows: having determined on a certain time interval the law of variation of flywheel angular momentum, which provides rotation of the SC body axes coordinate system at the required orbital angular rate, we can assure the nominal SC attitude in the orbital coordinate system by controlling the flywheels according to the above law.

In order to control SC attitude using the proposed method one needs to solve four major problems:

1. Create predictive model for flywheel angular momentum variation which will be used for generating the flywheel control law.
2. Determine initial conditions for the model – initial values of SC angular momentum vector and SC angular rate vector.
3. Determine the law of variation of the total disturbing torque vector acting on the SC.
4. Implement the obtained flywheel control law onboard the SC.

The ability to control SC attitude using its angular motion prediction is based on the exact knowledge of torques applied to the SC.

Estimation of inertia tensor of Yamal-200 was performed both in the daylight portions of the orbit, and when the geostationary satellite was in the Earth shadow. In the first case, taken into account from among the external torques affecting the satellite were gravitational torque and solar radiation pressure torque. In the second case, only gravitational torque was taken into account.

Gravitational torque components were defined through formulas

$$M_{g1} = v\sum_{i=1}^{3} x_i(x_2 I_{3i} - x_3 I_{2i}), \ M_{g2} = v\sum_{i=1}^{3} x_i(x_3 I_{1i} - x_1 I_{3i}),$$

$$M_{g3} = v\sum_{i=1}^{3} x_i(x_1 I_{2i} - x_2 I_{1i}), \ v = \frac{3\mu_E}{(x_1^2 + x_2^2 + x_3^2)^{5/2}}.$$

Here x_i are components of the geocentric radius vector of the satellite center of mass, I_{ij} are components of the satellite inertia tensor in the structural coordinate system, $I_{ij} = I_{ji}$ ($i, j = 1, 2, 3$), μ_E is the gravitational parameter of Earth.

The torque produced by the solar radiation pressure is approximated by the following expressions

$$M_{s1} = k(p_2 e_{si} - p_3 e_{si}), \ M_{s2} = k(p_3 e_{si} - p_1 e_{si}), \ M_{s3} = k(p_1 e_{si} - p_2 e_{si}).$$

Here $k = c^{-1}(r_*/\Delta)^2$, c is light velocity, r_* is mean radius of Earth orbit; Δ is the distance from the SC to the Sun; p_i are constant coefficients; e_{si} are components of the unit vector Satellite–Sun. Derivation of these expressions was done under the assumption that on the evaluation intervals of satellite angular momentum affected by the above external torques:

$$H_i = \sum_{j=1}^{12} F_{ij}(t)\alpha j \qquad (i = 1, 2, 3) \qquad (1)$$

where $\alpha_i = H_i(t_0)$ are initial conditions, t_0 is the specified point in time, $\alpha_4 = I_{11}$, $\alpha_5 = I_{12}$, $\alpha_7 = I_{22}$, $\alpha_8 = I_{23}$, $\alpha_9 = I_{33}$, $\alpha_{10} = p_1$, $\alpha_{11} = p_2$, $\alpha_{12} = p_3$. Functions $F_{ij}(t)$ are defined by initial problems for linear differential equations containing no parameters.

The software of the flywheel system, at certain points in time t_n ($n = 1, 2, \ldots, N$) permits measuring $H_i(t_n)$ component of its intrinsic angular momentum. The accuracy of estimating the angular momentum of Yamal-200 flywheels is 0.1 Nms. Accordingly, let us denote measurement results for these values as $H_i^{(n)}$. If the said measurements fall within a length of time, for which the values of attitude quaternion orientation are available, then, using formulas (1), they can be processed by some statistical method to determine vector $\alpha = (\alpha_1, \alpha_2, \ldots \alpha_{12})^T$. Since the values of $H_i(t_n)$ calculated from formulas (1) depend on α linearly, for processing we will use least squares method. Within the framework of that method, the estimate of vector α is its value, which minimizes the function

$$\Phi(\alpha) = \sum_{n=1}^{N} \sum_{i=1}^{3} \left[H_i^{(n)} - \sum_{j=1}^{3} F_{ij}(t_n)\alpha_j \right]^2.$$

In order to implement the described approach to evaluating the satellite inertia tensor, one needs to have a method of restoring its actual motion from telemetry data. Used for this purpose is a method based on the approximation of the satellite rotational motion with splines.

An analysis of obtained results indicates that it is possible to use the described models and estimates of the satellite momentum of inertia obtained through their use for controlling the mission. Further development of an adequate math model of rotational motion was related to updating the model of disturbances caused by solar radiation pressure, as well as models of the law of variation of the total angular moment of the SC and the law of variation of the angular momentum of flywheels.

An approach that was similar to the one described above was also used to solve other problems in satellite attitude control. For example, during the phase when the satellites were being brought to their orbital slots, an estimate of control torques from thrust modules of the satellites increased the accuracy of constructing the algorithm for selection and firing of thrusters for orbital correction.

This technology was then used for controlling the attitude of the Yamal-200 series SC. In the course of the SC mission there were situations when certain sensors for determining the satellite attitude could not operate in a stable manner for some reason or other. For example, a situation is possible where some glitches in the star tracker unit occur due to high density of charged particles from solar flares in the orbit of the SC. The drawbacks of the direction finder that is used onboard the SC include occurrences of operational problems during thunderstorms, heavy rain, snowfall. It is also worth noting that the Earth sensor cannot measure the coordinates of the center of Earth when it is located in the vicinity of local midnight during passage through Earth shadow. To make sure that the required attitude of the Yamal-200 series SC is maintained when the abovementioned operational problems occur the Prediction Mode was developed and is still used for controlling the Yamal-200 series SC, which is based on prediction of the SC angular motion.

In that mode the satellite attitude control is carried out using predictions of the total angular momentum. Considering that calculations of the total angular momentum use components of the inertia tensor that change in the course of satellite operation, in order to improve accuracy using the described approach, periodically, during dark and daylit portions of the orbit when solar arrays do not rotate and thrusters are not firing, estimation of SC inertia tensor was performed.

Yamal SC operational experience shows that, as a rule, after 4 hours of controlling in the Prediction Mode, attitude error exceeded the maximum allowable value of 0.3°. For controlling in the vicinity of the local midnight, during the intervals of short-time unavailability of measurements from star trackers caused by difficulties identifying stars, and passages of dark portions of the orbit the said duration of 4 hours of flight is sufficient.

But sometimes the period of unavailability of measurements from star trackers lasted for more than 24 hours. This happened multiple times during periods of high solar activity, which resulted in higher density of charged particles stream, affecting CCD matrix of star sensors. For controlling in that case, the Corrective Prediction Mode was developed, which involved periodic correction of attitude using information from all available sensors, including data from the specially designed hardware and software complex (H&SC) for determining the SC attitude from measurements of signals from onboard repeater, transmitted from ground satellite communications stations of the users of the frequency resource (Fig. 2).

H&SC is an integrated system of hardware and software, which enables spacecraft attitude determination and control without using standard sensors of the guidance, navigation and control system.

H&SC solves the problem of controlling the attitude of the satellite using prediction of the variation of the SC total angular momentum (Prediction Mode).

Fig. 2.   H&SC architecture

Two control options were developed:

1) Controlling without attitude correction using data from external sources. This mode is used to perform evaluation of initial values for kinematic parameters and subsequent attitude control based on variation of flywheel angular momentum.

2) Controlling with regular corrections of attitude using readings from standard sensors of the Guidance, Navigation and Control System (star tracker, Earth sensor, Sun sensor) or the system for determining the attitude from the signals of onboard repeater.

In the first option, measurements from standard attitude sensors and flywheels are used for initial alignment of attitude and angular rate of the SC. Subsequent attitude controlling is performed by the onboard computer based on predictions of the SC angular momentum variation. In the process, the onboard computer controls variation of the angular rate of flywheels in such a manner as to make sure that the SC total angular momentum in absolute coordinates remains the same as was formed based on the results of initial alignment. Performed in the course of carrying out the above task are:

- determining external disturbing torques affecting the satellite;
- determining the SC initial attitude using optical sensors;
- determining the angular momentum of the SC;
- filtering angular momentum values;
- real-time simulation of SC angular momentum taking into account external disturbing torques;
- calculating the vector of the required controlling torque;
- transmitting the values of the required controlling torque to the onboard computer, which subsequently generates commands to flywheels aimed at correcting the SC attitude.

In the second SC attitude control option in the Prediction Mode the satellite attitude and angular rate are regularly corrected from the Mission Control Center (MCC) based on measurements of these values made by standard sensors in the Guidance, Navigation and Control System (whenever such measurements are available) or on estimated attitude angles and angular rates calculated after a proper processing of signals received from onboard repeater.

When the Earth sensor measurements were available, controlling was performed based on roll and pitch angles measured by that sensor, and yaw angle estimated on measurements of the flywheel angular momentum variation and the Earth sensor measurements. Considering the fact that the Earth sensor remains operational throughout the entire flight, with the exception of some portions of orbit in the vicinity of the local midnight, such attitude correction was periodically performed during long intervals of controlling in the Adjustable Prediction Mode (APM).

A more reliable mode, which eventually became the primary mode, which was used continuously, with the exception of intervals of orbital correction, is the Adjustable Prediction Mode (APM), which uses regular correction from signals of ground stations for satellite communications located within the SC coverage area.

To implement the APM, the following tasks need to be performed:

- receiving signals from SC onboard repeater;
- scanning a specified frequency band to determine user signal levels;
- processing obtained values and calculating SC attitude angles;
- transmitting the values of the calculated attitude angles to the onboard computer, which subsequently generates commands to flywheels aimed at correcting the SC attitude and angular rate.
- Monitoring the SC attitude control.

The described attitude control method involves the use of:

- transmitting Ground Satellite Communications Stations (GSCS), selected according to criteria described below;
- repeater onboard the SC relaying signals from GSCS;
- receiving Central Ground Satellite Communications Station (CGSCS), which receives signals from onboard repeater;
- Ground Station of the Service Control Channel, intended for receiving from the satellite telemetry data required for calculations and transmitting controlling signals to the satellite;
- integrated software and hardware system intended for processing the signals received by CGSCS, and calculating SC kinematic parameters and control actions.

The telemetry data from the satellite received by the Ground Station of the Service Control Channel via control Work Station (WS) goes to the WS for determining kinematic parameters of the satellite (WS KP), where it is processed using specially developed software. In the process the signals are filtered, smoothed and interpolated/extrapolated in specified increments for certain points in time. Received via Central Ground Satellite Communications Station (CGSCS) are signals relayed by the onboard repeater, which are scanned by signal analyzer. After pre-processing these data are sent to WS KP, where they undergo additional processing (received signals are filtered, smoothed and interpolated/extrapolated in specified increments for certain points in time), satellite kinematic parameters are estimated, and control moments are calculated, which are

to be implemented by flywheels in order to maintain the required attitude.

For GSCS one could use regular stations of the users of the SC radio-frequency resource uplinking signals to the satellite. For Yamal-type SC such stations number in hundreds. Furthermore, no upgrading of the station or changing their characteristics is required. The only requirement consists in the absence of automatic gain control of the signal.

Selected in the initial phase are those transmitting GSCS of frequency resource users that are located as close as possible to the edge of directional pattern of the SC onboard repeater antennas. Signals from the users of the frequency resource are sent to the SC for their subsequent relaying. Installed at MCC is the CGSCS which receives relayed signals from transmitting GSCS stations. Signals received by CGSCS are sent to the Hardware and Software Complex, the main element of which is a spectrum analyzer. Operating according to a task that had been generated for it, H&SC selects out of the entire spectrum of signals relayed by the onboard repeater only those that had been selected during the preparatory phase and sends their characteristics to the server. At the server, the signals undergo the final processing, they are archived, and the power values for each selected signal are sent to automated work station for computing kinematic parameters, where the SC attitude and controlling torques needed to cancel deviations from the nominal attitude are calculated. The calculated attitude angles (or controlling torques) are sent to the automated WS for the SC control to be uplinked to the SC via the ground station of the Service Control Channel.

The satellite attitude is determined though analyzing variations in the power of the received signals. Key requirements for the signals from the transmitting stations are as follows:

- signal stability. Over long time intervals the signal shall not vary significantly when the satellite keeps nominal attitude, or its variation shall be predictable, and, therefore, can be described using some mathematical functions;
- signal variation caused by an attitude maneuver of the SC shall be described by a linear function (at least when its attitude changes within 0.5°, which is sufficient taking into account the pointing accuracy requirements);
- signals selected for evaluating roll angles shall vary to a larger degree during roll maneuvers and to a lesser degree during pitch and yaw maneuvers. Similarly, signals selected for evaluating pitch angles shall vary to a larger degree during pitch maneuvers and to a lesser degree during roll and yaw maneuvers, and so on.

To find those signals that meet the above requirements, the SC sequentially performed roll, pitch and yaw maneuvers.

Determined on the basis of maneuver results for the selected signals are functional relationships that describe their variation as a function of the satellite attitude angles:

$$\Delta P_i = a_{i1}\lambda_x + a_{i2}\lambda_y + a_{i3}\lambda_z \qquad (2)$$

where:

i – signal number, $i = 1, \dots ,N$; N is the number of GSCS stations involved, with N being equal to 20 to 30 in course of the flight.

a_{i1}, a_{i2}, a_{i3} are coefficients that characterize variation of the i-th signal when the SC performs yaw, roll, pitch maneuvers, respectively;

λ_x, λ_y, λ_z are angles of turn of the satellite about axes of yaw, roll, pitch, respectively.

In general, coefficients a_{i1}, a_{i2}, a_{i3} have different values in different directions and are functions of the angles of turn. Therefore, the first step toward the solution is to determine the most likely direction of turn for each of the axes. In this manner, the solution search domain is defined and coefficients of the system of equations (2) are determined, which, generally speaking, are functions of angles, but at small turn angles they can be considered constant.

In the matrix form, the system of equations describing signal variation with the satellite turning is as follows

$$A\Lambda = \Delta P,$$

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ \vdots & \vdots & \vdots \\ a_{N1} & a_{N2} & a_{N3} \end{pmatrix},$$

$$\Lambda = (\lambda_x, \lambda_y, \lambda_z),$$

$$\Delta P = (\Delta p_1, \Delta p_2, \dots \Delta p_N)^T, \qquad (3)$$

where:

A is the matrix of coefficients of the equation system; Λ is the sought after vector of the satellite turn angles, ΔP is the vector of the onboard repeater signal variation, elements of which are obtained from comparison of the current signal power value against its reference value determined at the start of the estimation process as the averaged power of each signal with satellite at nominal attitude (attitude angles are close to zero).

The sought after angles λ_x, λ_y, λ_z are defined in such a manner as to achieve the minimal root-mean-square error in equalities (2) or (3)

$$(A\Lambda - \Delta P)^T (A\Lambda - \Delta P)/N.$$

The following constraints are imposed on the solution of the problem:

- satellite attitude at the start of solving the problem of estimating the roll, pitch and yaw angles matches the nominal attitude;
- attitude control is implemented in such a manner as to make sure that the angles stay below 0.4°, which makes it possible to remain within the linear portion of the power vs. turn angle curve for the satellite, and thus improve accuracy when SC turns in positive and negative directions.

When determining the attitude from the onboard repeater signals, the SC orbital motion was taken into account.

The described solution for the satellite attitude control problem was implemented in both semi-automatic mode, that is, when calculations and uplinking of control signals to the satellite were performed from the WS of the MCC operator, and in automatic mode, when uplinked to

satellite were the calculated values for the satellite angular motion parameters, which were used by the onboard computer to calculate the control actions required to control the SC angular motion. The described control method was only used during passive phases of the flight. During corrective burns, standard attitude control sensors are used. This was mostly due to relatively longtime intervals between signal level measurements taken by the analyzers (10s).

V. SUMMARY

Operation of the ISS, the largest space vehicle in the history of spaceflight, required solving multiple problems related to in-flight updating its attitude and parameters: mass, moments of inertia, aerodynamic torques parameters, etc. For that purpose, engineering experiments "Tensor", "Izgib", "Environment-ISS" and others were set up onboard the ISS RS to conduct in-flight development of techniques for determining various characteristics and parameters of the ISS. This technology was then used for controlling the attitude of the Yamal-200 series SC. The use of the developed technology for controlling the attitude of the Yamal-202 SC based on predictions of its angular motion has enabled its continuing operation to this day, 20 years into its mission.

REFERENCES

[1] V.N. Branei "On the occasion of 20th anniversary of the launch of Russian telecommunications satellites Yamal-200", Space Engineering and Technology, no.4 (43), pp. 6-16, 2023.

[2] "Space Energia of Korolev. S.P.Korolev RSC Energia", Shans Publishers, p. 432, 2007.

[3] "S.P.Korolev Rocket and Space Corporation Energia in the first decade of the 21st century", S.P.Korolev RSC Energia, p.832, 2011.

[4] "S.P. Korolev Rocket and Space Corporation Energia at the turn of the century", 1996-2001.

[5] M.Y. Belyaev "Scientific Experiments Onboard Spacecraft and Orbital Stations", Mashinostroyenie, p. 264, 1984.

[6] M.Y. Belyaev "From the R-7 rocket and the first manned mission into space to the permanently manned orbital station", Gyroscopy and Navigation, vol. 29, no. 3 (114), pp. 96-121, 2021.

[7] M.Y. Belyaev "About one method of describing disturbed motion of the satellite", Proceedings of TsAGI, no. 6, pp. 48-54, 1974.

[8] M.Y. Belyaev, V.P. Semenko "About one method of numerical integration of differential equations", Space Research, no.2, pp. 300-301, 1976.

[9] M.Y. Belyaev, T.N. Tyan "Using gravity stabilization mode during conduct of experiments", Space Research, no.2, pp. 181-188, 1984.

[10] V.A. Sarychev, M.Yu. Belyaev, V.V. Sazonov, T.N. Tyan "Determining the motion of orbital complexes Salyut-6 and Salyut-7 with respect to the center of mass in gravity orientation from measurement data", Space Research, vol.23, issue 6, 1985.

[11] V.A. Sarychev, M.Yu. Belyaev, S.P. Kuzmin et al. "Determining the motion of orbital stations Salyut-6 and Salyut-7 with respect to the center of mass in the slow spinning mode from measurement data", Space Research, vol. 26, no. 3, pp. 390-405, 1988.

[12] M.Yu. Belyaev, V.A. Sarychev, V.V. Sazonov, N.I. Efimov, T.N. Tyan, E.K Sheffer., V.A. Sklyankin "Updating rotational motion of orbital station Salyut-7 from star photometer data", Space Research, vol. 27, issue. 4, pp. 528-544, 1989.

[13] M.Yu. Belyaev, N.I. Efimov, V.V. Sazonov "Determination of the attitude of the MIR orbital complex from indications of an optical star sensor", Space Research, vol. 33, no. 4, pp. 395-402, 1995.

[14] J.L. Farrel, J.C. Stuelpnagel et al. "A Least Squares Estimate of Satellite Attitude", SIAM Review, vol. 8, no. 3, pp. 384-386, 1966.

[15] V.V. Golubkov "Determining local attitude of spacecraft", Space Research, vol. 8, no. 6, pp. 811-822, 1970.

[16] M.Y. Katargin "Algorithm for root-mean-square estimation of spacecraft attitude and its errors", Space Research, vol. 24, no. 6, pp. 826-830, 1986.

[17] "Navigation support for the mission of orbital complex Salyut-6—Soyuz – Progress", Moscow, Nauka Publishers, p. 375, 1985.

[18] V.A. Sarychev, V.V. Sazonov, M.Yu. Belyaev, S.G. Zykov, E.Y. Chebukova "Estimating inertia tensor of the Mir orbital station from measurements of angular momentum of gyrodynes", Space Research, vol. 32, issue 4-5, pp. 22-42, 1994.

[19] V.V. Sazonov, M.Yu. Belyaev, S.G. Zykov "Study of the problem of estimating inertia tensor of the Mir orbital station from measurements of angular momentum of gyrodynes", Space Research, vol. 32, no. 3, pp. 3-16, 1994.

[20] M.Yu. Belyaev, S.G. Zykov, A.I. Manzhelei, D.N. Rulev, V.M. Stazhkov, V.P. Teslenko "Software for automated planning of research onboard orbital complex *Mir"*, Space Research, vol. 27, issue 1, pp. 126-134, 1989.

[21] V.V. Ryumin, M.Yu. Belyaev "Problems of control arised during the implementation of scientific research program onboard the multipurpose orbital station", Acta Astronautica, vol. 15, pp. 739-746, September 1987.

[22] Y.R. Banit, M.Yu. Belyaev, T.A. Dobrinskaya et al. "Determining inertia tensor of the International Space Station from telemetry data", Space Research, vol. 43, no. 2, pp. 135-146, 2005.

[23] V.V. Sevastianov, V.N. Branets, M.Yu. Belyaev, Y.R. Banit, V.R. Platonov, V.V. Sazonov "A study into the feasibility of controlling Yamal-200 SC using a math model of motion", Proceedings of the 14th international conference on integrated navigation systems, St.Petrsburg, pp. 196-203, 2007.

[24] Yu.R. Banit "A method of controlling the attitude of a geostationary spacecraft equipped with an onboard repeater", Patent for invention no.2761363, application no.2021106683, invention priority 15.03.2021, state registration date 07.12.2021, exclusivity period for the invention expires 15.03.2041, country – Russia.

[25] D.N. Sevastyanov, Yu.P. Banit, M.Yu. Belyaev "Updating dynamic characteristics of spacecraft in experiments onboard the ISS and applying this technique to controlling the mission of Yamal-200 SC", Space Engineering and Technology, no. 4 (43), pp. 100-114, 2023.

A Quality Evaluation Algorithm of Target Tracking Model Based on Observability

Zhong Wang
Bauman Moscow State Technical University
Moscow, Russia
wang.zhong.94@mail.ru

K.A. Neusypin
Bauman Moscow State Technical University
Moscow, Russia
neysipin@mail.ru

Bin He
Bauman Moscow State Technical University
Moscow, Russia
binhe93@njust.edu.cn

Hao Chen
Bauman Moscow State Technical University
Moscow, Russia
cheankh@student.bmstu.ru

Abstract—**This paper proposes a method for evaluating the quality of a target motion model based on the degree of observability (DoO) to enhance model matching accuracy in the presence of target motion model uncertainty. The Kalman filter-based DoO and its implementation are discussed, leading to the introduction of target motion model conformance criteria for quantitatively assessing the disparity between the current model and the actual target motion model. Simulation results indicate that DoO allows for a quantitative comparison of observability across different models, and the model conformance criteria based on DoO effectively indicate the degree of match between the current model and the real target motion model.**

Keywords—Maneuvering target tracking, Observability, Kalman filter

I. INTRODUCTION

Maneuvering target tracking has emerged as a prominent research focus in state estimation and information fusion[1]. It falls under the category of nonlinear filtering problems and holds significant importance in both national defense and in the field of civil applications[2]. To achieve accurate and stable tracking of maneuvering targets, it is crucial to develop a target motion model that accurately reflects the actual motion of the target.

Modeling the motion characteristics of a target involves describing the target's motion using motion equations. For non-maneuvering targets, linear models like constant velocity and constant acceleration can be used. However, as a target's maneuvering capability increases, its motion characteristics become non-linear. Using a linear motion model to describe this non-linear motion will result in significant errors, especially since targets often exhibit multiple maneuvering modes [3]. Therefore, accurately modeling the motion characteristics of a target is crucial in target tracking research. Various multi-model (MM) estimation methods, such as interactive multi-model (IMM) [4] and variable structure multi-model (VSMM) [5], have been developed to improve the precision and reliability of target tracking.

This paper delves into the characteristics of multi-model algorithms, with a specific focus on determining the weight of individual simple models within the multi-model framework. By establishing a motion model for the target in various motion states and employing the Kalman filter algorithm for tracking, a real-time quantitative evaluation method for assessing the quality of tracking models is developed by incorporating observability.

II. QUANTITATIVE EVALUATION METHOD FOR MOTION MODEL

A. The Degree of Observability

The concept of observability in a system refers to the ability to determine the initial state of the system based on measured values within a finite time[6], as first proposed by Kalman in 1960[7].

The analysis of observability in linear time-invariant systems is straightforward, while it becomes more complex for linear time-varying systems, requiring evaluation of the Gramian observability. However, the Gramian criterion cannot provide a complete indication of the observability of all state variables.

To better characterize the observability of each state variable in target tracking, the Degree of Observability (DoO) is introduced. This criterion helps identify which states are more accurately estimated and which are more challenging to estimate. Given that the estimability of states impacts filtering accuracy, the DoO can effectively reflect the influence of different abnormalities on filtering accuracy in real-time. In the context of a target tracking system, this concept can be mathematically expressed by the following equation:

$$
\begin{aligned}
x_k &= F_{k|k-1} x_{k-1} + G_{k-1} \omega_{k-1} \\
z_k &= H_k x_k + v_k
\end{aligned}
\tag{1}
$$

Where x_k is the vector of state, z_k is the vector of measurement, ω_k is the vector of input noise and v_k is the vector of measurement noise, the covariance matrices Q_k are and R_k, respectively, and ω_k is uncorrelated with v_k at any moment. $F_{k|k-1}$ is the state transition matrix, G_{k-1} is the system noise matrix, H_k is the measurement matrix.

Based on the above equations of state and observation, we analyze the observed data from time k to time k + n - 1, and can rewrite Eq. (1) as [6]

$$z_k = H_k x_k + v_k$$

$$z_{k+1} = H_{k+1} F_{k+1,k} x_k + H_{k+1} G_k \omega_k + v_{k+1}$$

$$\vdots$$

$$z_{k+n-1} = H_{k+n-1} F_{k+n-1,k+n-2} \cdots F_{k+1,k} x_k$$
$$+ H_{k+n-1} F_{k+n-1,k+n-2} \cdots F_{k+2,k+1} G_k \omega_k$$
$$+ \cdots + H_{k+n-1} G_{k+n-2} \omega_{k+n-2} + v_{k+n-1}. \qquad (2)$$

For the system represented by Eq. (2), the measured value z_k can be reformulated as

$$\mathbf{z}_k^* = \mathbf{O}_k x_k + \mathbf{v}_k^* \qquad (3)$$

where

$$\mathbf{z}_k^* = \begin{bmatrix} z_k \\ z_{k+1} \\ \cdots \\ z_{k+n-1} \end{bmatrix},$$

$$\mathbf{O}_k = \begin{bmatrix} H_k \\ H_{k+1} F_{k+1|k} \\ \cdots \\ H_{k+n-1} F_{k+n-1|k+n-2} \cdots F_{k+1|k} \end{bmatrix},$$

$$\mathbf{v}_k^* = \begin{bmatrix} v_k^+ \\ v_{k+1}^+ \\ \cdots \\ v_{k+n-1}^+ \end{bmatrix}$$

$$= \begin{bmatrix} v_k \\ H_{k+1} G_k w_k + v_{k+1} \\ \cdots \\ H_{k+n-1} F_{k+n-1|k+n-2} \cdots F_{k+2|k+1} G_k w_k + \cdots \\ + H_{k+n-1} G_{k+n-2} w_{k+n-2} + v_{k+n-1} \end{bmatrix}.$$

The matrix $\mathbf{O}_k$ is the observability matrix of the time-varying system. Eq. (3) can be reformulated as

$$x_k = \mathbf{O}_k^+ \mathbf{z}_k^* - \mathbf{O}_k^+ \mathbf{v}_k^* \qquad (4)$$

where $\mathbf{O}_k^+ = \left[\mathbf{O}_k^T \mathbf{O}_k \right]^{-1} \mathbf{O}_k^T$ denotes the generalized inverse matrix of $\mathbf{O}_k$.

Let $y_k = \mathbf{O}_k^+ \mathbf{z}_k^*$ and $\eta_k = \mathbf{O}_k^+ \mathbf{v}_k^*$, we can express them in a scalar form

$$y_k^i = a_{1,k}^i z_k + a_{2,k}^i z_{k+1} + \cdots + a_{n,k}^i z_{k+n-1}$$
$$\eta_k^i = a_{1,k}^i v_k^+ + a_{2,k}^i v_{k+1}^+ + \cdots + a_{n,k}^i v_{k+n-1}^+ \qquad (5)$$

where y_k^i is the i-th component of the vector y_k, η_k^i is the i-th component of the vector η_k, $a_{j,k}^i$ ($j=1,2,\cdots,n$) is the i-th line of $\mathbf{O}_k^+$.

Thus, the variance of the measurement noise η_k^i can be approximately described as

$$R_k^i = \left[\left(a_{1,k}^i \right)^2 + \left(a_{2,k}^i \right)^2 + \cdots + \left(a_{n,k}^i \right)^2 \right] R_k^0$$
$$= \sum_{j=1}^n \left(a_{j,k}^i \right)^2 R_k^0 \qquad (6)$$

Where R_k^0 is the initial variance of the interval measurement noise. Hence, we can obtain the DoO in scalar form as

$$D_k^i = \frac{E_k \left[\left(x^i \right)^2 \right] R_k^0}{E_k \left[\left(y^i \right)^2 \right] R_k^i} \qquad (7)$$

where

$$E_k \left[\left(x^i \right)^2 \right] = \frac{1}{n} \sum_{j=k}^{k+n-1} \left(x_j^i \right)^2$$

$$E_k \left[\left(y^i \right)^2 \right] = \frac{1}{n} \sum_{j=k}^{k+n-1} \left(y_j^i \right)^2$$

From Eq. (6), we know that the ratio of initial and current variances of observation noises is $\sum_{j=1}^n \left(a_{j,k}^i \right)^2$; thus, we can rewrite Equation (7) as [8-9]

$$D_k^i = \frac{E_k \left[\left(x^i \right)^2 \right]}{E_k \left[\left(y^i \right)^2 \right] \sum_{j=1}^n \left(a_{j,k}^i \right)^2} \qquad (8)$$

The defined Degree of Observability not only considers the observability criterion matrix, but also thoroughly evaluates the impact of dynamic errors in the estimation process and measurement noise on the system's reliability.

B. The model conformance criteria

To determine the weight of each simple model in the multi-model approach, it is essential to establish the motion model of the target in different motion states. The Kalman filter algorithm is utilized for tracking the moving target, and a real-time quantitative target tracking model quality evaluation method is developed by introducing the degree of observability criterion. The degree of observability incorporates the current measurement information, which highlights the disparity between the current model and the actual target motion model. When the target motion model conforms to the real model, the observability is large; Conversely, if the target motion model strays too far from the true motion model, the observability decreases. Leveraging this characteristic, and aiming to simplify the analysis of

model matching degree, this study establishes the model conformance criteria

$$\sigma_k^x = \frac{D_k^x}{D_k^x + D_k^y} \tag{9}$$

$$\sigma_k^y = \frac{D_k^y}{D_k^x + D_k^y} \tag{10}$$

$$\lambda_k = 1 - \left| \sigma_k^x - \sigma_k^y \right| \tag{11}$$

where D_k^x is the DoO of the state variable x at moment k and D_k^y is the DoO of the state variable y. The criteria specify the variation rule of state variables in the motion model over time. It states that for the same type of state variables, they exhibit the same variation rule. The model's conformance is high when it closely resembles the actual target motion model, and low when it deviates from it.

III. SIMULATION AND RESULTS

The Kalman filter algorithm is utilized to simulate the filtering process for single-model-based target tracking. Model 1 represents a constant velocity (CV) motion model, while model 2 represents a coordinated turn (CT) motion model. Each of these models is used to track the target independently. In the absence of prior information, the discrete state and measurement equations for the target are assumed to be represented as (1).

Taking the 2D case as an example, assuming that $x_k = \begin{bmatrix} x_k & \dot{x}_k & y_k & \dot{y}_k \end{bmatrix}^T$ is the state vector of the target at moment k and $z_k = \begin{bmatrix} x_k & y_k \end{bmatrix}^T$ is the observation vector, the state transfer matrices of the 2 models are

$$F_1 = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{12}$$

$$F_2 = \begin{bmatrix} 1 & \dfrac{\sin \omega T}{\omega} & 0 & \dfrac{\cos \omega T - 1}{\omega} \\ 0 & \cos \omega T & 0 & -\sin \omega T \\ 0 & \dfrac{1 - \cos \omega T}{\omega} & 1 & \dfrac{\sin \omega T}{\omega} \\ 0 & \sin \omega T & 0 & \cos \omega T \end{bmatrix} \tag{13}$$

$$H_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{14}$$

where T is the target adoption time interval, taking the value of 1s; ω is the target turning angle rate, taking the value of 1/30(rad°/s). The process noise covariance matrix of the CV and CT models is $Q^{CV} = G^{CV} diag(q_{CV1}, q_{CV2})$ $(G^{CV})^T$, $Q^{CT} = G^{CV} diag(q_{CT1}, q_{CT2}) (G^{CT})^T$

where

$$G^{CV} = G^{CT} = \begin{bmatrix} \dfrac{T^2}{2} & T & 0 & 0 \\ 0 & 0 & \dfrac{T^2}{2} & T \end{bmatrix}$$

$q_{CV1} = q_{CV2} = q_{CT1} = q_{CT2} = 0.01 m^2 \cdot s^{-4}$.

The measurement covariance matrix is $R = diag (100^2, 100^2)$. The initial state of the target radiation source is $x_0 = [80\ m, 400\ m/s, 80\ m, 600\ m/s]^T$.

First track the simple trajectory, in this paper we use linear trajectory and circular trajectory, use CV and CT model to track these two kinds of trajectory respectively, the conformity result is shown in Fig. 1 and 2, the tracking results from different motion models align closely with the model conformance criteria.

Fig. 1.   The conformance by tracking the linear trajectory.

Fig. 2.   The conformance by tracking the circular trajectory.

The motion trajectory comprises 220 points, divided into 4 segments: the first segment (points 1-20) involves uniform linear motion, the second segment (points 21-90) consists of a right turn maneuver, the third segment (points 91-160) involves a left turn maneuver, and the fourth segment (points 161-220) is uniform linear motion. This trajectory is illustrated in Fig. 3. Monte Carlo simulation experiment was repeated 100 times, and the model conformance curve is depicted in Fig. 4. The tracking results from different motion models align closely with the model conformance criteria.

Fig. 3.   True trajectory of the maneuvering target.

Fig. 4. The model conformance curve

IV. CONCLUSION

This paper presents a novel method for evaluating the quality of target motion models in maneuvering target tracking by leveraging the degree of observability (DoO). The proposed approach aims to enhance model matching accuracy in scenarios with uncertain target motion models. The key contribution of this study lies in utilizing DoO to quantitatively assess the disparity between the current model and the actual target motion model, and subsequently introducing a target motion model conformance algorithm based on this evaluation. To validate the efficacy of the method, a simulation experiment platform is constructed, and simulation experiments are conducted. The results confirm the effectiveness of the evaluation algorithm. Future research directions will focus on enhancing tracking accuracy in maneuvering target tracking systems through the exploration of multi-sensor fusion algorithms to bolster system robustness, as well as investigating the application of the target motion model quality evaluation method proposed in this paper to enhance the filtering accuracy of multi-model algorithms.

REFERENCES

[1] J. Ma and X. Guo, "Combination of IMM Algorithm and ASTRWCKF for Maneuvering Target Tracking," IEEE Access, vol. 8, pp. 143095-143103, 2020,

[2] Yin, Juqi, Zhen Yang, and Yazhong Luo. "Adaptive tracking method for non-cooperative continuously thrusting spacecraft." Aerospace 8.9: 244, 2021

[3] Na, Ki-In, Sunglok Choi, and Jong-Hwan Kim. "Adaptive target tracking with interacting heterogeneous motion models." IEEE Transactions on Intelligent Transportation Systems 23.11 (2022): 21301-21313.

[4] H. A. P. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with Markovian switching coefficients," IEEE Trans. Autom. Control, vol. AC-33, no. 8, pp. 780–783, Aug. 1988.

[5] X. R. Li and V. P. Jilkov, "Survey of maneuvering target tracking. Part V. Multiple-model methods," IEEE Trans. Aerosp. Electron. Syst., vol. 41, no. 4, pp. 1255–1321, Oct. 2005.

[6] K. Shen, T. Liu, Y. Li, N. Liu and W. Qi, "Robust Multi-Model Estimation for Reliable Relative Navigation Based on Observability and Abnormity Analysis," IEEE Transactions on Intelligent Transportation Systems, vol. 24, no. 5, pp. 5144-5158, May 2023

[7] R. E. Kalman, "A new approach to linear filtering and prediction problems," J. Basic Eng., vol. 82, no. 1, pp. 35–45, Mar. 1960.

[8] K. Shen, M. Wang, M. Fu, Y. Yang, and Z. Yin, "Observability analysis and adaptive information fusion for integrated navigation of unmanned ground vehicles," IEEE Trans. Ind. Electron., vol. 67, no. 9, pp. 7659–7668, Sep. 2019.

[9] Neusypin K. A., Selezneva M. S., Kai Shen. On Qualitative Characteristics of the State Variable Observability in Linear TimeVarying Models of Inertial Navigation Systems, Mekhatronika, Avtomatizatsiya, Upravlenie, 2018, vol. 19, no. 5, pp. 346—354.

Feedback Linearization and Adaptive Sliding Mode Control System for UAV

Bin He

Bauman Moscow State Technical University
Moscow, Russia
binhe93@njust.edu.cn

Zhong Wang

Bauman Moscow State Technical University
Moscow, Russia
wang.zhong.94@mail.ru

Neusypin K.A.

Bauman Moscow State Technical University
Moscow, Russia
neysipin@mail.ru

Mingming Zhang

Bauman Moscow State Technical University
Moscow, Russia
mz6250641@gmail.com

Abstract—**UAVs are typical complex nonlinear and parameter uncertain systems, and there is strong coupling between control channels, which poses a serious challenge to the design of flight control systems. In response to its nonlinear and coupling characteristics, the UAV dynamics model is decoupled into four independent linear time invariant systems using differential geometry theory by replacing the control signal input. Design attitude controllers for each channel separately using sliding mode control method. The simulation results show that the above linearization method can achieve 4-channel decoupling, and the designed control system meets the requirements of flight control performance.**

Keywords—Nonlinear systems, linearization, differential geometry theory, sliding mode control, control system

I. INTRODUCTION

In recent years, UAVs have received widespread attention due to their small size, low cost, high maneuverability, and wide adaptability to various scenarios. The flight control system [1] of a UAV is the center of the UAV, ensuring the completion of tasks and safe flight. It is also one of the key technologies in the development of UAVs. The dynamic model of UAVs has complex nonlinearity, time-varying parameters, and strong channel coupling, which poses a serious challenge to the design of flight control systems.

To design flight control systems, researchers generally used the theory of small disturbance [2] linearization to linearize the dynamics model of unmanned aerial vehicles in the early stages. When the system deviates from the equilibrium point, the system characteristics often no longer follow a linear variation relationship, and the linearized model is only an approximation, which will cause significant model errors and cannot meet the performance requirements of unmanned aerial vehicle flight [3]. The control system designed based on this linearization method is also difficult to achieve good control effects.

In terms of control system design, the PID design method [4] based on the classic proportional integral derivative controller has been widely used in the early days. However, with the continuous improvement of UAV control performance indicators, this method shows the following shortcomings: because PD parameters are designed based on the fixed controlled object, when there are time-varying parameters in the system, the PID parameters cannot be automatically adjusted according to the changes in system parameters, and therefore cannot guarantee accurate control. To improve the flight control performance of unmanned aerial vehicles, research on the application of some methods of modern control theory is gradually being carried out, such as the Linear Quadratic Regulator(LQR) method [5]. Meanwhile, thanks to the development of nonlinear control theory, researchers have explored its applications in the design of unmanned aerial vehicle flight control systems, such as backstepping control, sliding mode control, etc. However, backstepping control [6] and sliding mode control [7] suffer from differential expansion and chattering problems, which can be effectively improved by incorporating integral filters and integral sliding mode control. In recent years, some methods based on intelligent control [8] have also made progress in theoretical research, but they are rarely used in the design and application of unmanned aerial vehicle flight control systems. The reason is that intelligent control methods have complex algorithms, large computational load, and long execution time of onboard computers, which cannot guarantee the timeliness of control systems.

In summary, the linearization process of the model must be precise enough, and its dynamic process should be reflected in the controlled model. The control law should have strong robustness to time-varying system parameters and external disturbances, and should not be too complex, increasing the computational workload of onboard computers. Therefore, this article adopts differential geometry theory to linearize and decouple the dynamics model of unmanned aerial vehicles, obtaining four independent linear time invariant systems. The sliding mode control method is applied to complete the design of control laws for each channel. The control effectiveness of the control method was analyzed by comparing simulation results, and the effectiveness of the proposed control framework was verified.

II. UAV DYNAMICS MODEL

The quadrotor helicopter is assumed to be a rigid body, having 6 degrees of freedom and subject to external efforts. The model includes kinematics and dynamic equations. The dynamic equation of the UAV can be established as follows:

$$\begin{cases} \ddot{x} = [u_1(\cos\phi\sin\theta\cos\psi + \sin\phi\sin\psi)]/m; \\ \ddot{y} = [u_1(\cos\phi\sin\theta\sin\psi - \sin\phi\cos\psi)]/m; \\ \ddot{z} = [u_1\cos\phi\cos\theta - mg]/m; \\ \ddot{\phi} = [lu_2 + \dot{\theta}\dot{\psi}(I_y - I_z)]/I_x; \\ \ddot{\theta} = [lu_3 + \dot{\phi}\dot{\psi}(I_z - I_x)]/I_y; \\ \ddot{\psi} = [u_4 + \dot{\phi}\dot{\theta}(I_x - I_y)]/I_z, \end{cases} \tag{1}$$

In the formula, x, y, z represent the position of the center of mass of the UAV in the east north up inertial coordinate system; ϕ, θ, ψ represent the roll, pitching, and yaw angle; I_x, I_y, I_z represent the rotation inertia corresponding to the 3 axis of UAV. l represents the length from the center of the propeller to the origin of the body coordinate system; m represents the mass of the UAV. u_1, u_2, u_3, u_4 represent vertical, rolling, pitch and yaw motion control quantities; g is gravity acceleration.

Equation (1) can be rewritten as an affine nonlinear system:

$$\begin{cases} \dot{x} = f(x) + \sum_{i=1}^{4} g_i(x)u_i; \\ y = [h_1(x), h_2(x), h_3(x), h_4(x)]^T = [x, y, z, \psi]^T. \end{cases} \tag{2}$$

Where

$$x = [x, y, z, \dot{x}, \dot{y}, \dot{z}, \phi, \theta, \psi, \dot{\phi}, \dot{\theta}, \dot{\psi}, \zeta, \dot{\zeta}]^T = [x_1, \cdots, x_{14}]^T;$$

$$u = [\hat{u}_1, u_2, u_3, u_4]^T;$$

$$f(x) = \begin{bmatrix} x_4 \\ x_5 \\ x_6 \\ \dfrac{Cx_7 Sx_8 Cx_9 + Sx_7 Sx_9}{m}x_{13} \\ \dfrac{Cx_7 Sx_8 Sx_9 - Sx_7 Cx_9}{m}x_{13} \\ \dfrac{Cx_7 Cx_8}{m}x_{13} - g \\ x_{10} \\ x_{11} \\ x_{12} \\ k_1 x_{11} x_{12} \\ k_2 x_{10} x_{12} \\ k_3 x_{10} x_{11} \\ x_{14} \\ 0 \end{bmatrix}; \quad g_1(x) = \begin{bmatrix} 0_{13\times1} \\ 1 \end{bmatrix};$$

$$g_2(x) = \begin{bmatrix} 0_{9\times1} \\ k_4 \\ 0_{4\times1} \end{bmatrix}; \quad g_3(x) = \begin{bmatrix} 0_{10\times1} \\ k_5 \\ 0_{3\times1} \end{bmatrix}; \quad g_4(x) = \begin{bmatrix} 0_{11\times1} \\ k_6 \\ 0_{2\times1} \end{bmatrix};$$

$$k_1 = \frac{I_Y - I_Z}{I_X}; \quad k_2 = \frac{I_Z - I_X}{I_Y}; \quad k_3 = \frac{I_X - I_Y}{I_Z};$$

$$k_4 = l/I_X; \quad k_5 = l/I_Y; \quad k_6 = 1/I_Z.$$

III. Control Law Based on Feedback Linearization

The most natural approach for nonlinear systems is to linearize them, as there are abundant theories and mathematical tools available regarding linear systems. The classic linearization method is to make first-order linear approximations using Taylor expansion [9]. Although this method is simple and practical, its applicability is small, and the error increases with the increase of the action interval, always just an approximation. Feedback linearization method [10] is a set of precise linearization theories, which is a method based on differential geometry theory. If all input variables and all state variables can be linearized within the framework of feedback linearization, then the system is said to be able to achieve full state linearization; otherwise, it is considered partial state linearization. Its implementation methods mainly include coordinate transformation and state feedback. Therefore, feedback linearization methods may achieve full state or partial state linearization of the complete model, which provides ideas for solving problems.

A. The basic principle of feedback linearization

For the following affine nonlinear systems:

$$\begin{cases} \dot{x} = f(x) + \sum_{i=1}^{m} g_i(x)u_i \\ y_i = h_j(x), j = 1,2,\cdots,m \end{cases} \tag{3}$$

If $\exists x \in U$ and a series of non-negative integers $(r_1, r_2, \cdots, r_m)$, such that the following two conditions are met:

(1) $\forall x \in U$, $1 \leq i, j \leq m$ and $0 \leq k < r_i-1$,

$$L_{g_j} L_f^k h_i(x) = 0 \tag{4}$$

(2) $\forall x \in U$, Falb-Wolovich [11] matrix

$$\Gamma(x) = \begin{bmatrix} L_{g_1} L_f^{r_1-1}h_1(x) & \cdots & L_{g_m} L_f^{r_1-1}h_1(x) \\ L_{g_1} L_f^{r_2-1}h_2(x) & \cdots & L_{g_m} L_f^{r_2-1}h_2(x) \\ \cdots & \ddots & \cdots \\ L_{g_1} L_f^{r_m-1}h_m(x) & \cdots & L_{g_m} L_f^{r_m-1}h_m(x) \end{bmatrix}_{m\times m} \tag{5}$$

is non-singular. $L_g h(x) = \frac{\partial h(x)}{\partial x^T}g(x)$, represents the Lie derivative symbol.

So it is said that system (3) has relative orders r_1, r_2,$\cdots$,r_m. Consider the total relative order $r = \sum_{i=1}^{m} r_i$. If $r < n$, (where n is the system state dimension), the system can only be partially linearized and there exists a zero dynamic part. If $r = n$, then when the system satisfies:

a) Span $\Delta_i = span\begin{Bmatrix} g_1, \cdots, g_m, ad_f g_1, \cdots, \\ ad_f g_m, ad_f^i g_1, \cdots, ad_f^i g_m \end{Bmatrix}$ is non-singular, where $0 \leq i \leq n-1$;

b) $dim\Delta_{n-1} = n$;

c) $\Delta_i, 0 \leq i \leq n-2$ is involutive distribution. (Where, $ad_f g = \frac{\partial g}{\partial x^T}f - \frac{\partial f}{\partial x^T}g$; Involutive distribution refers to the fact that the rank sum of each component of the distribution itself is always equal to the rank of any two of its parentheses.)

When these three conditions are met, the system can be fully linearized.

For nonlinear systems (2), using the feedback linearization formula (4), obtain:

$$L_{g_j} L_f^k h_i(x) = 0, i = 1,2,3. \ j = 1,2,3,4 \ k = 0,1,2 \tag{6}$$

$$L_{g_j} h_4(x) = 0, j = 1,2,3,4 \tag{7}$$

And the Falb-Wolovich matrix

$$\Gamma(x) = \begin{bmatrix} L_{g_1}L_f^3h_1(x) & \dots & L_{g_4}L_f^3h_1(x) \\ L_{g_1}L_f^3h_2(x) & \cdots & L_{g_4}L_f^3h_2(x) \\ \dots & \ddots & \dots \\ L_{g_1}L_fh_4(x) & \cdots & L_{g_4}L_fh_4(x) \end{bmatrix}_{4\times4}$$

$$\begin{cases} \Delta_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \\ \Delta_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \end{cases}$$

According to feedback linearization theory, the nonlinear

$$= \begin{bmatrix} \frac{cosx_7cosx_9sinx_8+sinx_7sinx_9}{m} & \frac{k_4x_{13}(-cosx_9sinx_7sinx_8+cosx_7sinx_9)}{m} & \frac{k_5x_{13}cosx_7cosx_8cosx_9}{m} & \frac{k_6x_{13}(cosx_9sinx_7-cosx_7sinx_8sinx_9)}{m} \\ \frac{-cosx_9sinx_7+cosx_7sinx_8sinx_9}{m} & \frac{k_4x_{13}(-cosx_7cosx_9-sinx_7sinx_8sinx_9)}{m} & \frac{k_5x_{13}cosx_7cosx_8sinx_9}{m} & \frac{k_6x_{13}(cosx_7cosx_9sinx_8+sinx_7sinx_9)}{m} \\ \frac{cosx_7cosx_8}{m} & -\frac{k_4x_{13}cosx_8sinx_7}{m} & -\frac{k_5x_{13}cosx_7sinx_8}{m} & 0 \\ 0 & 0 & 0 & k_6 \end{bmatrix} \quad (8)$$

$$|\Gamma(x)| = \frac{k_4k_5k_6x_{13}^2cos(x_7)}{m^3} \quad (9)$$

$$L_{g_j}L_f^3h_i(x) \neq 0, \ i = 1,2,3. \ j = 1,2,3,4 \quad (10)$$

$$L_{g_j}L_fh_4(x) \neq 0. \ j = 1,2,3,4. \quad (11)$$

When $x_{13} \neq 0, x_7 \neq 0, |\Gamma(x)| \neq 0$.

So, the relative order of system (2) is $r_1 = r_2 = r_3 = 4$; $r_4 = 2$. The total relative order $r = \sum_{i=1}^4 r_i = 14 = n$. At the same time, it has been verified that the system (2) satisfies the three conditions mentioned earlier, so the system can be fully state feedback linearized.

Using Differential Homemorphism for Coordinate Transformation

$$z = \left[h_1(x), L_fh_1(x), \cdots, L_f^{r_1-1}h_1(x), \cdots, h_4(x), \cdots, L_f^{r_4-1}h_4(x)\right]^T_{14\times1} (12)$$

The original affine nonlinear system (2) is transformed into:

$$\begin{cases} \dot{z} = A_0z + B_0V \\ \xi = C_0z \end{cases} \quad (13)$$

Where

$$A_0 = \begin{bmatrix} 0_{4\times1} & \Delta_1 & 0_{4\times10} \\ 0_{4\times5} & \Delta_1 & 0_{4\times6} \\ 0_{4\times9} & \Delta_1 & 0_{4\times2} \\ 0_{2\times6} & 0_{2\times6} & \Delta_2 \end{bmatrix}_{14\times14}$$

$$B_0 = \begin{bmatrix} 0_{3\times1} & 0_{3\times1} & 0_{3\times2} \\ 1 & 0 & 0_{1\times2} \\ 0_{3\times1} & 0_{3\times1} & 0_{3\times2} \\ 0 & 1 & 0_{1\times2} \\ 0_{3\times1} & 0_{3\times1} & 0_{3\times2} \\ 0_{1\times2} & 1 & 0 \\ 0 & 0 & 0_{1\times2} \\ 0 & 0_{1\times2} & 1 \end{bmatrix}_{14\times4}$$

$$C_0 = \begin{bmatrix} 1 & 0_{1\times6} & 0_{1\times7} \\ 0_{1\times4} & 1 & 0_{1\times9} \\ 0_{1\times8} & 1 & 0_{1\times5} \\ 0_{1\times12} & 1 & 0 \end{bmatrix}_{4\times14}$$

relationship between the control law of the original system (2) and the system (13) can be obtained:

$$U = \Gamma(x)^{-1}(-b(x) + V) \quad (14)$$

Where

$$b(x) = \left[L_f^4h_1(x), L_f^4h_2(x), L_f^4h_3(x), L_f^2h_4(x)\right]^T \quad$$

B. *Controller design*

Based on feedback linearization, four independent channels are obtained. According to the control performance requirements of each channel, sliding mode control method is used to design controllers for four channels.

Specifically, for second-order systems:

$$\ddot{x} = v + d(t) \quad (15)$$

Among them, $d(t)$ is the external interference and satisfies $|d(t)| \leq D$. Error $e = x_d - x$. Design sliding mode function:

$$s = c_1e + \dot{e} \quad (16)$$

Taking sliding mode control law:

$$v = \ddot{x}_d + c_1\dot{e} + \eta_1 sgn(s) \quad (17)$$

Among them, sgn (.) is a sign function. The following constructs Lyapunov function to prove stability:

$$V = \frac{1}{2}s^2 > 0 \quad (18)$$

Seeking differentiation, obtaining:

$$\dot{V} = s\dot{s} = s\left(c_1\dot{e} + \ddot{x}_d - v - d(t)\right) = -\eta_1|s| - sd(t) \ (19)$$

When $\eta_1 \geq D, \dot{V} \leq 0$ satisfies the stability condition.

For the system:

$$x^{(4)} = v + d(t) \quad (20)$$

Design sliding mode function:

$$s = a_2e + b_2\dot{e} + c_2\ddot{e} + e^{(3)} \quad (21)$$

Taking sliding mode control law:

$$v = x_d^{(4)} + c_2e^{(3)} + b_2\ddot{e} + a_2\dot{e} + \eta_2 sgn(s) \quad (22)$$

When $\eta_2 \geq |d(t)|$, the stability condition is met.

Fig.1 The overall closed loop system

Finally, the closed loop system with the combined controller-observer-estimator is represented in figure (1).

IV. SIMULATION AND RESULTS

Simulation is carried out using the following quadrotor parameters: $I_x = 2.3 \times 10^{-3}$ kg·m^2, $I_y = 2.4 \times 10^{-3}$ kg·m^2, $I_z = 2.6 \times 10^{-3}$ kg·m^2, $m = 0.625$ kg, $l = 0.1275$ m, and $g = 9.8$ m/s^2. The initial state variable of the UAV is $x = 0_{4 \times 1}$, the expected state variable is $x_d = (50, 50, 50, \pi/6)$ and the simulation time is set to 50 s. And the sliding mode control is divided into two stages, the first stage (0-11s) is the approximation stage, and the second stage (11-50s) is the oscillation stage.

Fig. 2 The UAV location

Fig.3. The UAV attitude angle

V. CONCLUSION

Using differential geometry theory to linearize and decouple the dynamics model of unmanned aerial vehicles, three independent linear time invariant fourth-order systems and one independent linear time invariant second-order system were obtained. Sliding mode control method was applied to complete the design of control laws for each channel. The control effectiveness of the control method was analyzed by comparing simulation results, and the effectiveness of the proposed control framework was verified.

REFERENCES

[1] Kada, Belkacem, and Y. Ghazzawi. "Robust PID controller design for an UAV flight control system." In Proceedings of the World congress on Engineering and Computer Science, vol. 2, no. 1-6, pp. 1-6. 2011.

[2] Sadiq, Mohammad, Rameez Hayat, Kamran Zeb, Ahmed Al-Durra, and Zahid Ullah. "Robust Feedback Linearization based Disturbance Observer Control of Quadrotor UAV." IEEE Access (2024).

[3] Lee, Dasol, Hanseob Lee, Jaehyun Lee, and David Hyunchul Shim. "Design, implementation, and flight tests of a feedback linearization controller for multirotor UAVs." International Journal of Aeronautical and Space Sciences, 2017, pp. 740-756.

[4] Kada, Belkacem, and Y. Ghazzawi. "Robust PID controller design for an UAV flight control system." In Proceedings of the World congress on Engineering and Computer Science, vol. 2, no. 1-6, pp. 1-6. 2011.

[5] Yit, Kok Kai, Parvathy Rajendran, and Lim Kah Wee. "Proportional-derivative linear quadratic regulator controller design for improved longitudinal motion control of unmanned aerial vehicles." International Journal of Micro Air Vehicles 8, no. 1 (2016): 41-50.

[6] Chen, Fuyang, Wen Lei, Kangkang Zhang, Gang Tao, and Bin Jiang. "A novel nonlinear resilient control for a quadrotor UAV via backstepping control and nonlinear disturbance observer." Nonlinear Dynamics 85 (2016): 1281-1295.

[7] Zheng, En-Hui, Jing-Jing Xiong, and Ji-Liang Luo. "Second order sliding mode control for a quadrotor UAV." ISA transactions 53, no. 4 (2014): 1350-1356.

[8] Muthusamy, Praveen Kumar, Matthew Garratt, Hemanshu Pota, and Rajkumar Muthusamy. "Real-time adaptive intelligent control system for quadcopter unmanned aerial vehicles with payload uncertainties." IEEE Transactions on Industrial Electronics 69, no. 2 (2021): 1641-1653.

[9] Garcia, Gonzalo A., Shawn Keshmiri, and Daksh Shukla. "Nonlinear control based on H-infinity theory for autonomous aerial vehicle." In 2017 International Conference on Unmanned Aircraft Systems (ICUAS), pp. 336-345. IEEE, 2017.

[10] Voos, Holger. "Nonlinear control of a quadrotor micro-UAV using feedback-linearization." In 2009 IEEE International Conference on Mechatronics, pp. 1-6. IEEE, 2009.

[11] Ha, In Joong. Nonlinear decoupling theory with applications to robotics (system, decomposition, differential geometry). University of Michigan, 1985.

Controlling the Guidance of Quadrotor UAVs Group towards Moving Target

A.M. Popov
Baltic State Technical University "Voenmeh"
St. Petersburg, Russia
ORCID: 0000-0002-0732-9111

D.G. Kostrygin
Baltic State Technical University "Voenmeh"
St. Petersburg, Russia
ORCID: 0000-0002-5256-5180

V.Yu. Emelyanov
Baltic State Technical University "Voenmeh"
St. Petersburg, Russia

A.A. Shevchik
Baltic State Technical University "Voenmeh"
St. Petersburg, Russia
ORCID: 0000-0002-6223-0261

Abstract—**This paper is devoted to the problem of controlling a group of quadrotors while pointing at a maneuvering target. When solving the problem, the method of proportional guidance in 3D of the leader quadrotor was used. The other quadrotors of the group are guided to the target based on information about the trajectory of the leader's movement to the target. A consensus algorithm/protocol is used to implement the group pointing algorithm.**

Keywords — quadrotor grouping; proportional pointing; consensus; formation control

I. INTRODUCTION

The rapid development of unmanned aerial vehicle (UAV) technologies and control algorithms contributes to the expansion of their application areas. One of the current trends is the increasing autonomy of UAVs. UAVs that do not require human control and are capable of generating control signals based on information about the surrounding reality are being used for more and more tasks. Another trend is the use of groups of autonomous UAVs, including quadrotors. The increasing complexity of control implementation due to the need for coordinated behavior of vehicles (agents) is compensated by the efficiency of solving the tasks. The process of bringing a group of autonomous UAVs to a moving target point, when the trajectory of the agents is unknown in advance and determined by the movement of the target, is difficult from a scientific and technical point of view. Successful solution of such a problem can be useful when realizing the landing of a group of quadrotors on a moving platform, when delivering cargo and when escorting any objects.

The purpose of this paper is to develop an algorithm for pointing a group of quadrotors at a moving target.

The paper considers the situation when only one or several UAVs (leaders) have active targeting means, and the others are guided by receiving information about the leader's trajectory and moving along this trajectory, maintaining some predetermined configuration, possibly variable in time (swarm).

To solve such a problem, we need to consider and solve the following subproblems: control algorithm for an individual UAV, control algorithm for a group of UAVs (multi-agent control), and targeting algorithm for a leader UAV.

II. ALGORITHM DEVELOPMENT

A. Leader guidance algorithm

To mathematically describe the problem of pointing a single UAV in three-dimensional space, this paper uses a 5th order kinematic model [1]:

$$\dot{R} = -V \cos\theta_M \cos\phi_M,$$
$$\dot{\lambda}_y = \frac{V}{R} \sin\theta_M,$$
$$\dot{\lambda}_z = -\frac{V}{R} \cos\theta_M \sin\phi_M,$$
$$\dot{\theta}_M = \frac{a_z}{V} + \frac{V}{R} \cos\theta_M \sin^2\phi_M \tan\theta_L + \frac{V}{R}\sin\theta_M \cos\phi_M, \quad (1)$$
$$\dot{\phi}_M = \frac{a_y}{V\cos\theta_M} - \frac{V}{R}\sin\theta_M \sin\phi_M \cos\phi_M \tan\theta_L +$$
$$+ \frac{V}{R\cos\theta_M}\sin^2\theta_M \sin\phi_M + \frac{V}{R}\cos\theta_M \sin\phi_M,$$

where V – UAV velocity vector, R – distance from the UAV to the target along the line of sight, θ_L и ϕ_L – angles of inclination of the line of sight in the vertical and horizontal planes, respectively, σ – angle between the UAV velocity and the line of sight, θ_M и ϕ_M – angles of deviation of the velocity vector from the line of sight in two perpendicular planes, $\dot{\lambda}_y$, $\dot{\lambda}_z$ – components of the rotation velocity of the line of sight; a_y, a_z – accelerations of the UAV in the velocity coordinate system, acting as control signals. Figure 1 shows the coordinate systems.

The basic PPN pure proportional pointing law [2] for pointing in 3D is of the form [1]:

$$a_y = -NV\dot{\lambda}_y \sin\theta_M \sin\phi_M + NV\dot{\lambda}_z \cos\theta_M,$$
$$a_z = -NV\dot{\lambda}_z \cos\phi_M, \quad (2)$$

where N - navigation constant.

Figure 1 – Coordinate systems

In the present work, we propose to use a modified version of the guidance law [3]:

$$a_y = -(1 + \frac{N_m}{\cos\theta_M \cos\phi_M})V\dot{\lambda}_y \sin\theta_M \sin\phi_M +$$

$$+ (1 + \frac{N_m}{\cos\theta_M \cos\phi_M})V\dot{\lambda}_z \cos\theta_M, \quad (2)$$

$$a_z = -(1 + \frac{N_m}{\cos\theta_M \cos\phi_M})V\dot{\lambda}_y \cos\phi_M,$$

where N_m - navigation constant.

In [4], for system (1)-(3), an estimate of the time to reach the goal is obtained:

$$t_{go}(t) = \frac{R}{V}\left[1 + \frac{\sin^2\sigma}{2(2N-1)}\right], \quad (3)$$

where the angle σ is defined as $\cos\sigma = \cos\theta_M \cos\phi_M$.

B. Quadrotor control algorithm

We consider a model of a quadrotor in the NED coordinate system [5], shown in Figure 2. The equations of motion of the quadrotor in the SE (3) space have the form [6]:

$$\dot{x} = v,$$

$$m\dot{v} = mge_3 - fRe_3,$$

$$\dot{R} = R\hat{\Omega}, \quad (5)$$

$$J\dot{\Omega} + \Omega \times J\Omega = M,$$

where $x \in \mathbb{R}^3$ - position vector, $R \in SO(3)$ - directional cosine matrix, $\Omega \in \mathbb{R}^3$ - angular velocities, m - UAV mass, $J \in \mathbb{R}^{3\times3}$ - inertia matrix, $f \in \mathbb{R}$ - thrust force, $M \in \mathbb{R}^3$ - vector of control moments. The aerodynamic resistance of propellers and the UAV body [7-8] is not taken into account.

Figure 2 – Quadrotor model

To control the motion of the UAV (5) along the given trajectory $x_d(t) \in \mathbb{R}^3$ the geometric control algorithm from [6] is used:

$$f = (k_x e_x + k_v e_v + k_i \text{sat}_\sigma(e_i) + mge_3 - m\ddot{x}_d) \cdot Re_3,$$

$$M = -k_R e_R - k_\Omega e_\Omega +$$

$$+ (R^T R_c \Omega_c)^\wedge JR^T R_c \Omega_c + JR^T R_c \dot{\Omega}_c, \quad (6)$$

where $e_x = x - x_d$, $e_v = v - \dot{x}_d$, $e_i = \int_0^t e_v(\tau) + c_1 e_x(\tau)d\tau$,

$e_R = \frac{1}{2}(R_c^T R - R^T R_c)^\vee$, $e_\Omega = \Omega - R^T R_c \Omega_c$ are the error equations.

C. Multi-agent control based on consensus algorithm

A graph G can be represented as a pair (V, E), where $V = 1,2,...,n$ – is the set of nodes of the graph and $E \in V \times V$ – is the set of its edges. Moreover, each edge shows how a pair of different nodes are connected to each other. For example, edge (i, j) may indicate the ability of a node to receive information from node i. Let graph G contain n nodes. The adjacency matrix $A = [a_{ij}] \in \mathbb{R}^{N\times N}$ is defined as: $a_{ii} = 0$, $a_{ij} = 1$, if $(v_j, v_i) \in E$, otherwise $a_{ij} = 0$. In leader-follower problems [9], the adjacency matrix is formed as: $A = [a_{ij}] \in \mathbb{R}^{N+1 \times N+1}$, where the leader is $n+1$ nodes. In this case, row $n+1$ is zero, because the leader does not receive information from other nodes, and in column $n+1$ units stand only in those rows whose nodes receive information from the leader. An example of an information oriented graph is shown in Figure 3. Here, nodes 1-4 correspond to follower agents, and node 5 is the leader.

Let $\xi_i \in \mathbb{R}^3$ contain information about the position of the i-th agent. For agents with first-order dynamics in the leader's trajectory tracking problem $\xi^r \in \mathbb{R}^3$, the fundamental consensus algorithm [9]:

$$u_i = \rho_1 \sum_{j=1}^{n} a_{ij} \left[\dot{\xi}_j - \gamma \left(\xi_i - \xi_j \right) \right]$$
$$+ \rho_0 a_{i(n+1)} \left[\dot{\xi}^r - \gamma \left(\xi_i - \xi^r \right) \right], \quad i = 1, \ldots, n, \tag{7}$$

where $\rho_0 > 0, \rho_1 > 0, \gamma > 0$ are the coefficients of the algorithm. Consensus is reached asymptotically among all agents if for any $\xi_i(0)$ it is true that $\left\| \xi_i(t) - \xi_j(t) \right\| \to 0$, is true for all $i \neq j$ at $t \to \infty$. To solve the formation control problem of the agents, we introduce given offsets $\delta_i(t)$ of the position of the i-th agent from the given leader trajectory $\xi^r(t)$. Then the motion trajectories of agents $\xi_i(t)$ in (7) are replaced by:

$$\eta_i(t) = \xi_i(t) - \delta_i(t). \tag{8}$$

Figure 3 – Information graph

D. Algorithm for controlling a group of quadrotors

For targeting a group of quadrotors, the paper proposes a multi-agent control approach in the form of leader-follower formation control. The leader points to the target independently using the pointing algorithm (1), (3) and transmits information about its trajectory to all or part of the followers according to some known information graph. Initially, quadrotors have some position relative to the leader determined by offsets $\delta_{0i}, i = 1, \ldots, n$.

The final position of each quadrotor at the moment of hitting the target δ_{fi} is also specified. The current from the leader's trajectory required for the consensus algorithm (7)-(8) can be computed using expression (4) for the time to reach the goal:

$$\delta_i(t) = \delta_{fi} + \tau \delta_{0i},$$
$$\tau = \frac{t_{go}(t)}{t_{go}(t) + t}. \tag{9}$$

Then expressions (7)-(8) are used to calculate the errors $e_x(t), e_v(t)$ in the control law of the single quadrotor.

III. MODELING

To confirm the performance of the developed control algorithm, modeling was carried out in MATLAB/Simulink environment.

The task of pointing five quadrotors, a leader and four followers, was considered. The information graph is shown in Figure 3. Initial positions: leader [0; 0; 500] m; followers at x coordinate [240, -197, 470, -426] m, at y coordinate [1165, 404, 1550, 12] m, at z coordinate [725, 725, 500, 500] m. Initial velocities 10 m/s.

The formation of followers relative to the leader at the targeting endpoint: δ_{f1} = [0; 1; 0.5] m, δ_{f2} = [0; -1; 0.5] m, δ_{f3} = [0; 2; 0] m, δ_{f4} = [0; -2; 0] m.

Leader speed at homing is 30 m/s. Navigation constant $N_m = 3.5$.

Initial target position: [500; 0; 0]. Target velocity 20 m/s, target acceleration 0.5 m/s^2 - circular motion.

Coefficients of the consensus algorithm: $\rho_0 = 100, \rho_1 = 50, \gamma = 1$.

The results of the calculations are shown in Figures 4-5.

Figure 4 – Pointing results in the X-Y plane

Figure 5 – 3D pointing results

IV. CONCLUSION

In this paper, an algorithm for guiding a group of quadrotors to a target based on the leader proportional guidance method in 3D and a consensus algorithm for formation control has been developed. Computer simulation of the obtained algorithm has shown the effectiveness of the proposed approach when the target is maneuvering.

REFERENCES

[1] Seong-Ho Song and In-Joong Ha, A Lyapunov-like approach to performance analysis of 3-dimensional pure PNG laws, in IEEE Transactions on Aerospace and Electronic Systems, vol. 30, no. 1, pp. 238-248, Jan. 1994, doi: 10.1109/7.250424.

[2] P. Zarchan, Tactical and strategic missile guidance. American Institute of Aeronautics and Astronautics, 2012. doi: 10.2514/4.868948.

[3] A. M. Popov, D. G. Kostrygin, P. V. Krashanin and A. A. Shevchik, "Development of Algorithm for Guiding the Swarm of Unmanned Aerial Vehicles," 2022 29th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), Saint Petersburg, Russian Federation, 2022, pp. 1-4, doi: 10.23919/ICINS51784.2022.9815370.

[4] Shaoming He and Defu Lin. Three-Dimensional Optimal Impact Time Guidance for Antiship Missiles. Journal of Guidance, Control, and Dynamics 2019 42:4, 941-948

[5] Beard, R.W.; McLain, T.W. Small Unmanned Aircraft: Theory and Practice; Princeton University Press: Princeton, NJ, USA, 2012.

[6] F. Goodarzi, D. Lee and T. Lee, "Geometric nonlinear PID control of a quadrotor UAV on SE(3)," 2013 European Control Conference (ECC), Zurich, Switzerland, 2013, pp. 3845-3850, doi: 10.23919/ECC.2013.6669644.

[7] M. Faessler, A. Franchi and D. Scaramuzza, "Differential Flatness of Quadrotor Dynamics Subject to Rotor Drag for Accurate Tracking of High-Speed Trajectories," in IEEE Robotics and Automation Letters, vol. 3, no. 2, pp. 620-626, 2018, doi:10.1109/LRA.2017.2776353.

[8] Andrievsky, B.; Kuznetsov, N.; Popov, A. Algorithms for aerodynamic control of relative motion two satellites in a near circular orbit. Differencialnie Uravnenia i Protsesy Upravlenia 2020, pp. 28–58. (In Russian).

[9] Ren, W.; Beard, R.W. Distributed Consensus in Multi-Vehicle Cooperative Control; Springer: London, UK, 2008.

Adaptive Fault-Tolerant Attitude Control System for Small Spacecraft

I.V. Belokonov

Interuniversity Department of
Space Research
Samara University
Samara, Russia
belokonov.iv@ssau.ru

N.A. Elisov

Research Laboratory 102
Samara University
Samara, Russia
mr07th@gmail.com

A.V. Kramlikh

Interuniversity Department of
Space Research
Samara University
Samara, Russia
kramlikh.av@ssau.ru

I.A. Lomaka

Research Laboratory 102
Samara University
Samara, Russia
igorlomaka63@gmail.com

P.N. Nikolaev

Research Laboratory 102
Samara University
Samara, Russia
nikolaev.pn@ssau.ru

Abstract — **This paper develops a fault-tolerant attitude control system for small spacecraft. It is based on the reconfiguration of the software. The solution of the problem is divided into two tasks. The first one is sensor fault detection using the predicted measurements. The second task is design of the fault-tolerant feedback controller. It is also necessary to design attitude control.**

Keywords — *small spacecraft, angular motion, control system, fault tolerance, fault detection, attitude determination.*

I. INTRODUCTION

More and more small spacecraft are being launched these days. They solve many fundamental and applied problems. Often, to solve these problems, it is necessary to know and provide the required attitude of the spacecraft.

An important problem of attitude control system (ACS) design is its possible faults. The presence of a fault usually leads to mission failure. This problem is most relevant for small spacecraft with on-board commercial off-the-shelf (COTS) electronics. Usually, COTS electronics do not undergo the full set of ground tests. This is especially true for radiation resistance. Therefore, design of a fault-tolerant control system is a relevant problem.

The design of fault-tolerant control systems includes hardware, information, and software redundancy. Usually, hardware redundancy is used. However, it increases the mass and dimensions. Thus, this approach is not applicable to small spacecraft. Therefore, the development of special software can significantly improve the fault tolerance of the control systems.

This paper proposes an approach of fault detection using predicted measurements. The prediction is calculated by the adaptive on-board motion models. This approach increases computational costs. At the same time, it prevents the use of unreliable measurements.

Design of fault-tolerant ACS divided into two tasks:

- The first task is the spacecraft attitude determination when one or several sensors fail;

- The second task is the synthesis of the attitude control program in case of actuator faults.

For the first task, we assume that the small spacecraft receives information from magnetometers, angular rate sensors (ARS), sun sensors (SS), and a GNSS receiver. Attitude determination is based on various algorithms, e.g. one-time measurements algorithms [1-14] and Kalman filter [15-23]. This paper describes a technique for selecting an algorithm. The selection depends on the availability of measurements.

For the second task, we use an approach based on the representation of the control program as a Fourier series [24, 25]. The search of Fourier series coefficients is provided by the differential evolution algorithm [26]. This approach allows redistribution of the remaining resources from faulty actuators to faultless ones.

II. SENSOR FAULT DETECTION

A. *Mathematical models of small spacecraft motion*

A center of mass motion in an Earth-centered inertial (ECI) coordinate frame has the following view:

$$\ddot{\mathbf{r}} = -\frac{\mu}{r^3}\mathbf{r} + \mathbf{f}_{J2} + \mathbf{f}_A, \tag{1}$$

where $\mathbf{r}$ is radius-vector of small spacecraft; μ is gravitational standard parameter; $\mathbf{f}_{J2}, \mathbf{f}_A$ are gravitational and aerodynamic accelerations, respectively.

The kinematic equations of spacecraft angular motion are described as [27]:

$$\dot{\mathbf{q}} = \frac{1}{2}\mathbf{\Omega q}, \tag{2}$$

where $\mathbf{\omega} = \mathbf{\omega}^{bo} + \mathbf{\omega}^{oa}$ is the absolute angular velocity; $\mathbf{\omega}^{bo} = (\omega_x^{bo}, \omega_y^{bo}, \omega_z^{bo})^{\mathrm{T}}$ is the spacecraft angular velocity relative to the orbital coordinate frame; $\mathbf{\omega}^{oa} = \left(0, \sqrt{\frac{\mu}{r^3}}, 0\right)^{\mathrm{T}}$ is the orbital angular velocity relative to ECI;

$$\boldsymbol{\Omega} = \begin{bmatrix} 0 & -\omega_x^{bo} & -\omega_y^{bo} & -\omega_z^{bo} \\ \omega_x^{bo} & 0 & \omega_z^{bo} & -\omega_y^{bo} \\ \omega_y^{bo} & -\omega_z^{bo} & 0 & \omega_x^{bo} \\ \omega_z^{bo} & \omega_y^{bo} & -\omega_x^{bo} & 0 \end{bmatrix}$$ is the skew-symmetric

matrix.

Euler's dynamic equation has the following form:

$$\mathbf{I}\dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times \mathbf{I}\boldsymbol{\omega} = \mathbf{T_a} + \mathbf{T_g} + \mathbf{U} , \qquad (3)$$

where $\mathbf{I} = diag(I_x, I_y, I_z)$ is the tensor of inertia; $\boldsymbol{\omega} = (\omega_x, \omega_y, \omega_z)^{\mathrm{T}}$ is the absolute angular velocity of a spacecraft; $\mathbf{T_a} = \rho\, C_D\, S\left[\mathbf{d} \times \mathbf{V}^b\right]\left|\mathbf{V}^b\right|$ is the aerodynamic torque; $\mathbf{T_g} = \dfrac{3\mu}{r^3}\left(\boldsymbol{\eta}^b \times \mathbf{I}\boldsymbol{\eta}^b\right)$ is the gravitational torque; $\mathbf{U} = (u_x, u_y, u_z)^{\mathrm{T}}$ is the control; $\boldsymbol{\eta}^b$ is the local vertical in the body coordinate system (BCS); ρ is the density; C_D is the drag coefficient; S is the reference area; $\mathbf{d}$ is the center of pressure; V^b is the upstream velocity.

B. Mathematical models of measurements

The generalized model of vector sensors (magnetometer, sun sensors, etc.) can be written in the following form:

$$\mathbf{C}_v(t) = \mathbf{v}_{\text{meas}}(t) = \mathbf{R}(t)\mathbf{v}_{\text{orb}}(t) + \mathbf{w}_v(\sigma_v, t), \qquad (4)$$

where $\mathbf{R}(t)$ is rotation matrix from the orbital coordinate system (OCS) into BCS; $\mathbf{v}_{\text{orb}}(t)$ is the vector sensors in the OCS calculated in accordance with the given model (Sun model, Earth's magnetic field model, etc.); $\mathbf{w}_v(\sigma_v, t)$ is a noise of measurements; σ_v is the standard deviation.

The angular velocity measurement model has the following form:

$$\mathbf{C}_\omega(t) = \boldsymbol{\omega}_{\text{ИЗМ}}(t) = \boldsymbol{\omega}(t) + \mathbf{w}_\omega(\sigma_\omega, t), \qquad (5)$$

where $\mathbf{w}_\omega(\sigma_\omega, t)$ is the noise of measurements; σ_ω is the standard deviation.

C. Technique of fault detection of a sensor

The technique of fault detection of a sensor includes the following stages:

1) We identified the small spacecraft parameters. We did this by processing measurements from the previous flight stage [28, 29]. This stage is used for small spacecraft with variable mass-inertia characteristics.

2) Prediction of the center of mass motion parameters (Eq. (1)) and angular motion parameters (Eq. (2) and Eq. (3)) at the required time.

3) The measurement vector (further, vector of measurement model) is calculated in OBC.

4) The vector of the measurement model is translated from OCS into BCS in accordance with Eq. (4). The ARS vector of the measurement model is calculated by Eqs. (2),

(3) and (5). In further, the vector of the measurement model in BCS is the calculated measurement.

5) Assessment of difference (δ) between the calculated and measured values. Measurements are used in the control loop if $\delta < \delta_{\text{lim}}$. Otherwise, the measurements are not use and the attitude determination algorithm is reconfigured.

The value of δ_{lim} is defined for each measurement type with a given noise. Noise is defined by ground tests (ground calibration) or during flight missions. Otherwise, nominal values are used.

III. DESIGN OF THE FAULT TOLERANCE FEEDBACK CONTROLLER

The problem of on-board attitude determination is solved by a set of algorithms. They are based on both one-time measurements [1-14] and the Kalman filter [15-23].

Selection of the attitude determination algorithm is performed in accordance with stage 5. Solving the problem of attitude determination requires two steps. This is true for normal spacecraft operation [18]. In the first step, the QUEST algorithm uses the measurements of the magnetometer and SS. The termination criteria is the normalized quaternion. The normalized quaternion is also the initial information for the Kalman filter. The Kalman filter uses magnetometer and ARS measurements.

In the case of an ARS fault, the algorithm is similar to a normal operation. However, the Kalman filter uses only magnetometer measurements.

In the case of an SS fault, the QUEST algorithm uses the magnetometer's measurements. It also uses information about the navigation spacecraft geometric visibility. Furthermore, the algorithm performs according to the normal operation.

In the case of a magnetometer fault, the QUEST algorithm uses SS measurements. It also uses information about the navigation spacecraft's geometric visibility. The Kalman filter performs only by ARS measurements.

In the case of magnetometer and ARS faults, the attitude determination problem is solved until the GNSS receiver is operated. Continuous operation of the GNSS receiver is not considered. In this case, algorithms are applied that use information about the navigation spacecraft geometric visibility [30].

We solve the attitude determination problem by processing the sampled measurements. We perform this if there is at least one operating vector sensor [31, 32].

IV. DESIGN OF THE FAULT-TOLERANCE ATTITUDE CONTROL SYSTEM

The numerical approach [24, 25] is used to solve the problem of terminal control in the case of a fault of one or two actuators. The differential evolution algorithm is used in this approach [26].

The problem of terminal control is described as follows: it is necessary to reorient the spacecraft from its initial state:

$$\mathbf{q}(t_0) = \mathbf{q}_0, \quad \boldsymbol{\omega}(t_0) = \boldsymbol{\omega}_0 \qquad (6a)$$

to the desired ones:

$$\mathbf{q}(t_{\text{к}}) = \mathbf{q}_{\text{к}}, \quad \boldsymbol{\omega}(t_{\text{к}}) = \boldsymbol{\omega}_{\text{к}} \qquad (6b)$$

for fixed time interval $t_{\text{к}}$.

The reorientation problem with a faulted actuator is solved with the following structure (for each operating actuators):

$$U(t) = \frac{A_0}{2} + \sum_{n=1}^{8} A_n \cos\left(\frac{2\pi nt}{T} + \theta_n\right), \qquad (7)$$

where A_0, A_n, θ_n are searched parameters.

The differential evolution algorithm [26] was used to search for 34 coefficients A_0, A_n, θ_n (in the case of fault of one actuator) or 17 coefficients A_0, A_n, θ_n (in the case of fault of two actuators).

The task of searching the coefficients A_0, A_n, θ_n is reduced to the task of minimizing the following goal functions:

$$f(\mathbf{b}, t_{\text{f}}) = \arccos\left[\frac{1}{2}\left(\text{trace}\left(\mathbf{R}^{\text{T}}(\mathbf{q}_{\text{d}}(t_{\text{f}}))\mathbf{R}(\mathbf{q}(t_{\text{f}}))\right) - 1\right)\right] +$$
$$+ t_f \left|\boldsymbol{\omega}_{\text{d}} - \boldsymbol{\omega}(t_{\text{f}})\right|,$$

where $\boldsymbol{b} = (A_0, A_n, \theta_n)^{\text{T}}$ is the vector of the estimated parameters; $\mathbf{R}^{\text{T}}(\mathbf{q}(t_{\text{f}}))$ is the transition matrix from OCS to BCS (Eq. (6b)); $\mathbf{R}(\mathbf{q}(t_{\text{f}}))$ is the transition matrix from OCS to BCS calculated from the current normalized quaternion.

CONCLUSION

This paper proposes an approach to designing a fault-tolerant ACS for small spacecraft. The approach is based on the reconfiguration of on-board software. We propose a technique for sensor fault determination, which uses adaptive on-board motion models for prediction. The proposed technique prevents analytical redundancy of measurements. In other words, we may not implement on-board observers. It should be noted that the problem of observer synthesis has not been solved in the general case.

The fault tolerance of the feedback controller is provided by the reconfiguration of the algorithm depending on the available measurements. The fault tolerance of the control is provided by the representation of the control structure as the even Fourier series. Then, the problem is reduced to searching the Fourier series coefficients. This approach allows redistribution of control from faulty actuators to the operative ones.

REFERENCES

[1] F. L. Markley and J. L. Crassidis, Fundamentals of Spacecraft Attitude Determination and Control. New York, NY: Springer New York, 2014.

[2] J. R. Wertz, ed., Spacecraft Attitude Determination and Control, Vol. 73 of Astrophysics and Space Science Library. Dordrecht: Springer Netherlands, 1978.

[3] L. Markley, "Attitude determination using two vector measurements". In NASA Conference Publication; NASA: Washington, DC,USA, 1999; pp. 39-52.

[4] I. Y. Bar-Itzhack, and R. R. Harman, "Optimized TRIAD Algorithm for Attitude Determination," Journal of Guidance, Control, and Dynamics, Vol. 20, No. 1, January-February 1997, pp. 208-211.

[5] M. D. Shuster, "The optimization of TRIAD". The Journal of the Astronautical Sciences, Vol. 55, No 2, April – June 2007, pp. 245–257.

[6] H. D. Black, "A Passive System for Determining the Attitude of a Satellite," AIAA Journal, Vol. 2, July 1964, pp. 1350–1351.

[7] F. L. Markley, and M. Mortari, "Quaternion Attitude Estimation Using Vector Measurements, The Journal of the Astronautical Sciences, Vol. 48, Nos. 2 and 3, April–September 2000, pp. 359–380.

[8] D. Mortari, "EULER-2 and EULER-n Algorithms for Attitude Determination from Vector Observations," Space Technology, Vol. 16, Nos. 5–6, 1996, pp. 317–321.

[9] D. Choukroun, I. Y. Bar-Itzhack and Y. Oshman, 'Optimal-REQUEST Algorithm for Attitude Determination', May 2004 Journal of Guidance, Control, and Dynamics 27(3), pp. 418-425.

[10] D. Mortari, 'ESOQ: A Closed-Form Solution to the Wahba Problem'', Journal of Astronautical Sciences. 1997, vol. 45, pp. 195–205.

[11] D. Mortari, "ESOQ 2 Single Point Algorithm for Fast Optimal Spacecraft Attitude Determination", Am. Soc. Mech. Eng. 1997, vol. 95, pp. 817–826.

[12] A. V. Kramlikh, M. E. Melnik, and P. N. Nikolaev, "Attitude determination and stabilization algorithms of the samsat-218D nanosatellite", 23rd Saint Petersburg International Conference on Integrated Navigation Systems, ICINS 2016 - Proceedings. — 2016. – P. 366-372.

[13] F. L. Markley, "Attitude Determination Using Vector Observations: a Fast Optimal Matrix Algorithm". Jouranl of Astronaut Sci., 1993, vol. 41(2), pp. 261−280.

[14] D Cilden-Guler, E. S. Conguroglu, and C. Hajiyev, "Single-Frame Attitude Determination Methods for Nanosatellites". Metrology and Measurement Systems, 2017, vol. 24, pp.313–324.

[15] E. J. Lefferts, F. L. Markley, and M. D. Shuster. "Kalman Filtering for Spacecraft Attitude Estimation." Journal of Guidance Control and Dynamics 5 (1982), pp. 417-429.

[16] F. L. Markley, "Attitude Error Representations for Kalman Filtering". Journal of Guidance Control and Dynamics, vol. 26(2), pp. 311-317.

[17] J. D. Searcy, and H. J. Pernicka, "Magnetometer-Only Attitude Determination Using Novel Two-Step Kalman Filter Approach". J. Guid. Control Dyn. 2012, vol. 35, pp. 1639–1701.

[18] A. V. Kramlikh, P. N. Nikolaev, and D. V. Rylko "Onboard Two-Step Attitude Determination Algorithm for a SamSat-ION Nanosatellite", Gyroscopy and Navigation, 2023, vol. 14. Issue 2. № 2., pp. 138-153.

[19] D. Cilden-Guler, and C. Hajiyev, "SVD-Aided EKF for Nanosatellite Attitude Estimation Based on Kinematic and Dynamic Relations". Gyroscopy and Navigation, 2023, vol. 14, pp. 366–379.

[20] M. Ovchinnikov, and D. Ivanov, "Approach to study satellite attitude determination algorithms". Acta Astronaut., 2014, vol. 98, pp. 133-137.

[21] K. Xiong, T. Liang, and L. Yongjun, "Multiple model Kalman filter for attitude determination of precision pointing spacecraft". Acta Astronaut., 2011, vol. 68, pp. 843-852.

[22] J. L. Crassidis, and F. L. Markley, "Unscented filtering for spacecraft attitude estimation". J. Guid. Contr. Dynam, vol. 26(4), pp. 536-542.

[23] M. D. Pham, K. S. Low, S. T. Goh, and S. Chen, "Gain-scheduled extended Kalman filter for nanosatellite attitude determination system," IEEE Trans. Aerosp. Electron. Syst., vol. 51, no. 2, pp. 1017–1028.

[24] N. A. Elisov, A. V. Kramlikh, I. A. Lomaka, and D. P. Avariaskin, "An attitude control by the functional series in the problem of nanosatellite reorientation", Aerospace Science and Technology, 2023, vol. 132.

[25] N. A. Elisov, A. V. Kramlikh, and I. A. Lomaka, "Synthesis of Nominal Reorientation Trajectories of a Small Satellite in Case of Failure of One Actuator". Mekhatronika, Avtomatizatsiya, Upravlenie, 2023., vol. 24. Issue 11., pp. 608-615.

[26] R. Storn, and K. Price, "Differential Evolution - A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces". Journal of Global Optimization, 1997, no. 11, pp. 341-359.

[27] В. Н. Бранец, И. П. Шмыглевский Применение кватернионов в задачах ориентации твердого тела. М.: Наука, 1973. 320 с.

[28] I. V. Belokonov, and I. A. Lomaka, "Investigation of the possibility of determining the inertial characteristics and the angular velocity vector of chaotically rotating space debris object using a nanosatellite". 25th Saint Petersburg International Conference on Integrated Navigation Systems, ICINS 2018 - Proceedings.

[29] I. V. Belokonov, and I. A. Lomaka, "In-flight calibration of nanosatellites inertia tensor: The algorithm and requirements for on-board sensors". Proceedings of the International Astronautical Congress, IAC. — 2018. — Vol. 2018-October.

[30] I.V. Belokonov, A.V. Kramlikh, Space vehicle attitude control recovery procedure combining magnetometric and radionavigation measurements // Vestnik of Samara State Aerospace University. — 2007. — № 1 (12). — PP. 22-30.

[31] I. V. Belokonov, A. V. Kramlikh, I. A. Lomaka, and P. N. Nikolaev, "Reconstruction of a Spacecraft's Attitude Motion Using the Data on the Current Collected from Solar Panels". Journal of Computer and Systems Sciences International 2019. – Vol. 58. Issue 2., pp. 286-296.

[32] I. V. Belokonov, and I. A. Lomaka "Postflight Recovery of the Rotational Motion of a Small Space Vehicle from Solar Sensor Information". Journal of Computer and Systems Sciences International 2023. – Vol. 62. Issue 2. № 2., pp. 214-224.

Active Aerodynamic Stabilization of the Nanosatellite Angular Motion at Very Low Earth Orbit

N.A. Elisov

Research Laboratory 102
Samara University
Samara, Russia
Laboratory 549 Solar Astronomy
and Heliophysical Instrumentation,
Russian Space Research Institute,
Moscow, Russia
mr07th@gmail.com

A.V. Kramlikh

Interuniversity Department of
Space Research
Samara University
Samara, Russia
kramlikh.av@ssau.ru

I.A. Lomaka

Research Laboratory 102
Samara University
Samara, Russia
Laboratory 549 Solar Astronomy
and Heliophysical Instrumentation,
Russian Space Research Institute,
Moscow, Russia
igorlomaka63@gmail.com

Abstract—**This paper investigates stabilization of a nanosatellite planar angular motion by a deflectable tail. We use the control-Lyapunov function to synthesize the control law. The control law is applied to CubeSat-3U with various tail.**

Keywords—nanosatellite, very low Earth orbit, control-Lyapunov function, angular motion, direct simulation Monte-Carlo, differential evolution

I. INTRODUCTION

Very low Earth orbits (VLEO) allow us to obtain new scientific data. It is about the Earth's atmosphere and surface. VLEO has not been sufficiency studied enough due to the short lifetime of a spacecraft. VLEO has advantages. It has a higher resolution for remote sensing. It also has a lower probability of collision with space debris. Finally, it has simpler and faster deorbiting [1]. However, a spacecraft at VLEO faces higher drag. It has significant effects on the angular motion and life span of the spacecraft [1].

The analysis [2] of a space mission showed that the launch of a small spacecraft in VLEO is preferable. However, a small spacecraft experiences significant effects from aerodynamic and gravitational forces. They affect its angular motion [3].

Usually, the investigation of the angular motion is provided for a small spacecraft with fixed geometry. Therefore, layout gets special attention during design. This is to ensure passive stabilization [4–6].

Recently, researchers have begun to study active aerodynamic stabilization. One of the proposed methods is the rotating tail [7-9]. In [10], the deflectable tail was considered in the stabilization and reorientation problems. However, in [10], the control program used for the problem solution. It did not have feedback.

In this paper, the control law is derived based on the control-Lyapunov function. It was applied to nanosatellite with various tails (Fig. 1) in stabilization problem. The tails are 100x100 mm (type 1), 200x100 mm (type 2), and 300x100 mm (type 3).

Fig. 1. Nanosatellite with the tail: (a) type 1; (b) type 2 and (c) type 3

II. PROBLEM FORMULATION

A. Nanosatellite dynamics

The motion of a nanosatellite in a plane is described by the following system of equations [3]:

$$\dot{\alpha} = \omega, \tag{1}$$

$$\dot{\omega} = I_\tau^{-1}\left[T_a(\alpha,\delta) + k_g \sin(2\alpha)\right], \tag{2}$$

where α is the angle of attack; T_a is the aerodynamic torque; k_g is the gravitational torque coefficient; I_τ is the transverse moment of inertia.

The aerodynamic torque T_a has the following view:

$$T_a = C_\tau(\alpha,\delta) q_\infty S_{ref} l_{ref}, \tag{3}$$

where C_τ is the pitching moment coefficient; q_∞ is the velocity head; S_{ref} is the reference area; l_{ref} is the reference length.

The gravitational torque coefficient describes as:

$$k_g = 1.5\left(I_\tau - I_l\right)\omega_{orb}^2, \tag{4}$$

where I_l is the longitudinal moment of inertia; ω_{orb} is the orbital angular velocity.

B. Attitude control and its stability

The problem is described as follows: it is necessary to synthesize the control law of a nanosatellite motion stabilization. The controller is the tail. Stability analysis is provided by the direct Lyapunov method [11]. According to this method, the following conditions must be met [11]:

1. Lyapunov-candidate-function V must be globally positive definite except for the equilibrium point;

2. Lyapunov-candidate-function V must be constantly decreasing (or $\dot{V} < 0$), except for the equilibrium point.

Let us define the Lyapunov-candidate-function as follows:

$$V = 0.5\left[I_\tau \left(\omega - \omega_d\right)^2 + K\left(\alpha - \alpha_d\right)^2 \right], \qquad (5)$$

where α_d, ω_d are desired angle of attack angular velocity, respectively; K is the positive constant.

Thus, Eq. (5) is the globally positive definite. The derivative of Eq. (5) has the following view:

$$\dot{V} = I_\tau \tilde{\omega}\dot{\tilde{\omega}} + K\tilde{\alpha}\dot{\tilde{\alpha}}. \qquad (6)$$

Here $\tilde{\omega} = \omega - \omega_d$ and $\tilde{\alpha} = \alpha - \alpha_d$. Substituting Eq. (1) and Eq. (2) into Eq. (6) we obtain:

$$\dot{V} = \tilde{\omega}\left[T_a + k_g \sin\left(2\tilde{\alpha}\right) + K\tilde{\alpha} \right] \qquad (7)$$

To satisfy the second condition, we require that the aerodynamic torque changes in accordance with relation:

$$T_a = -k_g \sin\left(2\tilde{\alpha}\right) - K_a\tilde{\alpha} - K_\omega\tilde{\omega} \qquad (8)$$

where K_a, K_ω are positive constants.

Substituting Eq. (8) into Eq. (7), we obtain the following relation:

$$\dot{V} = \tilde{\omega}\left[-K_a\tilde{\alpha} - K_\omega\tilde{\omega} + K\tilde{\alpha} \right] \qquad (9)$$

We suppose that $K_a = K$. Then:

$$\dot{V} = -K_\omega\tilde{\omega}^2. \qquad (10)$$

Thus, Eqs. (6) and (10) guarantee the stable motion of a nanosatellite during its stabilization.

The coefficients K_a and K_ω is searched by the differential evolution algorithm [12]. The goal function has the following view:

$$\Phi\left(\delta\right) = K_T \left| T_a - k_g \sin\left(2\alpha\right)\right| + f, \qquad (11)$$

where K_T, is the weight coefficient; f is the penalty function:

$$f = \begin{cases} 0, & \dot{V}\left(t_i\right) < 0 \\ 10^6, & \dot{V}\left(t_i\right) > 0 \end{cases},$$
$$t_i = t_0 + ih,$$

where $t_i \in \left[0; \quad T\right]$ is the current time; h is the step; T is the modelling time.

The aerodynamic torque T_a is controlled by changing of the pitching coefficient moment C_τ. Let us rewrite Eq. (8) in accordance with Eq. (3):

$$C_\tau^{req} = \frac{-k_g \sin\left(2\tilde{\alpha}\right) - K_a\tilde{\alpha} - K_\omega\tilde{\omega}}{q_\infty S_{ref} l_{ref}}. \qquad (12)$$

The use of control law (12) is incorrect. This is because the pitching moment coefficient can change in the range of $C_\tau^{req} \in \left(-\infty; +\infty\right)$. In return, a nanosatellite has specific values of the pitching moment coefficient C_τ. Deriving the control law (12) through the deflection angle δ is very difficult. This is because the pitching moment coefficient C_τ depends on angle of attack α and deflection angle δ. Therefore, we must

restrict the required pitching moment coefficient C_τ^{req}. The pitching coefficient moment C_τ, is restricted by the brute force method (Fig. 2). The meaning of the brute force method is as follows:

1. The required pitching moment coefficient C_τ^{req} is defined in accordance with Eq. (12);

2. We calculate the possible pitching moment coefficients for the current angle of attack $\alpha(t)$. They are in the range of $\delta \in \left[0^o; 180^o\right]$ with a resolution 0.01^o;

3. We select the pitching moment coefficient C_τ. We do this based on the minimum of subtraction of the required and available coefficients.

Fig. 2. Restriction scheme of pitching moment coefficient C_τ

C. Nanosatellite aerodynamics

It is necessary to determine the aerodynamic characteristics of the nanosatellite for motion stabilization. The aerodynamic characteristics were calculated by the direct simulation Monte Carlo method [13]. The parameters of the considered nanosatellite are given in Table I.

TABLE I.　NANOSATELLITE CHARACTERISTICS

Characteristics	Value
Longitudinal moment of inertia I_τ, kg m^2	0.014
Transverse moment of inertia I_l, kg m^2	0.072
Reference length l_{ref}, м	0.3
Reference area S_{ref}, м^2	0.01
Tail length l_{tail}, m	[0.1; 0.2; 0.3]
Minimum angle of tail deflection δ_{min}, ^{o}	0
Maximum angle of tail deflection δ_{max}, ^{o}	180

The aerodynamic characteristics of the nanosatellite were calculated for flight altitude $H = 300$ km and date of 18.07.2023. The average daily chemical composition at this altitude and date is given in Table II [14].

TABLE II.　ATMOSPHERE COMPOSITION

Species	Mole fraction χ, %
Molecular nitrogen N_2	23.67
Molecular oxygen O_2	0.685
Atomic nitrogen N	1.633
Atomic oxygen O	73.41
Helium He	0.591
Hydrogen H	0.006
Argon Ar	0.001

Gas-surface interaction is described in accordance with the Cercignani-Lampis-Lord model [15-17]. The accommodation coefficients were taken equal to $\alpha_n = \sigma_\tau = 0.9$. The number of simulated particles was $N_{sim} = 3 \cdot 10^6$. The aerodynamic characteristics of the NS were obtained for the range of angles of attack $\alpha \in [-90; 90]^o$ with a resolution of 10° and tail deflection angles $\delta \in [0; 180]^o$ with a resolution of 15°.

Dependencies of the pitching coefficient moment C_τ on the angle of attack and deflection angle are given in Figs. 3-5. The highest value of the pitching moment coefficient is achieved at the deflection angle $\delta = 135^o$. It is worth noting, that tail type 3 has positive values of the pitching moment coefficient (i.e. the motion is unstable) for the range of $\delta \in \left[-45^o; 45^o \right]$.

Fig. 3. Dependence of the pitching coefficient moment on the angle of attack and deflection angle for tail type 1

Fig. 4. Dependence of the pitching coefficient moment on the angle of attack and deflection angle for the type 2

Fig. 5. Dependence of the pitching coefficient moment on the angle of attack and deflection angle for tail type 3

The pitching moment coefficients are then interpolated by bicubic polynomials [18]:

$$C_\tau \left(\alpha, \delta \right) = \sum_{i=0}^{3} \sum_{j=0}^{3} a_{i,j} \alpha^i \delta^j, \qquad (13)$$

where $a_{i,j}$ are polynomial coefficients.

III. RESULTS

The boundary conditions are given in Table III.

TABLE III. BOUNDARY CONDITIONS

Initial conditions		Boundary conditions	
Angle of attack α, °	Angular velocity ω, °/s	Angle of attack α, °	Angular velocity ω, °/s
20	0.5	0	0

The results of the problem solution are given in Fig. 6 and Fig. 7. The coefficient values of K_a и K_ω are presented in Table IV.

TABLE IV. RESULTS OF THE PROBLEM SOLUTION

Tail type	Coefficients	
	K_a	K_ω
Type 1	$5.243 \cdot 10^{-4}$	0.158
Type 2	$2.016 \cdot 10^{-5}$	0.0182
Type 3	$3.858 \cdot 10^{-5}$	0.0263

As shown in Fig. 6, tail type 1 cannot solve the stabilization problem because it has insufficient aerodynamic torque. On the other hand, tails type 2 and type 3 have sufficient aerodynamic torque. The stabilization time is 6764 and 5427 s for tails type 2 and type 3, respectively.

Fig. 6. Dependence of the angle of attack on time

The Lyapunov-candidate-function has constantly decreasing behavior (Fig. 7). In this regard, we can conclude that the control is globally asymptotically stable.

Fig. 7. Dependence of the Lyapunov-candidate-function on time

CONCLUSION

This paper considered the planar angular motion of a nanosatellite at very low Earth orbit. The control-Lyapunov function of the tail was obtained in stabilization problem. The

aerodynamic characteristics were calculated for the nanosatellite with three types of tails. At the same time, it revealed the following points:

- Tail type 1 does not have sufficient aerodynamic torque for the motion stabilization;

- The tails type 2 and type 3 solve the problem for 1.5 hours.

REFERENCES

[1] N.H. Crisp a, P.C.E. Roberts a, S. Livadiotti et al. The benefits of very low earth orbit for earth observation missions, Progress in Aerospace Sciences, 2020, vol. 117, p. 100619.

[2] A. Shao, E.A Koltz., J.R. Wertz Quantifying the Cost Reduction Potential for Earth Observation Satellites, AIAA Journal, 2014, vol. 4260, pp. 1-12.

[3] E.V. Barinova, I.V. Belokonov, I.A. Timbai, Motion Features of Aerodynamically Stabilized Cubesat 6U Nanosatellites, 29th St. Petersburg International Conference on Integrated Navigation Systems, St. Petersburg, CSRI Elektropribor, 2022, 4 p.

[4] Belokonov, I.V., Timbai, I.A. & Barinova, E.V. Design Parameters Selection for CubeSat Nanosatellite with a Passive Stabilization System. Gyroscopy Navig. 11, 149–161 (2020).
https://doi.org/10.1134/S2075108720020029.

[5] M.L. Psiaki, Nanosatellite attitude stabilization using passive aerodynamics and active magnetic torquing, Journal of Guidance, Control, and Dynamics, 2004, vol. 27, no. 3, pp. 347–355.

[6] S.A. Rawashdeh, J.E. Lumpp, Aerodynamic Stability for CubeSats at ISS Orbit, Journal of Small Satellites, 2013, vol. 2, no. 1, pp. 85-104.

[7] V. Canas, D. Gonzalez, J. Becedas et al., Attitude control for satellites flying in VLEO using aerodynamic surfaces, Journal of the British Interplanetary Society, 2020, vol. 73, no. 3, pp. 103-112.

[8] S. Livadiotti, N.H. Crisp, P.C.E. Roberts et al., Uncertainties and Design of Active Aerodynamic Attitude Control in Very Low Earth Orbit, 2022, vol. 45, no. 2, pp. 1-16.

[9] L. Chen, H. Gui, S. Xiao, Aerodynamic Attitude Control of Ultra-low Earth Orbit Satellite, Advances in Guidance, Navigation and Control, ICGNC 2022, Lecture Notes in Electrical Engineering, 2023, vol 845, pp. 5898-5908.

[10] E.V. Barinova, N.A. Elisov, I.A. Lomaka, A.V. Kramlikh, Angular Motion Control of a Low-Orbit Nanosatellite Using Aerodynamic Surfaces, 30th Anniversary Saint Petersburg International Conference on Integrated Navigation Systems, St. Petersburg, CSRI Elektropribor, 2023, 4 p.

[11] J. La Salle, S. Lefschetz, Stability by Liapunov's Direct Method with Applications, 1st ed., London: Academic Press, 1961.

[12] Storn, R., Price, K., Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces, Journal of Global Optimization, 1997, no. 11, pp. 341–359.

[13] G.A. Bird, Molecular Gas Dynamics and the Direct Simulation of Gas Flows. Oxford: Clarendon Press, 1994.

[14] J. M. Picone, A. E. Hedin, D. P. Drob. A.C. Aikin, NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues, Journal of Geophysical Research: Space Physics, 2002, vol. 107, SIA-15.

[15] R.G. Lord, Some extensions to the Cercignani-Lampis gas-surface scattering kernel, Journal of Physics of Fluids, 1991, vol. 3, pp. 1427-1433.

[16] R.G. Lord, Some further extensions of the Cercignani-Lampis gas-surface scattering kernel, Journal of Physics of Fluids, 1995, vol. 7, pp. 1159-1161.

[17] R.G. Lord, Application of the Cercignani-Lampis Scattering Kernel to Direct Simulation Monte Carlo Calculations, 17th International Symposium on Rarefied Gas Dynamics, Aachen, Rheinisch-Westfaelische Technische Hochschule Aachen, 1995, pp. 1427-1433.

[18] C.A. Hall, Natural cubic and bicubic interpolation, SIAM Journal on Numerical Analysis, 1973, vol. 10, pp. 1055-1060.

Algorithm for Trajectory Control of an Autonomous Quadcopter at the Final Section

A.M. Popov
Baltic State Technical University "Voenmeh"
St. Petersburg, Russia
ORCID 0000-0002-0732-9111

E.E. Vorobyeva
Baltic State Technical University "Voenmeh"
St. Petersburg, Russia

I.A. Yakovlev
Baltic State Technical University "Voenmeh"
St. Petersburg, Russia

D.G. Kostrygin
Baltic State Technical University "Voenmeh"
St. Petersburg, Russia
ORCID 0000-0002-5256-5180

Abstract — **The article is devoted to solving the problem of delivering a payload to a maneuvering platform. An algorithm has been developed for trajectory control of a quadcopter at the final section of the trajectory using the method of proportional guidance in 3D and a predictive model. Simulations were carried out, which showed good efficiency and robustness of the developed payload delivery algorithm.**

Keywords — quadrotor control, proportional pointing, predictive model

I. INTRODUCTION

Currently, a lot of research is being conducted to develop trajectory control algorithms for small unmanned aerial vehicles (UAVs). One of the most important challenges is the task of delivering payloads to a moving platform. Maneuvering targets create complexities that challenge existing autonomous navigation paradigms, requiring algorithmic solutions capable of real-time adaptation to unpredictable target/platform motions. There are quite a number of guidance methods available [1]. Their diversity is mainly determined by the fact that different vehicles have a different set of sensors and gauges available and, as a consequence, the type of control action should take into account the available information and hardware realization capabilities. The direct pointing method assumes that the longitudinal axis of the UAV coincides with the line of sight during the entire time of movement. The chase method implies the movement of the control object with a zero angle of anticipation, i.e., the UAV velocity vector must constantly coincide with the line of sight of the target. When implementing the parallel approach method, the control loop keeps the angle of the target line of sight constant.

Proportional guidance algorithms form a large group of methods. The Pure Proportional navigation (PPN) method is of the greatest interest because it was developed specifically for efficient target guidance and tracking of targets.

II. ALGORITHM DEVELOPMENT

A. Problem statement.

The purpose of this paper is to develop a proportional

guidance algorithm for an autonomous quadrotor to for delivering cargo to a moving platform. The PPN pure proportional guidance law is considered. It guarantees that the rate of change of the angle of the target line of sight is directly proportional to the angular velocity of the target. This mathematical relationship governs the continuous adjustment of the quadrotor's trajectory to minimize the target line-of-sight angle, guiding it to the predicted target position. The law provides guidance by "anticipating" the future position of the target based on its current motion, which demonstrates the effectiveness of the method against maneuvering objects.

The following approach is proposed in this paper: a virtual predictive model is created in the control loop in the form of a system of sixth-order differential equations. This model receives the necessary information from real UAV sensors and calculates the necessary control actions using the PPN guidance method. As a result, the model generates the trajectory along which the quadrotor should move in order to hit the target. The resulting trajectory is the input signal for the main control system of the quadrotor, which produces real values of thrust force and torques, ensuring movement along this trajectory. For this purpose, the control system must be highly efficient, allowing the quadrotor to perform complex maneuvers in space. The mathematical model of the quadrotor used for synthesizing the control loop must not have singular points.

B. Algorithm for controlling quadrotor motion along a given trajectory

The model of the quadrotor in the NED coordinate system [2] is shown in Figure 1. The equations of motion of the quadrotor in the SE(3) space have the form [3]:

$$
\begin{aligned}
\dot{x} &= v, \\
m\dot{v} &= mge_3 - fRe_3, \\
\dot{R} &= R\hat{\Omega}, \\
J\dot{\Omega} + \Omega \times J\Omega &= M,
\end{aligned}
\tag{1}
$$

where $x \in \mathbb{R}^3$ - position vector, $R \in SO(3)$ - directional cosine matrix, $\Omega \in \mathbb{R}^3$ - angular velocities, m - UAV mass, $J \in \mathbb{R}^{3 \times 3}$ - inertia matrix, $f \in \mathbb{R}$ - thrust force,

$M \in \mathbb{R}^3$ - vector of control moments. The aerodynamic drag of the propellers and the UAV body [4-5] is not taken into account.

Fig. 1. Quadrotor model [3]

A geometric control algorithm [6-7] from [3] is used to move the UAV (1) along a given trajectory $x_d(t) \in \mathbb{R}^3$:

$$f = \left(k_x e_x + k_v e_v + k_i \mathrm{sat}_\sigma(e_i) + mge_3 - m\ddot{x}_d\right) \cdot Re_3,$$
$$M = -k_R e_R - k_\Omega e_\Omega +$$
$$+\left(R^T R_c \Omega_c\right)^\wedge J R^T R_c \Omega_c + J R^T R_c \dot{\Omega}_c, \tag{2}$$

where $e_x = x - x_d$, $e_v = v - \dot{x}_d$, $e_i = \int_0^t e_v(\tau) + c_1 e_x(\tau) d\tau$,

$e_R = \dfrac{1}{2}\left(R_c^T R - R^T R_c\right)^\vee$, $e_\Omega = \Omega - R^T R_c \Omega_c$ the   error

equations. The required values of the rotation matrix R_c and angular velocity Ω_c are calculated by the formulas:

$$R_c = \left[b_{1_c}; b_{2_c}; b_{3_c}\right], \quad \hat{\Omega}_c = R_c^T \dot{R}_c$$
$$\dot{R}_c = \left[\dot{b}_{1_c}, \dot{b}_{2_c}, \dot{b}_{3_c}\right], \quad \ddot{R}_c = [\ddot{b}_{1_c}, \ddot{b}_{2_c}, \ddot{b}_{3_c}], \tag{3}$$
$$\Omega_c = \left(R_c^T \dot{R}_c\right)^\vee, \quad \dot{\Omega}_c = \left(R_c^T \ddot{R}_c - \hat{\Omega}_c^2\right)^\vee.$$

The construction of the $\left[b_{1_c}; b_{2_c}; b_{3_c}\right]$ vector system is shown in Figure 2.

Fig. 2. Vector system [3]

The vectors are calculated using the formulas:

$$b_{3_c} = -\frac{-k_x e_x - k_v e_v - mge_3 + m\ddot{x}_d}{-k_x e_x - k_v e_v - mge_3 + m\ddot{x}_d},$$
$$b_{2_c} = \frac{b_{3_c} \times b_{1_d}}{b_{3_c} \times b_{1_d}} \tag{4}$$
$$b_{1_c} = b_{2_c} \times b_{3_c}.$$

In equations (4), in addition to the given trajectory $x_d(t)$, it is necessary to specify a vector $b_{1_d}(t) \in \mathbb{R}^3$, which specifies the required orientation of the UAV. In this paper, we propose to always choose b_{1_d} in the direction of the desired velocity $\dot{x}_d$. With this approach, the geometric control algorithm (1)-(4) has high efficiency and no singular points. Figure 3 shows the simulation results for a given trajectory as a circle in the vertical plane. The quadrotor, starting from a 180-degree inverted position, further moves along the desired trajectory.

Fig. 3. Circular motion

C. UAV guidance algorithm

The PPN proportional pointing law in 3D [1] is used to point the quadrotor at the target:

$$a_{uav} = N\left[\Omega_{LOS} \times V_{uav}\right], \tag{5}$$

Where V_{uav} - vector of UAV velocity, Ω_{LOS} − vector of angular velocity of rotation of the target line of sight, N - navigation constant, a_{uav} - required control accelerations.

Since the control accelerations are very difficult to recalculate into real quadrotor controls f, M, this paper proposes to use a virtual predictive model of the form:

$$\ddot{x}_m = a_{uav}, \tag{6}$$

where $x_m(t) \in \mathbb{R}^3$ - the trajectory of motion to the target, which is created by the virtual model. Further we assume that $x_d(t) = x_m(t)$. Thus we set the desired trajectory for the UAV. Since the control algorithm (2)-(4) is very efficient and fast, the real trajectory of the UAV will converge to the target

$$\lim_{t \to \infty}\left\|x(t) - x_m(t)\right\| = 0. \tag{7}$$

III. MODELING

In order to confirm the performance of the developed control algorithms, modeling was carried out in MATLAB/Simulink package. The system (1) and the control law (2)-(6) were considered. Quadrotor parameters: mass = 5 kg, intertia $J = diag$ (0.05, 0.05, 0.08). Initial conditions for UAVs: $x(0) = [0,0,-250]$, $v(0) = [0,0,0]$. Maximum acceleration 14g. Navigation constant $N = 3$. Initial conditions for the target $x_t(0) = [500, 250, 0]$, $v_t(0) = [0, 0, 0]$. Velocity vector of the virtual model $v_m(0) = [20, 20, 0]$. Controller parameters (2):

$k_x = 10, k_v = 8, k_i = 10, c_1 = 1.5, \sigma = 10, k_R = 1.5, k_\Omega = 0.35.$

The results of the calculations are shown in Figures 4-7.

Fig. 4.   UAV and target trajectory

Fig. 5.   UAV speeds and targets

Fig. 6.   UAV acceleration

Fig. 7.   Guidance trajectory tracking errors

IV.   CONCLUSION

Computer simulation was carried out in the Matlab/Simulink system of various scenarios for the operation of the resulting trajectory control algorithm. The simulation results showed good performance and efficiency of the proposed approach. The use of geometric control made it possible to move quite accurately along the approach trajectories that were produced by the predictive system.

REFERENCES

[1]   Zarchan P. Tactical and Strategic Missile Guidance. Published by the American Institute of Aeronautics and Astronautics, 2012, 1095 p.

[2]   Beard, R.W.; McLain, T.W. Small Unmanned Aircraft: Theory and Practice; Princeton University Press: Princeton, NJ, USA, 2012.

[3]   F. Goodarzi, D. Lee and T. Lee, "Geometric nonlinear PID control of a quadrotor UAV on SE(3)," 2013 European Control Conference (ECC), Zurich, Switzerland, 2013, pp. 3845-3850, doi: 10.23919/ECC.2013.6669644.

[4]   M. Faessler, A. Franchi and D. Scaramuzza, "Differential Flatness of Quadrotor Dynamics Subject to Rotor Drag for Accurate Tracking of High-Speed Trajectories," in IEEE Robotics and Automation Letters, vol. 3, no. 2, pp. 620-626, 2018, doi:10.1109/LRA.2017.2776353.

[5]   Andrievsky, B.; Kuznetsov, N.; Popov, A. Algorithms for aerodynamic control of relative motion two satellites in a near circular orbit. Differencialnie Uravnenia i Protsesy Upravlenia 2020, pp. 28–58. (In Russian).

[6]   T. Lee, M. Leok, and N. H. McClamroch, "Geometric tracking control of a quadrotor UAV on SE(3)," in 49th IEEE conference on decision and control (CDC), pp. 5420–5425, IEEE, 2010.

[7]   T. Lee, M. Leok, and N. H. McClamroch, Nonlinear Robust Tracking Control of a Quadrotor UAV. Asian J Control, 2013.15: 391-408. https://doi.org/10.1002/asjc.567

Nadir Finding System with Infrared Temperature Sensors for CubeSat satellites

D. V. Akulin

Special Technology Center,
LLC
Saint Petersburg, Russia
alkuin.work@gmail.com

M. V. Mentyukov

Special Technology Center,
LLC
Saint Petersburg, Russia
maksim.mentukov.stc@yandex.ru

R.B. Goncharov

Saint Petersburg
Electrotechnical University
'LETI'
Saint Petersburg, Russia
rbgoncharov@stud.etu.ru

Abstract—**The paper presents a description of the system for determining the local vertical on board a nanosatellite, based on infrared temperature sensors. The main principles of the infrared temperature sensor operation and its application in the development of a spacecraft attitude control system are described. The results of laboratory tests of the system for determining the local vertical, as well as its trials in space are presented.**

Keywords—nanosatellite, CubeSat, the Earth's infrared radiation, infrared sensors, local vertical, nadir finding.

I. INTRODUCTION

At present, a new segment of nanosatellites such as CubeSat is being formed in the space industry. Relatively low cost of development and production make this type of spacecraft increasingly popular and available for mass use for research and commercial purposes. In this segment, it is important to miniaturize the functional units used in large spacecraft and to customize the technologies for the CubeSat format in order to reduce the cost of equipment, its weight and size, and power consumption.

One of the main functional systems of a satellite is its attitude determination and control system (ADCS) An important component of the ADCS is a system of sensors that provides the information on the spacecraft sensors that provides the information on the spacecraft's relative position in space. Satellites usually employ several types of sensors working on different physical principles.

The star tracker [1] receives an image of a starry arch fragment using a lens on an image sensing array, incorporated in the sensor. Search for star images is carried out in the obtained frame. The coordinates of the centers of the detected stars are estimated on the sensor array. After that, configuration of the stars is matched with the stars from the onboard star catalog. Using the matching stars, the sensor's attitude is calculated relative to the inertial coordinate frame fixed to the stationary stars. The star trackers make it possible to obtain the satellite's attitude with accuracy up to several arc seconds. The disadvantages of a star tracker are its large weight and size, expensiveness, array sensitivity to exposure to the Sun or other bodies; and lower quality

of attitude measurements during rapid rotation of the satellite.

The solar sensor [2] determines the directions to the Sun. The sensor provides the information in the form of unit vector projections on the axes of the sensor-fixed coordinate frame. Its operation is based on the conversion of information on the photodetector array exposure to the sunlight. The solar sensors have small size and weight, and low power consumption. However, the working area of the solar sensor is limited to the illuminated side of the orbit.

The magnetometer is intended for measuring the magnetic field induction vector. Since the magnetic field of the Earth is well-studied and there are precise models of average global magnetic field, the readings of magnetometers make it possible to determine the satellite's attitude relative to the Earth's magnetic field lines. The disadvantages of this type of sensors include their susceptibility to the distortions caused by the magnetic field induced by the onboard equipment. In addition, the readings of magnetometer depend on the sensor's housing temperature which varies on the illuminated and shadow sides of the orbit.

As for the infrared (IR) radiation sensors, they are small in size, immune to the electromagnetic field effect, and can operate on any side of the orbit.

The theoretical basis of the research consists of the works describing the experience in designing a system for the Earth's IR horizon detection. These research works [3–7] describe the principle of the IR sensor operation and its mathematical model. In their works, the authors consider different configurations of sensors to determine the nadir vector. Generally, it is proposed to solve the attitude problem by locating a number of sensors (4 to 9 ones) on one side of the satellite. The authors present the results of mathematical modeling and laboratory tests. The accuracy of the attitude measurements in these research works reaches about 0.5 deg. However, such a configuration strongly limits the field of view. Another disadvantage of the method proposed in the literature is that there is no open access to the results of in-flight experiments.

II. MATHEMATICAL MODEL OF NADIR FINDING SYSTEM (NFS)

A. IR temperature sensor

An IR temperature sensor [8] has several thermocouple elements connected in series to form a thermopile. The "hot" junctions of the thermocouple elements are connected to a thin IR radiation absorber on a silicon chip (Fig. 1, left). Due to the IR radiation exchange between the absorber and the object located in front of it, the temperature of the absorber increases or reduces, depending on the difference between its temperature and that of the object. Small mass of the absorber ensures quick thermal balance with the object, and its small thickness provides thermal insulation from the basic material of the chip. As a result, temperature gradient occurs between the absorber and the main part of the chip. The "cold" junctions of the thermocouple elements are located in the main part. The voltage generated on one thermocouple element is proportional to the temperature difference between two junctions. Total voltage of the thermopile is equal to the sum of voltages on all individual thermocouple elements. The temperature of the object is calculated from the sensor's array temperature measurements and the output voltage of the thermopile.

In our research we used the MLX90614 IR sensor made by Melexis (Fig. 1, right). This sensor has a wide field of view (130 deg.), and a wide range of operating temperatures.

Fig. 1. Diagram of the thermopile structure (right) and photo of IR sensor model (left)

B. NFS composition

NFS consists of six IR sensors fixed on six sides of the satellite. Each sensor records the integral temperature of the objects within its field of view. In the low near-Earth orbits, the Earth is the main object radiating in the IR spectrum. Therefore, the recorded temperature directly depends on the portion of the sensor's field of view, occupied by the Earth. This, in turn, can be interpreted as an angle between the sensor's sight axis and the direction to the nadir, i.e., the nadir angle. The presence of at least three nadir angles makes it possible to unambiguously determine the coordinates of the nadir vector in the coordinate frame fixed to the nanosatellite's body. Thus, the angles of the local vertical are determined in two steps. At the first step, the temperatures from the IR sensors are converted into the nadir angles; and at the second step, the nadir vector is calculated in the fixed coordinate frame.

C. Calculation of the nadir angle

In order to simplify the calculation of the nadir angle, we introduce the following assumptions:

- the IR radiation of the Earth is uniform in the spectrum under consideration. The influence of atmospheric and seasonal phenomena is neglected;

- the IR radiation of the space is constant. The influence of the Sun, the Moon and other celestial objects is negligible;

- the Earth has a spherical shape. The horizon observed by the sensor has a sharp circular edge.

Considering these assumptions, the temperature readings of the sensor can be deemed proportional to the area occupied by the Earth in the sensor's field of view (Fig. 2). The area is modeled as a region S where the projections of the sensor's field of view and the Earth onto a unit sphere with the center on the spacecraft overlap. Both projections are circles with known angular dimensions. The projection of the sensor's field of view has a radius R_s and the Earth disc projection has a radius R_e. The center of the sensor's field of view projection onto the unit sphere is the direction of the sensor's sight axis, and the center of the Earth disc projection onto the unit sphere is the direction of the nadir vector. The angle between the nadir vector and the sensor's sight axis is denoted as α. When $\alpha \geq R_e + R_s$, two circular projections do not overlap, which corresponds to the case when the sensor has not detected the Earth. When $\alpha \leq R_s - R_e$, the sensor's field of view is fully occupied by the Earth. The Earth's horizon is detected by the sensor when $R_s - R_e \leq \alpha \leq R_s + R_e$). The overlap area S in this range can be calculated as a function of α, R_s, R_e. For known values of R_s and R_e, this relationship helps to convert the IR sensor readings into the nadir angle. Thus, we have the dependency of the temperature readings of the sensor on the nadir angle. The inverse relationship can be obtained with large computational costs, so the nadir angle is calculated by interpolation, using a pre-calculated temperature characteristic (1).

Fig. 2. Projections of sensor's field of view and the Earth on a unit sphere.

$$T_s(\alpha) = \int_0^{\theta_0} \int_0^{2\pi} \frac{w(\theta)T(\theta,\varphi,\alpha)}{\Omega} d\theta d\varphi, \qquad (1)$$

where Ω is the area of the projection of the sensor's field of view onto the unit sphere;

θ_0 is the half-width of the sensor's field of view;

θ is the radial angle on the sensor's field of view;

φ is the azimuthal angle of the sensor's field of view;

$T(\theta, \varphi, \alpha)$ is the temperature at a point in the sensor's field of view.

D. Calibration of IR sensor's directional pattern (SDP)

Equation (1) contains a variable $w(\theta)$ which describes the non-uniformity of the sensor's readings in the field of view. This non-uniformity can be associated with the dependence of IR radiation intensity on the angle of incidence on the detector. Also, the distortions can be contributed by the light filter located within the sensor housing. Figure 3 shows a diagram of the sensor response depending on the radial θ and azimuthal φ angles. The darker the shade, the larger contribution is made by the given zone of the field of view in the sensor's reading.

Fig. 3. Representation of the sensor's directivity pattern divided into the zones of constant sensitivity

The SDP helps to determine the contribution in the integral temperature, made by an object located at a certain angular distance from the sensor's vision axis. Manufacturers often provide the SDP for a particular type of sensors. However, to improve the performance, it is sensible to calibrate each sensor and calculate its directional pattern.

Calibration should be carried out using a reference object with a high temperature relative to the environment, and with small angular dimensions compared to the sensor's field of view. The experimental unit consists of the following:

- a "point" source of IR radiation (an IR lamp in this case);

- a rotary mechanism to change the attitude of the IR sensor;

- a three-axis gyro, according to which the mutual attitude of the sensor and the "point" source of IR radiation is calculated.

The sensor readings are normalized within the range from 0 to 1. The relationship obtained this way is used as a directional pattern in (1).

Figure 4 shows the directional pattern provided by the manufacturer (blue solid line) and the values obtained experimentally for three sensors. It can be seen that the discrepancy between the experimental and rated data is about 10% within the range of angles from 30 to 60 deg. This confirms the suggestion that calibration is essential.

Fig. 4. Comparison of experimental and standard SDP

The altitude of a low near-Earth orbit of a nanosatellite is 400–600 km. In this orbit, the angular dimensions of the Earth will be about 130–140 deg. The directional pattern of the MLX90614 IR sensor is 130 deg wide. With such parameters, the temperature characteristic of the sensor (1) will be as shown in Fig. 5. According to the graph, the temperature values change little in the region from 0º to 20º C. The reason is that the Earth is fully in the field of view of the sensor. The temperature values also change little in the region from 110º C and above. This is due to the low sensitivity of the sensor at the edges of the directional pattern. Thus, of all the measured values of nadir angles, only those belonging to the range from 20º to 110º C are used in further solution.

Fig. 5. Temperature characteristic

To summarize the description of the mathematical model of IR sensor operation, it can be concluded that the sensor's temperature readings are the function of the following values:

- object's temperature T_o which is assumed to be constant;

- ambient temperature T_a which is assumed to be constant;

- sensor's directional pattern $w(\theta)$ obtained from the laboratory experiments for each sensor;

- angular dimension of the Earth projection R_e on the sensor's field of view, which depends on the satellite's altitude above the Earth surface at the moment.

E. Calculation of the nadir vector

To unambiguously determine the nadir vector, it is necessary to have at least three nadir angles for different sides of the satellite. In a general case, the problem consists in finding the area of intersection of several cones. The axis of each cone coincides with the sensor's vision axis. The angle between the generatrix and the cone axis is the nadir angle. The geometrical interpretation of the problem is presented in Fig. 6, where $\vec{f_1}, \vec{f_2}, \vec{f_3}$ are the vectors of direction of the sensors' fields of view, $\alpha_1, \alpha_2, \alpha_3$ are the nadir angles, and $\vec{n}$ is the nadir vector.

Fig. 6. Geometrical interpretation of the problem of determining the nadir vector

Analytically, this problem can be described by set of linear equations (2):

$$\begin{pmatrix} f_{1x} & f_{1y} & f_{1z} \\ \cdots & \cdots & \cdots \\ f_{kx} & f_{ky} & f_{kz} \end{pmatrix} \begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix} = \begin{pmatrix} \cos(\alpha_1) \\ \cdots \\ \cos(\alpha_k) \end{pmatrix} \qquad (2)$$

where α_k is the nadir angle;

$\vec{f_k} = (f_{kx}, f_{ky}, f_{kz})$ is the vector of direction of the vision axis of the k-th sensor;

$\vec{n} = (n_x, n_y, n_z)$ is the nadir vector.

Since in the general case the number of valid nadir angles can be more than three, the set of equations (2) is to be solved by the least squares method.

The orbital coordinate frame in which the nadir vector is calculated is defined as follows. Axis y is directed along the radius vector; axis x lies in the orbit plane in the direction of motion; and axis z complements the system making it right-handed; the datum point is located in the nanosatellite's center of the mass. The nadir vector in the given frame has the coordinates (0, -1, 0). Considering the direction cosine matrix [9 (page 339, formula 4.4.1)], it is possible to obtain the roll (3) and pitch (4) angles:

$$\gamma = \arcsin(-n_x); \qquad (3)$$

$$\vartheta = arctg\left(-\frac{n_z}{n_y}\right). \qquad (4)$$

As has been mentioned above, only the sensors for which the nadir angles lie within the range from 20º to 110º C are used in the solution. In Fig. 7, the overlap areas of the field of views of sensors installed on six sides of the nanosatellite are depicted. The color scale designates the number of sensors that are able to register the IR radiation of the Earth in different positions of the latter relative to the sensor's vision axis. Thus, the selected arrangement of sensors ensures full coverage of the Earth.

Fig. 7. Number of sensors registering the IR radiation at different attitudes of the nanosatellite relative to the Earth

III. EXPERIMENTAL RESULTS

A. Laboratory experiment

To test the NFS operation, a test bench consisting of the following parts has been developed (Fig. 8):

- a homogeneous heated object simulating the Earth's warm disc;

- homogeneous cold background simulating the outer space;

- nanosatellite mock-up installed on a two-axis rotary mechanism to simulate the elevation and to rotate the satellite mock-up;

- a three-axis gyroscope to calculate the exact attitude of the mock-up.

Fig. 8.   Experimental layout

The ground test bench approximately simulated the visible disc of the Earth and the position of the satellite. The IR simulator of the Earth was a circle one meter in radius; it was assembled from infrared membrane heaters. Cold background was a concrete wall on which the Earth simulator was installed. The nanosatellite mock-up was installed on the rotary mechanism at the distance of 0.5 m. At this distance, the angular dimension of the Earth simulator matches the angular dimension of the Earth in a low near-Earth orbit.

The coordinate frame in which the attitude was calculated was defined as follows. The datum point was located on the vertical axis of the rotary mechanism rotation, in the nanosatellite's center of the mass. Axis y is directed from the center of the Earth's IR radiation simulator, axis x is directed vertically upward, and axis z complements the frame up to right-handed one. Before conducting the main part of the experiment, a test record was done to determine the temperature of the Earth's IR radiation simulator (T_o) and the ambient temperature (T_a). Figure 9 shows the temperature measurements obtained from the IR sensors during the on-ground experiment. Based on the results of the test experiment, the temperature of the Earth's IR radiation simulator was chosen to be 42º C, and the ambient temperature— 24º C.

Fig. 9.   Readings of IR sensors during bench tests

Figures 10–11 present the results of the NFS bench-testing. The mock-up was rotated about axis OX. The initial roll angle was equal to –90 deg; the pitch angle was zero. The initial heading angle was 10 deg.

Fig. 10. Roll angle calculated from the readings of IR sensors and gyroscopes

Fig. 11. Pitch angle calculated from the readings of IR sensors and gyroscopes

According to the results of the series of tests, the roll angle error was 3.7 deg, and the pitch angle error—4.6 deg. The main reason for the errors is too close distance between the nanosatellite and the IR radiation source simulator (about 0.5 m). This location of the nanosatellite mock-up affects the experiment results as follows. As has been mentioned above, the temperature characteristic (1) depends on the angular dimensions of the object being registered. When the distance between the sensor and the simulator changes, the angular dimensions of the simulator in the sensor's field of view change as well. Since it is impossible to locate all sensors on the rotation axis of the rotary mechanism, the sensors happen to be at different distances from the simulator when the mock-up attitude changes. This can be clearly seen in Fig. 10. At the time points 20, 40, 60, 140 and 160 s, there are spikes in the roll angle measurements, caused by a change in the set of IR sensors involved in the solution.

The influence of this effect on the experimental results can be mitigated in the following ways. Firstly,

the mock-up attitude change should be limited to the conical rotation about the vertical axis with inclination not exceeding 15 deg. Secondly, the temperature characteristic should be re-calculated for each sensor depending on its position relative to the Earth's IR radiation simulator.

It should be noted that nonuniform ambient temperature causes errors in determining the nadir angle. This problem can be solved by measuring the ambient temperature in advance and taking it into account when determining the temperature characteristic (1).

B. In-flight experiment

In August 2022, a nanosatellite jointly developed by STC LLC and one of the Russian universities was launched into a low near-Earth orbit. During the flight, a series of tests was performed to check the NFS operation.

To assess the NFS performance, an attitude determination subsystem (ADSS) was used, which consisted of a three-axis gyroscope, a three-axis magnetometer, illumination sensors, and a navigation receiver. The latter provided the information on the nanosatellite position in the inertial space, based on which it was possible to determine the vector of direction to the Sun and the vector of magnetic induction of the Earth in the orbital coordinate frame. The illumination sensors determined the vector of direction to the Sun in the nanosatellite body-fixed coordinate frame. The magnetometer indicated the values of the Earth's magnetic induction vector in the nanosatellite body-fixed coordinate frame. The resulting vectors were used in the TRIAD algorithm [10] to determine the attitude of the nanosatellite.

Figure 12 shows the results of the NFS and ADS operation. Variations in the local vertical angles calculated from the readings of IR sensors (blue curve) are similar to the ADSS readings (red curve). The error in determining roll and pitch angles was no more than 10 deg throughout the series of the tests.

To determine temperature characteristic (1), the angular dimensions of the Earth were calculated depending on the orbit altitude provided by the navigation receiver. The Earth's temperature in the IR spectrum was taken equal to –18ºC [11], and the outer space temperature—equal to –270ºC. The sensor readings proved to differ from the temperatures predicted theoretically. The maximum registered temperature was –9ºC, and the minimum one was –136ºC. In Fig. 13, the minimum and maximum temperatures obtained in each experiment are presented. The discrepancy between the experimental and theoretical data can be explained by the fact that the sensor is not designed for such low temperatures. The minimum measurable temperature indicated in the manufacturer's documentation for the sensor is –80º C. On the other hand, during the flight the temperature of the satellite body varied from –10º to 25ºC. Thus, the light filter of the sensor could become a source of "parasitic" IR radiation, distorting the final signal.

Fig. 13. The minimum and maximum temperatures registered by the IR sensor in each experiment

Analysis of the data received from the satellite revealed the dependence of the sensor readings on the presence of the Sun in the field of view. Figure 14 shows the readings of the IR and illumination sensors. At the time point 80 s, the normalized values of solar radiation intensity drop to zero. This corresponds to the situation when the satellite goes into the shadow of the Earth. At the same moment, the readings of the IR sensors previously directed towards the Sun (sides y+, z+) experience a sharp change.

Fig. 14. Dependence of readings of IR sensors on solar illumination

In Fig. 15, the behavior of the local vertical angles in the described experiment is shown. The NFS readings are compared to those of the gyroscopes. At the time

Fig. 12. Angles of the local vertical, calculated from the NFS (blue) and ACS (red) readings

point when the nanosatellite is crossing the terminator line, change in the behavior of the local vertical angles is observed. At the same time, the roll and pitch angles calculated from the gyroscope readings retain their dynamics. This fact confirms that the NFS operation depends on the solar illumination.

Fig. 15. NFS behavior at the moment of crossing the terminator line

It is also worth mentioning that the difference between the readings of IR sensors on the shadow and Sun-illuminated sides of the orbit is 10ºC on average.

Based on the observations described above, the influence of the Sun should be taken into account in the mathematical model of IR sensor. The effect of the Sun can be compensated by calculating the temperature characteristic for each sensor in real time. In this case, characteristic (1) will take into account the nanosatellite's position in the orbit (shadow or illuminated side) and the presence of the Sun in the sensor's field of view.

IV. CONCLUSION

In the course of work, a possibility of using IR temperature sensors to create a system for determining the local vertical has been studied. A mathematical model of IR sensor has been developed. The main parameters affecting the IR sensor's readings have been identified. The sensors have been calibrated.

The NFS performance was studied by on-ground laboratory testing and during nanosatellite operation in the near-Earth orbit. As a result of in-flight trials, the readings of sensors registering the IR radiation of the outer space and the Earth have been obtained. It has been found that the sensor readings depend on the solar illumination. The angles of the local vertical have been obtained with accuracy up to 10 deg. With regard to the accuracy of determining the angles of the local vertical, it should be noted that the tests were not aided by any reference sources of attitude, such as star tracker or solar sensor. In our study case, the accuracy was estimated by the ADSS data where the error of attitude angles is about 5 deg.

In the future scheduled missions, the nanosatellites will be equipped with a precision attitude determination and control system based on a star tracker, which will improve the calculation of the NFS performance accuracy.

REFERENCES

[1] M. E. Prokhorov, A. I. Zakharov, A. V. Mironov, F. N. Nikolaev, andM. S. Tuchin, "Modern star trackers" Proc. 38th International Students' Scientific Conference "Physics of Space", Ural Federal University, 2009, pp. 170–186.

[2] G. Rufino and M. Grassi, "Multi-Aperture CMOS Sun Sensor for Microsatellite Attitude Determination", Sensors, 2009, vol. 9(6), pp. 4503-4524.

[3] S. W. Janson, B. S. Hardy, A. Y. Chin, D. L. Rumsey, D. A. Ehrlich, and D. A. Hikley, "Attitude control on the Pico Satellite Solar Cell Testbed-2", Proc. 26th Annual AIAA/USU Conference on Small Satellites, , Logan, UT, USA, August 13–16, 2012.

[4] T.Kuwahara, K. Fukuda, N. Sugimura, Y. Sakamoto, K.Yoshida, A. Dorsa, P.Pagani, and F. Z. Bernelli. "Design and Implementation of a Thermopile-Based Earth Sensor", Transactions of the Japan Society for Aeronautical and space Sciences, Aerospace Technology Japan, 2016, vol. 14, no. ists30, pp. 77–88.

[5] A. G. Sáez, J. M. Quero, and M. Angulo, "Earth Sensor Based on Thermopile Detectors for Satellite Attitude Determination", IEEE Sensors J., 2015, vol. 16(8), pp. 2260–2271.

[6] T. Nguyen, K.Cahoy, and A.Marinan, "Attitude Determination for Small Satellites with Infrared Earth Horizon Sensors", Journal of Spaecraft and Rockets, 2018, vol. 55(1), pp. 1–10.

[7] I.Vertat, R. Linhart, J. Masopust, A.Vobornik, and L. Dudacek, "Earth's thermal radiation sensors for attitude determination systems of small satellites", Contributions of the Astronomical Observatory SkalnatePleso, 2017, vol. 42(2), pp. 157–164.

[8] A. Graf, M. Arndt, M. Sauer, and G. Gerlach, "Review of micromachined thermopiles forinfrared detection", Measurement Science andTechnology, 2007, vol. 18(7), pp. R59–R75.

[9] G. I. Emel'yantsev and A. P. Stepanov, Integrirovannye inertsial'no-sputnikovye sistemy orientatsii i navigatsii (Integrated Itertial-Satellite Systems of Orietations and Navigation), St. Petersburg: Concern SRI Elektroprior, JSC, 2016.

[10] F. L. Markley, "Attitude determination using two vector measurements", Flight Mechanics Symposium, Greenbelt, MD, USA, May 1999, NASA Conference Publication NASA/CP-19989-209235, pp. 39–52.

[11] J. Harries, B. Carli, R. Rizzi, C. Serio, M. Mlynczak, L. Palchetti, T. Maestri, H. Brindley, and G. Masiello,"The far-infrared Earth",Reveiws of Geophysics, 2008, vol. 46(4), p. RG4004.

Technique for Detecting and Determining Ship Motion Parameters from Images of Wave Wakes in Stormy Conditions

P.K. Kuznetsov

Samara State Technical University,
Samara Scientific Center RAS
Samara, Russia
kurnesov@mail.ru

B.V. Martemyanov

Samara State Technical University,
Samara, Russia
bvmart@rambler.ru

Abstract — **The problem of detecting and determining the motion parameters of ships from images of wave wakes they create on a sea surface in conditions of heavy storm is considered. In order to solve the problem, video remote monitoring tools, operating in video and near-infrared ranges of electromagnetic radiation, are used to acquire a sequence of images of the scene of interest and, using a special procedure of fragmentary registration of the acquired images, pick out the wake waves in the images against a background of wind waves. A special filtering procedure is designed to determine the speed of motion of the detailed wake waves. Based on the analysis of the motion of the wake waves, information about the velocity and direction of the ship's motion is determined. The parameters of the proposed filter depend significantly on the values of the temporal and spatial characteristics of the wake and wind waves. The proposed registration technique is invariant to variations in density and contrast of the acquired images.**

Keywords— ship, image frames, detection, wave, wakes, velocity

I. INTRODUCTION

The task of detecting and determining ship motion parameters from wake wave images generated by ships in strong storm waves arises when searching for ships that are without radio communication, not registered in the AIS system or not interested in being detected and, therefore, pose a threat to shipping traffic. There is extensive literature on this subject. For example, Reference [1] considers the physical mechanisms of the formation of wakes and their imaging on optical and radar images, Reference [2] studies the process of formation of wakes and Kelvin waves from large-displacement ships, Reference [3] investigates the occurrence of turbulent vortices in wake waves. In [4–6], the authors of this paper propose a general method for calculating the vector field of motion velocities (optical flow) of images in the presence of random noise. The method is based on the procedure of high-precision matching (registration) of a sequence of image frames, and estimates of the error of such matching are obtained.

II. CHARACTERISTICS OF WAKE WAVE IMAGE RECOGNITION IMAGES ACQUIRED IN HEAVY STORM WAVES

The solution to the problem proposed here is based on the acquisition of a sequence of images of the scene of interest by remote sensing tools and the subsequent target filtering of the obtained images. Filtering is necessary because wind disturbance creates serious interference with the detection of wake waves. The technique of target filtering is based on the procedure of frame matching of the acquired images by fragments. Filtering is necessary because wind turbulence can cause serious interference in the detection of wake waves. To filter the sequence of images with wake wave images, we take into account the fact that there is a significant difference in the lifetime of wind waves and wake waves is used. Prior to filtering, preliminary estimates are obtained for the values of the correlation radius of the wind waves in space and the correlation radius of the lifetime of these waves. Wind waves live for units of seconds, while wake waves live for many tens of minutes. This fact is used here to detect and pick out wakes against a background of wind waves. Namely, it is proposed to obtain a sequence of images of the area of interest with a time interval between frames T, the value of which significantly (5-7 times) exceeds the correlation radius of the wind wave lifetimes, and then to perform the procedure of dense fragment-by-fragment registration of the obtained image frames. The size of the fragments is set so that it does not exceed the radius of spatial correlation of the wind waves.

Since the image fragments containing wind wave images are not correlated, it is highly unlikely that they will be superimposed. However, the fragments of wake wave images, which have a much larger lifetime correlation radius than wind waves and therefore retain their shape for a long time, are very likely to be superimposed.

There are inherent problems in combining wake wave images. For weak waves (no stronger than a 3–4-point storm), wake wave images have long stretches of linear shape. In this case, there is uncertainty in finding compatible fragments and, consequently, in determining the wave velocity by the matching method. In a weak storm, such linear regions are present over a large area of the turbulent wave images, which does not allow us to obtain sufficiently accurate estimates of the travel speed in these areas. The case of the heavy storm waves considered here is characterized by strong curvature and fracture of the wake waves due to the storm. This makes it possible to combine images of waves in the areas of curvature and fracture with a sufficiently small error, even in the presence of significant noise in the video data acquisition path and significant foreshortening distortions caused by changes in the imaging system's field of view when recording from a mobile base.

The result of combining pairs of consecutive images is a new image sequence containing only images of a travelling wave dynamically changing its position from frame to frame. Repeated application of the superposition operation to the image sequence synthesized in this way (containing only images of the wake wave) will allow us to obtain estimates of the true wave velocity vector. In turn, the obtained estimate of the wake velocity value will allow the estimation of the speed and direction of the ship's motion, even if there is no ship in the image.

An original procedure for high-precision image matching is outlined below. The procedure is based on the mathematical apparatus of distributions. The transition to the space of distributions allows us to correctly perform matching of foreshortened distorted images at low signal-to-noise ratio.

Figure 1 shows an example of a cargo ship wake image obtained at 8m resolution per pixel in 5-point storm conditions. The wake generated by the ship's propulsion and the divergent Kelvin waves are clearly visible.

III IMAGE REGISTRATION METHODOLOGY

It is proposed to solve the problem of detection and determination of the parameters of the wake motion using the procedure of fragment-by-fragment matching of a sequence of images of the scene of interest obtained by an Earth Remote Sensing (ERS) system. The matching procedure is performed, as mentioned above, in the space of distributions.

The translation of the problem into the space of distributions allows us to develop algorithms oriented towards real-time implementation for the matching of image fragments with low values of signal-to-noise ratio and significant relative displacements (up to 100 or more pixels). Furthermore, this approach allows to obtain a correct estimation of the matching error, which is inaccessible to other matching methods, such as optical flow or extreme correlation methods.

Let us suppose that the fragments of two images of the same scene, taken by the imaging system with an adjustable value of the time interval between the image frames (and possibly with a changed shooting angle), are combined. Let's set the time interval by condition:

$$Twk \gg T > Twind, \tag{1}$$

where Twk – correlation radius of the wake wave lifetime,

$Twind$ – correlation radius of the wind wave lifetime.

We will present the images (video signals) received by the left and right video cameras (L and R), taking into account the presence of random noise in the video frames, in the form (2), and will mark the video signals and their components with markers L and R: $\tilde{E}_L(x, t)$, $\tilde{E}_R(x, t)$:

$$\tilde{E}_\alpha(x, t) = E_\alpha(x, t) + h_\alpha(x, t), \qquad \alpha \in \{L, R\}, \tag{2}$$

where $x = (x, y)$; $E_\alpha(x, t)$ – video signal α with a noise $h_\alpha(x, t)$, we will assume that the noise in the video signals L and R are not correlated and that random functions $E(x, t)$, $h(x, t)$ are ergodic and mutually noncorrelated.

Fig. 1.   Image of the wake of a cargo ship at the mouth of the Pearl River, China; received by the 'GF-1' satellite. The wake waves and the cuspid (pointed) waves framing them are clearly visible. The diverging Kelvin waves are faintly visible. Image resolution 8m. per pixel. Storm surge 5 points

The random noise function $h(x, t)$ is centered and the conditions for the differentiability of its implementations (in the generalized sense) with respect to all its arguments are satisfied.

Let us define a distribution that we use in the registration process in the form of functional (3) (here, we omit the lower index that determines the left or the right image):

$$\tilde{F}(\tilde{E}(x, t), t) = \iint_D K(x)\, \tilde{E}(x - s(x, t), t) dx dy, \tag{3}$$

where $s(x, t) = (v(x, t), u(x, t))$ – the desired two-component vector of relative image displacement, $K(x)$ – the kernel of the functional, chosen according to the criterion of minimum computational complexity, has the form of a linear combination of regular pyramids; D – analysis window with a border without features.

The task of combining image fragments is to find, in the image R, such a fragment (conjugate) Ξ_R, that, in a certain sense, best matches some pre-selected fragment Ξ_L in the image L.

Since equation (3) contains two unknown components of the velocity vector, it is necessary to supplement this equation with other equations to obtain a solvable system of equations.

The additional equations are obtained as follows. Let us select a certain analysis area containing the fragment Ξ_L. in the image L and cover it with a system of windows Di, whose size does not exceed the spatial interval of the correlation of the wind waves. We will carry out a similar covering of the image R. Each window covers a certain fragment of the image, and then we organise an iterative process of combining adjacent fragments. The procedure is as follows

For each window D^i, by taking a derivative with respect to time, we obtain a relation of the form (4). That is the equation that functionally relates the parameters of images (2) with the relative displacement $v(x, t) = (v, u)$ of these images.

This equation (let's call it the FC equation) has the form:

$$\Delta \tilde{F} \cong - F_x(v) - F(v)_x, \tag{4}$$

where $F_x(s) = \iint_D \nabla(K(x))v(x, t_0) E(x, t_0)dxdy,$

$$F(v)_x = \iint_D K(x)E(x, t_0) \, div \, v(x, t_0) \, dxdy.$$

The proposed form of the FC equation eliminates the problem of differentiation by spatial coordinates the video signal (2) which is, in generally, non-differentiable.

On the coverage $\langle D^i \rangle$ of the neighborhood of some point x of the image $E(x, t)$ a FC equation of the following form is constructed:

$$\tilde{\mathbf{F}}_x^{\mathrm{T}}\dot{\tilde{\mathbf{F}}} = \tilde{\mathbf{F}}_x^{\mathrm{T}}\tilde{\mathbf{F}}_x\hat{v} , \qquad (5)$$

where $\dot{\tilde{\mathbf{F}}}, \tilde{\mathbf{F}}_x$ we obtain from (2, 3) by replacing the image $E(x)$ by the video signal $\tilde{E}(x,t)$; and $\hat{v} = v + \Delta v = [\hat{v} \ \hat{u}]^{\mathrm{T}} -$ column vector of the estimation for the vector v ; $\Delta v = [\Delta v \ \Delta u]^{\mathrm{T}} -$ the column vector of the estimation for an error in the calculation vector v.

From relations (4) and (5) we obtain FC equation:

$$\sum(H_t F_x + H_t H_x) = \sum(H_x)^2 + \Delta v \sum(\tilde{F}_x)^2 + v\sum(F_x H_x)+$$

$$+u\sum(F_x H_y + H_x H_y)+\Delta u\sum(F_x F_y + H_x H_y + F_x H_y + F_y H_x), (6)$$

where $H_\alpha^i = \iint_{D_i} K_\alpha^i(x) \, h(x, t) \, dxdy,$

$H_t^i = \iint_D K^i(x)h_t(x, t) \, dxdy; \alpha \in \{x, y\}.$

In (6), the summation is performed over all windows of the coverage $\langle D^i \rangle$.

It is shown that from the plausible assumption that the functions $\tilde{F}_x^i$ и $\tilde{F}_y^i$ as well as H_x^i и H_y^i are not correlated as a random function of the number of analysis windows in the coverage $\langle D^i \rangle$, the following estimate for the error Δv in the calculation of the velocity vector follows:

$$\Delta v \simeq -\,\mathbf{NSR}\,v, \qquad (7)$$

where $\mathbf{NSR} = [NSR_x \ NSR_y]$ – a row vector whose components characterise the noise/signal ratio:

$$NSR_x = (H_x)^2 \left(\sum(\tilde{F}_x)^2\right)^{-1},$$

$$NSR_y = \sum(H_y)^2 \left(\sum(\tilde{F}_y)^2\right)^{-1}.$$

The following possible estimation for the values of the functionals H_α takes place:

$$H_\alpha = F_\alpha - F_\alpha(E(x + s)), \ \alpha \in \{x, y\}. \qquad (8)$$

The estimate of the relative displacement vector obtained as a result of solving system (5) is used in the iterative procedure at the current iteration step as a value of a displacement of the fragment Ξ_R. The search for matching fragments is stopped when the registration error is less than a specified threshold value.

It follows from (7) that the value of the resulting image velocity error estimate is proportional to the motion velocity itself, i.e. a "velocity error" occurs. The velocity error can be reduced to zero by accurately aligning adjacent image fragments.

Fig. 2 shows an example of detecting a moving cargo ship from a wake image. Fig. 3 shows an example of the detection of an aircraft carrier and accompanying escort vessels. Figure 3 is supplemented by the velocity histograms of the wake and Kelvin waves generated by the cargo ship. The velocity histogram is constructed by calculating the velocities of image fragments using the proposed method.

Analysis of various aerial images shows that the area of the Kelvin wake is not large and does not represent more than 40% of the total area of the waves generated by the cargo ship. The visibility of internal waves generated only by cargo ships is extremely low due to the difficult conditions of their formation.

The ability to detect turbulent wakes is closely related to the spatial resolution of the image.

For example, low-resolution imagery is more suitable for ship detection, and ship wakes can be fully exploited to obtain ship information from high-resolution imagery.

The actual characteristics of ship wakes depend on the shape, speed and draught of the ship, as well as water depth, wind speed and direction, and surrounding wind waves, and their wakes are difficult to identify in low-resolution imagery. The difference between the reflectivity of the hull body and the surrounding water is greater and, strange as it may seem, liner elements are easier to distinguish in low resolution imagery than in high resolution imagery.

The existing Earth remote sensing spacecraft, such as the Aist-2 type, provide a great deal of data for detecting and recognising ships by the tracks they leave on the water surface. As an example of remote sensing imagery, Figure 2 shows images provided by the Aist-2.

Kelvin waves provide additional information about the ship's speed. The Kelvin angle opening range helps to determine a merging point of the Kelvin wave arms. The shift of the merging point at the frame time interval provides an information about ship's speed. Unfortunately, Kelvin waves are poorly observed and decay quite quickly. It is therefore more reliable to determine the parameters of the ship's speed vector by analyzing the characteristics of the wake.

Turbulent wakes should be prioritized according to their probability of occurrence, which is determined by the presence or absence of storm conditions in the observed water areas. High resolution imagery is more suitable for detecting bright, turbulent areas. An example of wake detection is shown in Fig. 3. The ship's speed can be determined from the motion velocity of the entire wake as an integral structure.

a)

Detection of the aircraft carrier.
(Possibly at a low speed)

Escort ship of the aircraft carrier

b)

Wake velocity mode

Kelvin wave velocity mode

Detection of the aircraft carrier escort ship by its wake

Fig. 3. An example of the acquisition of an image of a wake wave generated by an aircraft carrier and a destroyer escorting the aircraft carrier, with an image resolution of 8 m per pixel, in storm conditions of 5-6 points, with the construction of a histogram of the wake wave velocity

The authors intend to carry out experimental studies to confirm the main results of this work.

c)

Fig. 2. Images a), b). container ship imaged by GF-1 satellite in the coastal waters of Guangdong Province, China; image resolution is 16 m per pixel; c) the image synthesized by registration images a) and b).

REFERENCES

[1] Ermakov, S., I. Kapustin, and T. Lazareva, 2014: Ship wake signatures in radar/optical images of the sea surface: Observations and physical mechanisms. Proc. SPIE, 9240, 92400N,
    https://doi.org/10.1117/12.2067367.

[2] Gilman, M., A. Soloviev, and H. Graber, 2011: Study of the far wake of a large ship. J. Atmos. Oceanic Technol., 28, 720–733,
    https://doi.org/10.1175/2010JTECHO791.1.

[3] Golbraikh, E., A. Eidelman, and A. Soloviev, 2013: On helical behavior of turbulence in the ship wake. J. Hydrodyn., 25B, 83–90,
    https://doi.org/10.1016/S1001-6058(13)60341-8.

[4] Kuznetsov, P. K. New Method of Optical Flow Evaluation / P. K. Kuznetsov, B.V. Martemyanov // Proc. 24th Saint Petersburg International Conference on Integrated Navigation Systems 29 - 31 May 2017. - Saint Petersburg, Russia. - 2017. P. 27-31. ISBN 978-5-91995-044-8.

[5] Anshakov, G.P. and other. Recovery of Non-Visual Data from Images Acquired by Land-Survey Satellites/ G.P. Anshakov, G.N. Myatov, A.A. Yudakov, P. K. Kuznetsov, B.V. Martemyanov// Proc. 25th Anniversary Saint Petersburg International Conference on Integrated Navigation Systems 28 - 30 May 2018. - Saint Petersburg, Russia. – 2018. – P. 480-487.

[6] Kuznetsov, P.K., Technical vision of moving objects. Method for analysing the velocity field of a dynamic image / P.K. Kuznetsov, B.V. Martemyanov, V.I. Semavin // M.: Vestnik Computernykh i Informatsionnykh Tekhnologii. - 2014. – No. 1. - C. 3-9.

Indoor High Dynamic Positioning Technology of UWB/MIMU Integrated Navigation System with Extended Kalman Particle Filter

Chunfeng Gao

College of Advanced Interdisciplinary Studies, National University of Defense Technology; Nanhu Laser Laboratory, National University of Defense Technology Changsha,China neil1989@126.com

Wanqing Liu

1.College of Advanced Interdisciplinary Studies, National University of Defense Technology 2. Nanhu Laser Laboratory, National University of Defense Technology Changsha,China 131142114425@163.com

Guo Wei*

1.College of Advanced Interdisciplinary Studies, National University of Defense Technology 2. Nanhu Laser Laboratory, National University of Defense Technology Changsha,China nudtweiguo@163.com

Chengzhi Hou

1.College of Advanced Interdisciplinary Studies, National University of Defense Technology 2. Nanhu Laser Laboratory, National University of Defense Technology Changsha,China 821146885@qq.com

Wenjian Zhou

1.College of Advanced Interdisciplinary Studies, National University of Defense Technology 2. Nanhu Laser Laboratory, National University of Defense Technology Changsha,China tom7850@foxmail.com

Jiayi Cheng

1.College of Advanced Interdisciplinary Studies, National University of Defense Technology 2. Nanhu Laser Laboratory, National University of Defense Technology Changsha,China 1774394386@qq.com

Xu Zhu

1.College of Advanced Interdisciplinary Studies, National University of Defense Technology 2. Nanhu Laser Laboratory, National University of Defense Technology Changsha,China zhuxu22@nudt.edu.cn

Xudong Yu

1.College of Advanced Interdisciplinary Studies, National University of Defense Technology 2. Nanhu Laser Laboratory, National University of Defense Technology Changsha,China wind0909@163.com

Abstract—To improve the real-time positioning accuracy of indoor high dynamic carriers, an integrated navigation system based on EKPF(Extended Kalman Particle Filter) of UWB and MIMU is proposed in this paper. A tightly coupled framework is adopted to construct a nonlinear system model for UWB/MIMU data fusion, and a UWB/MIMU data fusion technology based on EKPF algorithm is proposed. The UWB/MIMU integrated navigation system has been developed and installed on unmanned aerial vehicles. To validate the effectiveness of the system, the high dynamic characteristics of drones were utilized to verify the positioning performance of the system in indoor NLOS (Non Line of Sight) environments. The experimental results show that the horizontal positioning accuracy of the system is better than 0.1m. Compared to pure UWB positioning systems, the accuracy has increased by 31%. The experimental results demonstrate that the proposed UWB/MIMU integrated navigation system achieves higher positioning accuracy and robustness compared to a pure UWB positioning system. The application of related technologies will provide solutions for high dynamic positioning of carriers in indoor buildings, large shelters, and large underground mines.

Keywords—Ultra-Wideband, Micro inertial measurement unit, Particle Filter, Indoor Positioning System

I. INTRODUCTION

UWB technology is a wireless carrier communication technology that uses a frequency bandwidth of more than 1GHz, which has strong penetration, low power consumption, good anti-multipath effect, high security, low system complexity, and can provide accurate positioning accuracy. It has a wide range of application scenarios in many fields, and it also provides the possibility to track the specific location of the target within a certain range.

In indoor or mining environments where GNSS(Global Navigation Satellite System) signals cannot be received, the positioning technology of high dynamic carriers in NLOS (Non Line of Sight) situations is a hot research topic in the field of navigation. A combined navigation and positioning solution based on UWB (Ultra-Wideband) and MIMU (Micro Inertial Measurement Unit) is capable of providing

real-time high-precision positioning information for the carriers, make the positioning of indoor high dynamic carriers possible by combining the technical advantages of inertial navigation with highly autonomous full parameter output and UWB ranging accuracy and strong communication capabilities.

II. FRAMEWORK OF UWB/MIMU INTEGRATED NAVIGATION SYSTEM

In the UWB positioning system, the number of UWB base stations is at least 3 to achieve 2D positioning of tags, and the number of UWB base stations is at least 4 to achieve 3D positioning of tags. In this paper, the tag is attached to the UAV platform, and the three-dimensional positioning of the tag needs to be realized, so the number of UWB base stations in the UWB positioning system is at least 4.

The spatial layout of the base station should ensure that the projection length of the base station in the XYZ axis is equal as much as possible, that is, all sides of the polyhedron with the base station as the vertex are equal. However, the base station layout can easily ensure that the XY axis projection length in the horizontal plane is equal, but in the Z-axis direction, due to factors such as the erection cost, it cannot be guaranteed that the Z-axis projection length is equal to the XY axis projection length, and the final base station layout is often flat, which will lead to the low positioning accuracy of the pure UWB positioning system in the Z-axis direction. Considering the above influencing factors, and improving the positioning accuracy of the system on the basis of controlling the cost as much as possible, the UWB positioning system used in this paper is composed of 6 base stations and 1 tag, and the base station layout is shown in Fig. 1.

Fig. 1. Schematic diagram of system layout

In the UWB positioning system, according to the right-hand system, the coordinate system is arranged in the form of ENU (Northeast Sky), in which the positive direction of the X-axis points to the east, the positive direction of the Y-axis points to the north, and the positive direction of the Z-axis points to the sky, and the coordinate system is named the navigation system to distinguish it from the carrier system of the UAV platform. Among them, A0, A1, A2, A3, A4, and A5 are UWB base stations, and T0 is a UWB tag, which is installed on the UAV platform. $A_0(0,0,z_l)$, $A_1(0,y_1,z_l)$, $A_2(x_2,y_1,z_l)$, $A_3(x_2,0,z_l)$, $A_4(0,y_4,z_h)$, $A_5(x_2,y_4,z_h)$ are the coordinates of each base station, $T_0(p_x,p_y,p_z)$ is the coordinates of the tag, The true distance between each base station and the tag is indicated separately by d_0,d_1,d_2,d_3,d_4,d_5,

The distance between each base station and the tag measured by the UWB positioning system is represented separately by r_0,r_1,r_2,r_3,r_4,r_5.

The UWB/MIMU integrated system adopts a tightly coupled framework, which does not require the UWB positioning system to output independent position information, but only needs to output the distance information from each base station to the tag. The position information of the INS output is calculated with the known coordinates of the UWB base station, and the distance from the label position of the INS output to each base station is calculated. The above information is used as input, and the fused navigation information is obtained through the UWB/MIMU data fusion algorithm, and the output process is shown in Fig. 2.

Fig. 2. Schematic diagram of the UWB/MIMU tight integrated systemframework

Because the close coupling only involves the processing of the ranging information output by the UWB positioning system, the navigation results of the independent UWB positioning system cannot be output. The advantages of tight coupling are: even if the individual base station in the UWB positioning system cannot output ranging information due to obstacles and other reasons, the integrated navigation system can still output navigation information, and there is no measurement correlation problem, and at the same time, under the same other conditions, compared with loose coupling, the combined navigation results output by tight coupling are more accurate and robust, therefore, this paper selects tight coupling as the framework of UWB/MIMU data fusion technology.

III. UWB/MIMU DATA FUSION ALGORITHM BASED ON EKPF

A. Filter equation of state for UWB/MIMU integrated navigation system based on EKPF

Although filtering algorithms such as Extended Kalman Filter (EKF) and Unscented Kalman Filter (UKF) can deal with nonlinear problems, they are still based on the assumption of Gaussian distribution, so they are not completely nonlinear non-Gaussian algorithms. So far, the true nonlinear non-Gaussian filter algorithm is the Particle Filter (PF) algorithm. The PF algorithm is a sequential Monte Carlo simulation method based on Bayesian estimation, and its core idea is to use some random samples, i.e., "particles", to represent the post-test probability density of random variables in the system, so as to obtain the approximate optimal numerical solution based on the physical model, instead of the optimal filtering of the approximate model, so it is suitable for the filtering of strong nonlinear non-Gaussian noise system models.

In order to solve the problem of particle degradation in the standard PF algorithm, and to consider the real-time requirements of navigation information in the process of autonomous take-off and landing of UAVs, the complexity of the algorithm should not be too high, and the importance density function is constructed based on EKF to move the particles to the high likelihood region, and the Extended Kalman Particle Filter (EKPF) algorithm is formed.

The filter state vectors are defined as follows:

$$X = \begin{bmatrix} p^{(n)} & v^{(n)} & q^{(n)} & b_f & b_\omega \end{bmatrix}^T \quad (1)$$

Among them, $p^{(n)} = \begin{bmatrix} p_x & p_y & p_z \end{bmatrix}^T$ is the position vector of the UAV under the system, $v^{(n)} = \begin{bmatrix} v_x & v_y & v_z \end{bmatrix}^T$ is the velocity vector of the UAV under the navigation system, $q^{(n)} = \begin{bmatrix} q_W & q_X & q_Y & q_Z \end{bmatrix}^T$ is the quaternion vector of the UAV attitude under the system. b_f is the zero-bias vector of the accelerometer, and b_ω is the zero-bias vector of the gyroscope.

Since the target is in low-speed motion and the positioning area is small, the influence of the earth's rotation and the earth's curvature can be ignored, and the equation of motion of inertial navigation can be simplified as follows:

$$\dot{X} = \begin{bmatrix} \dot{p}^{(n)} \\ \dot{v}^{(n)} \\ \dot{q}^{(n)} \\ \dot{b}_f \\ \dot{b}_\omega \end{bmatrix} = \begin{bmatrix} v^{(n)} \\ a^{(n)} \\ \frac{1}{2} q^{(n)} \otimes \begin{bmatrix} 0 \\ \omega^{(b)} \end{bmatrix} \\ -\frac{1}{\tau_f} b_f^b + \delta_f \\ -\frac{1}{\tau_\omega} b_\omega^b + \delta_\omega \end{bmatrix} = \begin{bmatrix} v^{(n)} \\ R_b^n \left(\tilde{f}^{(b)} - b_f - n_f \right) \\ \frac{1}{2} q^{(w)} \otimes \begin{bmatrix} 0 \\ \tilde{\omega}^{(b)} - b_\omega - n_\omega \end{bmatrix} \\ -\frac{1}{\tau_f} b_f^b + \delta_f \\ -\frac{1}{\tau_\omega} b_\omega^b + \delta_\omega \end{bmatrix} \quad (2)$$

Among them, $a^{(n)}$ it is the acceleration of the body under the **n** system, $\omega^{(b)}$ is the angular velocity vector under the **b** system, $\tilde{f}^{(b)}$ is the measured value vector of the accelerometer, $\tilde{\omega}^{(b)}$ is the measured value vector of the gyroscope. n_f and n_ω are Gaussian white noise of accelerometers and gyroscopes, respectively. $\otimes$ is quaternion multiplication. R_b^n is a rotation matrix from **b** to **n** series, which is calculated by the following formula.

$$R_b^w = \begin{bmatrix} 1 - 2q_Y^2 - 2q_Z^2 & 2q_X q_Y - 2q_W q_Z & 2q_X q_Z + 2q_W q_Y \\ 2q_X q_Y + 2q_W q_Z & 1 - 2q_X^2 - 2q_Z^2 & 2q_Y q_Z - 2q_W q_X \\ 2q_X q_Z - 2q_W q_Y & 2q_Y q_Z + 2q_W q_X & 1 - 2q_X^2 - 2q_Y^2 \end{bmatrix} \quad (3)$$

The recursive formula for the equation of state in discrete time is given by:

$$\begin{bmatrix} p_{k+1}^n \\ v_{k+1}^n \\ q_{k+1}^n \\ b_{f,k+1} \\ b_{\omega,k+1} \end{bmatrix} = \begin{bmatrix} I & I\Delta t & 0 & -R_b^n \frac{(\Delta t)^2}{2} & 0 \\ 0 & I & 0 & -R_b^n \Delta t & 0 \\ 0 & \Omega(-b_{\omega,k}) & 0 & 0 & 0 \\ 0 & 0 & 0 & e^{-\beta_1 \Delta t} & 0 \\ 0 & 0 & 0 & 0 & e^{-\beta_1 \Delta t} \end{bmatrix} \begin{bmatrix} p_k^n \\ v_k^n \\ q_k^n \\ b_{f,k} \\ b_{\omega,k} \end{bmatrix}$$

$$+ \begin{bmatrix} R_b^n \frac{(\Delta t)^2}{2} & 0 \\ R_b^n \Delta t & 0 \\ 0 & q_k^n \otimes \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{f}_b \\ \tilde{\omega}_b \end{bmatrix} + \begin{bmatrix} R_b^n \frac{(\Delta t)^2}{2} & 0 & 0 & 0 \\ R_b^n \Delta t & 0 & 0 & 0 \\ 0 & 0 & 0 & I \\ 0 & I & 0 & 0 \\ 0 & 0 & \Delta t \frac{Q_{2:4}}{2} & 0 \end{bmatrix} \begin{bmatrix} n_f \\ n_\omega \\ \delta_f \\ \delta_\omega \end{bmatrix} \quad (4)$$

B. Filter equation of measurement for UWB/MIMU integrated navigation system based on EKPF

According to the geometric relationship between the base station and the tag in the UWB positioning system in Fig. 1, the measurement equation can be obtained as follows:

$$\begin{bmatrix} r_0(k+1) \\ r_1(k+1) \\ r_2(k+1) \\ r_3(k+1) \\ r_4(k+1) \\ r_5(k+1) \end{bmatrix} = \begin{bmatrix} \sqrt{A^2 + B^2 + (C - z_l)^2} \\ \sqrt{A^2 + (B - y_1)^2 + (C - z_l)^2} \\ \sqrt{(A - x_2)^2 + (B - y_1)^2 + (C - z_l)^2} \\ \sqrt{(A - x_2)^2 + B^2 + (C - z_l)^2} \\ \sqrt{A^2 + (B - y_4)^2 + (C - z_h)^2} \\ \sqrt{(A - x_2)^2 + (B - y_4)^2 + (C - z_h)^2} \end{bmatrix}$$

$$+ \begin{bmatrix} \upsilon_0(k+1) \\ \upsilon_1(k+1) \\ \upsilon_2(k+1) \\ \upsilon_3(k+1) \\ \upsilon_4(k+1) \\ \upsilon_5(k+1) \end{bmatrix} \quad \begin{pmatrix} A = p_x(k+1) \\ B = p_y(k+1) \\ C = p_z(k+1) \end{pmatrix} \quad (5)$$

Among them, $\upsilon_i(k+1)(i = 0,1,2,3,4,5)$ is the measurement of noise.

C. Data processing processes of UWB/MIMU integrated navigation system based on EKPF

In order to solve the problem of particle degradation in the standard PF algorithm, and to consider the real-time requirements of navigation information in the process of autonomous take-off and landing of UAVs, the complexity of the algorithm should not be too high, and the importance density function is constructed based on EKF to make the particles move to the high likelihood region, and the Extended Kalman Particle Filter (EKPF) algorithm is constructed, and the steps of the algorithm are as follows

- Initialize ($k = 0$):

  For $i = 1, 2, \cdots, N$, according to the prior probability sampling $X_0^{(i)} \sim p(X_0)$ to calculation $\omega_0^{(i)} = p\left[Z_0 \big| X_0^{(i)} \right]$,

the total weight $\omega_T = \sum\limits_{i=1}^{N} \omega_o^{(i)}$ can be obtained, and the normalized weight $\omega_0^{(i)} = \dfrac{\omega_0^{(i)}}{\omega^T}$ can be calculated.

- Forecasts and updates($k \geq 1$):

  EKF filtering is performed:

$$
\left\{
\begin{aligned}
&\hat{X}_k^{(i)}(-) = f\left[\hat{X}_{k-1}^{(i)}\right] \\
&P_k^{(i)}(-) = F_k^{(i)} P_{k-1}^{(i)} \left[F_k^{(i)}\right]^T + Q_k \\
&K_k^{(i)}(-) = P_k^{(i)}(-)\left[H_k^{(i)}\right]^T \left\{H_k^{(i)} P_k^{(i)}(-)\left[H_k^{(i)}\right]^T + R_k\right\}^{-1} \\
&\hat{X}_k^{(i)} = \hat{X}_k^{(i)}(-) + K_k^{(i)}\left\{Z_k - h\left[\hat{X}_k^{(i)}(-)\right]\right\} \\
&P_k^{(i)} = P_k^{(i)}(-) - K_k^{(i)} H_k^{(i)} P_k^{(i)}(-)
\end{aligned}
\right.
\tag{6}
$$

$i = 1, 2, \cdots, N$

- For $i = 1, 2, \cdots, N$, sampling by importance density function $X_k^{(i)} \sim N\left[X_k^{(i)}; \hat{X}_k^{(i)}, \hat{P}_k^{(i)}\right]$, weights $\omega_k^{(i)} = \bar{\omega}_{k-1}^{(i)} p\left[Z_k \big| X_k^{(i)}\right]$ can be calculated, total weight $\omega_T = \sum\limits_{i=1}^{N} \omega_k^{(i)}$ can be obtained, and normalized weights $\omega_k^{(i)} = \dfrac{\omega_k^{(i)}}{\omega^T}$ can be calculated.

- If $N_{eff} < N_{th}$, resampling should be performed, and $\bar{\omega}_{k-1}^{(i)} = 1/N$ need to be reset.

- The output can be obtained:

$$
\left\{
\begin{aligned}
&\hat{p}\left(X_k \big| Z_k\right) = \sum_{i=1}^{N} \bar{\omega}_k^{(i)} \delta\left[X_k - X_k^{(i)}\right] \\
&\hat{X}_k = E\left(X_k \big| Z_k\right) \approx \sum_{i=1}^{N} \bar{\omega}_k^{(i)} X_k^{(i)}
\end{aligned}
\right.
\tag{7}
$$

- Move on to the next moment loop.

IV. UWB/MIMU INTEGRATED NAVIGATION SYSTEM IMPLEMENTATION AND EXPERIMENTAL VERIFICATION

In this paper, the UWB/MIMU integrated navigation system is constructed, and the UWB and MIMU sensor data are collected during the real UAV flight to verify the effectiveness of the UWB/MIMU integrated navigation system.

A. UWB/MIMU integrated navigation system equipment

The MIMU sensor used in this paper is the MPU6050 sensor produced by InvenSense, which is shown in Fig. 3(a), and its parameters are shown in Table 1. The UWB sensor selected in this paper is the LinkTrack P-B series produced by Nooploop Co., Ltd., as shown in Fig.3(b), and its ranging accuracy can reach the decimeter level, as shown in Table 2.

(a) MPU-6050   (b) LinkTrack P-B

Fig. 3.   MIMU sensors and UWB sensors

TABLE I.   MPU6050 MIMU SENSOR PARAMETER CONFIGURATION

Parameter	Index
Accelerometer range (g)	± 4
Accelerometer sensitivity (LSB/g)	8192
Gyroscope range (°/s)	± 500
Gyroscope sensitivity(LSB/°/s)	65.5

TABLE II.   LINKTRACK P-B UWB SENSOR PARAMETER CONFIGURATION

Parameter	Index
Maximum communication distance(m)	500
Ranging accuracy(cm)	10
Operating frequency band(MHz)	[3744, 4243], [4243, 4742]

In the indoor experiment, the true value of the position of the UAV during flight is provided by the lidar and the laser ranging sensor, wherein the structure of the lidar is shown in Fig. 4(a), and the lidar adopts the principle of time-of-flight (TOF) ranging, which can be carried out at 360° Omnidirectional laser scanning, ranging accuracy is 3cm, measuring radius is 25m, has the advantages of high positioning accuracy, good mapping effect, the laser radar is installed on the top of the UAV in the experiment, the plane position information of the UAV can be obtained, and the structure of the laser ranging sensor is shown in Fig. 4(b) As shown, the laser ranging sensor adopts the TOF ranging principle, the ranging accuracy is 1.5cm, and has the advantages of high ranging accuracy, low blind spot, adjustable field of view, etc., and the laser ranging sensor is installed at the bottom of the UAV in the experiment, and the height position information of the UAV can be obtained.

(a) N10 series LiDAR   (b) TOFSense series laser distance sensors

Fig. 4.   Reference information sensors

According to the model of the selected MIMU sensor and UWB sensor, this paper builds an unmanned aerial vehicle platform, which is composed of a carbon fiber frame, a flight control unit, a navigation unit, a power unit and a communication unit, and the physical structure of the unmanned aerial vehicle platform is shown in Fig. 5.

Fig. 5.   Drone platform

B. Indoor environment experimental site layout and experimental procedures

The size of the indoor experimental site is $8m \times 8m \times 2.5m$, and the position coordinates of the six UWB base station nodes are recorded as follows: $A0(-1.4,-1.4,1.5)$, $A1(-1.4,6.6,1.5)$, $A2(6.6,6.6,1.5)$, $A3(6.6,-1.4,1.5)$, $A4(-1.4,2.6,2.3)$, $A5(6.6,2.6,2.3)$, the specific layout of the UWB positioning system is shown in Fig. 6.

Fig. 6.   Schematic diagram of the layout of the UWB positioning system for indoor environment experiments

In the experiment, the real flight trajectory of the UAV is shown in Fig. 7, wherein, the red five-pointed star marks the starting point of the flight trajectory, the red six-pointed star marks the end of the flight trajectory, the UAV platform moves according to the square trajectory, and finally lands below the flight starting point, and the UAV flight time is 115 seconds.

Fig. 7.   Indoor environment experiment drone flight trajectory

In the experiment, the ranging data of each base station of the UWB positioning system is shown in Fig. 8, and it can be seen from the figure that the ranging output of the UWB system is relatively stable in the absence of NLOS propagation conditions, and there is no obvious outlier.

Fig. 8.   Indoor environment experiment drone flight trajectory

Figure 9 shows the position error pairing between the pure UWB algorithm and the EKPF-based UWB/MIMU data fusion algorithm.

(a) X-axis position error

(b) Y-axis position error

(c) Z-axis position error

Fig. 9.   Comparison of the position error of the indoor environment experiment

It can be seen from the figure that the position accuracy of the pure UWB algorithm in the XY axis direction is high, basically staying between ±0.2m, but the position error of the pure UWB algorithm in the Z-axis direction increases significantly, and the fluctuation is serious. In the indoor experiment in this paper, the height difference in the Z-axis direction is 0.8m, and the height difference in the X-axis and Y-axis directions is 8m and 4m, which makes the third column element in the UWB positioning matrix much smaller than the first two column elements, resulting in the position error of the pure UWB algorithm based on least squares method in the Z-axis direction is higher than that in the X-axis direction. The position error of the UWB/MIMU data fusion algorithm based on EKPF is on the same order of magnitude as that of the pure UWB algorithm in the XY axis direction, and it is improved, and the position error in the Z-axis direction is significantly reduced, and there is no problem that the position in the Z-axis direction cannot be solved.

TABLE III.   COMPARISON OF POSITIONING ERROR RMSE IN INDOOR ENVIRONMENT EXPERIMENTS

Algorithm	Position error		
	X-axis	*Y-axis*	*Z-axis*
UWB	0.114	0.109	1.016
EKPF	0.072	0.07	0.215

Table 3 shows the comparison results of the pure UWB algorithm and the EKPF algorithm on the RMSE value of the position error. Compared with the pure UWB algorithm, the RMSE values of the position error of the EKPF algorithm in the three axes are reduced by 36.8%, 35.8% and 78.8%, respectively, which fully proves the effectiveness of the UWB/MIMU integrated navigation system based on the EKPF algorithm.

In order to verify the anti-interference capability of the integrated navigation system, a non-line-of-sight (NLOS) environment was designed in the interior to fully block the propagation path between the UWB base station and the tag by obstacles. In this environment, obstacles such as cartons, plastic boxes, plastic stools, iron buckets, and iron plates are added to simulate the impact of different material obstacles on the UWB positioning system, and the experimental environment layout is shown in Fig. 10.

Fig. 10. Schematic diagram of the indoor NLOS environment experiment layout

During the experiment, the flight altitude of the UAV was reduced as much as possible to make it lower than the height of the obstacle, so that the obstacle could fully block the direct propagation path between the UWB base station and the tag, the ranging data of each base station of the UWB positioning system is shown in Fig. 11.

Fig. 11. UWB ranging data output for indoor NLOS environmental experiments

According to Fig. 11, due to the emergence of the NLOS environment, the UWB ranging data has obvious fluctuations, and the influence of NLOS error is more obvious when the UAV flight altitude is low, and the influence of NLOS is weakened when the UAV flight altitude rises.

Figure 12 shows the position error pairing between the pure UWB algorithm and the EKPF-based UWB/MIMU data fusion algorithm.

(a) X-axis position error

(b) Y-axis position error

(c) Z-axis position error

Fig. 12. Comparison of the position error of the indoor NLOS environment experiment

Compared with the indoor unobstructed environment, due to the aggravation of the jitter of UWB ranging, the jitter of the position error is more intense, and the position error of the pure UWB algorithm also increases due to the increase of the ranging error of individual base stations, and the positioning accuracy of the UWB/MIMU integrated navigation system is higher than that of the pure UWB algorithm.

TABLE IV.   COMPARISON OF POSITIONING ERROR RMSE IN INDOOR NLOS ENVIRONMENT EXPERIMENTS

Algorithm	Position error		
	X-axis	*Y-axis*	*Z-axis*
UWB	0.141	0.137	1.354
EKPF	0.081	0.084	0.215

Table 4 shows the comparison results of the pure UWB algorithm and the EKPF algorithm on the RMSE value of the position error in the indoor NLOS environment experiment, and it can be seen from the results that the NLOS environment has less impact on the UWB/MIMU integrated navigation system. Compared with the pure UWB algorithm, the RMSE values of the position error of the EKPF algorithm in the three axes are reduced by 42.5%, 38.7% and 84.1%, respectively.

V. CONCLUSION

In order to meet the navigation and positioning requirements of indoor highly dynamic carriers, this paper proposes an algorithm flow using EKPF algorithm for UWB/MIMU data fusion, and deduces the system composition, state equation and measurement equation in the UWB/MIMU system model. Experiments show that the combined system of EKPF algorithm can realize decimeter-level positioning under the condition of high dynamics of indoor carriers, which proves the effectiveness of the data fusion algorithm.

REFERENCES

[1] Z. Sahinoglu, S. Gezici, I. Guvenc. Ultra-Wideband Positioning Systems: Theoretical Limits, Ranging Algorithms, and Protocols [M]. New York: Cambridge University Press, 2008.

[2] H. Ding, L. He, L. Zheng, et al. A New RWLS Positioning Method Basing on Distance Information [J]. Procedia Computer Science, 2014, (34): 633~638.

[3] Bregar K, Mohorčič M. Improving Indoor Localization Using Convolutional Neural Networks on Computationally Restricted Devices [J]. IEEE Access, 2018, 6: 17429~17441.

[4] Angarano S, Mazzia V, Salvetti F, et al. Robust Ultra-Wideband Range Error Mitigation with Deep Learning at the Edge [J]. Engineering Applications of Artificial Intelligence, 2021, 102: 104278.

[5] Benini A, Mancini A, Marinelli A, et al. A Biased Extended Kalman Filter for Indoor Localization of a Mobile Agent using Low-Cost IMU and UWB Wireless Sensor Network [J]. IFAC Proceedings Volumes, 2012, 45(22): 735~740.

[6] Ascher C, Zwirello L, Zwick T, et al. Integrity Monitoring for UWB/INS Tightly Coupled Pedestrian Indoor Scenarios [C]//Proceedings of International Conference on Indoor Positioning and Indoor Navigation, 2011: 1~6.

[7] ang Y, Li X. An Improved Robust EKF Algorithm Based on Sigma Points for UWB and Foot-Mounted IMU Fusion Positioning [J]. Journal of Spatial Science, 2019, 66(2): 329~350.

[8] Krishnaveni B V, Reddy K S, Reddy P R. Indoor Tracking by Adding IMU and UWB Using Unscented Kalman Filter [J]. Wireless Personal Communications, 2022, 123(4): 3575~3596.

[9] You W, Li F, Liao L, et al. Data Fusion of UWB and IMU Based on Unscented Kalman Filter for Indoor Localization of Quadrotor UAV [J]. IEEE Access, 2020, 8: 64971~64981.

[10] Liu F, Li X, Wang J, et al. An Adaptive UWB/MEMS-IMU Complementary Kalman Filter for Indoor Location in NLOS Environment [J]. Remote Sensing, 2019, 11(22): 2628.

[11] Benzerrouk H, Nebylov A V. Robust IMU/UWB Integration for Indoor Pedestrian Navigation [C]//Proceedings of the 25th Saint Petersburg International Conference on Integrated Navigation Systems, 2018: 1~5.

[12] Zeng Q, Jin Y, Yu H, et al. A UAV Localization System Based on Double UWB Tags and IMU for Landing Platform [J]. IEEE Sensors Journal, 2023, 23(9): 10100~10108.

[13] Dong X, Gao Y, Guo J, et al. An Integrated UWB-IMU-Vision Framework for Autonomous Approaching and Landing of UAVs [J]. Aerospace, 2022, 9(12): 797.

Refining the Methods for Determining the Orbital Parameters from Planetary Images in the Vektor-T Space Experiment onboard the ISS

M.Y. Belyaev
S.P. Korolev Rocket and Space Corporation
Korolev, Moscow region, Russia
mikhail.belyaev@rsce.ru

P.A. Borovikhin
S.P. Korolev Rocket and Space Corporation
Korolev, Moscow region, Russia
pavel.borovikhin@rsce.ru

D.Y. Karavaev
S.P. Korolev Rocket and Space Corporation
Korolev, Moscow region, Russia
dmitry.karavaev@rsce.ru

Abstract — **The objective of the Vektor-T space experiment conducted onboard the International Space Station is to develop new methods for determining spacecraft motion. The paper discusses one of the areas on which this experiment is focused – determining and refining station position using Earth surface images taken with the onboard photospectral system. The paper also provides an example of applying the technique developed in the course of the Vektor-T experiment to determine the position of the Luna-25 robotic spacecraft from a photograph of the lunar surface it had taken.**

Keywords - International Space Station ISS, space experiment Vektor-T, orbit determination, autonomous navigation, space images, Luna-25

I. INTRODUCTION

When solving many of the problems in ballistics, mission control and conduct of science experiments (SE), a spacecraft (SC) is traditionally considered as a single mass point. However, the dimensions of modern orbital stations are so large, that one can no longer ignore them neither when predicting their motion, nor when conducting a series of science experiments onboard the station. The International Space Station (ISS) is comparable in size with a football field. Therefore, when calculating its orbit, one needs to take into account not only the motion of the ISS center of mass, but also the motion about its center of mass, as well as changes in the station shape, which occurs, for example, due to rotation of solar arrays.

In order to develop and refine new methods for determining motion of large SC, the SE Vektor-T was set up onboard the ISS, which additionally uses for orbit calculations the attitude data for the station and its elements, satellite navigation systems measurements, motion data from spherical satellites launched both outside and inside the station cabin, etc. It is worth noting that at present, special interest is shown toward developing autonomous methods for determining SC motion [1]. The current urgency to develop new autonomous navigation methods is associated with the need to ensure the reliable determination of SC motion in the event of any abnormal situation while flying in both near-Earth space and on a future mission to the Moon or Mars. Such autonomous navigation methods were actively tested on the Salyut-series space stations [2]. One of the focus areas of the Vektor-T space experiment deals with the development of a new autonomous method of motion determination, which involves obtaining data on the ISS position through processing images taken by the onboard cameras of the PhotoSpectral System (PSS). Setting up such a standalone system only became feasible fairly recently owing to a considerable progress in digital cameras, which occurred over the last two decades, and emergence of the public-domain integral photographic images of planetary surface in orthographic projection (orthophotomaps). It turned out that having a snapshot of a certain region and an orthophotomap covering that region, it is possible to determine, with an error acceptable for application tasks, the view point position (the point where the camera was located) in space at the time when the picture was taken.

If processing a series of sequentially taken images of the planetary surface results in having computed coordinates of a sufficiently large number of SC positions at the known points in time when the pictures were taken, this makes it possible to apply standard statistical methods in order to compute orbital parameters.

The technique for determining SC position that is being developed onboard the ISS can be used without any significant changes for navigational purposes during missions to the Moon and Mars. The feasibility of this is confirmed by an example of computing the radius vector for the Luna-25 robotic spacecraft from a well-known lunar image taken from onboard of the at spacecraft on August 17, 2023.

II. GENERAL APPROACH TO SOLVING THE PROBLEM OF DETERMINING THE SC SPATIAL POSITION FROM A PLANETARY SURFACE IMAGE TAKEN FROM ONBOARD THAT SC

As was noted in the paper [3], an optical navigational field can be highly informative, and, theoretically, its use makes it possible to achieve a high ultimate accuracy of SC positional data representation; however, it is important that high computing power hardware and high-resolution CCD optical instruments be used for this purpose.

Optical navigation systems include, among others, those that, in order to determine the SC position relative to a planet (during landing or orbital motion), use images of that planet surface taken from onboard the SC. In the opinion of the authors of the monograph [4], such a navigation system achieving the accuracy comparable with that of satellite radio navigation systems would usher in a new age in the development of autonomous navigation for SC. The said work proposes one of the versions of solving that problem based on automatic recognition of terrestrial landmarks by comparing Earth pictures taken from the orbit with reference images. Also to be used are measurements of SC altitude above the ocean and data from star trackers.

Also of interest for SC navigation is an approach proposed in [5] for aviation – used for achieving high-accuracy correction of navigational data for flying vehicles are methods of photogrammetry and optimal data filtration,

stereoscopic methods, also used are images of the terrain and single reference points, including contours.

Monograph [6] presents georeferencing of Earth surface pictures taken from onboard a SC with the use of special electronic maps, and involving the use of shore lines as reference points. In spite of the fact that [6] does not state the problem of autonomous navigation, it points out that corrections can be computed for the already available parameters describing the SC center of mass position and attitude, in order to update georeferencing parameters.

It is worth noting that both [5] and [6], along with the purely automatic image processing, also provide for human operator involvement in this process. For several years RSC Energia has been developing and refining processes for determining SC spatial position (in the form of radius vector) that are also based on automated georeferencing by human operator of the Earth surface images made from onboard the ISS.

In order to perform such georeferencing, one needs to have available, in addition to the digital planetary surface image itself, a ready-made orthophotomap of the photographed area. Since photograph itself is a centrally projected image (all the imaginary rays that form the image go through the focal point), double ratios theorem applies to it, and, therefore, using four selected common (reference) points it is possible to transform the photograph to the projection of the available orthophotomap.

The orthotransformation procedure consists in human operator identifying and selecting in the photograph which is being processed and in the reference orthophotomap of the area four common (reference) points; as a result, a special algorithm determines geographic coordinates of all the points in the photograph [7]. If need be, the practical accuracy of this referencing can be evaluated by superimposing the orthotransformed photograph on the orthophotoplan and determining spatial misalignments of the corresponding features in the two images.

Thus, as a result of orthotransformation each pixel in the photograph is mapped to a small spot (point) on the Earth surface with known geographic coordinates, which can be converted to any other coordinate system (CS) that is convenient for further calculations. Thereupon, it is easy to also determine the shortest linear distance between two points on the Earth surface that are seen in the photograph. In addition to this, for any pair of pixels in the image it is possible to determine the value of the angle between rays directed toward the corresponding points on the Earth surface from the focal point of camera's optical system.

It is easy to show that on the plane the segment OP (Fig. 1) will always be seen at the same angle φ from any point on a continuous line that goes through points O and P and is represented by the part of the circle, which in Fig 1 lies above the X-axis and is dotted. The radius and position of the center of that circle are determined by the value of the angle, as well as the length and position of the segment on the plane [8].

In three-dimensional space the surface from each point of which this segment can be seen at the specified angle is a closed torus, which is generated by rotation of the circle shown in Fig.1 about the segment OP.

Fig. 1. Possible observation points (dotted line), from which the segment OP is seen at the specified angle φ

One of the methods of using an orthotransformed photograph in order to determine the point from which the photograph was taken consists in selecting in the photo three pairs of widely spaced points, associating with each of these their own torus generated as stated above and their own equation of distance from an arbitrary point to the surface of that torus. Each of the equations is set to zero and expressed in terms of coordinates of one common CS, for example, Cartesian Greenwich one. Then, with a sufficiently high precision of determining coordinates of selected points and directions toward them, the point where the picture was taken is found using numerical methods as one of the solutions of the system of three non-linear equations. However, one needs to exclude those solutions that do not represent the true point from where the picture was taken (for example, by selecting an initial approximation, which is sufficiently close to the sought point).

If there is some way to determine the position of a certain plane in which the SC was located at the time of taking the picture, then the equation of the distance from that plane to an arbitrary point can be included in the system which defines the position of the point where the picture was taken, together with equations for the toruses (the number of which in this case can be reduced to two).

One can also tie with this known plane an orthogonal Cartesian CS, the Z-axis of which is perpendicular to the plane. Then one can express in terms of the coordinates of this CS the two equations of distances from an arbitrary point in the plane to intersections with this plane of the two toruses (generated from two pairs of points with known coordinates). Thus, the problem reduces to solving a system consisting of two non-linear equations.

If the segments connecting the selected pairs of points lie in the known plane (or are perpendicular to it), the intersections of their corresponding toruses with this plane are circular curves, and solutions of the corresponding systems of equations can be found using relatively simple analytical methods, the use of which will not cause the problems, which can potentially occur when numerical methods are used, that involve the search for initial approximations and convergence of iteration algorithms. Below is given an example of such approach, which is regularly used for determining the ISS position, in particular, within the framework of the SE Vektor-T.

An important element in implementation of autonomous navigation is the program that is currently used onboard the ISS for displaying to the crew the current trajectory and navigation situation, which is derived from the telecommunication monitoring and control system Sigma, which goes back to orbital station Mir, and was based, in its turn, on the autonomous navigation and control system Delta of orbital station Salyut-7. This program simulates the flight of the ISS, including such particulars as motion prediction, light and dark environment, view from the portholes, etc. Under good conditions (in the absence of ionospheric disturbances) the accuracy of the simulation is 1-2 s for several days without updating data from ground services.

In addition to this the program is a convenient platform for developing various technologies, including autonomous navigation, since it is expressly designed to support science programs, and its executable modules can be promptly changed and updated via radio link from the ground.

At present the program enables real-time control of various science equipments, including steerable platforms of the System for Orientation of Videospectral Equipment (SOVE) and an ultrasonic system for coordinate referencing of photographic images, which actually require solving autonomous navigation problems.

III.   THE USE OF A SPECIALLY SELECTED PLANE IN NAVIGATION PROCESSING OF AN IMAGE

As noted above, as a result of orthotransforming procedure it becomes possible to determine geographic coordinates of all the points shown in the image. If the georeferencing accuracy is sufficiently high, this makes it possible to use any fragments of the image in the algorithms determining the position of the point from where the image was taken. Let us consider how this can be used for selecting a special working plane in order to reduce the three-dimensional problem to a two-dimensional one.

Marked in the original image that has not yet been orthotransformed are pixels that form a circle inscribed into the image with the center that coincides with the center of the image. The radius of such a circle is equal to a half of the smaller side of the rectangular image frame. If the imaging was performed with deflection from nadir direction and the limb of the planet is outside the image frame, in the referenced orthotransformed image the circle is transformed into an elongated closed curve, which corresponds to the intersection of a cone (centered at the focal point of the optical system) with the shape of the planet. This curve looks like an ellipse, but is not an ellipse [9]. But the shape it circumscribes is always stretched in the direction of the largest tilt of the planetary surface with respect to the central axis of imaging (which is normal to the plane the camera image sensor) and is used for selection of the special working plane in the following manner.

The axis along which is stretched the shape circumscribed by that curve indicates the direction of the Earth surface tilt with respect to the central axis of imaging, and can be used for determining the plane in which further computations will be performed. Marked in the orthotransformed image are pixels that correspond to the center of the original image and the circle that was inscribed into it. Corresponding to these pixels are small spots (points) on the planetary surface with known coordinates. This makes it possible to calculate the distance from the point, which corresponds to the image center (let us denote it by B), to all the points which correspond to the circle. If the curve

corresponding to the circle is stretched, a point can be found in it (we shall call it A), which is farthest removed from point B. Then, on the diametrically opposite side from A one can select a point, the distance to which from the central point B is maximal in comparison with the neighboring points in the curve. Let us denote this point by C.

Selected as the working plane is the plane wherein lie points A, B and C. Lying in the same plane at the time when the picture was taken, must be the point, from which the picture was taken (focal point of the lens), which we shall denote by S. If the planetary surface is spherical, this plane goes though the center of the planet, and it also aligns with the prime vertical plane of the image, which is defined by points S, B and the point on the planetary surface which is nadir with respect to S.

Fig. 2, the plane of which coincides with the selected working plane, shows the scheme for determining the position of the point S from which the picture was taken. From that point segments AB and BC are seen at the same angle, that is, angles ASB and BSC have the same known value (let us denote it by α), which is determined by the geometrical properties of the photographic equipment and is equal to a half of the angular dimension of the narrow side of the photographic frame, because it was that frame in which the circle was inscribed before the orthotransformation procedure. Also known are coordinates of points A, B and C selected on the planetary surface. Dashed and dotted lines show fragments of circles from which segments AB and BC are seen at angle α. Instead one of them, one could also select the circle constructed on segment AC for its observation angle equal to 2α.

Fig. 2. Geometric chart showing how to determine position of point S from the known positions of points A, B and C

The two circles shown in Fig.2, under real geometrical conditions of imaging the planetary surface from an orbital SC, will have two intersection points, one of which coincides with point B (which can be used to verify the accuracy of the results), and the other coincides with point S. Thus it is not difficult to calculate the position of the point from which the picture was taken.

When using the geometric scheme shown in Fig.2, one should not necessarily use equations of the circles constructed on the specified line segments. Earlier, the following formula was derived [10] for angle *BCS* (let us denote it by γ):

$$\mathbf{tg}(\gamma) = \frac{\sin(\tau)\cos(2\alpha) - \cos(\tau)\sin(2\alpha)}{\cos(\tau)\cos(2\alpha) + \sin(\tau)\sin(2\alpha) + \dfrac{BC}{AB}}$$

where τ is the known value for angle *ABC*.

Since, taking this formula into account, all the angles of the triangle *BCS* are known, the position of point *S* is calculated from the known positions of points *B* and *C*.

However, in comparison with this solution, the approach that involves the use of the specially constructed circles is more universal even in a two-dimensional case. Thus, one can select line segments that do not have a common point and are visible from the point from which the picture was taken at different angles. This is possible if one succeeds in recognizing features with known coordinates that lie in the working plane but do not coincide with points *A*, *B* and *C*. In those cases the algorithm of solution with the use of circles does not significantly change from solution of the problem presented in Fig.2.

Some methods of image processing based on constructing special circles can simultaneously produce two simultaneous possible positions for the point from where the picture was taken that lie in the working plane. In those cases, additional methods are required in order to select the solution which corresponds to the actual point from where the picture was taken [10].

IV. DETERMINING THE ISS POSITION AND ORBIT USING IMAGES THAT WERE TAKEN FROM IT

As mentioned above, one of the focus areas of SE Vektor-T is development of a process for determining the ISS position and orbit by processing images of the Earth surface taken from the Russian Segment of the station. To achieve this objective, which includes search for 'bottlenecks' affecting the accuracy of the results, a camera that was rigidly secured (using FSS equipment) to a porthole inside the ISS cabin was preprogrammed to automatically take pictures of the Earth surface (as a rule, once per minute).

Currently being developed within the framework of Vektor-T experiment is the procedure presented in the previous section. One of its peculiarities is that the accuracy of the results goes down considerably when the points on the underlying surface that are being photographed lie close to nadir. Therefore, the camera was secured onto the porthole in such a manner as to make sure that the camera points in a direction that is at least 30 degrees away from nadir. For this purpose, the optical axis of the camera lens was tilted away from the ISS ground trace to the left or to the right depending on the predicted trajectory to make sure that as many images of the sunlit terrain that is potentially suitable for recognition are taken as possible.

Fig. 3 shows one of the images (with provisional ID number *54R0300*), which was taken in the course of an imaging session on June 28, 2023. During the first stage of processing a circle was inscribed into it as per the procedure.

Using special software, a ground-based operator recognized the terrain in the picture and matched four points in the original image to corresponding points in the orthophotomap of the Earth surface. All further actions were fully automatic. When converting geographic coordinates into Cartesian coordinates, Earth oblateness was taken into account as per geodetic system WGS 84.

The original image was transformed to orthographic projection and aligned with orthophotomap as shown in Fig.4. The stretched closed white curve corresponds to the circle, which, prior to orthotransforming, was inscribed into the original image shown in Fig.3. The white line segment marks where the Earth surface intersects with the plane that is close to the principal vertical plane of the image. Marked on this line segment are points that were selected in the same manner as in the previous section were selected points *A*, *B* and *C*. In Fig. 4 they are denoted, respectively, as *@_54R0300_A*, *@_54R0300_B* и *@_54R0300_C*.

Fig. 3. Image *54R0300* taken from onboard the ISS on June 28, 2023 at 16:20:00 UTC, with inscribed circle marked

Fig. 4. Orthotransformed image *54R0300* aligned with orthophotomap and marked out for further processing

The same three points are marked with red crosses in a less detailed map of the terrain (Fig. 5). The yellow line shows the ISS ground trace, that is, the projection of the ISS orbit known a priori, which served as a reference standard for estimating the accuracy of the calculations. The red circle on the ground race shows the ISS reference position (sub-satellite point) at the moment when image *54R0300* was taken. The ISS sub-satellite point calculated using that image is marked on the ground trace with a red cross labeled *@_54R0300_M* and partially covered by the red circle.

In this particular case the mismatch between calculated and reference sub-satellite points was 3 km. Mismatches between sub-satellite points obtained from processing the other 28 images taken during sessions conducted on June 28 and 30, 2023, turned out to be roughly the same and lay within the interval of 2.9–3.2 km. More considerable was the mismatch between the calculated and reference altitude positions of the ISS, which is about 6 km for image *54R0300*, and on average about 4 km for all the 29 images. Thus, the average mismatch of direct range was about 5 km.

Fig. 5. Comparison of the ISS reference position (marked with a red circle) with the position calculated from the orthotransformed image 54R0300

Used for calculation of the ISS position were methods presented in the previous section, taking into account atmospheric refraction and the altitudes of the selected areas of the Earth surface above sea level. The angle, at which the line segments formed by the points selected in Fig.4 were visible, was calculated based on the geometrical dimensions of the image sensor of the NikonD5 camera, which was used in the experiment, and the focal distance of the lens installed on that camera. To additionally verify the results, alternative calculation methods were used, including numerical methods. To obtain the final, more accurate, solution and additional step was taken, which is described below.

In the course of running the Vektor-T experiment, it was found that besides objective hurdles caused by conditions inherent in Earth imaging (cloud cover, large ocean surfaces, difficult mountainous terrain, refraction, etc.), the accuracy is also significantly affected by the mounting error, that is, by the fact that the optical axis of the lens does not go through the central pixel of the image (image sensor). Therefore, by specifying (in pixels) certain deflections from the center of the image, one can find such a variation for which, for example, two calculated angles formed by the optical axis and directions from the focal point toward two different points of the circle are closest to each other and to the specified value. Such a criterion turned out to be successful probably because usually optical components are machined by turning the workpiece, and deviations from ideal circles are negligible, and, therefore, in reality all the angles between the optical axis and the rays that go through the edges of the lenses virtually coincide.

The accuracy of the results obtained through the use of the procedure can be improved by calibration of the hardware parameters. For example, the calculation of the position of the point from which the picture was taken is significantly affected by the degree to which the actual focal distance of the optical system is close to the nominal value that is being used. For example, under conditions of the completed imaging sessions, an error of 1 mm in the focal distance of the lens would have resulted in 1 km error in determining the ISS attitude. And, of course, of crucial importance are image resolution and scale, and the accuracy of the orthophotomap used for georeferencing of the images.

The above methods do not require the knowledge of the point in time at which the picture was taken in order to determine the position of where the picture was taken from onboard a SC. Therefore, if the SC orbit is known, the calculated position of the point from where the picture was taken makes it possible to establish the time when the picture was taken, since it is possible to determine when the SC was in the vicinity of that point. That is why there have been multiple actual cases where the correct time when a picture was taken from onboard the ISS could be restored whenever it transpired during post-flight processing of images from different SEs that the readings of built-in camera clock were wrong. In certain cases, ground specialists managed to determine time readings bias from the images quickly enough to take it into account or remove it when planning the next Earth remote sensing session from onboard the ISS. The accuracy achieved in Vektor-T experiment is fairly sufficient for addressing such tasks.

If, on the other hand, the task is to determine the SC orbital parameters so that they can be used for predicting its motion, the requirement to exactly match the photographs and the times when they were taken becomes important. It is true that the use of classical analytical methods of determining the orbit from several positions of SC is unpractical if accuracy of the measurements is low. However, by using a sufficiently large number of images, one can calculate a series of successive values of the ISS radius vector for known points in time, and then, with the application of statistical methods, also determine the station orbit.

Most of the photographs of Earth obtained from automatic imaging with a camera attached to a porthole are not suitable for recognition and georeferencing, since they either show continuous cloud cover or some featureless terrain such as a desert or the surface of the ocean without islands.

Thus, the largest sequence of recognized and orthotranformed images which it was possible to obtain in one automatic imaging session amounts to only 30 photographs while the total number of images that were taken is 172 [8]. The values of radius vectors of the points where the pictures were taken that were obtained as a result of processing those 30 images made it possible to obtain a smoothing solution of the ISS motion equations, for which the positions of ascending nodes in the processing interval differed from positions of ascending nodes of the ISS reference orbit (exactly determined by other methods) by less than 1 km. However, in certain cases, discrepancies (deviations of the calculated positions of the points where the pictures were taken from reference solutions) had greater values. This is explained by an inaccuracy of time readings in photographing equipment which resulted in a bias error of an a priori unknown value. In order to reduce it (without using data about the reference orbit) a special correction of image time referencing was performed where the error in

determining the time of photography was introduced into the system of equations as a definable parameter. This made it possible to improve the accuracy of the ISS orbit determination.

The accuracy of orbital determination can be improved by using a similar correction of other values which need updating (for example, departures of photographic equipment parameters from nominal values) by means of introducing them into the system of equations as unknown parameters, as well as increasing the number of images from which the ISS positions are determined. The proportion of images that are suitable for recognition can be increased by intentionally selecting for imaging the areas with high recognition probability and low probability of cloud cover, which requires crew involvement or development of specialized software for imaging control using automatic image recognition. If approximate values of SC orbital parameters are known a priori, in order to refine them one can, based on the approximate orbit, schedule imaging of areas with low probability of cloud cover, which is estimated using statistical data or weather forecasts.

In the case of imaging the Moon, which has no bodies of water and no cloud cover, and has a lot of easily recognizable areas and landmarks such as craters, one can expect that the number of images unsuitable for navigation processing will be minimal even when performing automatic imaging with a camera that is rigidly fixed onboard a SC. It is also worth noting that when calculating circumlunar orbits one does not need to take into account atmospheric refraction and determine atmospheric drag factor.

The convenience of using this technology in manned missions has to do with the fact that there is no need to have special instrumentation for navigation such as star or sun trackers, because the crew will always have a camera and a laptop computer, however, at present one needs to involve an operator (a crewmember or a specialist on the ground) to process the images of Earth at the stage when the photographed terrain is being recognized and common (reference) points are being identified in the photo and orthophotomap. Transferring these functions to an autonomous program is complicated by such factors as cloud cover, changes in the appearance of Earth surface areas (for example, seasonal changes or those caused by human activities). As for the lunar images, the most important factor affecting their appearance are variations in lighting conditions, which makes it easier to recognize fragments of the lunar surface and gives hope that the proposed procedure for lunar missions can be made fully automatic. This is also confirmed by significant number of publications on the subject of developing, for navigational purposes, systems for recognition of such predominant lunar features as craters [11, 12].

V. Position estimate for robotic spacecraft Luna-25 made using a photo of lunar surface made from onboard the spacecraft

Considered among various navigation systems proposed for future lunar missions are visual navigation aids. One example of such an approach is autonomous optical navigation using lunar horizon and reference points on lunar surface employing special wide-angle cameras [13].

At the same time there is no doubt that in the course of crewed missions the imaging of the lunar surface for various purposes will be performed by the crew with hand-held cameras through portholes, and, possible, also with steerable cameras on the outer surface of the SC [14]. The above

procedure, whenever even a non-specialized computer is available (such as a laptop), can be used for navigational purposes and during imaging of this type, and not only in specially designed systems for optical navigation.

To verify theoretical feasibility of determining the SC position relative to the Moon from photos of the lunar surface taken with a hand-held camera through a porthole, calculations were performed earlier using photographs available online that had been taken during the US crewed program Apollo [10, 15]. In spite of modest quality of those digitized images, it was possible to orthotransform and use them for approximate estimation of the SC radius vectors at the time of photography.

In August of 2023, a photograph (Fig.6) was published [16] that was taken by the Service Television System STS-L from onboard spacecraft Luna-25 when it was flying in orbit around the Moon as its artificial satellite. To take the picture, the landing camera KAM-S was employed, performance data for which are given in [17]. The image resolution was primarily limited by the small focal distance (12.3 mm) of the camera's optical system. Nevertheless, the image was used to try out the procedure that is being developed in SE Vektor-T.

Fig. 6. Lunar picture taken from onboard Luna-25 on August 17, 2023 at 05:22:56 UTC

To perform selenographic referencing of the image, Mercator projected orthophotomap of the Moon found at [18] was used. The image, transformed to orthogonal projection and aligned with orthophotomap, is shown in Fig.7 Used a common (reference) points were centers of several small lunar craters. One can see that the orthophtomap has a lower resolution than the image, so one cannot expect the referencing to be highly accurate. This is borne out by the fact that, as can be seen from Fig.7, those points of the line that was obtained from the circle inscribed in the original image that were most distant from the central point are not diametrically opposed to each other.

Preliminary processing of the image in this particular case yielded the following results:

deflection of the optical axis from nadir: 19.87°;
line-of-sight range to the center of the image: 96.9 km;
altitude above the lunar surface: 91.5 km.

Fig. 7. Orthotransformed image aligned with orthophotomap and markings made during its automatic processing

These results correspond to two possible positions of the point from where the picture was taken in the plane that is close to the prime vertical plane of the image, with the following sub-satellite points in selenographic coordinate system:

logitude:-138.65°, lattittude: -75.67°;

longitude: -139.36°, lattitude: -77.73°.

Using methods that are being developed in SE Vektor-T (per scheme presented in Fig.2), selected out of these two options was the second one. The corresponding coordinates in Cartesian selenocentric selenoequatorial system:

x: -295.52 km, y: -253.68 km, z: -1785.56 km.

Their transformation to selenocentric geoequatorial coordinate system J2000 [19, 20] yields the following result:

X: -354.44 km, Y: 584.07 km, Z: -1695.03 km.

Thus, applying techniques developed in a space experiment onboard the ISS, it was possible to obtain a result even from processing a photograph taken from onboard Luna-25. However, this result should only be considered a rough estimate owing to a relatively low resolution of the images that were used.

VI. Summary

In view of current plans to launch mission to the Moon and Mars, as well as vulnerability of navigation satellite systems in near-Earth space, it is important to develop new methods for autonomous determination of SC motions. The described navigation techniques based on planetary images are implemented both as ground-based and onboard software. It was developed using math models of SC motion around the planet as well as in the course of SE Vektor-T and some other scientific experiments which involved observations of the Earth surface from the ISS. In addition to this, in certain cases the proposed algorithms were used for estimating SC position from the pictures of the lunar surface taken by that spacecraft.

The obtained results suggest that even when SC coordinates at the time when the SC were taking pictures were calculated from single images (including those that were taken for purposes other than navigation), those coordinates make it possible to update the already known parameters of the orbit (trajectory) of the SC, for example, in order to calculate corrective burns in the case of lunar missions or to adjust the timetable for observations of terrestrial targets for scientific experiments onboard an orbital station.

In the cases where a sufficiently large number of consecutive images of the planetary surface were taken, it becomes possible to use well-developed statistical methods for measurement results processing in order to determine the SC orbit, which was confirmed in the course of the experiment Vektor-T.

The proposed technology that is being developed onboard the ISS can be used for setting up an independent navigation system that can serve as, for example, a backup, emergency or verifying system supplementing the currently used navigation systems. However, its major objective is to create a standalone navigation system that does not need any specialized instrumentation, and can be used for both unmanned spacecraft and manned missions to the Moon and, eventually, to Mars.

References

[1] Peshekhonov V.G. High-Precision Navigation Independently of Global Navigation Satellite Systems Data // Gyroscopy and navigation. 2022. v. 30. No. 1(116). pp. 3–11. DOI: 10.17285/0869-7035.0084

[2] Gaushus E.V., Zybin Y.N., Legostaev V.P. / Autonomous navigation and control of the orbital station Salyut-7 // Space Research. 1986. v. 24. Issue 6. pp. 844–864.

[3] Pochukaev V.N. Basic conceptual foundations for trajectory and navigation support of robotic spacecraft missions. Moscow, Machinostroyeniye-Polyot, 2018. 156 p.

[4] Anshakov G.P., Golyakov A.D., Petrishchev V.F., Fursov V.A. Autonomous navigation of spacecraft. Samara: State Research and Production Rocket and Space Center TsSKB-Progress 2011. 486 p.

[5] Beloglazov I.N., Kazarin S.N., Kosyanchuk V.V. Data processing in iconic systems for navigation, guidance and remote sensing of terrain. Moscow: Fizmatlit, 2012. 368 p.

[6] Modern technologies for Earth Remote Sensing data processing. / S.V. Antonushkina, V.S. Gurov, N.A. Yegoshkin [et al.]; Under the editorship of V.V. Yeremeev. Moscow: Fizmatlit, 2015. 460 p.

[7] Results of monitoring of catastrophic floods in Krasnodar Territory from the board of ISS RS / G.I. Padalka, S.N. Revin, LV. Desinov [et al.] // Manned spaceflight. 2013. № 4(9). pp. 42–59.

[8] Belyaev M.Y., Borovikhin P.A., Karavaev D.Y., Sazonov V.V. A method for determining the orbit of the International Space Station using Earth surface photos taken from it // Proceedings of the 54th Tsiolkovsky Readings. Section "Problems of Rocket and Space Technology". 2020. pp. 5–16.

[9] Burdayev M.N. About the shape of the boundaries and sizes of the planetary surface areas of visual coverage from spacecraft // Manned spaceflight. 2014. № 3(12). pp. 71–75.

[10] Mikrin E.A., Belyaev M.Yu., Borovikhin P.A., Karavaev D.Yu. Determining orbit from earth and moon surface pictures taken by cosmonauts // Space Engineering and Technology. 2018. № 4(23). pp. 88–99.

[11] Maass B., Woicke S., Oliveira W. M., Razgus. B, Krüger H. Crater Navigation System for Autonomous Precision Landing on the Moon // Journal of Guidance, Control, and Dynamics. 2020. Vol. 43, No. 8. pp. 1414–1431. DOI: https://doi.org/10.2514/1.G004850

[12] Bobkov A.V., Xu Yang. Signature method for determination of the lunar lander position by video image // Engineering Journal: Science

and innovations. 2023, Issue 5. DOI: http://dx.doi.org/10.18698/2308-6033-2023-5-2278

[13] Zhukov B.S., Polyansky I.V., Zhukov S.B. Autonomous optical navigation at lunar orbits and during landing on the Moon using an extra wide angle camera // Current problems in Earth remote sensing. 2017. v. 14. No.3. pp. 24–35. DOI: 10.21046/2070-7401-2017-14-2-24-3

[14] Zhukov V.M. Analysis of the Conditions, Features and Ways of Performing the Visual Instrumental Observations and Photography on the Moon // Manned spaceflight. 2021. № 2(39). pp. 111–131. DOI 10.34131/MSF.21.2.111-131

[15] Mikrin E.A., Belyaev M.Y., Borovikhin P.A., Karavaev D.Y. A novel autonomous navigation technique using pictures in support of a circumlunar mission: Development testing aboard the ISS // 25th Saint Petersburg International Conference on Integrated Navigation Systems. ICINS 2018. Proceedings. 2018. pp. 1–7.

[16] Space Research Institute of the Russian Academy of Sciences, Moscow, Russia [online resource]//info and reference web portal. URL: https://iki.cosmos.ru/research/missions/luna-25

[17] Service Television System of Luna-25 SC. / G.A. Avanesov, F.V. Berezhkov, R.V. Bessonov [et al.] // Herald of Astronomy. 2021. v. 55. No. 6. pp. 601–617.

[18] GoogleMoon [online resource] // google.ru: Info and reference web portal URL: https://www.google.ru/moon/

[19] Engineering. Encyclopedia. Vol. IV – 22. Rocket and Space Technology. In 2 books. Book 1 / A.P. Adzhayan, E.L. Akim, O.M. Alifanov [et al.]; Under the editorship of Member of RAS V.P. Legostaev, Editor-in-Chief member of RAS K.V. Frolov, Moscow: Mashinostroeniye, 2012. 925 p.)

[20] Trajectory and navigation support of robotic spacecraft missions to Solar System bodies. / G.K. Borovin, Yu.F. Golubev, A.V. Grushevsky [et al.]; Under the editorship of Dr. Sci. Phys&Math A.G. Tuchin. Khimki JSC Lavochkin NPO, 2018, 336 p.

Optical-Inertial Positioning of Remotely Piloted Aircraft in the Loss of Satellite Information

N.P. Starostin

Ramenskoye Design Company, Ramenskoye,
Moscow Region, Russia,
e-mail: stark201288@gmail.com

A.V. Chernodarov

"NaukaSoft" Research & Production Association, Ltd.,
MAI University, Moscow, Russia,
e-mail: chernod@mail.ru

Abstract — The problem of optical-visual correction of inertial navigation systems (INS) in the event of loss or deterioration of satellite information is considered. The possibility of using optical-electronic means that are part of aircraft remote control systems for this purpose is being estimated. Such control is implemented by a human operator via a video transmission channel of an image of the earth's surface without the use of radar facilities. Algorithms for the formation of optical-inertial observations when sighting ground landmarks with known and unknown coordinates are presented. The results of mathematical modeling confirmed the fairly high potential accuracy of optical-inertial positioning with the extended Kalman filter in the errors estimation loop of the INS.

Keywords — *remotely piloted aircraft, inertial navigation system, global navigation satellite system, optical-electronic system, extended Kalman filter.*

I. INTRODUCTION

The current state of remotely piloted aircraft (RPA) [1] is characterized by the use of inertial and satellite navigation systems (ISNS) [2,3]. In such ISNS global navigation satellite systems (GNSS) provide high-precision positioning, and inertial ones (INS) - the determination of the angular orientation. However, the use of GNSS for aircraft positioning has a number of limitations related to the effects of natural and simulated interference [4-6]. At the same time, the presence of a video camera on board the aircraft and a channel for transmitting images of the Earth's surface for remote piloting makes it possible to expand the navigation capabilities of the ISNS based on additional optical-electronic positioning. Visible landmarks can be used for such positioning. The following approaches are traditional, using video information from optical-electronic systems (OES) for INS correction:

• correction of the INS by flying the aircraft over landmarks with known coordinates. This approach requires precise flight over a landmark, which cannot always be realized;

• INS correction using on-board rangefinders and measured sighting angles of a ground landmark. The inclusion of rangefinders in the ISNS structure is not always possible due to restrictions on the size and weight of onboard equipment of a RPA.

These approaches, in addition, do not provide an estimate of the errors in the orientation angles of the INS

when processing positional observations using the extended Kalman filter (EKF) [7].

It is proposed to perform INS correction by monitoring on the OES indicator screen images of landmarks that are at a visible distance from the RPA. The image obtained using a video camera is linked to a coordinate grid in the plane of the indicator. Taking into account the focal length of the video camera, the azimuth and elevation angle of the line of sight of the landmark will correspond to a point on the indicator screen. INS correction is implemented based on processing optical-inertial observations of viewing angles using EKF.

Based on known approaches to optical-electronic positioning, it is proposed to correct the INS using passive video location of ground landmarks without the use of rangefinders. The following options for the formation of optical-inertial observations for estimating errors of a strapdown INS (SINS) using the EKF are considered:

• observations in the form of the difference between the sighting angles predicted from SINS information and those measured using the onboard OES, namely: azimuth and elevation angle of a landmark in the coordinate system associated with the RPA. The coordinates of the landmark are assumed to be known. The video camera can be rigidly fixed in the body of the RPA. Then the viewing angles are determined in the coordinate system associated with the RPA. If the video camera is placed in a stabilized gimbal, then the viewing angles can be determined relative to the axes of the accompanying trihedron of the reference coordinate system;

• observations of sighting angles of a landmark with unknown coordinates. To implement such observations, it is proposed to carry out preparatory procedures related to determining the coordinates of the tracked landmark using the triangulation method for spaced positions of the RPA along the flight path.

The transition to optical-inertial positioning of the RPA is carried out when a loss of satellite information is detected [8] and there is an indicator of a localized visible landmark on the screen.

II. OPTICAL-INERTIAL OBSERVATIONS WHEN SIGHTING LANDMARKS WITH KNOWN COORDINATES

Optical-inertial observations can be formed taking into account the interconnection of projections of the line of sight

of a landmark on the axis of navigation trihedrons associated with the landmark and the video camera.

Let the unit vector $\bar{e}$ be directed along the sighting line of the landmark. Such a vector in projections on the axes of the *oxyz* coordinate system associated with the video camera will have the form

$$\bar{e}_{xyz} = [e_x e_y e_z]^{\mathrm{T}}, \quad (1)$$

where $e_x = \cos\alpha\cos\beta$; $e_y = \sin\beta$; $e_z = \sin\alpha\cos\beta$; (2)

α ; β are the azimuth and elevation angle of the landmark in the coordinate system associated with the video camera.

Taking into account relation (1), the values of viewing angles predicted from SINS information can be determined

$$\alpha = arctg(e_z / e_x) ; \quad (3)$$

$$\beta = \arcsin e_y, \quad (4)$$

where $\qquad \bar{e}_{xyz} = C\bar{e}_{ENH}, \quad (5)$

$\bar{e}_{ENH} = [e_E e_N e_H]^{\mathrm{T}}$ is a unit vector in projections on the axis of the accompanying trihedron *oENH* of the geodetic coordinate system;

$e_E = \Delta E / D$; $\qquad e_N = \Delta N / D$; $\qquad e_H = \Delta H / D$;

$D = \sqrt{\Delta E^2 + \Delta N^2 + \Delta H^2}$;

$\Delta E = (\lambda_{\mathrm{lm}}\text{-}\lambda)R_\lambda \cos\varphi$; $\Delta N = (\varphi_{\mathrm{lm}}\text{-}\varphi)R_\varphi$; $\qquad \Delta H = R_{\mathrm{lm}}\text{-}R$;

φ_{lm} ; φ is, respectively, the geodetic latitude of the landmark and the RPA; λ_{lm} ; λ is, respectively, the geodetic longitude of the landmark and the RPA; R_φ ; R_λ are the radii of curvature of the earth's ellipsoid; ΔH is the difference between the heights of the landmark and the RPA above the earth's ellipsoid; *C* is a direction cosine matrix (DCM) connecting the projections of the $\bar{e}$ vector in the *oxyz* and *oENH* bases. For these trihedrons, the traditional [9] sequence of rotations along orientation angles is adopted, taking into account the direction of the axes in the coordinate system associated with the RPA [10].

If the video camera is rigidly fixed in the body of the RPA, then the DCM *C* is determined by the angles of the true course, pitch and roll calculated by the SINS.

If the video camera is placed in a gimbal, then it can be stabilized in the horizontal plane in pitch, roll and azimuth according to SINS information. In this case, the DCM *C* is formed taking into account the orientation of the video camera only in azimuth.

In the presence of GNSS information, the SINS errors are estimated using EKF at the *i*th moment of time from inertial satellite observations

$$Z_{C(i)} = [\varphi_i \lambda_i h_i]^{\mathrm{T}}_{\mathrm{SINS}} - [\varphi_i \lambda_i h_i]^{\mathrm{T}}_{\mathrm{GNSS}} ; \quad (6)$$

$$Z_{V(i)} = [V_E V_N V_H]^{\mathrm{T}}_{(i)\mathrm{SINS}} - [V_E V_N V_H]^{\mathrm{T}}_{(i)\mathrm{GNSS}}. \quad (7)$$

When satellite information is lost and a landmark is detected, SINS errors are estimated from optical-inertial observations, which are formed as the differences between the landmark sighting angles calculated from SINS information and measured by the OES

$$Z_{\alpha\beta(i)} = [\alpha_i \beta_i]^{\mathrm{T}}_{\mathrm{SINS}} - [\alpha_i \beta_i]^{\mathrm{T}}_{\mathrm{OES}}. \quad (8)$$

The EKF uses both observation signals (6) − (8) and their models of the form $Z_i = H_i x_i + \vartheta_i$. For observations (8), the connection matrix H_i with the SINS error vector x_i is formed by varying relations (3), (4) according to the parameters included in them. Here ϑ_i is the vector of observation errors.

When placing a video camera in a gimbal [11], it is necessary to take into account landmark sighting errors in the observations due to inaccurate stabilization of the measurement plane of azimuth α and elevation angle β. The relationship between the errors in measuring the viewing angles of the landmark $\Delta\alpha$, $\Delta\beta$ and the errors x_δ of the independent elements of DCM *C*, calculated by SINS and used to stabilize the video camera, follows from relations (2), (5) and has the form

$$[\Delta\alpha\ \Delta\beta]^{\mathrm{T}}_{(i)\mathrm{OES}} = C^*_{\alpha\beta} H_{\alpha\beta} x_{\delta(\mathrm{SINS})}, \quad (9)$$

где $C^*_{\alpha\beta} = (C^{\mathrm{T}}_{\alpha\beta} C_{\alpha\beta})^{-1} C^{\mathrm{T}}_{\alpha\beta}$;

$$C_{\alpha\beta} = \begin{bmatrix} -\sin\alpha\cos\beta & -\cos\alpha\sin\beta \\ 0 & \cos\beta \\ \cos\alpha\cos\beta & -\sin\alpha\sin\beta \end{bmatrix}; \ H_{\alpha\beta} =$$

$$\begin{bmatrix} e_N & c_{2,2}e_N & -c_{1,2}e_E & e_H & -c_{2,1}e_E & c_{1,1}e_E \\ -c_{2,2}e_E & e_N & c_{0,2}e_E & e_{2,1}e_E & e_H & -c_{0,1}e_H \\ c_{1,2}e_E & -c_{0,2}e_E & e_N & -c_{1,1}e_E & c_{0,1}e_E & e_H \end{bmatrix}$$

$$x_{\delta(\mathrm{SINS})} = [\Delta c_{0,1} \Delta c_{1,1} \Delta c_{2,1} \Delta c_{0,2} \Delta c_{1,2} \Delta c_{2,2}]^{\mathrm{T}} ;$$

$c_{i,j}$ are the elements of DCM *C*.

III. OPTICAL-INERTIAL OBSERVATIONS WHEN SIGHTING LANDMARKS WITH UNKNOWN COORDINATES

The inverse problem of trajectory measurements is considered. It is known [12,13] that the direct problem of such measurements is related to determining the flight path of a RPA using two separated on the Earth's surface OES with known geodetic coordinates. When solving the inverse problem, the coordinates of a landmark are determined using trajectory measurements from the RPA. Such measurements include RPA coordinates and landmark sighting angles at two points on the flight path. The coordinates of the RPA are determined using SINS, and the sighting angles of the landmark are determined using the OES. The problem of optical-inertial positioning is solved by the triangulation method. To do this, it is necessary to place the video camera in a gimbal and stabilize it in the horizontal plane. It is also necessary to stabilize the gimbal

in azimuth relative to its orientation at the initial measurement point. The diagram of trajectory measurements when implementing this method is shown in Fig. 1.

Fig. 1. Diagram of trajectory measurements by the triangulation method

In accordance with Fig. 1, the linear coordinates of the landmark X, Y, Z in the o_1xyz coordinate system associated with the video camera can be calculated using the following relations

$$X = dtg\alpha_2 / (tg\alpha_2 - tg\alpha_1);$$

$$Y = d\sin\alpha_2 \sin(\beta_1 + \mu_1/2) / [\sin(\alpha_2 - \alpha_1)\cos(\beta_1 + \mu_1)];$$

$$Z = dtg\alpha_2 tg\alpha_1 / (tg\alpha_2 - tg\alpha_1),$$

where d is the distance between points o_1 and o_2 corresponding to the two camera locations; α_1; β_1; α_2 and β_2 are the viewing angles of the landmark measured by the video camera at points o_1 and o_2.

The vertex of the trihedron o_1xyz coincides with the initial coordinates of the video camera φ_{o_1} and λ_{o_1}, on the flight path $o_1 - o_2$.

When implementing the approach under consideration to positioning a landmark, it is necessary to match the X, Y, Z coordinates with their geodetic equivalents. To do this, the specified coordinates are projected onto the axes of the geodetic accompanying trihedron $oENH$ using the RPV orientation angles calculated by the SINS. Next, taking into account the average value of the radius vector of the location of the video camera on the trajectory section $o_1 - o_2$, the geodetic coordinates of the landmark are determined. As a result, the SINS correction can be performed taking into account the measured sighting angles of the landmark at point o_2 by processing observations (8) using EKF.

IV. ANALYSIS OF THE RESULTS OF STUDIES

The study of the potential capabilities of the RPV navigation complex, including GNSS, SINS and OES, was carried out on a semi-natural model of the SINS-500NS inertial-satellite navigation system [14] of medium accuracy class. The experiments were carried out using flight data recorded on flash memory from GNSS and an inertial measurement unit (IMU), including three accelerometers and three fiber-optic gyros placed along orthogonal axes. The flight experiment was carried out on a helicopter. The flight parameters of the helicopter and the RPA were assumed to be identical.

Fig. 2 shows the flight path of the helicopter in the plan defined by GNSS, where $\Delta\varphi_R = [\varphi(t) - \varphi(t_0)]R$; $\Delta\lambda_R = [\lambda(t) - \lambda(t_0)]R\cos\varphi$; R is the value of the radius vector of the SINS location.

Fig. 2. Flight path in the plan

When modeling, the coordinates φ_{lm} and λ_{lm} for a landmark were formed using GNSS signals, namely: $\varphi_{lm} = \varphi_{GNSS} + \Delta\varphi_{lm}$; $\lambda_{lm} = \lambda_{GNSS} + \Delta\lambda_{lm}$; $R_{lm} = R_E$; R_E is the value of the Earth's radius at the landmark location. It was assumed that the video camera was rigidly fixed in the body of the RPA. Observations (6), (7) were used in the inertial-satellite mode, and observations (8) were used in the optical-inertial mode.

Fig. 3 shows the circular error ΔS of estimating the location of the RPA in the inertial-satellite mode of determining coordinates with a correction step from GNSS of 10 second, where

$$\Delta S = \sqrt{\delta_\varphi^2 + \delta_\lambda^2}; \quad \delta_\varphi = (\varphi_{SINS} - \varphi_{GNSS})R;$$

$$\delta_\lambda = (\lambda_{SINS} - \lambda_{GNSS})R\cos\varphi_{GNSS}.$$

Fig. 3. Circular error of estimating the location of the RPA in the inertial-satellite mode

The results were obtained taking into account the write-off of the error estimates formed by the EKF of both the flight and navigation parameters of the SINS and the drifts of the sensors of the IMU. It can be seen that in this mode of reckoning flight parameters, the positional error does not exceed 20 meters.

Fig. 4 shows the circular error in estimating the location of the RPA in a combined inertial-satellite ($1000\sec < t < 2400\sec$) and inertial mode with optical-visual correction ($t > 2400\sec$). The optical-visual correction step is 300 seconds with root-mean-square errors in determining the azimuth and elevation angle of a landmark of 20 arc minutes.

Fig. 4. Circular error in estimating the location of the RPA in a combined inertial-satellite and inertial mode with optical-visual correction

It can be seen that in this case the circular positional error of the inertial coordinate reckoning does not exceed 200 meters. A slight decrease in the dynamics of the increase in SINS errors is associated with the refinement and write-off of estimates of drifts of the IMU sensors obtained in the process of optical-inertial positioning.

V. CONCLUSIONS

The optical-electronic means included in the on-board equipment of the RPA can be used not only for control via a video channel, but also for estimating SINS errors in the event of loss of satellite information using observations of ground landmarks with known and unknown coordinates. It is proposed to form observation signals as the differences between measured and predicted sighting angles of landmarks based on SINS information. To estimate SINS errors, such observations are processed using EKF.

The conducted studies confirmed the possibility of optical-inertial positioning of RPA based on SINS of average accuracy with errors along the flight route at the level of 200 meters, taking into account correction intervals of about 5 minutes and writing off estimates of drift of SINS sensors.

REFERENCES

[1] Beard, R., and McLain, T., Small unmanned aerial vehicles: theory and practice. Princeton, NJ, Princeton University Press, 2012.

[2] Emelyantsev, G.I., and Stepanov, A.P., Integrirovannyie inertsialno-sputnikovyie sistemyi orientatsii i navigatsii (Integrated Inertial-Satellite Orientation and Navigation Systems), Peshekhonov, V.G., Ed., St. Petersburg, CSRI Electropribor, 2016.

[3] Noureldin, A., Karamat, T., and Georgy, J., Fundamentals of Inertial Navigation, Satellite-based Positioning and their Integration, Heidelberg: Springer-Verlag, 2013.

[4] Peshekhonov, V.G., High-Precision Navigation Independently of Global Navigation Satellite Systems Data. Gyroscopy and Navigation, 2022, vol. 13, pp.1–6. https://doi.org/10.1134/S2075108722010059

[5] Schmidt, G.T., GPS Based Navigation Systems in Difficult Environments, Gyroscopy and Navigation, 2019, vol. 10, no 2, pp. 41 - 53.

[6] Chernodarov, A.V. Monitoring and Adaptive Robust Protection of the Integrity of GNSS/SINS Observations in Urban Environments // 11th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFERPROCESS 2022), IFAC-PapersOnLine, 2022, 55(6), pp. 378–383.

[7] Maybeck, P.S., Stochastic Models, Estimation and Control. N.Y., Academic Press, 1982, vol. 2.

[8] Chernodarov, A.V., Gorshkov, P.S., Patrikeev, A.P., and Polyakova, A.A., Flight Development of an Integrated Navigation System Based on MEMS Sensors, Resistant to Unstable Satellite Information, 31st Saint Petersburg International Conference on Integrated Navigation Systems, ICINS 2024, St. Petersburg, CSRI Electropribor, 2024.

[9] Bromberg, P.V., Inertial Navigation Systems Theory. Moscow, Nauka, GRFML, 1979.

[10] GOST 20058-80. Dinamika letatel'nyh apparatov v atmosfere. Terminy, opredeleniya i oboznacheniya (Aircraft Dynamics in Atmosphere. Terms, Definitions and Symbols), Moscow, Izdatelstvo gosstandartov, 1981.

[11] Rivkin, S.S., Stabilizatsiya izmeritelnyih ustroystv na kachayuschemsya osnovanii (Stabilization of Measuring Devices on a Swinging Base), Moscow, Nauka, GRFML, 1978.

[12] Krynetskiy, E.I., Alexandrovskaya, L.N., Sharonov, A.V., and Golubkov, A.S, Letnyie ispyitaniya raket (Missile Flight Tests), Krynetskiy, E.I., Ed., Moscow, Mashinostroenie, 1979.

[13] Gorshkov, P.S., Patrikeev, A.P., Kharkov, V.P., and Chernodarov, A.V., Inertial Satellite Compensation of Trajectory Instabilities of Optoelectronic Positioning Systems on a Swinging Base, 27th Saint Petersburg International Conference on Integrated Navigation Systems, ICINS 2020, St. Petersburg, CSRI Electropribor, 2020, IEEE, 9133859.

[14] Chernodarov, A.V., Gorshkov, P.S., Patrikeev, A.P., and Starostin N.P., Investigation of the Emergency Mode of the SINS-500NS Strapdown Inertial-Satellite Navigation System Based on Flight Data, 30th Saint Petersburg International Conference on Integrated Navigation Systems, ICINS 2023, St. Petersburg, CSRI Electropribor, 2023, IEEE Xplore: 06 July 2023.

Flight Testing of an Integrated Navigation System Based on MEMS Sensors, Resistant to Unstable Satellite Information

A.V. Chernodarov

"NaukaSoft" Research & Production Association, Ltd.,
MAI University, Moscow, Russia
e-mail: chernod@mail.ru

P.S. Gorshkov

"NaukaSoft" Research & Production Association, Ltd.,
Moscow, Russia
e-mail: contacts@naukasoft.ru

A.P. Patrikeev

"NaukaSoft" Research & Production Association, Ltd.,
Moscow, Russia,
e-mail: apatrikeev@naukasoft.ru

A.A. Polyakova

"NaukaSoft" Research & Production Association, Ltd.,
Moscow, Russia,
e-mail: apolyakova@naukasoft.ru

Abstract — **This paper is devoted to the problem of increasing the information integrity of navigation systems in conditions of instability of satellite data. The proposed solution to the problem is based on the inertial support of the global navigation satellite system (GNSS). Such support is associated with the determination of the angular orientation of the aircraft in normal mode, as well as with the localization and parrying of violations in the GNSS in emergency mode. Typical violations include abnormal interference or loss of satellite signals. The results of full-scale experiments with an inertial-satellite navigation system based on micro-electro-mechanical sensors are analyzed.**

Keywords — satellite navigation system, inertial navigation system, information integrity, extended Kalman filter, micro-electro-mechanical sensors.

I. INTRODUCTION

The current state of onboard equipment of mobile objects is characterized by the use of integrated inertial satellite navigation systems (ISNS) [1,2]. Integration is based on the difference in the spectral characteristics of errors of strapdown inertial (SINS) and global navigation satellite systems (GNSS). This difference provides mutual information support for SINS and GNSS in a real interference environment. At the same time, the problem of reducing the weight, dimensions and cost of the ISNS hardware while maintaining accuracy characteristics remains actual. A possible solution to this problem is associated with the inclusion of inertial measurement modules (IMUs) based on micro-electro-mechanical systems (MEMS) into the ISNS. However, MEMS sensors have low accuracy, a large insensitivity zone, and unstable parameters of the drift models of gyros and accelerometers. It seems that the contradiction between the necessary accuracy and the required mass and dimensional characteristics of the MEMS-based ISNS can be resolved by improving the software and mathematical support of the ISNS. One of the areas of such improvement is associated with estimating the errors of MEMS sensors and identifying their dynamic models in the presence of correct satellite information. Dynamic models take into account changes in the parameters of MEMS sensors during operation, and also provide the required accuracy characteristics of the ISNS in an autonomous inertial mode in the event of loss of GNSS signals [3]. It should be noted that in the process of factory bench calibration, as a rule, only the systematic errors of MEMS sensors are estimated. This calibration does not take into account MEMS errors that arise in the dynamic operating modes of the SINS associated with linear and angular accelerations.

Inertial information can be used to parry off natural and simulated interference in GNSS. At the same time, using satellite information, it is possible to restore the parameters of the SINS when they are lost in difficult abnormal operating conditions. Such conditions are associated, for example, with the loss of electrical power, as well as with exceeding the measurement ranges of the INS sensors. Exceeding the sensors measurement ranges may be due to aircraft maneuvering.

The purpose of the paper is to study the following modes using recorded MEMS and GNSS data:

• additional SINS alignment in flight using GNSS information;

• estimation of residual MEMS drifts in flight using GNSS information;

• on what time intervals the necessary accuracy of inertial calculation of flight parameters is ensured in the event of deterioration or loss of GNSS signals.

The achievement of this purpose is based on the use of inertial-satellite and inertial-geophysical observations, as well as the extended Kalman filter (EKF) [4,5] for estimating and predicting SINS errors.

The developed software and mathematical support of the ISNS was studied in post-processing using the recorded data. The following stages and operating modes of the ISNS were considered:

- preliminary ground-based MEMS calibration to estimation systematic sensor errors;
- a coarse initial alignment of SINS using signals from MEMS sensors by the analytical gyrocompassing method. It should be noted that in this mode, the true course is determined with a large error by the signals of the MEMS sensors;
- the fine initial alignment of SINS using observations of geophysical invariants and EKF;
- additional alignment of SINS using track angle from GNSS, inertial satellite observations and EKF;
- inertial-and-satellite navigational mode and autonomous inertial mode in the absence of GNSS signals. In semi-natural modeling, the absence of GNSS signals is implemented programmatically.

II. MEMS-Based Strapdown Inertial Satellite Navigation System as the Object of Studies

The ISNS based on MEMS LSM6DS33 is considered as an object of study [6]. ISNS signals are recorded on flash memory via an STM32 microcontroller [7] with a frequency of 1.666 kHz for MEMS and 5 Hz for GNSS. The indicated ISNS modules are presented in Figs. 1 and 2.

Fig. 1. MEMS board LSM6DS33 (10×23×3mm)

Fig. 2. Processor board with STM32 microcontroller (14x14mm)

The implementation of the software and mathematics support of SINS as part of the ISNS is based on solving the basic equation of inertial navigation [8,9]

$$\dot{V} = C_3^{\mathrm{T}} \bar{a}_{xyz} + \bar{g} - 2\bar{\Omega} \times \bar{V} - \bar{\omega} \times \bar{V} - \bar{\Omega} \times (\bar{\Omega} \times \bar{R}), \quad (1)$$

where $\bar{a} = [a_x a_y a_z]^{\mathrm{T}}$ is the vector of accelerometer signals in projections on the axes of the body-fixed frame $oxyz$;

$\bar{g} = \bar{g}_{\mathrm{gr}} - \bar{\Omega} \times (\bar{\Omega} \times \bar{R})$ is the vector of gravity acceleration; $\bar{g}_{\mathrm{gr}} = [g_{\mathrm{gr}(E)} g_{\mathrm{gr}(N)} g_{\mathrm{gr}(H)}]^{\mathrm{T}}$ is the vector of gravitational acceleration; $\bar{\Omega}_{ENH} = [\Omega_E \Omega_N \Omega_H]^{\mathrm{T}}$ is the vector of angular velocity of Earth rotation; $\Omega = \|\bar{\Omega}\|_2^2$; $\bar{R} = [0 \, 0 \, (R + h)]^{\mathrm{T}}$ is the radius vector of the location of the inertial measurement unit (IMU) as part of the MEMS. Such IMU includes three accelerometers and three angular velocity sensors (AVS), located along orthogonal axes; R; h is, respectively, the radius of the Earth and the height above the Earth's ellipsoid at the location of the IMU; $(\times)$ is the operator of vector product; $\bar{\omega}_{ENH} = [\omega_E \omega_N \omega_H]^{\mathrm{T}}$ is the vector of angular velocity flyby around the Earth, the elements of which are determined from the projections $V_E; V_N; V_H$ [8]:

$$\omega_E = -\dot{\varphi} = -V_N/R_\varphi; \quad \omega_N = \dot{\lambda}\cos\varphi = V_E/R_\lambda;$$

$$\omega_H = \dot{\lambda}\sin\varphi = V_E \mathrm{tg}\varphi/R_\lambda; \quad (2)$$

R_φ; R_λ are the radii of curvature of the earth's ellipsoid;

φ, λ are the geodetic latitude and longitude. In equation (1), the elements of all vectors, except $\bar{a}$, are determined in projections on the axes of the accompanying trihedron $oENH$ of the geodetic coordinate system.

The direction cosines matrix (DCM) C_3 characterizes the angular orientation of the IMU axes ox, oy, oz relative to the trihedron $oENH$. The C_3 matrix is determined from the solution of the Poisson equation [9], which has the form

$$\dot{C}_3 = \Pi_0 C_3 - C_3 P_3 \Pi_1 P_3^{\mathrm{T}}, \quad (3)$$

where

$$\Pi_0 = \begin{bmatrix} 0 & \dot{\Theta}_z & -\dot{\Theta}_y \\ -\dot{\Theta}_z & 0 & \dot{\Theta}_x \\ \dot{\Theta}_y & -\dot{\Theta}_x & 0 \end{bmatrix}; \quad P_3 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix};$$

$$\Pi_1 = \begin{bmatrix} 0 & -(\omega_H + \Omega_H) & (\omega_N + \Omega_N) \\ (\omega_H + \Omega_H) & 0 & -\omega_E \\ -(\omega_N + \Omega_N) & \omega_E & 0 \end{bmatrix};$$

$\dot{\bar{\Theta}} = [\dot{\Theta}_x \dot{\Theta}_y \dot{\Theta}_z]^{\mathrm{T}}$ is the AVS signals vector; $\bar{\Omega}_{ENH} = [0 \; \Omega\cos\varphi \; \Omega\sin\varphi]^{\mathrm{T}}$.

Using the elements $C_{3(i,j)}$ of the C_3 matrix, the orientation angles of the IMU relative to the trihedron $oENH$ are determined, namely: true heading ψ, pitch angle ϑ and roll angle γ:

$$\psi = arctg[-\frac{C_{3(1,0)}}{C_{3(1,1)}}]; \quad \vartheta = arctg\frac{C_{3(1,2)}}{\sqrt{C_{3(0,2)}^2 + C_{3(2,2)}^2}};$$

$$\gamma = arctg[-\frac{C_{3(0,2)}}{C_{3(2,2)}}]. \tag{4}$$

It can be seen from equations (1) – (3) that due to the drifts of the accelerometer and AVS signals, as well as due to the inaccurate determination of gravity acceleration, the SINS errors when integrating these equations have an increasing, oscillatory character with the Schuler period [10]. SINS errors are estimated and compensated during the initial alignment using geophysical invariants [11], and in navigation mode using information from GNSS.

A typical scheme for estimating the SINS error vector, based on the use of GNSS and EKF, is presented in Fig. 3,

Fig. 3. Typical scheme for estimating the SINS error vector

where Y is a vector of flight and navigation parameters (FNP) formed at the i-time; $\hat{x}_{i/i-1}$; $\hat{Y}_{i/i-1}$; $\hat{x}_{i/i}$; $\hat{Y}_{i/i}$ are, respectively, predicted and adjusted estimates of the parameters and error vector of the SINS; $h(y)$ is a function that matches the parameters of the SINS and the external corrector; CC is a coordinate converter; $^\wedge$ is the symbol for estimate; $Z_i = [Z_{k(i)} Z_{V(i)}]^T$ is a vector of inertial-and-satellite observations, the elements of which at the i-th moment of time have the form

$$Z_{k(i)} = [\varphi_i \lambda_i h_i]^T_{SINS} - [\varphi_i \lambda_i h_i]^T_{GNSS}; \tag{5}$$

$$Z_{V(i)} = [V_E V_N V_H]^T_{(i)SINS} - [V_E V_N V_H]^T_{(i)GNSS}. \tag{6}$$

A rough initial alignment of SINS is performed using signals from calibrated MEMS sensors. In the fine initial alignment mode, observations of geophysical invariants associated with the immobility of the SINS base are used, namely: zero relative velocities, constant coordinates of the initial alignment point (IAP) and projections of the angular velocity vector of the Earth's rotation on the axes of the inertial coordinate system $O_I XYZ$

$$Z_{k(i)} = [\varphi_i \lambda_i h_i]^T_{SINS} - [\varphi_i \lambda_i h_i]^T_{IAP}; \tag{7}$$

$$Z_{V(i)} = [V_E V_N V_H]^T; \tag{8}$$

$$Z_{\bar{\Omega}(i)} = \Delta\hat{\bar{\Omega}}_i - \Delta\bar{\Omega}_{XYZ(i)}, \tag{9}$$

where $\Delta\hat{\bar{\Omega}}_i = \int_{t_{i-1}}^{t_i} C_0^T(\tau)\dot{\hat{\Theta}}(\tau)d\tau$;

$\Delta\bar{\Omega}_{XYZ(i)} = [0 \vdots 0 \vdots \Omega\Delta t_i]^T$ is the vector of rotation angles of the Earth during the $\Delta t_i = t_i - t_{i-1}$ time in projections on the axes of the inertial coordinate system;

$C_0 = C_3 P_3 C_2 C_1$ is a matrix characterizing the orientation of the IMU relative to the Earth centered inertial frame;

$$C_1 = \begin{bmatrix} \cos\Omega t & \sin\Omega t & 0 \\ -\sin\Omega t & \cos\Omega t & 0 \\ 0 & 0 & 1 \end{bmatrix};$$

$$C_2 = \begin{bmatrix} \cos\lambda\cos\varphi & \sin\lambda\cos\varphi & \sin\varphi \\ -\sin\lambda & \cos\lambda & 0 \\ -\cos\lambda\sin\varphi & -\sin\lambda\sin\varphi & \cos\varphi \end{bmatrix}.$$

The implementation of the scheme presented in Fig.3 is based on the SINS and GNSS error models. The SINS errors model is obtained by varying equations (1) - (3) according to the parameters included in them. Error models of MEMS sensors are described by first-order shaping filter equations [12]. For the presented equations, the SINS error vector includes 18 parameters, namely: errors in the reckoning of projections of the trajectory velocity vector on the axes of the accompanying trihedron $oENH$; errors in reckoning of latitude, longitude and altitude above the Earth's ellipsoid; errors in reckoning of independent elements of the C_3 matrix and drifts of MEMS sensors. The observation errors associated with GNSS are treated as Gaussian. The presented equations form the basis of the regular ISNS mode. If satellite information is lost or deteriorated, the ISNS must switch to the inertial FNP reckoning mode. Such a transition can be realized by monitoring the corresponding diagnostic parameters [13].

When the ISNS switches to the autonomous inertial navigation mode, the error estimates of the MEMS sensors, formed before the loss of satellite information, and also refined during the processing of inertial-and-geophysical observations of the form [11] are compensated:

$$Z_{\bar{\Theta}(i)} = \int_{t_{i-1}}^{t_i} \{C_{0(i)}^T \dot{\bar{\Theta}}(\tau)d\tau - C_{2(i)}^T [\bar{\omega}_i(\tau) + \dot{\bar{\delta}}_i(\tau)]\}d\tau -$$
$$-[0 \vdots 0 \vdots \Omega\Delta t_i]^T, \tag{10}$$

where $\dot{\bar{\delta}} = [\dot{\delta}_E \dot{\delta}_N \dot{\delta}_H]^T$ is the vector of the angular velocity of rotation of the IMU relative to the accompanying trihedron $oENH$;

$$\dot{\delta}_E = \dot{\vartheta}\cos\psi - \dot{\gamma}\sin\psi\cos\vartheta;$$

$$\dot{\delta}_N = \dot{\vartheta}\sin\psi + \dot{\gamma}\cos\psi\cos\vartheta;$$

$$\dot{\delta}_H = \dot{\psi} + \dot{\gamma}\sin\vartheta.$$

The derivatives for the orientation angles are determined from equations (3), (4).

III. ANALYSIS OF THE RESULTS OF STUDIES

Before conducting the experiments, preliminary calibration of the MEMS sensors was performed under stationary conditions to estimate systematic drifts. The timing diagram of operation of the ISNS included the following stages: coarse initial alignment of SINS-MEMS ($t = 0 \div 100$sec); fine initial alignment (100sec $< t \leq$ 256sec) with observation (7), (8) processing step 1Hz; inertial-and-satellite navigational mode (256sec $< t \leq$ 1080sec) with observation (5), (6) processing step 1Hz; autonomous inertial navigation mode (1080sec $< t \leq$ 1260sec) with observation (10) processing step 1Hz. The initial alignment was carried out in an accelerated version, taking into account the additional alignment using the track angle from GNSS in the inertial-and-satellite navigational mode. It was assumed that GNSS signals were unavailable at the final part of the trajectory for 3 minutes. In this section of the trajectory, inertial dead reckoning of flight parameters was performed with compensation of predicted SINS-MEMS error estimates obtained in the inertial-and-satellite navigational mode.

Figs. 4, 5 show the circular position errors of ISNS ΔS, namely: in Fig. 4, taking into account the compensation of sensors drifts estimates stored at $t = 1080$sec; in Fig. 5, taking into account observation (10) and compensation of predicted $\hat{x}_{i/i-1}$ estimates in inertial mode, where

$$\Delta S = \sqrt{\delta_\varphi^2 + \delta_\lambda^2} \ ; \ \delta_\varphi = (\varphi_{\text{SINS}} - \varphi_{\text{GNSS}})R_\varphi \ ;$$

$$\delta_\lambda = (\lambda_{\text{SINS}} - \lambda_{\text{GNSS}})R_\lambda \cos\varphi_{\text{GNSS}} \ ;$$

R_φ ; R_λ are the radii of curvature of the earth's ellipsoid.

ΔS, m

Fig. 4. Circular error of ISNS with compensation of sensors drifts estimates stored at $t = 1080$ sec in inertial mode

ΔS, m

Fig 5. Circular error of ISNS taking into account the use of observation (10) and compensation of the predicted estimates $\hat{x}_{i/i-1}$ in inertial mode

It can be seen that compensation of the predicted SINS-MEMS error estimates in the section of autonomous inertial dead reckoning of flight parameters made it possible to reduce the SINS circular position error by almost two orders of magnitude on a time interval of 3 minutes.

IV. CONCLUSIONS

To ensure the resistibility of the integrated navigation system on MEMS sensors to the instability of satellite information, it is necessary to promptly detect abnormal GNSS signals and switch to an autonomous inertial mode of operation of the ISNS. Taking into account the low accuracy characteristics of MEMS sensors, their inclusion in the ISNS can be based on the implementation, inter alia, of the following procedures for combined information processing:

- ground-and-onboard calibration of MEMS sensors [14];

- estimation of MEMS sensor drifts in inertial-satellite mode and compensation of their predicted values in the mode of autonomous dead reckoning of flight parameters in the absence of GNSS information;

- estimation and compensation of SINS errors in autonomous mode by processing inertial-and-geophysical observations of the form (10) using EKF.

REFERENCES

[1] Emelyantsev, G.I., and Stepanov, A.P., Integrirovannyie inertsialno-sputnikovyie sistemyi orientatsii i navigatsii (Integrated Inertial-Satellite Orientation and Navigation Systems), Peshekhonov, V.G., Ed., St. Petersburg, CSRI Electropribor, 2016.

[2] Noureldin, A., Karamat, T., and Georgy, J., Fundamentals of Inertial Navigation, Satellite-based Positioning and their Integration, Heidelberg: Springer-Verlag, 2013.

[3] Peshekhonov, V.G., High-Precision Navigation Independently of Global Navigation Satellite Systems Data. Gyroscopy and Navigation, 2022, vol. 13, pp.1–6. https://doi.org/10.1134/S2075108722010059

[4] Maybeck, P.S., Stochastic Models, Estimation and Control, New York., Academic Press, 1982, vol. 2.

[5] Stepanov, O.A., Osnovyi teorii otsenivaniya s prilozheniyami k zadacham obrabotki navigatsionnoy informatsii. Ch.1. Vvedenie v teoriyu otsenivaniya (Fundamentals of Estimation Theory with Applications to Navigation Information Processing Problems. Introduction to the Estimation Theory. P.1), St. Petersburg, CSRI Electropribor, 2017.

[6] LSM6DS33. Micro-emiconductor.com›datasheet..LSM6DS33TR.pdf

[7] STM32. ru.wikipedia.org›STM32

[8] Babich, O.A., Obrabotka informatsii v navigatsionnykh kompleksakh (Data Processing for Navigation Complexes), Moscow, Mashinostroenie, 1991.

[9] Savage, P.G., Strapdown Analytics. Maple Plan, Minnesota, Strapdown Associates Inc., 2007, Part 1.

[10] Klimov, D.M., Inertsialnaya navigatsiya na more (Inertial Navigation at Sea), Moscow, Nauka, 1984.

[11] Chernodarov, A.V., Patrikeev, A.P., and Khalyutina, O.S., Sposob povysheniya tochnosti besplatformennoj inercial'noj navigacionnoj sistemy (Method for Improving the Accuracy of Strapdown Inertial Navigation System), RF Patent № 2725029, 2020.

[12] Stepanov, O.A., Osnovyi teorii otsenivaniya s prilozheniyami k zadacham obrabotki navigatsionnoy informatsii. Ch.1. Vvedeniye v teoriyu filtratsii (Fundamentals of Estimation Theory with Applications to Navigation Information Processing Problems. Introduction to the Filtering Theory. P.2), St. Petersburg, CSRI Electropribor, 2012.

[13] Chernodarov, A.V., Monitoring and Adaptive Robust Protection of the Integrity of GNSS/SINS Observations in Urban Environments, 11th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFERPROCESS 2022), IFAC-PapersOnLine, 2022, 55(6), pp. 378–383.

[14] Chernodarov, A., Patrikeev, A., In-Motion Calibration and Testing of MEMS Sensors Using a Reference Inertial Satellite Navigation System, International Seminar on Electron Devices Design and Production, SED 2021, IEEE Proceedings, 2021, 9444510.

On the Navigation Problem of Unmanned Wheeled Agricultural Machinery Using MEMS-Based INS, GNSS and Odometry

D.I. Smolyanov

Lomonosov Moscow State University,
JSC «Cognitive»
Moscow, Russia
danil.smolyanov@yandex.ru

A.A. Golovan

Lomonosov Moscow State University
Moscow, Russia
aagolovan@yandex.ru

Abstract—The paper describes the development, testing, and implementation of navigation algorithms for unmanned wheeled agricultural vehicles when the navigation equipment consists of a low-grade inertial navigation system (INS) with microelectromechanical sensors (MEMS), a receiver of Global Navigation Satellite Systems (GNSS), and odometry data — odometer-derived velocity and a steering sensor.

The navigation problem is one of the most important in the operation of unmanned agricultural machinery. The positioning accuracy of the corresponding equipment needs to be a few centimetres for the treatment of most agricultural crops. High accuracy is also important to reduce fuel costs and ensure safe operation of agricultural machinery for humans. The reliability of the navigation solution depends on its stability to outliers and short-term losses of GNSS data.

The paper considers two approaches to GNSS/INS/Odometry fusion: loose and tight integrations. Loose integration implies two simultaneously operating estimation algorithms, and tight integration – one algorithm with extended state vector.

Also, the work is devoted to a method of INS velocity fusion based on kinematic model properties of navigation objects – so-called component-wise ZUPT (Zero-Velocity Update Technology). Mentioned fusion does not require new sources of navigation information, but is based on the property of the movement of the ground vehicle. The relevant covariance analysis shows the observability of accelerometer biases in fusion proposed. The paper presents two types of equipment configuration that are structurally different: non-articulated (classic) and articulated. The classic configuration implies a four-wheeled vehicle, which is steered by the front or the rear axle. The articulated configuration implies two frames connected by a joint, each frame has two wheels fixed to the corresponding frame. In this configuration, one can make a turn by changing the angle between the frames.

Keywords — INS, MEMS, odometry, ZUPT, UGV, unmanned vehicle, agricultural machinery

I. INTRODUCTION

JSC «Cognitive» is engaged in the development and production of motion control systems for agricultural machinery with a high level of autonomy. In the operation of unmanned agricultural machinery, the navigation problem plays the most significant role. The positioning accuracy of the corresponding equipment needs to be a few centimetres for the treatment of most agricultural crops. High accuracy also helps to reduce fuel costs and ensure safe operation of agricultural machinery for humans [1]. In addition, a navigation solution must be reliable, which means that the solution is resistant to outliers and short-term losses of GNSS data, especially in the case of a low-precision INS.

In addition, the system provides a special mode of operation without GNSS. In this case, the INS-GNSS integration positional solution is not used, the latter does not eliminate the need of solving the navigation problem, even when using a low-grade MEMS-based INS. In particular, estimates of the constant odometry coefficients are required for satisfactory solution of the automatic control problem. The appropriate integration filter delivers mentioned estimates.

II. PROBLEM STATEMENT

Navigation equipment includes a low-grade INS, a single antenna GNSS receiver (subsequently referred to as GNSS), an odometer, and a steering sensor. There is a case when odometer data are not available for some technical reasons. This case requires a separate examination.

INS is a MEMS-based (rate noise spectral density is 0.01 °/s/√Hz, acceleration noise spectral density is 190 µg/√Hz, zero-rate output (ZRO) drift is ±0.8 °/s). The main feature of the use of MEMS sensors is the strong influence of temperature on their readings, so the navigation system includes a thermal stabilization system for MEMS sensors, which makes it possible to increase the accuracy and stability of their readings [2, 3]. The inertial sensors and the GNSS antenna are located on the roof of the vehicle. In addition, the system includes a stereo camera installed on the roof of the vehicle. Owing to machine vision technologies, the unmanned control system is capable to determine the boundary of a ploughed area, edge, row, windrow etc. and control the vehicle without GNSS. However, accurate estimates of the odometry coefficient constants are essential for a satisfactory solution to the automatic control problem. The same applies to accelerometer biases and angular rate sensors (ARS) drifts, which are determined during the process of solving the navigation problem. The use of the stereo camera for navigation is beyond the scope of this work.

Fig. 1. Camera view to the boundary of a ploughed area

The wheeled agricultural machinery under examination includes tractors and combine harvesters. Most tractors and combines have the classic configuration, i.e., a four-wheeled vehicle, which is steered by the front or the rear axle. Most tractors are steered by the front axle, while the rear axle steers most combines. Tractors with the articulated configuration require a separate examination. The articulated configuration implies two frames connected by a joint, each of which has two wheels fixed to the corresponding frame. In this configuration, one can make a turn by changing the angle between the frames. An example of this configuration is the Kirovets K-7M tractor manufactured by the JSC «Petersburg Tractor Plant».

Fig. 2. Kirovets K-7M with the Cognitive Agro Pilot system

The steering sensor operation is based on the Hall effect. In the case of the classic configuration, the sensor is installed on the steering axle and connected to one of the steering wheels. In the case of articulated composition, the sensor is installed on the joint.

III. THE KINEMATIC MODELS

The models under consideration postulate the position of the heading point. The heading point is a point whose velocity vector has direction along the longitudinal axis of the vehicle without any lateral slip. This property can be described by a non-holonomic constraint of the following form

$$\dot{x}_1 cos\psi - \dot{x}_2 sin\psi = 0, \tag{1}$$

where x_1, x_2 are the horizontal coordinates, ψ is the heading angle. This property is used to derive velocity aiding measurements in the event of a loss of GNSS signal. We will discuss this in more details below. In addition, this property can be used for aiding of the heading angle, because the track angle of this point almost coincides with the heading angle.

A. The bicycle model

The bicycle kinematic model [4, 5] is a sufficiently good approximation of classic vehicles without lateral slip.

In the bicycle model, two wheels on the same axle are replaced by one virtual wheel on both the turning and the non-turning axles. A rigid rod replaces the vehicle body. It is assumed that there is no lateral slip.

The bicycle model postulates the location of the heading point on the centre of the non-turning axle.

Fig. 3. The bicycle model

The basic kinematic equation of the bicycle model is [4, 5]

$$\dot{\psi} = \frac{V}{L} tg\delta \cdot (cos\theta \, cos\gamma), \tag{2}$$

where ψ, θ, γ are the heading, pitch and roll angles respectively, δ is the steering angle, V is the heading point velocity, L is the rod length (wheelbase).

B. The articulated model

The articulated kinematic model [6, 7] is a sufficiently good approximation of articulated vehicles without lateral slip.

Fig. 4. The articulated model

In the articulated model, each frame and each axle are replaced by rigid rods. Each axle has two wheels which cannot turn with respect to the corresponding frame. We assume that there is no lateral slip.

The articulated model postulates two heading points, that are the centres of the front and rear axles. Further, we will only consider the front one.

The basic kinematic equation of the articulated model is [6, 7]

$$\dot{\psi}_1 = \left(\frac{sin\delta}{l_2+l_1 cos\delta}V_1 + \frac{l_2}{l_2+l_1 cos\delta}\dot{\delta}\right)cos\theta \, cos\gamma, \tag{3}$$

where ψ_1, θ, γ are the heading, pitch and roll angles of the front frame respectively, δ is the steering angle (the angle between the frames), V_1 is the velocity of the front heading point, l_1, l_2 are the lengths of the front and rear frames respectively.

IV. THE NAVIGATION PROBLEM

The traditional approach to solving this problem proposes the usage of geodetic reference frame as navigation frame in INS dead reckoning. Next, the fusion of INS and GNSS data is implemented through feedbacks based on the estimates of the corresponding Kalman filter (KF). Methodically, the fusion comes down to solving the estimation problem (for illustration in continuous time).

$$\dot{y} = Ay + q, z = Hy + r, \qquad (4)$$

where y is the state vector, A is the matrix that corresponds to the INS linear error equations used and to the model of the odometer and steering instrumental errors, q, r are vector-valued zero-mean white-noise processes, z is the measurement vector, generated by GNSS and odometer data, H is the corresponding observation matrix.

A. Odometry errors model

The linear error model for steering δ is assumed to have a null bias k_0^δ, a scale factor k_1^δ and a stochastic error $\Delta \delta_s$:

$$\delta' = \delta + k_0^\delta + k_1^\delta \delta + \Delta \delta_s . \qquad (5)$$

Note that in the case of the classic composition, the steering sensor is connected to only one of the steering wheels. Therefore, it is important to consider the Ackermann principle [8], which states that during a turn, the inner wheel of a vehicle turns through a larger angle than the outer wheel, as it travels along a shorter arc.

The model of the odometer velocity measurement errors under consideration includes the scale factor error k^v and the misalignment angles $\varkappa_1, \varkappa_3$ between the «measuring» axis of the odometer and the INS instrument axes [9]:

$$V_s' = V_z + V'(-\varkappa_3, k^v, \varkappa_1)^T + \Delta V_s^s , \qquad (6)$$

where $V_s' = (0, V', 0)^T$ is the vector of the velocity measurement from the odometer in the vehicle's frame Ms, V_z is the ideal velocity vector in the INS frame Mz, ΔV_s^s is the stochastic error.

Coefficients k_0^δ, k_1^δ, k^v, $\varkappa_1$, $\varkappa_3$ are assumed to be constant.

B. Loose INS/Odometry integration

The loose INS/Odometry integration implies two simultaneously working reckonings — inertial and odometric, as well as two integrated estimation problems. Inertial reckoning, as already mentioned, is based on the traditional approach. Odometric reckoning is reduced to the numerical integration of kinematic equations (2) and (3), supplemented by positional equations

$$\dot{x}_1 = V_1 sin\psi_1, \dot{x}_2 = V_1 cos\psi_1, \qquad (7)$$

in the case of a bicycle model, the indices for V and ψ are omitted.

In addition to two reckonings, two estimation problems are solved: inertial with 18-dimensional state vector

$$y_1 = (\Delta x^T, \delta V^T, \alpha_1, \alpha_2, \beta_3, k^v, \varkappa_1, \varkappa_3, v_0^T, \Delta f_0^T)^T, \qquad (8)$$

and odometric with 5-dimentional state vector

$$y_2 = (\Delta x^*_1, \Delta x^*_2, \Delta \psi^*, k^{v*}, k_0^\delta, k_1^\delta), \qquad (9)$$

where Δx^T is INS position errors, δV^T is the vector of INS dynamic velocity errors, α_1, α_2 indicate the deflections of virtual horizon, β_3 is the azimuth attitude error, k^v is the odometer scale factor error, $\varkappa_1, \varkappa_3$ are the misalignment angles between the «measuring» axis of the odometer wheel and the INS instrument axes, v_0 is the vector of the ARS drifts, Δf_0 is the biases vector of the accelerometers.

The two state vectors shown have common components, but the values of the estimates of these components may differ; the subscript (*) is used to demonstrate this fact.

The dynamic equations for y_1 are based on traditional INS error equations [10]. Dynamic equations for y_2 are based on linearized odometric reckoning error equations (for bicycle model as an example)

$$\begin{aligned} \Delta \dot{x}^*_1 &= V sin(\psi) k^{v*} + V cos(\psi) \Delta \psi^* \\ &\quad + sin(\psi) \Delta V^s, \\ \Delta \dot{x}^*_2 &= V cos(\psi) k^{v*} - V sin(\psi) \Delta \psi^* \\ &\quad + cos(\psi) \Delta V^s, \\ \Delta \dot{\psi}^* &= \frac{Vtg\delta}{L} k^{v*} + \frac{V}{Lcos^2\delta} k_0^\delta + \frac{V\delta}{Lcos^2\delta} k_1^\delta + \\ &\quad + \frac{tg\delta}{L} \Delta V^s + \frac{V}{Lcos^2\delta} \Delta \delta^s. \end{aligned} \qquad (10)$$

The dynamic equations for the articulated model are obtained similarly and have a similar structure, but are omitted here due to their cumbersomeness.

Both estimation problems allow GNSS aiding. In addition, one estimation problem can use estimates from another as aiding measurements, which, in particular, allows (as practice has shown) to better estimate the scale factor and the zero offset of the steering angle sensor.

This approach can be called traditional, combining two traditional approaches to ground-based navigation — inertial and odometric.

The advantage of this approach is the independent execution of two parallel estimation tasks, which increases the reliability in the event of failure of one of the components

C. Tight INS/Odometry integration

The tight INS/Odometry integration implies one reckoning — inertial dead reckoning. In this case, the estimation problem is also single, but with an extended state vector y

$$y = (\Delta x^T, \delta V^T, \alpha_1, \alpha_2, \beta_3, k^v, \varkappa_1, \varkappa_3, k_0^\delta, k_1^\delta, v_0^T, \Delta f_0^T)^T. \qquad (11)$$

Odometry data is used as aiding measurements in this approach. The odometer velocity measurement model [7] is

$$\underline{z}_v = \underline{V}'^{ins} - \underline{V}'^{odo} = V'^{odo} \cdot (-\varkappa_3, k^v, \varkappa_1)^T + \underline{\Delta V_s} , \qquad (12)$$

where $\underline{V}'^{ins}$ is the estimated velocity vector, V'^{odo} is the scalar odometer measurement, $\underline{V}'^{odo} = (0, V'^{odo}, 0)^T$, $\underline{\Delta V_s}$ is the stochastic velocity error vector.

The steering angle sensor measurement model is

$$z_\delta = \omega'^{ins} - \omega'^{odo} = -\nu_3 - \Delta\dot\psi, \tag{13}$$

where ω'^{ins} is the «vertical» angular rate in the vehicle's body frame (ARS measurement), $\omega'^{odo} = \dot\psi$ appears from equation (2), ν_3 is the drift of the «vertical» ARS, and $\Delta\dot\psi$ appears from equation (10). By «vertical» axis we mean a vehicle-related axis that is approximately vertical when the roll and pitch angles are small.

Tight integration with odometry also enables INS-GNSS aiding.

The advantage of this approach is that it eliminates the additional computational costs associated with the redundancy of estimation problems that arise from a loose integration approach.

D. Equivalence of loose and tight integration

Note that the above loose and tight integration approaches are equivalent in terms of the behaviour of the accelerometers biases and odometry coefficients estimates.

Fig. 5. Estimate of $k_0{}^\delta$ zero offset of the steering angle sensor

Fig. 6. Estimate of $k_1{}^\delta$ scale factor of the steering angle sensor

In addition, the accelerometer biases and odometry coefficients are estimated together even without the use of GNSS.

V. COMPONENT-WISE ZUPT

Here, component-wise ZUPT (Zero-Velocity Update Technology) means aiding by zero components of the heading point velocity vector. This method allows aiding even in the absence of odometer measurements, and, moreover, does not require stopping, unlike the classic ZUPT method.

Let us recall that the velocity vector of the heading point is directed along the longitudinal axis of the vehicle, i.e. $V_s^p = (0, V^p, 0)^T$. Where V^p is the unknown longitudinal velocity. Let us present a vector of a virtual velocity measurement (lateral and «vertical» velocities in vehicle-related frame are equal to zero)

$$z_s = A_{sx}\tilde V_x - V_s^p - \hat\omega_s \Delta R_s, \tag{14}$$

where A_{sx} is the transition matrix from the geodetic reference frame Mx to the vehicle's body frame Ms, $\tilde V_x$ is the current estimate of the velocity vector of the INS centre in the frame Mx, $\hat\omega_s$ is a skew-symmetric matrix that corresponds to the angular rate vector ω_s, ΔR_s is the vector connecting the INS centre with the heading point. The angular rate is measured by the ARS and corrected for drift.

On the basis of (8) or (11) one can form the following model for velocity-derived measurement vector

$$z_s = A_{sx}\delta V + \Delta V_s, \tag{15}$$

where ΔV_s is the stochastic error of the imaginary velocity measurement.

The velocity V^p of the heading point is unknown, therefore, for aiding the inertial reckoning, we propose to use only the first (lateral) and the third (vertical) components of the equations (14) and (15). This approach can be treated as an aiding by the zero components of the velocity vector V_s^p.

Note that the aiding proposed does not require any additional sensors. At the same time (as practice shows) the mentioned type of aiding provides a significant improvement in the performance of the navigation algorithm when GNSS data are not available. So in this case, one can obtain adequate estimates of the accelerometers biases in the complete absence of GNSS and odometer data. Certainly, one can use this approach for wheeled vehicles when there are no odometer data available.

A. Covariance analysis of observability

To study the observability of accelerometer biases, we created a digital simulator for trajectory motion parameters, tractor kinematics, and inertial sensors.

Three types of trajectories were simulated: movement along a straight line, movement along a circle with a constant radius, and movement along a periodic curve. The readings of inertial sensors were simulated for these trajectories, taking into account accelerometer biases and noise. Using the simulated inertial sensor data, the navigation algorithm described in Chapter IV was executed, where only the component-wise ZUPT was used in the update step of Kalman filter

Let us consider the covariance of the estimates of accelerometer biases for each of the three types of trajectories. See the relevant plots below.

Fig. 7. Covariances of accelerometer biases estimates on a straight line

Fig. 8. Covariances of accelerometer biases estimates on a circle

Fig. 9. Covariances of accelerometer biases estimates
on a periodic curve

One can see that the «horizontal» accelerometer biases are unobservable on a straight line and on a circle, but are observable on a periodic curve.

VI. CONCLUSION

We present two approaches for integrating INS and odometry data: loose and tight integration. We demonstrate that these two methods are equivalent in terms of their behaviour with respect to accelerometer biases and odometry coefficient estimates, regardless of the kinematic model used (classical or articulated).

We present the so-called component-wise ZUPT approach for INS aiding, which is based on kinematic models of the movement for ground wheeled vehicles. This approach does not require any additional navigation sensors.

Based on covariance analysis, we demonstrate the observability of accelerometer biases in periodic curve trajectories, and their unobservability in a straight line and a circle of constant radius.

REFERENCES

[1] Tormagov, T.A., Generalov, A.A., Shavin, M.Y. et al. Motion Control of Autonomous Wheeled Robots in Precision Agriculture. Gyroscopy Navig. 13, 23–35 (2022). DOI 10.17285/0869-7035.0083

[2] D. B. Pazychev and R. N. Sadekov, "MEMS Temperature Stabilization," 2020 27th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), St. Petersburg, Russia, 2020, pp. 1-4, doi: 10.23919/ICINS43215.2020.9133899

[3] D. B. Pazychev, K. S. Bakulev and R. N. Sadekov, "Low-Cost Navigation System for UAV," 2023 30th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), Saint Petersburg, Russian Federation, 2023, pp. 1-6, doi: 10.23919/ICINS51816.2023.10168469

[4] Polack, Philip & Altché, Florent & Novel, Brigitte & de La Fortelle, Arnaud. (2017). The kinematic bicycle model: A consistent model for planning feasible trajectories for autonomous vehicles? 812-818. 10.1109/IVS.2017.7995816.

[5] Min, Haigen & Wu, & Cheng, & Zhao. (2019). Kinematic and Dynamic Vehicle Model-Assisted Global Positioning Method for Autonomous Vehicles with Low-Cost GPS/Camera/In-Vehicle Sensors. Sensors. 19. 5430. 10.3390/s19245430.

[6] Nayl, Thaker. (2013). Modeling, control and path planning for an articulated vehicle.

[7] Delrobaei, Mehdi & McIsaac, Kenneth. (2011). Design and Steering Control of a Center-Articulated Mobile Robot Module. Journal of Robotics. 2011. 10.1155/2011/621879.

[8] Zhao, Jing-Shan & Liu, Zhi-Jing & Dai, Jian. (2013). Design of an Ackermann Type Steering Mechanism. Journal of Mechanical Engineering Science. 227. 10.1177/0954406213475980.

[9] Nikitin I.V. The navigation problem of a ground object based on strapdown INS and odometer data (PhD thesis). 2016

[10] Vavilova N.B., Golovan A.A., Parusnikov N.A. Mathematical foundations of inertial navigation systems. Moscow University Publishing House. 2020

An Accurate Method Applying Inertial Sensors for Axial Calibration of Spherical PTZ Camera

B.F Zhu

Jimei University

L.N Chen

University of Science and
Technology of China (USTC)

G.Y Shi

Peking University,
Chinese University of Hong Kong

Abstract—This paper proposed a simpler, direct and reliable axial calibration approach applying the high-precision inclination sensor, which is mounted on the rotating mechanical platform of the spherical PTZ Camera and collected data at the demanding action to calculate the coordinate rotation matrix. This method reduced the transfer relations between the different frames of axes, and solved problems in the indirect axial calibration, such as complex deployment, cumbersome calibration process, too many transfer relations and unreliable reference source. Additionally, we optimized the alignment algorithm based on the dual-vector and circle fitting algorithm, etc. To verify the feasibility of the algorithm, we implemented the simulation experiment whose results showed that the accuracy of this method can reach 0.02°, which meets the demand for high accuracy and stability in the application of infrared intelligent spherical PTZ Cameras. In general, this paper lays the foundation for the combination of optical and inertial devices in the visual security monitoring.

Keywords—Spherical PTZ Camera, Inertial Sensor; Phase-Variant Axial Calibration; Error Compensation; High Precision Visual Security Monitoring

I. INTRODUCTION

Spherical PTZ Camera is a camera device that integrates all-in-one zoom camera module, spherical rotating platform and embedded system. It can achieve panoramic monitoring and provide fast detail positioning function [1-3]. However, it was a tremendous hassle to achieve and keep consistency among sensors based on different principles. This rather contradictory result may be due to the strong vibration and shock existing in the actual working condition and the long-term operating conditions. The perception layer of spherical PTZ cameras, simply assembled with photoelectric encoder and decoder, and inclination sensors, was extremely prone to malfunctions. In addition, during the production of lens modules, it is difficult to keep the image sensor plane perpendicular to the mechanical axis used to polish the edges of the lens. To elucidate, there is an eccentricity tolerance between the optical axis and the mechanical axis, which leads to a large error in the end measurement results. Therefore, the mixing of inertial sensors and optical devices exerts a powerful effect upon high stability and precision of visual security monitoring through condition monitoring and fault diagnosis of rotating machinery. Most studies in the field of calibration with the two kind of equipment have only focused on method using external measurements, such as pictures from multiple locations, output from theodolite. Through the pre-set marking points on the rotating machinery, the angles were measured indirectly, and then the error could be calculated with the projection transformation relationship and equivalent algorithms [4-6]. Such approaches, however, have failed to address the following problems:

（1）   It makes no attempt to use a standardized and direct manner of calibration between mechanical and optical structures. Limited to the projection relationship, the requirements of external optical equipment were generally complicated. It brings about difficulties in promotion to engineering applications;

（2）   It fails to quantify the transfer relations in the calibration and equivalence calculation process. It is that, in each link, there are inevitable errors affecting the calibration accuracy of the whole system;

（3）   It does not offer an adequate explanation for the true value of the on-site calibration. Unfortunately, most of them were unilaterally, even completely, referenced to the results from the optical measurement. This induced that the accuracy and credibility of the calibration were still in doubt;

（4）   The weakness of existing research is to face the error superficially, rather than the underlying causes. Furthermore, it resulted that the current method has poor applicability and low efficiency in the calibration of rotating machinery [7-10].

As previously stated, this paper has attempted to provide a simpler, direct and reliable Standard Operating Procedure (SOP) to calibrate multiple orthogonal systems. Only via this SOP can the system autonomously measure the angle from the optical axis to the horizontal straight line under any position of the spherical PTZ Camera. From another perspective, we introduced a three-axis accelerometer screwed to the camera module and embedded double-vector attitude fixing algorithm and the long-term drift error calibration algorithm, thereby reducing the transmission relationship between the frame of axes and increasing the motion angle accuracy of the spherical PTZ Camera. This approach is best illustrated by the case of the Farmland protection PTZ Camera released by Hikvision in 2023. The machine calibrated by this method can reduce the monitoring deviation from a few hundred meters to within the meter level.

II. SYSTEM ARCHITECTURE DESIGN AND SELECTION

The high-precision inclinometer used in this study, also known as the digital biaxial inclinometer, has high cost-effectiveness, high long-term stability, small temperature drift, ease of use, and strong resistance to external interference. The measurement range is ±90°, with a maximum accuracy of 0.02°, and the working temperature is -40℃ -+85℃. The inclinometer integrates MEMS technology and digital output technology, and selects high-precision MEMS accelerometers and high-resolution differential digital-to-analog converters, with built-in automatic compensation and filtering algorithms. Thus, it is able to eliminate errors caused by physical field

changes to the greatest extent possible, and convert changes in the static gravity field into changes in inclination angle. Figure shows the device design diagram and bare board physical diagram, which can be screwed at any position of the rotating machinery.

We fixed the inertial inclinometer on the sheet metal of the camera module, with a slight inclination generally within 5~15 °. When the spherical PTZ camera rotated to a relative position, the data from the inclinometer are collected to calculate the pitch angle of the optical axis relative to the absolute horizontal plane. In addition, it is essential to align the measurement frame of axes of the sensor with the mechanical rotation axis, at the spatial and temporal levels, respectively through installation error calibration and zero offset calibration. To better integration of sensors with the overall system, as Figure 1 shows that the tilt sensor is modularized in the form of a bare board in the product development process, supporting specific installation structures and data interfaces.

Fig. 1. Schematic of sensor mounting and bare board.

III. Algorithm design

High precision algorithm design is urgently needed to ensure design safety. Various scenarios require spherical PTZ cameras to maintain high accuracy and stability, especially in environments with large temperature changes and strong shaking. The prime contributing factor is that the inertial measurement principles bring inclinometers strong anti-interference ability and autonomy, as well as advantages such as short preparation time and strong resistance to vibration and impact. Therefore, it is suitable for completing monitoring tasks with multiple points, parameters, and functions on the mechanical platform, thereby achieving precise control.

First and foremost, it is imperative to calibrate and compensate for the error angle between the inclination sensor measurement axis and the rotating machinery. Unfortunately, the influence of machining errors and installation errors leads to deviations in the axial measurement of rotating machinery and inclination sensors. The reason for this part is to accurately measure the axial deflection angle of rotating machinery, that is, the projection of the spatial angle of the measurement frame of axes relative to the reference frame of axes in the load system. In addition to the deterministic error calibration mentioned above, real-time compensation is also required for random drift caused by uncertain factors. This should be drawn attention for inertial measurements, as their errors accumulate over time, making it difficult to maintain high-precision positioning and orientation.

Before designing the algorithm, frame of axes should be established for devices with different principles and a unified reference frame of axes should be found. From a theoretical perspective, gravitational acceleration and Earth's rotational velocity are natural physical references for inertial sensors, which can be measured and determined through accelerometers and gyroscopes, respectively. As shown in Figure 2, during the

operation of the entire machine, there are geodetic frame of axes (1), mechanical platform fixed frame of axes (2), mechanical platform rotation frame of axes (3), sensor measurement frame of axes (4), and optical axis measurement frame of axes (5). The commonality is that in the optical axis measurement coordinate system, the rays directed towards the front of the lens can also be projected onto the geodetic coordinate system, thereby calculating the angle between the rays along the direction of gravity in two positions.

Fig. 2. Schematic diagram of the frame of axes involved in the operation of the machine, geodetic frame of axes (1), mechanical platform fixed frame of axes (2), mechanical platform rotation frame of axes (3), sensor measurement frame of axes (4), optical axis measurement frame of axes (5).

A. Dual Vector Positioning Algorithm to Calibrate misalignment error

Rotating machinery can rotate around both horizontal and vertical axes, and the following provisions need to be made before calibration can begin: the angle of rotation around the horizontal axis is stated as the pitch angle. The angle of rotation about the vertical axis is referred to as the heading angle; this is shown by the green (horizontal) and red (vertical) lines in Figure 3.

Fig. 3. Schematic diagram of the pitch heading angle of the dual-vector attitude fixing algorithm.

During the calibration process, central system collects the three-axis outputs of the sensor and the data are averaged over a stationary period of time.

Based on the above provisions, the ectopic orthogonal axis calibration method applicable to rotating machinery is developed according to the following steps:

(1) Control the rotating machinery to be located at a fixed heading angle and keep the heading angle unchanged, set the pitch angle to the usual initial position of the equipment movement (given by the control mechanism), and record the three-axis sensor outputs;

(2) Maintaining the heading angle the same as in (1), controlling the pitch angle within the movable range, measuring multiple positions, and recording the three-axis sensor output for each position;

(3) Keeping the heading angle the same as in (1), set the pitch angle to a common termination position for device motion and record the triaxial sensor output;

(4) Set the pitch angle to a position in the middle of the range of motion, keep it fixed, control the heading angle within the movable range, measure multiple positions, and record the three-axis sensor output for each position.

Through the above prescribed actions, the sensors screwed on the rotating machinery collect the corresponding data, and the calibration matrix can be calculated. This type of problem is essentially to solve the orientation relationship between two spatial frame of axes, generally assuming that there are two non-collinear reference vectors, the two rotary axes of the equipment as a reference vector, according to its projected coordinates under the two frame of axes to find the transformation relationship, which can be presented as a directional cosine matrix, noted $\mathbf{C}_b^n$, assuming that the known rotary axes in the geodetic frame of axes under the frame of axes for the coordinate:

$$\mathbf{V}_1^n = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$$
$$\mathbf{V}_2^n = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \tag{1}$$

Correspondingly, based on the data from the sensors fixed on the rotating machinery, the coordinates of the rotating axis under the carrier system are obtained as:

$$V_1^b = Mat\_Vt$$
$$V_2^b = Mat\_Hz \tag{2}$$

Since direction cosine matrices are generally three-dimensional square matrices, it requires to construct another vector equation, i.e., to cross-multiply the two known vectors and then to unite the three vector equations:

$$\begin{bmatrix} V_1^n & V_2^n & V_1^n \times V_1^n \end{bmatrix} = C_b^n \begin{bmatrix} V_1^b & V_2^b & V_1^b \times V_1^b \end{bmatrix} \tag{3}$$

The algorithm for solving the two-vector fixed-posture has been obtained so far, which results in

$$C_b^n = \begin{bmatrix} \left(V_1^n\right)^T \\ \left(V_2^n\right)^T \\ \left(V_1^n \times V_2^n\right)^T \end{bmatrix}^{-1} \begin{bmatrix} \left(V_1^b\right)^T \\ \left(V_2^b\right)^T \\ \left(V_1^b \times V_2^b\right)^T \end{bmatrix} \tag{4}$$

After figuring out the direction cosine matrix for the coordinate transformation, the coordinate measured on one system are transformed to another frame of axes, e.g., the values measured in the coordinate system A can be expressed in the following equation：

$$f^a = K^{-1}N_A - \nabla^a \tag{5}$$

Obtaining that the direction cosine matrix from the a-system to the coordinate system B is C_a^b, then:

$$f^a = K^{-1}N_A - \nabla^a$$
$$f^b = \left(C_b^a\right)^T f^a$$
$$\nabla^b = \left(C_b^a\right)^T \nabla^a \tag{6}$$
$$K_A = \left[K\left(C_a^b\right)^T \right]^{-1}$$

Due to the fact that, yet, some of the vectors originated from the motion of the rotating mechanical platform, there is a certain measurement error, and also a certain amount of vector error, including modulus error and direction error, which makes the calculation results fail to strictly meet the requirements of unit orthogonalization. In this regard, the algorithm adopted data preprocessing and specific position correction. Before the execution of the SOP, all the vectors that need to be solved are processed for orthogonalization and unitization, and corrected at the limited points. Ultimately, the rotation matrix is calculated and corrected by the approaches stated above.

Fig. 4. Schematic diagram of the computational flow of the algorithm.

Fig. 5. Schematic diagram of the computational flow of the algorithm.

B. Circle Fitting by Nonlinear Least Squares to Compensate for Random Errors

From the aspect of engineering realization, according to the motion characteristics of machinery, fully autonomous random error calibration function is essential. It is that Spherical PTZ cameras are mostly installed in remote areas, which makes manual installation and maintenance difficult and costly. By collecting data from machinery at a constant speed (1 °/s), the system calculates the center and radius of the rotating circular surface based on algorithms; for convenience, the requirements of SOP are consistent with the axial calibration in A. After outputting the calibration matrix in this way, it can be saved to the on-chip system.

In addition, in algorithm implementation, it is crucial to have a mathematical foundation for removing outliers and fitting the overall curve. Referring to Practical Methods of Optimization, the most suitable scenario for the Gaussian Newton method is when the sum of squares is small when the minimum value is taken, or when the nonlinearity of the data

is mild. At the same time, to avoid the problem of outlier sensitivity in nonlinear least squares using the Gaussian Newton minimization algorithm, the fitting problem can be simplified towards the direction of linear test squared.

To summarize briefly, the core concept of the algorithm, which describes geometric relationships, is that "the collected data should be distributed on multiple concentric unit circles.". The advantage of the algorithm is that it avoids the sensitivity of conventional fitting methods to outliers and identifies the point at which the minimum value is taken, which is the calibration reference position. In order to prevent the different testing processes of the centers of the two rotating circles of the ball machine, two inclinometers were used and multiple sets of paths were arranged. This algorithm used as many movement routes as possible to get closer to the movement trajectory and mechanical zero position.

Fig. 6. Data acquisition process and upper computer software interface diagram.

In the light of the aforesaid, through the A and B mentioned above, the calibration and compensation of two types of errors, which are prone to occur when inertial sensors are applied to spherical mechanical and optical equipment, have been achieved from both spatial and temporal perspectives, as the

IV. SIMULATION VERIFICATION

At this point in writing, the existing calibration solutions still have issues with unreliable and unstable reference values. Emerging as the times require, a high-precision three-axis rate rotary table is used as the reference source in the actual verification process, and this kind of rotary table is the most ideal equipment for large-scale and multi-functional inertia test, and the three-axis rate rotary table contains three frames, which are the outer, middle and inner frames respectively, and the sensors will be mounted and fixed on the inner frame, and the measured objects can be moved in any direction of angular velocity in space due to the three frames constituting the universal bracket. The vertical three-axis rotary table has an outer frame for azimuth, a middle frame for pitch, and an inner frame for roll. The rotary table is used as the only reference source, and by installing multiple sensors of the same type on the rotary table, calibration is carried out to eliminate device zero bias, scale factor and cross-axis errors, thus ensuring the accuracy of the calibration results.

Fig. 7.   Physical drawing of simulation verification rotary table.

A. *Dual vector attitude verification*

In the process of demonstrating the technical effect, the rotary table is used as a rotating mechanical platform to be monitored, and multiple sensors are installed in different positions on the work place as shown in the figure. The rotary table control software enables the rotary table to perform the corresponding actions in accordance with the prescribed calibration process and cooperate with the upper computer for data acquisition and decalculation, so as to complete the calibration task.

Fig. 8.   Simulation verification of calibration results.

The calibration results are shown in Fig.8 and Table 1, the collection and processing software calculated the calibration matrix. After calibration, the central system collected the data of the sensor mounted on the reference platform again, and compared the angular reference value. The results show that, in the rotary table for the position of 0°, 15°, 30°, 45°, 60°, 75° and 90°, the D-value between the output of the sensor and the reference source value are all within 0.03°, which is in line with the requirement of monitoring accuracy of industrial equipment.

TABLE I.   AXIAL CALIBRATION VERIFICATION EXPERIMENT RESULTS

Group	Control angle (°)	Sensor Output (°)	Reference source output (°)	Error (°)
1	0	0.0000	0.0000	0.0000
2	15	15.0100	14.9999	0.0101
3	30	30.0070	30.0000	0.0007
4	45	44.9880	45.0000	0.0120
5	60	59.9730	59.9999	0.0269
6	75	74.9750	75.0000	0.0250
7	90	89.9900	90.0000	0.0100

The reason for adopting this verification method is because turntable and common rotating machinery have common characteristics, are around a certain axis in space for high and low angle action, so after the turntable calibration and calibration of the sensor, in the actual use of the process, can be installed through Fig. 9 mechanical relationship, the rotating machinery as a carrier, the equivalent replacement for the turntable, through the transfer of the relationship to ensure that the installation of sensors on the rotating machinery can output the same accurate measurement data.

Fig. 9.   Accuracy Traceability Flowchart.

As the external excitation brought by the rotating machinery to the sensor is unknown and its true value cannot be measured, this paper adopts the method of rotary table accuracy verification and consistency verification of actual occasions. The results maximize the assurance of the stability and accuracy of the sensor in the field of measurement.

B. *Circle Fitting Verification*

The four sets of test results for the two sensors are shown in the table, from which it can be seen that the basic distribution of the collected data points is on a concentric multiple unit circle.

TABLE II.   FOUR SETS OF TEST RESULTS FOR SENSORS

Experiment number	Coordinates of the center of the circle x	Coordinates of the center of the circle y	Radius length
1-1	-0.0025	0.0245	1.0034
1-2	-0.0016	0.0704	1.0990
1-3	-0.0032	0.0012	1.0096
1-4	-0.0032	-0.0016	1.0130
1-5	-0.0019	0.0346	0.9998
1-6	-0.0014	0.0782	1.0492
1-7	0.0020	0.0111	1.0040
1-8	-0.0019	0.0107	1.0044

Optical control bit value	Pitch reference angle (°)	Sensor 1 (°)	Sensor 2 (°)	Accuracy (°)
0	0	0.4770	0.4610	0.0160
150	15	15.5110	15.5250	0.0140
300	30	30.6320	30.6550	0.0230
450	45	45.7630	45.7970	0.0340
600	60	60.9060	60.9340	0.0280
750	75	76.1090	76.1510	0.0420
800	80	81.1660	81.2000	0.0340
900	90	88.7840	88.7630	0.0210

To elucidate, Circle Fitting eliminated long-term random error and avoided the sensitivity to outliers in conventional fitting methods. It identified the point at which the minimum value. In the verification process, we used two sensors choreographed with multiple sets of paths, which can maximize the number of movement routes, so that the result could more closely approximate the movement trajectory of the movement and the mechanical zero position. This process could prevent the spherical PTZ Camera from having a different center of gravity for the two rotating circles.

C. Verification of simulation results

In order to avoid the mutual influence of error correction caused by redundant operations and affect the accuracy of the final result, by executing A and B in different orders on the embedded system, it is concluded that the process of calibrating the random bias and then the mounting error is adopted to verify the simulation results; As an insurance scheme, the total station is used as the verification equipment. The points are marked according to the pitch reference angle. By controlling the mechanical photoelectric coding disc position and aligning the reference datum, the sensor values are recorded respectively. Through the actual test results, it is concluded that the actual accuracy is between 0.02-0.04° after the zero bias calibration and installation error calibration are executed successively, the accuracy of the large angle (60°-90°) is decreased, and the accuracy of the main monitoring angle (15°-60°) is between 0.02-0.03, which is in line with the nominal requirements.

V.   CONCLUSION

In this paper, based on the inertial sensor, a simple, direct and reliable phase-variant axial calibration method is proposed to complete the phase-variant axial calibration task through the dual-vector attitude algorithm, circle fitting algorithm and so on. It has definitely improved the accuracy of the spherical PTZ Camera, and the accuracy of the sensor's monitoring range as high as 0.02°, by the simulation and modelling experiments, which satisfied the high accuracy for the spherical PTZ camera.

REFERENCES

[1] Li Y, Pan L, Cheng T. A Camera PTZ Control Algorithm for Autonomous Mobile Inspection Robot[C]//International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering.IEEE, 2021.

[2] Mao K, Xu Y, Wang R, et al. A General Calibration Method for Dual-PTZ Cameras Based on Feedback Parameters[J]. Applied Sciences, 2022, 12(18): 9148.

[3] Yu Y. Technology Development and Application of IR Camera: Current Status and Challenges[J]. Infrared and Millimeter Wave, 2023, 1(1): 1-7.

[4] Zhang J, Zhang Q, Jiang T, et al. Vibration displacement measurement method based on vision Gaussian fitting and edge optimization for rotating shafts[J]. Measurement, 2024: 114699.

[5] Tiboni M, Remino C, Bussola R, et al. A review on vibration-based condition monitoring of rotating machinery[J]. Applied Sciences, 2022, 12(3): 972.

[6] Brito L C, Susto G A, Brito J N, et al. An explainable artificial intelligence approach for unsupervised fault detection and diagnosis in rotating machinery[J]. Mechanical Systems and Signal Processing, 2022, 163: 108105.

[7] Tiboni M, Remino C, Bussola R, et al. A review on vibration-based condition monitoring of rotating machinery[J]. Applied Sciences, 2022, 12(3): 972.

[8] Ma Y, Wen G, Cheng S, et al. Multimodal convolutional neural network model with information fusion for intelligent fault diagnosis in rotating machinery[J]. Measurement Science and Technology, 2022, 33(12): 125109.

[9] Kateris D, Moshou D, Pantazi X E, et al. A machine learning approach for the condition monitoring of rotating machinery[J]. Journal of Mechanical Science and Technology, 2014, 28: 61-71.

[10] Heng A, Zhang S, Tan A C, et al. Rotating machinery prognostics: State of the art, challenges and opportunities[J]. Mechanical systems and signal processing, 2009, 23(3): 724-739.

Navigation Systems of Unmanned Aerial Vehicles with Artificial Intelligence Units

Konstantin Veremeenko
Flight-Navigation and Information-Measuring Systems Department
Moscow Aviation Institute (National research university)
Russia
nio3@mai.ru

Roman Zimin
Flight-Navigation and Information-Measuring Systems Department
Moscow Aviation Institute (National research university)
Russia
ziminry@mai.ru

Andrei Pronkin
Flight-Navigation and Information-Measuring Systems Department
Moscow Aviation Institute (National research university)
Russia
pronkinan@mai.ru

Maksim Zharkov
Flight-Navigation and Information-Measuring Systems Department
Moscow Aviation Institute (National research university)
Russia
zharkovmv@mai.ru

Ivan Kuznetsov
Flight-Navigation and Information-Measuring Systems Department
Moscow Aviation Institute (National research university)
Russia
kuznetsovim@mai.ru

Abstract—The high interest in improving navigation accuracy and UAV autonomy, which has increased many times in recent years, forces developers to look for new resources to improve the characteristics of the UAV. One of the relatively new and potentially promising approaches to the UAV navigation system (NS) algorithms design can consider the use of artificial intelligence techniques. The article describes the architecture of the UAV NS algorithmic support with artificial intelligence modules in its structure (AI UAV NS). Artificial neural networks (ANN) can be used with varying degrees of integration into the system structure. The most obvious advantages of ANN, such as the mathematical models adaptability, the lack of necessity to use a priori data on its parameters, the absence of requirements for linearization allow them to be used for a wide range of tasks – from approximating error models of strapdown inertial navigation systems (SINS), designing its algorithms, to implementing an alternative to an integrated solution using Kalman filter (KF) with various degrees of integration into a system – from complete replacement of the KF, estimation and compensation of errors in the KF states during the outages of signals from the main corrector to predicting measurements received at the KF input. Nevertheless, the presented advantages of ANN are unrealizable without properly conducted preliminary training, and the choice and design of the NS architecture, as well as the selection of training samples and flight paths of UAS, is a separate scientific task requiring detailed study. As part of the implemented experimental research methodology, preliminary results were obtained reflecting the accuracy characteristics of the navigation complex with artificial intelligence modules structurally included in its composition.

Keywords—Navigation system, artificial intelligence, machine learning, EKF, complex information processing

I. INTRODUCTION

The integration techniques for traditional navigation systems of unmanned aerial vehicles (UAVs) used today are usually based on the use of the Kalman filter, which estimates the state vector, including errors in the navigation system subsystems [1, 2]. In this case, both loosely and tightly coupled integration techniques are used [3, 4]. The functionality of the UAV navigation system (NS), depending on the completeness of the on–board composition of its measuring part, may, among other things, allow to ensure the requirements of regulatory documents for the possibility of integrating UAVs into the common airspace, and the use of special algorithms – to be able to detect organized interference, both spoofing and jamming [5]. It should be specifically noted that such requirements appear, among other things, in the documents of the federal project on unmanned aircraft systems (UAS FP) [6].

To date, modern and promising navigation and attitude systems for mobile objects are, generally, an integrated NS, the algorithmic support of which carries out complex processing of measuring information received from on-board sensors and systems. The core of such an NS is usually an inertial navigation system (INS), due to a number of advantages:

- high information content (output of a complete list of motion parameters: coordinates, velocities, attitude parameters),
- autonomy,
- high frequency of data output.

However, a significant disadvantage for a number of applications is the unlimited growth of errors over time. All other sensors and systems from the integrated NS are

commonly referred to as INS correctors. Depending on the type of correctors, the following types of correction can be distinguished:

- positional (using coordinate information),

- using velocity information (an information about linear velocity projections),

- angular type (using information about the parameters of the UAV angular attitude). The best error estimation performance of the INS and its sensors (angular rate sensor (ARS) and accelerometers) can be obtained when all types of correction are implemented simultaneously [4].

II. PROPOSED APPROACH

This part of the paper discusses NS basic equations and parameters to be estimated, the state vector, the UAV NS structure and the architecture of the neural network used in the UAV NS structure.

A. NS basic equations

The object of research in the proposed work was an integrated inertial-satellite navigation system, which contains a tactical accuracy class inertial measurement unit (IMU) with a GNSS receiver operating in different modes. In this case, the positional-velocity and at the same time angular type of correction of the INS is considered.

The equation of the system state is presented as follows:

$$\dot{\bar{X}} = F \cdot \bar{X} + G \cdot \bar{W} , \qquad (1)$$

where F is the dynamics matrix, $\bar{X}$ is the state vector, G is the system noise matrix, $\bar{W}$ is the system noise vector.

The system state is estimated using $\bar{Z}$ measurements vector.

The equation of the system state is presented as follows:

$$\bar{Z} = H \cdot \bar{X} + \bar{V} , \qquad (2)$$

where H is the measurement matrix, $\bar{V}$ is the measurements noise vector.

It is proposed to include the following parameters in the composition of the state vector [7]:

$$\bar{X} = \begin{bmatrix} X_I & \Delta \end{bmatrix}^T , \qquad (3)$$

where $X_I = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & \alpha & \beta & \gamma \end{bmatrix}$ are the INS errors in coordinates, velocities projections and attitude angles definitions; $\Delta = \begin{bmatrix} \Delta\Omega_1 & \Delta\Omega_2 & \Delta\Omega_3 & \Delta n_1 & \Delta n_2 & \Delta n_3 \end{bmatrix}$ are the constant components of the ARS and accelerometers errors (when expanding the model of sensor errors, the state vector can also be expanded by the corresponding components).

The dynamics matrix F is formed based on the error model of the INS and its sensors. The system noise vector includes random components of the errors of the INS sensors, and the measurements noise vector includes random components of the correctors errors.

The measurement vector is full-format and has the following form:

$$\bar{Z} = \begin{bmatrix} \delta\Phi & \delta U & \delta\Psi \end{bmatrix}^T , \qquad (4)$$

where $\delta\Phi = \begin{bmatrix} \delta\varphi^I - \delta\varphi^C & \delta\lambda^I - \delta\lambda^C & \delta h^I - \delta h^C \end{bmatrix}$ is the difference between the INS and corrector coordinates outputs, $\delta U = \begin{bmatrix} \delta U_E^I - \delta U_E^C & \delta U_N^I - \delta U_N^C & \delta U_Z^I - \delta U_Z^C \end{bmatrix}$ is the difference between the INS and corrector velocities outputs, $\delta\Psi = \begin{bmatrix} \delta\psi^I - \delta\psi^C & \delta\vartheta^I - \delta\vartheta^C & \delta\gamma^I - \delta\gamma^C \end{bmatrix}$ is the difference between the INS and corrector yaw, pitch and roll angles outputs respectively.

To estimate the state vector, a variant of the discrete Kalman filter is used, formulated in the so-called Joseph form [7], which provides increased protection against computational divergence.

B. UAV NS structure

As already mentioned, the paper considers an integrated navigation system consisting of an inertial navigation system and several GNSS receivers. The INS is the core of such an integrated system, and GNSS receivers are correctors used to estimate and compensate for the INS errors.

In the structure of the general INS algorithm, algorithms for determining navigation parameters (navigation algorithm) and attitude parameters (attitude algorithm) can be distinguished. Within the work, various applications of neural networks in the task of approximating nonlinear equations of the INS navigation coordinates were investigated. Figures 1, 2 show two schemes for applying and modeling the work of a neural network designed to approximate the algorithm for INS coordinates determination.

The first of the schemes implements an algorithm for the INS ideal operation, the second is a functioning algorithm that takes into account the signals that compensate for the impact of the errors of the INS on its output navigation parameters. In the first case (Fig. 1), the neural network approximates the transition matrix corresponding to the equations of inertial navigation. The functioning of the neural network in the configuration mode is as follows: vectors consisting of navigation coordinates P_k, projections of linear velocities V_k, attitude parameters A_k and accelerometers n_k and ARSs ω_k measurements are received at the input of the neural network, without taking into account measurement errors (ideal values of projections of specific force and absolute angular velocity) at time t. The output of the neural network (navigation parameters) is compared with the precise solution of the INS operation equations for the time $t + \Delta t$. The neural network learning algorithm adjusts its parameters in such a way as to minimize the residual ΔP_k, ΔV_k, ΔA_k between the output of the neural network and the precise solution formed by the possible values of navigation coordinates and measured values when modeling the equations of their ideal equation. In the main mode of operation, navigation parameters and measurement signals are received at the input of the neural network, and navigation and attitude parameters for the next moment are the output. Thus, the neural network approximates the algorithm of the INS ideal operation.

Fig. 1. Block diagram of the INS ideal operation algorithm approximation by the ANN

Fig. 2. Block diagram of the INS ideal operation algorithm approximation by the ANN, taking into account the errors

In the second case (Fig. 2), when training a neural network, vectors are fed to its input, formed not from ideal parameter values, but from measured ones $P_k + \Delta P_k$, $V_k + \Delta V_k$, $A_k + \Delta A_k$, $n_k + \Delta n_k$, $\omega_k + \Delta\omega_k$, i.e. including measurement errors. The output of the neural network adjusts to the ideal values of navigation and attitude parameters. In this case, the parameters of the neural network will take into account the errors of the measurers. This scheme can be used in the integrated NS to predict navigation parameters during periods of GNSS signals loss. At the same time, the neural network learning process can take place continuously while the GNSS generates accurate coordinates and velocities. At the same time, the decomposed versions of the neural network can individually solve attitude and navigation tasks (the functional algorithm block can be both an attitude algorithm and a navigation algorithm).

Other schemes are also possible in which the neural network approximates not the algorithm for calculating the INS coordinates, but the nonlinear equations of the errors of the INS, i.e. the outputs of the neural network are coordinate determination errors [8-11]. In this case, manufacturing errors of the INS elements can be interpreted as configurable parameters of the neural network: the magnitude of synaptic connections or its shifts.

C. Neural network architecture

The neural network structurally included in the integrated NS under consideration was designed using LSTM architecture (Figure 3) and was trained using a high-quality dataset. The target data was obtained using the output of the software package, which is described in the next section.

Fig. 3. LSTM ANN cell architecture, here X_t is the input time step, h_t is the output, C_t is the cell state, f_t is the forget gate, i_t is the input gate, o_t is the output gate, $\hat{C}_t$ is the internal cell state. Operations inside dark grey circle are pointwise.

III. EXPERIMENTAL RESULTS

Due to the long training time of the considered architecture of the neural network, it was decided to decompose training samples. As such, coordinates, projections of linear velocities and orientation angles were used, coming from an integrated navigation system outputs, the algorithms of which are implemented in the software, and its functionality is given below.

The raw measurements of the inertial measurement unit (IMU performance are presented in Table 1) and three Novatel OEMV GNSS receivers recorded onboard a driving car were processed together with the GNSS base station (BS) measurements (Novatel OEM 729) in the post-processing software, which is the part of the MAI Navigation System Car Testbed [12, 13].

TABLE I.   LITEF UIMU LCI PERFORMANCE

Gyroscope Performance	
Input range	±800 deg/sec
Rate bias	< 1.0 deg/hr
Rate scale factor	100 ppm (typical)
Angular random walk	<0.05 deg/√hr
Accelerometer Performance	
Range	±40 g
Scale factor	250 ppm (typical)
Bias	<1.0 mg

The following algorithms are implemented in the software:

- autonomous strapdown INS;

- position (single, differential and RTK mode), velocity and attitude angles calculated in the processing of multiple GNSS onboard receivers (OR) and BS pseudorange, doppler and carrier phase raw measurements;

- loosely and tightly coupled integration of IMU, multiple GNSS OR and BS using the extended Kalman filter for position, velocity, attitude angles, IMU errors, and GNSS carrier phase ambiguity estimates.

To create an additional independent source of navigation information, the Testbed also includes NovAtel Inertial Explorer (Canada) post-processing software [14], widely used in the field of geodesy and navigation. NovAtel Inertial Explorer has similar functionality to the one described above by MAI, but operates on the basis of unpublished algorithms from Novatel.

Fig. 4.   The trajectory of the test car drive

The results obtained in post-processing mode of a pre-trained neural network using a high-quality dataset are shown in Figures 5, 6.

In which $\Delta\lambda$, $\Delta\phi$, Δh are the errors in determining longitude, latitude and altitude, expressed in meters, respectively, ΔV_x , ΔV_y , ΔV_z are the errors in determining projections of linear velocities on the corresponding axes, expressed in meters per second, $\Delta\psi$, $\Delta\theta$, $\Delta\gamma$ are the errors in determining the attitude angles – yaw, pitch and roll, respectively, expressed in degrees.

Fig. 5.   NN performance (2D coordinates)

Fig. 6.   NN performance compared to INS standalone

As can be seen from the figures, in the mode of approximation by the neural network of the integrated navigation system output at intervals of GNSS signals loss within 200 seconds (for coordinates and projections of linear velocities) and within 400 seconds (for attitude angles), it is possible to get the accuracy levels acceptable for solving the UAV navigation and attitude task – with maximum error magnitudes up to 1 m in coordinates, up to 0.25 m/s according to linear velocity projections and up to 0.15 degrees at orientation angles, with the values of mathematical expectation – 1.3832 m, 0.8724 m, 0.7205 m, 0.008 m/s, 0.001 m/s, 0.075 m/s, 0.0035 deg., 0.0047 deg., 0.0041 deg., and with rms values – 0.4971 m, 0.6641 m, 0.6067 m, 0.0373 m/s, 0.0794 m/s, 0.0432 m/s, 0.0521 deg., 0.0191 deg., 0.0199 deg. in longitude, latitude and altitude, corresponding to projections of linear velocities and angles of course, pitch and roll, respectively.

IV. CONCLUSION

The results of the study demonstrate the possibility of using the proposed approach to solve the problems of the UAV navigation and attitude determination, while operating an integrated NS with an artificial intelligence module in its structure under conditions of a sufficiently long loss of GNSS signals. In the future, it is planned to investigate the effect of the trajectory influence on the operability and achievable levels of accuracy of the neural network, structurally a part of the UAV integrated NS.

REFERENCES

[1]   O. Stepanov, A. Motorin «Models and methods in estimation tasks in the processing of navigation information» // Materials of the XI conference in memory of the outstanding designer of gyroscopic devices N. N. Ostryakov, St. Petersburg, pp. 285-288, October 2020

[2] K. Veremeenko, K. Sypalo, D. Kozorez «Modeling of the processes of functioning of integrated aircraft systems based on the methodology of object-oriented analysis and design» // Moscow, p. 93, 2010.

[3] Noureldin A., Karamat T., Georgy J. «Fundamentals of Inertial Navigation, Satellite-based Positioning and their Integration» // Springer Heidelberg New York Dordrecht London, p. 324, 2013.

[4] M. Zharkov, K. Veremeenko, D. Antonov, I. Kuznetsov «Attitude determination using ambiguous GNSS phase measurements and absolute angular rate measurements» // Gyroscopy and navigation, Vol. 26, № 3(102), pp. 54-68, 2018.

[5] M. Zharkov, K. Veremeenko, I. Kuznetsov, A. Pronkin, «Global Navigation Satellite System Spoofing Detection in Inertial Satellite Navigation Systems» // Inventions p.8, 2023. https://doi.org/10.3390/inventions8060158

[6] Decree of the Government of the Russian Federation dated 06/21/2023 No. 1630-r «On approval of the strategy for the development of unmanned aircraft of the Russian Federation for the period up to 2030 and for the future up to 2035»

[7] Veremeenko, K.K., Zharkov, M.V., Kuznetsov, I.M. et al. «Strapdown Inertial Navigation System Transfer Alignment: Algorithmic Features and Simulation Performance Analysis» // Russ. Aeronaut. 63, 618–626 (2020).

[8] W. Fang, J. Jiang, S. Lu, Y. Gong, Y. Tao, Y. Tang, P. Yan, H. Luo, J. Liu «A LSTM Algorithm Estimating Pseudo Measurements for Aiding INS during GNSS Signal Outages» // MDPI Remote Sens, Switzerland, p.24, 2020.

[9] R. Sharaf, A. Noureldin «Online INS/GNSS integration with a radial basis function neural network» // IEEE Aerosp. Electron. Syst. Mag., pp.8–14, 20, 2005.

[10] R. Sharaf, A. Noureldin, «Sensor integration for satellite-based vehicular navigation using neural networks» // IEEE Trans. Neural Netw., pp. 589–594, 18, 2007.

[11] M. Jaradat, M.Abdel-Hafez «Non-linear autoregressive delay-dependent INS/GNSS navigation system using neural networks» // IEEE Sens. J,. pp. 1105–1115, 17, 2017.

[12] M. Zharkov, K. Veremeenko, I. Kuznetsov, A. Pronkin «Experimental Results of Attitude Determination Functional Algorithms Implementation in Strapdown Inertial Navigation System» // Sensors 2022, 22, 1849

[13] MAI Navigation and Attitude Determination Systems Research and Development Laboratory. Available online: https://mai.ru/science/studies/equipment/index.php?ELEMENT_ID=167952 (accessed on 20 April 2024)

[14] NovAtel Inertial Explorer. Available online: https://novatel.com/products/waypoint-post-processing-software/inertialexplorer (accessed on 20 April 2024)

Receiving Astronomical Measurements in a Strapdown Astroinertial Navigation System for Atmospheric Applications

N.N. Vasilyuk

SPC Elektrooptika, LLC
Moscow, Russia
ORCID: 0000-0003-2317-8066

Abstract — **The astronomical measurement refers to the attitude of the optoelectronic astrovision unit, obtained as a result of processing digital images of observed stars. By itself, the astronomical measurement is useless for the navigation relative to the surface of Earth. However, with the aid of additional information, useful navigation parameters may be extracted from the astronomical measurement. Astronomical measurement is distorted by various physical effects associated with distant stars observation through the atmosphere of moving Earth. These effects are taken into account when obtaining astronomical measurement by using additional navigation information from the strapdown inertial measurement unit. This additional information is also used to extract the useful navigation parameters from the astronomical measurement.**

Keywords — star tracker, astronomical measurements, Wahba problem, astroinertial navigation system, motion blur correction, refraction, aberration, distortion

I. INTRODUCTION

A strapdown astro-inertial navigation system (AINS) consists of an astrovision unit (AVU) and an inertial measurement unit (IMU) (Fig. 1.). AINS also contains a computer equipped with a time keeper. The time keeper contains a stable fixed frequency master oscillator. The initial time value of this time keeper is set synchronously with the UTC(SU) time scale at the moment when the AINS turns on.

Fig.1.   The layout of the strapdown AINS. 1,2,3 – AVU cameras; 4 – AVU base; 5 – IMU.

The AVU is used to obtain digital images of stars and consists of several digital cameras with sequential numbers $i = 1\ldots I$, where I is the number of cameras. The cameras are rigidly attached on a common base so that their fields of view do not intersect. The exposure starting instants of individual frames in all cameras are synchronized with each other. Each camera i is associated with a right orthogonal coordinate frame CF^i (Camera Frame). The directions to the observed stars are determined relative to this frame. The IMU consists of three gyroscopes and three accelerometers and measures vectors of an angular rate and a proper acceleration of the AINS. Projections of the angular rate and proper acceleration vectors are measured relative to the right orthogonal coordinate frame MF (Measurement Frame) associated with the IMU. The digital values of the time instants of receiving synchronous frames in the AVU and the time instants of the IMU measurements are determined from the readings of the time keeper.

All AINS components are assembled into a single mechanical structure that has no moving parts. This layout of the AINS keeps the relative attitude of the cameras within the AVU and the attitude of the AVU relative to the IMU unchanged. The constancy of the relative attitude of the cameras allows to consider the AVU as a kind of "virtual camera" with a complex virtual field of view. This field of view is constructed by computationally combining the observation results obtained by individual physical cameras. The right orthogonal coordinate frame VF (Virtual Frame) is associated with the virtual camera. The attitude parameters of the CF^i relative to the VF are calibrated with the error of $\sigma \approx 2''$ and specified as orthogonal matrices $\{\mathbf{C}_{VF}^{CF^i}\}_{i=1}^{I}$.

The constancy of the attitude of the AVU relative to the IMU allows to use the astronomical observations of the "virtual camera" for correcting of increasing errors of an inertial navigation. The attitude parameters of VF relative to MF are calibrated with the error of $\sigma \approx 6''$ and specified as the orthogonal matrix $\mathbf{C}_{MF}^{VF}$.

In the AINS, the astronomical measurement refers to the values of the parameters of the instantaneous attitude of the AVU relative to the inertial coordinate system associated with the stars. The attitude of the AVU is determined from images of stars obtained by individual AVU cameras. The digital value of the moment in time to which the astronomical measurement refers is called the measurement epoch. The measurement epoch is determined by the readings of the time keeper and is designated as t_n, where $n = 1\ldots$ - is the sequential number of the epoch.

When they talk about the position and speed of the AINS, they mean the position and speed of the origin of the MF relative to the Earth's surface. The AINS attitude refers to the attitude of the MF relative to the local ENU (East-North-Up) coordinate frame. This attitude is represented in terms of heading ψ_n^{MF}, pitch ϑ_n^{MF} and roll γ_n^{MF} angles of the MF per epoch t_n. Since the MF and the VF are connected through the matrix $\mathbf{C}_{MF}^{VF}$, the MF attitude angles are one-to-one converted to the VF attitude angles ψ_n^{VF}, ϑ_n^{VF}, γ_n^{VF}.

To obtain one astronomical measurement, frames with stars images are subjected to computational processing, consisting of several successive stages. At the first stage, the consequent frames are accumulated with a motion blur correction. Frames are accumulated separately for each physical camera of the AVU. From the accumulated frame, the direction vectors of the observed stars are calculated.

At the second stage, various types of refraction that arise when observing stars through a gaseous medium are corrected in the direction vectors. After that, the detected stars are recognized, i.e. establishing a correspondence between a digital image of a star and an entry in a star catalogue.

At the third stage, the speed aberration from the orbital motion of Earth in the direction vectors of the recognized stars, calculated from the star catalog, is taken into account. As a result, we obtain two vector bundles of the same dimension, for which we can set the Wahba problem [1]. The solution of the Wahba problem gives one astronomical measurement containing the attitude parameters of the AVU.

II. OBTAINING THE ASTRONOMICAL MEASUREMENT

A. Detection of the Stars Images

The image of the sky area obtained by each camera of the AVU contains blurred images of stars and an image of an optical background. The background is formed by the scattering of sunlight in the atmosphere. The strapdown principle of the AINS design requires to use cameras with a relatively wide field of view, at least 10°. The background illumination of the image in the camera of this kind will not allow to detect the blurred star's image in one frame (Fig. 2.). Therefore, sequential frames should be accumulated before detecting of stars images. The blur correction is performed in each frame under accumulation. To do this, the blur trajectory - the curve along which the star's image moves during the frame's exposure, is calculated in each pixel. The blur trajectory is calculated from the IMU measurements obtained during the exposure. This trajectory serves as the basis for synthesizing the impulse response of a two-dimensional digital correction filter matched with the motion blur [2].

The image inside each frame is distorted by a lens distortion, which is corrected using individually calibrated intrinsic parameters. Lens distortion must be taken into account when synthesizing the impulse response of the correction filter matched with the blur. Otherwise, serious losses of energy of the useful signal from the star occur, making the accumulation pointless [3].

As a result, the accumulated image, in which the background noise is smoothed out and the intensity of the

star images is increased is formed for each camera (Fig. 2.). The accumulated image is tied to the instant of the start of exposure of one selected frame within the accumulated sequence. This instant is taken as the epoch of the astronomical measurement.

Fig.2. Motion blur correction in the image of a single star, observed against the background of the daytime sky. On the left is an image of a star taken in the single frame. On the right is the result of accumulating of 40 consecutive frames with the star's images. The black spot is the starting point of the blur trajctory, the black line is the blur trajectory, the red square is the brightest pixel of the image.

The number of accumulated frames is calculated based on the optical background level. For calculation purposes, it is assumed that the optical background value has a Poisson distribution. As a rule, for stable detection of the brightest pixel, it is enough that the signal-to-noise ratio in this pixel is at least 5 [4]. The typical number of accumulated frames for various conditions of daytime observations lies in the range of 40...300 [5].

If R_n^i stars images have been detected in the frame of camera i, accumulated to the epoch t_n, then the bundle of unit direction vectors $\{\mathbf{s}_{CF^i\,n,r^i}^i\}_{r^i=1}^{R_n^i}$ is calculated from the brightest pixels inside these images. Each vector in this bundle is specified by its projections $\mathbf{s}_{CF^i\,n,r^i}^i = [s_{x\,CF^i\,n,r^i}^i \quad s_{x\,CF^i\,n,r^i}^i \quad s_{x\,CF^i\,n,r^i}^i]^{\mathrm{T}}$ on the CF^i. This vector is directed from the projective center of the camera i towards the star $r^i = 1...R_n^i$. This bundle is distorted by various types of refraction.

B. Refractions Correction

The AVU in the AINS for atmospheric applications is covered by a transparent porthole to protect it from effects of incoming air flow. The use of the porthole requires filling the internal volume of the AINS with a dry gas, followed by sealing. I.e. the atmospheric AVU observes stars from its own gaseous atmosphere, different from the atmosphere of Earth. As a result, the light beam coming from the star to the AVU is successively affected by three types of refraction: astronomical, aerodynamic and internal (Fig. 3.). Types of refraction are corrected against the direction of the light beam from the star, i.e. in the direction from the camera to the star [6].

To correct the refraction, the boundary between two media is assumed to be flat. Let starlight come from a gaseous medium I with a refractive index n, refract at the flat boundary and enter to a gaseous medium II with a refractive index n'. In the medium II, the star is visible with a direction vector $\mathbf{s}'_{RF}$, projected onto an arbitrary right-handed

orthogonal coordinate system RF (Reference Frame). The direction vector $\mathbf{s}_{RF}$, with which the star is visible in the medium I, is obtained from the visible vector $\mathbf{s}'_{RF}$:

$$\mathbf{s}_{RF} \equiv \mathrm{refr}(n, n', \mathbf{u}_{RF}, \mathbf{s}'_{RF}) = \frac{n'}{n}(\mathbf{I}_3 - \mathbf{u}_{RF}\mathbf{u}_{RF}^{\mathsf{T}})\mathbf{s}'_{RF} +$$

$$+\mathbf{u}_{RF}\sqrt{1 - \left(\frac{n'}{n}\right)^2 + \left(\frac{n'}{n}\mathbf{u}_{RF}^{\mathsf{T}}\mathbf{s}'_{RF}\right)^2},$$

where refr(...) is the function of the refraction correction; $\mathbf{u}_{RF}$ is the unit vector of the normal to the flat boundary, directed into the medium I (outward, towards the star) and specified by its projections on the RF.

Fig.3. Refraction when observing stars in the atmospheric AINS.

Internal refraction is the refraction on two sides of the porthole, inner and outer. The inner side separates the gas atmosphere, with a refractive index of μ, inside the AINS and the transparent porthole's material. The outer side separates the porthole's material and the aerodynamic boundary layer, with a refractive index n^{bl}, on the surface of the porthole. The value of μ is determined by the pressure and temperature of the gas inside the AINS. The value n^{bl} depends on the observation altitude and on the magnitude and direction of the AINS airspeed. The sides of the porthole are parallel and its material can be excluded from consideration. The refraction will be determined by the boundary between the boundary layer and the internal atmosphere of the AINS. Internal refraction is corrected in each direction vector: $\mathbf{s}^i_{CF^i\,n,r^i}$:

$$\mathbf{s}^{inr,i}_{CF^i\,n,r^i} = \mathrm{refr}(n^{bl}, \mu, \mathbf{u}^{hole,i}_{CF^i}, \mathbf{s}^i_{CF^i\,n,r^i}),$$

where $\mathbf{u}^{hole,i}_{CF^i}$ is the unit vector normal to the surface of the camera i porthole.

Aerodynamic refraction occurs at the boundary between the boundary layer and the surrounding steady atmosphere with a refractive index of $n(H)$, where H is the observation altitude above the sea level. For computational purposes, the boundary layer within the camera's field of view is modeled by an optical wedge. The attitude of the outer plane of the wedge, formed in front of the camera i, is described by the unit normal vector $\mathbf{u}^{bl,i}_{CF^i}$, defined by its projections onto the CF ^{i}. Refraction is corrected in each vector $\mathbf{s}^{inr,i}_{CF^i\,r^i}$:

$$\mathbf{s}^{bl,i}_{CF^i\,n,r^i} = \mathrm{refr}(n(H), n^{bl}, \mathbf{u}^{bl,i}_{CF^i}, \mathbf{s}^{inr,i}_{CF^i\,n,r^i}).$$

Bundles $\{\mathbf{s}^{bl,i}_{CF^i\,n,r^i}\}_{r^i=1}^{R^i_n}$, received from all cameras $i = 1 \dots I$ at the measurement epoch t_n, are projected onto VF:

$$\mathbf{s}^{bl,i}_{VF\,n,r^i} = \mathbf{C}^{CF^i}_{VF}\mathbf{s}^{bl,i}_{CF^i\,n,r^i}.$$

The result is the combined vectors bundle $\{\mathbf{s}^{bl}_{VF\,n,r}\}_{r=1}^{R^{tot}_n}$, where $R^{tot}_n = \sum_{i=1}^{I} R^i_n$ is the total number of stars detected by all cameras; $r = r^i + \sum_{q=1}^{i-1} R^q_n$ - sequential number of the star r^i in the combined bundle.

Astronomical refraction occurs when stars are observed through the atmosphere of Earth. This refraction is corrected in the vectors bundle $\{\mathbf{s}^{bl}_{VF\,n,r}\}_{r=1}^{R^{tot}_n}$. If angles ψ^{VF}_n, ϑ^{VF}_n, γ^{VF}_n are known, then this type of refraction is corrected similarly to the previous cases. The correction is performed in each vector $\mathbf{s}^{bl}_{VF\,n,r}$ taking into account the unit refractive index of an airless space outside the atmosphere of Earth (above 100 km):

$$\mathbf{s}^{astro}_{VF\,n,r} = \mathrm{refr}(1, n(H), \mathbf{u}^{vert}_{VF}, \mathbf{s}^{bl}_{VF\,n,r}),$$

where $\mathbf{u}^{vert}_{VF} = \mathbf{S}^{ENU}_{VF}(\psi^{VF}_n, \vartheta^{VF}_n, \gamma^{VF}_n)[0\ 0\ 1]^{\mathsf{T}}$ is the direction vector of the local vertical in the point of observation, projected onto the VF; $\mathbf{S}^{ENU}_{VF}(\psi^{VF}_n, \vartheta^{VF}_n, \gamma^{VF}_n)$ is the coordinates transformation matrix from the ENU to the VF, calculated from angles ψ^{VF}_n, ϑ^{VF}_n, γ^{VF}_n.

The refraction model used here (within a plane-parallel atmosphere) does not depend on either an azimuth of an observed star or a heading angle of a virtual camera. Therefore, it is possible to construct the algorithm for converting the vector $\mathbf{s}^{bl}_{VF\,n,r}$ to the $\mathbf{s}^{astro}_{VF\,n,r}$, depending only on the attitude of the "virtual camera" relative to the local vertical, i.e. depending only on the angles ϑ^{VF}_n, γ^{VF}_n [6].

The effects of aerodynamic and astronomical refractions are demonstrated as families of plots of angles between direction vectors before and after refraction correction. To estimate the refractive index n^{bl}, the simplest model of a laminar boundary layer formed on a surface of a flat plate moving with an air speed V_{AS} at a height H is used [7]. The family of graphs of the aerodynamic refraction angles $\Delta\zeta^{bl,i}_{n,r^i} = \arccos(\mathbf{s}^{bl,i\ \mathsf{T}}_{CF^i\,n,r^i}\mathbf{s}^{inr,i}_{CF^i\,n,r^i})$ for the bundle of vectors of dimension $R^i_n = 21$, received by camera i, is shown on Fig.4. The family of graphs of the astronomical refraction angles $\Delta\zeta^{astro}_{VF\,n,r} = \arccos(\mathbf{s}^{bl\ \mathsf{T}}_{VF\,n,r}\mathbf{s}^{astro}_{VF\,n,r})$ for the same bundle is shown in Fig.5.

As a result of these transformations, the bundle $\{\mathbf{s}^{astro}_{VF\,n,r}\}_{r=1}^{R^{tot}}$ of direction vectors of stars is obtained, which could be detected by the "virtual camera" in a vacuum.

Fig.4. Dependences of the aerodynamic refraction angle on the AINS airspeed at various observation altitudes for the camera with the field of view of 10° and the zenith angle of 35°.

Fig.5. Dependences of the angle of astronomical refraction on the observation altitude for various zenith angles of the camera with the field of view of 10°.

C. Taking into Account the Relativistic Velocity Aberration for Recognized Stars

The AINS is involved in the orbital motion of Earth, so visible directions to stars are distorted by a relativistic velocity aberration. The magnitude of the velocity aberration lays in the range $0''\ldots21''$ and depends on the direction of observation. The presence of aberration does not interfere with the recognition of stars, detected at the previous stages of obtaining the astronomical measurement. As a result of recognition, the bundle $\{\mathbf{s}^{astro}_{VF\,n,r}\}_{r=1}^{R_{tot}}$ is opposed to the bundle $\{\mathbf{b}_{BCRS\,n,r}\}_{r=1}^{R_{tot}}$ of the same dimension. The coordinates of the vectors in this bundle are calculated in the inertial coordinate system BCRS (Barycentric Celestial Reference System) from the catalog coordinates of the recognized stars.

When obtaining the astronomical measurement, the presence of velocity aberration must be taken into account. To do this, the bundle $\{\mathbf{b}_{BCRS\,n,r}\}_{r=1}^{R_{tot}}$ is converted to the bundle $\{\mathbf{g}_{GCRS\,n,r}\}_{r=1}^{R_{tot}}$, specified in the inertial coordinate system GCSR (Geocentric Celestial Reference System). The transformation uses the velocity vector $\mathbf{v}_{\oplus n}$ of the Earth's center of mass relative to the BCRS at the epoch t_n, calculated from the ephemerides of Earth [8]. The transformed vector $\mathbf{g}_{GCRS\,n,r}$ takes into account only the orbital motion of Earth. Aberration from the daily rotation of Earth and annual parallaxes of stars are not taken into account, which gives the error in coordinate transformation no more than $1''$ [9]:

$$\mathbf{g}_{GCRS\,n,r} = \left(1 + \frac{\mathbf{v}_{\oplus n}^{\mathsf{T}}\mathbf{b}_{BCRS\,n,r}}{c}\right)^{-1} \times$$

$$\times \left(\frac{\mathbf{b}_{BCRS\,n,r}}{\gamma_n} + \frac{\mathbf{v}_{\oplus n}}{c} + \frac{\gamma_n - 1}{\gamma_n}\frac{(\mathbf{v}_{\oplus n}^{\mathsf{T}}\mathbf{b}_{BCRS\,n,r})\mathbf{v}_{\oplus n}}{v_{\oplus n}^2}\right),$$

where $v_{\oplus n} = |\mathbf{v}_{\oplus n}|$; $\gamma_n = (1 - v_{\oplus n}^2/c^2)^{-1/2}$; c is the velocity of light in vacuum.

The astronomical measurement at the epoch t_n is obtained from the solution of Wahba's problem for the catalog $\{\mathbf{g}_{GCRS\,n,r}\}_{r=1}^{R_{tot}}$ and measured $\{\mathbf{s}^{astro}_{VF\,n,r}\}_{r=1}^{R_{tot}}$ vector bundles [1]. The result of this solution is the attitude of the VF relative to the GCRS, expressed as the orthogonal matrix $\mathbf{S}^{GCRS}_{VF}(t_n)$. The error of astronomical measurement is determined by the spatial configuration of vector bundles and the calibration errors of the cameras intrinsics [10].

III. USE OF ASTRONOMICAL MEASUREMENTS

In the process of inertial navigation, the position, speed and attitude of the MF relative to any coordinate system rigidly connected to Earth are determined. Therefore, the measured attitude of the AVU relative to the GCRS must be converted to the attitude of the IMU relative to the GCRS, i.e. the coordinate transformation matrix $\mathbf{S}^{GCRS}_{MF}(t_n)$ from the GCRS to the MF should be obtained:

$$\mathbf{S}^{GCRS}_{MF}(t_n) = \mathbf{C}^{VF}_{MF}\,\mathbf{S}^{GCRS}_{VF}(t_n).$$

The GCRS is not connected to Earth in any way, except for the single moment in time t_n. At the instant t_n the position and velocity vector of the GCRS origin coincide with the position and velocity vector of the center of mass of Earth. Therefore, matrices $\mathbf{S}^{GCRS}_{MF}(t_n)$, by themselves, are completely useless. They need to be tied to the surface of Earth using some additional information that cannot be obtained from the astronomical observations.

The numerical value of the measurement epoch t_n is synchronized with the UTC(SU) time scale. This value makes it possible to determine the rotation of Earth relative to the GCRS. This rotation is expressed as the orthogonal coordinate transformation matrix $\mathbf{S}^{ITRS}_{GCRS}(t_n)$ from the ITRS (International Terrestrial Reference System) to the GCRS. ITRS is constantly attached to Earth and participates in its daily rotation and orbital motion. The origin of the ITRS is

located at the center of mass of Earth. The matrix $\mathbf{S}_{GCRS}^{ITRS}(t_n)$ is calculated from the rotation model of Earth, regularly updated by the IERS (International Earth Rotation Service) [11]. Using this matrix, it is possible to calculate the attitude matrix of the IMU relative to Earth. This matrix links the astronomical measurements to the body of Earth, but is still useless for navigating relative to its surface:

$$\mathbf{S}_{MF}^{ITRS}(t_n) = \mathbf{S}_{MF}^{GCRS}(t_n)\mathbf{S}_{GCRS}^{ITRS}(t_n).$$

If the angles ϑ_n^{MF} and γ_n^{MF} of the MF inclination relative to the local vertical at the point of observation are known, then we can write the system of trigonometric equations for three unknown angles: ψ_n^{MF}, latitude B_n and longitude L_n [12]:

$$\mathbf{S}_{MF}^{ITRS}(t_n) = \mathbf{S}_{MF}^{ENU}(\psi_n, \vartheta_n, \gamma_n)\mathbf{S}_{ITRS}^{ENU}(B_n, L_n)^{\mathsf{T}},$$

where $\mathbf{S}_{MF}^{ENU}(\psi_n^{MF}, \vartheta_n^{MF}, \gamma_n^{MF})$, $\mathbf{S}_{ITRS}^{ENU}(B_n, L_n)$ are the standard transformation matrices from the ENU to the coordinate systems specified in the footnotes.

Thus, adding the AINS tilt angles to the astronomical measurement allows us to determine the unknown heading angle, latitude and longitude of the AINS. The same system of equations makes it possible to determine three unknown attitude angles of the AINS (heading, roll, pitch), if we add the latitude and the longitude of the point of observation to the astronomical measurement. The height of the AINS above the surface of Earth is not determined from the astronomical measurement.

When converting the astronomical measurement to the geographic coordinates, the factor, equal to the radius of Earth $R_{\oplus} \approx 6371$ km, is introduced. This factor converts the angular error θ of the astronomical measurement to the linear error in determining the position of the AINS $\delta x \approx \theta R_{\oplus}$. For the characteristic error $\theta \approx 15''...25''$ of the astronomical measurement of the AVU, equipped with a single camera, the linear error of the AINS position will be $\delta x \approx 460$ m...620 m. If the AVU uses two cameras with non-overlapping fields of view, the situation with errors improves, but not radically. Now the error of astronomical measurement is $\theta \approx 10''$, and the linear error of the AINS position is $\delta x \approx 310$ m.

If the attitude angles of the AINS relative to Earth are determined from the astronomical measurement, then the errors in these angles include errors in specifying the additional information. Thus, 1 mile of AINS coordinate error gives $1'$ attitude angle's error (or 100 m of coordinate error $\approx 3''$ attitude angle's error).

Another source of errors that occurs when using the astronomical measurements to navigate relative to the surface of Earth is the time keeper. If δk is the relative instability of the frequency of the master oscillator, then after the time T of continuous operation of the AINS the error $\delta\varphi_{\oplus} = (\delta t + T\delta k)\,\Omega_{\oplus}$ in calculating the angle of rotation of Earth will accumulate, where δt is the error in the initial synchronization of the time keeper; $\Omega_{\oplus} \approx 15''/\sec$ is the Earth's rotation rate. This error will be included in the errors in determining both geographic coordinates and the AINS attitude angles. The instability of a temperature-compensated quartz oscillator is about $\delta k \approx 10^{-6}$. This instability leads to

the growing attitude error $1.3''/$day or to the position error of 40 m/day. The initial synchronization error $\delta t \approx 1\sec$ leads to the attitude error of $15''$ or the position error of 463 m.

IV. CONCLUSION

Astronomical measurements of the AVU are the measurements of the AVU's attitude relative to stars. By itself, this attitude is useless for navigation relative to the surface of Earth. In order to bring the practical significance for the astronomical measurements, they need to be tied to the surface of Earth using additional information received from external sources. If such information contains inclination angles of the AINS relative to the local vertical at the point of observation, then the geographic coordinates and heading of the AINS can be calculated from the astronomical measurements. If the geographic coordinates of the point of observation are available, then three attitude angles of the AINS relative to the surface of Earth can be calculated from the astronomical measurements. The height is not determined from the AVU's astronomical measurements.

The astronomical measurements are extremely inconvenient for accurate navigation relative to surface of Earth. On the one hand, the attitude of the AVU relative to the stars must be measured with an error of order of several arc seconds in a practically meaningful range of observation conditions (time of a day, inclination angles and angular velocity of the AINS, etc.). On the other hand, these precision optical attitude measurements lead to extremely low accuracy in determining the observer's coordinates in comparison with satellite navigation measurements.

The astronomical measurements can be useful in determining the precise attitude of the AINS relative to Earth. If the coordinates of the point of observation (with an error of several tens of meters) and the precise observation time are available, the error in determining the attitude of the AINS will be several tens of arc seconds. Naturally, such an error in determining the attitude can only be achieved with correct consideration of all kinds of effects associated with the propagation of starlight through the atmosphere of Earth.

REFERENCES

[1] F.L. Markley, "Attitude determination using vector observations and the singular value decomposition," Journal of The Astronautical Sciences 1988, vol. 36, no. 3, pp. 245–258.

[2] N.N. Vasilyuk, "Calculation of Motion Blur Trajectories in a Digital Image as a Special Problem of Inertial Navigation," Gyroscopy and Navigation 2023, vol. 14, no. 1, pp. 75–88, doi: 10.1134/S2075108723010078.

[3] N.N. Vasilyuk, "Taking into account the lens geometric distortion during accumulation of blurred star images in an astro-inertial attitude sensor," Computer Optics 2024, 48(3), pp. 465-459, doi: 10.18287/2412-6179-CO-1408.

[4] G.A Avanesov, R.V. Bessonov, V.A. Vavaev, E.A.Mysnik and others, "Airborne strapdown stellar-inertial navigation system," Mechanics, control and informatics (in books) 2011, № 2. pp. 13-35, EDN OJSJNB.

[5] N.N. Vasilyuk, "Accumulation of Motion-Blurred Star Images Obtained by a Strapdown Astroinertial Navigation System under Daytime Conditions," 2023 30th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), Saint Petersburg, Russia, 2023, pp. 1-5, doi: 10.23919/ICINS51816.2023.10168426.

[6] N.N. Vasilyuk, "Vectorial correction of the atmospheric refraction for an intra-atmospheric star sensor," Aerospace instrument-making 2022, № 9, pp. 31–44, doi: 10.25791/aviakosmos.9.2022.1299.

[7] H. Schlichting and K. Gersten, "Boundary-Layer Theory," Springer, 2000.

[8] Laboratory of Ephemeris Astronomy. Ephemeris EPM. URL: https://iaaras.ru/en/dept/ephemeris/epm/ (Access date: 06.04.2024 г.).

[9] N.N. Vasilyuk, "Vector correction of the stellar aberration for an intra-atmospheric star tracker," Aerospace instrument-making 2022, № 10, pp. 17–31, doi: 10.25791/aviakosmos.10.2022.13.

[10] N.N. Vasilyuk, "Star tracker error model taking into account calibration errors of intrinsic parameters of the digital camera," Gyroscopy and Navigation 2024, vol. 15, no 1.

[11] IERS Earth orientation data [Электронный ресурс]. URL: https://www.iers.org/IERS/EN/DataProducts/EarthOrientationData/eop.html. (Access date: 06.04.2024 г.).

[12] G.A Avanesov, R.V. Bessonov, A.N. Kurkina and others, "Development of the authonomous strapdown astroinertial navigation system,", 2012, 19th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), Saint Petersburg, Russia, 2012, pp. 147-162, EDN SHEBIB.

Research on GPS-assisted inertial/starlight high-precision information fusion algorithm

Mingming Zhang
Department of Informatics and control
Bauman Moscow State University
Moscow, Russia
mz6250641@gmail.com

K. A. Neusypin
Department of Informatics and control
Bauman Moscow State University
Moscow, Russia
neysipin@mail.com

Bin He
Department of Informatics and control
Bauman Moscow State University
Moscow, Russia
binhe93@njust.edu.cn

Abstract—When the high altitude and long flight UAVs carry out the tasks of high-altitude camera and regional mapping, they have higher requirements for navigation and positioning and attitude accuracy. GPS can obtain high accuracy position information, and with the development of CCD technology, the on-board starlight tracker can also obtain high accuracy attitude information. By utilizing the information fusion method, the information fusion of Jetion inertial guidance, star sensor and GPS can obtain high-precision attitude and positioning information. The error of Jetlink inertial guidance system is selected as the state of the combined navigation system, and in the combined inertial/satellite navigation algorithm, the difference between the latitude, longitude and altitude information given by the inertial guidance system and the corresponding information given by the GPS receiver is utilized to output the position observation; and the difference between the velocities given by the inertial guidance system and the GPS receiver is utilized to output the velocity observation. In the combined inertial/starlight navigation algorithm, the attitude matrix output from the star sensitizer and the equivalent attitude matrix calculated from the inertial guidance output are utilized to construct the measurements. Thereby, an airborne inertial/satellite/starlight high-precision combined navigation method is designed using the federal filtering technique. The results show that the information fusion algorithm of inertial guidance, GPS, and star sensitizer proposed in this paper can obtain high-precision attitude, velocity, and position information.

Keywords—combined navigation, star sensor, information fusion, IMU, GPS

I. Introduction (Heading 1)

High-altitude long-range UAVs have high requirements for navigation and positioning and attitude accuracy when performing tasks such as high-altitude videography and area mapping. The Global Positioning System (GPS) plays a key role in providing accurate position information, but its performance may be limited when GPS signals are interfered with. Inertial navigation relies on the data measured by gyroscopes and accelerometers, and the motion parameters such as position, velocity and attitude of the carrier are solved by inertial navigation, which has the advantages of no need for external information, strong anti-interference ability, good concealment, complete navigation information and high data update rate [1]. However, the integral operation in the inertial solving will lead to the accumulation of navigation error over time, and for the long-distance and long-time motion platforms, it is necessary to use other navigation information to correct the error.

GPS can obtain high-precision position information, and with the development of CCD technology, the airborne star tracker can also obtain high-precision attitude information. Using the information fusion method to fuse Jetlink inertial guidance, star sensitizer and GPS can obtain high-precision attitude and positioning information. In contrast, starlight navigation utilizes the constellations in the sky as reference points to determine the position of the receiver. Starlight navigation has outstanding advantages such as high positioning accuracy, strong autonomy, and good anti-interference. Combining starlight navigation with inertial navigation and satellite navigation for combined navigation can effectively make up for the shortcomings of inertial guidance/satellite combined navigation.

For this reason, based on Jetlink inertial guidance, satellite navigation and starlight navigation systems, this paper investigates the inertial/satellite/starlight high-precision combination navigation method, which realizes high-precision attitude fixing by using the federated filtering technique to overcome the problem of the inertial guidance system's measurement error accumulating over time, and also obtains the vehicle's attitude with respect to the local geographic coordinate system, so as to achieve the purpose of complementing its strengths and shortcomings [2]-[5]. Since both the inertial guidance system and the star sensor do not radiate any information to the outside, have strong anti-jamming property and are completely autonomous, the GPS-assisted inertial/star-sensitive high-precision attitude fixing method researched in this paper has the significant advantages of high precision, good concealment and strong autonomy, etc. The GPS-assisted inertial/star-sensitive high-precision attitude fixing method researched in this paper has the significant advantages of high precision, good concealment and strong autonomy.

II. Airborne inertial/satellite/starlight combined navigation system model

In the airborne GPS/inertial/starlight combination navigation system, the Jetlink inertial guidance system has the advantages of comprehensive navigation parameters, strong maneuvering and tracking capability, timely and continuous output, good concealment and strong anti-jamming capability, and is thus used as the basic system of the combination navigation [6]. In the starlight navigation system, the CCD star sensitizer with the highest accuracy of attitude measurement is adopted, which can output high-precision carrier attitude information in real time. Among them, the navigation coordinate system adopts the east-north-sky geographic coordinate system, the inertial guidance system outputs the carrier's position, speed and attitude information, the satellite navigation system outputs the carrier's position information, and the starlight navigation

system outputs the transformation matrix of the star-sensitive coordinate system relative to the inertial coordinate system.

Firstly, based on Jetlink inertial guidance, the compensation algorithm after inertia/satellite and the attitude solution algorithm for inertia/starlight combined navigation are designed to calculate and obtain the two sets of local optimal estimates of the state of the combined navigation system. Then, the two sets of local optimal estimates are sent to the main filter for global information fusion using the federal filtering technique to obtain the global optimal estimate of the system state. Finally, the obtained global optimal estimates of the Jet-Link inertial guidance error state are used to correct the error of the Jet-Link inertial guidance system in real time, and the corrected Jet-Link inertial guidance output is used as the output of the combined navigation system [7]-[9]. Therefore, the principle of this inertial/satellite/starlight combined navigation system is shown in Fig. 1.

Fig. 1. GPS-assisted inertial/starlight high-precision information fusion modeling

III. ERROR MODELING AND INFORMATION FUSION ALGORITHMS FOR NAVIGATION SYSTEMS

A. Error modeling for starlight navigation systems

The measurement accuracy of the CCD star sensor in the starlight navigation system is very high and the measurement error does not accumulate over time, however, the installation error of the star sensor must be strictly calibrated. In particular, the star sensor is affected by the external temperature during the missile flight, and there will be a large error between the actual installation matrix and the laboratory calibration parameters, which will seriously affect the star navigation accuracy. Therefore, the installation error of the star sensitizer needs to be considered in the starlight navigation system, and the installation error of the star sensitizer along the three directions of the carrier x, y, and z, $\delta A_i (i = x, y, z)$, can be considered as a random constant value, i.e.

$$\delta \dot{A}_i = 0 \ (i = x, y, z) \tag{1}$$

B. Jet-Link inertial guidance system error model

There are many sources of errors in the Jetlink inertial guidance system, and this paper focuses on the inertial device errors and the resulting errors in the Jetlink inertial guidance system. After the inertial device errors are calibrated and compensated for, there are mainly some random drifts left that can not be calibrated, among which the gyro errors mainly include constant value drifts and white noise, and the accelerometer errors mainly include random constant value errors and white noise. As the inertial device itself has errors, this leads to some errors within the Jetlink inertial guidance system and its output navigation

parameters, specifically including inertial device errors, platform attitude angle errors, velocity errors and position errors.

C. Combined navigation system equation of state

The state equations use the basic navigation parameter error equations of the Jetlink inertial navigation system and the error equations of the inertial instrumentation, which specifically include the mathematical platform attitude errors ϕ_E, ϕ_N, ϕ_U of the Jetlink inertial guidance, the velocity errors δv_E, δv_N, δv_U, the position errors δL, $\delta \lambda$, δh, the gyroscope random constant value drifts bx by bz, the accelerometers random constant value errors ∇_{bx}, ∇_{by}, ∇_{bz}, and the star sensitizer mounting errors δA_x, δA_y, δA_z, i.e., the state vector of the combined navigation system X is

$$\begin{aligned} X = [&\phi_E, \phi_N, \phi_U, \delta v_E, \delta v_N, \delta v_U, \delta L, \delta \lambda, \delta h \\ &\varepsilon_{bx}, \varepsilon_{by}, \varepsilon_{bz}, \nabla_{bx}, \nabla_{by}, \nabla_{bz}, \delta A_x, \delta A_y, \delta A_z]^T \end{aligned} \tag{2}$$

According to the system error model, combined with the state vector X, the state equation of the combined inertial/satellite/starlight navigation system can be listed as

$$\dot{X}(t)_{18 \times 1} = A(t)_{18 \times 18} X(t) + G(t)_{18 \times 9} W(t)_{9 \times 1} \tag{3}$$

where $F(t)$ is the system state matrix, $G(t)$ is the system noise driven array, and $W(t)$ is the system white noise.

D. Measurement equations

In the GPS-assisted inertial/starlight combined navigation system, there are three groups of observations to choose from: the first group is the position observation, i.e., the difference between the latitude, longitude, and altitude information given by the inertial guidance system and the corresponding information given by the GPS receiver; the second group is the velocity observation, i.e., the difference between the velocities given by the inertial guidance system and the GPS receiver; and the third group is the attitude observation, i.e., the difference between the attitudes given by the inertial guidance and the starlight tracker.

$$Z(t) = \begin{bmatrix} v_{eI} - v_{eG} \\ v_{nI} - v_{nG} \\ v_{uI} - v_{uG} \\ (L_I - L_G)R_n \\ (L_I - L_G)R_e \cos L \\ h_I - h_G \\ \psi_{rI} - \gamma_{rS} \\ \theta_{pI} - \theta_{pS} \\ \psi_{hI} - \psi_{hS} \end{bmatrix} = \begin{bmatrix} \delta v_e + M_e \\ \delta v_n + M_n \\ \delta v_u + M_u \\ R_N \delta L + N_e \\ R_M \cos L \delta \lambda + N_n \\ \delta h + N_u \\ \delta \gamma + O_r \\ \delta \theta + O_p \\ \delta \psi + O_h \end{bmatrix}$$

$$= \begin{bmatrix} H_v(t) \\ H_p(t) \\ H_a(t) \end{bmatrix} X(t) + V(t) \tag{4}$$

O_r, O_p, O_h are the errors of the star tracker roll angle, pitch angle and heading angle. M_e, M_n, M_u are the velocity errors of GPS along the east, north and sky directions; N_e, N_n, N_u are the position errors of GPS along the east, north and sky directions.

The attitude angle error is obtained from the measurement information, while the platform error angle is used in the system state variables, so it is necessary to convert the attitude angle error into the Jetlink inertial guidance mathematical platform error. According to the definition of attitude angle and the conversion relationship between platform coordinate system and geographic coordinate system:

$$C_b^n = C_p^n \times C_b^p = \begin{bmatrix} 1 & -\phi_u & \phi_n \\ \phi_u & 1 & -\phi_e \\ -\phi_n & \phi_e & 1 \end{bmatrix} \times$$
$$\begin{bmatrix} \cos\gamma_I \cos\psi_I + \sin\gamma_I \sin\theta_I \sin\psi_I & \cos\theta_I \sin\psi_I & \sin\gamma_I \cos\psi_I - \cos\gamma_I \sin\theta_I \sin\psi_I \\ -\cos\gamma_I \sin\psi_I + \sin\gamma_I \sin\theta_I \cos\psi_I & \cos\theta_I \cos\psi_I & -\sin\gamma_I \sin\psi_I + \cos\gamma_I \sin\theta_I \cos\psi_I \\ -\sin\gamma_I \cos\theta_I & \sin\theta_I & \cos\gamma_I \cos\theta_I \end{bmatrix}$$
(5)

where C_b^n is the attitude transfer matrix between the carrier system and the geographic system，P denotes the platform coordinate system, γ_I, θ_I, ψ_I are the roll, pitch and heading angles output by the Jetlink inertial guidance, and there are

$$\begin{cases} \gamma_I = \gamma + \delta\gamma \\ \theta_I = \theta + \delta\theta \\ \psi_I = \psi + \delta\psi \end{cases}$$
(6)

The obtained linearized measurement equation for the attitude angle

$$\begin{bmatrix} \delta\gamma \\ \delta\theta \\ \delta\psi \end{bmatrix} = \frac{-1}{\cos\theta} \begin{bmatrix} \sin\psi & \cos\psi & 0 \\ \cos\psi\cos\theta & -\sin\psi\cos\theta & 0 \\ \sin\psi\sin\theta & \cos\psi\sin\theta & -\cos\theta \end{bmatrix} \begin{bmatrix} \phi_e \\ \phi_n \\ \phi_u \end{bmatrix}$$
(7)

E. GPS information delay processing

GPS positioning system belongs to active positioning, first by the ground control center through the geosynchronous satellite to the ground users send positioning inquiry signal, the user receives the signal and then sends out a positioning application, is after a series of signal request, calculation and transmission when the user receives the signal has generated a delay delay time for the user signal request to receive the signal time interval. Especially for highly dynamic users this delay phenomenon can lead to serious position errors.

Star Sensor is a high-precision attitude measurement instrument which shoots the starry sky through CCD elements, and from the photographed stars, it screens out the useful stars through certain conditions to measure their positions relative to the spacecraft, and compares them with the position parameters of the star in the ephemeris to determine the attitude of the spacecraft. The measurement data is large, and the processing and identification of the data can only be done by a computer. Therefore, the star sensitizer has a complex structure and consumes a lot of power, and it takes a certain amount of time to give a measurement result.

In the multi-information fusion process, the Kalman filter must utilize the measured values at the same moment in time

for filtering. However, there is a delay in the output information of the GPS receiver, which will definitely affect the accuracy of the measurement information, so it is necessary to design an algorithm to compensate for the hysteresis.

Assuming that the satellite information is obtained at the moment of t_{kp}, the lag time is Δt, and the output period of the Jetlink solver is T, the time sequence is shown in Fig. 2. It can be seen in the figure that the GPS information obtained at t_{kp} is actually the information at the time t_{kp} - Δt, therefore, to perform the combined filtering at the time t_{kp}, the actual GPS information at the time t_{kp} must be obtained.

Fig. 2. Timing diagram of SINS, GPS with measurement lag

Considering the high accuracy of the Jetlink inertial guidance used in the high-altitude long-haul UAV, its position and velocity accuracy in a short period of time is high, assuming that the information outputted by the Jetlink inertial guidance system in each cycle of the time period from the moment of t_{kp} - Δt to the moment of t_{kp} is as follows: v_{ei}, v_{en}, v_{ui} denote the eastward, northward, and celestial velocities outputted in each cycle, a_{ei}, a_{en}, a_{ui} denote the motion acceleration information of the corresponding cycle, L_i, h_i denote the latitude and altitude of the corresponding cycle.

Velocity Compensation Calculation Formula:

$$\begin{cases} \Delta v_e = \sum_{i=t_{kp}-\Delta t}^{t_{kp}} a_{ei} \times T \\ \Delta v_n = \sum_{i=t_{kp}-\Delta t}^{t_{kp}} a_{ni} \times T \\ \Delta v_u = \sum_{i=t_{kp}-\Delta t}^{t_{kp}} a_{ui} \times T \end{cases}$$
(8)

Positional Compensation Calculation Formula:

$$\begin{cases} \Delta\lambda = \sum_{i=t_{kp}-\Delta t}^{t_{kp}} \frac{1}{(R_{Ni}+h_i)\cos L_i} v_{ei} \times T \\ \Delta L = \sum_{i=t_{kp}-\Delta t}^{t_{kp}} \frac{1}{R_{Mi}+h_i} v_{ni} \times T \\ \Delta H = \sum_{i=t_{kp}-\Delta t}^{t_{kp}} a_{ui} \times T \end{cases}$$
(9)

Where Δv_e, Δv_n, Δv_u denote the eastward, northward, and skyward velocity compensation in the Δt time period, respectively; $\Delta \lambda$, ΔL, and ΔH denote the compensation of longitude, latitude, and altitude in the Δt time period, respectively. By utilizing the above formula, the actual GPS output information at the time of tkp can be obtained.

For the star sensitizer, the delay correction method is to use an array M to write down all the state information of SINS at each moment within 1s, and then randomly combine it with the SINS information to filter when the star sensitizer has measurement information output. The subfilter first generates a random number random- k from 0 to 1, and then Kalman filters the SINS/star sensitizer subfilter at random- k by combining the measurement information of the star sensitizer and the SINS state information at random- k. The estimated state value and mean value of the subfilter are then calculated. The state estimates and the mean square error array of the SINS/star sensitizer subfilter at the random- k time are calculated by Kalman filtering. The system state transfer array is then utilized for time updating.

F. Filtering algorithm

The combined inertial/GPS/starlight navigation system suffers from the problem of GPS measurement lag while at the same time the outputs of the GPS and star sensitizer are not synchronized. The Kalman filter update process can be divided into two steps, i.e., time update and measurement update. It can be seen that the Kalman filter time update is independent of the measurement information, and therefore, the time update can be performed independently. Based on this property, a combined asynchronous SINS/GPS/starlight filtering mechanism is developed.

When there is no output of star sensitizer and GPS receiver information, the time update of the Kalman filter is performed only according to equation (10), using the properties of the system state transfer matrix;

Time Update:

$$\begin{cases} X_{k,k-1} = 0 \\ P_{k.k-1} = \Phi_{k,k-1} P_{k-1.k-1} \Phi_{k,k-1}^T + \Gamma_{k,k-1} Q_{k-1} \Gamma_{k,k-1}^T \end{cases} \quad (10)$$

Measurement Updates:

$$X_{k,k} = K_k Z_k$$
$$K_k = P_{k,k-1} H_k^T [H_k P_{k,k-1} H_k^T + R_k]^{-1} \quad (11)$$
$$P_{k,k} = [I - K_k H_k] P_{k,k-1} [I - K_k H_k]^T + K_k P_k K_k^T$$

When new measurement information is available for the star-only sensitizer, The horizontal position information from the output of the inertial guide and the output information from the star sensitizer are utilized to solve the carrier attitude and perform the attitude combination to complete the time update of the Kalman filter and the metrology update;

At the same time, there are GPS information and star sensitizer information output, so the horizontal position information output from the inertial guidance and the star sensitizer information are used to solve the carrier attitude, combine the attitude, velocity, and position, and complete the time updating and measurement updating of the Kalman filter.

IV. EXPERIMENTAL ANALYSIS

Fig. 3. Position error for combined inertial/GPS/starlight navigation

Fig. 4. Position error curve

From the simulation curves, it can be seen that all the errors of the combined system converge rapidly without divergence and the steady state error values are satisfactory. Because of the addition of the star sensor, we can see that the attitude information of the system has been greatly improved. However, the speed and position of the system do not improve significantly because the high accuracy of the star sensitizer information has little effect on its correction, which mainly depends on the accuracy of SINS and BeiDou positioning system.

ACKNOWLEDGMENT *(Heading 5)*

that the asynchronous centralized filtering algorithm proposed in this chapter can effectively fuse multiple information to obtain high-precision attitude and position information, and can effectively solve the problem of GPS measurement output lag.

REFERENCES

[1] Bin Zhu, Juan Zheng. "New development of inertial navigation and guidance technology in the united states," Aerospace China，Vol. 3 pp.15-16, 2008.

[2] Hong song Du, Shu ming Yuan. "CNS/GPS point calibration algorithm for INS," Optics & Optoelectronic Technology, Vol.13, pp. 50-54, 2015.

[3] R. Wang, Z. Xiong, J. Liu. "A new tightly-coupled INS/CNS integrated navigation algorithm with weighted multistars observations," Proceedings of the Institution of Mechanical Engineers Part G Journal of Aerospace Engineering, Vol. 230, pp. 698-712, 2016.

[4] X. L. Ning, G. J. Zhang, M. Z. Gui. "A fast calibration method of the star Sensor installation error based on observability analysis for the tigh tly coupled SINS/ CNS integrated navigation system," IEEE Sensors Journal, Vol.18, pp. 6794-6803, 2018.

[5] M. RANGASWAMY, H. MICHELSJ. "Adaptivesignalprocessinginnon-Gaussiannoisebackgrounds," Ninth IEEE Signal Processing Workshop on Statistical Signaland Array Processing, Vol.1109, pp. 53-56, 1998.

[6] K. Ito, K. Hoshinoo, M. Ito. "Differential positioning experiment using two geostationary satellites," IEEE Transactions on Aerospace and Electronic Systems, Vol.35, pp. 866-877, 1999.

[7] Congwei Hu, Wu Chen, Yongqi Chen. "Adaptive Kalman filtering for vehicle navigation," Journal of Global Positioning Systems, Vol.2, pp. 42~47, 2003.

[8] Yang Zhang，Zhanrong Jing. "Design and Experiment of SINS/ GPS Integrated Navigation System," Journal of Computers, Vol. 8, pp.76 － 83, 2013.

[9] Ｒong Wang，Zhi Xiong. "SINS/GPS/CNS Information Fusion System Based on Improved Huber Filter with Classified Adaptive Factors for High- speed UAVs," Position Location and Navigation Symposium (PLANS), IEEE, pp. 441 － 446, 2012.

Assessing the influence of temperature on the zero offset of silicon accelerometers

Vladislav Lopatin

Russian Metrological Institute of Technical Physics and Radio Engineering VNIIFTRI, Moscow region, Mendeleevo, Russia
lopatin@vniiftri.ru

Vladimir Pudlovsky

Russian Metrological Institute of Technical Physics and Radio Engineering VNIIFTRI, Moscow region, Mendeleevo, MPEI, Moscow, Russia
pudlovskiy@vniiftri.ru

Oleg Denisenko

Russian Metrological Institute of Technical Physics and Radio Engineering VNIIFTRI, Moscow region, Mendeleevo, Russia
denisenko@vniiftri.ru

Abstract — **The purpose of the work is to experimentally evaluate the influence of temperature on the zero offset of silicon accelerometers, taking into account the assessment of changes in the frequency of the reference oscillator as part of a strapdown inertial navigation system. An analysis of the results of processing two series of experiments showed a more significant dependence of the zero offset of the ADXL320 silicon accelerometers due to a change in the frequency of the reference oscillator when its temperature changes as part of the navigation system than a change in the temperature of the accelerometers alone.**

Key words - accelerometer, temperature zero offset, inertial measurement unit, frequency instability of the reference oscillator

I. Введение

Currently, platform or strapdown inertial navigation systems (SINS) integrated with a receiver of global navigation satellite systems (GNSS) are used to navigate various vehicles [1, 2]. The main components of SINS are accelerometers and gyroscopes. A well-known feature of SINS is the increase in the error in determining coordinates and velocities over time. The magnitude of these errors is influenced by the instrumental errors of inertial sensors, the accuracy of their installation in the block of sensitive elements (SE) and the quality of the initial installation of SINS [2-6]. Assessing the influence of instrumental errors on the accuracy of determining motion parameters allows us to correctly formulate the requirements for accelerometers and gyroscopes. This ensures the required quality of operation of the entire inertial navigation system [7, 8].

The literature identifies zero offsets, scaling factor conversion errors, and random components in the form of white noise. [3]. The zero offset is an additive component of the error of the primary measurements of the sensors [3, 6].

A preliminary analysis of the literature [2-8] on the study of error components of SINS SEs (accelerometers, gyroscopes) showed that the influence of instability of the frequency of the reference oscillator (RO) as part of the SINS on the errors of its measurements is practically not taken into account.

At the same time, only two works analyzed the influence of the instability of the time scale of the inertial measurement unit on the SINS errors [9, 10].

It should be noted that the RO is one of the key elements of almost all types and designs of GNSS receivers and largely determines the accuracy of this equipment, as well as other onboard navigation systems and complexes [1]. Currently, most of the RO used in GNSS receivers and SINS are quartz..

The influence of instability of the RO in GNSS receivers on the error in determining coordinates and corrections to the time scale has been described and studied by many authors [1, 11]. Today, GNSS receiver designers can consciously select the type and characteristics of the RO based on their requirements.

The purpose of the work is to experimentally evaluate the effect of temperature on the zero offset of a block of three silicon accelerometers, taking into account the assessment of changes in the frequency of the RO as part of the SINS.

II. EXPERIMENTS TO ESTIMATE THE ACCELEROMETER ZERO Offset

A. Description of the experimental stand

A special stand was created to conduct experiments to evaluate the effect of temperature on the zero offset of silicon accelerometers of the ADXL320 type [12] as part of the Javad SINS SINS in the laboratory of the Russian Metrological Institute of Technical Physics and Radio Engineering VNIIFTRI. The block diagram of the stand and the view of the SE in the thermostat are shown in Figure 1,2.

Fig. 1. Block diagram of the experimental stand

The following abbreviations are used in Figure 1: ADC – analog-to-digital converter; RO – reference generator; QSF – quantum standard of frequency; PC – personal computer.

As can be seen from the block diagram of the stand, the use of a thermostat allows one to evaluate the influence of temperature on the RO as part of the SINS and/or accelerometers while simultaneously monitoring the temperature both inside and outside the thermostat.

Using this stand, from November 2023 to January 2024, two series of experiments were carried out to assess the effect of temperature on the relative displacement of the zeros of stationary accelerometers, the block of which was installed horizontally.

Fig. 2.   View of the SE in the thermostat

B. Estimation of the zero offset of accelerometers when the temperature of only the RO in SINS changes

In the first series of experiments, the temperature of the accelerometer block was stabilized using a thermostat, and the temperature of the SINS computing block together with the reference generator was changed by changing the ambient temperature within the range of 20,12…25,15°C. The temperature of the SINS sensitive elements inside the thermostat was measured using a PT1000 temperature sensor, and the temperature outside the thermostat and the SINS computing unit was recorded using a precision temperature meter MIT8.30. The measurement time was several tens of minutes. At the same time interval, using a Pendulum CNT-90XL frequency meter, the frequency value of the RO in the SINS was measured relative to an external QFS of the NAP-KPN type [13]. Simultaneously with measurements of the temperature and frequency of the RO, measurements of a stationary block of silicon accelerometers were recorded using a PC connected to the SINS interface. Thus, the zero offset of these sensitive elements was assessed when the temperature of the SINS computing unit changed.

The results obtained in one of the experiments of the first series are presented in figure 3. The acceleration values presented in this figure and below correspond to the absolute values of the accelerations along the axes, taking into account the subtraction of the first value in a series of measurements. Thus, the results presented are incremental accelerations relative to the first values in a series of measurements.

Fig. 3.   Estimation of zero offset of thermally insulated accelerometers when the reference generator is heated

C. Estimation of the zero offset of accelerometers when the temperature of only accelerometers changes

In the second series of experiments, only the block of sensitive elements inside the thermostat was subjected to periodic heating within the range of 25.17...30.1°C. As before, temperature measurements were carried out inside and outside the thermostat, and the frequency value RO of the SINS was also measured relative to the external QSF. Recording of readings from a stationary block of silicon accelerometers, as in the first series, was carried out using a PC on a single time scale. The measurement results obtained in one of the experiments of the second series are presented in Figure 4.

Fig. 4.   Estimation of zero offset during heating of accelerometers and thermal stabilization of the RO

III. PROCESSING THE RESULTS

The sample correlation coefficient was chosen to assess the proximity of the zero offset of the accelerometers and the temperature change inside and outside the thermostat with the accelerometer block, as well as the frequency of the RO built into the SINS.

The values of the sample correlation coefficient of the above measured values were calculated identically in accordance with [14]. As an example, the expression for

calculating the sample correlation coefficient between the accelerometer zero offset and the change in the frequency of the RO of the SINS:

$$\mathbb{R} = \frac{\sum\limits_{i}(a_i - \bar{a}) \cdot (f_i - \bar{f})}{\sqrt{\sum (a_i - \bar{a})^2 \sum (f_i - \bar{f})^2}}, \qquad (1)$$

where a_i - acceleration value measured at the i-th moment of time; $\bar{a}$ - average acceleration calculated for a given sample; f_i - frequency value of the RO in SINS, measured at the i-th moment of time; $\bar{f}$ - average frequency of the reference oscillator in SINS for a given samples.

To assess the reliability of the sampling coefficient using formula (1), the following expression was used [15]:

$$s_{\mathbb{R}} = \sqrt{\frac{1 - \mathbb{R}^2}{n - 2}}, \qquad (2)$$

where n – number of measurements.

A correlation coefficient close to 1 will indicate a strong linear dependence of the zero offset of stationary accelerometers and the frequency of the RO in the SINS, the change of which is mainly determined by the change in the temperature of this RO.

IV. DISCUSSION

The results of processing two series of experiments when the temperature changes only the RO in the SINS (1st series) and the temperature changes only the accelerometer block (2nd series) are summarized in Tables 1 and 2.

Table 1 presents the calculated values of the sample correlation coefficients of the acceleration increment relative to the change in the temperature of the SINS (T_1) and relative to the change in the temperature of the accelerometer block inside the thermostat (T_2).

Table 2 presents the calculated values of the sample correlation coefficients of the acceleration increment relative to the change in the frequency of the RO as part of the SINS.

TABLE I.

Series of meas./ axis	Num. of meas.	Range of change T1, ºC	$\mathbb{R}_{T1}$	Range of change T2, ºC	$\mathbb{R}_{T2}$
Heating if RO					
1.1/x	2000	20.12-25.15	0.97	25.25-24.8	0.42
1.1/y	2000	20.12-25.15	0.96	25.25-24.8	0.36
1.1/z	2000	20.12-25.15	0.97	25.25-24.8	0.38
1.2/x	1000	20.12-22.48	0.98	24.8-25.1	0.33
1.2/y	1000	20.12-22.48	0.97	24.8-25.1	0.40
1.2/z	1000	20.12-22.48	0.97	24.8-25.1	0.38
Heating of accelerometers					
2.1/x	864	21.4-21.6	0.09	25.17-28.11	0.28
2.1/y	864	21.4-21.6	0.08	25.17-28.11	0.29
2.1/z	864	21.4-21.6	0.04	25.17-28.11	0.40
2.2/x	464	21.6-21.7	0.08	25.61-30.1	0.35
2.2/y	464	21.6-21.7	0.05	25.61-30.1	0.15
2.2/z	464	21.6-21.7	0.03	25.61-30.1	0.33

TABLE II.

Series of meas./ axis	Num. of meas.	Range of change T, ºC	$\mathbb{R}$	$s_{\mathbb{R}}$	Interval estimation (p=0.99) for $\mathbb{R}$
1	**Heating if RO**				
1.1/x	2000	20.12-25.15	**0.93**	0.008	0.91…0.95
1.1/y	2000	20.12-25.15	**0.95**	0.007	0.93…0.97
1.1/z	2000	20.12-25.15	**0.92**	0.009	0.90…0.94
1.2/x	1000	20.12-22.48	**0.97**	0.008	0.95…0.99
1.2/y	1000	20.12-22.48	**0.97**	0.008	0.95…0.99
1.2/z	1000	20.12-22.48	**0.96**	0.009	0.94…0.98
2	**Heating of accelerometers**				
2.1/x	864	25.17-28.11	0.19	0.033	0.1…0.27
2.1/y	864	25.17-28.11	0.09	0.034	0.02…0.18
2.1/z	864	25.17-28.11	0.16	0.034	0.07…0.25
2.2/x	464	25.61-30.1	0.26	0.045	0.14…0.38
2.2/y	464	25.61-30.1	0.026	0.046	0…0.14
2.2/z	464	25.61-30.1	0.23	0.045	0.11…0.31

A comparative analysis of the sample correlation coefficients presented in Table 1 shows a significant influence on the zero offset of the accelerometers of the temperature RO of the SINS, and not the temperature of the accelerometer block inside the thermostat.

Analysis of the results of measurement processing (Table 2), obtained in the 1st series of experiments, shows a strong connection between the zero offset of stationary accelerometers and the change in the frequency of the RO of the SINS, when the temperature of this generator changes. On the other hand, heating only the stationary block of accelerometers has virtually no effect on the zero offset of the accelerometers, provided that the frequency of the built-in reference oscillator (2 series) is stable.

It should be noted that if the zero offset of the accelerometers (1st series of experiments) is interpreted due only to changes in the temperature of the SINS as a whole, then the observed zero offset of the ADXL320 type accelerometer (see Fig. 3) correspond to the zero offset versus temperature from the specification for this type accelerometers [12].

CONCLUSION

Analysis of the results of both series of experiments has shown that for silicon accelerometers ADXL320, a change in the frequency of the RO as part of the SINS Javad with a change in temperature has a stronger effect on the zero offset of the accelerometer than a change in the temperature of the sensitive elements alone.

It should be assumed that the discovered effect of correlation between the zero offset of the sensitive elements and the change in the frequency of the RO as part of the SINS will be observed for other types of inertial sensors that use ADCs and OR without stabilizing the frequency.

For sensitive elements with a discrete (digital) interface, the models of systematic and noise errors of the SINS must be supplemented with parameters that take into account the change in the frequency value of the reference oscillator and the shift in the time scale of the measuring unit. Without an adequate model of SINS measurements and their errors in a common time scale with GNSS signal receivers, the effect of close integration cannot be obtained.

REFERENCES

[1] A.I. Perov "GLONASS. Modernization and development prospects", Moscow:Radiotekhnika, 2020. 1072 p. (in Russian).

[2] Kalchenko A.O. "Zadacha kalibrovki beskardannoy inertsial'noy navigatsionnoy sistemy v polete pri pomoshchi informatsii ot sputnikovoy navigatsionnoy sistemy (The problem of calibrating a gimballess inertial navigation system in flight using information from a satellite navigation system)", Ph.D. Thesis: Lomonosov Moscow State University, Moscow, Russia, 2016. 105 p. (in Russian).

[3] Chelpanov I.B., Kozlov D.P., Kochetkov A.V. "Creation of complex system of the organization of calibration of micromechanical sensors", "Naukovedenie", Vol. 8, No. 1 (2016) http://naukovedenie.ru/PDF/21TVN116.pdf

[4] Dzuev A.A. "Invariantnaya kalibrovka bloka akselerometrov besplatformennykh inertsial'nykh navigatsionnykh sistem (Invariant calibration of the accelerometer unit of strapdown inertial navigation systems)", Ph.D. Thesis: BMSTU, Moscow, Russia, 2017. 156 p. (in Russian).

[5] Kutovoy D.A. "Metodiki otsenki pogreshnostey, prisushchikh bloku chuvstvitel'nykh elementov v sostave besplatformennykh inertsial'nykh navigatsionnykh sistem (Methods for assessing errors inherent in a block of sensitive elements as part of strapdown inertial navigation systems)", Ph.D. Thesis: BSTU «VOENMECH» named after D.F.Ustinov, St.Petersburg, Russia, 2022. 146 p. (in Russian).

[6] Motorin A.V. "Identifikatsiya modeley pogreshnostey navigatsionnykh datchikov i sredstv korrektsii metodami nelineynoy fil'tratsii (Identification of error models of navigation sensors and correction tools using nonlinear filtering methods)", Ph.D. Thesis: ITMO University, SPb., Russia, 2017. 134 p. (in Russian).

[7] A.M. Lestev, M.V. Fedorov, S.D. Evstafiev, "Identification of the noise structure of micromechanical inertial transducers of motion parameters". Radiopromyshlennost, 2019, vol. 29, no. 2, pp. 69–75 (In Russian), doi: 10.21778/2413‑9599‑2019‑29‑2-69-75.

[8] S.V. Andreyev, V.V. Ilinykh, O.A. Ilinykh, M.S. Chertkov, A.V. Klyuchnikov, "Estimation of inertial sensors error impact on platformless inertial navigation system accuracy". Journal of «Almaz – Antey» Air and Space Defence Corporation. 2018; no2, pp. 29-34. (In Russian), doi:10.38013/2542-0542-2018-2-29-34.

[9] N.I. Krobka, "On the influence of non-ideal onboard time scale on the structure of error equations and the accuracy of strapdown inertial navigation systems", 21st Saint Petersburg International Conference on Integrated Navigation Systems, pp. 361-364, 2014.

[10] R. Kulikov, O. Denisenko, I. Merkuryev and O. Glukhov, "Modeling the Influence of Time Scale Instability on Inertial Navigation Error," 29th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), Saint Petersburg, Russian Federation, 2022, pp. 1-4, doi: 10.23919/ICINS51784.2022.9815338.

[11] A.Y. Shatilov, "Reference Oscillator Short-Term Drift as it's Sensed by GNSS Receiver", Proceedings of the 27th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2014), Tampa, Florida, September 2014, pp. 2625-2634.

[12] ADXL320 Silicon Accelerometer Specification: Accessed: 02.03.2024, https://www.alldatasheet.com/datasheet-pdf/pdf/105155/AD/ADXL320.html.

[13] Rubidium subminiature quantum frequency standart NAC-CPT. Accessed: 02.03.2024, https://www.vniiftri.ru/catalog/products/pribory/sredstva-izmereniy/etalony-vremeni-i-chastoty/standart-chastoty-rubidievyy-sverkhminiatyurnyy-kvantovyy-na/.

[14] National standard GOST R ISO 3534-1-2019 Statistical methods. Vocabulary and conventions. Part 1. General statistical terms and terms used in probability theory. (in Russian).

[15] Karasev A.I. "Probability theory and mathematical statistics", Moscow, "Statistics", 1979. 279 p. (in Russian).

Algorithmic support of a tightly coupled navigation system of a ground-based mobile object

V.A. Pogorelov

Aircraft engineering
Don State Technical University
Rostov-on-Don, Russia
vadim.pogorelov.rnd@gmail.com

Abstract — **The problem of integrating a Strapdown Inertial Navigation Systems (SINS) with a global navigation satellite system (GNSS) and a Doppler speed sensor (DSS) for ground-based mobile objects (GBO) has been solved. In order to increase the noise immunity of the integrated navigation system under interference conditions, a tightly coupled scheme was chosen that does not impose a strict limit on the number of visible satellites. A stochastic model of the navigation system in the form of an "observer object" has been constructed for the selected instrument composition and estimation algorithms have been synthesized to ensure the required accuracy of GBO navigation both in integrated mode and in the absence of GNSS signals. The results of simulation modeling are presented.**

Keywords — strapdown inertial navigation systems, global navigation satellite system, Kalman filter

I. Introduction

The problem of creating GBO navigation systems that ensure the required accuracy of determining all navigation parameters in both autonomous and integrated modes (using GNSS signals) is highly relevant [1, 2].

Solving this problem becomes significantly more difficult under conditions of interference of various physical natures and especially when GNSS signals disappear in conditions of dense urban development, complex terrain, tunnels and interference caused by the passage of the signal through the ionosphere, troposphere [3].

The accuracy of GNSS also depends on the composition of the satellite constellation, ephemeris support, measurement noise of the GNSS receiver and GBO speed, and possible re-reflections of the GNSS signal at the receiving point [4]. The low rate of information output by modern GNSS receivers also hinders the effective solution of the GBO navigation task [5].

Intellectual interference is a separate problem [6-8]. To date, a number of methods have been developed to simulate true GNSS signals [9]:

- emission of rangefinder codes similar to GNSS signals, leading to incorrect estimates of the consumer's location and time;

- emission of signals from non-existent satellites;

- overlapping of the visible satellite signal with noise from the false satellite signal.

The existing disadvantages of GNSS and SINS lead to the widespread use of integrated navigation systems. In well-known practical applications, the following schemes for integrating GNSS readings with SINS are usually used [10, 11]:

1) separate,

2) loosely coupled,

3) strongly coupled;

4) a deep integration mode.

To solve the problem of GBO navigation, we will consider a strongly coupled scheme, since it does not involve significant modification of the instrument complex. In addition, to implement it, it is enough to have the readings of only one satellite [10-17]. Modern navigation algorithms based on a strongly coupled scheme involve the use of linear SINS error equations [10, 11]. Their application leads to methodological errors due to the loss of nonlinear relationships. This approach ensures the required accuracy of solving the problem of navigation of a wide class of GBO only for a limited time interval [10].

The purpose of the work. Based on a strongly coupled scheme for integrating stochastic readings of GNSS, INS and DSS, to work out a new algorithm for solving the GBO navigation problem that does not require linearization of its state equations.

II. Construction of stochastic equations of the state vector of the GBO navigation system in the form of an "object"

We will introduce into consideration the following coordinate systems (CS): the instrument CS (ICS) $J\,0xyz$, Earth-centered inertial (ECI), the Earth-centered earth-fixed (ECEF) and the accompanying CS (ACS) $OXYZ$. The orientation of the entered CS is given in [18].

The GBO navigation system includes SINS, built on the basis of three orthogonal accelerometers and three orthogonal angular velocity sensors (AVS), a two-axis DSS, and a GNSS signal receiver.

The interference of inertial and non-inertial sensing elements will be approximated by white Gaussian noise with zero mean and known intensity. Next we will use Euler angles as kinematic parameters. Their application is not a fundamental limitation for the proposed approach. If necessary, other kinematic parameters can be used instead of Euler angles, for example, the Rodrigue-Hamilton parameters, the matrices of guide cosines, the Cayley-Klein parameters, the Lushe numbers, the vector of final rotation, etc. [18]

The orientation of the instrument trihedron relative to the inertial one is described by a system of kinematic Euler equations [18]

$$\begin{bmatrix} \dot{\alpha} \\ \dot{\beta} \\ \dot{\gamma} \end{bmatrix} = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \dfrac{\sin\gamma}{\cos\alpha} & \dfrac{\cos\gamma}{\cos\alpha} & 0 \\ \sin\gamma tg\alpha & \cos\gamma tg\alpha & 1 \end{bmatrix} \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} = \Phi_{\alpha}(\alpha,\gamma)\omega_J , \qquad (1)$$

where α, β, γ – Euler angles,

$$\Phi_{\alpha}(\alpha,\gamma) = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \dfrac{\sin\gamma}{\cos\alpha} & \dfrac{\cos\gamma}{\cos\alpha} & 0 \\ \sin\gamma tg\alpha & \cos\gamma tg\alpha & 1 \end{bmatrix}, \quad \omega_J = \begin{bmatrix} \omega_x & \omega_y & \omega_z \end{bmatrix}^T - \text{the}$$

vector of the angular velocity of rotation ICS relative to ECI.. To determine, ω_J we will use the indications $\mathbf{Z_d} = \begin{bmatrix} Z_{dx} & Z_{dy} & Z_{dz} \end{bmatrix}^T$ AVS .

Following [19], the AVS output signals are represented as

$$\omega_J = \mathbf{Z_d} - \mathbf{m_d} - \mathbf{W_d} , \qquad (2)$$

where $\mathbf{m_d} = \begin{bmatrix} m_{dx} & m_{dy} & m_{dz} \end{bmatrix}^T$ – the expectation vector of the zero offset AVS; $\mathbf{W_d} = \begin{bmatrix} W_{dx} & W_{dy} & W_{dz} \end{bmatrix}^T$ – the AVS interference vector approximated by NW with zero mean and intensity matrix $\mathbf{D_d}$.

Taking into account (2), the Euler equations (1) take the form (3)

$$\begin{bmatrix} \dot{\alpha} \\ \dot{\beta} \\ \dot{\gamma} \end{bmatrix} = \Phi_{\alpha}(\alpha,\gamma)(\mathbf{Z_d} - \mathbf{m_d} - \mathbf{W_d}) . \qquad (3)$$

The orientation of ACS relative to ECI can also be found by integrating the Euler equations:

$$\begin{bmatrix} \dot{\varphi} \\ \dot{\lambda} \\ \dot{\chi} \end{bmatrix} = \begin{bmatrix} -\cos\chi & -\sin\chi & 0 \\ \dfrac{-\sin\chi}{\cos\varphi} & \dfrac{\cos\chi}{\cos\varphi} & 0 \\ -\sin\chi tg\varphi & \cos\chi tg\varphi & -1 \end{bmatrix} \begin{bmatrix} \omega_{XS} \\ \omega_{YS} \\ \omega_{ZS} \end{bmatrix} = \Phi_{\varphi}(\varphi,\chi)\omega_S , \quad (4)$$

where φ – GBO latitude, λ – GBO longitude, χ – azimuth of axis Y in ACS, $\Phi_{\varphi}(\varphi,\chi) = \begin{bmatrix} -\cos\chi & -\sin\chi & 0 \\ \dfrac{-\sin\chi}{\cos\varphi} & \dfrac{\cos\chi}{\cos\varphi} & 0 \\ -\sin\chi tg\varphi & \cos\chi tg\varphi & -1 \end{bmatrix}$,

$\omega_S = \begin{bmatrix} \omega_{XS} & \omega_{YS} & \omega_{ZS} \end{bmatrix}^T$ – angular velocity vector in ACS.

To build the projection ω_{XS} and ω_{YS} angular velocity of ACS, we will use the DSS readings Z_{Dx} and Z_{Dy}.

Projections V_x, V_y, expressed in terms of DSS readings have the form:

$$V_x = Z_{Dx} - W_{Dx} , \qquad (5)$$
$$V_y = Z_{Dy} - W_{Dy} ,$$

Measurement interference included in (5) DSS W_{Dx}, W_{Dy} can be described by a system of differential equations (6)

$$\dot{W}_{Dx} = f_{Dx}(W_{Dx},t) + f_{Dx0}(W_{Dx},t)\xi_{Dx} , \qquad (6)$$
$$\dot{W}_{Dy} = f_{Dy}(W_{Dy},t) + f_{Dy0}(W_{Dy},t)\xi_{Dy} ,$$

where $f_{Dx}, f_{Dy}, f_{Dx0}, f_{Dy0}$ – known functions, ξ_{Dx}, ξ_{Dy} – NW with zero average and known intensities D_{Dx} and D_{Dy}.

For the final definition of the ω_{XS} and ω_{YS} projections we will use projection V_z, which is not known yet. As a result, we will have:

$$\begin{bmatrix} \omega_{YS} \\ \omega_{XS} \end{bmatrix} = (r+h)^{-1} \begin{bmatrix} -C_{11} & -C_{21} & -C_{31} \\ C_{12} & C_{22} & C_{32} \end{bmatrix} \begin{bmatrix} Z_{Dx} - W_{Dx} \\ Z_{Dy} - W_{Dy} \\ V_z \end{bmatrix}, \quad (7)$$

where C_{ij} – ij - the matrix component $C(\alpha,\beta,\gamma,\varphi,\lambda,\chi)$, determining the orientation of the instrument trihedron relative to the maintainer

$$C(\alpha,\beta,\gamma,\varphi,\lambda,\chi) = D(\alpha,\beta,\gamma)\mathbf{B^T}(\varphi,\Omega t,\chi),$$

$$D = \begin{bmatrix} \sin\beta\sin\alpha\sin\gamma + \cos\beta\cos\gamma & \cos\alpha\sin\gamma & \cos\beta\sin\alpha\sin\gamma - \sin\beta\cos\gamma \\ \sin\beta\sin\alpha\cos\gamma - \cos\beta\sin\gamma & \cos\alpha\cos\gamma & \cos\beta\sin\alpha\cos\gamma + \sin\beta\sin\gamma \\ \sin\beta\cos\alpha & -\sin\alpha & \cos\beta\cos\alpha \end{bmatrix},$$

$\mathbf{B} = D(-\varphi,\Omega\tau + \lambda,-\chi)$.

From the basic equation of inertial navigation, written for the rotating Earth, we determine the ω_{ZS} and V_z projections in the form of:

$$a_x = \dot{V}_x + (2C_{(2)}\Omega_S + C_{(2)}\omega_S)V_z - (2C_{(3)}\Omega_S + C_{(3)}\omega_S)V_y + C_{(1)}\mathbf{g_S},$$
$$a_y = \dot{V}_y - (2C_{(1)}\Omega_S + C_{(1)}\omega_S)V_z + (2C_{(3)}\Omega_S + C_{(3)}\omega_S)V_x + C_{(2)}\mathbf{g_S} \quad (8)$$
$$a_z = \dot{V}_z + (2C_{(1)}\Omega_S + C_{(1)}\omega_S)V_y - (2C_{(2)}\Omega_S + C_{(2)}\omega_S)V_x + C_{(3)}\mathbf{g_S},$$

where $\mathbf{a} = \begin{bmatrix} a_x & a_y & a_z \end{bmatrix}^T$ – acceleration vector in ICS,

$\Omega_S = \begin{bmatrix} \Omega_X & \Omega_Y & \Omega_Z \end{bmatrix}^T$ – the vector of the angular velocity of the Earth rotation, which projections for the selected orientation of the ACS axes can be represented as:

$$\Omega_X = -\Omega\cos\varphi\sin\chi ,$$
$$\Omega_Y = \Omega\cos\varphi\cos\chi , \qquad (9)$$
$$\Omega_Z = \Omega\sin\varphi ,$$

$\mathbf{g_S} = \begin{bmatrix} g_X & g_Y & g_Z \end{bmatrix}^T$ – the gravity acceleration vector in ACS, which projections for the selected ACS orientation have the form:

$$g_X = -\Omega^2(r+h)\cos\varphi\sin\varphi\sin\chi ,$$
$$g_Y = \Omega^2(r+h)\cos\varphi\sin\varphi , \qquad (10)$$
$$g_Z = -\Omega^2(r+h)\cos^2\varphi - g_0(r,h,\varphi),$$

$C_{(i)}, i = \overline{1,3}$, – i - the row of the matrix $\mathbf{C}$.

From the second equation (8) we find the projection ω_{ZS}

Accelerometer output signal Z_y has the form:

$$a_y = Z_y - m_y - W_y , \qquad (11)$$

where Z_y – the output signal of the accelerometer, which sensitivity axis is directed along the axis y ICS,

m_y – mathematical expectation of accelerometer zero offset Z_y, W_y – interference at the output of an accelerometer which sensitivity axis is directed along the axis y ICS.

Taking into account the DSS readings Z_{D_x}, Z_{D_y} and (6), projections $\dot{V}_x, \dot{V}_y$ can be written as (12)

$$\dot{V}_x = \dot{Z}_{D_X} - f_{D_X} - f_{Dx0}\xi_{D_X}, \tag{12}$$

$$\dot{V}_y = \dot{Z}_{D_Y} - f_{D_Y} - f_{Dy0}\xi_{D_Y}.$$

The performed constructions allow us to record the angular velocity projection ACS ω_{Z_S} in the form of::

$$\omega_{Z_S} = \left(C_{13}V_z - C_{33}\left(Z_{D_x} - W_{D_x}\right)\right)^{-1}\left\{\dot{Z}_{D_y} - f_{D_y} - \right.$$
$$- f_{Dy0}\xi_{D_y} - Z_y + m_y + W_y - \tag{13}$$
$$- \left[2\mathbf{C_{(1)}}\boldsymbol{\Omega_S} + (r+h)^{-1}\left(\mathbf{C_{12}C_{T(1)}} - \mathbf{C_{11}C_{T(2)}}\right)\begin{bmatrix} Z_{D_x} - W_{D_x} \\ Z_{D_y} - W_{D_y} \\ V_z \end{bmatrix}\right]V_z +$$
$$+ \left[2\mathbf{C_{(3)}}\boldsymbol{\Omega_S} + (r+h)^{-1}\left(\mathbf{C_{32}C_{T(1)}} - \mathbf{C_{31}C_{T(2)}}\right)\begin{bmatrix} Z_{D_x} - W_{D_x} \\ Z_{D_y} - W_{D_y} \\ V_z \end{bmatrix}\right]\times$$
$$\times\left(Z_{D_x} - W_{D_x}\right) + \mathbf{C_{(2)}gs}\right\} = \Gamma_Z + \boldsymbol{\Gamma_{0Z}}\begin{bmatrix} W_y \\ \xi_{D_y} \end{bmatrix},$$

where

$$\Gamma_Z = \left(C_{13}V_z - C_{33}\left(Z_{D_x} - W_{D_x}\right)\right)^{-1}\left\{\dot{Z}_{D_y} - f_{D_y} - Z_y + m_y - \right.$$
$$- \left[2\mathbf{C_{(1)}}\boldsymbol{\Omega_S} + (r+h)^{-1}\left(\mathbf{C_{12}C_{T(1)}} - \mathbf{C_{11}C_{T(2)}}\right)\begin{bmatrix} Z_{D_x} - W_{D_x} \\ Z_{D_y} - W_{D_y} \\ V_z \end{bmatrix}\right]V_z +$$
$$+ \left[2\mathbf{C_{(3)}}\boldsymbol{\Omega_S} + (r+h)^{-1}\left(\mathbf{C_{32}C_{T(1)}} - \mathbf{C_{31}C_{T(2)}}\right)\begin{bmatrix} Z_{D_x} - W_{D_x} \\ Z_{D_y} - W_{D_y} \\ V_z \end{bmatrix}\right]\times$$
$$\left(Z_{D_x} - W_{D_x}\right) + \mathbf{C_{(2)}gs}\right\}$$

$$\Gamma_{0Z} = \left(C_{13}V_z - C_{33}\left(Z_{D_x} - W_{D_x}\right)\right)^{-1}\left[1 - f_{Dy0}\right] = \Gamma\left[1 - f_{Dy0}\right],$$

$\mathbf{C_{T(i)}}, i = \overline{1,3}$, – i-the row of the matrix $\mathbf{C^T}$.

Using the DSS readings, we obtain the equation for determining the height h (14)

$$\dot{h} = V_Z = C_{13}\left(Z_{D_x} - W_{D_x}\right) + C_{23}\left(Z_{D_y} - W_{D_y}\right) + C_{33}V_z. \tag{14}$$

To construct the GBO equations in a closed form, we find the projection V_z from the third equation (8)

$$\dot{V}_z = Z_z - m_z - W_z - \left(2\mathbf{C_{(1)}}\boldsymbol{\Omega_S} + (r+h)^{-1}\left|\mathbf{C_{12}C_{T(1)}} - \mathbf{C_{11}C_{T(2)}}\right|\times\right.$$
$$\times\begin{bmatrix} Z_{D_x} - W_{D_x} \\ Z_{D_y} - W_{D_y} \\ V_z \end{bmatrix} + C_{13}\Gamma_z + C_{13}\boldsymbol{\Gamma_{0z}}\begin{bmatrix} W_2 \\ \xi_{D_y} \end{bmatrix}\right)\left(Z_{D_y} - W_{D_y}\right) +$$

$$+ \left(2\mathbf{C_{(2)}}\boldsymbol{\Omega_S} + (r+h)^{-1}\left|\mathbf{C_{22}C_{T(1)}} - \mathbf{C_{21}C_{T(2)}}\right|\begin{bmatrix} Z_{D_x} - W_{D_x} \\ Z_{D_y} - W_{D_y} \\ V_z \end{bmatrix}\right) +$$
$$+ C_{23}\Gamma_z + C_{23}\boldsymbol{\Gamma_{0z}}\begin{bmatrix} W_2 \\ \xi_{D_y} \end{bmatrix}\right)\left(Z_{D_x} - W_{D_x}\right) - \mathbf{C_{(3)}gs}. \tag{15}$$

Combining the equations (1-15), we write the SINS state equations in the form:

$$\begin{bmatrix} \dot{\alpha} \\ \dot{\beta} \\ \dot{\gamma} \end{bmatrix} = \boldsymbol{\Phi_\alpha}(\alpha, \gamma)(\mathbf{Z_d} - \mathbf{m_d} - \mathbf{W_d}),$$

$$\begin{bmatrix} \dot{\varphi} \\ \dot{\lambda} \\ \dot{\chi} \end{bmatrix} = \boldsymbol{\Phi_\varphi}\begin{bmatrix} (r+h)^{-1}\begin{bmatrix} -C_{11} & -C_{21} & -C_{31} \\ C_{12} & C_{22} & C_{32} \end{bmatrix}\begin{bmatrix} Z_{D_x} - W_{D_x} \\ Z_{D_y} - W_{D_y} \\ V_z \end{bmatrix} \\ \hline \Gamma_Z + \boldsymbol{\Gamma_{0Z}}\begin{bmatrix} W_y \\ \xi_{D_y} \end{bmatrix} \end{bmatrix},$$

$$\dot{V}_z = Z_z - m_z - W_z - \left(2\mathbf{C_{(1)}}\boldsymbol{\Omega_S} + (r+h)^{-1}\left|\mathbf{C_{12}C_{T(1)}} - \mathbf{C_{11}C_{T(2)}}\right|\times\right.$$
$$\times\begin{bmatrix} Z_{D_x} - W_{D_x} \\ Z_{D_y} - W_{D_y} \\ V_z \end{bmatrix} + C_{13}\Gamma_z + C_{13}\boldsymbol{\Gamma_{0z}}\begin{bmatrix} W_2 \\ \xi_{D_y} \end{bmatrix}\right)\left(Z_{D_y} - W_{D_y}\right) +$$

$$+ \left(2\mathbf{C_{(2)}}\boldsymbol{\Omega_S} + (r+h)^{-1}\left|\mathbf{C_{22}C_{T(1)}} - \mathbf{C_{21}C_{T(2)}}\right|\begin{bmatrix} Z_{D_x} - W_{D_x} \\ Z_{D_y} - W_{D_y} \\ V_z \end{bmatrix}\right) +$$
$$+ C_{23}\Gamma_z + C_{23}\boldsymbol{\Gamma_{0z}}\begin{bmatrix} W_2 \\ \xi_{D_y} \end{bmatrix}\right)\left(Z_{D_x} - W_{D_x}\right) - \mathbf{C_{(3)}gs},$$

$$\dot{h} = C_{13}\left(Z_{D_x} - W_{D_x}\right) + C_{23}\left(Z_{D_y} - W_{D_y}\right) + C_{33}V_z,$$

$$\dot{W}_{D_x} = f_{D_x}\left(W_{D_x}, t\right) + f_{Dx0}\left(W_{D_x}, t\right)\xi_{D_x},$$

$$\dot{W}_{D_y} = f_{D_y}\left(W_{D_y}, t\right) + f_{Dy0}\left(W_{D_y}, t\right)\xi_{D_y}.$$

Or in vector-matrix form

$$\dot{Y} = \mathbf{F}(\mathbf{Y}, t) + \mathbf{F_0}(\mathbf{Y}, t)\boldsymbol{\xi}, \tag{16}$$

where

$$\boldsymbol{\xi} = \begin{bmatrix} \mathbf{W_d^T} & W_y & W_z & \xi_{D_x} & \xi_{D_y} \end{bmatrix}^T,$$

$$\mathbf{F_0}(\mathbf{Y}, t) = \begin{bmatrix} -\boldsymbol{\Phi_\alpha} & 0 & 0 & 0 & 0 \\ \hline 0 & \boldsymbol{\Phi_\varphi}\Gamma & 0 & 0 & \boldsymbol{\Phi_\varphi}\Gamma \\ \hline 0 & \Pi & -1 & 0 & \Pi \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & f_{Dx0} & 0 \\ \hline 0 & 0 & 0 & 0 & f_{Dy0} \end{bmatrix},$$

$$\Pi = C_{13}\Gamma\left(Z_{D_y} - W_{D_y}\right) + C_{23}\Gamma\left(Z_{D_x} - W_{D_x}\right).$$

$$F=\begin{bmatrix} \Phi_\alpha(\mathbf{Z_d}-\mathbf{m_d}-\mathbf{W_d}) \\ \hline \Phi_\varphi \quad \begin{bmatrix} (r+h)^{-1}\begin{bmatrix} -C_{11} & -C_{21} & -C_{31} \\ C_{12} & C_{22} & C_{32} \end{bmatrix}\begin{bmatrix} Z_{D_x}-W_{D_x} \\ Z_{D_y}-W_{D_y} \\ V_z \end{bmatrix} \\ \hline \Gamma_Z+\Gamma_{\mathbf{0_Z}}\begin{bmatrix} W_y \\ \xi_{D_y} \end{bmatrix} \end{bmatrix} \\ \hline Z_z-m_z-\left(2\mathbf{C_{(1)}}\Omega_\mathbf{S}+(r+h)^{-1}\left[C_{12}\mathbf{C_{T(1)}}-C_{11}\mathbf{C_{T(2)}}\right]\times \right. \\ \times\begin{bmatrix} Z_{D_x}-W_{D_x} \\ Z_{D_y}-W_{D_y} \\ V_z \end{bmatrix}+C_{13}\Gamma_z+C_{13}\Gamma_{\mathbf{0_Z}}\begin{bmatrix} W_2 \\ \xi_{D_y} \end{bmatrix}\right)(Z_{D_y}-W_{D_y})+ \\ +\left(2\mathbf{C_{(2)}}\Omega_\mathbf{S}+(r+h)^{-1}\left[C_{22}\mathbf{C_{T(1)}}-C_{21}\mathbf{C_{T(2)}}\right]\begin{bmatrix} Z_{D_x}-W_{D_x} \\ Z_{D_y}-W_{D_y} \\ V_z \end{bmatrix}+ \\ +C_{23}\Gamma_z+C_{23}\Gamma_{\mathbf{0_Z}}\begin{bmatrix} W_2 \\ \xi_{D_y} \end{bmatrix}\right)(Z_{D_x}-W_{D_x})-\mathbf{C_{(3)}gs} \\ \hline C_{13}(Z_{D_x}-W_{D_x})+C_{23}(Z_{D_y}-W_{D_y})+C_{33}V_z \\ \hline f_{D_x}(W_{D_x},t) \\ \hline f_{D_y}(W_{D_y},t) \end{bmatrix}$$

From the first equation (8) we obtain the equation of the "observer":

$$Z_x=\dot{Z}_{D_x}-f_{D_x}+\left(2\mathbf{C_{(2)}}\Omega_\mathbf{S}+(r+h)^{-1}\left(C_{22}\mathbf{C_{T(1)}}-C_{21}\mathbf{C_{T(2)}}\right)\times\right.$$

$$\times\begin{bmatrix} Z_{D_x}-W_{D_x} \\ Z_{D_y}-W_{D_y} \\ V_z \end{bmatrix}+C_{23}\Gamma_Z\Bigg)V_z-$$

$$-\left(2\mathbf{C_{(3)}}\Omega_\mathbf{S}+(r+h)^{-1}\left(C_{32}\mathbf{C_{T(1)}}-C_{31}\mathbf{C_{T(2)}}\right)\times\right.$$

$$\times\begin{bmatrix} Z_{D_x}-W_{D_x} \\ Z_{D_y}-W_{D_y} \\ V_z \end{bmatrix}+C_{33}\Gamma_Z\Bigg)(Z_{D_y}-W_{D_y})+$$

$$+\mathbf{C_{(1)}gs}+m_x+W_x-f_{Dx0}\,\xi_{D_x}+$$

$$+\left(C_{23}V_z-C_{33}(Z_{D_y}-W_{D_y})\right)\Gamma_{\mathbf{0_Z}}\begin{bmatrix} W_y \\ \xi_{D_y} \end{bmatrix}.$$

Or in a vector-matrix form

$$Z=Z_x-\dot{Z}_{D_x}-m_x=H(Y,t)+\mathbf{H_0}(\mathbf{Y,t})\zeta, \qquad (17)$$

where

$$H(Y,t)=-f_{D_x}+\left(2\mathbf{C_{(2)}}\Omega_\mathbf{S}+(r+h)^{-1}\left(C_{22}\mathbf{C_{T(1)}}-C_{21}\mathbf{C_{T(2)}}\right)\times\right.$$

$$\times\begin{bmatrix} Z_{D_x}-W_{D_x} \\ Z_{D_y}-W_{D_y} \\ V_z \end{bmatrix}+C_{23}\Gamma_Z\Bigg)V_z-$$

$$-\left(2\mathbf{C_{(3)}}\Omega_\mathbf{S}+(r+h)^{-1}\left(C_{32}\mathbf{C_{T(1)}}-C_{31}\mathbf{C_{T(2)}}\right)\times\right.$$

$$\times\begin{bmatrix} Z_{D_x}-W_{D_x} \\ Z_{D_y}-W_{D_y} \\ V_z \end{bmatrix}+C_{33}\Gamma_Z\Bigg)(Z_{D_y}-W_{D_y})+\mathbf{C_{(1)}gs},$$

$$\mathbf{H_0}(Y,t)=\left[1\vdots\left(C_{23}V_z-C_{33}(Z_{D_y}-W_{D_y})\right)\Gamma\vdots-f_{Dx0}\vdots\right.$$

$$-f_{Dy0}\left(C_{23}V_z-C_{33}(Z_{D_y}-W_{D_y})\right)\Gamma\right]$$

$$\zeta=\left[W_x\ W_y\ \xi_{D_x}\ \xi_{D_y}\right]^T.$$

III. SYNTHESIS OF A NONLINEAR STOCHASTIC FILTER OF THE AUTONOMOUS SINS, TAKING INTO ACCOUNT THE CORRELATION OF THE "OBJECT" AND THE "OBSERVER" NOISE

Since in the considered autonomous GBO SINS, the "object" and the "observer" noises are correlated, we represent (16) for the autonomous evaluation of the SINS state vector as (18)

$$\dot{\mathbf{Y}}=\mathbf{F}(\mathbf{Y,t})+\mathbf{F_{01}}(\mathbf{Y,t})\xi_1+\mathbf{F_{02}}(\mathbf{Y,t})\zeta, \qquad (18)$$

where $\mathbf{F_{01}}(\mathbf{Y,t})=\begin{bmatrix} -\Phi_\alpha & 0 \\ \hline 0 & 0 \\ \hline 0 & -1 \\ \hline 0 & 0 \\ \hline 0 & 0 \\ \hline 0 & 0 \\ \hline 0 & 0 \end{bmatrix}$, $\xi_1=\left[\mathbf{W_d^T}\ W_z\right]^T$,

$$\mathbf{F_{02}}(\mathbf{Y,t})=\begin{bmatrix} 0 & 0 & 0 & 0 \\ \hline 0 & \Phi_\varphi\Gamma & 0 & \Phi_\varphi\Gamma \\ \hline 0 & \Pi & 0 & \Pi \\ \hline 0 & 0 & 0 & 0 \\ \hline 0 & 0 & f_{Dx0} & 0 \\ \hline 0 & 0 & 0 & f_{Dy0} \end{bmatrix}.$$

Let's introduce intermediate variables

$$\gamma_1(\mathbf{Y,t})=\left[\mathbf{F_{01}}(\mathbf{Y,t})\mathbf{F_{02}}(\mathbf{Y,t})\right],\ \gamma_2(\mathbf{Y,t})=\left[0\vdots\mathbf{H_0}(\mathbf{Y,t})\right],$$

$$\eta=\left[\xi_1^T\ \zeta^T\right]^T. \qquad (19)$$

Taking into account (18) and (19), the equations of state of the autonomous SINS, written in the form of "object" (16) and "observer" (17), take the canonical form (20)

$$\dot{\mathbf{Y}}=\mathbf{F}(\mathbf{Y,t})+\gamma_1(\mathbf{Y,t})\eta, \qquad (20)$$

$$Z=H(\mathbf{Y,t})+\gamma_2(\mathbf{Y,t})\eta,$$

And the equations of nonlinear stochastic filtering of autonomous SINS take the form (21) [20]

$$\hat{\dot{\mathbf{Y}}}=\mathbf{F}(\hat{\mathbf{Y}},t)+\mathbf{K}(\hat{\mathbf{Y}},t)\left[Z-H(\hat{\mathbf{Y}},t)\right], \qquad (21)$$

$$\mathbf{K}(\hat{\mathbf{Y}},t)=\left\{\mathbf{R}\frac{\partial\mathbf{H}^\mathbf{T}(\hat{\mathbf{Y}},t)}{\partial\hat{\mathbf{Y}}}+\gamma_1(\hat{\mathbf{Y}},t)\mathbf{D}\gamma_2^\mathbf{T}(\hat{\mathbf{Y}},t)\right\}\left(\gamma_2(\hat{\mathbf{Y}},t)\mathbf{D}\gamma_2^\mathbf{T}(\hat{\mathbf{Y}},t)\right)^{-1}$$

$$\dot{\mathbf{R}}(\hat{\mathbf{Y}},t)=\frac{\partial\mathbf{F}(\hat{\mathbf{Y}},t)}{\partial\hat{\mathbf{Y}}}\mathbf{R}(\hat{\mathbf{Y}},t)+\mathbf{R}(\hat{\mathbf{Y}},t)\frac{\partial\mathbf{F}^\mathbf{T}(\hat{\mathbf{Y}},t)}{\partial\hat{\mathbf{Y}}}+\gamma_1(\hat{\mathbf{Y}},t)\mathbf{D}\gamma_1^\mathbf{T}(\hat{\mathbf{Y}},t)-$$

$$-\mathbf{K}(\hat{\mathbf{Y}},t)\left(\gamma_2(\hat{\mathbf{Y}},t)\mathbf{D}\gamma_2^\mathbf{T}(\hat{\mathbf{Y}},t)\right)\mathbf{K}^\mathbf{T}(\hat{\mathbf{Y}},t)$$

where

$$\hat{\mathbf{Y}}_0=\mathbf{M}(\mathbf{Y_0}),\ \mathbf{R_0}=\mathbf{M}\left\{\left(\mathbf{Y_0}-\hat{\mathbf{Y}}_0\right)\left(\mathbf{Y_0}-\hat{\mathbf{Y}}_0\right)^\mathbf{T}\right\},$$

$$\mathbf{D}=\begin{bmatrix} D_d & 0 & 0 \\ 0 & D_a & 0 \\ 0 & 0 & D_{D_x} \end{bmatrix}.$$

The application of equations (21) solves the problem of autonomous estimation of angular and linear phase variables

of the SINS state vector. However, during over a long period of time, the use of (21) leads to an increase in estimation errors of all the components of the SINS state vector. In this regard, in order to solve the problem of navigation over a long time interval, it becomes necessary to use the readings of the GNSS signals receiver [21].

IV. SYNTHESIS OF A STOCHASTIC MODEL OF SATELLITE MEASUREMENTS

In order to construct an observation equation that provides the required accuracy of estimating the entire state vector of the GBO navigation system, we will use GNSS code and Doppler measurements.

According to [4], the equations of code measurements have the form:

$$Z_R = \sqrt{(\xi_c - \xi)^2 + (\eta_c - \eta)^2 + (\zeta_c - \zeta)^2} + W_{ZR}, \quad (22)$$

where $\xi_c, \eta_c, \zeta_c -$ satellite coordinates in ECEF, $\xi, \eta, \zeta -$ GBO coordinates in ECEF, $W_{ZR} -$ interference of code measurements described by NW with zero mean and variance $D_{ZR}(t)$.

And for Doppler measurements Z_V:

$$Z_V = \left[(\xi_c - \xi)(V_{\xi c} - V_\xi) + (\eta_c - \eta)(V_{\eta c} - V_\eta) + (\zeta_c - \zeta)(V_{\zeta c} - V_\zeta) \right] \times \quad (23)$$

$$\times \left(\sqrt{(\xi_c - \xi)^2 + (\eta_c - \eta)^2 + (\zeta_c - \zeta)^2} \right)^{-1} + W_{ZV},$$

where $V_{\xi c}, V_{\eta c}, V_{\zeta c} -$ projections of the satellite velocity vector on the ECEF axis, $V_\xi, V_\eta, V_\zeta -$ projections of the GBA velocity vector on the ECEF axis.

To express the phase variables SINS through GNSS readings, we express the orientation of ACS through ECI [20]

$$\xi = (r+h)\cos\varphi\sin\lambda, \quad \eta = (r+h)\sin\varphi, \quad \zeta = (r+h)\cos\varphi\cos\lambda \quad (24)$$

and substitute the result in (22). As a result, we get:

$$Z_R = \sqrt{(\xi_c - (r+h)\cos\varphi\sin\lambda)^2 + (\eta_c - (r+h)\sin\varphi)^2 +} \quad (25)$$

$$\overline{+ (\zeta_c - (r+h)\cos\varphi\cos\lambda)^2} + W_{ZR}.$$

The vector $\mathbf{V_G} = [V_\xi \, V_\eta \, V_\zeta]^T$ can be expressed in terms of a vector $\mathbf{V_S} = [V_X \, V_Y \, V_Z]^T$ as:

$$\mathbf{V_S} = \mathbf{B}(\varphi, \lambda, \chi)\mathbf{V_G}. \quad (26)$$

It follows from (26) that

$$\mathbf{V_G} = \mathbf{B}^T(\varphi, \lambda, \chi)\mathbf{V_S}. \quad (27)$$

Taking into account (27), the information model of Doppler measurements takes the form (28):

$$Z_V = \left[(\xi_c - (r+h)\cos\varphi\sin\lambda)(V_{\xi c} - \mathbf{B}^T_{(1)}(\varphi, \lambda, \chi)\mathbf{V_S}) + \right.$$

$$+ (\eta_c - (r+h)\sin\varphi)(V_{\eta c} - \mathbf{B}^T_{(2)}(\varphi, \lambda,)\mathbf{V_S}) +$$

$$\left. + (\zeta_c - (r+h)\cos\varphi\cos\lambda)(V_{\zeta c} - \mathbf{B}^T_{(3)}(\varphi, \lambda, \chi)\mathbf{V_S}) \right] \times \quad (28)$$

$$\times \sqrt{(\xi_c - (r+h)\cos\varphi\sin\lambda)^2 + (\eta_c - (r+h)\sin\varphi)^2 +}$$

$$\overline{+ (\zeta_c - (r+h)\cos\varphi\cos\lambda)^2} + W_{ZV},$$

or in a compact form

$$Z_R = H_R(\varphi, \lambda, \chi, h) + W_{ZR}, \quad (29)$$

$$Z_V = H_V(\varphi, \lambda, \chi, h, V_X, V_Y, V_Z) + W_{ZV},$$

where $\mathbf{B}^T_{(i)}(\varphi, \lambda, \chi) - i-$ the row of the matrix $\mathbf{B}^T(\varphi, \lambda, \chi)$

or in vector-matrix form

$$\mathbf{Z}^{(c)} = \begin{bmatrix} Z_R \\ Z_V \end{bmatrix} = \begin{bmatrix} H_R(\varphi, \lambda, \chi, h) \\ H_V(\varphi, \lambda, \chi, h, V_X, V_Y, V_Z) \end{bmatrix} +$$

$$+ \begin{bmatrix} W_{ZR} \\ W_{ZV} \end{bmatrix} = \mathbf{H}^{(c)}(\mathbf{Y}, \mathbf{t}) + \zeta^{(c)} \quad (30),$$

In a compact form (30) takes the form (31):

$$\mathbf{Z_k^{(c)}} = \mathbf{H}^{(c)}(\mathbf{Y}, \mathbf{k}) + \zeta_\mathbf{k}^{(c)}, \quad (31)$$

where $k = 1, 2 \ldots$

The constructed observer (31) allows, following [22], to write down the equations of stochastic filtering of the integrated navigation system

$$\hat{\mathbf{Y}}(\mathbf{t_k} + \mathbf{0}) = \hat{\mathbf{Y}}_{\mathbf{k0}} + R(\mathbf{t_k} + \mathbf{0}) \frac{\partial \mathbf{H}^{(c)T}(\hat{\mathbf{Y}}_{\mathbf{k0}}, \mathbf{k})}{\partial \hat{\mathbf{Y}}} \mathbf{D_c^{-1}} \times$$

$$\times \left[\mathbf{Z_k^{(c)}} - \mathbf{H}^{(c)}(\hat{\mathbf{Y}}_{\mathbf{k0}}, \mathbf{k}) \right] \quad (32)$$

$$\mathbf{R}^{-1}(\mathbf{t_k} + \mathbf{0}) = \mathbf{R_{\kappa0}^{-1}} + \frac{\partial \mathbf{H}^{(c)T}(\hat{\mathbf{Y}}_{\mathbf{k0}}, \mathbf{k})}{\partial \hat{\mathbf{Y}}} \mathbf{D_c^{-1}} \frac{\partial \mathbf{H}^{(c)}(\hat{\mathbf{Y}}_{\mathbf{k0}}, \mathbf{k})}{\partial \hat{\mathbf{Y}}}.$$

The filter (21) is used in the sections of the SINS autonomous traffic when there is no GPS signal, and the filter (32) when there is a GNSS receiver reading. At the same time, the final estimates (21) were the initial conditions (32). Accordingly, after the GNSS signals disappeared, the final conditions (32) were the initial conditions for (21).

The use of stochastic filters (21) and (32) makes it possible to solve the problem of integrating SINS with GNSS using a strongly coupled scheme.

V. SIMULATION RESULTS

In order to test the effectiveness of the proposed approach to the construction of an integrated navigation system, simulation modeling of equations (21) and (32) in the ENGEE environment was carried out over a time interval of $t \in [0; 2000]$ seconds.

The ENGINE program is specially designed for simulation modeling of dynamic systems using the methodology of model-oriented design [23].

The initial conditions for modeling an instrument trihedron relative to an inertial one were defined as:

$$\alpha_0 = \frac{\pi}{3}, \beta_0 = -\frac{\pi}{5}, \gamma_0 = \frac{\pi}{4}.$$

The initial conditions for modeling the accompanying trihedron relative to the inertial one were set as follows:

$$\lambda_0 = \frac{\pi}{5}, \varphi_0 = \frac{\pi}{4}, \chi_0 = \frac{\pi}{3}$$

The parameters of the linear motion of the GBO were set by differential equations in projections on the ASC axis

$$\dot{V}_X = 200\exp(-0.1t) + 10\cos(0.25t),$$

$$\dot{V}_Y = 100\exp(-t) + 5\sin(0.25t),$$

$$\dot{V}_Z = 10\exp(-t) + \sin(0.1t).$$

The noise intensities included in the stochastic filtering equations (21), (32) were determined based on the achieved level of instrumentation and were calculated for:

accelerometers $-\left(10^{-5}\,\text{m/s}^2\right)^2$, DSS - $\left(0.5\text{m/s}\right)^2$, AVS – $\left(10^{-7}\,1/\text{s}\right)^2$, GNSS code measurements $-\left(5\text{m}\right)^2$, Doppler measurements of GNSS - $\left(0.2\text{m/s}\right)^2$.

During simulation GNSS signals were absent on the 100th, 300th, 700th, 1200th and 1500th seconds in a time interval of $10 - 30$ seconds.

The fourth-order Runge-Kutta method was used to integrate the filtering equations (21). The integration step was chosen Δt=0.01s equal as quite sufficient for navigation and control of most GBOs.

The solution (32) was implemented in steps of 1 sec. This step corresponds to the rate of information output by most modern GNSS receivers.

The estimation errors were calculated as the difference between the GBO model motion (obtained without taking into account the noise of the sensitive elements) and the estimates obtained at the output of the filtering algorithm. At the end of the modeling interval, the estimation errors were: by orientation angles – 0.1%, according to the projections of the GBO velocity vector $V_Z - 7\%$, $V_X - 5,8\%$, $V_Y - 5\%$, by longitude – $3{*}10^{-7}$ radian, by latitude – $3,5{*}10^{-7}$ radian.

Thus, the results of simulation modeling indicate the possibility of effective use of the algorithm described in the article for integrating inertial and radio engineering sensing elements according to a strongly coupled scheme in GBO navigation systems for various purposes.

VI. CONCLUSION

Based on the use of continuous and discrete stochastic filters, the problem of integrating SINS with GNSS using a strongly coupled scheme is solved. The proposed algorithm provides an assessment of all phase variables of the GPS navigation system both in the presence of GNSS signals and in their absence.

The application of the proposed algorithm in existing and promising GBO navigation systems will, on the one hand, improve the accuracy of their positioning in conditions of internal and external disturbances, and, on the other hand, it can be implemented with minimal modifications to the measuring complex of the integrated navigation system under study.

REFERENCES

[1] D. Dardari, M. Luise, E. Falletti, "Satellite and Terrestrial Radio Positioning Techniques. A Signal Processing Perspective", Oxford, Academic Press, 2011.

[2] S.A. Adebayo, I.A.A Azeez, "Review of global navigation satellite systems (GNSS) and its application", International Journal of Scientific and Engineering Research, 2021. vol 12 (4), pp. 1042-1049.

[3] "Undestending GPS. Principles and applications", Ed. by D. Kaplan, C.J. Hegarty. Second edition, Artech House, 2006.

[4] "Springer Handbook of Global Navigation Satellites Systems", Ed. by J.G. Peter Teunissen, Oliver Montenbruck, Springer International Publishing AG, 2017.

[5] B. L. Stevens and F. L. Lewis, "Aircraft Control and Simulation", John Wiley & Sons, Second edition, 2003.

[6] D. Shepard, J. Bhatti, T. Humphreys, "Drone Hack", GPS World Aug. 2012, pp. 30-33.

[7] Congressional Budget Office. The Global Positioning System for Military Users: Current Modernization Plans and Alternatives., Oct. 2011.

[8] A. R. Baziar, M. Moazedi, M. R Mosavi, "Analysis of single frequency GPS receiver under delay and combining spoofing algorithm", Journal of Wireless Personal Communications, 2015, vol. 83, no. 3, pp. 1955–1970.

[9] S. Mahmood, "Analysis of Differential Global Positioning System (DGPS) Techniques and GPS Jamming on Precision Guided Munition (PGM)" Performance, NATO/AGARD MSP Meeting, Technologies for Precision Air Strike Operations in Rapid Reaction and Localized Conflict Scenarios, Seville, Spain, AGARD CP-576, June 1996.

[10] G. Schmidt, R. Phillips, "INS/GPS Integration Architecture Performance Comparisons", NATO RTO Lecture Series, RTO-EN-SET-116, Low-Cost Navigation Sensors and Integration Technology, Mar. 2011.

[11] G. Schmidt, R.Phillips, "INS/GPS Integration Architectures", NATO RTO Lecture Series, RTO-EN-SET-116, Low-Cost Navigation Sensors and Integration Technology, Mar. 2011.

[12] I. N. Rozenberg, S.V. Sokolov, V.I. Umanskij and V.A. Pogorelov, "Theoretical Basis of Close Integration of Inertial Satellite Navigation Systems", Moscow: Publ. Fizmatlit, 2018.

[13] S.V.Sokolov, V.A. Pogorelov, "Stochastic Evaluation, Control and Identification in High-precision Navigation Systems", Moscow: Publ. Fizmatlit, 2016.

[14] S. V. Sokolov, V. A. Pogorelov, A. L. Okhotnikov "Stochastic Estimation of Angular Orientation Parameters of a Moving Object by Inertial-Satellite Measurements" Optoelectronics, Instrumentation and Data Processing. 2023. Vol. 59, no. 5, pp. 602-611.

[15] S. V. Sokolov, V. A. Pogorelov, A. B. Shatalov, "General Solution of the Problem of Nonlinear Filtering of the Orientation Parameters of the Antenna Complex by Inertial Measurements", Russian Aeronautics, 2021, vol.64, no.1, pp. 132-141.

[16] S. V. Sokolov, V. A. Pogorelov, A. B. Shatalov, "Stochastic Estimation of Orientation Parameters of an Antenna Complex based on Strapdown Inertial System Measurements", Measurement Techniques. 2020, vol. 63, no. 2. pp. 87-95.

[17] S. V. Sokolov, V. A. Pogorelov, A. V. Sukhanov, M. V. Kurinenko, "Robust nonlinear filtering of navigation satellite measurements" Telecommunications and Radio Engineering. 2023. vol. 82, no. 2, pp. 1-15.

[18] A. Yu. Ishlinskii, "Orientation, Gyroscopes and Inertial Navigation" Moscow, Nauka, 1976.

[19] O.S Salychev, "Verified Approaches to Inertial Navigation", Moscow, BMSTU Press, 2017.

[20] I.Rusnak, "Optimal State Estimation of Nonlinear Dynamic Systems, Nonlinear Systems - Modeling, Estimation, and Stability", InTech, 2018, pp 159-178.

[21] P. G Savage, "Strapdown Inertial Navigation Integration Algorithm Design Part 1: Attitude Algorithms", Journal of Guidance, Control, and Dynamics, 1998, vol 21(1), pp 19–28.

[22] V. I.Tikhonov, "Statistical Radioengineering", Moscow, Radio I Svyaz, 1982.

[23] D.A. Balakin, A.A. Kalshchikov, E.V. Shalimova, "Spectral analysis of signals in the mathematical environment of Engee, Moscow, Publishing House of the MPEI, 2024.

Analysis of Chaotic Modes of Perturbed Motion of Solids under the Action of Piecewise Continuous Perturbations

A. Chuiko

Department of Flight Dynamics and Control Systems
Samara National Research University
Samara, Russia
tchuikoalisa@yandex.ru

Abstract— **The paper considers the free attitude motion of a gyrostat satellite under the action of periodic piecewise continuous perturbations. The chaotic dynamics of the satellite is analyzed using the analytical Melnikov method. The fact of occurrence of chaotic modes of the gyrostat satellite motion using Poincaré sections is proved. And an algorithm for numerical analysis of chaotic dynamics is also developed.**

Keywords— gyrostat-satellite, separatrix, Melnikov method, Poincaré section, chaotic dynamics.

I. INTRODUCTION

Chaotic modes of motion are among the important and under-researched phenomena in the perturbed dynamics of gyrostat-satellites. These regimes can cause unforeseen situations during space missions or lead to their complete failure. Therefore, it is important to consider such regimes in the design of real space systems.

The main reason for the emergence of chaotic dynamics is the complication of the phase space structure in the neighborhood of unperturbed separatrix. The chaotic behavior of the system can be detected using the Melnikov method [1]. It is based on obtaining a function whose zeros indicate the fact of multiple crossing of split separatrix. The research of a symmetric gyrostat satellite with internal harmonic perturbations using the classical Melnikov method is considered in [2]. In [3,4] a formalism for computing the Melnikov function for multidimensional systems taking into account different types of perturbations was constructed. Melnikov-Wiggins method was applied in [5-7]. For example, a case with small dynamic asymmetry of the rotor is considered in [5]. In [6] the models of the dynamics of angular motion of a spacecraft with a rotor under different types of perturbations and asymmetry are considered in detail. The work [7] is a continuation of [6], an overview of chaotic dynamics is given and the Melnikov-Wiggins formalism is applied. In [8] a modified Melnikov method is used to research the dynamics of nanosatellite orientation with perturbation dissipation.

Also, various numerical methods are used to analyze chaotic dynamics, in addition to analytical study. In [2,5,8], the method of Poincaré sections is used to detect chaotic layers near unperturbed separatrix. In [9], a method for numerical and analytical computation of perturbed separatrix based on the application of Matrizant theory was developed. That allows numerically to construct the decoupled forms of separatrix with accuracy up to the first order of smallness of the perturbation.

In this paper, a multivariate analysis of chaotic modes of motion under the action of internal piecewise continuous perturbations is carried out. Such interference may occur in the internal motor circuit of the system due to errors in the angular velocity sensor. A comparison of approaches to the realization of the Melnikov method when approximating the perturbation by Fourier series and representing it as a piecewise continuous function is given. A numerically-graphical algorithm for determining the chaotisation of the dynamics has developed.

II. MATHEMATICAL MODEL

The free attitude motion of a gyrostat satellite relative to the center of mass consisting of a carrier body and a rotor is considered. The carrier body has a triaxial inertia tensor and the rotor is a dynamically symmetric body. We introduce the following coordinate systems:

$OXYZ$ is an inertial coordinate system; $Ox_1y_1z_1$ the connected principal system of coordinates of the carrier body; $Ox_2y_2z_2$ the connected principal system of coordinates of the rotor. The Oz_1 and Oz_2 axes coincide.

Using Euler's dynamic equations, the motion with respect to the center of mass is described for a system with four degree-of-freedom [2]:

$$\begin{cases} A\dot{p} + (C_2 - B)qr - q\Delta = 0, \\ B\dot{q} + (A - C_2)pr - p\Delta = 0, \\ C\dot{r} + \Delta + (B - A)pq = 0, \\ \dot{\Delta} = M_\Delta. \end{cases} \quad (1)$$

where p, q, r are components of angular velocity of the carrying body in the coordinate system; σ -the rotor angular velocity relative the carrier body; $I_2 = diag[A_2, B_2, C_2]$ - inertia tensors of the supporting body in the connected coordinate system $Ox_2y_2z_2$; $I_1 = diag[A_1, B_1, C_1]$ -rotor inertia tensors in the coupled coordinate system $Ox_1y_1z_1$ $A = A_1 + A_2, B = A_1 + B_2, C = C_1 + C_2$ - are the principal moments of inertia of the system of coaxial bodies in the coordinate system connected with the carrier body; M_Δ - is the internal torque of the coaxial bodies interaction; $\Delta = C_1(r + \sigma)$ the longitudinal angular moment of the rotor.

The satellite-gyrostat is a dynamically symmetric body hence the moments of inertia $A_1 = B_1$. Suppose that there is no interaction between the two bodies, then there exist

explicit analytic heteroclinic solutions [2], which are necessary to apply the classical Melnikov method:

$$\overline{p}(t) = \pm\sqrt{\frac{C_2(B-C_2)}{A(A-B)}}y(t), \overline{q}(t) = \pm\sqrt{s^2 - k^2(y(t)+\Delta\beta)^2},$$

$$\overline{r}(t) = y(t) + \frac{\Delta}{(B-C_2)}, y(t) = \frac{4a_0 E(y_0^\pm)e^{\left(\mp\frac{M\sqrt{a_0}}{k^2}t\right)}}{\left[E(y_0^\pm)e^{\left(\mp\frac{M\sqrt{a_0}}{k^2}t\right)} - a_1\right] - 4a_2 a_0}.$$

With a set of constants that depend only on the initial conditions of motion and moments of inertia:

$$\Delta = const > 0; a_2 = -k^2; a_1 = -2\Delta\beta k^2; a_0 = s^2 - k^2\Delta^2\beta^2;$$

$$y_0^\pm = \pm\frac{s}{k} - \Delta\beta; H = 2T(A-\tilde{D}) + \Delta^2\beta;$$

$$\tilde{D} = \frac{\Delta^2 a}{2T} + B; M = \frac{(A-C_2)}{B}\sqrt{\frac{C_2(B-C)}{A(A-B)}};$$

$$E(y_0^\pm) = \frac{1}{y_0^\pm}(2a_0 + a_1 y_0 + 2\sqrt{a_0}\sqrt{a_2(y_0)^2 a_1 y_0 + a_0});$$

$$a = \frac{C_1 C_2 + (B-C_2)(C_1-B)}{(B-C_2)C_1}; b = \frac{C_1 C_2 + (A-C_2)(C_1-A)}{(A-C_2)C_1};$$

For the purpose of further analysis and numerical modeling in order to apply Melnikov's method [1], we will pass to the Hamiltonian form of writing down the equations in Andoyer–Deprit variables [10]. This will allow the transition from a system with four degrees-of-freedom to a system with one degree-of-freedom, which is necessary to apply the classical Melnikov method. The motion of the system is described by the angles l, φ_2, φ_3 which characterize the rotations with respect to the axis OZ in the directions of the kinetic momentum of the system and the axis $Oz_{1,2}$, respectively, and their generalized impulses L, I_2, I_3. Then the equation of the system in Andoyer–Deprit variables looks as follows:

$$\begin{cases} \dot{L} = -\dfrac{\partial H}{\partial l} = f_L(l,L) + \varepsilon g_L(t), \\ \dot{l} = -\dfrac{\partial H}{\partial L} = f_l(l,L) + \varepsilon g_l(t). \end{cases} \quad (2)$$

with functions

$$f_L(l,L) = -\frac{\partial H_0}{\partial l} = \alpha(I_2^2 - L)\sin l\cos l,$$

$$f_l(l,L) = \frac{\partial H_0}{\partial L} = L\left[\frac{1}{C_1} - \frac{\sin^2 l}{A} - \frac{\cos^2 l}{B}\right] - \frac{\Delta}{C_1},$$

$$g_L = -\frac{\partial H}{\partial l}, g_l = -\frac{\partial H}{\partial L}.$$

where $\alpha = \dfrac{1}{B-A}$.

III. MELNIKOV METHOD

Consider the perturbed motion of the system at the moment when the spinning of the rotor is completed and it has assumed a constant angular velocity and a constant

angular momentum corresponding to it. A perturbing internal moment occurs in the engine control system:

$$M_\Delta = \mu\arcsin(\sin[\nu t]) \quad (3)$$

when μ - small parameter, ν is the frequency of the disturbance.

A. Approximation of the perturbation by Fourier series

Since the perturbation (3) is a piecewise continuous periodic function, it is decomposable in Fourier series. After substituting the perturbation into the last equation of the system (1), the analytical solution for the rotor kinetic momentum is obtained:

$$\Delta = \Delta_C + \frac{\mu}{\nu}\sum_{n=0}^{N}b_n\cos\left[(2n+1)\nu t\right] \quad (4)$$

where $b_n = -\dfrac{4(-1)^n}{\pi(2n+1)^3}$ - Fourier series coefficient, Δ_C is the constant of integration.

The Melnikov function for the considered system with account of the perturbation (4) takes the form:

$$M(t_0) = \varepsilon\nu\alpha\int_{-\infty}^{\infty}(I_2^2 - L^2)\sin l\cos l\sum_{n=0}^{N}b_n\cos\left[(2n+1)\nu(t+t_0)\right]dt =$$

$$= \varepsilon\nu\alpha AB\sum_{n=0}^{N}b_n\left\{J_1^{(n)}\cos\left[(2n+1)\nu t_0\right] - J_2^{(n)}\sin\left[(2n+1)\nu t_0\right]\right\}.$$

where $J_1^{(n)} = \int_{-\infty}^{+\infty}\overline{g}(t)\cos\left[(2n+1)\nu t_0\right], J_2^{(n)} = \int_{-\infty}^{+\infty}\overline{g}(t)\sin\left[(2n+1)\nu t_0\right],$

$$\overline{g}(t) = \overline{p}(t)\overline{q}(t) = \pm\sqrt{\frac{C_2(B-C_2)}{A(A-B)}}\sqrt{s^2 - k^2(y(t)+\Delta\beta)^2}y(t), \varepsilon = \frac{\mu}{C_2\nu^2}.$$

Figure 1 shows the integrals $J_1^{(n)}$ and $J_2^{(n)}$, represented as areas of curvilinear trapezes bounded by integrand functions. The function $\overline{g}(t)$ (Fig. 1a) is odd and decays rapidly to zero value. The integrals $J_1^{(n)}$ as seen in Fig.1b tend to zero, and the integrals $J_2^{(n)}$ (Fig.1c) tend to some constant R_n.

Fig. 1. Improper integrals

Then the Melnikov function takes the next form:

$$M(t_0) = \varepsilon v \alpha A B \sum_{n=0}^{N} R_n b_n \sin\left[(2n+1)v t_0\right] \quad (5)$$

Function (5) has a simple polyharmonic form, which implies the presence of an infinite number of simple zeros. This proves the fact of splitting of separatrix at the presence of a small periodic perturbing internal momentum and hence the appearance of a chaotic layer near separatrix.

B. The representation of the perturbation as a piecewise continuous function

In additional to the analysis of the perturbation by Fourier series approximation, we can consider the perturbation (3) as a piecewise continuous function. And then (3) can be written in the form:

$$M_\Delta = \mu(-1)^{n+1}(\pi n - t v), \frac{\pi}{2v}(2n-1) < t \le \frac{\pi}{2v}(2n+1),$$

where $n \in \mathbb{R}$.

The Melnikov function after transformations looks as follows:

$$M(t_0) = \varepsilon v \alpha A B \left\{ J_1\left(\pi(-1)^{n+1} n - v t_0 (-1)^{n+1}\right) + \right.$$

$$+ J_2\left(\pi t_0 (-1)^{n+1} n - (-1)^{n+1} \frac{t_0^2 v}{2}\right) - J_3 \frac{v}{2}(-1)^{n+1} \right\},$$

$$\frac{\pi}{2v}(2n-1) < t \le \frac{\pi}{2v}(2n+1).$$

where $J_1 = \int_{+\infty}^{-\infty} \bar{g}(t) t \, dt, J_2 = \int_{+\infty}^{-\infty} \bar{g}(t) \, dt, J_3 = \int_{+\infty}^{-\infty} \bar{g}(t) t^2 \, dt.$

The function $\bar{g}(t)$, as it was said earlier, is odd and decays rapidly to zero, hence the integral J_2 tends to zero. The integral J_1 as the area of a curvilinear trapezoid bounded by the integrand tends to some constant R, and the integral J_3 as the area of a curvilinear trapezoid bounded by the integrand tends to zero.

Finally, the Melnikov function assumes the form:

$$M(t_0) = \varepsilon v \alpha A B R\left(\pi(-1)^{n+1} n - v t_0 (-1)^{n+1}\right),$$

$$\frac{\pi}{2v}(2n-1) < t \le \frac{\pi}{2v}(2n+1) \quad (6)$$

where $n \in \mathbb{R}$.

Function (6) has an infinite number of simple zeros, which, as it was said earlier, shows the fact of separatrix splitting and the appearance of the chaotic layer.

IV. NUMERICAL ANALYSIS

In addition to the analytical study, numerical tools are used to detect the chaotic behavior of the system. Numerical integration of system (2) under the action of perturbation of the form (3), was carried out with the given parameters of the system:

$$A_1 = 5 \kappa z \cdot m^2, C_1 = 4 \kappa z \cdot m^2, A_2 = 15 \kappa z \cdot m^2, B_2 = 8 \kappa z \cdot m^2,$$

$$C_2 = 6 \kappa z \cdot m^2, I_2 = 20 \frac{\kappa z \cdot m^2}{c}, \varepsilon = 0,05.$$

The chaotic layer can be illustrated using Poincaré cross sections. Poincaré sections show the phase space of the system at times with integer perturbation period. Comparing the Poincaré sections in the Andoyer–Deprit variable space for unperturbed (Fig. 2) and perturbed (Fig. 3) systems, one can observe the appearance of a chaotic layer near the separatrix.

Fig. 2.   Poincaré section of the unperturbed system

Fig. 3.   Poincaré section of the perturbed system

The chaotic nature of the system, is also illustrated using the perturbed phase trajectory. In Fig. 4 shows a change in the behavior of the phase trajectory near the separatrix in the presence of perturbations, expressed in a chaotic change of phase zones from oscillation to rotation.

Fig. 4.   Phase trajectories, 1 - unperturbed motion, 2 - perturbed motion

To analyze chaotic dynamics, an algorithm is developed to plot the evolution of the phase trajectory near the separatrix in the presence of perturbations. The numerical solution of the system of differential equations (2) in the absence of perturbations is obtained. Each point of the obtained phase trajectory serves as a set of initial conditions for further integration of system (2) with perturbation (3) on the time interval $t \in [0,T]$. After integration in the forward time direction we obtain an unstable manifold of the phase trajectory, and in the reverse direction we obtain a stable manifold. The obtained solutions allow us to determine all possible intersections of the split manifold. Figure 5 shows the multiple intersection of the stable and unstable manifolds of the perturbed trajectory taken near the separatrix.

Fig. 5.   Split phase trajectory, 1 - unstable manifold, 2 - stable manifold

CONCLUSION

In this research the dynamics of gyrostat satellite motion under the action of piecewise continuous perturbations has been considered. The emergence of chaotic modes of motion under the action of internal periodic piecewise continuous perturbations is shown analytically using the Melnikov method. The analysis when describing the perturbations in the form of Fourier series and representing them as piecewise defined functions gave the same result, the presence of zeros of the Melnikov function, but the analysis when decomposing the perturbations into Fourier series is simpler to implement, since the representation of such functions as piecewise continuous functions is not always possible. The chaotic motion was also confirmed numerically using Poincaré cross sections and phase trajectories. A numerical algorithm for the construction of split phase trajectories was developed.

REFERENCES

[1]   Mel'nikov V.K. [On stability of the center under time-periodic perturbations]. Trudy Moskovskogo matematicheskogo obshchestva. 1963, Vol 12, P. 1– 56.

[2]   Doroshin, A.V. Heteroclinic dynamics and attitude motion chaotization of coaxial bodies and dual-spin spacecraft (2012) Communications in Nonlinear Science and Numerical Simulation, 17 (3), pp. 1460-1474.

[3]   Holmes P. J., Marsden J. E. (1983), Horseshoes and Arnold diffusion for Hamiltonian systems on Lie groups, Indiana Univ. Math. J. 32, pp. 273-309.

[4]   Wiggins S. (1988), Global Bifurcations and Chaos : Analytical Methods (Applied mathematical sciences : vol. 73). Springer-Verlag.

[5]   Aslanov, V.S., Doroshin, A.V. Chaotic dynamics of an unbalanced gyrostat (2010) Journal of Applied Mathematics and Mechanics, 74 (5), pp. 524-535.

[6]   Anton V. Doroshin, Heteroclinic chaos and its local suppression in attitude dynamics of an asymmetrical dual-spin spacecraft and gyrostat-satellites. The part I main models and solutions, Communications in Nonlinear Science and Numerical Simulation (2016), 31 (1-3), pp. 151-170

[7]   Anton V. Doroshin, Heteroclinic Chaos and Its Local Suppression in Attitude Dynamics of an Asymmetrical Dual-Spin Spacecraft and Gyrostat-Satellites. The Part II The heteroclinic chaos investigation, Communications in Nonlinear Science and Numerical Simulation (2016), 31 (1-3), pp. 171-196

[8]   Doroshin A.V., Eremenko A.V. (2023) "Heteroclinic chaos detecting in dissipative mechanical systems: Chaotic regimes of compound nanosatellites dynamics". Communications in Nonlinear Science and Numerical Simulation, 2023

[9]   Anton V. Doroshin, Computing the Heteroclinic Orbits Splitting in Systems Phase Spaces via the Matricant Method (2018) Lecture Notes in Engineering and Computer Science Volume 1, 2018, Pages 428-433

[10] Arkhangel'skii Yu.A. Analiticheskaya dinamika tverdogo tela [Analytical dynamics of a solid body]. Moscow, Nauka Publ., 1977, P. 17-22.

The Problem of Intercepting a Group of Targets in Sun Synchronous Orbits of an Artificial Earth Satellite Taking into Account the Second Zonal Harmonic in an Pulse Formulation

A.S. Samokhin
38 laboratory ICS RAS
Moscow, Russia
samokhin@ipu.ru,
ORCID: 0000-0002-0821-050X

M.A. Samokhina
38 laboratory ICS RAS
Moscow, Russia
ph@ipu.ru,
ORCID: 0000-0002-7043-706X

Abstract—**The paper develops a technique for intercepting targets moving along prescribed trajectories in the field of attraction of a single attracting center in space. The problem is considered in a pulse formulation. The interception is performed by a single controlled object. The problem formulation is taken from 9th global trajectory optimization competition. The problem of bypassing a group of targets is considered as a dynamic traveling salesman problem.**

Keywords—flight chain, Lambert problem, dynamic traveling salesman problem, impulsive formulation, Kessler syndrome, trajectory optimization, spacecraft flight optimization

I. INTRODUCTION

The problem of alternating traversal of a set of targets by a group of guided objects – interceptors is currently topical. Controlled objects can be both manned and unmanned. A visit to one target may consist of servicing it, such as making repairs, or delivering cargo to the target, such as refueling. In the space it is also possible to extract resources, collect samples or install special equipment, see [1] for problem statements.

Such tasks involve the development of rather cumbersome missions in which the order of interceptors to targets and the trajectories themselves must be determined, and they can be broken down into subtasks:

1) building an elementary operation – servicing one regular target with one interceptor;

2) planning of individual bypass plans – construction of flight chains for one controlled object;

3) implementation of the entire mission by the interceptor team – planning a set of plans to bypass the entire set of targets.

Various merit functions or ordered vector criteria may be considered in different problem statements, e.g., it may be necessary to visit all targets at least once or to visit the maximum number of targets (it is not always possible to intercept all targets), each target may be necessary to visit at most once or at least twice [1].

This paper investigates the problem from the 9th international Global Trajectory Optimization Competition (GTOC) [2].

It is believed that the safety of low-Earth orbit operation is undermined by the possibility of the Kessler syndrome.

In this case, we consider a set of 123 orbital objects – space debris whose removal will restore the possibility of safe operation in this orbital environment and reliably prevent the Kessler syndrome.

The competition required planning a series of missions capable of removing all specified critical debris in sun-synchronous orbits while minimizing total cumulative cost. The point of servicing the target in the context of the competition was to deliver a special package to the debris, taking it later automatically out of the orbit of an artificial Earth satellite.

The study develops the construction of an elementary operation and flight chain for a single guided object.

II. PROBLEM STATEMENT

The orbits of the 123 debris to be intercepted are given in osculating elements. This makes it possible to calculate the coordinates and velocities of the debris analytically after numerically solving the Kepler equation for any moment in time.

The orbits of the debris themselves are close to circular, and differ little in altitude and inclination, but have significantly different longitudes of the ascending node.

All debris must be intercepted. Any number of spacecrafts can be used to intercept, but it is mostly advantageous to use as few as possible. If some debris is not part of the series of missions built, it is assumed to be eliminated by a separate single mission, which is inefficient according to the task optimization criterion. The maximum fuel reserve of a single spacecraft is limited, and it is impossible for one spacecraft to fly around all the debris.

Each spacecraft is a material point, the equations of motion take into account the 2nd zonal harmonic, responsible for the non-sphericity of the Earth.

The system of differential equations describing the passive motion of the spacecraft between control actions is as follows:

$$\ddot{x} = -\frac{\mu x}{r^3}\left\{1 + \frac{3}{2}J_2\left(\frac{r_{eq}}{r}\right)^2\left(1 - 5\frac{z^2}{r^2}\right)\right\},$$

$$\ddot{y} = -\frac{\mu y}{r^3}\left\{1 + \frac{3}{2}J_2\left(\frac{r_{eq}}{r}\right)^2\left(1 - 5\frac{z^2}{r^2}\right)\right\},$$

$$\ddot{z} = -\frac{\mu z}{r^3}\left\{1 + \frac{3}{2}J_2\left(\frac{r_{eq}}{r}\right)^2\left(3 - 5\frac{z^2}{r^2}\right)\right\},$$

where x, y, z – spacecraft coordinates, r – distance to the Earth's center, μ – gravitational parameter of the Earth, r_{eq} – average equatorial radius of the Earth, J_2 – given coefficient of the second zonal harmonic of the Earth's gravitational potential.

The spacecraft is controlled by pulse actions, after each pulse the mass of the spacecraft changes according to Tsiolkovsky's formula.

The merit function of the problem J :

$$J = \sum_{i=1}^{n}\left[c + \alpha\left(m_{0_i} - m_{dry}\right)^2\right]$$

reflects the final cumulative cost of all missions and is subject to a fixed multiplier $\alpha = 2.0 \cdot 10^{-6}\dfrac{MEUR}{kg^2}$, the cost of mass of each i-th spacecraft for the mission $m_{0_i} - m_{dry}$, the number of the spacecrafts n and the bonus coefficient c, which responds for the time of finding a solution to the problem.

A minimum of 5 days must be spent on each debris to install the special package. After leaving one debris, the spacecraft must fly to the next one within 30 days. The total mission duration is limited to 8 years.

The current research considers the construction of chains of consecutive flights to the debris by a single spacecraft.

III. SOLUTION METHOD

The described problem is NP-hard and it was considered as a dynamic traveling salesman problem (MT-TSP) [3]. Finding the trajectory between two debris was chosen as an elementary operation, which was determined in 2 stages:

1) Fast calculation of the transition between two debris in a simplified formulation in the central Newtonian force field.

2) Trajectory refinement by the shooting method taking into account the 2nd zonal harmonic using the modified Newton method and the Runge-Kutta method of order 8 based on the Dorman-Prince calculation formulas with automatic step selection. External optimization based on the gradient method can be used to refine the start and finish moments of time.

The first stage is necessary to gain an understanding of which debris is feasible to fly between at all, since highly energy disadvantageous flights will not yield an acceptable aggregate solution.

Two possibilities for implementing the stage 1 maneuver were considered in this work. First, a discrete enumeration of start and finish points in the orbits of the debris, between

Fig. 1. Schematic of a 3-pulse maneuver. Here o_s – start orbit, o_f – finish orbit, l – line of nodes, Δ_0 – first maneuver impulse, taking the spacecraft from the orbit of the first debris, Δ_1 – second impulse, turning angle of the orbit plane, Δ_2 – third impulse, equalizing the spacecraft speed with the speed of the second debris, two elliptical pieces of the flight path between the debris are shown in red and orange.

which the flight was planned, was performed with a given time step.

Then, based on the obtained points, the corresponding Lambert problems were formed and solved on the basis of the universal Kepler equation to obtain two-pulse trajectories. This approach turns out to be justified when the transition is close to coplanar. And, secondly, the construction of a three-pulse transition with an additional impulse on the line of nodes for orbital rotation, schematically depicted in Fig. 1.

Lambert problems are convenient because they have solutions for any correct set of parameters, their solving can be reduced to finding the root of a nonlinear function from one variable, and do not produce any computational difficulties. In this paper, the Lambert problems were solved numerically by a modified Newton method based on the universal Kepler equation [4].

The first stage of the elementary operation construction was used to construct the flight chains, while the second stage was used to refine the resulting trajectories so that they satisfy the complicated mathematical model of the original problem.

MT-TSP problem of the intercept chain construction was solved by beam search, which is some combination of depth and width search, as it was necessary to reduce the search considerably. It is not possible to solve the problem by a complete search of all possible combinations in a reasonable time.

For the selected first debris, a tree of possible further flights was built based on the calculation of the corresponding elementary operations, but only a limited number of branches promising in terms of local and total characteristic costs of flights were unfolded.

The tree is completed when the spacecraft's fuel supply is depleted. After that, the elementary operations that make up the flight chain are recalculated in the original more accurate mathematical model.

IV. RESULTS

The paper develops a method of elementary operation construction in the conditions of the considered problem formulation and flight chain for one spacecraft.

At first, elementary operations are calculated as fast as possible in a simplified mathematical model for performing the global optimization. After that the beam search is used to reduce the enumeration.

As a result, for each launching debris, promising flight chains are calculated in terms of the cost of characteristic speed, which are further recalculated taking into account the second zonal harmonic.

The authors have implemented the program complex in C language for the problem solving.

V. FURTHER INVESTIGATIONS

The development of the study is the construction of a database of successful target overflight chains from which a series of single missions can be developed to eliminate all 123 dangerous debris. This requires combining successful chains of debris overflights by different spacecrafts so that none are visited twice, but all are visited at least once.

One of the options could be a sequential application of the algorithm developed in this paper with consideration among the set of targets each time only those objects that are not yet involved in the already constructed chains.

Also, the idea of exchanging debris between spacecraft is promising, for example, it can be realized on the basis of a genetic algorithm, in order to find the best functional solution to the problem. In this case, it is possible to remove duplicate objects from the constructed chains and add to the already constructed chains inconvenient debris that have not been visited in the course of the solution.

REFERENCES

[1] URL: http://sophia.estec.esa.int/gtoc_portal/

[2] D. Izzo, M. Maertens, "The Kessler Run: On the Design of the GTOC9 Challenge," Acta futura, vol. 11, 2018, pp. 11–24. Doi: 10.5281/zenodo.1139022.

[3] C.S. Helvig, G. Robins, A. Zelikovsky, "Moving target TSP," Lecture Notes in Comput. Sci., vol. 1461, 1998, pp. 453–464.

[4] A.S. Samokhin, M.A. Samokhina, "Verification of the second-order optimality conditions in the modeling of the SC expedition with the returning to the Earth based on two Lambert's problems solving," Advances in the Astronautical sciences, vol. 161, 2018, pp. 843–862.

A Circular Perimeter Breakthrough in a Threat Environment

M.A.Samokhina

38 laboratory ICS RAS

Moscow, Russia

ph@ipu.ru, ORCID: 0000-0002-7043-706X

A.S. Samokhin

38 laboratory ICS RAS

Moscow, Russia

samokhin@ipu.ru, ORCID: 0000-0002-0821-050X

Abstract—**The paper considers the problem of classifying the evasive detection paths of a controlled object in a threat environment. The threat environment is represented by several stationary sensors. The goal of the object is to get to the perimeter defined by a circle as stealthily as possible. The problem is formalized as an optimal control problem. The Pontryagin extremals are constructed numerically by the shooting method.**

Keywords—threat environment, detection risk functional, trajectory optimization, optimal control, perimeter breach.

I. Introduction

The relevance of the work is related to the recent widespread use of autonomous vehicles in the tasks of breaching the protected perimeter, the continuing development of this field and the corresponding improvement of security systems.

We consider a conflict environment [1] represented by fixed sensors (detectors) and a moving controlled mobile object (CMO).

It is assumed that the location of the sensors, which form a threat map, is known to the CMO [2]. The evasive mobile object route and movement parameters are chosen in such a way as to maximize the stealth of movement, so the CMO minimizes the integral functional of the detection risk [3-4].

The problem in a continuous formulation for CMO transition between given points is considered in [5], the idea of its discretization is presented in [6]. The inverse problem of deploying sensors to counteract the covert movement of the CMO is investigated in [7]-[8].

II. Problem Statement

The formulation is a development of the problem considered in [5].

There are some known number $N \geq 1$ of fixed sensors – points S_i on the plane. The detection fields of the sensors are considered circular. The problem consists in controlling the CMO, which is a material point that must move at a distance not less than a given distance $R < \infty$ from its known initial position for a fixed time $T < \infty$, the motion scheme is shown in Fig. 1 (a)-(b).

We consider a motion model in which the CMO is controlled by the modulus and direction of the velocity vector. The material point is controlled so as to minimize the integral functional of the detection risk:

$$\int_0^T \left(\sum_{i=1}^N q_i \frac{v^2(t)}{r_i^2(t)} \right) dt \to \min , \qquad (1)$$

where $v(t)$ – modulus of CMO velocity vector, $r_i(t)$ – distances to points S_i at the moment of time t; values $q_i > 0$ – weight coefficients of points S_i influence. The location of the points S_i and their corresponding values q_i are given and known to the CMO.

We introduce a Cartesian coordinate system (CS) in the plane so that its origin coincides with the CMO start point. Let the points – sensors S_i have coordinates (a_i,b_i), $i=1,\ldots,N$ in the introduced coordinate system.

Then the motion of the CMO can be described by the system:

$$\begin{cases} \dot{x} = v \cos \varphi, \\ \dot{y} = v \sin \varphi, \end{cases} \qquad (2)$$

where (x, y) – CMO coordinates in the introduced CS.

The problem is considered in two formulations. In the first one the control angle $\varphi(t)$ in (2) is selected from the half-interval $[0, 2\pi)$ at any given time $t \in [0,T]$. And in the second one the angle $\varphi(t)$ is chosen fixed on the whole trajectory, which means that the angle $\varphi(t) = const \quad \forall t \in [0,T]$ on each trajectory, and the motion is along a straight line.

The velocity of the material point is bounded:

$$0 \leq v \leq v_{\max} . \qquad (3)$$

At some point of time no later than T the CMO must reach the boundary of the region – the perimeter, which is a circle of radius R centered at the origin. The point at which the perimeter is reached is optimized.

III. Solving Method

Both problems are formalized as optimal control problems and L.S. Pontryagin's maximum principle is applied to them. It is assumed that at no moment of time the coordinates of the material point coincide with the coordinates of the sensors.

Conjugate variables p_x, p_y are introduced. The system of necessary optimality conditions is written out and analyzed.

a) CMO trajectories, first formulation.

b) CMO trajectories, second formulation.

c) CMO velocity, trajectories, second formulation.

Fig. 1. Solutions of the problem of path classification in a threat environment. In figures a), b), a black square marks point where the CMO movement starts, a black circle marks the circular perimeter, red points correspond to sensors, and colored curves show the trajectories of the CMO movement. The velocity graphs in figure c) correspond to the trajectories in figure b).

For the first formulation with a variable $\varphi(t)$ on the trajectory, we obtain the following boundary value problem

$$\begin{cases} \dot{x} = \hat{v}\cos\hat{\varphi}, \\ \dot{y} = \hat{v}\sin\hat{\varphi}, \\ \dot{p}_x = -\sum_{i=1}^{N} q_i \dfrac{\hat{v}^2}{\left((x-a_i)^2+(y-b_i)^2\right)^2}\cdot(x-a_i), \\ \dot{p}_y = -\sum_{i=1}^{N} q_i \dfrac{\hat{v}^2}{\left((x-a_i)^2+(y-b_i)^2\right)^2}\cdot(y-b_i), \end{cases} \quad (4)$$

where:

$$\hat{v} = \min\left(\dfrac{\|p\|}{\sum_{i=1}^{N}\dfrac{q_i}{\left(x-a_i\right)^2+\left(y-b_i\right)^2}}, v_{\max}\right), \quad (5)$$

control $\hat{\varphi}$ is determined from the relations $\cos\hat{\varphi} = \dfrac{p_x}{\|p\|}$, $\sin\hat{\varphi} = \dfrac{p_y}{\|p\|}$, and $\|p\| = \sqrt{p_x^2 + p_y^2} \neq 0$.
Boundary conditions:

$$\begin{cases} x(0) = 0, \\ y(0) = 0, \end{cases} \qquad \begin{cases} x(T)p_y(T) - y(T)p_x(T) = 0, \\ x^2(T) + y^2(T) = R^2. \end{cases} \quad (6)$$

To investigate the second formulation with a constant $\varphi(t)$ on the trajectory, we pass to the polar CS. Let us denote the distance from the origin to the CMO by r . Then the boundary value problem has the following form:

$$\begin{cases} \dot{r} = \hat{v}, \\ \dot{p}_r = -\sum_{i=1}^{N} q_i \dfrac{\hat{v}^2(r - d_i\cos(\hat{\varphi}-\alpha_i))}{\left(r^2 - 2rd_i\cos(\hat{\varphi}-\alpha_i) + d_i^2\right)^2}, \\ \dot{w} = \sum_{i=1}^{N} q_i \dfrac{\hat{v}^2 r\sin(\hat{\varphi}-\alpha_i)}{\left(r^2 - 2rd_i\cos(\hat{\varphi}-\alpha_i) + d_i^2\right)^2}, \end{cases} \quad (7)$$

where d_i – distance from the origin to the i -th sensor, α_i – angular position of i -th sensor,

$$\hat{v} = \min\left(\dfrac{p_r}{\sum_{i=1}^{N}\dfrac{q_i}{r^2 - 2rd_i\cos(\varphi-\alpha_i) + d_i^2}}, v_{\max}\right), \quad (8)$$

moreover, new conjugate variable $p_r \geqslant 0$. Boundary conditions:

$$\begin{cases} r(0) = 0, \\ w(0) = 0, \end{cases} \qquad \begin{cases} r(T) = R, \\ w(T) = 0. \end{cases} \qquad (9)$$

The variable w is introduced in (7), (9) to determine the value of the unknown parameter $\hat{\varphi}$, which is found from the integral condition $w(T) = 0$:

$$\frac{\partial \mathcal{L}}{\partial \varphi} = 0 \Leftrightarrow \int_0^T \left(\sum_{i=1}^N q_i \frac{\hat{v}^2 r \sin(\hat{\varphi} - \alpha_i)}{\left(r^2 - 2rd_i \cos(\hat{\varphi} - \alpha_i) + d_i^2 \right)^2} \right) dt = 0.$$

The obtained boundary value problems of the maximum principle (4)-(6) and (7)-(9) are solved by the shooting method with 2 shooting parameters. The series of arising Cauchy problems are solved numerically by the explicit Runge-Kutta method with automatic step selection.

The problem under consideration is multi-extremal. In order to obtain all possible trajectories, the shooting parameters are enumerated on a grid from the set of their admissible values with successive chopping.

IV. Results

The calculations were performed for both problem formulations with different number of sensors. As a result of numerical modeling, maps of optimal paths for different sensor locations are constructed, and then the best in terms of functionality are selected among them.

Fig. 1 shows the results for solving the problem of breaking through a circular perimeter of radius $R = 1$ under the restrictions $v_{max} = 0.5$, $T = 10$ for the case of 10 sensors located at the points: (0.1, 0.3), (0.2, -0.6), (0.7, 0.7), (0.3, 0.2), (-0.7, -0.3) (-0.4, 0.5), (-0.3, -0.9), (-0.2, 0.9), (-0.3, -0.3), (-0.8, 0.0) with equal impacts $q_i = 1, i = 1..10$.

Bolder trajectories correspond to solutions with better value of the functional (1). It may be noted that on the best trajectory in Fig. 1 b) the speed increases monotonically. At the same time, on many other trajectories, the CMO slows down as it approaches the sensors and accelerates as it moves away from them.

In Fig. 1 a) one can see the trajectories crossing the perimeter. The following statement is true.

Theorem. The trajectories $(\hat{x}, \hat{y}, \hat{v}, \hat{\varphi})$ on which the CMO crosses the perimeter and does not complete its movement yet are not optimal.

Proof. Let the first intersection of the perimeter occur at the moment of time $\tau < T$: $r(\tau) = R$. Consider the CMO control coinciding with the $(\hat{v}, \hat{\varphi})$ on $[0, \tau]$, and then put $v = 0$, φ – arbitrary on $(\tau, T]$, i.e. the CMO stands still just after reaching the perimeter. Then such a controlled process satisfies constraints (2)-(3), so it is admissible in the original optimal control problem, but the corresponding functional (1) value better (less) on the new constructed trajectory.

V. Conclusion

The investigated problem of optimal control of a mobile object in a threat environment has been solved in two formulations. The authors have implemented the corresponding program complex in C language. The paper presents Pontryagin extremals for the problem of circular perimeter breakthrough for the case of 10 sensors.

The conducted research admits development. It is possible to consider the problem in which the perimeter is a curve rather than a circle, as well as the problem in which the boundary of the perimeter is not stationary in time. Also, it is possible to take into account the restriction on the path length, to add moving detectors, and to search for analytical solutions for special cases of the problem.

References

[1] T. G. Abramyants, A. A. Galyaev, E. P. Maslov, I. M. Rudko, V. P. Yakhno, "Evasion of a moving object in a conflict environment from detection by the system of heterogeneous observers," Management and navigation of moving objects, no. 2, pp. 31–37, 2015.

[2] A. Dogan, U. Zengin, "Unmanned Aerial Vehicle Dynamic-Target Pursuit by Using Probabilistic Threat Exposure Map," Journal of Guidance, Control and Dynamics, vol. 29, no. 4, pp. 723–732, 2006, doi: 10.2514/1.18386.

[3] L. P. Sysoev, "Kriteriy veroyatnosti obnaruzheniya na traektorii v zadache upravleniya dvizheniem obyekta v konfliktnoy srede [The probability criterion for detection on the trajectory in the task of controlling the movement of an object in a conflict environment]," Problemy upravleniya – Control Sciences, no. 6, pp. 64-70, 2010.

[4] A. A. Galyaev, E. P. Maslov, "Optimization of laws of mobile object evasion from detection," Izv. RAS. Theory and Control Systems, no. 4, pp. 52–62, 2010, doi: 10.1134/S1064230710040064.

[5] M. A. Samokhina, A. A. Galyaev, "Constructing a Map of Locally Optimal Paths for a Controlled Moving Object in a Threat Environment," Control Sciences, no. 1, pp. 75–85, 2024, doi: 10.25728/cs.2024.1.8.

[6] A. A. Galyaev, A. S. Samokhin, M. A. Samokhina, "O reshenii zadachi postroeniya traektorii upravlyaemogo podvizhnogo obyekta v konfliktnoj srede pri pomoshchi diskretizacii [On solving the problem of constructing the trajectory of a controlled moving object in a conflict environment using discretization]," Matematicheskaya teoriya upravleniya i ee prilozheniya: MTUIP-2022. Materialy 15-oj multikonferencii konferencii po problemam upravleniya – Mathematical Control Theory and Its Applications: ICINS-2022. Proceedings of the 15th multi-conference conference on management problems, St. Petersburg, pp. 63–65, 2022.

[7] A. A. Galyaev, A. S. Samokhin, M. A. Samokhina, "On the problem of optimal placement of sensors on the plane," Mathematical control theory and its applications Conference proceedings, St. Petersburg, Concern CSRI Elektropribor, JSC, pp. 372–375, 2020, doi: 10.1088/1742-6596/1864/1/012075.

[8] A. Galyaev, A. Samokhin, M. Samokhina, "Application of the gradient projection method to the problem of sensors arrangement for counteraction to the evasive object," 28th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), Saint Petersburg, Russia, pp. 1–3, 2021, doi: 10.23919/ICINS43216.2021.9470857.

Application of Artificial Intelligence to Prevent Aircraft Collisions on the Airfield Surface

V.I. Baburov
JSC Navigator
St. Petersburg, Russia
info@navigat.ru

S.V. Baburov
JSC Navigator
St. Petersburg, Russia
info@navigat.ru

N.V. Ivantsevich
JSC Navigator
St. Petersburg, Russia
info@navigat.ru,
Baltic State Technical University
VOENMEH
St. Petersburg, Russia
sciencebstu@spb.su

K.V. Koshelev
JSC Navigator
St. Petersburg, Russia
info@navigat.ru

V.V. Khudoshin
JSC Navigator
St. Petersburg, Russia
info@navigat.ru

Abstract—**The paper describes traffic observation at the airfield for the aircraft collision prevention system. Application of neural network is shown to be effective for aircraft collision prevention during taking off, landing, and moving on the runway and taxiway.**

Key words — algorithm, collision prevention, airfield surface, neural network

I. INTRODUCTION

Convenient and safe piloting of the modern aircraft strongly depends on the performance of avionics and the onboard software. Now, advanced electronics allows designing high-performance multifunctional radioelectronic systems.

An example of this avionics system is a Russian airborne collision avoidance system [1]. It is intended to enhance the flight safety by avoiding the midair collision risk, warning on ground proximity and reactive windshear, improving the crew air situation awareness, and to provide safe landing and driving on the airfield surface.

The appearance of the collision prevention system is shown in Fig. 1.

Under highly intensive traffic, the collision risk during taking off, landing, driving on the runway and taxiway is increased, even in international airports equipped with the required air traffic monitoring and control systems. A tragic incident in Tokyo, Japan, in January 2024 [2] highlights the significance of using additional systems to warn the crew on threats on the airfield surface.

Fig. 1.   Airborne collision avoidance system kit.

Figure 2 shows the map of the airport, where the accident took place.

Fig. 2 Scheme of collision in Haneda airport.

To avoid these accidents, it is proposed to use the systems warning the pilots on hazards on the airfield surface, such as Surface Indications and Alert System (SURF IA) within the Automatic Dependent Surveillance Broadcast (ADS-B) technology [3].

Figure 3 shows the navigation screen with the active SURF IA. The screen displays a map of the airfield with runways and taxiways. The relative positions of aircraft are shown with graphic symbols with color gradation depending on the degree of danger to the own aircraft. If a dangerous situation occurs, an information message is displayed.

Fig. 3.   Cockpit screen with active SURF IA.

II.   Problem statement

The input data for SURF IA are the position and velocity of own aircraft and all other air vehicles at the airport, and the airport map. Based on these data the algorithm generates several types of messages alerting the pilots on possible collision with the other aircraft. The messages are displayed on the indicator screen in the cockpit.

It should be mentioned that the message generation mode depends on the airport topology. In Fig. 4, the own aircraft is shown with the blue symbol, and the intruder, with the red one.

Fig. 4.   Mutual positions of the air vehicles.

Two cases of mutual positions of air vehicles moving at the airfield are shown. In both cases, **A** and **B,** own aircraft moves along the runway and intruder moves along the taxiway with the same directions and speeds. The difference is that in case **B** the taxiway does not merge with the runway. In accordance with SURF IA logic, in case **A** the pilots should see an alert message on the collision threat, and in case **B**, should not, because in this case the mutual positions of the runway and taxiway excludes the possibility of aircraft physical contact.

III.   Algorithm implementation

Analyzing the requirements to the application algorithm [3] suggests that two algorithmic procedures should be developed to provide its good performance. The first procedure uses the point coordinates on the airfield surface as an input and outputs the number of an object in the airfield area the point belongs to. By the objects we mean the runways and taxiways. The input data for the second procedure are the vector describing the geometrical segment in the airfield plane

and the number of object in the airport area. The procedure outputs the answer to the question if all the segment points belong to the particular object. The second procedure reduces to the first one, so it would be sufficient to describe the implementation of the first procedure only.

The procedure algorithm can be implemented using a neural network consisting of six neurons shown in Fig. 5. Input neurons have two inputs corresponding to Cartesian coordinates on the plane. Hidden neurons have three inputs. Below we will denote each artificial neuron with the number of its parameters. For example, the output neuron of the network shown in Fig. 5 is assigned number 3 (two inputs plus the activation coefficient).

Fig. 5.   Neural network configuration.

The specific configuration of the neural network is written as a string of numbers (each being the number of parameters for particular neuron), written layerwise from the input layer to the output layer. Thus, the network shown in Fig. 5 is denoted as {3,3,3,4,4,3}.

We chose the hyperbolic tangent as an activation function. For example, the signal of the output neuron is determined by (1):

$$th(w_0 + w_1 x_1 + w_2 x_2), \qquad (1)$$

where x_1, x_2 are the input signals, and w_0, w_1, w_2 are the weight coefficients of the artificial neuron.

To detail the neural network parameters optimization, we introduce the airfield model shown in Fig. 6. The model contains three objects, and the neural network should be able to recognize the airfield field itself in addition to these three objects.

Fig. 6.   Airfield model containing three objects.

The problem is therefore formulated as follows. For the airfield map in Fig. 6, it is needed to find a configuration of the neural network and optimize its parameters so that it accepts the Cartesian coordinates of the point on the airport

surface as an input and outputs the number of the object the point belongs to.

For convenience, we divide the problem into two parts, namely, we design neural networks of the first and second types. The mission of the first type network is to produce the number of an object (1, 2, or 3) on the airfield surface from the inputted coordinates of the point belonging to one of these objects. The second type network answers the question if the point belongs to one of the objects 1, 2, or 3, or to object 4. In other words, it provides a yes or no answer to the question if the point with particular coordinates belongs to one of the circles. It can be easily seen that both networks jointly solve the common problem formulated earlier. First, we ask the second type network if the point belongs to any circle, and if the answer is positive, we input the point coordinates to the first type network, which, in its turn, answers the question, which particular circle (object 1, 2, or 3) the point belongs to.

To create the artificial neural network of the first and second types we wrote scripts in the software environment for statistical computing R [4]. The network parameters were optimized with the function optim(…, method = "BFGS"), included in one of R standard libraries. The graphic images generated by the optimized neural networks have also been designed in R. The network training is performed as follows. The airport area is uniformly covered with the grid of points. Each point belongs to some airport object. Thus, a set of examples for training the artificial neural networks is formed. An example of product of optimized first type network is shown in Fig. 7. Interestingly, the efficiency (probability) of this network is 100%.

Fig. 7.   Airfield model built by neural network of the first type.

The parameters of the neural network are optimized by minimizing some function. These functions are different for the networks of the first and second types. For the first type network, the function to be minimized looks like (2):

$$F(\{w\}) = \sum_{\{x,y\}}(NN - c_1)^2 + (NN - c_2)^2 + (NN - c_3)^2, \qquad (2)$$

where $\{w\}$ is a set of network parameters to be optimized, NN is the network output at a certain point on the plane, coefficients c_1, c_2, c_3 being (-0.5, 0.0, 0.5), respectively, are the centers of attraction for the points belonging to the objects. The sum refers to the coordinates of all points of the grid covering the surface of the airport, with only the points belonging to the object contributing to the term of the attraction center. The diagram shown in Fig. 6 is constructed for all the plane points based on the proximity of the optimized network output to the particular attraction point. In other words, if the output is within $\epsilon =$

0.2 vicinity of particular attraction center, the point is considered to belong to the object.

The function for optimizing the second type network is constructed in similar way and has two centers of attraction (-1, 1) for the plane points belonging and non-belonging to the circles, respectively. Thus, all the plane points providing negative signal at the network output are considered to belong to the circles. The results provided by the optimized second type networks are demonstrated in Figs. 8, 9, and 10.

Fig. 8.   Airfield model constructed by the second type neural network with configuration {3,3,3,4}.

Fig. 9.   Airfield model constructed by the second type neural network with configuration {3,3,3,3,3,3,3,3,3,3,3,13}.

Fig. 10.   Airfield model constructed by the second type neural network with configuration {3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,26}.

Efficiencies of the neural networks with these configurations are 84%, 90%, and 94%, respectively. Generally, the efficiency of optimization is determined by the probability that the point with random coordinates on the plane falls within the specified area. For the first type network, the efficiency is the limiting ratio of the points on the plane correctly identified by the network to all the points falling within the circles. For the second type network, the efficiency is the limiting ratio of the correctly identified points to all the points.

It can be easily seen that that the functional of the second type neural network can be implemented with a bit map, where ones correspond, for example, to the airfield objects, and zeros, to the rest part. The option is selected in the product based on the resource consumption.

IV. CONCLUSIONS

The function realization in the form of a trained artificial neural network provides higher speed of subroutine performance. Additional advantage is an economy of machine memory, because a set of network coefficients occupies less memory space than aviation database containing the airport maps and control tools.

It is proposed to build artificial neural networks for all the airports, that is, about ten thousand totally. The database containing all the required topographic data on the airports should be provided for network training. The topographic data in the database should be updated. In case of changes in the database, the neural networks should be retrained with account for the relevant corrections. The software producer is responsible for the database relevance and regular updates.

Recommendations of SURF IA are not mandatory for the pilots. The messages generated by the application are information messages, unlike the instructions provided by Traffic Collision Avoidance System (TCAS), and don't exclude the need for the pilots to observe the air situation using the standard observation tools.

Storing the map information in the form of neural network coefficients corresponds to NoSQL style of database management. In other words, this database is not a relational one. It produces both pros and cons. This data store will produce a faster response to query, but will require more efforts to enter or correct the data as compared to relational database.

REFERENCES

[1] V.I. Baburov., N.V. Ivantsevich, and V.V. Khudoshin, "Operation of an aircraft collision avoidance system for relative navigation", Proceedings of the 30th Saint Petersburg International Conference on Integrated Navigation Systems, 2023.

[2] https://www.bnnbloomberg.ca/jal-flight-516-was-cleared-to-land-before-fiery-tokyo-collision-1.2017156.

[3] RTCA DO-323. Safety, Performance and Interoperability Requirements Document for Enhanced Traffic Situational Awareness on the Airport Surface with Indications and Alerts (SURF IA). Washington: RTCA, 2010. 580 p.

[4] www.r-project.org.

Estimation Of Characteristics For Different GNSS Augmentation Systems

I. V. Kotov
IAC PNT
JSC «TsNIIMash»
Korolev, Russia
KotovIV@tsniimash.ru

A. A. Arzhannikov
IAC PNT
JSC «TsNIIMash»
Korolev, Russia
AA.Arzhannikov@glonass-iac.ru

V. D. Glotov
IAC PNT
JSC «TsNIIMash»
Korolev, Russia
GlotovVD @tsniimash.ru

The modern GNSS consumer requires more accurate and reliable navigation. To meet such needs, there are several different solutions, one of which is Space Based Augmentation Systems (SBAS). The purpose of this work is an estimation of characteristics for the following currently operating SBAS systems: ANGA (Africa - Europe), BDSBAS (China), EGNOS (Europe), GAGAN (India), MSAS (Japan), SouthPAN (Australia – New Zealand), UKSBAS (Great Britain), WAAS (USA).

I. INTRODUCTION

Calculations were made on the selected daily interval on December 15, 2023 in accordance with the documentation [1]. A similar work on the assessment of the System for Differential Corrections and Monitoring (SDCM) was previously published by the IAC PNT [3]. The current study is a continuation of the started work with more systems. The IAC PNT processes arrays of SBAS corrections in SISNeT (Signal-in-Space through Internet) format. On its basis files with ephemeris and clock offsets, and ionosphere maps of Total Electron Content (TEC) in standard GNSS format files (SP3 and IONEX) are generated. Using these SP3 files, spatial distribution maps of visible space vehicles (SV) number, position and horizontal Dilution of Precision (PDOP and HDOP) over the daily interval were estimated. Based on the calculated IONEX files, SBAS ionosphere accuracy was evaluated by pairwise comparison method using global ionosphere maps (GIM) from IAC PNT [2, 4] and four other analysis centers (CODE, ESOC, JPL, WHU). Also, as a reference the accuracy of the GPS broadcasted Klobuchar ionosphere model was estimated. Since SBAS coverage areas are different, for the ionosphere assessment common zones from 40 to 50 degrees north latitude were selected for seven SBAS systems and for SouthPan (Australia) from 40 to 50 degrees south latitude.

II. RESEARCH RESULTS

Based on the research results, it was found that the considered SBAS systems augment 30-32 GPS space vehicles. The average simultaneously broadcasted augmented SV number was 16-18, the average observed SV number was 9-10, the average PDOP value was 1.6-1.7, the average HDOP value was 0.8-0.9.

The RMS pairwise comparison results of the SBAS ionosphere models and global ionosphere maps are presented in Tables 1, 2 in TEC units (1 TECu is 10^{16} electrons/m^2, which corresponds to 0.16 m for the L1 navigation signal). Tables shows three values: RMS value relative to the IAC PNT ionosphere map, the minimum and maximum RMS values relative to one of ionosphere maps from different analysis centers (IAC PNT, CODE, ESOC, JPL, WHU).

TABLE I. RMS PAIRWISE VALUES OF THE SBAS IONOSPHERE MAPS RELATIVE TO GLOBAL IONOSPHERE MAPS (IN TECU), NORTH

Type	Map	North (from +40 to +50 degrees)		
		IAC PNT	*MIN*	*MAX*
GIM	CODE	1,95	1,57	2,63
	ESOC	1,63	1,63	2,27
	IAC PNT		1,63	2,16
	JPL	1,94	1,79	2,45
	WHU	2,04	1,57	2,73
SBAS	ANGA	2,16	2,16	2,73
	BDSBAS	2,20	2,15	2,43
	EGNOS	1,48	1,48	2,10
	GAGAN	2,80	2,80	4,02
	MSAS	3,20	3,10	3,54
	UKSBAS	3,55	3,49	4,14
	WAAS	2,11	2,11	2,80
Klobuchar	GPS	7,16	6,71	7,70

TABLE II. RMS PAIRWISE VALUES OF THE SBAS IONOSPHERE MAPS RELATIVE TO GLOBAL IONOSPHERE MAPS (IN TECU), SOUTH

Type	Map	South (from -50 to -40 degrees		
		IAC PNT	*MIN*	*MAX*
GIM	CODE	3,57	2,02	4,00
	ESOC	3,61	2,02	3,82
	IAC PNT		3,57	4,44
	JPL	3,71	2,86	3,71
	WHU	3,67	2,24	4,35
SBAS	SouthPAN	4,44	3,48	4,44
Klobuchar	GPS	6,57	5,36	6,57

III. CONCLUSION

As noted earlier, SBAS systems differ from each other and have different coverage areas. The above assessment is limited and requires further additions. The comparison results show that the SBAS ionosphere maps accuracy is comparable to the accuracy of global ionosphere maps made by the IAC PNT and other analysis centers, and is significantly better than the GPS broadcasted Klobuchar

model. On considered latitudes the value of global ionosphere maps RMS ranged from 1.6 to 2.7 TECu (from 0.25 to 0.43 m) in the northern hemisphere and from 2.0 to 4.4 TECu (0.32 to 0.71 m) in the southern. The RMS value of seven SBAS systems in the northern hemisphere ranged from 1.5 to 4.1 TECu (0.24 to 0.66 m) and for SouthPan (Australia) the range was from 3.5 to 4.4 TECu (0.55 to 0.71 m) in the southern. For comparison, the RMS value of GPS broadcasted Klobuchar ionosphere model ranged from 5.3 to 7.7 TECu (0.85 to 1.23 m).

Accomplished work allows us to assert that the selected SBAS systems parameters are within acceptable values. Obtained results open the way for further research on SBAS systems.

REFERENCES

[1] Minimum operational performance standards for Global Positioning System / Wide Area Augmentation System airborne equipment. RTCA/DO-229C. November 28, 2001

[2] Section "Ionosphere" on the website of the IAC KVNO JSC TsNIIMash [Electronic resource]. URL: https://www.glonass-iac.ru/iono/ (date: 01.02.2024).

[3] I.V. Kotov, A.A. Arzhannikov, V.D. Glotov. Characteristics estimation of the GLONASS augmentation system // Cosmonautics and rocket engineering. 2023. № 4 (133). P. 67–69. URL: https://glonass-iac.ru/about/publications/detail.php?ID=4104 (date: 01.02.2024).

[4] A. A. Arzhannikov, V. D. Glotov, V. V. Mitrikas. Estimation of Differential Code Biases and Ionosphere Maps Using GNSS // Transactions of IAA RAS. — 2022. — Issue 60. — P. 3–11. URL: https://glonass-iac.ru/about/publications/detail.php?ID=4023 (date: 01.02.2024).

Current IAC PNT information support services based on Internet technologies for GNSS users

V. D. Glotov
IAC PNT
JSC «TsNIIMash»
Korolev, Russia
GlotovVD@tsniimash.ru

A. A. Arzhannikov
IAC PNT
JSC «TsNIIMash»
Korolev, Russia
AA.Arzhannikov@glonass-iac.ru

S.I. Baturin
IAC PNT
JSC «TsNIIMash»
Korolev, Russia
SI.Baturin@glonass-iac.ru

E.V. Bakaeva
IAC PNT
JSC «TsNIIMash»
Korolev, Russia
E.Bakaeva@glonass-iac.ru

V.L. Lapshin
IAC PNT
JSC «TsNIIMash»
Korolev, Russia
V.Lapshin@glonass-iac.ru

S.D. Zhilenko
IAC PNT
JSC «TsNIIMash»
Korolev, Russia
SD.Zhilenko@glonass-iac.ru

Activities on information support based on Internet technologies for users of global navigation satellite systems (GNSS) began at the Information and Analytical Center for Coordinate-Time and Navigation Support (IAC PNT, Korolev) in the 1997. Currently this work is carried out within the Roscosmos Space State Corporation contract via https://glonass-iac.ru website, which has first positions in search engines (Yandex and Russian-language segment of Google) for GLONASS requests in Russian and English languages.

I. GENERAL INFORMATION

At the same time IAC PNT exercise powers of Roscosmos Space State Corporation specified in Decree of the Russian Federation government of April 30, 2008 № 323:

- providing civilian (including foreign) users information about the current and predicted state of the GLONASS system orbital constellation spacecraft;

- introduction of the GLONASS system navigation technologies and augmentations in the interests of the socio-economic development of the Russian Federation and in the interests of civilian and scientific purposes (including foreign);

- informational support for navigation activities by posting on its official website information about navigation services provided in accordance with the standards of public services and these standards.

IAC PNT provides continuous online information support for Russian and foreign civilian users of the GLONASS system in five languages. Information support is based on the principle of providing up-to-date information for users of any level of competence - from reference materials to detailed precision estimates of the accuracy of ephemeris and time information of all GLONASS and GPS spacecraft within the framework of the service based on the performance monitoring and verification system [1] created by order of Roscosmos Space State Corporation

II. GNSS PERFORMANCE MONITORING

Performance monitoring and verification system service provides users with the results of evaluation of the main functional characteristics that determine the potential user accuracy of GLONASS and GPS systems: transformation parameters of terrestrial reference frames and time scales of GLONASS and GPS, errors of clock corrections and ephemeris information for all spacecraft of the systems, errors of location and time determination due to the error of the space segment of the systems, estimates of the accuracy of ephemeris information IAC PNT, including relative to foreign centers of analysis on the interval of three months.

The service based on a mobile measuring and diagnostic laboratory provides testing of navigation equipment samples in real operating conditions, monitoring of the radio-navigation situation in a given area of Russia, detection and localization interference sources for GNSS signals. The navigation equipment test bench in kinematic mode using a robot manipulator allows to determine millimeter errors of navigation equipment positioning in dynamic mode, including the use of high-precision augmentations.

III. DEVELOPMENT DIRECTIONS

IAC PNT is developing its own analytical products, including a global ionospheric map [2] of world-class accuracy that can be used to assess the accuracy of ionospheric models transmitted as part of the signals of all GNSS and space-based functional augmentation systems (SBAS).

One of the promising of IAC PNT service developments is the assessment of characteristics of correction information dataflow of Russian and foreign functional augmentation services based on high-precision technologies such as SBAS and PPP (Precise Point Positioning). In addition, IAC PNT is currently developing and experimentally testing the service of testing multi-system navigation equipment on the basis of all operating satellite navigation systems: GLONASS, GPS, Beidou, Galileo, QZSS.

IV. ACCESS TO INFORMATION

Access to information on IAC PNT service capabilities and results is organized using IAC PNT website and FTP-server, which contains data on current status, health flags and almanacs of all GNSS, final and preliminary

characterization estimates based on ephemeris-time accuracy analysis as well as on IAC PNT own analytical products.

In addition to that, IAC PNT website provides access to the test reports of the mobile measurement and diagnostic laboratory evaluations of GNSS user characteristics in real operating conditions, and also provides access to a wide range of tools for testing of GNSS, monitoring of GNSS user characteristics, including the use of corrective information of augmentations.

REFERENCES

[1] Section "Performance monitoring and verification system" on the website of the IAC KVNO JSC TsNIIMash [Electronic resource]. URL: https://www.glonass-iac.ru/skph/ (date: 01.02.2024).

[2] Section "Ionosphere" on the website of the IAC KVNO JSC TsNIIMash [Electronic resource]. URL: https://www.glonass-iac.ru/iono/ (date: 01.02.2024).

Comparison of methods for evaluating navigation support for the GLONASS system

T.A. Brovko
Radio systems department
NRU "MPEI"
Moscow, Russia
brovkota@mpei.ru

A.P. Malyshev
Radio systems department
NRU "MPEI"
Moscow, Russia
malyshevap@mpei.ru

V.B. Pudlovsky
Radio systems department
NRU "MPEI"
Moscow, Russia
pudlovskyvb@mpei.ru

Abstract— **The paper presented a method of estimating the performance characteristics of the GLONASS system, in terms of estimating the coordinate error, which takes into account the time and space correlation of the systematic components of the pseudorange errors. The comparison of three methods of coordinate error estimation is carried out, including the methodology of the GLONASS open service performance standard. The possibility of using simulation model of GLONASS system for reproduction of current values of errors of coordinates estimation and estimation of characteristics of errors of navigation estimations is shown.**

Keywords— GLONASS, position navigation and time error, dilution of precision

I. INTRODUCTION

With increasing accuracy of ephemerides and clock support of the orbital constellation (OC), the greatest contribution to the error of navigation support is made by the environment of propagation of radionavigation signals from navigation space vehicles (SV) to the user equipment (UE). The complexity of the processes in various GLONASS subsystems, the dependence of measurements on a large number of factors does not allow for the required studies of the accuracy characteristics of the system based on certain analytical ratios.

An effective way is the application of methods of mathematical modelling of individual elements of the GLONASS system, which are used to solve a variety of problems. The existing models do not always take into account the interrelationships between individual GLONASS subsystems and elements, leading to distorted results, in particular when assessing the characteristics of navigation estimations. Therefore, it seems relevant to develop a full-scale simulation model of the GLONASS.

The key characteristic of the GLONASS system for the user is the error of position navigation and time (PNT) estimations. However, in a number of reports [1-4] and regulatory documents [5], the error of PNT is interpreted as an error of positioning due to the space segment (hereinafter referred to as SISRE), depending on the dilution of precision, determined by the composition of the OC.

The purpose of this work is to validate the simulation model of the GLONASS as an alternative option for estimating the error of the PNT, and to compare the proposed method with the methods used to assess the quality of the basic navigation service of the GLONASS system.

To approve the work of the simulation model we used the method of comparing the errors of PNT results reproduced by the model and obtained from the results of processing real measurements from the IGS IRKJ00RUS [6] station.

II. THE DATA USED AND THE METHODS OF PROCESSING

A. The data used

To reproduce the errors of PNT in the simulation model of the GLONASS system on the day of 10.01.2024, the following data were used:

− Individual SISRE SV's visible at the selected simulation interval in BRDC-IAC_240110.etd file [7];

− raw measurement files in RINEX format for UE in Irkutsk irkj0100.24o [6];

− broadcast GLONASS ephemerides summarized in the Information and Analytical Center (IAC) for a given date in RINEX .*g [7] brdc0100.24g.;

− file for a' posteriori processing of coordinates and clock shifts of SV's Sta30s22963.sp3 [7];

− global ionospheric map file in the IONEX format [7].

B. Estimation of PNT errors based on real and model data

To validate the accuracy of the simulation model, it was tested by comparing it with the errors in coordinate estimation derived from real data using the RTKLib software [8].

On the day of January 10, 2024, for the IRKJ00RUS station in Irkutsk, the simulation was run under the following conditions:

- 24 GNSS SV's positioned according to the GLONASS almanac;

- UE operating on standard precision signals in the L1 frequency band;

- pseudorange measurement error due to multipath propagation of radionavigation signals is not modelled;

- reproduction of tropospheric and ionospheric refraction effects based on established algorithms [9];

- least squares algorithm used for navigation solutions [10];

- the errors in navigation solutions were assessed relative to the known coordinates of the station.

The pseudorange measurement error due to OC for the i-th observed SV's was reproduced in the simulation model as follows:

$$\Delta R_i = \frac{(\Delta S_i, r_{i,ENU})}{\left\| r_{i,ENU} \right\|} + \Delta \tau_i \cdot c, \tag{1}$$

where ΔS_i - current difference in coordinates of the SV center of mass in the topocentric coordinate system (ENU)

obtained from *.sp3 and .*24g files using known algorithms [10,11]; $r_{i,ENU}$ - vector of the true range to the i-th SV in ENU, $\Delta\tau_i$ - the current clock difference of the i-th SV according to the data in the *.sp3 and *.24g files using the known algorithms [11, 12]; c - speed of light.

In order to synchronize the simulated measurement errors with the real data, they were reproduced only for those SV's for which measurements were available in the irkj0100.24o observation file at the specified points in time.

To compare the degree of coincidence of the obtained realizations of errors in estimating coordinates from real measurements and using the simulation model of the GLONASS system, a sample correlation coefficient was used [12]:

$$\rho = \frac{\sum_i (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \qquad (2)$$

x_i - the value of one of the coordinates in the TCS, obtained from the simulation model at the i-th moment in time; $\bar{x}$ - the mean error value of one of the coordinates, calculated for the given sample.; y_i - The error value of one of the coordinates, obtained using real measurements at the i-th moment in time; $\bar{y}$ - The average error of one of the coordinates, obtained using real measurements, calculated for a given sample.

C. Options for converting SISRE the error of navigation and time support

At present, the GLONASS Open Service Performance Standard (OS PS) has been adopted as the main document approving the methodology for assessing the PNT support [5]. In this document, the focus is on such a characteristic as SISRE, which characterizes only errors for the OC. For the users, data on SISRE are published on the website of the Information and Analytical Center (IAC) for PNT support [14]. However, these data are not enough to assess the error of the user's coordinates. For the correct assessment of this error, data on atmospheric refraction (in the troposphere and ionosphere) are required, which are specified in Appendix B of the GLONASS Open Service Performance Standard [5].

However, there are no official data on refractive values, including troposphere, which makes it difficult to comprehensively assess the errors of the PNT support , according to section B.6 in Appendix B of the the GLONASS OS PS [5]. For this reason, in order to compare the errors in estimating coordinates obtained using the the GLONASS OS PS method and using the simulation model, we used data only for SISRE from the BRDC-IAC_240110.etd file at the selected interval without taking into account other components of the total pseudorange error budget. Thus, the current estimate of coordinate errors according to the IAC data, in this paper, was calculated for the following two options:

1) Using data from IAC according to the following formula:

$$\Delta X_1 = (H^T \cdot H)^{-1} \cdot H^T \cdot \Delta_{SISRE}, \qquad (3)$$

where ΔX_1 - the vector of the error in the user's coordinates in the ENU; H - matrix of direct cosines of the lines of sight user – SV [10]; Δ_{SISRE} - vector obtained by interpolating values equivalent to the pseudorange error from the BRDC-IAC_240110.etd file for all observed SV's at a given time.

2) *Using the SISRE averaging over the entire OC.*

$$\Delta X_2 = \begin{bmatrix} \Delta E_2 \\ \Delta N_2 \\ \Delta U_2 \end{bmatrix} = \begin{bmatrix} \Gamma_{EDOP} \\ \Gamma_{NDOP} \\ \Gamma_{VDOP} \end{bmatrix} \cdot \sigma_{SISRE}, \qquad (4)$$

$$\sigma_{SISRE} = \frac{1}{N} \sum_{i=1}^{N} \Delta_{SISRE,i}, \qquad (5)$$

where $\Gamma_{EDOP}, \Gamma_{NDOP}, \Gamma_{VDOP}$ - the value of the delusion of precision based on the coordinates in the ENU; σ_{SISRE} - the average SISRE value for the entire exhaust gas at the current moment, determined by the formula (5); $\Delta_{SISRE,i}$ - the value of the pseudorange equivalent error from the BRDC-IAC_240110.etd file for the i-th NSA from the OG.

To compare the degree of coincidence of the obtained realizations of the errors in the estimated coordinates only due to SISRE according to the data of the IAC and the simulation model of the GLONASS system, a sample correlation coefficient (2) was also used.

III. RESULTS

To assess the correctness of the GLONASS simulation model, comparative results obtained using real measurements and modeling are presented below.

To illustrate the change of PNT errors over time, the values in Figure 1 show the implementations of estimating the altitude error over an interval of four hours from the selected day, obtained using a simulation model (blue line) and based on real data (red line).

Fig. 1. Implementation of current values of altitude estimation error..

To illustrate the change of PNT errors only due to the influence of SISRE over time, Figure 2 shows the realisations of altitude error over a four-hour interval of the selected day, obtained using a simulation model (blue line), using individual (3) and average (4,5) SISRE values (red and yellow lines, respectively).

Fig. 2. Current implementation of altitude estimation error values (UP).

Distribution of errors only due to the effect of SISRE on the daily interval of the selected date, obtained using the simulation model (blue) and using individual and average SISRE values (red and blue, respectively) are shown in Figure 3.

Fig. 3. Distribution of coordinate estimation errors along the East axis.

Distribution of coordinate errors along the East axis are shown in Figure 4. Similarly, distribution of coordinate errors along the North axis are shown in Figure 4.

Fig. 4. Distribution of coordinate estimation errors along the North axis.

Fig. 5. Distribution of coordinate estimation errors along the Altitude axis.

Table 1 shows the values of the sampling correlation coefficients obtained by (2) for three processing options:

1) Between the realisations of coordinate errors obtained through simulation model and with the use of real GNSS measurements;

2) Between the realisations of coordinate errors obtained through simulation considering only the errors of the SISRE, calculated using formula (3);

3) Between the realisations of coordinate errors obtained through simulation considering only the errors of the SISRE, calculated using formula (4,5).

TABLE I.   THE VALUE OF THE CORRELATION COEFFICIENT

Option of processing	Correlation coefficient			Number of measurements
	E	*N*	*U*	
1	0.87	0.84	0.95	2880
2	0.47	0.48	0.90	2880
3	-0.47	0.00	0.04	2880

IV. DISCUSSION OF RESULTS

A comparative analysis of the coordinate errors reproduced by the simulation model and obtained from real data showed a high degree of correlation between these errors (Table 1, option 1), which can be seen in Figure 1.

These data characterize a sufficient degree of adequacy of the GLONASS simulation model to reproduce the coordinate errors at the selected modeling interval, as well as the completeness of the accepted budget components of the errors used in the modeling.

However, the degree of correlation of the error results only due to SISRE (Figures 2-5), as a result of modeling and using the the GLONASS OS PS methodology, turned out to be not so high compared to the degree of correlation of total errors relative to real measurements (Table 1, options 2 and 3).

In comparison with the simulation results of errors obtained by the the GLONASS OS PS method alone, the average value of the coordinates shifts to the range of positive values, especially when using the average values for the EG SISRE obtained by formulas (4,5).

The analysis of distributions in Figures 3-5 shows the presence of a shift to the positive values for the distribution of PNT errors obtained by GLONASS OS PS methodology (4, 5). This can be explained by the peculiarities of the expression used in GLONASS OS PS [5] to calculate the pseudorange error in Appendix B.6

CONCLUSION

The possibility of using a simulation model of the GLONASS system to reproduce the current values of errors in the estimation of coordinates is shown.

A high degree of correspondence between the PNT errors obtained by modeling and real data makes it possible to use a simulation model to assess the quality of navigation support based on GLONASS signals.

The comparative analysis of the obtained results showed a significant discrepancy between the values of estimation of IEE errors obtained by GLONASS OS PS methodology and with the use of the simulation model of the GLONASS.

REFERENCES

[1] S.N. Karutin, A.N. Kuzenkov, Scientific and Technical Problems of GLONASS Development in Modern Conditions,

[2] A.E. Tyulin, V.V. Betanov, Guaranteed Coordinate-Time and Navigation Support of Satellite System Information Consumers, Legal Informatics No 2, pp. 1-11. 4-16, 2020.

[3] Boyarkeeva O.V. Application of Simulation Modeling Methods for Studying the Accuracy of Unsolicited Trajectory Measurements on GLONASS Navigation Satellites. diss. Novosibirsk. FSUE "SNIIM", 2011. 125.

[4] A.A. Arzhannikov, V.D. Glotov, A.L. Kozhinov, V.V. Mitrikas, S.A. Panov, I.O. Skakun, V.V. Yanishevsky, On the ways to improve the accuracy characteristics of the GLONASS system, Cosmonautics and rocket engineering, 2 (125), pp. 1. 25-37, 2022.

[5] Open Service Performance Standard (OS PS) Revision 2.2 (6.2020) Korolev 2020

[6] International GNSS Service (IGS). Официальная страница. URL: https://network.igs.org/ (date of access: 25.04.2024). - Text: electronic.

[7] FTP archive Applied Consumer Center GLONASS Information and Analytical Center for Coordinate-Time and Navigation Support - ftp://ftp.glonass-iac.ru (date of access: 25.04.2024). - Text: electronic.

[8] System of Differential Correction and Monitoring of SDKM. Interface control document. Radio signals and digital information composition of the functional complement of the GLONASS Differential Correction and Monitoring System, Revision 2.0 MOSCOW 2020

[9] RTKLIB: An Open Source Program Package for GNSS Positioning. URL: https://rtklib.com/ (дата обращения: 25.04.2024). - Текст: электронный.

[10] GLONASS. Modernization and development prospects. Monograph / Edited by A.I. Perov. Moscow, Radiotekhnika Publ., 2020. – 1072 p.: ill. ISBN 978-5-93108-198-4

[11] Global Navigation Satellite System GLONASS. Interface control document. Radio navigation signal in the L1, L2 bands (version 5.1). Moscow: OAO "Russian Space Systems", 2008.

[12] J. Sanz Subirana, J.M. Juan Zornoza and M. Hernández-Pajares. GNSS Data Processing, Vol. I: Fundamentals and Algorithms (ESA TM-23/1, May 2013). ISBN 978-92-9221-886-7.

[13] GOST R ISO 3534-1-2019 Statistical methods. Dictionary and symbols. Part 1. General statistical terms and terms used in probability theory.

[14] Applied Consumer Center GLONASS Information and Analytical Center for Coordinate-Time and Navigation Support: official site. - 2022. - URL: https://glonass-iac.ru/ (accessed 20.04.2024). - Text: electronic.

Ionospheric model for navigational GLONASS-definitions

Dmitrii Trofimov
Saint Petersburg State University
Saint Petersburg, Russia
d.trofimov@spbu.ru

Sergey Petrov
Saint Petersburg State University
Saint Petersburg, Russia
s.d.petrov@spbu.ru

Ilya Chekunov
Bauman Moscow State Technical University
Moscow, Russia
onip4@mail.ru

Abstract—**The article describes an approach to solving the problem of the absence of data on the state of the ionosphere in the GLONASS navigation almanac. This leads to the fact that in the case of single-frequency observations, the ionospheric delay is not corrected, for which makes single-frequency GLONASS measurements obviously less accurate than single-frequency observations of other systems, primarily GPS. It is proposed to construct a fully autonomous model of the ionosphere based on the decomposition of global maps of the total electron content of the ionosphere into spherical functions. The optimal depth of decomposition and methods for modeling decomposition coefficients are determined.**

Keywords——— ionosphere, total electron content, GLONASS, spherical functions

I. INTRODUCTION

One of the main requirements for global navigation satellite systems (GNSS) is their ability to make accurate and reliable measurements based on the received navigation signal, without a source of additional information. In this part, the domestic GLONASS system has disadvantages compared to other GNSS. The reason for this is the absence of ionospheric data in the GLONASS navigation almanac, which is transmitted as part of the navigation signal. All systems are built on a similar principle; coordinates are determined based on rangefinder measurements of distances to navigation spacecraft. The influence of the ionosphere is removed by transmitting the ranging signal at two frequencies. However, a receiver that handles signals at two frequencies is an expensive equipment; there is a class of cheap, single-frequency receivers. Due to its price, which may not exceed hundreds of rubles or several dollars, this equipment is extremely widespread and is actively used in unmanned vehicles and other applications. In the GPS system, ionospheric correction of single-frequency observations is performed based on information transmitted in the almanac, which contains the parameters of the ionospheric model. As stated above, there is no such information in the GLONASS navigation almanac. In multi-system receivers that operate on the signal of several GNSS, information from the GPS navigation almanac is used to obtain a GLONASS navigation solution in the case of single-frequency observations to take into account the ionospheric delay. However, such a solution leads to some vulnerability, since the unavailability of a GPS signal, for example due to the use of electronic warfare, directly affects the accuracy of navigation determinations based on single-frequency GLONASS observations. Thus, there is no independence of single-frequency GLONASS observations, which reduces the suitability and competitiveness of the GLONASS system in the field of low-cost equipment. It should be noted that errors in determining the ionospheric delay are the main source of errors in navigation and coordinate determinations from single-frequency GNSS observations and can reach values of the order of 30-40 meters.

This situation needs a solution. Let us consider possible approaches to solving the problem. The most radical and sure way to solve the problem is to include data on the ionosphere in the navigation message. However, this solution seems at the same time the most unlikely. Changing the navigation message format will disrupt the continuity of the system; old equipment will either become useless or will require a software update. This is a decision that affects the fundamental foundations of the system. Other approaches are either to create a special autonomous model of the total electron content (TEC) of the ionosphere, which can be used as part of additional software for processing GLONASS navigation definitions, or to create a special service for determining the ionospheric delay. It seems most relevant and necessary to develop an autonomous TEC model, since it is this path that will allow solving the problem of precise determination of coordinates solely on the basis of the received GLONASS signal. A special service for determining delay is also necessary, but its use requires a connection with it, which is not always possible.

At the moment, one of the frequently used methods for determining the ionospheric delay in the conditions of post processing of observations is the use of global TEC maps provided by various GNSS data processing centers, for example, the Russian Information and Analysis Center for Positioning, Navigation and Timing (IAC PNT) or others, for example, the European Center for Orbit Determination in Europe (CODE). Numerous software designed for processing observations usually read standard map data storage formats without any problems. Thus, a possible result of creating a TEC model may be a technique for autonomously constructing a global TEC map of the ionosphere. The most promising approach seems to be the following. Decompose global ionospheric TEC maps into spherical functions, thus, at any time, the spatial distribution of TEC (with a certain accuracy) can be specified by a finite set of coefficients for the corresponding spherical functions. Since we already have a fairly long accumulated array of data on the ionosphere (CODE has been providing its data since 1999), for each of these coefficients, in principle, it is possible to build a model that, based on date and time and 11-year cycle of solar activity, will allow adequately predict it. Thus, to solve our problem, it is necessary to determine the approximate accuracy of the model, based on it, determine the optimal depth of decomposition and, accordingly, the number of decomposition coefficients. The Klobuchar model used in GPS can be used as a guide to the accuracy of which to take as a basis, so it seems necessary to

consider it in detail. But first, let's look at global TEC maps and the basic mathematical principles of our approach.

II. GLOBAL TEC MAPS

The main data on which we build our model are global TEC maps of the ionosphere. Currently, map data is provided by various GNSS data processing centers. In addition to the above-mentioned IAC PNT and CODE, these maps are also built by Wuhan University (PRC) or the Jet Propulsion Laboratory (NASA JPL) in the USA. All these maps are built on the same principle. A single-layer ionosphere model is assumed, that is, it is assumed that all free electrons are located in a thin layer at a certain height above the Earth's surface. As a rule, the height of this thin layer is assumed to be 450 km (in the IAC PNT and CODE maps). On the grid, with a step of 5° in longitude (longitude from -180° to 180°) and 2.5° in latitude (from -87.5° to 87.5°), TEC values are given for each point in units of 0.1 TECU (Total electron content unit).

Fig. 1. An example of a global TEC map reconstructed from spherical expansion coefficients up to degree 15 (top). The original map from the IAC PNT is in the center. Below is a TEC map constructed based on the Klobuchar model by the Astronomical Institute of the University of Bern. The maps represent the spatial distribution of TEC in TECU at 12:00 UTC on January 10, 2024.)

These maps are based on observations from a global network of GNSS stations. It is necessary to understand that the distribution of these stations is uneven, far from optimal, which, accordingly, affects the accuracy of the constructed TEC maps. According to [1-4], the root mean square error of

TEC maps reaches 3 TECU over continents and over oceans it can be even greater. This value gives us a certain lower limit of accuracy. It is hardly possible to construct a mathematical model of the ionosphere that will be more accurate than the provided global ionospheric TEC maps, so it is built on their basis. On the other hand, the different accuracy of maps over continents and over oceans theoretically makes it possible to build local models of the ionosphere that work in some selected continental zones.

III. SPHERICAL ANALYSIS OF GLOBAL TEC MAPS

We decompose global TEC maps into spherical functions based on the standard mathematical apparatus described in [5]. However, let us briefly outline the main essence of the method.

Any value $y(\lambda, \varphi)$ on the surface of a sphere, with coordinates λ - longitude and φ - latitude, can be represented as a sum:

$$y(\lambda, \varphi) = \sum_{n=0}^{n_{max}} \sum_{m=0}^{n} \tilde{P}_{nm} \sin\varphi \, (a_{nm} \cos m\lambda + \\ + b_{nm} \sin m\lambda), \tag{1}$$

where $\tilde{P}_{nm}$ - are normalized Legendre polynomials of degree n and order m, a_{nm} and b_{nm} are decomposition coefficients. In turn, the normalized Legendre polynomials are defined as:

$$\tilde{P}_{nm} = N_{nm} P_{nm}, \tag{2}$$

Where N_{nm} is the normalizing function and P_{nm} are the Legendre polynomials.

The decomposition was carried out up to degree 15 and order 15 inclusive. This decomposition is complete; the difference between the maps reconstructed by the coefficients of decomposition by spherical functions and the original ones is minimal, and is on the order of 0.1 TECU. These degrees and order of decomposition are also specified by the procedures for constructing global TEC maps. Thus, we can set an upper bound for the decimposition in terms of spherical functions. However, this upper limit seems excessive. Already, to a large extent, the coefficients for functions of degree 15 have near-zero values and a chaotic, noisy behavior. It most likely makes sense to set the degree of decomposition used in the offline model to a lower value. Modeling all degrees of decomposition may make sense for another task, namely creating a service for accurate TEC forecast. An example of the original TEC map and the map reconstructed based on the decomposition performed is shown in Fig. 1. To develop the methodology, we used data from the domestic center IAC PNT. However, unfortunately, IAC PNT data are available only starting from 2018, and to build a model it is desirable to use the longest data series possible. Given the obvious influence of solar activity on the state of the ionosphere, it is desirable that the processing period covers several cycles of solar activity. Therefore, further work was carried out on the basis of CODE data available since 1999, thus analyzing data for the last full 25 years. To set the optimal decomposition depth for constructing an autonomous ionospheric delay model, it is necessary to consider the Klobuchar model, which we rely on in terms of accuracy

IV. KLOBUCHAR MODEL

The Klobuchar model was proposed [6] in the 1980s, and as mentioned above, is a working model of the ionosphere for single-frequency GPS observations. This model refers to single-layer models of the ionosphere, that is, it is believed that all free electrons are located in a thin layer, at an altitude of 350 km above the Earth's surface. This model is focused on directly calculating the delay, so the parameters that are set there are not the electron concentration, but the delay itself. It is believed that the night, unexcited ionosphere gives a signal delay of 5 ns. An area of increased ionization, which is caused by solar radiation, moves along the day side. This area, or rather the delay caused by the passage of a signal through this area, is modeled as a bell-shaped profile, described as part of a sinusoid (from 0 to π). The parameters of this profile are included in the transmitted navigation message. On the one hand, this model is simple, has a clear and understandable physical meaning; long-term experience shows that for single-frequency GPS observations, this model makes it possible to halve the error caused by the delay in the ionosphere.

It should also be noted that the results of the Klobuchar model can be recalculated into a global ionospheric TEC map provided in a standard format. International processing centers such as IGS make such maps publicly available. An example of such a map is shown in Fig. 1. Let us try to estimate what depth of expansion (the maximum degree of the spherical function) will be optimal compared to the Klobuchar model. Let us compare the ionospheric delay at the zenith for a point with coordinates 60º north latitude and 30º east longitude, which corresponds to St. Petersburg (the coast of the Gulf of Finland in the Lisiy Nos area). On the one hand, we will compare the delay obtained from the Klobuchar model, on the other hand, the delay obtained from the synthesized map, depending on the maximum degree of decomposition used. This comparison is made in Table 1, delays are given at 12:00 UTC January 10, 2024.

TABLE I. DELAY COMPARISION

Degree	RMD, m	Zenith ionospheric delay, in m		
		Synthesized map	Klobuchar	CODE map
0	3.16	4.78		
1	2.19	5.36		
2	1.31	4.85		
3	1.01	4.10		
4	0.96	3.92		
5	0.80	3.95		
6	0.77	4.11		
7	0.73	4.18		
8	0.63	4.51		
9	0.55	4.75	7.42	5.07
10	0.51	5.22		
11	0.45	5.22		
12	0.35	5.06		
13	0.31	5.23		
14	0.20	5.17		
15	0.07	5.07		

From Table 1 it can be seen that decomposition by any degree, starting from the second, gives a more accurate value than the Klobuchar model if we take the original TEC map as a reference value. The maximum deviation from the reference value is 1.15 m for the degree of decomposition 4, for degrees of decomposition starting from 9 it does not exceed 0.5 m. The root mean square deviations (RMD) of the synthesized maps from the original one are also given. The standard deviation of the Klobuchar model from the TEC map is 2.17 m for the given date. That is, even synthetic maps with very limited degree of decomposition give a standard deviation almost twice as good as the Klobuchar model. Thus, we can assume that the ionosphere model, built on the basis of the decomposition of global TEC maps into spherical functions, even at low degrees of expansion (n = 7 - 8), will have potential accuracy no worse than the Klobuchar model, and subject to the development of an adequate model of the spherical coefficients decomposition may even surpass it.

V. MODELING THE IONOSPHERE BASED ON SPHERICAL DECOMPOSITION COEFFICIENTS

Fig. 2. TEC map built on the basis of an autoregressive model for expansion up to degree 15 (top) to degree 9 (in the center) and the original TEC map from CODE at 12:00 UTC January 10, 2024 (bottom).

Let's try to model the coefficients of spherical decomposition. To begin with, we will resort to purely mathematical modeling. Let us assume that the process of changing coefficients can be described by an autoregressive process model. Based on the original CODE data for the

period from 1999 to 2023 (109572 source maps), an autoregressive model of spherical decomposition coefficients was built using the Burg method [7]. The model order is p = 600. Since the time resolution of our initial data is 2 hours, it turns out that when calculating the coefficients, data for the last 50 days are used. It is of interest to check the adequacy of this model.

Based on this model and initial data ending on December 31, 2023, the ionospheric TEC map was predicted for 12:00 UTC on January 10 2024. The results are presented in Fig. 2. Maps were constructed for decomposition up to degree 15, up to degree 9, and a map for January 10 from CODE, built on the basis of observations, is also shown.

The RMD of the predicted maps was 1.55 m for a map constructed using decomposition up to degree 9 and 1.48 m for a map constructed using decomposition up to degree 15. The delay value itself at the zenith was 2.80 m in 1 case, 2.76 m in the second. The RMD of the map constructed based on the Klobuchar model is 2.17 m. It can be seen that the data predicted on the basis of the autoregressive model for the spherical decomposition coefficients more accurate than according to the Klobuchar model, there are fewer standard deviations of such maps, these maps better reflect the spatial features of the TEC distribution.

The further direction of improving the model is to replace the coefficient model from autoregressive to trigonometric and to select the optimal degree of decomposition. Such a model will be quite suitable for autonomous use.

VI. CONCLUSION

A model of the ionosphere was constructed based on the decomposition of global TEC maps into spherical functions.

Even a limited decomposition gives a zenith delay value that is more than 2 times more accurate than the Klobuchar model. A forecast model for the global TEC map has been constructed, which makes it possible to predict the vertical ionospheric delay 10 days ahead with an accuracy of approximately 1.5 times more accurate than the Klobuchar model. It is planned to create an improved model suitable for use in embedded receiver software.

REFERENCES

[1] Q. Zhang, Q. Zhao, "Global Ionosphere Mapping and Differential Code Bias Estimation during Low and High Solar Activity Periods with GIMAS Software," in *Remote Sensing*. vol. 10, No.5, p. 705, 2018, https://doi.org/10.3390/rs10050705

[2] J. Zhao, M. Hernández-Pajares, Z. Li, N. Wang, and H. Yuan, "Integrity investigation of global ionospheric TEC maps for high-precision positioning," *Journal of Geodesy*, vol. 95, no. 3, Springer, 2021. doi:10.1007/s00190-021-01487-8

[3] B. Milanowska, P. Wielgosz, A. Krypiak-Gregorczyk, W. Jarmołowski, "Accuracy of Global Ionosphere Maps in Relation to Their Time Interval," in *Remote Sensing*, vol 13. no. 18, 3552, 2021 https://doi.org/10.3390/rs13183552

[4] N. Wang, Y. Yuan, Z. Li, O. Montenbruck, and B. Tan, "Determination of differential code biases with multi-GNSS observations," *Journal of Geodesy*, vol. 90, no. 3, Springer, pp. 209–228, 2016. doi:10.1007/s00190-015-0867-4.

[5] V.V.Vityazev. *Analysis of astrometric catalogs using spherical functions*. Saint Petersburg, Russia: St. Petersburg State University Publishing House. 2017.

[6] J. A. Klobuchar, "Ionospheric Time-Delay Algorithm for Single-Frequency GPS Users," in *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-23, no. 3, pp. 325-331, May 1987, doi: 10.1109/TAES.1987.310829.

[7] J. P. Burg, "Maximum entropy spectral analysis," *37th Ann. Int. Mtg. Soc. Explor. Geophys.*, 1967.

Assessing Radio Navigation Systems through Flight Tests

I. Kopylov
M.M. Gromov Flight Research Institute
s. Zhukovsky, Russia
e-mail nio9@lii.ru

E. Kharin
M.M. Gromov Flight Research Institute
s. Zhukovsky, Russia
e-mail nio9@lii.ru

V. Kopelovich
M.M. Gromov Flight Research Institute
s. Zhukovsky, Russia
e-mail nio9@lii.ru

A.Yakushev
M.M. Gromov Flight Research Institute
s. Zhukovsky, Russia
e-mail nio9@lii.ru

E. Gorskiy
M.M. Gromov Flight Research Institute
s. Zhukovsky, Russia
e-mail nio9@lii.ru

V. Ilyin
M.M. Gromov Flight Research Institute
s. Zhukovsky, Russia
e-mail nio9@lii.ru

Abstract – **The report outlines the means and methods for assessing radio navigation systems as part of modern integrated aircraft flight and navigation equipment. The report gives examples of assessment results based on flight test data for each system under consideration.**

Key words – radio navigation equipment, flight tests, means and methods for assessing equipment characteristics.

I. INTRODUCTION

To assess radio navigation equipment for compliance with specified requirements, *M.M. Gromov Flight Research Institute* has developed special methods and means [1–4].

The onboard trajectory measurement complex (OTMC) developed by the Flight Research Institute is employed for recording radio navigation data and determining actual trajectory parameters. The OTMC determines highly accurate trajectory parameters using the differential mode of satellite navigation systems.

Flight and navigation equipment of modern aircraft comprises different radio navigation systems (RNS). RNS functioning is based on measurements of parameters of electromagnetic signals that are emitted by special airborne or ground-based equipment. Most aircraft use the following radio navigation and landing equipment:

- automatic direction finder (ADF);
- short-range radio navigation (SRNAV) system: domestic SRNAV system, VOR system (Very High Frequency Omni-Directional Range) and DME system (Distance Measuring Equipment) standardized by ICAO;
- Doppler navigation system (DNS);
- long-range radio navigation system (LRNAV).

For now, the DNS and LRNAV systems are mostly used in state aviation.

II. AUTOMATIC DIRECTION FINDER

Automatic Direction Finder (ADF) serves as stand-alone angle measurement equipment integrated into airborne radio navigation systems. The ADF receives signals from medium-frequency locators and broadcasting radio stations and continuously measures the relative bearing (RB) of radio station. Aircraft crew uses the ADF to handle piloting tasks, including approach maneuvers.

The relative bearing of radio station measured by the ADF is determined as the angle between the longitudinal axis of the aircraft and the direction to a ground-based radio station (RS). The angle is measured in the clockwise direction and changes from 0 to 360° (Fig. 1).

Fig. 1.   Relative bearing of radio station.

The ADF requirements define distances from a radio station for RB calculations by the system as well as the RB parameter measurement accuracy. The ADF range at altitude of 1000 m shall be not less than 150 km. In flight inbound the station (RB values in the neighbourhood of 0º) or outbound the station (RB values in the neighbourhood of 180º), the RB absolute error shall not exceed 3º with the probability of 0.95. For other RB values, the error shall be within ±5°.

Flight testing of the ADF shall be performed within the coverage area of a homing station with known operating characteristics. Figure 2 shows an example of ADF evaluation for an aircraft flying over the radio station. Time values expressed in seconds from the beginning of the day are plotted on the abscissa axis. RB values in degrees and actual value RB_A are plotted on the left side on the ordinate axis. Error values ΔRB and tolerance in degrees are plotted on the right side. Actual values RB_A were calculated after flight, based on aircraft position coordinates, ground-based radio station coordinates and true aircraft heading. The tolerance is equal to ±3º. The test flight was the flight inbound the station. When flying over the station, the RB signal was emitted with an uncertainty attribute. After flyover, the RB was in the neighborhood of 180º. The moment of RS flyover is marked with a green vertical line as shown in Fig. 2. Within the entire flight segment at RB values near 0 and 180º, the error ΔRB shown as a red line was in the range of −3º to 3º while less

than 5 % of the total amount of errors ΔRB obtained was beyond tolerance.

Fig. 2.   Example of RB parameter assessment during RS flyover.

III.   SHORT-RANGE RADIO NAVIGATION SYSTEM

The short-range radio navigation system (SRNAV) is designed for measuring the aircraft azimuth and slant range referenced to ground-based radio beacons located at points with known coordinates. Therefore, the SRNAV comprises azimuth and distance measuring channels, each including the ground-based radio beacons and related avionics. Radio beacons are deployed along air routes and at aerodromes. Since the SRNAV operates in the VHF band, its maximum range depends on the line-of-sight range (Fig. 3).

Fig. 3.   Graphical representation of the aircraft flight at altitude H and distance D from a ground-based radio beacon.

The domestic short-range UHF radio navigation system and VHF and UHF VOR/DME systems standardized by ICAO are the most common systems. The VOR system determines the aircraft azimuth referenced to a ground-based radio beacon while DME is intended to measure the distance to a radio beacon. Ground-based VOR и DME radio beacons may be deployed separately or co-located.

The SRNAV azimuth channel is based on time or phase methods. A VOR station generates an azimuth signal whose informative parameter depends on the azimuth of the receiver point. Airborne equipment determines the moment of azimuth reference by the reference signal transmitted from a radio beacon. Azimuth information is indicated by the time or phase shift of the received signal relative to the reference signal. Domestic SRNAV systems use the time method. The phase method is used in foreign systems (VOR).

The distance measuring channel in domestic (SRNAV) and foreign (DME) short-range navigation systems operates on the basis of the time method. Distance information is indicated by the difference between the moment of measurement of an interrogation signal and the moment of reception of a return signal by airborne equipment.

Flight tests of SRNAV and VOR/DME equipment involve ground-based radio beacons adjusted as appropriate. With its role as the onboard measurement system, the OTMC records VOR/DME information. As part of the OTMC, the SNS receiver and the ground base station are used to obtain actual values of trajectory parameters in the differential mode. Based on flight tests of VOR/DME equipment, the following problems can be solved:

− determining the range from a ground-based radio beacon;

− assessment of radio beacon distance measurement error;

− assessment of azimuth measurement error;

− determination of functional capabilities at large elevations;

− assessment of VOR/DME signals for the presence of signal failures and overshooting.

VOR/DME is assessed using the developed flight test practice. For distance determination, a test flight is at the cruise altitude is conducted inbound and outbound the radio beacon. Distances from the radio beacon, within which the equipment starts or stops determining the azimuth and distance, are recorded as the system range.

Errors of VOR/DME parameters are assessed at a distance of around 20 to 350 km from the radio beacon to flight altitudes of around 10,000 m. Post-flight data processing includes the preprocessing procedure to obtain physical values of azimuth and range represented in from of tables and diagrams, conduct a quick analysis and select flight segments for secondary processing. During secondary processing, azimuth and distance values measured by the equipment are compared with their actual values. Azimuth and distance measurement errors are determined at a frequency of at least 1 Hz.

$$\Delta A_i = A_i - A_{i,\text{A}}, \quad \Delta D_i = D_i - D_{i,\text{A}}.$$

Figures 4 and 5 show examples of the flight test results for estimating azimuth error VOR and range error DME, respectively. Depending on the distance to the VOR beacon, the VOR azimuth values and the actual azimuth values are shown. The tolerance for the azimuth error ΔA does not depend on the range and is shown in Fig. 4 in the form of two parallel blue lines. The estimation of expectation M for 38 values of azimuth error in the example under consideration is equal to 1.16°; the estimation of the standard deviation σ is equal to 1.38°. Therefore, $M + 2\sigma < 5°$, which meets the specified requirements.

Fig. 4. Example of assessment of VOR azimuth measurement error ΔA.

Fig. 5. Example of assessment of DME distance error ΔD.

All values of DME distance error ΔD shown in Fig. 5, remain within the tolerable margin. The estimation of expectation is 0.054 km, and the estimation of RMS σ is equal to 0.053 km. That is why the value of $|M| + 2\sigma$ equal to 0.16 km is not greater than the required value of 0.37 km.

In the scanning mode, the applicable airborne DME system can determine the distance to several ground-based beacons. If the number N of radio beacons, relative distances to which are D_i, $i = 1, \dots , N$, is not less than 2, the current aircraft position coordinates (φ, λ) can be determined by distances only. Calculations include flight altitude H and coordinates of radio beacons $(\varphi_{bi}, \lambda_{bi}, H_{bi})$. Before calculations, all measured distances D_i shall be normalized to the fixed moment of time T of coordinate calculation. Cartesian coordinates of the radio beacon (x_b, y_b, z_b) and the aircraft (x, y, z) are calculated by the following formulas:

$$x_b = \left(a\xi_b + H_b\right)\cos\varphi_b \cos\lambda_b$$
$$y_M = \left(a\xi_b + H_b\right)\cos\varphi_b \sin\lambda_b \qquad (1)$$
$$z_M = \left(a\xi_b\left(1 - e^2\right) + H_b\right)\sin\varphi_b$$
$$x = \left(a\xi + H\right)\cos\varphi \cos\lambda$$
$$y = \left(a\xi + H\right)\cos\varphi \sin\lambda \qquad (2)$$
$$z = \left(a\xi\left(1 - e^2\right) + H\right)\sin\varphi$$

where $\xi_b = \dfrac{1}{\sqrt{1 - e^2 \sin^2 \varphi_M}}$, $\xi = \dfrac{1}{\sqrt{1 - e^2 \sin^2 \varphi}}$,

e is the Earth ellipsoid eccentricity.

The relationship between relative distances D_i and aircraft position coordinates is as follows:

$$\left(x - x_{bi}\right)^2 + \left(y - y_{bi}\right)^2 + \left(z - z_{bi}\right)^2 = D_i^2, i = 1,...,N \quad (3)$$

In conjunction with the equations (1), (2), the equations (3) constitute a system for unknown coordinates (φ, λ). If $N = 2$, the number of unknown quantities $(x, y, z, \varphi, \lambda)$ matches the number of equations. As a rule, unknown quantities can be found using the system of equations with the iterative method applied. The aircraft position coordinates calculated at the previous moment of time are selected as initial estimate. If the number of known distances to beacons is greater than 2, the number of equations becomes greater than the number of unknown quantities, and the solution is found using the least square method.

Since relative distances D_i contain errors, calculated current coordinates also contain a random error. Let us determine the characteristics of the measurement error of the current aircraft position calculated by two DME distances ($N = 2$). To simplify calculations, we will handle the problem in plane and assume that distances D_i are much greater than their errors ΔD_i.

$$D_i = D_i^o + \Delta D_i, \qquad \Delta D_i << D_i^o, \quad i = 1, 2.$$

Assuming that random errors ΔD_1 and ΔD_2 are independent, RMS errors $\sigma_{\Delta D}$ of distance are taken as identical values for each of two beacons. If DME meets the requirements, $\sigma_{\Delta D}$ does not exceed 0.185 km. RMS errors $\sigma_{\Delta R}$ of current position determination by two distances are calculated by the following approximation formula:

$$\sigma_{\Delta R} = \sqrt{2}\,\frac{\sigma_{\Delta D}}{\sin B} \qquad (4)$$

In the equation (4), B is the angle between the directions of beams from the current aircraft position to two ground-based radio beacons, which provide information. If DME meets the requirements, $\sigma_{\Delta D}$ does not exceed 0.185 km.

Thus, the position determination error is proportional to the radio beacon distance measurement error and depends on the angle of intersection of lines of position B. The RMS position determination error is minimal when the directions of lines from the aircraft to radio beacons are perpendicular ($B = 90°$). If we limit the interval of values of the angle B and determine the aircraft position by two DME distances when the value of sin B is near 1, the RMS error ΔR is relatively small. Let us assess the error $\sigma_{\Delta D} = 0.185$ km. At $B = 90°$, we will get $\sigma_{\Delta R} = 0.262$ km from the equation (4). Therefore, with the probability of 0.95, the error of position determination by two DME distances will not exceed $2\sigma_{\Delta R} = 0.52$ km. The method of position determination by two and more distances can be used for navigation and as the corrector in the flight management system, even in those flight segments where the requirements for navigation and flight management are sufficiently tough, for example, in the aerodrome zone when executing letdown procedures.

As an example, Fig. 6 shows the position error in the flight management system when using two DME distances as the corrector after post-flight test data processing. Within the entire flight segment under consideration, absolute values of coordinate error values do not exceed 270 m.

Fig. 6.   Position errors ($\Delta\varphi$, $\Delta\lambda$) of the flight management system in the mode of position correction by two DME distances.

IV.   DOPPLER NAVIGATION SYSTEM

Airplane and helicopter flight and navigation equipment includes the Doppler navigation system (DNS) or the airborne radar based on the Doppler effect. The DNS is designed for continuous automatic calculation of the components of the ground speed vector and drift angle. Drift angle (DA) is determined as the angle between the longitudinal axis of the aircraft and its flight direction relative to the ground surface. Drift angle is the result of sideslip motion and cross wind. Fig. 7 shows DA without sideslip.

Fig. 7.   Drift angle without sideslip.

The DNS system uses the Doppler effect for its operation. The Doppler effect is the difference between the received frequency of the ground reflected signal compared to the emitted signal frequency, and the difference depends on the aircraft ground velocity. To measure the ground speed, the DNS system is equipped with a pencil-beam antenna system that uses three or four beams for its operation. The reflected signal in each beam is Doppler-shifted with the shift value being directly proportional to the projection of the aircraft speed vector onto this beam. To measure the speed vector, only three noncoplanar beams may be used. The use of four beams provides certain redundancy without significant complication of the design.

DNS errors depend on the surface background. DNS requirements usually differentiate overland and oversea flights. According to the requirements, the DNS shall measure the ground speed with the probability of 0.95 and the measurement error shall not exceed 0.5 % of the actual speed; the drift angle measurement error shall not exceed 0.2° over land and 0.3° over sea. DNS switching to the Memory mode is allowed at roll angles over 30°.

To assess measured components of ground speed and drift angle, we should know actual values of the aircraft ground speed vector and the angular position, i. e. heading

angle Ψ, roll angle γ and pitch angle υ, as well as DNS antenna setting angles. Based on flight test data, DNS errors are assessed on the basis of satellite-based trajectory measurement aids using the OTMC. Based on the data of SNS differential mode, we can determine three speed components (V_N, V_E, V_H) within the axes of the geographic moving trithedral. Assumed to be actual values, ground speed V_A, track angle TK_A, drift angle DA_A are calculated.

$$V_A = \sqrt{V_N^2 + V_E^2},$$
$$\text{tg } TK_A = \frac{V_E}{V_N}, \tag{5}$$
$$DA_A = TK_A - \psi_A$$

In the equation (5), Ψ_A is the true heading value calculated with the help of integrated processing of data from the SNS receiver and inertial navigation system based on the Kalman filter. DNS errors in determination of ground speed and drift angle are calculated with the help of the differences

$$\Delta V_G = \Delta V_G^{DNS} - V_A,$$
$$\Delta DA = \Delta DA_G^{DNS} - DA_A, \tag{6}$$

where all the right-hand members of equation (6) are normalized to the same time value. To assess the components of DNS speed, the speed vector projections (V_{NA}, V_{EA}, V_{HA}) taken as actual values are recalculated with the help of actual values of heading angle Ψ_A, roll angle γ_A and pitch angle υ_A in the projection of (V_{XA}, V_{YA}, V_{ZA}) within the coordinate system bound to the aircraft construction lines. DNS errors in determination of ground speed vector components are calculated as follows:

$$\Delta V_X = V_X - V_{XA},$$
$$\Delta V_Z = V_Z - V_{ZA}.$$

Let us give an example of assessment of DNS output parameters in one of straight-line overland flight segment at a constant speed. Fig. 8 depicts the resulted DNS errors of drift angle ground speed projections V_X and V_Z.

Fig. 8.   DNS parameter errors in the flight segment.

The segment duration is around 1.5 hour. Time values are plotted on the abscissa axis. The figure shows DNS parameter error graphs and true heading values. In the flight segment under consideration, the expected value M

of the drift angle error is equal to −0.25°, the RMS deviation σ of error is equal to 0.30°, the value of |M| + 2σ is equal to 0.85°.

V. LONG-RANGE RADIO NAVIGATION SYSTEMS

Long-range radio navigation systems (LRNAV) are designed for determining aircraft current coordinates by signals from reference ground stations.

The main mode of operation of long-range radio navigation systems is the time-difference mode. Measurements of distance difference are based on an airborne receiver which determines the time difference in propagation of radio waves from two ground stations (master and slave). The master station signal is received by an aircraft at the certain moment of time T_m (Fig. 9).

Fig. 9. Aircraft position determination using LRNAV stations

The slave station signal reaches an aircraft at the moment of time T_N. The baseline delay B_N and the delay τ_n being constant for this station are selected in such a way that T_M is less than T_N for any receiver point. The difference between time values T_N and T_M defines the difference in distances between the master station with number N and the aircraft D_N.

$$D_N - D_M = c\,(T_N - T_M).$$

Here c is the radio wave propagation speed.

The line of equal differences in distance is the hyperbolic line and the point of intersection of two hyperbolic lines of equal differences in distances matches the aircraft position. Aviation currently uses pulse-phase and phase LRNAV systems. Pulse-phase radio navigation systems (PPRNS) are widely used for solving flight navigation problems. Such systems include domestic Tropik-2 and Tropik-2P LRNAV systems and foreign Loran-C system. PPRNS operate in the long-wave range with the radiated signal frequency of 100 kHz (wavelength of ~ 3 km).

For a few years, domestic radio industry has developed the following navigation user equipment operating by PPRNS signals: article A-711 (Kremniy) with coordinate converter (article A-713), articles A-723, A-727 and A-737И (Kvitok), as well as equipment PPA-S/V (jam-resistant airborne receiver for airplanes and helicopters) and PNAP-PLA (jam-resistant navigation user equipment of manned aircraft). The RMS position error in the PPRNS time-difference mode is 100–800 m depending on the geometry factor and known conditions

of radio wave propagation. Implementation of the differential mode can significantly increase the position determination accuracy that can reach 10–50 m. Phase radio navigation systems (PRNS) operate in the very low frequency (VLF) band. Such systems include the domestic Marshrut system and the foreign Omega system (currently not used). To resolve the phase ambiguity, phase radio navigation systems emit signals at three frequencies. For example, the foreign-made Omega PRNS employed VLF band frequencies of 10.2 kHz, 11.33 kHz, and 13.6 kHz. The domestic PRNS uses frequencies close to these values. As examples of domestic airborne navigation PRNS signal user equipment, we can mention articles A-723, A-727, A-722, PPA-S/V, PNAP-PLA. According to PRNS data, the RMS position error can be 1÷4 km. Implementation of the differential mode can improve accuracy characteristics.

The characteristics of radio navigation systems are assessed during flight tests using the methods and means developed by Research Unit NIO-9 at JSC Flight Research Institute n. a. M. M. Gromov. These methods and means are presented in proceedings of the International Conference on Integrated Navigation Systems (MKINS) for the period of 2005–2023. Flight research and testing of long-range radio navigation systems was conducted by means of flying laboratories at the Flight Research Institute n. a. M. M. Gromov and prototype aircraft. Tests were conducted in LRNAV operational areas during flights in the Flight Research Institute's aerodrome area and during en-route flights. Onboard LRNAV equipment information was recorded by the onboard trajectory measurement complex (OTMC). The data in standard and differential SNS modes were taken as actual aircraft position coordinates.

Fig. 10 shows graphs illustrating the in-flight errors of PPRNS A-723 equipment onboard a Tu-154M flying laboratory in the aerodrome zone operated by the Flight Research Institute n. a. M. M. Gromov. The obtained errors have constant components. With consideration given to these components and implementation of the differential mode of operation of article A-723, errors can be significantly decreased.

Fig. 10. LRNAV coordinate measurement errors.

VI. CONCLUSIONS

The developed flight test technology with the use of the onboard trajectory measurement complex allows

evaluation of performance characteristics of flight and navigation equipment and their systems.

Technology implementation allows running flight tests and research of radio navigation systems such as ADF, SRNAV, VOR/DME, DNS, LRNAV and other radio navigation aids.

The application of developed means and methods for flight testing of radio navigation and landing equipment makes it possible to reduce manpower and time expenditures required for flight tests.

REFERENCES

[1] **Kharin, Ye.G.** Flight Test Technologies for Aircraft Equipment with the Use of Airborne Trajectory Measuring Systems./ Ye.G. Kharin, I.A. Kopylov/ – M.: Izd-vo MAI-PRINT, 2012. – 360 p.

[2] **Kharin, Ye.G.** Assessment Methods and Means for Aircraft Piloting and Navigation Equipment During Flight Tests. /Ye.G. Kharin, I.A. Kopylov, V.G. Polikarpov, V.A. Kopelovich./ – XV Saint Petersburg International Conference on Integrated Navigation Systems. 26–28 May 2008, St. Petersburg, Russia, pp. 208–219.

[3] **Gorskiy, Ye.B.** Trajectory Measurements During Monitoring and Testing of Ground-Based Radio Equipment and Airborne Equipment of Instrument Landing Systems. /Ye. B. Gorskiy, Ye. G. Kharin, I. A. Kopylov, V. A. Kopelovich, A. V. Yasenok./ – XXVI Saint Petersburg International Conference on Integrated Navigation Systems. 27–29 May 2019, St. Petersburg, Russia, pp. 91–94.

[4] **Avgustov, L.I.** Aircraft Navigation in Near-Earth Space. /L.I. Avgustov, A. V. Babichenko, M. I. Orekhov, S. Ya. Sukhorukov, V. K. Shkred./ Edited by G. I. Dzhandzhgava – 2nd revised edition, M.: Izdatel'stvo Grani Uspekha, 2022. – 548 p.

Information Characteristics of Working Satellite Constellations in Relative Positioning Using Two GNSS under the Conditions of User Rolls

V.I. Baburov
«Navigator» JSC
Saint Petersburg, Russia
info@navigat.ru

N.V. Vasilyeva
«Navigator» JSC
Saint Petersburg, Russia
info@navigat.ru

N.V. Ivantsevich
«Navigator» JSC
Saint Petersburg, Russia
info@navigat.ru
BGTU VOENMEH
Saint Petersburg, Russia
sciencebstu@bstu.spb.su

Abstract—The possibility of using the information redundancy of the combined constellation of two satellite systems to increase the accuracy of relative positioning of aircraft moving each with an arbitrary roll is studied.

Keywords—*global navigation satellite systems, accuracy of relative positioning, rolls of the first and the second aircraft, simulation modeling.*

I. INTRODUCTION

Global satellite navigation systems are currently used on all types of modern vehicles to determine the coordinates and speed of their change due to the globality and accessibility of the navigation field they create anywhere in the globe and surface space. At the same time, in order to ensure high performance of these characteristics, certain requirements must be met for the navigation receivers and the surrounding space, the absence of interference and other factors. These conditions are listed in the interface control documents for the systems. Further, we will call such conditions standard or regular. However, these conditions are not met for all vehicle modes of moving. In such cases, the accuracy of GNSS positioning or positioning in general may not be guaranteed, for example, when each of the aircraft is moving with a roll or pitch. To increase the positioning accuracy in emergency situations to the values guaranteed by a single GNSS under standard conditions, methods can be used to increase the information redundancy of the working satellite constellations. One of these methods is based on the joint processing of navigation measurements from two GNSS, for example, GLONASS and GPS. In this case the satellites number of a working constellation increases, the information of the constellation increases, the entropy decreases, and the information capacity of each satellite decreases.

The report studies the possibility of using the information redundancy of the working constellation formed by two satellite systems to increase the accuracy of relative positioning of aircraft moving each with an arbitrary roll to values corresponding to the standard GLONASS positioning, as indicated in the interface control document for the satellite navigation system [1].

II. INFORMATION CRITERION FOR THE QUALITY OF THE WORKING SATELLITE CONSTELLATION

There are well-known publications in which the criterion of the number of navigation satellites in a working constellation is taken as an information criterion of the possibility of using data from navigation satellite systems. According to this criterion, the maximum permissible value of the entropy increase is set, considering the number of satellites in the working constellation as a "cyber-physical parameter" [2]. However, this criterion does not define the positioning accuracy, characterized by the size of the uncertainty zone of the measured navigation parameters, and it is not suitable for absolute and relative positioning. Therefore, we will use the criterion accuracy of positioning.

The navigation satellite information received by users spaced apart for relative positioning will be characterized by the amount of information and Shannon entropy. These information characteristics will change when the rolls of the first and/or the second user occurred. A change in these characteristics uniquely determines a change in the accuracy of navigation definitions. It can be shown [3,4] that, under certain conditions, the accuracy of relative positioning is determined by the geometric factor of the simultaneously visible satellites constellation. The change in the entropy of the working constellations of navigation satellites when user has roll will be equal to the change in the entropy of the geometric factors of these working constellations.

III. A METHOD FOR STUDYING GNSS WORKING SATELLITE CONSTELLATIONS FOR TWO USERS WITH INDEPENDENT ROLLS

The research of information characteristics of relative positioning was carried out by the method of mathematical simulation using the Monte Carlo method. When relative positioning, it is necessary to use navigation satellites located in the joint radio visibility zone of two users. When determining the visible simultaneously satellites, one of the users was randomly located on the Earth's surface, and the time of its positioning was chosen equiprobably from the repeatability interval of the satellite configuration. The second user was located at some distance ΔL from the first one in an randomly direction. The time of navigation definitions on the second user could differ from the time of the first user by the value ΔT. In addition, each of the users had their own given roll. The roll direction of each user was chosen equally probable from the interval [0°, 360°]. Positioning by both users were carried out using the combined GPS+GLONASS constellation [5]. The minimum permissible elevation angle of the navigation satellites was 5°. The sample size was 100,000.

Geometric factors were calculated based on constellation included the visible simultaneously satellites by two users.

Based on the data obtained, probabilistic distributions of geometric factors were constructed for different values of the roll angles of users, with simultaneous and time- and space-shifted measurements. In addition, the number of tests in which the calculated value of the geometric factor PDOP exceeded the permissible value of six was recorded. The statistical characteristics of the obtained distributions were determined: the mean and maximum values, the standard deviation.

IV. The Composition of the GLONASS+GPS Working Constellation foe Relative Positioning Depending on the User Rolls

In the work, the number of satellites from the joint radio visibility zone of two users was determined, depending on the values of their roll angles β_1 and β_2. Figure 1 shows the probability distributions of the visible satellites number in the GLONASS+GPS working constellation with the roll of the first user β_1 equal to 20° and four values of the roll angles of the second user: $\beta_2 = 5°, 10°, 15°, 20°$. The distance between users was $\Delta L = 0$ km.

Fig. 1. Probability distribution of the number of visible satellites for two users with roll (β_1=20º, ΔL=0 km)

Figure 2 shows the probability distributions of the number of visible satellites with the same roll of the first user β_1 equal to 20° for the distance between users $\Delta L = 50$ km. As follows from Figs. 1 and 2, the distance between users of 50 km practically does not change the probabilistic characteristics of the distributions.

Fig. 2. Probability distribution of the number of visible satellites for two users with roll (β_1=20º, ΔL=50 km)

For comparison, Fig. 3 shows similar characteristics of visible satellites for cases where for $\Delta L = 50$ km and the roll β_1 of the first user is twice less and equal 10°.

Fig. 3. Probability distribution of the number of visible satellites for two users with roll (β_1=10º, ΔL=50 km)

Comparing the distributions of the number of simultaneously visible satellites for two users shown in Figs. 1–3, it can be concluded that the influence of the roll has a more significant effect on the probabilistic characteristics of the distributions than the distance between users.

V. Geometric Factors of GLONASS+GPS Working Constellations for Relative Positioning Depending on the User Rolls

For the situations listed above, the geometric factors of the GLONASS+GPS working constellations were calculated depending on the rolls and the spatial and temporal discrepancy between the measurements of the two users in relative positioning. Statistical distributions of geometric factors were constructed and statistical characteristics of the distributions were determined. Due to the large volume of the data obtained, only the most interesting of them are graphically presented. Figure 4 shows the PDOP distributions of the GLONASS+GPS working constellations for relative positioning of two users with a roll equal 20°. At the same time, in the first case, both users were spaced 100 km apart and carried out navigation definitions with a time shift of 100 seconds, and in the second case, the spatial separation was 1 km and navigation measurements were carried out simultaneously. For comparison, Fig. 4 shows the distribution of geometric factors calculated for users having the same spatial and temporal shift as in the first case, but in the absence of user rolls.

Fig. 4. Probability distribution of PDOP for relative positioning of two aircraft under rolls

Statistical characteristics of the distributions of geometric factors of the GLONASS+GPS working constellations were obtained with relative navigation definitions of two aircraft with a roll. PDOP statistics calculated for different values of

the roll angles of the aircraft β_1 and β_2 are presented in the table 1 and 2. The roll angles β_1 and β_2 corresponded to the practical tasks of approach and landing of various types of aircraft. The data from the table 1 correspond to the case when users spaced 100 km apart (ΔL=100 km) performed navigation determinations with a time shift of 100 seconds (ΔT=100 s). The last column of the table shows the number of tests n^* in which the calculated value of the geometric factor PDOP exceeded the maximum value allowed for navigational determinations, equal to six, $n^*\{PDOP>6\}$.

TABLE I.   PDOP STATISTICS FOR RELATIVE POSITIONING OF TWO AIRCRAFT UNDER A ROLL AT $\Delta L = 100$, $\Delta T = 100$

β_1	β_2	Mean	RMS	n^*
0º	0º	1.171	0.125	0
20º	20º	1.721	0.358	19

The PDOP statistics calculated for different values of the roll angles of the aircraft β_1 and β_2 for the case of ΔL=0 and ΔT=0 is presented in Table 2

TABLE II.   PDOP STATISTICS FOR RELATIVE POSITIONING OF TWO AIRCRAFT UNDER A ROLL AT $\Delta L = 0$, $\Delta T = 0$

β_1	β_2	Mean	RMS	n^*
0º	0º	1.159	0.122	0
5º	5º	1.261	0.146	0
5º	20º	1.492	0.225	1
10º	20º	1.551	0.257	4
15º	20º	1.624	0.302	6

Based on the simulation results shown in Fig. 4 and in Tables 1 and 2, it can be concluded that with the considered values of the parameters ΔL, ΔT, β_1 and β_2, the space-time parameters have significantly less influence on the distributions of geometric factors than the rolls of the first and second users.

Let's introduce a notation for the geometric factor: PDOP = PDOP (x_1; x_2; x_3; x_4), where x_1, x_2, x_3, x_4 are ΔL, km; ΔT, c; β_1, deg; β_2, deg, respectively. Then the space-time equivalent of changing the geometric factor in the absence of user rolls has the form:

$$PDOP\ (100;\ 0;\ 0;\ 0) \sim PDOP\ (0;\ 100;\ 0;\ 0).$$

Changes in the user's roll angle from 0° to 20° increase the average PDOP value by no more than 1.5 times, and the

standard deviation by 3 times. The number of rejected working constellations according to the criterion $\{PDOP>6\}$ is no more than 20 out of 100,000, i.e. no more than 0.02 %.

VI. CONCLUSION

Based on the simulation results, the permissible values of deviations of the parameters characterizing the relative positioning of aircraft in difficult conditions using two GNSS have been determined. These parameters include the distance between two users, the roll of the first user, the roll of the second user, and the time shift between user's moments of positioning. The spatial-temporal-angular equivalent of the deterioration in the accuracy of relative navigation satellite definitions during user's rolls has been determined.

These studies of the structural and information characteristics of the global navigation field created by two satellite radio navigation systems make it possible to expand the range of tasks solved during navigation of various classes of users, including the relative positioning of aircraft in difficult flight conditions. Modeling of possible situations confirms the conclusion that the information criterion used in this work is more effective for estimating information losses under abnormal conditions of satellite navigation definitions using two GNSS than the criterion based on the analysis of the number of navigation satellites in the working constellation. This criterion is directly related to the main parameter of navigation equipment – the accuracy of navigation definitions.

REFERENCES

[1] Global navigation satellite system GLONASS, Interface control document, edition 5.1, Moscow: RNII KP, 2008

[2] E.S. Basan, E.S.Abramov, A.G. Basyuk, N.A. Sushkin Metod obnaruzheniya atak na sistemu navigacii BPLA/ Informatika i avtomatizaciya, 2021, Issue 20 Vol. 6, 1368-1394. DOI: 10.15622/ia.20.6.7.

[3] Setevye sputnikovye radionavigatsionnye sistemy (Network satellite radio navigation systems), Shebshaevich, V.S., Dmitriev, P.P., Ivantsevich, N.V., et al., Moscow: Radio and communication, 1993, second edition.

[4] Sovmestnoe ispol'zovanie navigatsionnykh polei sputnikovykh radionavigatsionnykh sistem i setei psevdosputnikov (Joint using navigation fields of satellite radio navigation systems and pseudosatellite networks) Baburov, V.I., Ivantsevich, N.V., Vasilyeva, N.V., Panov, E.A., St. Petersburg, Agency «RDK-Print», 2005.

[5] Information and analysis center for positioning, navigation and timing. Available at: https://www.glonass-iac.ru.

Research of Neural Network Model for Predicting Satellite Emergency Situations Based on Telemetry Data

A. Voronov
Laboratory of System Identification
United Institute of Informatics
Problems NAS Belarus
Minsk, Belarus
voronov@lsi.bas-net.by

A. Moroz
Dept. of Information Management
Systems
Belarussian State University
Minsk, Belarus
marozas@bsu.by

P. Zhuk
Dept. of Information Management
Systems
Belarussian State University
Minsk, Belarus
paskam1note@gmail.com

Abstract—**This paper is about research of neural network model used for monitoring telemetry data of space nanosatellite. There are two main tasks: prediction of the telemetry parameters of nanosatellite subsystems and anomalies detection by an example of power supply system.**

Keywords—neural network, time series forecasting, telemetry from power supply system of nanosatellite

I. INTRODUCTION

Space telemetry is a set of techniques allowing inspecting the state of spacecraft onboard subsystems [1] and it is very important for their operation. Space telemetry data are as usual multivariate time series. As a rule, a telemetry system contains a data source, which is usually a sensor that converts the measured parameters into electrical signals. The paper considers the telemetry of a nano spacecraft of the Cubesat type. An important characteristic of any space system, particularly its orbital part, is reliability and resilience to failures and abnormal situations and the possibility of recovery. The main difficulties in solving diagnostic problems are noisy signals as well as stochastic nature of sensor signals. Some sensors can filter noise. The forecasting problem of telemetry data is difficult, but not a new one [1-2]. And new monitoring methods are needed to detect abnormal situations in telemetry data. Forecasting a multivariate time series based on the known history of sequence can predict or estimate the new number in the sequence. Recently, neural network (NN) models [3-8], statistical models and ML models have been widely used to solve the problem of forecasting a multivariate time series. There are various types of models such as the regression model, autoregressive model (AR), artificial neural network (ANN), exponential smoothing (ES), Markov chain, classification and regression trees (CART), support vector machine (SVM), genetic algorithm (GA), transfer function (TF), fuzzy logic (FL), and fundamental basic models. To improve the accuracy of the forecast, several models are used in parallel or sequentially. In a state-of-art survey for forecasting researchers, we can find the following classifications of models: univariate models (PatchTST (2023), NHits (2022), NLinear (2022), NBeats (2020)); multivariate time series forecasting (TSMixer (2023), AutoFormer (2021), TFT (2020), DeepAR (2018)); probabilistic forecasting (D3VAE (2023), TimeGrad (2021), Transformer MAF (2021), Quantile Networks (2021), spatio-temporal forecasting STDMAE (2023), STEP (2022), A3T-GCN (2021), and Evolve GCN (2019). Neural network models utilize a framework for organizing complex algorithms for transformation of data. The precision of the analysis is enhanced by incorporating the characteristics of spacecraft subsystems within the NN structure, which includes the creation of specialized hybrid NN architectures, that merge two well-known models to efficiently address ectively solve a real-life manufacturing process. The initial architecture is designed for data preprocessing and feature extraction; the subsequent architecture is for decision-making based on specific task at a hand (segmentation, identification, classification, forecasting). Different configurations of neural networks for this purpose include: multilayer perceptron, convolutional neural network (CNN), self-organizing map, long-term memory, SVM, recurrent NN, etc. Additionally it is possible to create ensembles of neural networks (ENN), which are groups of NNs that collectively make decisions by averaging the outputs of individual NNs, thereby enhancing the overall quality of the results. A model that is trained at a specific moment in time - and the decision line created by this model - might not accurately represent the resent condition of nature because of shifts in fundamental surroundings. This type of environment is known as a nonstationary environment, and the problem of learning in such an environment is often called the challenge of concept drift. Concept drift issues are typically linked to learning in a step-by-step manner, as new information is introduced over time. And the accuracy of the timeseries multivariate forecasting increases if we have good time series data quality.

II. DATA PREPARATION AND ANALYSIS

The training sample must be representative. A representative set determines the ability to generalize the results of a research of a certain sample on the entire population from which it was collected. In analysis, it is important to understand how representative the data is. This means having enough varied examples during training that accurately show the rules and patterns the model needs to identify as it learns. The following principles must be observed: sufficiency, diversity, and uniformity of class representation [8-10]. At the stage of preprocessing, scaling, transformation of nominal variables, removal of outliers, and analysis of correlation matrices are carried out. Variables must be converted to the values from -1 to 1. Outliers are removed. Often, the initial step in the analysis is to compute the correlation matrix of all variables and check for significant correlations. Perform an analysis of correlation matrix for Cubesat telemetry values with the call sign RS20S. This Cubesat is developed using the Geoscan 3U satellite base and this platform consist of power supply

system, communication system and control of the spacecraft payload, also provides conditions for the equipment operation during the flight. Telemetry data for this Cubesat we download from the website of SatNOGS project.

the problem of data processing, consider the time series of voltage values for the one month, Fig. 4. The horizontal lines in Fig.4 indicate the upper and lower thresholds of permissible voltage. To address the issue of data processing

	Current of consumption	Current from solar panels	Voltage, #1	Voltage, total	CPU Load	VoltageTotalLabel	Voltage#1Mark	Commu Number of reloads
Current of consumption	1.000000	0.035835	-0.048517	-0.046919	0.050962	-0.005847	-0.005847	-0.076728
Current from solar panels	0.035835	1.000000	-0.006756	-0.000729	-0.035100	0.007459	0.007459	-0.017550
Voltage, #1	-0.048517	-0.006756	1.000000	0.990507	0.104808	-0.442096	-0.442096	-0.490529
Voltage, total	-0.046919	-0.000729	0.990507	1.000000	0.098149	-0.429971	-0.429971	-0.510738
CPU Load	0.050962	-0.035100	0.104808	0.098149	1.000000	-0.017589	-0.017589	-0.013286
VoltageTotalLabel	-0.005847	0.007459	-0.442096	-0.429971	-0.017589	1.000000	1.000000	-0.013761
Voltage#1Mark	-0.005847	0.007459	-0.442096	-0.429971	-0.017589	1.000000	1.000000	-0.013761
Commu Number of reloads	-0.076728	-0.017550	-0.490529	-0.510738	-0.013286	-0.013761	-0.013761	1.000000

Fig.1. Correlations matrix between pairwise combinations of values from sensors related to the power supply system of CubeSat

	Current of consumption	Current from solar panels	Voltage, #1	Voltage, total	SP+X Temperature	SP-X Temperature	SP+Y Temperature	SP-Y Temperature	SP+Z Temperature	SP-Z Temperature	Cell One Temperature	Cell Two Temperature	CPU Load
Current of consumption	1.000000	0.035835	-0.048517	-0.046919	-0.008322	-0.006525	-0.008378	-0.007304	nan	-0.028530	-0.057035	-0.050974	0.050962
Current from solar panels	0.035835	1.000000	-0.006756	-0.000729	-0.514330	-0.481828	-0.550937	-0.518327	nan	-0.037186	-0.075793	-0.085686	-0.035100
Voltage, #1	-0.048517	-0.006756	1.000000	0.990507	0.118580	0.175039	0.127800	0.151619	nan	-0.031695	-0.008420	-0.030155	0.104808
Voltage, total	-0.046919	-0.000729	0.990507	1.000000	0.120092	0.176964	0.128363	0.153766	nan	-0.039067	-0.021949	-0.045129	0.098149
SP+X Temperature	-0.008322	-0.514330	0.118580	0.120092	1.000000	0.850407	0.859164	0.820989	nan	-0.010235	0.056123	0.057678	0.047287
SP-X Temperature	-0.006525	-0.481828	0.175039	0.176964	0.850407	1.000000	0.832494	0.819180	nan	0.015501	0.077382	0.075559	0.057970
SP+Y Temperature	-0.008378	-0.550937	0.127800	0.128363	0.859164	0.832494	1.000000	0.759870	nan	0.020908	0.081253	0.090267	0.063315
SP-Y Temperature	-0.007304	-0.518327	0.151619	0.153766	0.820989	0.819180	0.759870	1.000000	nan	0.011643	0.083284	0.081759	0.063110
SP+Z Temperature	nan	nan	nan	nan	nan	nan	nan	nan	nan	nan	nan	nan	nan
SP-Z Temperature	-0.028530	-0.037186	-0.031695	-0.039067	-0.010235	0.015501	0.020908	0.011643	nan	1.000000	0.466711	0.486662	-0.009178
Cell One Temperature	-0.057035	-0.075793	-0.008420	-0.021949	0.056123	0.077382	0.081253	0.083284	nan	0.466711	1.000000	0.862881	0.017353
Cell Two Temperature	-0.050974	-0.085686	-0.030155	-0.045129	0.057678	0.075559	0.090267	0.081759	nan	0.486662	0.862881	1.000000	0.013111
CPU Load	0.050962	-0.035100	0.104808	0.098149	0.047287	0.057970	0.063315	0.063110	nan	-0.009178	0.017353	0.013111	1.000000
OBC Number of reloads	-0.019073	0.015114	-0.592586	-0.616022	-0.191271	-0.273835	-0.216236	-0.224529	nan	0.046760	0.104363	0.103005	-0.014721
Commu Number of reloads	-0.076728	-0.017550	-0.490529	-0.510738	-0.167415	-0.227658	-0.191949	-0.184380	nan	0.046642	0.097817	0.093325	-0.013286
RSSI	0.002816	-0.006369	0.159126	0.163512	0.060955	0.093053	0.062804	0.073905	nan	0.004652	-0.052973	-0.039000	0.069919
VoltageTotalLabel	-0.005847	0.007459	-0.442096	-0.429971	-0.005827	-0.006752	-0.007567	-0.007453	nan	0.005397	0.003076	0.003007	-0.017589
Voltage#1Mark	-0.005847	0.007459	-0.442096	-0.429971	-0.005827	-0.006752	-0.007567	-0.007453	nan	0.005397	0.003076	0.003007	-0.017589

Fig.2. Correlations matrix between pairwise combinations of values from all sensors of CubeSat

Then construct correlations matrix between pairwise combinations of values from sensors related to the power supply system, Fig. 1. The values of the same nature correlate, as you can see in Fig. 2. If you think that diagrams are more useful, you can see a correlation diagram of voltage values with the values from other sensors in Fig. 3. To solve

and forecasting, we use the apparatus of TensorFlow library, for example, MinMaxScaler is applied. It scales the values within a given range because it preserves the relative proportions of feature values. Also, TensorFlow allows you to explicitly create a training model using the Keras library, which is very convenient for both LSTM and deep LSTM.

III. RESULTS OF TRAINING A NEURAL NETWORK MODEL

We use a dLSTM NNs as a basic NN model [9-11] and configure dLSTM with the ADAM optimizer. Then we change the following hyperparameters: the number of training cycles, the volume of the training dataset, and the sequence length during the empirical research.

Fig. 3. Example of correlation diagram of voltage sensor and signals from other sensors

Fig. 4. Example of graph of voltage in Cubesat power supply system and Predicted value

We use mean squared error (MSE) and the number of correctly identified abnormal situations as metric in our research. The maximum correct prediction of abnormal situations – 83.7 % was achieved with the following combinations of hyperparameters: the number of training cycles – 5, the volume of the training dataset – 8, the sequence length – 8 and the MSE error value is 0.02864.The graph of the obtained and predicted voltage in Fig. 4. A new direction for further research is adaptation of the developed neural network model to run on CPU or embedded devices using the quantization mechanism for the considered model.

In conclusion: the dLSTM neural network model, like other complex neural network models adapted for use in forecasting, outperform simple models. Complex neural network models are used successfully to solve problems in computer vision and natural language processing.

REFERENCES

[1] Hao Quan., "Short-Term load and wind power forecasting using neural network-based prediction intervals," IEEE Transactions on Neural Networks and Learning Systems., 2017, vol. 25, Iss. 2., ISSN: 2162-237X., pp. 303 – 315.

[2] M. Valipour, "Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir," Journal of Hydrology 476 (2013), pp. 433–441.

[3] N. Kourentzes, "Neural network ensemble operators for time series forecasting", Expert Systems with Applications, July 2014, vol.41, Iss. 9., ISSN: 0957-4174., pp. 4235–4244.

[4] R. Elwell, R. Polikar, "Incremental learning of variable rate concept drift", MCS, volume 5519 of Lecture Notes in Computer Science., Springer 2009., pp. 142–151.

[5] S. Abrar, M. Samad, "Perturbation of deep autoencoder weights for model compression and classification of tabular data," Neural Networks, 2022, vol. 156, pp. 160–169.

[6] M. Zongwei, D. Sagnik, S. Christopher, R. Liu, J. Bi, P. Balyan, Y. Liu, "A review of statistical methods used for developing large-scale and long-term PM2.5 models from satellite data", Remote Sensing of Environment, 2022, vol. 269, pp. 112–118.

[7] C. Napoli, G. De Magistris, C. Ciancarelli, F. Corallo, F. Russo, D. Nardi, "Exploiting wavelet recurrent neural networks for satellite telemetry data modeling, prediction and control," Expert Systems with Applications, 2022, vol. 206, pp. 117–131.

[8] L. Yang, T. Gao, Y. Lu, J. Duan, T. Liu, "Neural network stochastic differential equation models with applications to financial data forecasting," Applied Mathematical Modelling, 2023, vol. 115, pp. 277–299.

[9] K.D. Ashutosh, "Study and analysis of SARIMA and LSTM in forecasting time series data," Sustainable Energy Technologies and Assessments, 2021, No 47, pp. 101-114.

[10] K. Swingler, Applying Neural Networks. A practical Guide. SanFrancisco:Morgan Kaufmann, 1996.

[11] dLSTM: a new approach for anomaly detection using deep learning with delayed prediction https://link.springer.com/article/10.1007/s41060019001860. Date of access :15.04.2024

[12] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

Algorithms for Three-Dimensional Reconstruction of Electron Concentration Fields in the Ionosphere using Data from the Global Navigation Satellite System

P.A. Khmarski

Microwave electronics and radio tomography laboratory
Institute of Applied Physics of the National Academy
of Sciences of Belarus
Minsk, Republic of Belarus
https://orcid.org/0000-0003-3404-3917

A.O. Naumov

Microwave electronics and radio tomography laboratory
Institute of Applied Physics of the National Academy
of Sciences of Belarus
Minsk, Republic of Belarus
https://orcid.org/0000-0002-4624-9261

Abstract—**The features of the algorithms construction for three-dimensional reconstruction of electron concentration fields in the ionosphere based on the data of the global navigation satellite system are presented. The results of measurements of the satellite precision positioning system of the Republic of Belarus and navigation data of high-orbit navigation satellites were used as input data and GPS navigation data**

Keywords— ionosphere, radio tomography, total electron content, vertical total electron content, global navigation satellite system, navigation signal

I.   INTRODUCTION

The ionosphere is a highly ionized layer of the Earth's atmosphere, the state of which plays an important role in radio communications, radio navigation and other systems using electromagnetic waves [1]. Reliable information about the three-dimensional structure of electron concentration in the ionosphere makes it possible to improve the accuracy of positioning and navigation, as well as to provide more reliable radio communication under conditions of ionospheric variability [2, 3]. The advent of global navigation satellite systems (GNSS) has provided access to an extensive network of satellites emitting signals at different frequencies that interact with the ionosphere during their passage through it [1, 2]. At present, many effective methods for the three-dimensional reconstruction of electron concentration fields in the ionosphere are known, primarily based on the use of various modifications of the iterative algebraic reconstruction algorithm [4, 5]. In spite of this, a number of problems related to the reconstruction still remain unsolved. Among them we should mention [3, 6, 7, 8]: the choice of the optimal voxel size and geometry, the choice of the method of smoothing the reconstructed data, the justification of the input constraints and taking into account the a priori information about the state of the ionosphere, the multidimensional optimization of the parameters of the developed reconstruction structure. The proposed paper considers the developed algorithms and examples of three-dimensional reconstruction of electron concentration fields in the ionosphere based on the results of measurements of the satellite precision positioning system of the Republic of Belarus.

II.   THREE-DIMENSIONAL RECONSTRUCTION OF ELECTRON CONCENTRATION FIELDS IN THE IONOSPHERE USING DATA FROM THE GLOBAL NAVIGATION SATELLITE SYSTEM

The GNSS-based ionospheric radio tomography process requires image reconstruction in the region of interest by analyzing a subset of integral projections (Figure 1).

Fig. 1.   Principle of ionosphere radio tomography based on data from global navigation satellite systems

These projections are measured at different angles or ray paths. The mathematical inverse problems associated with image reconstruction by measurements within or around the study area are difficult because the information sought, such as the electron density distribution, is not directly available. Instead, only certain transformations or projections of this information are available. In practice, these measurements are incomplete and noisy samples of information [1, 4, 5, 10, 11]. In the case of GNSS-based ionospheric tomography, the presence of constraints such as a limited number of satellites and ground stations and a limited reception aperture makes this problem ill-posed and requires the consideration of additional a priori information for its solution.

To reconstruct three-dimensional images of the ionosphere electron density, it is necessary to carry out observations of the linear integrals of its electron density at different positions and orientations through the probed region. These linear integrals are measured in the form of total electron content (TEC), which can be written for an arbitrary moment of time t as [1]:

$$TEC(t) = \int_l Ne(t,l)dl \ , \qquad (1)$$

where Ne(l, t) is the electron density of the ionosphere, and l specifies the ray path between the satellite and the receiver. In detail, all the stages of determining the total electronic content, including the calculation of TEC by code and phase pseudo-distance, the elimination of cycle slip, and the calculation of differential code delays, are described in the authors' works [9, 10, 11].

Within the framework of the computed tomography approach, and taking into account that the electron density distribution is stable over a selected time period, the ionosphere is partitioned into a regular grid of n voxels. In the reconstruction, the electron density is assumed to be constant in each voxel, so the continuous density distribution Ne(l) is discretized by a column vector x of dimension n×1. The set of TEC measurements is expressed by a column vector y of dimension m×1. After discretization, expression (1) reduces to the form:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \boldsymbol{\varepsilon} \ , \qquad (2)$$

where A is a matrix of dimension m×n, relating the electron density distributions to the TEC measurements;

$\boldsymbol{\varepsilon}$ – vector-column of observation noise of dimension m×1.

The problem of radio tomography in this case is reduced to determining the unknown distribution of electron density in a given sensing area using the known values of TEC y measurements and the coefficient matrix A.

As a result of comparative studies, the Landweber reconstruction algorithm [6, 7], which is characterized by the lowest computational complexity and high convergence rate, was chosen as the basis for the developed reconstruction algorithm. The developed algorithm was modified for the initial data and is reduced to the calculation of the following expression:

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \beta\boldsymbol{\lambda}_k\tilde{\mathbf{A}}^T\left(\tilde{\mathbf{y}} - \tilde{\mathbf{A}}\mathbf{x}^k\right), \qquad (3)$$

where $\mathbf{x}^k$ is the vector of electron concentration values in the m×n grid nodes at the k-th iteration;

$\tilde{\mathbf{A}}$ - extended coefficient matrix consisting of matrix A augmented by a constraint matrix accounting for the smoothness of the electronic concentration;

$\tilde{\mathbf{y}}$ - extended observation vector;

$\boldsymbol{\lambda}_k$ - relaxation coefficient, which takes into account the a priori distribution of electron concentration;

β - weighting coefficient.

The realized algorithm is reduced to the following steps (see Fig. 2):

- initialization of λk with data on the a priori distribution of electron concentration;
- calculation of the vector y for a given moment of time;
- calculation of the components of the matrix A;

- calculation of the matrix with constraints B [7] and its addition to the matrix A;
- setting initial values of x0 taking into account IRI-2016 data [13] and Chapman constraints for night conditions [1] and exponential distribution [14] for day conditions, provided that there is no a priori information;
- execution of expression (3) until the required accuracy is achieved.

Fig. 2. Structural diagram of the algorithm of simultaneous algebraic reconstruction with input constraints

III. EXAMPLE OF RECONSTRUCTION OF ELECTRON CONCENTRATION FIELDS OVER THE TERRITORY OF THE REPUBLIC OF BELARUS DURING THE GEOMAGNETIC STORM ON NOVEMBER 5, 2023

Figures 3-5 show examples of operation of the algorithm for reconstruction of electron concentration fields in the ionosphere over the territory of the Republic of Belarus before, during and after the geomagnetic storm on November 5, 2023. The reconstruction was carried out at the time of maximum electron concentration (at 11:40 UTC) for three days: on the eve, during and one day after the geomagnetic storm.

The three-dimensional reconstruction was performed based on the results of the full electron content for 90 ground stations. To ensure the necessary amount of data, the time discretization step was chosen equal to 15 minutes. Figures 3-5 show examples of cross sections of the reconstructed area in three planes (in the horizontal plane at altitudes of 200, 300, and 400 kilometers, in the plane with constant latitude 54°, in the plane with constant longitude 27.5°), as well as the total vertical total electron content obtained by summing the reconstructed field by vertical columns.

It can be seen that the day before the geomagnetic storm, inhomogeneities appear in the ionosphere - a positive ionospheric storm (Fig. 3). On the day of the geomagnetic storm, the TEC value increases significantly (up to 70 TECU, Fig. 4), and the next day, there is a significant decrease of TEC (approximately by a factor of 3, Fig. 5), which is referred to in the literature as a negative ionospheric storm.

Fig. 3. Reconstruction of the electron concentration field before a geomagnetic storm.

Fig. 4. Reconstruction of the electron concentration field on the day of a geomagnetic storm.

Fig. 5. Reconstruction of the electron concentration field one day after a geomagnetic storm.

IV. CONCLUSION

The features of the construction of algorithms for three-dimensional reconstruction of electron concentration in the ionosphere are considered. The main scientific problems that arise in the implementation of this algorithm are: the choice and justification of the size and geometry of the voxel; the choice of the method of algebraic reconstruction and the used method of stopping the reconstruction procedure; the choice of the method of smoothing the reconstructed data; the consideration of a priori information in the reconstruction algorithms; the optimization of the parameters of the chosen reconstruction structure.

REFERENCES

[1] Kunitsyn V.E., Tereshchenko E.D., Andreeva E.S. Radiotomography of the ionosphere. Moscow: Fizmatlit, 2007. 693 p.

[2] Hofmann-Wellenhof B., Lichtenegger H., Wasle E. GNSS – Global Navigation Satellite Systems. GPS, GLONASS, Galileo, and More. Springer, 2008. xxix, 516 p.

[3] Materassi M., Forte B., Coster A., Skone S. The Dynamical Ionosphere a Systems Approach to Ionospheric Irregularity. Elsevier, 2020. 323 p.

[4] Artemiev V.M., Naumov A.O., Radio tomography of the electron concentration field in the ionosphere based on the Kalman filter, Proceedings of the National Academy of Sciences of Belarus. Physical-technical series. 2012. 2. C. 86–92.

[5] Artemiev V.M., Naumov A.O., Stepanov V.L., Murashko N.I. Method and Results of Real Time Modeling of Ionosphere Radiotomography on the Basis of the Kalman Filter Theory // Journal of Automation and Information Sciences, 40(2), 2008, p. 52–62.

[6] Herman G.T. Fundamentals of Computerized Tomography, Image Reconstruction from Projections, Springer, New York, 2009. 297 p.

[7] Hobiger T., Kondo T., Koyama Y. Constrained simultaneous algebraic reconstruction technique (C-SART) – a new and simple algorithm applied to ionospheric tomography. Earth Planet. Sp 60. 2008. P.727–735.

[8] Wen D., Liu S., Tang P. Tomographic reconstruction of ionospheric electron density based on constrained algebraic reconstruction technique. GPS Solut. 14. 2010. P. 375–380.

[9] Naumov A., Khmarskiy P., Byshnev N., Piatrouski M. Methods and software for calculating total electronic content based on GNSS data. 7th Advanced Engineering Days, 1-2 July. 2023. Mersin, Türkiye. 2023. P.158-160.

[10] Naumov A., Khmarskiy P., Byshnev N., Piatrouski M. Methods and software for estimation of total electron content in ionosphere using GNSS observations. Engineering Applications. Vol. 2. no.3. P. 243–253.

[11] Naumov A., Khmarskiy P., Byshnev N., Piatrouski M. Determination of total electron content in the ionosphere over the territory of the Republic of Belarus based on global navigation satellite systems data. Proceedings of the National Academy of Sciences of Belarus. Physical-technical series. 2024. Vol. 69, № 1. P. 53–64.

[12] Ivanov, V.B., Zatolokin, D.A. & Gorbachev, O.A. Comparing models of total electron content in the ionosphere for GLONASS. Gyroscopy Navig. 2017. 8. P. 295–299.

[13] International Reference Ionosphere. URL: https://ccmc.gsfc.nasa.gov/modelweb/models/iri2016_vitmo.php. Дата обращения 07.06.2023.

[14] Stankov S. M., Stegen K., Muhtarov P., Warnant R. Local ionospheric electron density profile reconstruction in real time from simultaneous ground-based GNSS and ionosonde measurements. Advances in Space Research. 2011. Vol. 47. no. 7. P. 1172-1180.

Collaborative Positioning with Global Navigation Satellite Systems

S. Dolin

Department of Space and Physical Geodesy
Siberian State University of Geosystems and
Technologies
Novosibirsk, Russia
ORCID: 0000-0002-0686-2272

L. Lipatnikov

Department of Space and Physical Geodesy
Siberian State University of Geosystems and
Technologies
Novosibirsk, Russia
ORCID: 0000-0002-5082-4062

Abstract— **The article is devoted to the implementation and experimental study of the concept of collaborative positioning with global navigation satellite systems intended to provide global coverage of a high-precision navigation field while significantly reducing the load on the telecommunications network and ensuring the initialization time and accuracy of determining coordinates characteristic of the RTK method for satisfying modern requirements of new categories of high-precision GNSS users.**

Keywords—GNSS, PPP, collaborative positioning, extended Kalman filter, moving-baseline

I. INTRODUCTION

The present study dwells upon a concept of collaborative positioning with global navigation satellite systems (GNSS) proposed by Chris Rizos, Joel van Cranenbroeck et al. [1]. The concept implied that each GNSS user could provide corrections in real time to other users so that together they would create a new ground-based dynamic infrastructure for high-precision positioning. The concept can be developed further and implemented differently, depending on the particular ways the corrections are distributed and applied.

Implementation of the concept can be especially promising in the context of combining Precise Point Positioning (PPP) method and relative real-time positioning method known as Real-Time Kinematic (RTK), comprising the benefits of both methods, which is considered in the present study.

The potential key benefits are global coverage of high-precision seamless navigation field; reduction of PPP convergence time to RTK initialization time, decreasing data traffic in comparison with transmitting observations space representation (OSR) corrections form continuously operating reference stations (CORS) to all users constantly [3; 4].

Process of collaborative positioning must be fully automatic to satisfy the needs of new categories of users and applications, such as robots and users of mass market positioning-based applications, described in [5]

II. IMPLEMENTATION OF COLLABORATIVE POSITIONING

PPP is proposed to be used as the main method to provide global coverage and minimize the data traffic, and RTK corrections can be provided by one user to another on demand for fast initialization or reinitialization of PPP Kalman filter.

Implementation of collaborative positioning implies changing the way GNSS receivers interact with each other. Thus "base" and "rover" are no longer considered as constant receiver types but as receiver's current role or status. Apart from them a new "candidate" status has been added. The following scheme of interaction between them is considered.

"Rover" is the default status. It is assigned to the user's navigation equipment (UNE) immediately after it is turned on or when its positioning accuracy is poor or not known. After getting this status, rover sends its approximate coordinates and a request for RTK correction to the cloud data center (CDC).

"Candidate" status is assigned to the UNE if standard deviation of its coordinates is below a certain level (20 cm for test purposes) and if the user has given consent to transfer one's high-precision coordinates to other users.

"Base" status is assigned to a candidate by CDC if there is any rover nearby that require corrections. Base transmits a stream of OSR corrections - measurement data and high-precision coordinates in RTCM3 format to rover through the CDC.

Refinement of the user's location is carried out by partial substitution of *a priori* information in the extended Kalman filter of the PPP method with values, obtained from RTK solution in real time.

The scheme of interaction between UNE and CDC implemented in a prototype as presented in Fig. 1.

Fig. 1. Prototype service of collaborative positioning.

Platform as a service (PaaS) implementation has been chosen to provide scalability and reliability. CDC in the form of PaaS allows deploying system on different virtual machines within one application, operating in one cloud service. User acess is provided through dedicated public IP adress. Automatic load balance is performed. Cloud platform is scalable. Virtual machines are deployed in such a way to eliminate a single point of failure.

The CDC data base contains information about UNE status, coordinates in ITRF, their standard deviations, kinematic status (still or in motion), receiver type.

The decision if a UNE needs RTK correction to (re)initialize PPP Kalman filter is made according to the following criteria:

1) rover's PPP solution has not converged (approached the highest level of precision),

2) available alternative (RTK solution) is significantly more precise than PPP solution.

The first criterion is based on imperical observation that in PPP kinematic mode the standard deviation of each coordinate converge finally to a level of few centimeters and do not change significantly after that. For that purpose the rate of standard deviation decrease is compared to imperical value $L = 0.02$ for data sampling interval 1s , as in (1):

$$\frac{\left|SMA_{m_t} - \overline{\sigma}_{PPP_t}\right|}{SMA_{m_t}} > L \qquad (1)$$

where SMA_{m_t} – simple moving average of average standard deviations with the sampling width m at time

$t : SMA_{m_t} = \frac{1}{m} \sum_{i=n-m+1}^{n} \overline{\sigma}_{PPP_i}$;

$\overline{\sigma}_{PPP_t}$ – average standart deviation at moment t :

$\overline{\sigma}_{PPP_t} = \frac{\sigma_X + \sigma_Y + \sigma_Z}{3}$.

If the result of the condition (1) holds, it is considered that PPP solution has not converged. Then OSR corrections are requested from the most apropriate candiate, turning it into base, in order to calculate RTK solution for the rover as an alternaitve to PPP solution.

OSR corrections from base include GNSS measurements and precise coordinates in RTCM3 format, which are used in relative real-time positioning based on double differrence (*DD*) phase measurements model: regular RTK if the base ia a CORS or in moving-baseline mode if the base is moving. In either case coordinates or trajectory of the base in real time are expected to be highly accurate. In both cases we further refer the acquired solution as RTK solution.

The second criterion is applied to check if RTK solution is more precise than previously acquired PPP solution, as in (2):

$$\overline{\sigma}_{PPP_t} > \overline{\sigma}_{DD_t} \qquad (2)$$

where $\overline{\sigma}_{DD_t}$ – average standart deviation of RTK estimates of coordinates at time t :

$\overline{\sigma}_{DD_t} = \frac{\sigma_X + \sigma_Y + \sigma_Z}{3}$;

$\overline{\sigma}_{PPP_t}$ – average standart deviation of PPP estimates of coordinates at time t .

When both criteria are met, PPP solution is recalculated as follows. A priori coordinates and their variance at prediction step for the current moment t are substituted with values from RTK solution, after that PPP measurement update is recalculated. This allows to draw PPP estimates of coordinates closer to RTK solution at moment t, and, what is more important, it facilitattes gaining weight of current estimates of phase ambiguities and wet troposheric dealy, so that PPP solution can converge faster and become stable.

III. EXPERIMENTS

To evaluate performance of the collaborative positioning preliminary pseudo-kinematic and kinematic tests were carried out. Pure RTK, PPP, and collaborative positioning solutions were calculated simultaneously in real time and then compared. For the PPP method, real-time differential code biases from the SSR correction stream were considered [1; 6].

Pseudo-kinematic tests were caried out using International GNSS Service and EUREF CORS providing real-time data streams as rovers. Pseudo-kinematic mode means that though sites are Earth-fixed, processing is carried out as they were mobile. Known positions at current epoch were obtained from ITRF2020 catalogue and used in accuracy evaluation for calculating root mean square error using Gauss formula, as in (3):

$$m = \sqrt{\frac{\left[\Delta^2\right]}{n}} \qquad (3)$$

Results of accuracy assessment for four ground points are presented in Table 1. The duration of the experiment was one day, including six four-hour sessions for each station. Data sampling rate is one second.

TABLE I.   ACCURACY ASSESSMENT FOR GROUND POINTS, M

Point / distance to "Candidate" (km)	Methods		
	Collaborative positioning	*PPP*	*RTK*
WGTN / 320 km	E: 0.17 N: 0.14 U: 0.30	E: 0.16 N: 0.15 U: 0.45	E: 0.51 N: 0.40 U: 0.78
BORJ / 80 km	E: 0.12 N: 0.09 U: 0.29	E: 0.20 N: 0.22 U: 0.51	E: 0.09 N: 0.07 U: 0.15
KARL / 125 km	E: 0.12 N: 0.10 U: 0.20	E: 0.19 N: 0.14 U: 0.38	E: 0.19 N: 0.16 U: 0.26
RDSD / 550 km	E: 0.17 N: 0.17 U: 0.37	E: 0.18 N: 0.20 U: 0.64	E: 2.48 N: 1.40 U: 3.90

Also, the preliminary kinematic experiments in urban conditions were carried out with two receivers (UNE sets) installed on a car roof as shown in Fig. 2. The first UNE set capable of implementing collaborative positioning included the u-Blox ZED F9P microcircuit [2] and the Harxon HX-CSX601A dual-frequency survey antenna. The second UNE set included a surveying grade receiver (Stonex S800A), which was used for acquiring a reference trajectory.

Fig. 2.   Mobile installation

Based on the results of the test, a positive effect is noted associated with fast reinitialization of the solution after passing under bridges and within tunnels, as in Fig. 3

Fig. 3.   Restoration of the solution (White point: collaborative positioning; Green point: reference track from relative method)

Results of accuracy assessment for mobile receiver from collaborative positioning, PPP and RTK methods, presented in table 2. RMS estimates also include 15.7 cm shift in horizontal plane between the two antennae phase centers. Vertical shift was corrected.

TABLE II.   ACCURACY ASSESSMENT FOR MOBILE RECEIVER

Type of accuracy	Methods		
	Collaborative positioning	*PPP*	*RTK*
RMSE (68,27%), m	*E: 0.40* *N: 0.56* *U: 1.03*	E: 0.88 N: 1.35 U: 1.39	E: 0.39 N: 0.60 U: 3.16

IV. RESULTS

Based on the results of the experiments, we should recognize the positive effect with the reduction of the convergence time and improvement of accuracy to a level comparable with RTK. And also, the effect of fast reconvergence of the solution after the car passes under an obstacle was noticed; the tracking is restored almost instantly and the accuracy improves significantly if compared to PPP. At the same collaborative positioning time proved to be suitable for application at long distances from base stations, unlike RTK. More kinematic tests is planned.

V. DISCUSSIONS

The results of the preliminary tests of collaborative positioning show that it may be used to achieve global coverage of high-precision navigation field while reducing the initialization time to same of RTK as well as for the case of reinitialization of the solution after an interruption in the first seconds of receiving new measurements. It also can be fully automated such to be suitable for applications such as precise navigation of robots and new mass-market positioning-based applications.

REFERENCES

[1] Dolin, S.V. Application of Differential Code Biases in Multi-GNSS Measurements in Real-Time Precise Point Positioning. Gyroscopy Navig. 13, 276–282 (2022). https://doi.org/10.1134/S2075108722040034

[2] J. van Cranenbroeck and C. Rizos, "The "uberization" of the GNSS Positioning Infrastructure", FIG Working Week 2019, Hanoi, Vietnam, April 22–26, 2019.

[3] A.P. Karpik, A.V. Mareev and D.S. Mamaev, 'Free software for geodetic monitoring Moncenter", Vestnik SSUGT (Siberian State University of Geosystems and Technologies). 2022. Vol. 27. no. 5. – P. 43-54. – DOI: 10.33764/2411-1759-2022-27-5-43-54

[4] G. Wübbena, M. Schmitz and A. Bagge "PPP-RTK: Precise Point Positioning Using State-Space Representation in RTK Networks", Presented at the 18th International Technical Meeting, ION GNSS-05, September 13-16, 2005, Long Beach, California, DOI: https://doi.org/10.1007/s10291-023-01428-2

[5] V. Janssen, "A comparison of the VRS and MAC principles for network RTK" International Global Navigation Satellite Systems Society IGNSS Symposium 2009, Holiday Inn Surfers Paradise, Qld, Australia 1 – 3 December, 2009, pp 1-13.

[6] Xingxing Wang, Sheng, C., Yu, B. et al. Rapid Re-Convergence of Real-Time Dynamic Precise Point Positioning by Adding Velocity Constraints. Gyroscopy Navig. 13, 283–293 (2022). https://doi.org/10.1134/S2075108722040125

[7] EUSPA EO and GNSS Market Report 2024 issue 2, EUSPA 2023, ISBN 978-92-9206-079-4, DOI: 10.2878/73092

Algorithms for Localization and Matching of Traffic Lights in Driver Assistance Systems

N.S. Guzhva

Postgraduate student of NITU MISIS,
Cognitive Technologies
Moscow, Russia
n.guzhva@cognitivepilot.com

R.N. Sadekov

Professor of the Department of NITU MISIS,
Cognitive Technologies
Moscow, Russia
0000-0001-6286-358X

Abstract – The paper considers the task of identification (detection, filtering and map-matching) of traffic lights on an image sequence in tram driver assistance systems. The solution of the problem is described from the formation of 3D measurements from the output of a neural network (detections) for each camera to the comparison of the results with the map and planning to pass the intersection. The novelty of the work is the use of 3D measurements, the UKF filter to process the results of detections, and the procedure of combining the results of two cameras. The presented algorithm and its modifications are evaluated on real data, proving its effectiveness.

Keywords: computer vision, identification, traffic light recognition, driver assistance system

I. INTRODUCTION

Progress in the development of unmanned technology [1-3] has led to the introduction of these technologies and on public transport, recently on the streets of St. Petersburg began to operate systems to assist the driver of the streetcar [4, 5].

Traffic light recognition system is an integral part of autopilots. When driving on city streets, a car operating in robot mode must recognize traffic lights when approaching an intersection, classify their signals, and compare the results with a map and plan. Thus, the system should perform a complete stop of the tram (at the stop line) in case the driver continues to move at a prohibited traffic signal.

Despite the rapid development of deep learning algorithms, the solution to the problem of recognizing traffic lights is still fraught with many difficulties due to changing weather conditions and traffic at different times of the day, a wide variety of signals and types of traffic lights, "heap" placement of traffic lights at intersections and the presence of numerous "false" light sources (lighting, car headlights, etc.), the location of traffic lights in different parts of the road scene, a decrease in the brightness of signals during prolonged operation.

The above circumstances require the development of complex algorithms for recognizing traffic lights in the image, filtering them and comparing them with the map. The necessity of using a map is explained by the task of selecting a traffic light or a group of traffic lights that regulate the passage in a given direction at an intersection. The information about traffic lights is usually stored in digital maps. In this paper, it is assumed that a map of traffic lights location, can be obtained by digitizing a satellite image and does not require expensive "lidar" markup.

Unlike the known ones, in this paper, the measurements are 3D detections of the neural network obtained from the image of each camera. The cameras used have significantly different internal parameters, which leads to the need to process measurements with different accuracy of their filtering and comparison with the map.

The paper consists of several sections, which describe the problem statement, the stages of its solution: detection of traffic lights in the image, filtering of the obtained measurements using UKF filter, comparison of measurements from each camera, determination of the signal class, comparison of the traffic light position with the map, and evaluation of the accuracy of the proposed approach.

II. ANALYSIS OF EXISTING APPROACHES

The traffic light detection problem is usually formulated and solved on an image, and a priori map information can be used. The information from the digital map is used to constrain the search area of traffic lights in the image. Detection and classification networks, which are currently implemented based on deep neural networks, are used to detect the position of a traffic light and determine its signal.

The problem of traffic light detection can be solved for both night [6] and daytime [7] conditions, to take into account the features associated with the small size of the traffic light in the image [7]. Use information from the map [8], [9]. In addition, lidars [10], stereo cameras [6] or multiple cameras [8] can be used to solve the problem of map matching, data projection and reprojection. Problem formulations can also consider different locations: urban environment [11] or freeway [7]; types of traffic lights: automobile or tram [9].

The overall solution of the problem is usually composed of solving the following subproblems:

- detection and classification of traffic signals on the available image;

- tracking of traffic lights between frames and between sensors;

- filtering of measurements and comparison of the obtained measurements with the map.

A brief review of works on the proposed topic is presented in Table 1.

Source	Sensors and features	Detection	Tracking and filtering	Digital map
[8]	Stereo camera, odometry and inertial module. The traffic lights in the image are extremely small in size	YOLO, traffic lights with arrows, separate neural network for classification	Based on a separate neural network	Not used
[9]	Lidar and three cameras. Unmanned tram. Suburban.	SSD detector. Tram traffic lights	Hidden Markov model. Cross-camera processing	High definition lidar map, using lidar to generate search areas
[10]	Lidar, one camera	YOLO, two types of traffic lights	Accumulation of detections in buffer, use of clustering	High definition lidar map, using lidar to generate search areas
[12]	Multiple cameras with different focal lengths, GPS, inertial module, urban environment	YOLOv5 for each camera, automobile traffic lights	Hungarian algorithm and hidden Markov model	Digital map to generate search areas
[13]	Stereocamera, GPS	Neural network	Kalman filter and Hungarian algorithm for traffic light map formation	Not used

The analysis of works has shown that the solution of the traffic light recognition problem is based on multistage algorithms. In this case, the currently implemented algorithms use information from two or more identical (to form stereo) cameras or lidars, which significantly facilitates and increases the accuracy of the problem solution. At the detection stage, a priori information from the digital map is most often used, which is also obtained on the basis of high-precision measurements. However, the known works do not include solutions using cameras with significantly different focal lengths, new neural network architectures capable of generating 3D information about the road scene. In addition, the solutions are rarely tested in urban environments (as the most difficult case), and are not related to the task of passing an intersection based on the detection of only relevant traffic lights.

III. PROBLEM STATEMENT AND STAGES OF ITS SOLUTION

The tram driver assistance system consists of several cameras installed in the tram driver's cabin. Each camera is associated with a coordinate system xyz centered at the point C. Cameras are considered calibrated, with known internal $K_m^{<cam/vr>}$ and external $R_m^{<cam/vr>}, T_m^{<cam/vr>}$ installation parameters relative to the tram coordinate system, where m is the camera number, Figure 1. The beginning of the tram coordinate system (hereinafter vehicle related - VR) coincides with the center of the front bumper and is on the ground. In the VR coordinate system the primary analysis of the road scene is performed. In addition, the tram's navigation system determines its position and orientation in the global metric coordinate system (GM). GM coordinate system allows to determine the position of any object in space based on the vector $s_t^{gm} = [x_t, y_t, z_t, \alpha_t, \beta_t, \gamma_t]$, where x_t, y_t, z_t – object coordinates, $\alpha_t, \beta_t, \gamma_t$ – orientation angles. It is obvious that for the transition from cam to GM coordinate system, it is necessary to sequentially perform the transitions from cam to VR and from VR to GM.

The coordinates of the traffic lights in the GM coordinate system are known in advance. We denote them as a vector $o_b^{gm} = [x, y, z, \gamma]$, where b – traffic light index, x, y, z – coordinates and γ – yaw. At each intersection, the traffic lights are grouped into groups. Under the group of traffic lights is understood a set of traffic lights (main and duplicate) that regulate traffic at the intersection and have a common stop line for stopping a vehicle. To reduce the designations we will not introduce an additional index of the group of traffic lights.

Recognition of traffic lights is performed by a neural network based on the images formed by each camera. The network is multi-headed (multitasking). Each of the network heads has its own purpose. We can distinguish: solving the problem of segmentation, 3D detection, embedding formation and others. Figure 2 shows the result of the network. The architecture of such a network is described in detail in [4], so it is not considered here.

Figure 1. Image of tram and VR coordinate system (a)

At each moment of time, the neural network receives an image I_t^m as input. The neural network finds objects: vehicles (including cars and trucks, public transport), pedestrians, traffic signs and traffic lights and determines their spatial position described by bounding boxes (further 3d detections or simply detections)

Figure 2. Visualization of road scene recognition (3D detections - cars (yellow), traffic lights (red), segmentation area - curbs (red), tramway (turquoise))

We denote the detection of traffic light j at time t, for camera m as $r_j^{t,m}$. The measurement vector $r_j^{t,m}$ includes:

- coordinates (pixels) [u, v] of the traffic light center in the image I_t^m;
- depth of the traffic light center d in the camera coordinate system, meters;
- orientation in euler angles (pitch, roll, yaw);
- physical dimensions of the object: width (w), length (l) and height (h), meters;
- unique identifier of the traffic light area - embedding (E).

All the above parameters except u, v are generated in the cam coordinate system.

It is required to determine the correspondence between the *j*-th detection and the ordinal number of the traffic light b, based on the measurements $r_j^{t,m}$. The following steps are performed to solve the task:

Tracking - matching traffic light detections between frames for each camera separately, i.e., finding the correspondence between $r_{j1}^{t,m}$ and $r_{j2}^{t+1,m}$.

Combining detection results from different camera for the same traffic light into a single measurement, i.e., finding a correspondence between $r_{j1}^{t,m}$ and $r_{j2}^{t,m+1}$;

Traffic light position filtering - determining the coordinates of traffic lights in the GM coordinate system by several measurements;

Comparison of the obtained coordinates based on the filter results with the traffic lights available on the map;

Classification - determining the class of a traffic light signal from the image to make a decision about the possibility of passing;

Let's consider the above stages in more detail.

IV. TRACKING

In modern computer vision tasks tracking (tracking) of objects is performed using embeddings. Embeddings are understood as a vector of features stable in time and space, which characterizes an object (region) in the image. The neural network used in this paper generates an embedding - E_j - for each *j*-th detected object (traffic light). The dimensionality of the embedding vector is 128. To determine the fact of similarity of two areas in the images at the moments of time t and t-1 a cosine similarity measure is used, given by the expression of the following form

$$C = \frac{E_i E_j}{|E_i||E_j|}$$

where E_i, E_j - the embeddings of objects i and j at times t and t-1.

The larger the value of C to one, the higher the similarity between the embeddings and as a consequence the more clearly the two traffic light regions in the image sequence correspond to each other. To find the matching pairs between i and j. We perform pairwise computation of the cosine measure for all measurements obtained at time t and t-1 and select pairs with maximum C values.

V. MERGING OF DETECTION RESULTS BETWEEN CAMERAS

In order to solve the traffic light recognition problem reliably, it is necessary to recognize both main and duplicate traffic lights at the same time. In addition, it is necessary to recognize traffic signals in advance to ensure timely braking of the vehicle if necessary. To this end, the project uses two cameras, with different focal lengths, differing by a factor of about two. One of the cameras is set to work with distant objects, the second with near objects. Images from each camera are independently processed by a neural network, so the results of recognizing objects and traffic lights obtained from the two cameras must be combined, i.e. find a match between them. Since tracking is most often understood as a procedure of inter-frame tracking rather than inter-camera tracking, the term "merging" is used to avoid confusing the reader. The data pooling problem is not trivial due to the rather large number of variables and can be solved using the following distance functions:

- Cosine embedding distance. This technique is widely used in practice and does not depend on the camera settings. However, embeddings may coincide for similar objects that are in different parts of the image and the embedding may change from frame to frame (e.g. due to flashing traffic lights).

- A measure of Intersection over Union (IoU) 2D by transferring the recognition results from one image to another. An understandable and geometrically well interpretable metric that is widely used in practice. However, its use requires highly accurate calibration parameters for both cameras and depth knowledge for transferring detection regions.

- Computation of IoU for 3D detections. The neural network outputs 3D detections, so it is possible to use this information for intersection verification and not be tied to the image. Obviously this reduces ambiguity when objects in the frame are side by side. However, because the depth of objects from different cameras may differ greatly (due to the large difference in focal length), there may be no intersection of cuboids in space at all, or on the contrary there may be a false intersection.

- Euclidean distance (L2) between centers of objects (in space). An understandable and popular approach, but due to errors in predicted depth, objects may also not coincide.

- Mahalanobis distance, which is similar to Euclidean distance, but allows to take into account the error in the depth of the objects for each camera. This approach allows to take into account the covariance (noise) of depth determination by the neural network.

- Distance between the centers of 2D detections by transferring the coordinates of the centers of selected regions to one image and calculating the distance. In some cases it is difficult to take into account the difference in pixels for distant objects, because the size of the boxes changes greatly in one image relative to the second image. In addition, the metric depends on the parameters of both cameras.

As can be seen from the analysis, there are quite a lot of ways to "combine" measurements. To choose the optimal strategy of merging measurements or their combinations it is necessary to proceed from the following factors: metrics should produce normalized values in the range from 0 to 1, when using several metrics it is necessary to determine the coefficients of their influence..

Based on the experiments performed, a final metric was created that combines data on the mutual location of the centers of 2D objects, embeddings and IoU of 2D objects. With the latter two components, weights are introduced to adapt their contribution to the final metric. Thus, the final correspondence metric (w) can be calculated, based on a weighted sum of the cosine sum of the distance between embeddings (d_emb) and the inversion of the IoU index (1 - iou), with the importance of each component determined by the weighting factors (w_emb and w_iou, respectively). In doing so, the equation takes into account a threshold value for the distance in pixels between the centers of 2D boxes (d_pixel) to screen out further false comparisons. Through numerous experiments, the weights were found to be in the range: weight w = 0.6...0.7 for embeddings, weight w = 0.3...0.4 for IoU.

VI. OPTIMAL FILTERING OF TRAFFIC LIGHTS LOCATION COORDINATES

The results obtained directly from a neural network often contain a significant level of random inaccuracies or "noise", which requires additional filtering. The Kalman filter is used for this purpose. The filtering of traffic light positions can be performed using a UKF filter based on the following equations.

The equation of state contains a constant model

$$s_{t+1}^{gm} = s_t^{gm} + w_t^{gm}$$

where w_t^{gm} – model noise, $s_t^{gm} = [x_{gm}, y_{gm}, z_{gm}, w, l, h, yaw_{gm}]$ – state vector.

The measurement equation has the form

$$z_t^{vr} = H(s_t^{gm}) + v_t^{vr}$$

where $z_t^{vr} = r_j^t = [u_{cam}, v_{cam}, d_{cam}, yaw_{vr}, w, l, h]$. The transition from the state vector to the measurement vector is performed using the internal camera calibration parameters and matrices $R_m^{<gm/vr>}, T_m^{<gm/vr>}$ determined by the navigation system.

VII. MAP MATCHING

Based on the results of the filter operation, a vector of traffic lights coordinates is formed, which is further compared with the information on the location of traffic lights available in the digital map.

The digital map is a database that stores the coordinates of traffic lights location, route sections they regulate, classes of signals to allow passage of route sections and other service information.

The comparison of the vectors of coordinates of traffic lights locations with the calculated values is performed using the Mahalanobis distance, which is determined by the following equation

$$d_{ib} = \sqrt{(s_i^{gm} - o_b^{gm})P^{-1}(s_i^{gm} - o_b^{gm})}$$

where o_b^{gm}- coordinate vector for the b-th traffic light from the map, s_i^{gm}- coordinate vector for the i-th detected traffic light, P^{-1} - covariance matrix for the detected traffic light. A "cost" matrix with values of a metric inversely proportional to the Mahalanobis distance is compiled for the distances found. Next, the best measurement-map pairs are determined using the Hungarian algorithm.

VIII. CLASSIFICATION OF TRAFFIC LIGHT SIGNALS

The classification problem is solved using a deep neural network. The input of the network is a fragment of traffic light image formed by the detector, and the output of the network is the class of traffic light signal. The traffic light signal classifier contains 25 classes. The network architecture consists of two sequential blocks: feature extraction and class prediction. A neural network of RegNetY architecture [17] is used for feature extraction. Since the data in the collected dataset is not balanced, the following Focal Loss [18] is used in training such a network, which allows dynamic scaling of the loss of the cross-entropy function. The scaling factor decreases to zero as the confidence of the neural network in the correctness of the predicted class increases.

Markup of image fragments containing traffic lights uses a system of classes that takes into account: type of traffic lights (automobile, tram, pedestrian), orientation (direction), number of sections (5-section, 3-section, 2-section), as well as special symbols and their combinations (straight, left, straight and left, etc.).

Examples of traffic lights of different classes are shown in the figure.

Figure 3. Some classes of traffic light signals

IX. ACCURACY ASSESSMENT

The accuracy of the system operation is evaluated in conjunction with the control system [14], [15]. Such

experiments are primarily required for autopilot solutions, as they determine the speed and safety of the robot car movement on city streets. Usually, test routes containing intersections are used for these purposes. In this case, the result of correct recognition of traffic signals, stopping at the intersection in case of a forbidden signal and driving through a permissive traffic signal is evaluated. Unfortunately, it is not possible to automatically evaluate the accuracy of a driver assistance system with active traffic signal control because drivers often stop the tram before the stop line and start driving when the permissive traffic signal is illuminated. Therefore, the accuracy of the system was assessed using data recorded when the tram was running on routes and manual checks of normal and abnormal operation.

A comprehensive evaluation of the system can be determined using the target and actual system actions as a function of the traffic signal:

– STOP at stop signal;

– MOVING at permissive signal.

In the context of passing an intersection, the most critical action is STOP, so it is taken as the basic ("positive" from the point of view of the inconsistency matrix). These events allow us to form the following matrix of situation correspondence, presented in Table 2.

TABLE 2. MATRIX OF EVENT CORRESPONDENCE

EVENT	Target action	Actual system action
TN	MOVING	MOVING
TP	STOP	STOP
FN	STOP	MOVING
FP	MOVING	STOP

Thus, the following events are generated: TP - stopping at a forbidden signal; TN - moving at a permissive signal; FP - stopping at a permissive signal; FN - moving at a forbidden signal, the most critical error from the point of view of traffic safety, since it allows moving at a red light.

The accuracy evaluation was performed by manually analyzing the video data and the data attributed by the telemetry system for the routes presented in Table 3. The accuracy was evaluated for two algorithms:

– using one camera and UKF filter to filter and match traffic lights (A1).

– using two cameras and a UKF filter to filter and match traffic lights (A2).

TABLE 3. INFORMATION ON THE ROUTES USED FOR ACCURACY ASSESSMENT

Routes	Route length, km
36 route	8 runs, 17.43 km
24 route	3 runs, 14.67 km
27 route	1 run, 18.04 km

TABLE 4. ACCURACY EVALUATION OF THE ALGORITHMS

| Events | Algorithms | |
	A_1	A_2
TP	181	192
TN	183	180
FP	57	60
FN	16	5
Precision	76.05 %	76.19 %
Recall	91.88 %	97.46 %

Table 4 presents the results of the system accuracy estimation. Both basic TP, TN, FP, FN and derived parameters Precision, Recall are calculated. These parameters are calculated using standard equations

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

Based on the logic of the problem being solved, the Recall parameter is the highest priority, because passing the intersection on a red traffic signal is critical. From the presented analysis we can see that when using two cameras, the quality of the system improves in terms of completeness by 6%, which is explained by a better ability to detect distant "duplicate" traffic lights.

Figure 4 shows the results of detection of traffic lights in different locations of St. Petersburg. Bounding boxes show all objects recognized by the neural network, with cars in orange, traffic lights in red and green, blue dots - position of traffic lights on the map. It can be seen that the objects recognized by the neural network have a connection with the points on the ground, which indicates that the matching task has been successfully executed.

Figure 4. Fragments of visualization of traffic lights recognition results

X. CONCLUSION

The problem of traffic light recognition has been addressed for quite a long time, and there is currently no absolutely reliable solution due to a large number of external factors affecting the quality of recognition. The use of multiple cameras/lidars, deep neural networks and high-precision maps significantly improve the quality of the problem. The paper describes an algorithm that solves the problem of recognizing traffic lights in real time with high accuracy in modern tram driver assistance systems. The evaluation of the accuracy of the proposed

identification algorithm together with the tram control system using two cameras allows to achieve the parameters of accuracy 76.19% and completeness 97.46%. The existing errors of the system are currently associated with the recognition of primarily tram traffic lights (due to their low expressiveness and burnout during operation). Implementation and use of this algorithm in practice allows to reduce the requirements to the map markup and provides high recognition accuracy in real time.

REFERENCES

[1] D. I. Smolyanov and A. A. Golovan, "On the Navigation Problem of Unmanned Wheeled Agricultural Machinery Using MEMS-Based INS, GNSS and Odometry," 2023 30th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), Saint Petersburg, Russian Federation, 2023, pp. 1-3, doi: 10.23919/ICINS51816.2023.10168339.

[2] Ali, B., Sadekov, R.N. & Tsodokova, V.V. A Review of Navigation Algorithms for Unmanned Aerial Vehicles Based on Computer Vision Systems. Gyroscopy Navig. 13, 241–252 (2022). https://doi.org/10.1134/S2075108722040022.

[3] Golovan A.A., M.N. Drobyshev, D.I. Smolyanov "Decomposed algorithm for processing raw satellite navigation measurements // 2022. Т. 23. № 4. P. 23-35. DOI: 10.21455/gr2022.4-2.

[4] N. S. Guzhva, V. E. Prun, V. V. Postnikov, M. G. Lobanov, R. N. Sadekov and D. L. Sholomov, "Using 3D Object Detection DNN in an Autonomous Tram to Predict the Behaviour of Vehicles in the Road Scene," 2022 29th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), Saint Petersburg, Russian Federation, 2022, pp. 1-6, doi: 10.23919/ICINS51784.2022.9815388.

[5] N. S. Guzhva, B. Ali, K. S. Bakulev, R. N. Sadekov and A. V. Sholokhov, "Evaluating the Accuracy of Tram Positioning System in High-Rise Building Environment Using Data from Visual Geoinformation Systems," 2023 30th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), Saint Petersburg, Russian Federation, 2023, pp. 1-5, doi: 10.23919/ICINS51816.2023.10168407.

[6] Haltakov, V., Mayr, J., Unger, C., Ilic, S. (2015). Semantic Segmentation Based Traffic Light Detection at Day and at Night. In: Gall, J., Gehler, P., Leibe, B. (eds) Pattern Recognition. DAGM 2015. Lecture Notes in Computer Science(), vol 9358. Springer, Cham. https://doi.org/10.1007/978-3-319-24947-6_37.

[7] M. Weber, P. Wolf, and J. M. Zo ̈llner, "Deeptlr: A single deep convolutional network for detection and classification of traffic lights," in 2016 IEEE Intelligent Vehicles Symposium, IV 2016, Gotenburg, Sweden, June 19-22, 2016, 2016, pp. 342–348

[8] Behrendt K., Novak L., Botros R. "A deep learning approach to traffic lights: Detection, tracking, and classification", 2017 IEEE International Conference on Robotics and Automation (ICRA), 2017, pp. 1370–1377

[9] A.W. Palmer, A. Sema, W. Martens, P. Rudolph and W. Waizenegger, "The Autonomous Siemens Tram," 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), Rhodes, Greece, 2020, pp. 1-6, doi: 10.1109/ITSC45102.2020.9294699.

[10] Possatti, Lucas C. et al. "Traffic Light Recognition Using Deep Learning and Prior Maps for Autonomous Cars", 2019 International Joint Conference on Neural Networks (IJCNN), 2019, pp. 1-8.

[11] Wang, Q.; Zhang, Q.; Liang, X.; Wang, Y.; Zhou, C.; Mikulovich, V.I. Traffic Lights Detection and Recognition Method Based on the Improved YOLOv4 Algorithm. Sensors 2022, 22, 200. https://doi.org/10.3390/s22010200

[12] Wu S. et al. aUToLights: A Robust Multi-Camera Traffic Light Detection and Tracking System //arXiv preprint arXiv:2305.08673. – 2023.

[13] Mentasti S, Simsek YC and Matteucci M (2023) Traffic lights detection and tracking for HD map creation. Front. Robot. AI 10:1065394. doi: 10.3389/frobt.2023.1065394.

[14] J. Kim, H. Cho, M. Hwangbo, J. Choi, J. Canny and Y. P. Kwon, "Deep Traffic Light Detection for Self-driving Cars from a Large-scale Dataset," 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 2018, pp. 280-285, doi: 10.1109/ITSC.2018.8569575.

[15] Li Z, Zeng Q, Liu Y, Liu J, Li L. An improved traffic lights recognition algorithm for autonomous driving in complex scenarios. International Journal of Distributed Sensor Networks. 2021;17(5). doi:10.1177/15501477211018374.

[16] M. S. Chelva, S. V. Halse and B. K. Ratha, "Object tracking in real time embedded system using image processing," 2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES), Paralakhemundi, India, 2016, pp. 1840-1844, doi: 10.1109/SCOPES.2016.7955763.

[17] Radosavovic I. et al. Designing network design spaces //Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. – 2020. – pp. 10428-10436.

[18] Lin T. Y. et al. Focal loss for dense object detection //Proceedings of the IEEE international conference on computer vision. – 2017. – С. 2980-2988.

The Research Progress of Autonomous Navigation Technology For Unmanned Transport Aircraft In Satellite-Denied Environments

Shanwu Wang
Flight-Control and Navigation Tech Research Lab
Aerospace Times Feipeng Co., Ltd.
Beijing, China
505-wangshanwu@htsdfp.com

Bingchuan Zhang
Flight-Control and Navigation Tech Research Lab
Aerospace Times Feipeng Co., Ltd.
Beijing, China
zhangbingchuan01@126.com

Qing Guo
Chief Engineer Office
Aerospace Times Feipeng Co., Ltd.
Beijing, China
gqcnsd@126.com

Shichao Li
Flight-Control and Navigation Tech Research Lab
Aerospace Times Feipeng Co., Ltd.
Beijing, China
18332300241@163.com

Chuang Xu
Flight-Control and Navigation Tech Research Lab
Aerospace Times Feipeng Co., Ltd.
Beijing, China
xuchuang@hit.edu.cn

Abstract—**With the development of the logistics industry, the operating scenarios for unmanned transport aircraft are constantly expanding. Unmanned aircraft have put forward higher demands for navigation technology in satellite-denied environments and have become hot topics and challenges in current research. Traditional navigation systems that combine satellite navigation with inertial navigation require high performance from satellite navigation. Once satellite signals are blocked or interfered with, the navigation accuracy and stability of unmanned aircraft will be severely affected. Therefore, it is of great significance to research and develop autonomous navigation technology for unmanned aircraft in satellite-denied environments. Key technologies currently being researched for autonomous navigation of unmanned aircraft include positioning-attitude measurement and route planning. This paper discusses the characteristics of navigation technology requirements for large fixed-wing unmanned transport aircraft. It analyzes the aforementioned key technologies from the perspectives of technical implementation system architecture, the challenges of technical implementation, characteristics of emerging technologies, and the practical applications of autonomous navigation systems that integrate new technologies in unmanned transport aircraft. In terms of positioning and attitude measurement, effective combinations of high-altitude image matching technology, information extraction technology based on elevation maps, and multi-source information fusion navigation technology can effectively compensate for the deficiencies of traditional navigation technologies and gradually become the mainstream of research. In terms of route planning, combining global static route planning with local dynamic route planning algorithms can assist unmanned transport aircraft in flying safely and efficiently in complex environments. In the future, with the continuous advancement of information technology and the development of intelligent algorithms, autonomous navigation technology for unmanned aircraft will evolve towards higher accuracy, stronger stability, and better real-time performance, providing strong support for the rapid development of unmanned aircraft logistics transportation industry.**

Keywords—Unmanned Transport Aircraft, Autonomous Navigation Technology, Satellite Denied Environments, Positioning-Attitude Measurement, Route Planning.

I. INTRODUCTION

In recent years, the unmanned transportation industry has been rapidly developing. According to the "Global Unmanned Aerial Vehicle Logistics and Transportation Market Report" released by the global market research firm Markets and Markets, the global logistics UAV market is projected to grow to $29.06 billion by 2027, with a compound annual growth rate of 21.01% during the forecast period [1]. Compared to traditional logistics transportation methods, unmanned transport vehicles offer advantages such as high efficiency, speed, flexibility, safety, and low cost [2]. Additionally, the operating environment of unmanned transport vehicles presents characteristics such as large spatial coverage, variable weather conditions, complex ground environments, and complicated air traffic management. To ensure the safe and efficient operation of UAVs, autonomous navigation technology for UAVs has undergone rapid development, with particular emphasis on autonomous navigation technology in GNSS-denied.

Autonomous navigation technology for UAV can estimate the position, velocity, attitude, and other state information of UAVs in real-time without relying on external support, using only the perception and computational equipment carried onboard. It also enables dynamic adjustments to flight paths. Traditional UAV autonomous navigation is achieved through the fusion of GNSS and Inertial Navigation Systems. Under GNSS completely denied conditions, the positioning accuracy of inertial navigation systems is high in the short term. Cumulative errors over time can lead to decreased system stability, making it difficult to ensure flight safety [3].

To address the aforementioned issues, this paper introduces the development process of autonomous navigation technology for UAVs, analyzes the characteristics of navigation technology requirements for large fixed-wing unmanned transport aircraft, and comprehensively reviews the key technologies of

autonomous navigation and path planning for UAVs under satellite denial conditions. It also discusses solutions for large fixed-wing unmanned transport aircraft based on application scenarios, providing references for related research and applications.

II. THE DEVELOPMENT STAGES OF UAV AUTONOMOUS NAVIGATION TECHNOLOGY

The development of UAV autonomous navigation technology has undergone several stages, gradually evolving and improving with the continuous advancement of technology.

A. Early Stage: Exploration and Initiation

During the early stage of UAV autonomous navigation technology development, researchers primarily focused on addressing how UAVs could navigate and locate themselves without human intervention. During this period , autonomous navigation technology heavily relied on Inertial Navigation Systems (INS), which estimated the speed and position of UAVs using sensors such as gyroscopes and accelerometers. However, due to the accumulation of errors over time, the positioning accuracy of INS was limited.

B. Mid-stage: Integration of Multiple Technologies

As technology progressed, researchers began to explore the integration of various navigation technologies to enhance the accuracy and reliability of UAV autonomous navigation. For example, the introduction of the Global Positioning System (GPS) provided UAVs with absolute positioning information, significantly reducing the error accumulation of INS. Additionally, visual navigation technology started to be applied in UAV autonomous navigation, enabling UAVs to perceive and locate their positions more accurately by identifying and analyzing visual features in the environment.

C. Modern Stage: Intelligence and Deep Learning

Entering the modern stage, UAV autonomous navigation technology has embraced the wave of intelligence and deep learning. By equipping advanced sensors and computational devices, UAVs can acquire and process a large amount of environmental information in real-time, thereby achieving higher-level autonomous navigation functions. At the same time, the application of deep learning enables UAVs to autonomously learn and recognize obstacles and threats in the environment, thereby better performing path planning and obstacle avoidance.

D. Future Outlook: More Autonomous and Intelligent

Looking ahead, UAV autonomous navigation technology will continue to evolve towards greater autonomy and intelligence. On the one hand, with the continuous improvement of sensor technology and computational capabilities, UAVs will be able to acquire richer environmental information and achieve higher accuracy positioning and navigation. On the other hand, through technologies such as deep learning and artificial intelligence, UAVs will better understand and adapt to

complex and dynamic environments, enabling higher-level autonomous decision-making and actions.

III. ANALYSIS OF NAVIGATION TECHNOLOGY REQUIREMENTS FOR LARGE FIXED-WING UNMANNED TRANSPORT AIRCRAFT

Large fixed-wing unmanned transport aircraft possess several advantages, including high payload capacity, low transportation costs, high transportation efficiency, and short transportation times. Additionally, their requirements for takeoff and landing conditions are much lower compared to those of civil cargo aircraft. In current unmanned transport operations, large fixed-wing unmanned transport aircraft typically carry payloads exceeding 1000 kilograms, have a range of over 1000 kilometers, cruise at speeds exceeding 150 kilometers per hour, require takeoff and landing distances of no more than 500 meters, and are capable of meeting the demands of inter-city logistics transportation, transportation of agricultural and sideline products in remote areas, and emergency delivery of supplies to islands.

a.*Early Stage* b.*Mid-stage*

c.*Modern Stage* d.*Future Outlook*

Fig. 1. The Development Stage of Auto-Navigation-Tech for UAV

Due to the operational requirements of large fixed-wing unmanned transport aircraft in various scenarios, their operating environment is more complex than that of civil cargo aircraft. This complexity includes:

A. Complex and variable weather conditions.

Large fixed-wing unmanned transport aircraft often operate over long distances, traversing diverse weather environments at takeoff and landing sites as well as throughout flight paths. Therefore, these aircraft require dynamic path planning capabilities and autonomous landing capabilities to adapt to changing weather conditions. As shown in Fig. 2, during UAV route planning and flight operations, avoiding areas with variable weather conditions needs to be considered.

Fig. 2. Setting flight routes for unmanned transport aircraft requires consideration of complex and variable weather conditions

B. Diverse geographical environments.

Flight routes for large fixed-wing unmanned transport aircraft often traverse diverse geographical environments.

a. Simulating unmanned transport aircraft flying in rainy and snowy weather

b.Simulating unmanned transport aircraft flying in valleys

Fig. 3. Simulating unmanned transport aircraft flying in valleys.

Apart from regions with relatively flat terrain such as plains and deserts, these routes typically also include areas with varying elevations such as mountains and plateaus, as shown in Fig. 3. Therefore, unmanned aircraft need to possess terrain elevation sensing capabilities and the ability to autonomously plan paths based on terrain features.

C. Unpredictable electromagnetic environment.

Various electronic information systems operating in frequency bands, such as ground-based radar systems, communication radio stations, and electro-optical equipment, as well as terrestrial and maritime clutter, cloud and rain particles, and land-sea-air platforms, all constitute the physical entities generating electromagnetic interference signals for unmanned aircraft [4]. Due to factors such as the distribution of electronic equipment, operating frequencies, radiation power, radiation modes, geographic and meteorological conditions, the electromagnetic environment during the flight of large fixed-wing unmanned aircraft exhibits complex characteristics, manifesting as diverse, dynamic, and adversarial electromagnetic interference [5]. Large fixed-wing unmanned transport aircraft typically fly at altitudes ranging from several thousand to tens of thousands of feet, making satellite navigation, radio communication, and other systems more susceptible to electromagnetic interference.

Fig. 4. During UAV flights, the satellite navigation status can be affected by electromagnetic interference.

The electromagnetic environmental factors affecting satellite navigation systems can be divided into unintentional interference, intentional interference, and disturbances caused by natural environmental factors [6]. Unintentional interference and disturbances from natural environmental factors are unavoidable and often result in signal loss and navigation state fluctuations. Intentional interference mainly includes jamming and spoofing [7,8]. For instance, as shown in Fig. 4, during the flight of unmanned aircraft, satellite navigation signals are subjected to varying degrees of electromagnetic interference, leading to a decrease in the number of satellites available for navigation to below 6, and a significant increase in PDOP values, which can affect the positioning accuracy of the unmanned aircraft. As

depicted in Fig. 5, during the flight of unmanned aircraft, the number of satellites available for navigation and PDOP values deteriorate similarly. The baseline length of the satellite navigation dual antennas experiences a sudden change. More critically, there is a sudden change in the output of the satellite navigation measurement for attitude and heading angles, reducing the accuracy of the unmanned aircraft's heading angle and posing a threat to flight safety. Therefore, when satellite navigation is affected by electromagnetic interference, it is necessary to utilize scene matching technology to detect interference in satellite navigation and to correct real-time deviations in satellite navigation positions.

Fig. 5. The satellite navigation's measurement of attitude and heading can be affected by electromagnetic interference during UAV flight.

IV. KEY TECHNOLOGIES FOR AUTONOMOUS NAVIGATION OF UNMANNED AIRCRAFT UNDER SATELLITE NAVIGATION DENIED CONDITIONS

Key Technologies for Autonomous Navigation of Unmanned Aircraft in Satellite Navigation Denied Environments mainly comprise two aspects: perception and positioning technology, and path planning technology. Among these, perception and positioning technology primarily include information extraction based on elevation maps, image matching navigation technology, and multi-mode integrated navigation technology.

A. Information-Extraction-Tech Based on Elevation-Maps

With the rapid advancement of drone technology, the demand for navigation in various complex environments is growing. Elevation maps, as important data sources reflecting ground height information, provide crucial support for drone navigation and decision-making. Information extraction technology based on elevation maps mainly consists of two aspects: the construction and updating of elevation maps, and the extraction and analysis of terrain features. By processing and analyzing elevation data, key terrain features can be extracted, providing important foundations for drone path planning, positioning, and decision-making.

1) Construction and Updating of Elevation Maps

The construction of elevation maps is a crucial aspect of UAV navigation technology. It provides detailed terrain information to UAVs, including ground elevation, slope, ridges, valleys, and other features, aiding UAVs in terrain perception. Constructing elevation maps typically relies on remote sensing technology and ground measurement data. Remote sensing utilizes methods such as satellite

imagery and radar scanning to obtain ground elevation data, while ground measurement data is obtained through precise measurements using ground equipment such as LiDAR and total stations. After processing and analysis through Geographic Information Systems (GIS) and other methods, these data can generate high-precision elevation maps. Additionally, UAVs themselves can carry elevation measurement devices such as LiDAR or ultrasonic sensors for real-time elevation data collection.

Reference [9] presents the design and implementation of a rapid construction system for locally dense elevation maps with feasible domain information, as well as an online construction system for globally consistent dense elevation maps over large areas. Reference [10], building upon RTAB-Map visual odometry, implements the construction of a 2.5D elevation grid map, extending the elevation map construction module for RTAB-Map. However, this form of environmental map is more suitable for autonomous localization and navigation of robots or other ground-based unmanned platforms. It is worth mentioning that RTAB-Map is a classic solution in vision-based SLAM, primarily comprising robust visual odometry. It also provides three forms of map construction: dense point cloud maps, 2D occupancy grid maps, and Octomap (3D occupancy grid maps). Among these, Octomap is particularly useful for reflecting obstacle information in three-dimensional space and is commonly used in UAV navigation.

Reference [11] utilizes UAV remote sensing technology for high-quality terrain image acquisition in complex mountainous environments. It employs Structure from Motion (SfM) and Multi View Stereo (MVS) algorithms to construct 3D scenes, and integrates visual interpretation, GIS terrain feature extraction, and analysis methods to achieve comprehensive analysis of terrain feature parameters such as elevation distribution, slope, and aspect. The process of constructing the 3D scene is illustrated in Fig. 6.

a. Building sparse point clouds   b. Three-dimensional grid

c. Texture mapping   d. Construction of 3D-scenes

Fig. 6.   The process of constructing 3D scenes.

Reference [12] utilized an aircraft remote sensing system, combined with real-time pose compensation technology, to acquire centimeter-level precise orthoimage data, thereby obtaining high-definition three-dimensional real scene images. Subsequently, LiDAR360 was employed to generate digital elevation models, digital surface models, and canopy height models to obtain useful terrain and forest information, and to construct a three-dimensional GIS spatial database.

Traditional aerial photography, while capable of swiftly and efficiently acquiring geographic information data, demands costly equipment and imposes high requirements on takeoff locations and flight airspace. These factors to some extent limit its application in constructing large-scale elevation topographic maps. Reference [13] employed unmanned aerial vehicles (UAVs) with oblique photogrammetry, pseudo-ground flight techniques, and 3D modeling to construct a three-dimensional model of the test area, upon which elevation topographic maps were created. This UAV pseudo-ground flight oblique photogrammetry technique enables large-scale mapping in complex areas.

With the rapid development and improvement of surveying satellite technology, data collection for large-scale topographic mapping can now be achieved through satellite remote sensing photography. However, this method is susceptible to weather conditions and terrain factors, leading to suboptimal results in elevation topographic mapping. Furthermore, regular updates of elevation maps are crucial, especially in regions where geographical conditions change rapidly, to maintain their accuracy and usability. Methods for updating elevation maps primarily include regularly collecting new remote sensing data, conducting field measurements using unmanned aerial vehicles or ground equipment, and dynamically updating maps using machine learning algorithms, which involve local or global adjustments to the maps. These methods effectively update elevation maps to adapt to rapid changes in terrain.

2) Extraction and Analysis of Terrain Features

Terrain feature extraction involves identifying key elements such as peaks, valleys, saddles, ridges, slopes, as well as the overall trends and local variations of terrain from elevation maps. This typically entails image processing techniques and computer vision algorithms such as filtering, edge detection, region segmentation, and feature point extraction.

Reference [14] investigated elevation data extraction methods for contour maps, primarily utilizing the secondary development platform ArcGIS Engine of the geographic information system (GIS) software ArcGIS. Firstly, contour maps are rasterized, and then converted into point feature data based on grid units. Finally, each point is assigned a reasonable elevation value through interpolation methods. It's worth noting that not every point on the contour map has an elevation value. Point distribution is spatially correlated, with nearby points having similar elevation values. To avoid local influences, the inverse distance weighted interpolation (IDW) method is chosen for elevation data extraction. Reference [15] focusing on underwater terrain-aided navigation (UTAN), studied real-time underwater terrain models for extracting local underwater terrain features. They developed a UTAN digital terrain map (DTM) and interpolation reconstruction method based on multi-beam echo sounders (MBES). Terrain Matching Diagram of UAV is illustrated in Fig. 7.

Reference [16] proposes a method for extracting and matching terrain features from sparse point cloud data,

addressing the robust matching of three-dimensional features under flat terrain conditions. This method is not only applicable to terrain-related navigation for planetary rovers and landers but also can be utilized in object recognition and SLAM. Reference [17] introduces a method for quickly extracting Digital Elevation Model (DEM) data from scanned contour maps using contour rendering.

Fig. 7. Terrain Matching Diagram of UAV [15]

Further analysis after terrain feature extraction may include statistical characteristics of the terrain, evaluation of terrain complexity, and other aspects to provide more comprehensive terrain information. Analyzing terrain features helps UAVs understand the flight environment and predict potential risks such as terrain obstacles and airflow variations. Additionally, terrain features can provide important basis for path planning and navigation decision-making, enabling the selection of safer and more efficient flight trajectories. Reference [18] has designed a complete automated solution based on digital map data reading, data processing, key data extraction, geometric parameter calculation and analysis, and model display.

3) The Application Trends of Elevation Maps in UAV Navigation

The application of elevation maps in UAV navigation is extensive and profound, particularly in complex terrain environments. Taking large fixed-wing unmanned transport aircraft as an example, during flight missions in mountainous and plateau regions, elevation maps can provide precise terrain information, supporting terrain perception and autonomous path planning. This enables avoidance of potential danger areas such as mountains and valleys, selecting safe flight corridors. Throughout the flight, UAVs can adjust flight altitude and heading in real-time based on elevation maps to adapt to terrain changes.

To meet the flight demands of UAVs in mountainous and undulating terrain, especially to avoid the impact of obstacles such as mountains on UAVs, elevation maps can also be used for terrain-following flights. This allows UAVs to fly close to the ground or maintain a certain safe altitude, thus achieving the technology of terrain-following flight in complex terrain.

Many researchers have conducted studies on low-altitude terrain-following flight technology and its applications based on elevation map information[19-26] to enhance the safety of UAV flights. It is worth noting that Digital Elevation Model (DEM) data lacks elevation information

for vegetation, buildings, etc., compared to Digital Surface Model (DSM) data. Therefore, before conducting UAV terrain-following flight operations based on elevation map extraction technology, it is advisable to establish DSM image data containing elevation information of tall buildings, power lines, vegetation, communication towers, etc., within the terrain-following flight area through pre-surveying operations to prevent collision accidents caused by altitude issues during UAV terrain-following flight operations. Reference [27] uses existing three-dimensional surface data DSM to keep the UAV at a constant height relative to ground targets, overcoming problems such as large terrain height differences and occlusions by prominent surfaces, thus enabling safer identification of steep slopes and hazardous rock masses. Reference [28-31] utilize existing DSM elevation map data to implement terrain-following oblique photogrammetry technology, applied in the production of large-scale topographic maps in mountainous areas.

In addition, elevation maps can be combined with other navigation technologies to form a multi-source information fusion navigation solution. For example, combining visual navigation technology with elevation maps can assist in the extraction and matching of visual features using terrain information provided by elevation maps. Similarly, integrating inertial navigation technology with elevation information can correct and compensate for inertial measurement data, thereby improving navigation accuracy and stability. Reference [32] addresses the autonomous search for emergency landing sites for UAVs in outdoor environments. Based on DEM digital elevation maps, it proposes a rapid search algorithm for emergency landing sites based on neural networks. By leveraging deep learning visual navigation technology and constructing an image segmentation deep network, it achieves fast and stable search results.

B. Image Matching Navigation Technology

Image matching navigation is divided into two types: terrain contour matching navigation and scenery matching navigation. The key data for image matching navigation is referred to as a digital map. Since surface features generally do not change much, surface images are pre-taken and stored in the UAV. When the UAV flies over, navigation is carried out by identifying the original image and current surface features to determine the flight position.

1) Terrain Matching Technology

Terrain matching navigation technology relies on terrain data along the flight route as a basis. During flight, real-time altitude data is collected using onboard radar, radio altimeter, barometric altimeter, etc. The elevation profile of the flight path projected onto the ground is calculated and compared with the gridded digital map elevation data stored on board using relevant algorithms to determine the position of the aircraft and make navigation adjustments accordingly. This technology is suitable for navigation in hilly and mountainous terrain with significant undulations. Therefore, terrain matching navigation is also an application of the aforementioned elevation map information extraction technology in the

field of UAV autonomous navigation and terrain-following flight, providing strong support for UAV applications in complex terrain navigation, aerial photography, geological exploration, etc.

Reference [33] designed a UAV terrain matching terrain-following flight method based on DEM data, which enables the UAV to make corresponding adjustments according to the terrain undulations, ensuring a fixed flight altitude difference and achieving terrain-following flight, thereby improving the quality of aerial photographs and the safety of UAVs. Reference [34] studied DEM acquisition methods and proposed a method for scene-assisted UAV sequence image positioning based on elevation information constraints, demonstrating higher accuracy compared to positioning methods based on plane scene reconstruction. Reference [35-36] investigated underwater terrain matching navigation technology, similar to aircraft terrain matching. When an AUV crosses a terrain-adapted area, the terrain elevation values of measurement points are obtained through depth sensors, and the optimal pose of the AUV is determined by comparing them with the reference terrain elevation map stored in advance using terrain matching algorithms.

However, terrain matching technology also has some limitations. For example, in mountainous areas with drastic terrain changes or in plains with poor terrain features, the positioning accuracy may be affected. Therefore, terrain matching technology needs to be integrated with other navigation methods to improve overall navigation performance.

2) Scenery Matching Technology

Scenery matching navigation is a visual navigation technology that utilizes image sensors to capture images of the surrounding area during flight or near the target area. These images are then matched with stored reference images to obtain aircraft position data. This method is suitable for navigation in flat areas with varying features and minimal elevation changes. Scenery matching navigation, with its high resolution, autonomous navigation capabilities, and compact size, has gained increasing attention in the field of precise navigation. It has become an important research direction in the field of autonomous navigation and positioning of unmanned aerial vehicles (UAVs) in anti-GPS-denied environments [37-38].

The guidance accuracy of scenery matching depends on the real-time image quality, reference image accuracy, and performance of the matching method. This method is essentially a target positioning method based on template matching. Unlike terrain matching, the guidance accuracy of scenery matching is relatively higher by an order of magnitude. Therefore, while terrain matching guidance is generally used for mid-course guidance, scenery matching guidance is typically used for terminal guidance [39]. Although scenery matching navigation demonstrates high flexibility and outstanding autonomy, in practical operations, UAVs usually operate in open environments where real-time images captured may vary significantly. Especially in complex scenes, real-time images and reference images are often captured under different conditions such as different times, sensors, scenes, angles, and weather conditions, leading to differences in grayscale, clarity, geometric distortion, and occlusions between the two images. These significant differences pose significant challenges to the positioning algorithms of scenery matching navigation.

Currently, the mainstream approach is to research methods for registering heterogeneous images [40-44]. Reference [45], a new method proposed. enhances the internal features of target images using a dual-stream deep network, enabling precise matching of aerial images obtained in different environments. The general process of registration for heterogeneous image matching is illustrated in Fig. 8. Initially, corresponding reference images are selected from the reference image database based on the positioning results of the inertial navigation system. Simultaneously, real-time images are captured by the UAV through onboard sensors. Due to the drift error of the inertial navigation system, there exist geometric differences between the reference and real-time images, necessitating image registration methods to align them. After alignment and obtaining affine transformation parameters, the UAV's position can be calculated. The center of the real-time image represents the position of the UAV. Using the affine transformation parameters, the position of the real-time image center on the reference image can be calculated. Since the reference image contains geographic information, the position of the UAV can be determined accordingly.

Fig. 8. Positioning system based on heterogeneous image matching

The traditional research on image-matching navigation mainly focuses on area-adaptive navigation, with few methods capable of continuous image matching navigation across adaptive and non-adaptive areas, resulting in poor continuity and to some extent limiting its application scope. To address this issue, it is generally necessary to study various forms of sensors and try multiple research methods. Image-matching visual devices can include sensors in forms such as monocular, binocular, multiocular, fisheye, etc., with visible light, infrared, laser, etc. Research methods include feature extraction, feature matching, optical flow analysis, bionics, and various combined navigation methods. Reference [46-47] studied image matching algorithms for UAV target positioning in night vision environments, solving the problems of low feature detection repeatability and low feature matching accuracy in infrared and visible light image matching.

In addition, the navigation systems of large unmanned transport aircraft have high engineering requirements for

real-time image matching due to their fast flight speeds. Poor real-time performance cannot be compensated for even with high precision and reliability of matching results. During the matching process, it is usually necessary to search for matches over a large range in the reference image, resulting in large computational complexity. Therefore, improving the real-time performance of image matching is crucial. Some researchers have conducted related studies on how to optimize algorithms to improve matching speed and accuracy and reduce computational complexity to meet real-time [48-49]. Reference [50] proposed a new fast and robust template matching framework for multi-mode remote sensing image registration. It uses a template matching strategy to detect correspondences between images and defines a fast similarity measure based on feature representation in the frequency domain using Fast Fourier Transform (FFT), improving computational efficiency. Reference [51] proposed a new hybrid keypoint selection method and verified that applying appropriate keypoint selection techniques can improve the accuracy and efficiency of UAV image matching and positioning results. Reference [52] proposed a matching algorithm that divides the feature neighborhood into concentric circles based on feature points and improves the generation of feature descriptors in the Scale Invariant Feature Transform (SIFT) algorithm. It also filters matching point pairs to improve the algorithm's real-time performance and accuracy.

3) The Application Trend of Image Matching Technology in UAV Positioning

The application of image matching technology in UAV positioning has become increasingly widespread [53-55]. By integrating ground scene recognition, terrain matching, and image matching technologies, UAVs can achieve precise positioning and navigation in complex environments. This navigation method not only improves the accuracy and stability of UAV positioning but also provides strong support for UAV applications in agriculture, environmental protection, rescue operations, logistics, and other fields.

With the development of artificial intelligence technologies such as deep learning, the performance of image matching algorithms has been significantly improved. In recent years, the development of deep learning technology has provided strong technical support for ground scene recognition [56-59]. Training deep neural networks can efficiently and accurately identify ground images, greatly improving the accuracy and stability of UAV positioning. Reference [60] proposes a UAV autonomous navigation and obstacle avoidance method based on deep reinforcement learning algorithm—FRDDM-DQN. Based on the DQN, this algorithm introduces an optimized Faster R-CNN model to extract obstacle information from images and achieves better and more efficient training results through the proposed experience pool experience admission mechanism. This research achievement realizes image matching without prior maps, expands the application scenarios of UAVs, and further enhances the ability of UAVs to autonomously execute tasks.

High-precision digital scene maps obtained through machine learning algorithms such as deep learning and reinforcement learning serve as the basis for visual scene matching navigation. Constructing structured digital scene map data with multiple scales, resolutions, and temporalities, even SLAM three-dimensional mapping, and dynamically updating elements such as digital scenes, airport runways, ground signs, buildings, towers, and signal lights in the map can meet the accuracy and integrity requirements of scene matching navigation in unfamiliar environments.

C. Multimodal fusion navigation technology

Unmanned aerial vehicles (UAVs) may experience positioning failure in heavily interfered environments during flight missions. In such cases, reliance on other sensors is necessary to infer the UAV's position, ensuring safe landing or return. Inertial sensor navigation systems, as autonomous navigation devices, operate without relying on external radio signals, providing inherent resistance to interference and spoofing. However, under inertial navigation mode, positioning errors can be significant, necessitating combination with other navigation methods such as radio navigation or terrain recognition navigation. Because inertial navigation can provide a wealth of navigation parameters and full attitude information, it is typically the primary component in multi-modal fusion navigation systems.

To address situations where satellite navigation is interfered with or fails, the system is designed with solutions such as image matching, radio navigation, and laser-assisted inertial navigation. Multimodal navigation devices are added to adaptively improve navigation in unfamiliar environments, enabling navigation in straits, islands, coastlines, and hilly terrains. This multimodal navigation technology can adapt to satellite navigation failures, utilizing "pure inertial" navigation, "image + inertial" combination navigation, and "radio + inertial" combination navigation techniques. As shown in Fig. 8, for redundant satellite navigation equipment and inertial navigation equipment, design redundancy switching priority to achieve automatic switching in case of navigation equipment failure.

D. Autonomous Trajectory Planning for Unmanned Transport Aircraft (UTA)

Autonomous trajectory planning is crucial for enabling unmanned transport aircraft to navigate efficiently and safely in diverse environments. It involves generating an optimal path that adheres to performance constraints and ensures flight safety, essentially solving an optimization problem under multiple constraints.

1) Constraints

Unmanned transport aircraft face various influencing factors during flight and planning, necessitating compliance with multiple constraints:

- **Self-performance constraints**: These include minimum trajectory segment length, minimum turn radius, maximum pitch angle, minimum and maximum flight altitudes, and maximum trajectory length.

- **Mission constraints**: These involve completion time, starting point, target point, fixed heading angle, and other mission-specific requirements.

- **Environmental constraints**: These encompass various threat field constraints, no-fly zone restrictions, and the UAV's compliance with penetration requirements. In battlefield environments, attention must also be paid to the UAV's stealth capabilities.

Fig. 9. Multimodal Navigation Technology Roadmap Diagram

2) Trajectory Planning Algorithm

Trajectory planning algorithms can be broadly categorized into traditional and modern intelligent algorithms.

a) Traditional Trajectory Planning Algorithm.

Traditional algorithms include graph search algorithms, spatial sampling algorithms, and potential field methods. While they offer good search performance and accuracy, their limitations become apparent in complex environments:

- **Graph search algorithms**: These algorithms, like the A* algorithm [61], excel in simple environments but struggle with increased search nodes and reduced real-time performance.

- **Spatial sampling algorithms**: These algorithms, like the Rapidly Exploring Random Tree (RRT), efficiently explore the search space but may overlook optimal paths due to their randomness.

- **Potential field methods**: These algorithms, like the Artificial Potential Field (APF) method [62], provide fast planning and safe results but can get stuck in local optima in complex environments.

b) Modern Intelligent Trajectory Planning Algorithm

Modern algorithms address the shortcomings of traditional methods, enabling efficient planning in complex environments:

- **Genetic Algorithm (GA)** [63]: Inspired by evolution, GA uses genetic operators to iteratively evolve an optimal trajectory based on a fitness function. While robust and suitable for complex scenarios, GA can struggle with local search and convergence speed in later stages.

- **Ant Colony Optimization (ACO)** [64]: Taking inspiration from ant foraging behavior, ACO uses pheromone updates to guide the "ant colony" towards the optimal path. Its positive feedback mechanism and ease of implementation make it popular. However, ACO can be susceptible to local optima and slow convergence.

- **Particle Swarm Optimization (PSO)**: Inspired by the social behavior of bird flocks and fish schools, PSO utilizes information sharing among "particles" to iteratively improve their positions and converge towards the optimal trajectory. PSO offers good balance between exploration and exploitation but may suffer from parameter sensitivity.

TABLE I.   ALGORITHM COMPARISON

Title	Advantages	Disadvantages
A-Star algorithm	Good search performance and high accuracy	efficiency decreases significantly as nodes increasing
Artificial Potential Field algorithm	software runs at high speed and good real-time performance	Easy to fall into local optima
Ant Colony Optimization	Strong global search capability	Slow convergence speed and easy to fall into local optimal solutions
Genetic Algorithm	Strong scalability and easy to combine with other algorithms	The implementation is relatively complex;The algorithm parameters are set based on experience

V. CONCLUSIONS AND PROSPECTS

The future research directions for autonomous navigation technology in large fixed-wing unmanned transport aircraft can be summarized as follows:

1) Multi-UAV Cooperative Navigation: The technology of cooperative navigation for unmanned aerial vehicles (UAVs) integrates sensor information and navigation data

carried on individual UAVs. Through cluster information sharing and optimization of inter-vehicle observations, more accurate navigation positioning information is provided for individual UAVs within the UAV cluster, enabling autonomous navigation of the cluster. The cluster navigation technology for large fixed-wing unmanned transport aircraft can significantly enhance the robustness and accuracy of navigation for UAV clusters, thereby improving the payload capacity of UAVs. Research focus is on leveraging communication among UAV clusters to enhance overall performance. Key areas of future research include filter-based cooperative localization optimization, graph-based cooperative localization optimization, and optimization algorithms based on machine learning. Filter-based cooperative localization optimization often utilizes distributed Kalman filtering and particle filtering, known for their low algorithmic complexity and wide applicability. Research in this area focuses on addressing issues such as system time delay, methods for compensating for state estimation errors [65], and convergence speed of initial states [66]. Graph-based cooperative localization optimization involves constructing factor graphs from historical data of cluster systems, establishing optimized cost functions, and iteratively minimizing linear errors within the system to achieve the optimal solution. By leveraging graph optimization methods combined with distance information between UAVs, this approach enables dense distribution of UAVs in a cluster for cooperative navigation [67]. Network factor graph algorithms are employed for distributed cooperative navigation to handle complex navigation scenarios while reducing network communication data [68]. Optimization algorithms based on machine learning leverage historical and current data on individual UAV parameters, sensor data, environmental data, etc., providing advantages in real-time performance and system energy consumption [69]. This area has become a research hotspot due to its potential for modeling and analyzing large volumes of data.

2) Integration of Multiple Technologies to Enhance the Intelligence of Navigation Systems: Research focuses on the application of machine learning, artificial intelligence, and other technologies in autonomous navigation systems for fixed-wing unmanned transport aircraft, aiming to improve the intelligence and adaptability of navigation systems. Historical weather data, terrain characteristics, and UAV flight data from the flight area are utilized to train models for predicting the optimal flight paths of UAVs. These models enable real-time online adjustments to the flight paths, allowing UAVs to better adapt to complex flight environments and mission requirements.

3) UAV Fault Prediction and Diagnosis: Due to the multitude of factors affecting UAV system safety and the nonlinearity between influencing factors and evaluation criteria, fault detection in large fixed-wing unmanned transport aircraft systems is highly complex. Utilizing methods such as Markov models [70], comprehensive fuzzy assessment [71], and deep learning [72], the state parameters of UAV systems are analyzed to predict the probability of UAV system failure. This provides a theoretical basis for preventive maintenance and fault localization.

The continuously increasing demand for navigation accuracy in large fixed-wing unmanned transport aircraft, coupled with increasingly complex navigation environments, poses numerous challenges for autonomous navigation systems. On one hand, multisensor fusion technology requires stronger sensor performance and platform computing power. Visual technology demands high image quality and imposes certain restrictions due to environmental and camera requirements. Radar equipment offers relatively high precision but is bulky and has limited perception range. On the other hand, the level of intelligence in autonomous navigation systems needs further enhancement. In addition to positioning the UAV itself during operation, there is a need to perceive surrounding obstacles and other UAVs, whether static or dynamic, requiring advanced algorithmic capabilities to support these tasks.

REFERENCES

[1] Wang Xiaomeng. Cargo drones: a new option for battlefield delivery [N]. PLA Daily, 2023-11-24(09-11).

[2] Ren Xuan, Huang Hui, Yu Shaowei, et al. Review on vehicle-UAV combined delivery problem[J]. Control and Decision, 2021, 36(10):2313-2327.

[3] Ni Anqing, Li Jun. Review on the algorithm of combined navigation of global navigation satellite system and inertial navigation system [J]. Equipment Machinery, 2023(01),4-8.

[4] Wang Liandong, Shen Xujian, Zhou Bo. Review on cognitive complexity of electromagnetic environment [J]. Aerospace Electronic Warfare, 2020, 36(2):1-5.

[5] Chen Hang-yong, Zhang Dian-zong, et al. Review on Battlefield Electromagnetic Environment Quantitative Analysis [J]. Electronic Information Warfare Technology, 2010,25(4):44-51.

[6] Zhang Tong. Research on Jamming Detection and Recognition technology of satellite navigation system [D], Thesis for Master Dgress, ShanDong University, 2023.

[7] F. Wang, H. Li and M. Lu, GNSS Spoofing Detection Based on Unsynchronized Double-Antenna Measurements[J], in IEEE Access, 2018 vol. 6, pp. 31203-31212.

[8] P. Bethi, S. Pathipati and A. P, Stealthy GPS spoofing: spoofer systems, spoofing techniques and strategies [C], 2020 IEEE 17th India Council International Conference (INDICON), New Delhi, India, 2020, pp. 1-7.

[9] Pan Yiyuan. Dense elevation mapping and place recognition [D], Master's thesis, Zhejiang University, 2021.

[10] Yue Yunxiao, Ren Jiahao. Research on topographic environment elevation mapping based on RTAB-Map visual odometry [J]. Industrial Control Computer, 2023, 36(3):42-44.

[11] Bi Rui, Gan Shu, Yuan Xiping et al. 3D scene construction of complex landform based on UAV remote sensing [J]. Mountain Research, 2022, 40(1):151-164.

[12] Peng Cheng, China Patent CN113806897A, 202110996284 .4, 2021.

[13] He Yarui. A three-dimensional modeling study of oblique photogrammetry of drones in the western alpine region [D], Master's thesis, East China University of technology, 2022.

[14] Zhang Juntao, Guan Qun. Research on AE based method for extracting elevation data from contour maps [J]. Information and Communications, 2013(10):14-15.

[15] Chen Pengyun, Zhang Pengfei, Chang Jianlong, et al. Fast Extraction of Local Underwater Terrain Features for Underwater Terrain-Aided Navigation [J]. Journal of Marine Science and Application, 2019, 18(3):9.

[16] Tian Yang, Li Guoqing, Song Xin. A Novel 3D Terrain Feature Detecting and Matching Method [J]. Journal of Astronautics, 2018, 39(6):7.

[17] Qiao Lijun. A terrain map elevation data extraction technique based on contour rendering [J]. Small And Medium-Sized Enterprise Management And Technology, 2011.

[18] Ye Songtao. Research on geometric information extraction technology in digital maps [D], Master's thesis, Xidian University, 2017.

[19] Wang Tiehao, China Patent CN115390580A, 202210744671.3, 2022.

[20] Tan Jiaying, Zhou Liwei, Cai Huanqing, et al. Ground simulation flight technology of composite wing unmanned aerial vehicle based on improved RRT algorithm [J]. High Voltage Engineering, 2023, 49(S01):210-214.

[21] Zhang Yilin. Application of UAV ground imitation flight in large scale mapping in complex mountainous areas [J]. Heilongjiang Science, 2023, 14(10):114-116.

[22] Song Jingyuan. Application of drone ground simulation flight technology in emergency investigation of individual landslides in reservoir areas [J]. Gansu Water Resources And Hydropower Technology, 2023, 59(10):58-60.

[23] Sun Xuechao. Research on terrain following technology of plant protection UAV based on millimeter wave radar [D], Master's thesis, Hangzhou Dianzi University, 2018.

[24] Chen Chao, China Patent CN112000129A, 202010903956 .8, 2020.

[25] Liu Bo. Design and implementation of a ground navigation station system for a certain type of unmanned aerial vehicle [D], Master's thesis, Northwestern Polytechnical University, 2007.

[26] He Taotao, Wang Lei. Construction method and application of uav dem based on InSAR [J]. Electric Power Survey & Design, 2023(S01):76-81.

[27] Pang Xin, Yuan Ming, Lu Yuan, et al. Rapid identification method for the dangerous rock mass of a high-steep slope based on UAV LiDAR and ground imitation flight [J]. Bulletin of Geological Scienceand Technology, 2023, 42(6): 21-30.

[28] Zheng Chaofeng, Wang Dongye. Research on the application of geomorphic flight technology in large scale topographic mapping in mountainous areas [J]. Sci-Tech Innovation & Productivity, 2023, 44(9): 142-144.

[29] Ye Xueqi, Ku Ruihong. Research on the application of oblique photography based ground simulation flying technology in large-scale topographic mapping in mountainous areas [J]. Technology Innovation and Application, 2023, 13(25):189-192.

[30] Dai Ming, Liu Xiaowen, Wu Yingmei. Research on the application of oblique photography and ground simulation flight technology in mountain terrain mapping [J]. Modern Information Technology, 2023, 7(16):121-124.

[31] Huang Lizhang. The application of unmanned aerial vehicles simulating ground flight in oblique photogrammetry of complex terrain reservoirs [J]. Survey World, 2022(004):000.

[32] Liu Bing, Fang Yuan, Yang Baihan, Gao Shiqi. Potential landing sites research of UAV through digital elevation model [J]. Avionics Technology, 2023, 54 (4): 29-35.

[33] Yu Bin, China Patent CN113625731A, 202110836272 .5, 2021.

[34] Li Xiaoxue. Study on scene aided location method for ground target based on UAV image sequences [D], Master's thesis, Graduate School of National University of Defense Technology, 2015.

[35] Gao Jingxuan, Zhang Ya, Sun Fengsheng, Ku Ruihong. Review for AUV seabed terrain aided navigation methods [J]. Ship Engineering, 2023, 45(2):167-176.

[36] Wu Yuanfeng. Auxiliary navigaion method based on underwater terrain image registration[D], Master's thesis, Harbin Engineering University, 2016.

[37] Ji Xiang, Huan Junwei, Liang Nan, et al. UAV positioning simulation method based on scene matching [J]. Journal of System Simulation, 2023, 26(6):1291-1296.

[38] Li jiasong, Li Minglei, Wei Dazhou, et al. Research on scene matching navigation method under GPS-denied environments [J]. Application of Electronic Technique, 2022, 48(3) : 88-93.

[39] Hu Shuyan, Wang Shoudao, Zhu Long, Issues and solutions for UAV visual-based navigation in GNSS-denied environments, 10th China Aviation Society Youth Science and Technology Forum, Nan Chang, Chinese Society Of Aeronautics and Astronautics, 2022, pp. 1089–1096.

[40] Yi Kuiying, Huang Guan, Huang Zhaoyou. SAR platform's scene matching positioning method based on generalized pseud-range [J]. Radar Science and Technology, 2021, 19(5) : 589-597.

[41] Huang Zhaoyou. Research on scene matching navigation technology based on multi-modality Image Maching and generalized pseud-range positioning[D], Master's thesis, Nanjing Research Institute of Electronics Technology, 2021.

[42] Wang Li. Research on multi-sensor matching based on line features [D], Master's thesis, University of Chinese Academy of Sciences, 2015.

[43] Luo Shibin, Liu Haijiao, Hu Maoqing, et al. Review of multi-modal image matching assisted inertial navigation positioning technology for unmanned aerial vehicle [J]. Journal Of National University Of Defense Technology, 2020, 42(6) : 1-10.

[44] Xu Xuyang. Remote sensing image matching of different sources [D], Master's thesis, Jilin University, 2021.

[45] Jae-Hyun Park, Woo-Jeoung Nam, Seong-Whan Lee. A two-stream symmetric network with bidirectional ensemble for aerial image matching [J]. Remote Sensing, 2020, xx(5) : 1-21.

[46] Wang Qi. Research of macthing methods of infrared and visible images based on the similarity of global structures [D], Doctoral dissertation, Dalian University of Technology, 2020.

[47] Li Chenyang. Research on image matching algorithm of UAV target positioning in night vision environment [D], Master's thesis, Harbin University of Science and Technology, 2020.

[48] Chen Fang. Research on the real-time and accurate scene matching algorithm based on local invariant features [D], Master's thesis, Nanjing University of Aeronautics and Astronautics, 2010.

[49] Li Zhulin, Zhang Genyao. Research of scene matching algorithm based on regions [J]. Application Research of Computers, 2012, 29(9) : 3578-3580.

[50] Ye Yuanxin, Lorenzo Bruzzone, Shan Jie, et al. Fast and robust matching for multimodal remote sensing image registration [J]. IEEE Transactions On Geoscience And Remote Sensing, 2019, 57(11): 9059-9070.

[51] Vahid Mousavi, Masood Varshosaz, Fabio Remondino. Using information content to select keypoints for UAV image matching [J]. Remote Sensing, 2021, 13(1302) : 1-27.

[52] Zhao Qibing, Wang Yangzhu, Hu Yonghao. Remote sensing image matching for UAVs based on improved sift algorithm [J]. Electronics Optics & Control, 2012, 19(3) : 36-39, 55.

[53] Yang Wei. Research on image matching method for scene perception [D], Master's thesis, University of electronic science and technology of China, 2017.

[54] Song lin. Research on the key technologies of vision navigation for UAV in flight [D], Doctoral dissertation, Northwestern Polytechnical University, 2015.

[55] Wan Chenghao. Research on scene matching algorithms for UAV visual positioning [D], Master's thesis, PLA Strategic Support Force Information Engineering University, 2018.

[56] [56] WEI Hongan, SHAN Xiaojun, ZHENG Ke, et al. Review of SAR and Optical Image Registration Based on Deep Learning [J]. Radio Engineering, 2021, 51(12) : 1363-1372.

[57] Shen Xuelun. Cross-domain image matching based on deep learning [D], Master's thesis, Xiamen University, 2019.

[58] Zheng Bowen. Fast aerial image matching based on deep learning [D], Master's thesis, Dalian University of Technology, 2020.

[59] Wei Shaowei. Registration of cross-spectral images based on deep learning [D], Master's thesis, Xidian University, 2020.

[60] Wang Fei, Zhu Xiaoping, Zhou Zhou, et al. Deep-reinforcement-learning-based UAV autonomous navigation and collision avoidance in unknown environments [J]. Chinese Journal of Aeronautics, 2024, 37(3): 237–257.

[61] Zhang Yidou, Zhao Yan, Wei Tong. Improved A ∗ algorithm for obstacle avoidance path planning strategy of the blind [J]. Aero Weaponry, 2017(3): 86 -92.

[62] Han Yao, Li Shaohua. UAV path planning based on improved artificial potential field [J]. Systems Engineering and Electro nics, 2021(11): 3305 -3311.

[63] Duan Yuqian, He Jiali. Genetic algorithm and its modification [J]. Proceedings of the CSU EPSA, 1998(1): 39 -52. (in Chinese).

[64] Guo Qimin, Zhang Peng, Wang Yingyang, et al. Review of intelligent algorithms for autonomous path planning of UAVs [J]. Aero Weaponry, 2023, 30(3): 29 -40.

[65] Qu Y, Zhang Y. Fault-tolerant localization for multi-uav cooperative flight [C], IEEE/ASME International Conference on Mechatronics & Embedded Systems & Applications. IEEE , 2010: 131-136.

[66] Li S , Coppola M , De Wagter C ,et al. An autonomous swarm of micro flying robots with range-based relative localization [J], 2020.

[67] Chen Mingxing, Xiong Zhi, Liu Jianye, et al. Distributed collaborative navigation method for UAV swarm based on factor graph [J]. Journal of Chinese Inertial Technology, 2020, 28(4): 456-461.

[68] Li Q, Gama F, Ribeiro A, et al. Graph neural networks for decentralized multi-robot path planning [C]. In 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 11785-11792).

[69] Cheng N, Wenchao X, Weisen S, et al. Air-ground integrated mobile edge networks: architecture, challenges and opportunities [J]. Communications Magazine, IEEE, 2018, 56(8): 26-32.

[70] Mao Lei, Wang Xiaofang, et al. Fault prediction of a certain type of UAV system based on Markov chain [J], Inner Mongolia Science Technology & Economy, 2022(15):125-126.

[71] Sun Longlong; Jin Xiaobo; Feng Jinqi; Jia Jinyuan. UAV safety assessment technology based on functional failure impact analysis and multi model fusion [A], Proceedings of the 19th China Aviation Measurement and Control Technology Annual Conference [C]; 2022.

[72] QIU Ling, ZHANG Ansi, ZHANG Yu, et al. Application method of knowledge graph construction for uav fault diagnosis [J], Computer Engineering and Applications, 2023,59(9):280:288.

Information-Measuring System of a Rotation-Stabilized Spacecraft

V.V. Matveev

Laboratory of Inertial Sensors of Primary Information,
Orientation and Navigation Systems
Tula State University
Tula, Russia
matweew.valery@yandex.ru

M.D. Kirsanov

Laboratory of Inertial Sensors of Primary Information,
Orientation and Navigation Systems
Tula State University
Tula, Russia
scientific-lab-tulsu@yandex.ru

M.G. Pogorelov

Laboratory of Inertial Sensors of Primary Information,
Orientation and Navigation Systems
Tula State University
Tula, Russia
scientific-lab-tulsu@yandex.ru

I.A. Bechler

Laboratory of Inertial Sensors of Primary Information,
Orientation and Navigation Systems
Tula State University
Tula, Russia
scientific-lab-tulsu@yandex.ru

A.N. Khomyachkova

Laboratory of Inertial Sensors of Primary Information,
Orientation and Navigation Systems
Tula State University
scientific-lab-tulsu@yandex.ru

Abstract— **The paper examines the information-measuring system of a rotation-stabilized spacecraft consisting of a module of gyroscopes and magnetometers. A method is proposed for estimating the displacement of the gyroscope zero signal based on data from magnetometric sensors.**

Keywords— rotating spacecraft, gyroscope, magnetometer

I. INTRODUCTION

Stabilization of spacecraft (SC) by rotation has certain advantages compared to other types of stabilization [1, 2]: reduced energy consumption for control, inertia-free precessional motion of the spacecraft in the first approximation, long-term orientation of solar panels to maximum illumination, etc. However, over time, the angular velocity rotation of the spacecraft will gradually decrease due to energy dissipation in the joints of the structure, the influence of the Earth's magnetic field, changes in the moment of inertia and other disturbances. In this regard, the information-measuring system of the spacecraft must generate data on the angular velocity of its own rotation in order to subsequently use this information to create control signals.

The information-measuring system of a spacecraft usually includes a gyroscopic angular velocity sensor, which can be implemented using micro-electromechanical systems (MEMS) technology. It is known that MEMS gyroscopes, as low-accuracy sensors, have significant instrumental errors. The purpose of this work is to describe a method for compensating for the offset of the zero signal of a gyroscopic angular velocity sensor based on magnetometer readings.

II. INFORMATION AND MEASUREMENT SYSTEM OF SPACECRAFT

Figure 1 shows the information-measuring system of the spacecraft, containing a block of primary information sensors, including a MEMS gyroscope with a measuring axis Z, as well as magnetometric sensors with measuring axes X and Y, respectively. The angular velocity of the spacecraft is measured by a MEMS gyroscope, the signal of which enters the amplification-converting path designed to control the flywheel engine. The flywheel is the executive element of a closed system for stabilizing the angular velocity of the spacecraft. If the angular speed of the spacecraft decreases, it is necessary to increase the angular speed of the flywheel, and the direction of its rotation should be opposite to the direction of rotation of the spacecraft body. As the angular velocity of the spacecraft increases, the flywheel, on the contrary, must slow down. However, this mode is unlikely, since the disturbing moments acting on the spacecraft are braking [3]. A decrease in the angular velocity of the spacecraft's own rotation will have a negative effect on its dynamics and, ultimately, can lead to loss of stability. In addition, the accuracy of spacecraft stabilization is determined by the characteristics of the MEMS gyroscope, in particular the magnitude of the shift of its zero signal.

It is important to take into account that in order to compensate for the decrease in the speed of the spacecraft's own rotation, it is necessary to constantly increase the speed of rotation of the flywheel. However, this increase is limited by the maximum capabilities of the electric motor and at some point, the flywheel will reach a state of saturation and the system will become inoperable. To return the system to working condition, an additional unloading system with actuators of a different type is required, for example, through jet nozzles.

Magnetometers, in addition to the main function of measuring the direction and magnitude of the Earth's magnetic field, are also used to correct the gyroscope signal.

Fig. 1. Functional diagram of the information-measuring system

III. METHOD FOR ESTIMATION THE ZERO SIGNAL OF A GYROSCOPE

The essence of the method for determining the offset of the gyroscope zero signal is as follows. During one-way rotation of the spacecraft, the measuring axes of the magnetometers move relative to the geomagnetic induction vector. As a result, magnetometers produce amplitude-modulated signals at the spacecraft rotation frequency. Analysis of the spectral characteristics of magnetometer signals using the fast Fourier transform allows one to estimate the rotation frequency of the spacecraft. It is important to note that the intrinsic displacements of magnetometer zero signals do not have a significant effect on the accuracy of determining the rotation frequency. This is due to the fact that quasi-constant shifts of magnetometer zero signals will be located in the low-frequency region of the spectrum. Flight calibration of magnetometers is considered in [4]. The spacecraft rotation frequency obtained in this way is used to estimate the displacement of the gyroscope zero signal.

IV. RESULTS OF EXPERIMENTAL STUDIES

A MEMS module of sensitive elements containing gyroscopic angular velocity sensors and magnetometers was studied. The module was installed on a rotating stand, simulating the rotation of the spacecraft body at an angular velocity of about 20 °/s. The output signals of magnetometers are shown in Figure 2.

Fig. 2. Output signals of magnetometers on a rotating base

To increase the accuracy of the functioning of the information-measuring system, the displacements of

magnetometer zeros were previously eliminated by centering their output signals plotted on the XY plane (Fig. 3).

Fig. 3. Elimination of offsets of zero signals

Next, their amplitude spectral characteristics were determined, from which the rotation frequency of the stand was estimated (Fig. 4), amounting to 19.975 °/s. The angular velocity of the stand, measured from the MEMS gyroscope signal, was 20.149 °/s. Since the undercompensation of magnetometer zero signals and the inaccuracy of setting their scale factors do not have a significant effect on the calculation of the stand rotation frequency by the spectral method, we can assume that magnetometric measurements are more accurate. By looking for the difference between the angular velocity of the stand thus found and the signal of the MEMS gyroscope, we can establish the offset of its zero signal.

Fig. 4. Spectral characteristics of magnetometric sensors

The estimated offset of the gyroscope's zero signal was 0.174 °/s, with the true value obtained at a stationary position being 0.2 °. The estimation error was 15%.

CONCLUSION

The information-measuring system of a rotation-stabilized spacecraft is considered. A method is shown to compensate for the offset of the zero signal of a MEMS gyroscope based on an analysis of the spectral characteristics of magnetometric sensors.

REFERENCES

[1] Sirotin A. N. On special cases of one problem of optimal control of the angular motion of a symmetrical spacecraft stabilized by rotation // Proceedings of MAI. 2017. No. 96.

[2] Biggs, James & Horri, Nadjim. (2012). Optimal geometric motion planning for a spin-stabilized spacecraft. Systems & Control Letters. 61. 609–616. 10.1016/j.sysconle.2012.02.002.

[3] Artyukhin Yu. P., Kargu L. I., Simaev V. L. Control systems for spacecraft stabilized by rotation. M.: Nauka, 1979. 287 p.

[4] Rozin P.E., Simonov A.V., Gordienko E.S., Zaiko Yu.K. Calibration of the magnetometer of the Descartes spacecraft in flight // Proceedings of MAI. 2022 No. 124 DOI: 10.34759/trd-2022-124-19

Adaptive Control of Small Satellites Constellation in Projection Orbits Using Implicit Reference Model

I. Kostin
Baltic State Technical University «VOENMEH»
Saint Petersburg, Russia
ORCID: 0000-0001-9459-1929

A.A. Shevchik
Baltic State Technical University «VOENMEH»
Saint Petersburg, Russia
ORCID: 0000-0002-6223-0261

J. Fadeeva
Baltic State Technical University «VOENMEH»
Saint Petersburg, Russia
ORCID: 0000-0002-2304-9806

I.A. Yakovlev
Baltic State Technical University «VOENMEH»
Saint Petersburg, Russia
ORCID: 0000-0001-9627-816X

Abstract—**The problem of group control of small satellites of variable mass is considered. Employing the speed-gradient design method, a combined adaptive control law is proposed. The estimation of the unknown satellite mass and direct adaptive control with an implicit reference model concept are realized. A consensus protocol is used to improve the accuracy of the required group formation.**

Keywords—satellite formation, adaptive control, consensus, projected circular orbit

I. Introduction

When solving a number of research tasks, such as Earth sensing, satellite imagery, etc., it turns out to be useful to use a small spacecraft group in Earth orbit. It is necessary to create and maintain the required formation shape. There are several satellite constellation formation methods for creating time-invariant and time-varying configurations. Among the groupings with a constant structure, a class stands out that uses projected circular orbits (PCO) to form the structure. This approach assumes that a group of spacecraft has a leader and follower satellites that surround the leader.

A significant difficulty turns out to be ensuring a given control accuracy when some spacecraft parameters are significantly uncertain. Having its drawbacks, adaptive control is one of the most frequent solutions when the parameters of the control object are uncertain. Limited performance, caused by the need to have time for controller coefficients self-tune, is an insignificant problem when small satellites control.

The purpose of this paper is to develop an adaptive algorithm for group of small satellites with variable mass control. The relevance of this work is determined by the possibility of reducing the requirements to the completeness of current information about the parameters of spacecraft of the group.

II. Mathematical Model

To describe the motion of satellites, an inertial coordinate system XYZ, is used, the origin of which is associated with the Earth's center of mass.

The position vector of the master satellite $\vec{r}_c$ and latitude angle of the chief θ are used to set the position of the leader satellite in the orbit. Latitude angle of the leader θ is the angle between the direction to the pericenter of the orbit and the position vector of the leader satellite $\vec{r}_c$ with its vertex at the Earth's center of mass. The vector $\vec{r}_c$ and the angle θ define the absolute position of the origin of the mobile non-inertial (relative) coordinate system xyz. The mobile coordinate system is rigidly connected to the leader satellite. The axis x is directed along the radius-vector $\vec{r}_c$, the axis y lies in the orbit plane and is perpendicular to the radius-vector in the direction of motion, the axis z completes the triple of vectors to the right. The absolute position of the follower satellite is given by the vector $\vec{r}$, and the position in the relative coordinate system is given by the vector $\vec{\rho}$.

The coordinate systems are shown in Figure 1.

Fig. 1. The coordinate systems.

The motion of the follower satellite is calculated relative to the leader satellite [2]. A general non-linear equation of relative motion is expressed as in [2]:

$$\begin{bmatrix} \ddot{x} - 2\dot{\theta}\dot{y} - \ddot{\theta}y - \dot{\theta}^2 x + \dfrac{\mu}{\gamma}x + \dfrac{\mu}{\gamma}r_c - \dfrac{\mu}{r_c^2} \\[2mm] \ddot{y} + 2\dot{\theta}\dot{x} + \ddot{\theta}x - \dot{\theta}^2 y + \dfrac{\mu}{\gamma}y \\[2mm] \ddot{z} + \dfrac{\mu}{\gamma}z \end{bmatrix} = \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} \qquad (1)$$

$$\ddot{r}_c = r_c \dot{\theta}^2 - \frac{\mu}{r_c^2}, \quad \ddot{\theta} = -\frac{2\dot{r}_c \dot{\theta}}{r_c}, \qquad (2)$$

where x, y and z are state variables to describe the relative position vector $\vec{\rho}$ in x, y and z axes, respectively; a_x, a_y and a_z are the orbital perturbation terms, such as the aspherical geopotential perturbation, thrust, air drag, and solar radiation pressure. $a_j = (f_j + d_j)/m_f \quad \forall j = \{x, y, z\}$, where m_f denoting the mass of the deputy spacecraft and f_j denoting the control input applied by the deputy spacecraft, d_j is for the perturbation term, μ is a gravitational parameter. Finally, γ is defined as:

$$\gamma \equiv |\vec{r}_c + \vec{\rho}| = \left((r_c + x)^2 + y^2 + z^2\right)^{3/2}. \qquad (3)$$

It is convenient to use the matrix form (1):

$$\dot{\mathbf{x}} = \mathbf{A}(\mathbf{x}, t)\mathbf{x} + \mathbf{B}\mathbf{u}, \qquad (4)$$

where state variables are the relative positions and relative velocities of the deputy satellite with regard to the leader satellite.

The term $\left(\dfrac{\mu}{\gamma}r_c - \dfrac{\mu}{r_c^2}\right)$ from (1) can be expressed with multiplier separation $\mathbf{x}$ in the following form [3-4]:

$$\frac{\mu}{\gamma}r_c - \frac{\mu}{r_c^2} = \mu\left(\frac{r_c}{\left(1 + \dfrac{2}{r_c}x + \dfrac{x^2 + y^2 + z^2}{r_c^2}\right)^{3/2}}\frac{1}{r_c^3} - \frac{1}{r_c^2}\right) =$$
$$= \frac{\mu}{r_c^2}\left(\left[1 - \left(-\frac{2}{r_c}x - \frac{x^2 + y^2 + z^2}{r_c^2}\right)\right]^{-\frac{3}{2}} - 1\right). \qquad (5)$$

Let us define

$$\xi \equiv -\frac{2}{r_c}x - \frac{x^2 + y^2 + z^2}{r_c^2} = \left(-\frac{2}{r_c} - \frac{x}{r_c^2}\right)x +$$
$$+ \left(-\frac{y}{r_c^2}\right)y + \left(-\frac{z}{r_c^2}\right)z. \qquad (6)$$

Then, by negative binomial series we obtain the following expression:

$$\frac{\mu}{\gamma}r_c - \frac{\mu}{r_c^2} = \frac{\mu}{r_c^2}\left((1 - \xi)^{-\frac{3}{2}} - 1\right) = \frac{\mu}{r_c^2}\left(1 + \frac{3}{2}\Psi\xi - 1\right) =$$
$$= \frac{3}{2}\frac{\mu}{r_c^2}\Psi\xi, \qquad (7)$$

where

$$\Psi \equiv 1 + \psi_1 + \psi_2 + ..., \quad \psi_1 \equiv \frac{\left(\dfrac{3}{2} + 1\right)}{2}\xi, \quad \psi_2 \equiv \frac{\left(\dfrac{3}{2} + 2\right)}{3}\psi_1\xi, ...$$

Then, taking into account (6), (7):

$$\frac{\mu}{\gamma}r_c - \frac{\mu}{r_c^2} = -\frac{3}{2}\frac{\mu}{r_c^4}\Psi\left((2r_c + x)x + y^2 + z^2\right). \qquad (8)$$

III. ADAPTIVE CONTROL

A. Control goal

Let the master satellite move freely along its trajectory around the Earth, and let the follower satellite move along a given trajectory Let the trajectory be given by the functions:

$$\mathbf{x}_d(t) = \begin{bmatrix} x_d(t) \\ y_d(t) \\ z_d(t) \end{bmatrix}, \quad \dot{\mathbf{x}}_d(\mathbf{t}) = \begin{bmatrix} \dot{x}_d(t) \\ \dot{y}_d(t) \\ \dot{z}_d(t) \end{bmatrix}, \quad \ddot{\mathbf{x}}_d(t) = \begin{bmatrix} \ddot{x}_d(t) \\ \ddot{y}_d(t) \\ \ddot{z}_d(t) \end{bmatrix}. \quad (9)$$

Let us introduce the following error vector:

$$\mathbf{e}_1 \triangleq \mathbf{x} - \mathbf{x}_d, \quad \mathbf{e}_2 \triangleq \dot{\mathbf{x}} - \dot{\mathbf{x}}_d, \quad \mathbf{e} = [\mathbf{e}_1, \mathbf{e}_2] \qquad (10)$$

and let us denote the output signals of the control object as a linear combination of these errors:

$$\mathbf{y} = \boldsymbol{\alpha}\mathbf{e}_1 + \mathbf{e}_2, \qquad (11)$$

where $\boldsymbol{\alpha} = [\alpha_x, \alpha_y, \alpha_z]$ is the vector of positive constants.

Taking into account the new notations, the system (1) will take the following general form:

$$\dot{\mathbf{e}} = \bar{\mathbf{A}}(\mathbf{e}, \mathbf{t})\mathbf{e} + \bar{\mathbf{B}}\bar{\mathbf{u}}, \quad \mathbf{y} = \mathbf{C}\mathbf{e}, \qquad (12)$$

where matrices:

$$\bar{\mathbf{A}}(\mathbf{e}, t) = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ a_{41} & a_{42} & a_{43} & 0 & a_{45} & 0 \\ a_{51} & a_{52} & 0 & a_{54} & 0 & 0 \\ 0 & 0 & a_{63} & 0 & 0 & 0 \end{bmatrix}, \qquad (13)$$

$$\bar{\mathbf{B}} = \frac{1}{m_f}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}^{\mathrm{T}}, \qquad (14)$$

$$\mathbf{C} = \begin{bmatrix} \alpha_x & 0 & 0 & 1 & 0 & 0 \\ 0 & \alpha_y & 0 & 0 & 1 & 0 \\ 0 & 0 & \alpha_z & 0 & 0 & 1 \end{bmatrix}, \qquad (15)$$

где

$$a_{41} = \dot{\theta}^2 - \frac{\mu}{\gamma} + \frac{3}{2}\frac{\mu}{r_c^4}\Psi(2r_c + x), \quad a_{51} = -\ddot{\theta},$$

$$a_{42} = \ddot{\theta} + \frac{3}{2}\frac{\mu}{r_c^4}\Psi y, \quad a_{52} = \dot{\theta}^2 - \frac{\mu}{\gamma}, \quad a_{43} = \frac{3}{2}\frac{\mu}{r_c^4}\Psi z,$$

$$a_{63} = -\frac{\mu}{\gamma}, \quad a_{54} = -2\dot{\theta}, \quad a_{45} = 2\dot{\theta}.$$

Let the target function be:

$$Q_t(\mathbf{e}_1, \mathbf{y}) = \frac{1}{2}\mathbf{e}_1^T\mathbf{e}_1 + \frac{m_f}{2}\mathbf{y}^T\mathbf{y}. \qquad (16)$$

The control goal can be formulated as follows:

$$\lim_{t \to \infty} Q_t(\mathbf{e}_1, \mathbf{y}) = 0. \qquad (17)$$

B. Control loop synthesis

We will consider the inputs in the form:

$$
\begin{aligned}
u_1 &= -\hat{m}\left(k_x y_x - \ddot{x}_d(t)\right) + \hat{k}_x y_x, \\
u_2 &= -\hat{m}\left(k_y y_y - \ddot{y}_d(t)\right) + \hat{k}_y y_y, \qquad (18)\\
u_3 &= -\hat{m}\left(k_z y_z - \ddot{z}_d(t)\right) + \hat{k}_z y_z,
\end{aligned}
$$

where $\left(y_x, y_y, y_z\right)$ are components of vector $\mathbf{y}$, $\hat{m}, \hat{k}_x, \hat{k}_y, \hat{k}_z$ are estimated parameters and $k_x > 0, k_y > 0, k_z > 0$ are constant parameters.

It was shown in [5] that for the system (12) with inputs (18) and known value of m_f it is always possible to find parameters $\hat{m}, \hat{k}_x, \hat{k}_y, \hat{k}_z, k_x, k_y, k_z$ such that control goal (17) is achieved and the closed-loop system is exponentially stable.

To obtain the adaptive tuning laws for the parameters of the control law (18), the speed-gradient design method is used [6–7].

The time derivative of the goal function (16) can be written as follows:

$$\dot{Q}_t = \mathbf{e}_1^T\mathbf{e}_2 + m_f\mathbf{y}^T\dot{\mathbf{y}}. \qquad (19)$$

The vector of function (19) has a gradient with respect to the adjustable parameters:

$$\nabla\dot{Q}_t = \begin{bmatrix} -k_x y_x^2 + y_x\ddot{x}_d - k_y y_y^2 + y_y\ddot{y}_d - k_z y_z^2 + y_z\ddot{z}_d \\ y_x^2 \\ y_y^2 \\ y_z^2 \end{bmatrix} \qquad (20)$$

Then tuning algorithms for variables are:

$$
\begin{aligned}
\dot{\hat{m}} &= -\gamma_m\left(-k_x y_x^2 + y_x\ddot{x}_d - k_y y_y^2 + y_y\ddot{y}_d - k_z y_z^2 + y_z\ddot{z}_d\right), \\
\dot{\hat{k}}_x &= -\gamma_x y_x^2, \quad \dot{\hat{k}}_y = -\gamma_y y_y^2, \quad \dot{\hat{k}}_z = -\gamma_z y_z^2,
\end{aligned} \qquad (21)
$$

where $\gamma_m, \gamma_x, \gamma_y, \gamma_z$ are the constant adaptation gains.

IV. ADAPTIVE CONTROL OF GROUP OF SATELLITES

A. Basic information on consensus algorithm

We define graph G as a pair (V, E), where $V = 1,2,...,n$ is a set of nodes, and $E \in V \times V$ is a set of edges in which each edge is represented by an ordered pair of different nodes. For example, edge shows that node j j can receive information from node i. Let graph G contain n nodes. The adjacency matrix $A = \left[a_{ij}\right] \in \mathbb{R}^{N \times N}$ is defined as: $a_{ii} = 0$, $a_{ij} = 1$, if $\left(v_i, v_j\right) \in E$, and otherwise $a_{ij} = 0$.

Assume that $\xi_i \in \mathbb{R}$ contains information about the state of the i-th agent. For information states with first-order dynamics, the following fundamental first-order consensus algorithm is known [8]:

$$\dot{\xi}_i = u_i, \quad u_i = \rho_0\sum_{j=1}^{n} a_{ij}\left(\xi_i - \xi_j\right), \quad \rho_0 > 0 \text{ is a certain}$$

constant.

Consensus is achieved asymptotically among all agents when for any $\xi_i(0)$ it is valid that $\left\|\xi_i(t) - \xi_j(t)\right\| \to 0$, for all $i \neq j$ as $t \to \infty$.

B. Adaptive control of satellites formation based on the consensus protocol

Extended control algorithm for i-th agent with additional control signals can be written as follows [8]:

$$
\begin{aligned}
v_{1i} &= u_1 - \rho_x\sum_{j=1}^{N} a_{ij}\left(y_{1i} - y_{1i}\right), \\
v_{2i} &= u_2 - \rho_y\sum_{j=1}^{N} a_{ij}\left(y_{2i} - y_{2i}\right), \qquad (22)\\
v_{3i} &= u_3 - \rho_h\sum_{j=1}^{N} a_{ij}\left(y_{3i} - y_{3i}\right),
\end{aligned}
$$

where u_1, u_2, u_3 are adaptive control components, that can be calculated by formula (18), $\rho_x > 0$, $\rho_y > 0$, $\rho_h > 0$ are parameters of consensus algorithm, a_{ij} are adjacency matrix coefficients.

Control (22) improves formation positioning accuracy in adaptive control of small satellites with uncertain masses.

V. SIMULATIONS

A computer simulation was performed in the Matlab/Simulink system for various scenarios of the created adaptive control algorithm.

The motion of four satellites along a PCO-orbit of radius 1 km was considered. The following parameters were assumed for the calculations: $\mu_e = 398.600\,\text{km}^3\,\text{sec}^{-2}$; perigee radius $R_p = 6971$ km; eccentricity $e = 0.2$, $\Omega = \omega = i = M = 0$ rad.

The initial states of the satellites (km): $x_{10} = 1$, $x_{20} = 0.7$, $x_{30} = -1$, $x_{40} = -0.5$, $y_{10} = 2$, $y_{20} = -2$, $y_{30} = -2$, $y_{40} = 1$, $z_{10} = 2$, $z_{20} = 2$, $z_{30} = -2$, $z_{40} = -2$.

The reference trajectory is given by the relations:

$$\begin{bmatrix} x_d \\ y_d \\ z_d \end{bmatrix} = \begin{bmatrix} (r_d/2)\sin(nt+\phi) \\ r_d\cos(nt+\phi) \\ r_d\sin(nt+\phi) \end{bmatrix}, \qquad (23)$$

where r_d is PCO-orbit radius, n is average angular velocity of the leader, $\phi = \pi/2$ is phase shift between vehicles in orbit.

The disturbance model including gravitational disturbances, atmospheric drag, and solar wind effects was specified as follows [9]:

$$\begin{bmatrix} d_x \\ d_y \\ d_z \end{bmatrix} = 1.2 \times 10^{-3} \begin{bmatrix} 1 - 1.5\sin(nt) \\ 0.5\sin(2nt) \\ \sin(nt) \end{bmatrix}, \qquad (24)$$

The initial masses of the satellites are 15 kg and decrease linearly with velocities: -0.68, -0.45, -0.3, and -0.5 (kg/h), respectively. Also, during the simulation, the second, third and fourth satellites lose 3, 5 and 2 kg, respectively. The graphs of the vehicles' masses changes are shown in Figure 2.

Fig. 2. Graphs of satellites' masses changes during simulation.

In (11) the following output parameters were used: $\alpha_x = \alpha_y = \alpha_z = 8.5 \cdot 10^{-4}$. The constant parameters of the control law (18) were taken: $k_x = k_y = k_z = 0.0255$. The parameters of the adaptation algorithm (21) were set: $\gamma_m = 10^6$, $\sigma_m = 8 \cdot 10^{-9}$, $\gamma_x = \gamma_y = \gamma_z = 5 \cdot 10^{-5}$, $\sigma_x = \sigma_y = \sigma_z = 8 \cdot 10^{-9}$. The initial values of the estimated parameters were set: $m_0 = 12.5$ kg, $\hat{k}_{x0} = \hat{k}_{y0}, = \hat{k}_{z0} = -135$. The controls were limited within ± 0.2 H.

The parameters of the consensus protocol were set: $\rho_x = \rho_y = \rho_z = 15.5$. The adjacency matrix was taken:

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}.$$

Simpler versions of the information graphs under dynamic topology changes are considered in article [10].

The trajectory of a satellites group after simulation is presented in Figure 3. The projection of the trajectory is shown in Figure 4.

VI. Conclusion

In this paper, an algorithm for adaptive flight control for the group of variable mass small satellites in projection orbits was developed. The group control is implemented using a decentralized approach based on a consensus algorithm. Computer simulation of the obtained algorithm has shown

the effectiveness of the proposed approach under conditions of variable mass of small satellites.

Fig. 3. The trajectory of a satellites group.

Fig. 4. The projection of the trajectory of a satellites group.

References

[1] Andrievsky, B.; Kuznetsov, N.; Popov, A. Algorithms for aerodynamic control of relative motion two satellites in a near circular orbit. Differencialnie Uravnenia i Protsesy Upravlenia 2020, pp. 28–58. (In Russian).

[2] K. Alfriend, .S. Vadali, P. Gurfil. Spacecraft Formation Flying. Butterworth-Heinemann: Oxford, 2010.

[3] H. E. Park, S. Y. Park, and K. H. Choi, "Satellite formation reconfiguration and station-keeping using state-dependent Riccati equation technique". Aerospace Science and Technology 2011, vol. 15, pp. 440–452.

[4] A. M. Popov; I. Kostin; J. Fadeeva; B. Andrievsky. Development and Simulation of Motion Control System for Small Satellites Formation. Electronics 2021, 10, 3111. https://doi.org/10.3390/electronics10243111

[5] Popov, A.M.; Kostrygin, D.G.; Shevchik, A.A.; Andrievsky, B. Speed-Gradient Adaptive Control for Parametrically Uncertain UAVs in Formation. Electronics 2022, 11, 4187.

[6] Fradkov, A.L. Speed-gradient scheme and its application in adaptive control problems. Autom. Remote Control 1980, 40, 1333–1342.

[7] Andrievsky, B.R.; Fradkov, A.L. Speed Gradient Method and Its Applications. Autom. Remote Control 2021, 82, 1463–1518.

[8] Ren, W.; Beard, R.W. Distributed Consensus in Multi-Vehicle Cooperative Control; Springer: London, UK, 2008.

[9] Godard; K. Kumar. Fault Tolerant Reconfigurable Satellite Formations Using Adaptive Variable Structure Techniques. J. Guid. Control Dyn. 2010, 33, 969–984.

[10] A. M. Popov, D. G. Kostrygin, P. V. Krashanin and A. A. Shevchik, "Development of Algorithm for Guiding the Swarm of Unmanned Aerial Vehicles," 2022 29th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), Saint Petersburg, Russian Federation, 2022, pp. 1-4, doi: 10.23919/ICINS51784.2022.9815370.

Experimental Estimation of the Dynamic Error of Navigation Aids Using the Spectral Method

D.G. Gryazin

Concern CSRI Elektropribor, JSC,
ITMO University
Saint Petersburg, Russia
volnagdg@yandex.ru

O.O. Belova

Concern CSRI Elektropribor, JSC,
Saint Petersburg, Russia
olga_o_belova@mail.ru

Abstract—**The spectral methods for estimating the dynamic errors of devices and systems under near-real conditions are proposed. The methods are based on reproducing the oscillations in a given frequency spectrum and are adapted for use on testing equipment available at Concern CSRI Elektropribor. The methods were applied to experimentally evaluate the dynamic errors of electronic roll and pitch inclinometer and magnetic compass. The laboratory results are consistent with the earlier field results. The dynamic error of the electronic roll and pitch inclinometer has been estimated with a traditional method - by setting successive harmonic oscillations with the known frequency and amplitude. It is shown that the error obtained with the traditional approach does not fully reflect the real error under the operating conditions. It is concluded that to estimate the dynamic errors of devices in laboratory conditions, the inputs typical of specific operating conditions should be reproduced.**

The developed methods simplify the experimental estimation of dynamic error when monitoring the characteristics of serial products, reduce test time from several hours to 15-20 minutes and provide more reliable results.

Keywords - dynamic error, spectral density, operating conditions, tests, electronic roll and pitch inclinometer, magnetic compass

I. Introduction

Dynamic error is determined by the difference between the measuring instrument error in dynamic mode and the static error at the current time [1]. For the navigation devices operating in fast changing external conditions, the dynamic errors often determine their performance.

The most complete information on the device dynamic error can be obtained by experimental studies of its dynamic characteristics. However, classical known methods for determining the dynamic characteristics of systems are rather time-consuming. Moreover, they do not consider the specific operating modes of the instruments on board the ship. The shipboard attitude determination devices are mostly affected by polyharmonic disturbances with a limited frequency spectrum. The real error in this case may greatly differ from the error throughout the bandwidth determined by the device frequency characteristics.

II. Problem Statement

Analysis of performance testing methods for navigation devices shows that their dynamic errors are generally not standardized or studied during acceptance tests due to the complex experimental work, long measurements, and the lack of dynamic test procedures. However, it is these

characteristics that in most cases determine the performance of the developed equipment.

The most widely used and easily implemented method for dynamic error estimation is based on comparing the parameters of successively recorded harmonic inputs at a selected frequency in a certain bandwidth with the parameters of the device output signal [2]. However, this method fails to estimate the dynamic errors of devices and systems exposed to disturbances typical of their operating conditions, such as rolling and pitching under waves with the parameters characterizing the real operating conditions. In this case, the harmonic inputs should be set successively at each frequency throughout the studied frequency range, which seems expensive.

Obviously, in serial production of navigation devices, the dynamic errors should be estimated in minimal time in order to reduce costs. To do this, the disturbing inputs should not be simulated successively at each frequency, rather, they should be made pseudo-random, stationary, and corresponding to a given frequency spectrum. The frequency spectrum should be selected based on the device real operating conditions. The spectral characteristics of the error should be calculated by switching from the time to the frequency domain using the Fourier transform. It is important to know whether the processes at the input and output of the system are correlated. Correlated processes is the most common case, since in estimating the device characteristics, the input (command) and output (device signal) depend on each other. However, noncorrelated signals are also possible, for example, the sum of the useful signal and random noise from external influences.

The objective of the research is to develop methods for estimating the dynamic errors of navigation devices, which refine the error value in specific operating conditions and reduce the time of experimental operations in serial production.

III. Spectral Methods for Experimental Estimation of Dynamic Errors of Navigation Aids

The paper proposes spectral methods for experimental estimation of dynamic errors of navigation aids considering the degree of correlation of input and output processes.

Consider the developed method for estimating the dynamic error for correlated signals [3].

Let $x(t)$ be the input signal from the test bench (reference input), $y(t) = x(t) + \varepsilon(t)$ is the device output signal, where $\varepsilon(t)$ is its dynamic error. Then $\varepsilon(t) = y(t) - x(t)$.

The device input and output signals are correlated, so the correlation error function $R_\varepsilon(\tau)$ is determined by (1):

$$R_\varepsilon(\tau) = R_x(\tau) + R_y(\tau) - R_{xy}(\tau) - R_{yx}(\tau), \quad (1)$$

where $R_x(\tau)$ is the autocorrelation function of the test bench signal, $R_y(\tau)$ is the autocorrelation function of the device output signal, $R_{xy}(\tau)$ and $R_{yx}(\tau)$ are the cross-correlation functions of the test bench and device signals, τ is the correlation interval.

The error spectral density $S_\varepsilon(f)$ is defined as a one-way Fourier transform of the error correlation function [4]:

$$S_\varepsilon(\omega) = 2\int_0^\infty R_\varepsilon(\tau)e^{-j\omega\tau}d\tau . \quad (2)$$

As a result, we obtain the device error spectral density, which characterizes its error variance throughout the frequency range. It is defined as the area under the curve of the resulting spectral density (3):

$$D = \sigma^2 = \frac{1}{\pi}\int_0^\infty S(\omega)d\omega . \quad (3)$$

The root-mean-square deviation of the device dynamic error is determined as $\sigma = \sqrt{D}$.

If the input is not correlated with the output signal, the variance of the sensor or module error will be equal to the difference between the variances of the output and input signals $D_\varepsilon = D_y - D_x$. Thus, the error variance can be defined as the difference between the areas under the spectral density curves for the input and output signals.

Note that the appearance of the spectral density graph characterizes the device dynamic error and reveals, e.g., low-frequency drift or high-frequency noise.

IV. VERIFICATION OF DEVELOPED SPECTRAL METHODS

We have verified the proposed methods. The studies were carried out in the laboratory using test bench equipment providing oscillations with a given period and amplitude, and polyharmonic oscillations with a given frequency spectrum. The dynamic yaw error of the magnetic compass and dynamic error of the electronic roll and pitch inclinometer [5] during roll were estimated in laboratory studies. The units under test were placed on the test platform. When testing the magnetic compass, the test bench axis was positioned vertically to simulate the ship yawing; when testing the roll and pitch inclinometer, it was positioned horizontally to simulate roll. Note that during the magnetic compass tests it was needed to exclude additional magnetic fields affecting the device readings. For this purpose, the test bench with the installed compass was installed inside a shielded electromagnetic chamber/stand, where different values of the vertical and horizontal components of the Earth's magnetic field could be simulated.

At the first step of experimental studies, the dynamic error of the roll and pitch inclinometer was estimated with the traditional method, assuming deterministic approach to testing: successive comparison of harmonic oscillations of known frequency and amplitude set by the test bench and the obtained device measurements. The tests were carried out under roll and pitch periods from 3 to 10 s and amplitudes of ±2°, ±5° and ±10°. The RMS deviation under the specified frequencies at different amplitudes ranges from 0.028° to 0.32°.

At the second step, we estimated the dynamic errors of the magnetic compass and roll and pitch inclinometer using the proposed spectral method and compared the laboratory results with the field test results. Based on field data, the amplitude spectra of the ship's roll, pitch, and yaw were constructed. Next, the amplitudes and periods of the harmonic components of the test bench oscillations were selected to simulate the conditions that were closest to real ones (Fig. 1). Note that the issue of methods for smoothing the spectral density graph when selecting the parameters of the harmonic components remains open [6]. When constructing the amplitude spectra, we have achieved the required smoothness of the curves, and at the same time have not greatly distorted the variance.

Fig. 1. Amplitude spectra constructed from real samples (dashed line) and from bench data (solid line).

The devices under study were installed on the test bench, and polyharmonic oscillations in the preset frequency spectrum were simulated with simultaneous recording of data from the test bench and devices (roll and pitch inclinometer and magnetic compass). Based on these samples, the error correlation functions were calculated using (1). After their Fourier transformation, graphs of the error power spectral density for the electronic inclinometer and magnetic compass were constructed (Fig. 2).

The error variance of the devices was determined as the area under the indicated graphs. The RMS deviation was 0.36° for the inclinometer, and 1.2° for the magnetic compass. The RMS deviations obtained in field studies by comparison with reference instruments was 0.28° for the inclinometer, and 1.68° for the magnetic compass. Analysis of the findings showed good agreement between the experimental and field test results. Moreover, the dynamic error of the roll and pitch inclinometer, calculated using the deterministic approach, significantly differs from the values obtained using the spectral approach and field experiment, i.e., traditional methods fail to estimate the real error of the devices under their operating conditions. Therefore, in experimental laboratory studies of the errors of navigation devices, it is reasonable to set oscillations in the frequency

spectrum rather than reproduce individual harmonic oscillations with the known frequency and amplitude.

Fig. 2. Error spectral densities for the inclinometer (a) and magnetic compass (b).

V. CONCLUSIONS

The proposed method for experimental estimation of the dynamic errors of navigation devices based on spectral methods significantly reduces the test time (up to 15-20 minutes), which is extremely important for serial-produced products, and provides more reliable results for the specific operating conditions of the devices.

REFERENCES

[1] RMG 29-2013. State system for ensuring the uniformity of measurements. Metrology. Basic terms and definitions. Moscow: Standartinform, 2014.

[2] V.A. Granovskiy, Dinamicheskie izmereniya: Uchebnoe posobie (Dynamic measurements: Study Guide), Leningrad: SZPI Publ., 1987.

[3] D.G. Gryazin and O.O. Belova, "Experimental method for estimating the dynamic error of devices and sensors under their operating conditions", Scientific and Technical Journal of Information Technologies, Mechanics and Optics, 2023, vol. 23, no 3. pp. 652–658.

[4] A.A. Sveshnikov, Prikladnye metody teorii sluchainykh funktsii (Applied Methods of the Random Functions Theory), Moscow: Nauka Publ., 1968.

[5] D.G. Gryazin and T.V. Paderina, "Development of a MEMS-based electronic inclinometer free from the effect of translational accelerations", Gyroscopy and Navigation, 2023, vol. 14, pp. 129–137.

[6] Hangfang Zhao, Lin Gui, "Nonparametric and parametric methods of spectral analysis", MATEC Web of Conferences 283, 07002 (2019) FCAC 2018/

Algorithm for Constructing a Digital Model of the Subsurface Layer of a Vehicle's Route Using a Neural Network in a Georadar Measurement Model

R.R. Bikmaev

Navigation and geodetic systems dept.

ANO «Institute of Engineering Physics»

Serpukhov, Russian Federation

e-mail: naviserp5@iifmail.ru

Abstract — **A digital model of the subsurface layer is proposed that increases the sensitivity of vehicle state error detection and the correction accuracy of the dead reckoning.**

Keywords — digital subsurface model, GPR, radarogram, measurement model.

I. INTRODUCTION

In dead reckoning (DR) used in autonomous mobile vehicles, gyroazimuth drift with a quadratic time-increasing law has the greatest influence on the increase of on-board navigation error [1]. To correct the error, navigation information is integrated with global position data from the Satellite Navigation System (SNS). Currently, there are many solutions to combine navigation information and gyroazimuth drift estimation. For example, one approach is to apply the Error State Extended Kalman Filter to estimate the biases of accelerometers and gyroscopes with slow dynamics [2]. Correction of the growing error in the DR remains relevant due to the fact that offline global position data are determined from analyzing the gradient of the Earth's geophysical fields. The radio locational contrast of the Earth's subsurface has a natural physical field gradient and is stable over long periods of time. The study of the radio locational contrast of the Earth's subsurface is used by geophysics to determine the structure of the Earth's surface, to excavate underground objects, and to inspect road structures or pavements. Ground penetrating radar (GPR) with pulsed or step-frequency emission of electromagnetic signals is used for this purpose [3].

II. APPROACHES

At the present time an approach has been proposed to estimate the state of a moving vehicle by comparing the current array of radar reamers contrast received from GPR with antenna array with a priori map of radar reamers [4]. In this approach, preliminary GPR measurements with eleven channels were used for map preparation which were weighted averages of the previous measurement and converted to a "grid". Electromagnetic contrast measurements were referenced to geodetic location data obtained by the SNS. Current GPR measurements underwent identical signal processing and conversion to a "grid" map to determine the location. Region of interest with linked SNS data was defined on the map where vehicle was highly likely to be located for place recognition. A location was searched in a defined array of radar reamers by calculating the correlation coefficient (rough search). Next, a particle swarm optimization algorithm was applied to determine the three orientation angles and refine the position by calculating the vehicle position (accurate search). The method used allowed us to obtain navigation data of global position every second when linked with SNS data.

In the approach described in [5], which used the same GPR with antenna array, the data from the sensor were integrated with inertial navigational system measurement and odometer by applying the Kalman filter. Moreover, the data from GPR was combined by separate filtering due to absence of global location data in a priori map in places of poor satellite signal reception. When comparing estimates in different weather conditions, GPR measurements obtained in rain and snow were less correlated with the a priori map than in clear weather condition.

In another study [6], a method was proposed in which single-channel GPR data were used to correct the increasing error of the inertial system by calculating the correlation of scalar radargram data obtained from GPR measurements of the same track section in different directions. To obtain navigation data a neural network was trained to predict the relative motion of the robot between inconsistent pairs of radargram images. The correction of the autonomous navigation system was carried out as a logical derivation of the maximum posterior probability by optimizing the hidden states and constraints of the factor graph. The contribution of this approach to the presented work is data obtained synchronously from measurement sensors (odometer, IMU, GPR and total station) and published in the public domain of the Internet [7].

The key scientific results of the proposed approach are:

- GPR measurement model, that establish the dependence of radargram illumination gradient variation on vehicle geodetic location data and velocity;

- construction of a digital model of subsurface layer of a vehicle route on the basis of determining the parameters of combined neural network (convolutional and recurrent) by machine supervised learning;

- an algorithm for correcting the increasing error of the DR using a digital subsurface model of the vehicle route and a Kalman filter with error state estimation.

III. THE PROBLEM STATEMENT

Assume that the vehicle starts moving from a starting point in the navigation coordinate system $OXYZ$. The on-board navigation system determines the distance increment by means of an odometer and the orientation by an inertial measurement unit (IMU). The vehicle is equipped with a

single-channel GPR that receives the reflected electromagnetic signal from the Earth's subsurface in the form of amplitude reports. They are converted into a radargram, which is deployed in a two-dimensional image plane by signal reflection depth and path travelled. For the reference estimation of the vehicle trajectory, measurements of its position in the geodetic coordinate system are carried out using a total station in the auto surveillance mode.

A. Features of condition estimation

The general idea of using Error State Extended Kalman Filter is to represent the true state as a composite of the nominal state x **,** as a large-valued signal, and the error state δx, as a small-valued signal. The nominal state integrates into an IMU navigation data without taking into account model disadvantages. They accumulate in the form of noise in the error state δx, which is estimated by linear-Gaussian filtering.

B. Nominal state

Nominal state x is a vector $\begin{bmatrix} p_x & p_y & p_z & v_x & v_y & v_z & q_w & q_x & q_y & q_z \end{bmatrix}^T$, that varies in position, velocity and orientation (in quaternions) in the motion model presented below:

$$p_k = p_{k-1} + v_{k-1}\Delta t + (R_{GL}a_{k-1} - g)\frac{\Delta t^2}{2}$$

$$v_k = v_{k-1} + (R_{GL}a_{k-1} - g)\Delta t$$

$$q_k = \Omega(q(\omega_{k-1}\Delta t))q_{k-1}$$

where $a = \begin{bmatrix} a_x & a_y & a_z \end{bmatrix}$ - accelerations, $\omega = \begin{bmatrix} \omega_x & \omega_y & \omega_z \end{bmatrix}$ - angular velocities, measured by IMM.

C. Error state

Error state vector $\delta x = \begin{bmatrix} \delta p_x & \delta p_y & \delta p_z & \delta v_x & \delta v_y & \delta v_z & \delta\theta_x & \delta\theta_y & \delta\theta_z \end{bmatrix}^T$ describes the orientation in Euler angles and varies according to the following dependence:

$$\delta x_k = F_{k-1}\delta x_{k-1} + L_{k-1}n_{k-1}, \text{ where transformation matrix}$$

$$F_{k-1} = \begin{bmatrix} I_{3\times3} & I_{3\times3}\Delta t & 0 \\ 0 & I_{3\times3} & -[R_{GL}a_{k-1}]_\times\Delta t \\ 0 & 0 & I_{3\times3} \end{bmatrix}, \text{ covariance matrix}$$

$$L_{k-1} = \begin{bmatrix} 0 & 0 \\ I_{3\times3} & 0 \\ 0 & I_{3\times3} \end{bmatrix}, \text{ IMU measurement noise}$$

$$n_k \propto N\left(0, Q_k = \Delta t^2 \begin{bmatrix} \sigma_{acc}^2 & \\ & \sigma_{gyro}^2 \end{bmatrix}\right).$$

D. Prediction step

The covariance matrix of state estimation error is defined using the following equation:

$$\breve{P}_k = F_{k-1}P_{k-1}F_{k-1}^T + L_{k-1}Q_{k-1}L_{k-1}^T.$$

E. Correction step

At the correction step, the error state vector is determined by calculating the Kalman gain coefficient:

$$K_k = \breve{P}_k H_k^T (H_k \breve{P}_k H_k^T + R)^{-1}, \text{ where } R \text{ - covariance}$$

matrix of measurement noise of correction sensors.

The error state vector is calculated as $\delta x_k = K_k(y_k - x_k)$, where y_k - sensor measurement vector. Correction of the prediction step is determined by the following dependencies:

$$\hat{p}_k = \breve{p}_k + \delta p_k, \hat{v}_k = \breve{v}_k + \delta v_k, \hat{q}_k = \Omega(q(\delta\theta))\breve{q}_k$$

The covariance matrix is updated by the formula: $\hat{P} = (I - K_k H_k)\breve{P}_k(I - K_k H_k)^T + K_k R K_k^T$.

After the correction step, the error state vector is reset to zero.

F. Measurement model

For the dead reckoning system measurement vector y_k is determined as follows:

$$y_k = \begin{bmatrix} p_x & p_y & 0 & v_x & v_y & 0 & \theta_x & \theta_y & 0 \end{bmatrix}.$$

Accordingly, the measurement matrix H will be as follows:

$$H = Diag\begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}.$$

GPR receives a reflected electromagnetic signal from the subsurface layer of the Earth in the form of amplitude reports, which are converted into a 2-dimensional image, one axis of which is determined by the depth of signal reflection, and the second - by the travelled path. The depth of signal reflection is not informative for determining the navigation model of the measurement due to the motion over the Earth's surface, but does affect the representation of the illumination gradient.

Due to the lack of a coupling equation between GPR measurements and navigation data, it was necessary to synthesise an algorithm for constructing a digital subsurface model (DSM) of the vehicle route, that can be used to determine the required measurement vector y_k (Fig.1).

Fig. 1. The problem statement

IV. Solution

The algorithm for constructing the DSM of the vehicle route includes three stages:

- submap generation from reflected electromagnetic signals from the Earth's subsurface;

- place recognition from geo-referenced submap by using Siamese convolutional neural network model;

- orientation prediction from a sequence of vehicle orientations using a model of a gated recurrent neural network.

A. Submap generation

Submaps in the form of short radarograms (images) (Fig. 2) were synthesized from reflected electromagnetic signals from the Earth's subsurface using the algorithm given in [7].

The resulting digital subsurface models in the form of submap images were georeferenced to total station geodetic measurements. Each submap included an average of 25 sequences from the vehicle orientation.

Fig. 2. The submap image is sampled at 350 ms, shifting with a window size of 14 ms (same as the total station measurement frequency).

B. Place recognition

A similarity function s was used to recognize the place, which determines the distance between feature vectors $g_{1,L-1}$ and $g_{2,L-1}$, obtained from two "twin" neural networks using sigmoid in activation at the output layer.

$$s = \sigma(\sum_j (g_{1,L-1}^j - g_{2,L-1}^j)^2) \qquad (1)$$

where L is the number of hidden layers.

C. Orientation prediction

For orientation prediction, a gated recurrent neural network with 25 modules according to the number of latent orientation states analyzed was used, which defined orientation as a state $c^{<t>}$ according to the following dependency:

$$c^{<t>} = \Gamma_u \circ \tilde{c}^{<t>} + (1 - \Gamma_u) \circ c^{<t-1>} \quad (2)$$

whrre Γ_u - hidden state variation gate, $\tilde{c}^{<t>}$ - intermediate state, $c^{<t-1>}$ - previous state.

D. Measurement model

The synthesized algorithm for the construction of DSM (Fig.3) made it possible to determine the measurement vector y_k as follows:

$$y_k = \begin{bmatrix} p_x & p_y & p_z & 0 & 0 & 0 & \theta_x & \theta_y & \theta_z \end{bmatrix}.$$

Accordingly, the measurement matrix was:

$$H = Diag\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}.$$

Fig. 3. Digital subsurface model

V. Experiment

A. Experimental data

CMU-GPR data [7] contains measurements of accelerometers and gyroscopes in three axes, magnetometer, encoder and GPR, obtained by a special experimental vehicle SuperVision. The dataset with the largest odometric error was selected for the experiment.

B. Experiment results

To estimate the algorithm for constructing the DMS of vehicle route, weight parameters for two neural network models were obtained by machine supervised learning. Learning quality was reviewed on test data using well-known metrics: accuracy, precision and recall. The selected weight parameters of the models were used in data processing in two modes: with application of only the dead reckoning system in the correction step in the Kalman filter and with application of GPR and dead reckoning system in the correction step. The results obtained from the two modes are shown in Fig. 4.

For quantitative estimation, the standard deviation of the vehicle position in the geodetic coordinate system was calculated. The results of determination are given in Table 1.

TABLE I. Standard diviation with and without GPR

Trajectory types	Standard deviation		
	X	Y	Z
Without GPR	1,11 м	2.13 м	0.33 м
With GPR	0.13 м	0.21 м	0.05 м

Fig. 4.   Comparison of measured, estimated and reference trajectories

In view of applying Gaussian filtering to the error state, the dynamics of position and orientation error were analysed in two modes. Comparative graphs are given in Figs. 5 and 6.

Fig. 5.   Comparison of position error in two modes (maximum error covariance in red, error amplitude in blue)

Fig. 6.   Comparison of orientation error in two modes (maximum error covariance in red, error amplitude in blue)

VI. CONCLUSION

By adding GPR data to the Kalman filter correction step, which were processed by the DMS construction algorithm, the sensitivity of the system to location and orientation errors has been improved as shown in Figs. 5 and 6.

REFERENCES

[1] Kotov, N.I., Fizicheskie i teoreticheskie osnovy moreplavaniya: Teksty lektsii. RF. (Physical and Theoretical Foundations of Navigation: Lectures). Serpukhov: Minoborony MO RF, 2007.

[2] Brigadnov I., Lutonin A., Bogdanova K. "Error State Extended Kalman Filter Localization for Underground Mining Environments", 2023.

[3] Daniels, D. J. Ground penetrating radar, volume 1, Institution of Engineering and Technology. London, United Kingdom. 2004

[4] M. Cornick, J. Koechling, B. Stanley, and B. Zhang, "Localizing Ground Penetrating RADAR: A Step Toward Robust Autonomous Ground Vehicle Localization," Journal of Field Robotics, pp. 82–102, 2016.

[5] T. Ort, I. Gilitschenski, and D. Rus, "Autonomous navigation in inclement weather based on a localizing ground penetrating radar," IEEE Robotics and Automation Letters,pp. 3267–3274, 2020.

[6] A. Baikovitz, P. Sodhi, M. Dille, and M. Kaess, "Ground encoding: Learned factor graph-based models for localizing ground penetrating radar," in arXiv, 2021.

[7] A. Baikovitz, P. Sodhi, M. Dille, and M. Kaess, "CMU-GPR Dataset: Ground Penetrating Radar Dataset for Robot Localization and Mapping," in arXiv, 2021

Method of Combined Micronavigation Measurements in a Quasi-Continuous Wave Radar during a Survey of Airspace against the Background of the Earth

V. Kovregin
St. Petersburg State University of Aerospace Instrumentation
St. Petersburg, Russia
kovregin@mail.ru

G. Kovregina
St. Petersburg State University of Aerospace Instrumentation
St. Petersburg, Russia
g_kovregina@mail.ru

Abstract – A method is presented for measuring micronavigation parameters in a radar, combined with a view of the airspace "against the background of the earth" with quasi-continuous radiation. The effectiveness of using measurements obtained by this method to correct the micronavigation system of a radar in case of radio interference to satellite navigation channels is shown.

Keywords – multifunctional radar; quasi-continuous radiation; inertial-Doppler micronavigation

I. INTRODUCTION

The architecture of modern multifunctional aircraft-based radars (MFRs), as a rule, assumes a "built-in" inertial-satellite micronavigation system (MNS), used to determine the trajectory parameters of the antenna phase center (PCA) and the orientation of the main beam (MB) of the MFR antenna radiation pattern at relatively short intervals (no more than 10-15 s) when synthesizing the aperture and coherent accumulation of a "weak" signal from an airborne object [1-3]. If there is interference with satellite navigation systems (SNS), there is a need to use alternative/additional noise-resistant and sufficiently high-precision corrective sensors in the information structure of the MNS. For example, in work [4], the Doppler speed and drift angle meter (DSDA) from the carrier's navigation complex was considered in this capacity. However, in a number of practical cases, the effective integration of DSDA and MNS is difficult due to delays in the interface, measurements, asynchrony in complex observations, high parametric/statistical uncertainty in the DISS error model, the distance of the DSDA measurement center from the PCA and due to errors in coordinate transformations. The works [1,5,6, etc.] consider the capabilities of the radar itself as a sensor of (micro)navigation information in special operating modes, similar to DSDA. With a radiation pattern with one main beam, such modes are applicable only at those stages of flight when the radar is not used to solve its main tasks - viewing the ground and airspace. Currently, methods of radar measurements and their applications in the aspect of micronavigation "against the background" of solving problems of viewing the earth's surface – at a low pulse repetition rate – are being actively developed. For example, works [5,6] show methods for measuring speed towards the center of a mapped area of the earth's surface with a potential accuracy of up to 0.001 m/s and using such measurements for additional correction of the MNS.

In the MFR, a number of air-to-air tasks are solved when detecting / "capturing" the trajectory / tracking a target (air object) against the background of the ground, i.e. MB directed downwards. There is a known [7] method for measuring the speed and range of a PCA to the ground in the GL direction, implemented with several bursts of pulses with quasi-continuous radiation and with chirp modulation of the carrier wave over a burst of pulses. But in many practical cases, such problems are solved with typical quasi-continuous radiation, without the use of chirps. The specified method [7], therefore, cannot be applied. The purpose of the work is to present: • a method of combined (micro)navigation radar measurements of the FCA speed/range to the ground in the GL direction, performed simultaneously with a review of the airspace "against the background of the earth" with typical quasi-continuous radiation; • schemes, basic procedures for complex processing of inertial radar (micro) navigation information using these measurements and MNS inertial channel data. *The ultimate pragmatic goal* is to expand the information content/functionality of the MFR due to accompanying additional measurements of the PCA (micro)navigation parameters; increasing the stability and accuracy of MFR micronavigation in conditions of radio interference.

II. FORMULATION OF THE PROBLEM

The MNS, starting from the initial alignment mode and further in the inertial satellite (micro)navigation mode, functions continuously, regardless of the state and operating modes of the radar. At the same time, an interface is provided for the interaction of the MNS with the standard navigation complex and with the radar (including modules for controlling the direction and parameters of radiation, reception, primary processing and time-frequency changes of signals). During the operation of the MNS in flight, long intervals are possible when the SNS channels of the MNS and the standard navigation complex are inoperative or produce unreliable data due to radio interference. As an example, Fig. 1 shows the number of satellites recorded by the SNS receiver in the MNS during a real flight: with signs of reliability of position-velocity data (green); in their absence (red). The work examines the "red" sections of the flight, starting from the moment the unreliability of SNS data is revealed. Options are possible when: the MFR is not involved in the main functions, it can be used before/after this moment specifically in the interests of micronavigation under various radiation, including quasi-continuous radiation (hereinafter referred to as standard radiation); at the moment

of "SNS failure", the MFR performs a line-by-line survey of the airspace with a narrow main beam in the standard radiation mode at wavelength λ in the X-band or implements target observation/tracking procedures. The work considers radiation with a high-average pulse repetition frequency (PRF) Fr, i.e. in the range of 60-100 kHz [1] and at lower PRFs, up to the upper limit of the range for average PRFs of 10-40 kHz [1].

Fig. 1.   Number of satellites during the operation of the MNS with signs (green)/without signs (red) of information reliability

In the first and second variants (when the radar operation cyclogram includes bursts of pulses specially allocated for micronavigation measurements), it is possible to control the orientation of the MB and the parameters of emission/reception/processing of signals. In this case, the angle of inclination of the MB exceeds half of its width in the vertical plane, and the area of the earth's surface irradiated by the MB ("trace" of the MB) is limited.

In a radar, when viewing airspace "against the background of the earth", a similar orientation of the MB and radiation parameters are set based on the logic/cyclogram of solving the main problem. The ability to control them to improve the accuracy/stability of related micronavigation measurements is no longer available.

In general, two situations are considered: • a procedure for measuring/estimating the specified trajectory parameters of the PCA, "built-in" into the cyclogram of the main task, with "own" control of the parameters of the orientation of the MB/emission/reception/processing of the signal, independent of the main task of the radar; • a procedure for measuring/estimating these parameters that is completely combined with the main task, performed simultaneously with the task, but without the above "independent" control. It is also necessary to evaluate the accuracy of radar micronavigation measurements, the feasibility of their use for additional correction of the MNS in a complex radio interference environment, and to determine schemes and basic correction procedures based on complex processing of inertial and radar measurements.

In general, the formulation of the problem comes down to the development of methodological support for the implementation in the radar of an additional channel of micronavigation measurements and "in-depth" integration of the radar and the MNS.

III. THEORETICAL AND EXPERIMENTAL JUSTIFICATION AND ESSENCE OF THE METHOD OF COMBINED MICRONAVIGATION MEASUREMENTS

In the general case, powerful reflections from the ground along the side lobes and the MB appear in all discrete samples $d\tau_i$ of the ambiguous signal delay (ASD) in the "working" zone of the pulse repetition period (PRP) $T_r=1/F_r$. A signal from a "point" air object (target) at a distance R with an ASD τ will be observed only in one discrete ASD. As an example, Fig. 2 shows the results of spectral processing of a real signal from the ground and the target - in the form of a superposition of Doppler spectra obtained over all $d\tau_i$ ($i=1, ..., n$) in the "working" zone of the PRP, where $d\tau_i \approx \tau_{pulse}/3$, τ_{pulse} is the pulse duration. In this case, at the half-power level, the width of the MB in the vertical plane is $\approx 8^0$, the azimuthal width is $\approx 2^0$, and the angle of inclination of the MB is $\approx 6^0$.

Fig. 2.   Superposition of Doppler spectra of signals from the ground and the target, received in all discrete $d\tau_i$ of the "working" zone of the PRP.

Within the boundaries of the MB, based on reflections from its trace on the ground in each $d\tau_i$ ($i=1, ..., n$), a superposition of signal spectra from several ($j>1$) bands on the surface is carried out, which have different areas dA_j and correspond to different distances from the PCA to the center of band and these distances are multiples of value (c $T_r/2$), where c is the speed of radio waves.

Ratio of signal C from one section (band) with area dA at distance R to noise N [1]:

$$C / N = \frac{P_{AVE} G_T G_R \lambda^2 \sigma^o dA}{(4\pi)^3 R^4 L_C k T_S B_n}, \tag{1}$$

where P_{AVE} is the average transmitter power; G_T/G_R - gain of the transmitting and receiving antenna in the direction of the site; σ^0 - backscatter coefficient; L_C - collateral losses in radar; k – Boltzmann's constant; T_S - system noise temperature; $B_n = \Delta f$ – Doppler filter width. The area of such a plot:

$$dA = \frac{c\tau_{pulse}}{2\cos\alpha} R\theta_\alpha, \tag{2}$$

where α is sliding angle; θ_α - MB width in azimuth at half power level. In the general case, in each dτi the ratio (C/N) will be determined as the integral for equation (1) over the Doppler frequency and range for each position of the section dA_j within the MB.

In general, theoretical aspects and a number of experiments have shown: • the maxima in the spectra of all $d\tau_i$ ($i=1, ..., n$) in the "working" zone of the PRP vary in a narrow frequency band (see Fig. 2) even with a relatively wide MB in elevation; • among these maxima, the presence of an "explicit" global amplitude maximum at a certain $d\tau_k$,

corresponding to the distance to the ground in the direction of the axis (direction of maximum) of MB, is not ensured. For example, according to the results of processing, a real signal from the ground in the direction of the MB in Fig. 3 shows the maxima in the spectra for all $d\tau_i$ (i=1, ..., n at n=13) in the "working" zone of the PRP. In this case: • they appear only in two adjacent filters; in most cases (>60%), the serial numbers of these filters are equal to the calculated (reference) serial number obtained on the basis of velocity estimates in the MNS in the inertial-satellite mode and when they are recalculated to the direction of the MB axis.

Fig. 3. Maxima in the spectra for $d\tau_i$ depending on their serial number i (number of discrete distance) in the PRP working area

An unambiguous estimate of the Doppler frequency (DF) $F_{D(i)}$ by the filter number with a maximum in the signal spectrum for element $d\tau_i$ (and speed $V_{R(i)} = 0.5\ \lambda\ F_{D(i)}$) in the general case is not related to a specific distance to the ground/inclination angle within the framework of the MB, since it is the result of a superposition of the signal spectra from several spaced sections (for j>1). The same is true for estimating frequency/speed based on the filter number with a global maximum among all spectra observed in the PRP working area. In such a situation, direct measurement of the distance to the ground along the MB axis is impossible and it is advisable to obtain a nonparametric statistical estimate of the frequency F_D^* from a sample $\{F_{D(i)}, i=1, ..., n\}$, for example, in form of a median, average, α-truncated average and etc. When estimating F_D^* in the form of an average frequency value $F_D^*=$ ave $\{F_{D(i)}, i=1, …, n\}$, the theoretical standard deviation (SD) σ_{fdave} for estimation errors:

$$\sigma_{fdave} = \frac{1}{\sqrt{n}}\sigma_{fd},\qquad(3)$$

where $\sigma_{fd}=\sigma_{fd(i)}\approx0{,}3\Delta f$ - SD for instrumental frequency measurement errors by filter number with width $\Delta f= F_r/N_{FFT}$; N_{FFT} - number of samples (base) of the fast Fourier transform.

At different angles of inclination of the MB, a series of experiments were performed with several bursts of pulses for each angle and with the determination of velocity values for each burst based on: a) the average frequency F_{Dave}; b) F_{DGM} frequency, determined by the number of the filter with a global maximum amplitude within the PRP. At the "reference" speed (according to SNS data), the errors in determining the speed based on F_{Dave} and F_{DGM}, their average values and standard deviation were obtained from a series of bursts of 2048 pulses at each angular position of the MB. The results of these experiments are shown in Fig. 4. In this case, the average value of the error in speed measured based

on F_{Dave} is $\approx$ in the range from 0.1 m/s to 0.23 m/s, and the standard deviation is $\approx$ in range from 0.2 m/s to 0.35 m/s.

Fig. 4. Average values and ±3SD intervals for speed determination errors based on the average frequency and global maximum

It is obvious that a radar corrector in this design is several times inferior to the SNS in accuracy, but, as studies have shown, in conditions of radio interference it ensures stable operation of the MNS even in the long-term absence of reliable SNS data. Its potential accuracy can be increased by increasing the PRP the number of bursts of pulses emitted in one direction of the MB. The work proposes an unconventional approach to increasing the accuracy and expanding the information content of the radar channel of micronavigation measurements - based on adaptive-robust control of the direction of the MB, parameters during radiation, processing and measurement of signals during the interaction of the MNS and the radar. The essence of the proposed method is that the current data of the MNS with known parameters of the radar radiation pattern make it possible to determine: • such a combination of the orientation parameters of the MB axis and the PRP, in which there will be a global maximum in the signal spectrum for the $d\tau_k$ band corresponding to the distance to the ground along the MB axis, and $d\tau_k$ is guaranteed to be in the working area of the PRP; • boundaries for notch filtering of signals from the ground along the side lobes of the radar.

Depending on the height H, orientation, width of the MB and radiation parameters, the size of the MB trace can be such that the superposition of spectra in $d\tau_i$ (i=1, ..., n) is absent or does not affect the manifestation of the global maximum in $d\tau_k$. Distances R_0 along the MB axis, along the upper R_{max} and lower R_{min} boundaries of the MB, as well as their difference $\Delta R = (R_{max}-R_{min})$ and the length L of the irradiated area:

$$R_0 = \frac{c}{2}(r_0 T_r + \tau_0) = \frac{H}{\sin\beta_0};\qquad(5)$$

$$R_{max} = \frac{c}{2}(r_{max}T_r + \tau_{max}) = \frac{H}{\sin\beta_{min}} = \frac{H}{\sin(\beta-\Delta\beta)};\ (6)$$

$$R_{min} = \frac{c}{2}(r_{min}T_r + \tau_{min}) = \frac{H}{\sin\beta_{max}} = \frac{H}{\sin(\beta+\Delta\beta)};\ (7)$$

$$\Delta R = \frac{c}{2}[T_r(r_{max} - r_{min}) + (\tau_{max} - \tau_{min})] =$$
$$=H[\cos ec(\beta-\Delta\beta) - \cos ec(\beta+\Delta\beta)];\qquad(8)$$

$$L = H(\operatorname{ctg}\beta_{\min} - \operatorname{ctg}\beta_{\max}) =$$
$$= \frac{c}{2}\left[(r_{\max}\cos\beta_{\min} - r_{\min}\cos\beta_{\max})T_r + \tau_{\max}\cos\beta_{\min} - \tau_{\min}\cos\beta_{\max n}\right]$$

where r and τ are the PRP and ASD index; $\beta > 0$ and $\Delta\beta > 0$ – the angle of inclination and half the width of the MB at the elevation angle. The values of the ASD are needed at $r_0 = r_{\max} = r_{\min} \rightarrow (\tau_{\min} = 0; \tau_{\max} = T_r)$ or at $r_{\max} = r_0 + 1$, $r_{\min} = r_0 - 1 \rightarrow (\tau_{\min} = T_r - t_{bzr}; \tau_{\max} = \tau_{pulst} + t_{bzl})$, where t_{bzr}/t_{bzl} – receiver blanking time before/ after radiation. Then from (8) it follows:

$$\Delta R \le \frac{c}{2}T_r \tag{9}$$

or $\Delta R \le \frac{c}{2}(T_r + \tau_{pulse} + t_{bzr} + t_{bzl})$. (10)

It is necessary to guarantee the position τ_0 in the working area of the PRP

$$\tau_0 \in [(\tau_{pulce} + t_{bzl}); (T_r - t_{bsr})], \tag{11}$$

taking into account the accuracy of MNS data. Error δR_0 distance (5) with a known PRP index:

$$\delta R_0 = \frac{c}{2}\delta\tau_0 = R_0\operatorname{ctg}\beta_0\delta\beta_0 + \frac{\delta H}{\sin\beta_0}, \tag{12}$$

$$\delta\tau_0 = \frac{2}{c}[R_0\operatorname{ctg}\beta_0\delta\beta_0 + \frac{\delta H}{\sin\beta_0}] \approx 3\sigma_{\tau 0}, \tag{13}$$

where $\delta\tau_0$ is the ASD error; $\delta\beta_0$, δH – errors in the angle of inclination and height; $\sigma_{\tau 0}$ is the standard deviation of the ASD measurement errors. Then condition (11) is represented as

$$[\tau_0 \pm k(\frac{2}{3c}[R_0\operatorname{ctg}\beta_0\delta\beta_0 + \frac{\delta H}{\sin\beta_0}])] \in [(\tau_{pulse} + t_{bzl}); (T_r - t_{bzr})], \tag{14}$$

where k is the coefficient defining the confidence interval.

Thus, it is possible to determine the number of the filter f_{Amax} with the global maximum amplitude and the number $k_{\tau Amax}$ of the ASD band $d\tau$ in the PRP where this maximum was identified. Accordingly, direct primary measurements of the ASD τ, frequency F_D and the vector of measurements of the micronavigation parameters of the PCA are formed: the distance R to the ground in the direction of the land area giving the maximum echo signal within the MB; speed F_D in this direction. The same parameters are determined on the basis of inertial measurements in the MNS. The observation vector – as the difference between inertial and radar measurements – is processed based on Kalman filtering procedures to estimate/compensate for inertial channel errors.

IV. CONCLUSION

A method for (micro)navigation measurements of the trajectory of the phase center of a radar antenna with quasi-continuous radiation based on signals from the ground in the MB is proposed. The method is compatible with the review of airspace "against the background of the earth"; it can be implemented simultaneously or with an additional burst of pulses "built into" the cyclogram during the review. In the first case, there is no control of the MB and radiation parameters in the interests of micronavigation. This ensures

speed measurement with relatively low accuracy. When a number of criteria are met, speed is measured with increased accuracy, as well as the distance to the area of the earth's surface with the most intense signal. In the second case, the fulfillment of these criteria and high-precision measurements of distance and speed are ensured by adaptive-robust control of the MB orientation and/or radiation parameters.

REFERENCES

[1] Spravochnik po radiolokatsii (Handbook of Radiolocation), Ed. M. I. Skolnik. Translated from English. In 2 books. Book 2. Moscow: Tekhnosfera, 2015. – 680 s.

[2] A.V. Chernodarov, A.P. Patrikeev, V.N. Kovregin, G.M. Kovregina, I.I. Merkulova, Flight development of a distributed inertial satellite micronavigathion system for synthetic-aperture radar, Nauchniy vestnik MGTU GA, 2017, no. 1 pp. 222-231.

[3] V. Kovregin, G. Kovregina, Adaptive-robust methods for detecting, capturing and tracking hovering, low- and high-speed objects in integrated radar-inertial systems with quasi-continuous radiation. URL: https://ieeexplore.ieee.org/document/9470845 (Date of access: 17.10.2021).

[4] A.V. Chernodarov, V.N. Kovregin, G. M. Kovregina. Inertial-Satellite Micro-Navigation System for Radar with Synthetic Aperture and Doppler Channel for Measuring Trajectory Velocity. 2023 30th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS) pp. 1-3.

[5] V.U. Sevastianov, M.V. Stepenov. Izmerenie skorosti dvizheniya nositelya v monoimpul'skom radiolokatore s sintezirovannoj aperturoj antenny, Radiotekhnika i elektronika, Moscow, Radiotekhnika i elektronika, 2007, T. 52, no 8, pp.944-948 (in Russian).

[6] M Krasil'shchikov, M.N., Kozorez, D.A., Sypalo, K.I. et al. High accuracy positioning of phase center of multifunction airborne radar antenna. Gyroscopy Navig. 4, 164–173 (2013).
https://doi.org/10.1134/S2075108713030073.

[7] V. Kovregin, G. Kovregina, Method for observing a hovering object against the background of the earth in radars with quasi-continuous chirp radiation. 30th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), Saint Petersburg, Russian Federation, 2023, pp. 1-4, doi: 10.23919/ICINS51816.2023.10168478.

Methods and Algorithms for Active Adaptive-Robust Tracking of an Aerial Object and Accompanying (Micro)Navigation Measurements in Quasi-Continuous Wave Radars

V. Kovregin
St. Petersburg State University of Aerospace Instrumentation
St.-Petersburg, Russia
kovregin@mail.ru

G. Kovregina
St. Petersburg State University of Aerospace Instrumentation
St.-Petersburg, Russia
g_kovregina@mail.ru

Abstract – **Methods and algorithmic structures are presented for simultaneous solution of problems: active tracking of an air target based on adaptive-robust procedures for controlling quasi-continuous radiation and Kalman filtering of ambiguous measurements; measuring micronavigation parameters of the radar phase center using echo signals from the ground received along the side lobes of the radar radiation pattern.**

Key words – quasi-continuous radar radiation; target tracking, interfering reflections from the ground, micronavigation, radio interference

I. INTRODUCTION

Currently, in «air-to-air» missions solved by multifunctional radars, the urgent problem is to increase the stability and accuracy of target tracking in a complex signal-noise/interference environment, when maneuvering the target and/or carrier and long intervals between radio contacts. In such conditions, as a result of the initial observation – detection/identification/measurement of target trajectory parameters – there is a high probability of "anomalous" range/relative speed measurements, which include: error in the pulse repetition period (PRP) index when estimating the range; error when estimating speed not from the "fuselage", but from the "rotational" component of the "complex" Doppler spectrum of the target. Traditional tracking methods [1] are not robust in relation to anomalous initial data obtained after the observation stage. Typically, with PRP $T_r=const$, only the "drift" of the ambiguous delay for the target signal (ASD) due to the target's movement is monitored, the PRP changes only occasionally to ensure observability of the ASD τ in the "working" zone of the repetition period, and the error according to the PRP index is preserved. Anomalies ("outliers") are also possible in ASD measurements, leading to divergence of standard error filtering procedures, for example, the Kalman filter. Such destabilizing factors lead to a sharp decrease in accuracy and, ultimately, to a breakdown in tracking. There is a practical need to develop new methods/algorithms for target tracking that are effective in the event of such anomalies in the source data and current measurements. In addition, the problematic situation during tracking is significantly aggravated by the fact that when there is organized interference with satellite navigation systems, the accuracy/operability of typical inertial satellite micronavigation systems in radars (MNS) is quickly lost. Currently, *the problem* of providing an alternative (additional) noise-resistant correction of MNS

with a sufficient level of accuracy is relevant. As part of its solution, a number of works [1-5, etc.] consider the use of the radar itself as a corrector for MNS.

For example, when tracking a target by range: with a constant PRP, only the "drift" of the ambiguous delay of the target signal (ASD), caused by the relative movement of the target, is monitored, and the error of the PRP index is preserved. In ASD measurements, "outliers" are also possible, leading to practical divergence of traditional procedures for statistical filtering of ASD errors. These destabilizing factors lead to a significant decrease in accuracy and, ultimately, to a "breakdown" of tracking. In general, all known methods of (micro)navigation radar measurements are based on the use of echo signals from a section of the earth's surface irradiated by the main beam (MB) of the radar pattern (RP). When tracking a target in the "upper hemisphere," these methods cannot be used simultaneously with the tracking step, since echo signals from the ground are received only via side lobes (SL), and their information content for micronavigation has not yet been sufficiently studied.

The purpose of the work is to present unconventional methods/algorithms: • active adaptive-robust (AR) target tracking in quasi-continuous radiation (QCR) mode; • radar (micro)navigation measurements based on signals from areas of the earth irradiated by the SL while tracking a target in the "upper hemisphere". The ultimate pragmatic goal is to increase noise immunity, accuracy, expand the information content of the radar in the tasks of tracking a MB target in the "upper hemisphere" and (micro)navigation support for the radar in conditions of radio interference

II. ESSENCE OF THE METHOD OF ACTIVE ADAPTIVE-ROBUT TRACKING AN AIR TARGET

In air-to-air missions, the logic of radar operation usually includes an observation stage followed by a transition to an auto-tracking mode or episodic tracking of a target "on the route." The works [2, 6, 7, etc.] present a number of methods for adaptive-robust observation with QCR and QCR with chirp of a carrier wave in a burst of pulses (QCR -chirp) - in a single- and multi-target situation in the MB, for targets with a narrow spectrum and/or dispersed ("complex"), with limited or extended ranges of relative speeds. To describe the essence of the active AR tracking method proposed in this work, it is sufficient to consider a single-target situation in

the MB, with a typical QCR for an approaching target with a narrow Doppler spectrum located outside the spectral zone of "interfering" reflections from the ground (IRG). All the basic methodological provisions and analytical dependencies presented below with such a "simplified" formulation of the problem, however, can be generalized to solve more complex problems of AR tracking - for a variety of different-speed targets with different types of spectrum. In the QCR-chirp mode, the essence and mathematical content of the proposed tracking method does not change.

In AR observation [2,6-8], the tasks of detecting and "capturing" a target at a range with QCR are reduced to a single formalized formulation and are solved on the basis of unified methods within the framework of the concept of AR control of radiation parameters and evaluation of ambiguous measurements. The statement and approach to solving the tracking problem are similar. In general, the only differences are: • in the levels of a priori uncertainty in the distance to the target (according to the PRP indices) in the source data; • in the different "genesis" of the used estimates of the ASD τ, formed on the basis of quasi-measurements and/or real measurements of the ASD.

For example, if the target is not detected at the PRP value T_r, then the initial range of PPI indices [r_{min}; r_{max}] for the current distance $R = 0.5\ c\ (r\ T_r + \tau)$, where c is the radio wave speed, r is the PRP index, will be determined from an a priori wide range - from the minimum R_{min} to the maximum R_{max} distance of target detection. In this case, the assessment of the ASD is based on quasi-measurement, assuming that the ASD of the target is located in the "blind" zone of the PRP. After the disclosure of "blind distances" with uncertainty reduction and a "rough" estimate of the distance to the target in the initial data for "capturing" the trajectory, there will already be a significantly reduced range of PRP indices and one real measurement of the ASD. After capture, the source data for tracking also contains a real measurement of the ASD, but the range of PRP indices degenerates - the PRP index r is theoretically considered to be precisely defined. In the general case, AR control of the QCR involves maximizing the module of the increment of the PRP, at which the ASD τ in the current PRP T_r will be guaranteed to be "observed" in the "working" zone of PRP [t_{BZ1}; $T_r - t_{BZ2}$], where t_{BZ1} and t_{BZ2} are the blanking time of the receiver after and before the start of radiation. The PRP index is considered to be precisely determined when, as a result of several iterations, an increment of PRP is achieved that exceeds the maximum total error for the difference between two measurements of the ASD and the prediction of the ASD drift in the interval between the emission of bursts of pulses. The relationship for predicting the position of the ASD τ^* in the PRP at a given increment $\Delta T_{r\ i/i-1} = T_{r\ i} - T_{r/i-1} = \Delta T_{r\ SP\ i/i-1}$:

$$\tau^* = \tau_{i/i-1} - \Delta r_{i/i-1} T_{r\ i-1} - r_{i/i-1} \Delta T_{ri/i-1}, \qquad (1)$$

where $\tau_{i/i-1}$, $r_{i/i-1}$ are the extrapolated values of the ASD and the PRP index at a constant PRP, taking into account the drift of the ASD due to the movement of the target and the radar in the interval between the emission of pulse bursts; $\Delta r_{i/i-1} = r_i - r_{i/i-1}$ - change in the PRP index due to the increment in PRP.

Calculation of the PRP increment at a given $\tau^* = \tau^*_{SP}$:

$$\Delta T_{ri/i-1} = \frac{\tau_{i/i-1} - \tau^* - \Delta r_{i/i-1} T_{r\ i-1}}{r_{i/i-1}}. \qquad (2)$$

In case of anomalous completion of the "capture", the obtained estimates for the PPI indices r_{i-1} and $r_{i/i-1}$ are characterized by an integer error $|\delta r| \geq 1$.

Eliminating such errors during active AR tracking involves:

• specified range [$r_{min\ i/i-1} = r_{i-1} - |\delta r|$; $r_{max\ i/i-1} = r_{i/i-1} - |\delta r||$], which determines a priori the level of robustness of the target AR tracking to anomalous errors according to the PRP index. As model/natural experiments have shown, errors in the PKP index during AR capture in the worst case are $\delta r = \pm 1$;

• setting the required position of the ASD $\tau^* = \tau^*_{CWZ\ i}$ in the center of the working area in the PRP $T_{r\ i}$ at $\tau_{i/i-1}$, $r_{i/i-1}$;

• search based on expressions (1), (2) of the maximum modulo PRP increment:

$$\Delta T_{ri/i-1} = \max\{|\Delta T_{ri/i-1}(NdT_r, M\Delta r_{i/i-1})|\},$$
$$N = 1, 2, \ldots; M = 0, 1 \ldots, \qquad (3)$$

where dT_r is the discreteness of the PRP implementation. Moreover, all values from the series of ASD forecasts of the form (1)

$$\tau^*(J) = \tau_{i/i-1} - \Delta r_{i/i-1} T_{r\ i-1} - J\Delta T_{ri/i-1}; J = r_{min\ i/i-1}, \ldots, r_{i/i-1}, \ldots, r_{max\ i/i-1}$$

must be in the working area of one PRP. The interval $\Delta\tau^*$ between "neighboring" forecast estimates of the ASD will be equal to the increment of the PRP (modulo):

$$\Delta\tau^* = |\tau^*(J+1) - \tau^*(J)| = |\Delta T_{ri/i-1}|. \qquad (4)$$

• determination minimum error in forecasting the NPC relative to the current measurement of the NPC τ_i and the current PPI index r_i corresponding to this minimum:

$$|\nu_i|, r_i \leftarrow \min\{|\tau_i - \tau^*(J)|\}. \qquad (5)$$

• calculation of variances of forecast errors, measurement of the ASD and deviations ν_i, estimation of the ASD based on the AR Kalman filtering procedure, for example, using statistical criteria χ^2 - distribution when analyzing the quadratic form of the normalized residual [3], determination of the current distance based on the obtained estimates of the PRP index and ASD.

Thus, active AR tracking of a target in range carries out a distributed simultaneous forecast of the ASD for several trajectories at different PRP indices, while ensuring the observability of all ASD within the working area of one PRP. The correspondence between the current measurement and the ASD forecast is performed according to the criterion of the minimum error of the ASD forecast model. Based on the selected predictive model and the measurement of the ASD, the Kalman-type AR filtering step is implemented. As a result, the stability of the tracking is ensured in the presence of anomalous errors in the initial data and measurements, and the accuracy of range estimation is increased to a level exceeding the instrumental accuracy of the ASD measurement.

Figure 1 shows the characteristic results of modeling active AR tracking after target detection/acquisition with errors in the PRP index up to ± 2 (i.e., up to several km), confirming the effectiveness of the proposed method.

Fig. 1. Errors of active AR tracking in range in the presence of anomalous estimates at the stage of detecting/capturing the target trajectory

Fig. 2. Dependence of amplitude on angles in the azimuthal/vertical planes

Fig. 3. View of the antenna spatial pattern

III. ESSENCE OF THE METHOD OF (MICRO)NAVIGATIONAL MEASUREMENTS COMBINED WITH TARGET TRACKING

When observing/tracking a MB target in the "upper hemisphere", the radar receives signals from areas of the ground irradiated only in SL directions. These signals, traditionally designated as IRG by SL, with QCR and QCR - chirp are "powerful" even at high radar altitudes and are excluded from processing by notch filtering [2, 6, 7]. Active AR tracking of a target in the SOI-chirp mode is considered. In works [2, 6, 7, etc.] for QCR -chirp, a technology is shown for estimating the position/width of the spectrum of the IRG from SLs directed downwards - using data from the MNS and the parameters of the orientation/angular width of the SL for a calibrated antenna radiation pattern (by level SL and directions of their maxima). After calculating the boundaries of the spectrum of IRG signals from the SL, their rejection was performed. The proposed method of combined micronavigation measurements in a radar, on the contrary, is based on a "fine" spectral analysis of precisely the IRG signals from the SL. Their processing is carried out in parallel with the processing of signals received via the MB when searching/observing/tracking a target. The logic of solving these main problems of the radar is subject to AR control of the direction of the MB and the parameters of the QCR-chirp - PRP T_{ri}, , the slope of the chirp S_i, the duration τ_{pulse} and the number of N_{FFT} pulses in a burst of pulses (the basis of the fast Fourier transform (FFT)). To describe the essence of the proposed method, a typical pattern of a phased array antenna is considered. In Fig. 2 shows the distribution of the MB and SL amplitudes depending on the angles in the azimuthal and elevation planes with the MB axis directed perpendicular to the antenna surface. In Fig. 3 shows a view of the spatial radiation pattern of the radar antenna. To describe the essence of the proposed method, it is quite sufficient to consider the situation when IRG are formed only by SLs with downward-directed axes located in the same vertical plane. The width of the side lobes in this plane is a few degrees, and the azimuthal width is several times smaller. For such "narrow and flat" SLs, the influence of the azimuthal width is insignificant and is not taken into account in the relations presented below.

In the normal terrestrial coordinate system $OXYZ$, the phase center of the radar at height H moves with a constant horizontal speed W in the direction of the OX axis. The azimuthal angle α of the axes of all considered SLs is the same. The areas of the earth's surface irradiated by the SL ("traces" of the SL) do not intersect, their centers are located on the same straight line. The surface of the earth is assumed to be flat and homogeneous (in terms of reflection). With known parameters of the QCR-chirp and the orientation of the SL for the echo signal from each such section, using the MNS data, the following can be calculated [2, 6, 7]: "central" values of the Doppler frequency (DF) $F_{D(0)}$ and rangefinder-Doppler frequency (RDF) $F_{RD(0)}$, corresponding to the direction of the SL axis (direction of the maximum); boundaries of the DF and RDF spectra within a given SL. For QCR-chirp mode:

• the spectrum of powerful IRGs manifests itself in all elements of the ASD resolution with a width $\delta\tau$ in the working area of the PRP, being the result of a superposition of RDF spectra for signals from several strips of the earth's surface spaced apart from each other and which intersect the "trace" of the SL;

• in the distance range $[R_{min}; R_{max}]$ to the near and far boundaries of the "trace" of the SL for each j-th strip of the RDF value $F_{RD(j)}$ and the difference $\Delta F_{RD(k, j)} = F_{RD(k)} - F_{RD(j)}$ for the k-th and j-th strips are defined by the expressions:

$$F_{RD(j)} = \frac{2}{\lambda} W \cos\alpha \cos\beta_{(j)} + ST_{ri}r_{i(j)} + S\tau; \quad (6)$$

$$\Delta F_{RD(k,j)} = \frac{2W\cos\alpha}{\lambda}\left[\cos\beta_{(k)} - \cos\beta_{(j)}\right] + ST_{ri}\left[r_{i(k)} - r_{i(j)}\right] =$$

$$= \frac{2W\cos\alpha}{\lambda}\left[2\sin\frac{\beta_{(k)}+\beta_{(j)}}{2}\sin\frac{\beta_{(j)}-\beta_{(k)}}{2}\right] + ST_{ri}\left[r_{i(k)}-r_{i(j)}\right], \quad (7)$$

where λ – wavelength; τ – ASD for the "middle" line of the band under consideration, and $\tau = \tau_{(k)} = \tau_{(j)}$; $\beta_{(k)}$ and $\beta_{(j)}$ – angles of inclination to the k-th and j-th stripes located at distances $R_{(k)} = 0.5\ c\ (r_{i(k)}T_{ri} + \tau)$ and $R_{(j)} = 0.5\ c\ (r_{i(j)}T_{ri} + \tau)$.

From (6) and (7) it is clear that in each discrete element of the ASD, due to the overlap of the RDF spectra of signals from different strips, in the resulting RDF spectrum local amplitude maxima will be observed, significantly separated in frequency (due to the difference in distances and directions). For example, even without taking into account the speed component in expression (7), with typical values of PRP $T_{ri} = 3*10^{-5}$ s and chirp slope $S_i = 20*10^6$ Hz/s, the frequency interval between "neighboring" the maximums ($\Delta F_{RD\ (k,\ k-1)} = F_{RD(k)} - F_{RD(k-1)}$) will be at least 600 Hz - at least 70 filters with a width of ≈ 8 Hz for the FFT base 4096. This interval significantly exceeds the width of the RDF spectrum of the echo signal from each distance band on the surface within the SL [1]. This makes it possible to quite simply identify local and global amplitude maxima when analyzing the resulting RDF spectrum in each discrete element of the ASD observed in the PRP. The number of the ASD element Nd is determined, in which the largest amplitude maximum appeared and the corresponding value of the RDF $F_{RD\max\ (0,\ Nd)}$ of the form (6), where: ASD $\tau = \tau_{(0)}$ is estimated from the measured Nd and its position in the PRP; PPI index $r_{i(j)} = r_{i(0)}$ is determined using data from MNS:

$$r_{i(j)} = r_{i(0)} = \mathrm{int}\left[\frac{2}{c}\frac{H}{\sin\beta_{(0)}T_{ri}}\right], \quad (8)$$

where $\beta_{(0)}$ is the calculated angle of inclination of the SL axis.

By measuring the ASD and assessing the PRP index of type (8), the distance $R_{(0)}$ is determined, the inclination angle is adjusted - towards the center of the "working" strip on the earth's surface, corresponding to Nd:

$$R_{(0)} = \frac{c}{2}\left[T_{ri}r_{i(0)} + \tau_{(0)}\right], \quad (8)\quad \beta_{(0)} = \arcsin\frac{H}{R_{(0)}}. \quad (9)$$

Based on the measurements of the RDF and ASD, the radial speed $V_{R(0)}$ of the radar in a given direction within the SL is determined:

$$V_{R(0)} = W\cos\alpha\cos\beta_{(0)} = \frac{\lambda}{2}\left[F_{RD\max(0,Nd)} - ST_{ri}r_{i(0)} - S\tau\right]. \quad (10)$$

In addition, from (7) follows a difference method for measuring the projection V_α of the ground speed W of the phase center of the radar onto the known azimuthal direction of the SL axis and, in fact, estimating the value of the ground speed itself in the form $W = (V_\alpha / \cos\alpha)$:

$$V_\alpha = W\cos\alpha = \frac{\lambda}{2}\left[\frac{\Delta F_{RD(k,j)} - ST_{ri}\left[r_{i(k)} - r_{i(j)}\right]}{\cos\beta_{(k)} - \cos\beta_{(j)}}\right] =$$

$$= \frac{\lambda}{4}\left[\frac{\Delta F_{RD(k,j)} - ST_{ri}\left[r_{i(k)} - r_{i(j)}\right]}{\sin 0.5(\beta_{(k)}+\beta_{(j)})\sin 0.5(\beta_{(j)}-\beta_{(k)})}\right]. \quad (11)$$

IV. CONCLUSION

The proposed method of active AR tracking provides resistance to anomalous measurements/estimates of range parameters in the initial data (due to errors in trajectory capture), and during the tracking process (due to extrapolation errors during rare "polls" of the target, delays in the interface, and computational failures, anomalous primary measurements of the signal). Along with noise immunity, the accuracy of estimating target trajectory parameters based on AR Kalman filtering of ambiguous primary measurements increases. Against the background of the implementation of this method when tracking a target in the "upper hemisphere", the possibility of simultaneously performing measurements of (micro)navigation parameters of the radar phase center at each tracking step is shown. An unconventional method is presented for carrying out such accompanying measurements of the range/speed of a radar relative to the ground - based on clutter signals from the earth's surface received by the radar along the side lobes of the radiation pattern. This method provides an expansion of the information content of the radar, primarily for using the radar as a noise-resistant and sufficiently accurate correction channel for a typical micronavigation system in the event of radio interference.

REFERENCES

[1] Spravochnik po radiolokatsii / Pod red. M. I. Skolnika. Per. s angl. Pod obshchey red. V. S. Verby. V 2 knigakh. Kniga 2. Moskva: Tekhnosfera, 2015. – 680 s (in Russian).

[2] Kovregin V.N., Kovregina G.M. Method for observing a hovering object against the background of the earth in radars with quasi-continuous chirp radiation. 2023 30th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), pp. 1-4, doi: 10.23919/ICINS51816.2023.10168478.

[3] A.V. Chernodarov, V.N. Kovregin, G. M. Kovregina. Inertial-Satellite Micro-Navigation System for Radar with Synthetic Aperture and Doppler Channel for Measuring Trajectory Velocity. 2023 30th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS) pp. 1-3.

[4] V.U. Sevastianov, M.V. Stepenov. Izmerenie skorosti dvizheniya nositelya v monoimpul'skom radiolokatore s sintezirovannoj aperturoj antenny, Radiotekhnika i elektronika, Moscow, Radiotekhnika i elektronika, 2007, T. 52, no 8, pp.944-948 (in Russian).

[5] Krasil'shchikov, M.N., Kozorez, D.A., Sypalo, K.I. et al. High accuracy positioning of phase center of multifunction airborne radar antenna. Gyroscopy Navig. 4, 164–173 (2013). https://doi.org/10.1134/S2075108713030073.

[6] Kovregin V.N., Kovregina G.M., Murzaev A.S. A Unified Method for Observation of an Air Object with a Complex Spectrum in Radar with Quasi-Continuous Radiation. 2022, 29th Saint Petersburg International Conference on Integrated Navigation Systems (ISINS), pp. 1-4.

[7] V.N. Kovregin, G.M.Kovregina, A.S. Murzaev. Adaptivno-robastnoe vserakursnoe nablyudenie raznotipnyh ob"ektov v glavnom luche radara s kvazinepreryvnym LCHM-izlucheniem, Radiotekhnika, Moscow, Radiotekhnika, 2023. T. 87, no. 1, pp. 69−78 (in Russian).

Modeling of Movement-Related Evoked Potentials by Identification of The Fitzhugh-Nagumo Model

A. Rybalko

Institute for Problems in Mechanical Engineering,
The Russian Academy of Sciences
Saint Petersburg, Russia
https://orcid.org/0009-0003-9168-0514

Artur Sagatdinov

Institute for Problems in Mechanical Engineering,
The Russian Academy of Sciences
Saint Petersburg, Russia
https://orcid.org/0009-0000-4654-0537

Abstract—**The paper is devoted to movement-related evoked potentials modeling, which is an important problem for the development of brain-computer interfaces. The evoked potentials are modeled by the FitzHugh-Nagumo mathematical model of neuronal activity and its identification algorithm based on the speed gradient method.**

Key words—identification, speed gradient method, neural activity modeling, FitzHugh-Nagumo model

I. INTRODUCTION

Today one of the main research directions in neuroscience is the development of brain-computer interfaces [1], which allow the human brain to interact directly with external devices (such as computers or robots), for example, through the electrical activity of the brain obtained using an electroencephalogram (EEG). However, due to the noisy and ambiguous nature of the EEG signal, as well as the immense complexity of the human brain, the problem of identifying patterns indicating a person's intentions is extremely challenging. A significant portion of existing solutions are based on the application of machine learning methods to recognize movement-related evoked potentials (or the readiness potential), which are oscillations in the EEG signal of a certain type that occur when a subject is going to make a movement or even just thinks about it.

The possibilities of using the FitzHugh-Nagumo model [2, 3] of neuronal activity to model evoked potentials are studied in this paper. The parameters of the FitzHugh-Nagumo model are chosen by an adaptive identification algorithm [4, 5] based on the speed gradient method [6]. This algorithm was applied to fragments of the EEG signal corresponding to evoked potentials before moving the right or left hand. The biological plausibility of the FitzHugh-Nagumo model gives hope that detection of relationships between model's parameter values and evoked potentials will enable to use the algorithm [4] to improve the accuracy of brain-computer interfaces.

II. MODELING OF EVOKED POTENTIALS

The transformed FitzHugh-Nagumo model is considered [4]:

$$y'' = \theta_1^* y' + \theta_2^* (y^3)' + \theta_3^* y + \theta_4^* y^3 + \theta_5^*, \qquad (1)$$

where $y(t)$ is the electric potential of a single neuron or a group of neurons, θ_i^*, $i \in 1{:}5$ – unknown parameters. The problem is to find such unknown parameters estimates of (1), $\theta_i(t)$, that would provide the identification goal for an EEG signal, $\hat{y}(t): y(t) - \hat{y}(t) \to 0$ *for* $t \to \infty$. The EEG signal is a difference between signals from two electrodes, which are located on a surface of a subject's head. To solve this problem, the identification algorithm from [4] is used.

The EEG signal fragment corresponding to the movement-related evoked potential is considered as $\hat{y}(t)$. This fragment starts 1.5 seconds before the movement and ends at the moment when the movement begins [1]. The strong noise level of the EEG signal and the presence of artifacts (associated, for example, with eyes blinking) requires serious preprocessing, which is described in detail in [7].

Figures 1 and 2 show the results of applying the algorithm from [4] to the C1-C2 signal before the left hand movement. The graphs show that the identification goal is achieved quite fast and with high accuracy.

Figures 3 and 4 illustrate similar graphs but for the C4-Cz signal associated with the subject's right hand movement. The identification goal is also achieved quickly. The issue of the relationship between the simulated processes and the obtained parameter estimates, $\theta_i(t)$, requires further study.

Fig. 1. Adaptive estimates of the model parameters (1) obtained using the algorithm [4] when tuning (1) to the C1-C2 EEG signal corresponding to the evoked potential before the left hand movement.

Fig. 2. The fragment of the C1-C2 EEG signal corresponding to the evoked potential before the left hand movement (the dotted line) and the result of its modelibg by (1) with adaptive parameters (Fig. 1) (the solid line).

Fig. 3. Adaptive estimates of the model parameters (1) obtained using the algorithm [4] when tuning (1) to the C4-Cz EEG signal corresponding to the evoked potential before the right hand movement.

Fig. 4. The fragment of the C4-Cz EEG signal corresponding to the evoked potential before the right hand movement (the dotted line) and the result of its modelibg by (1) with adaptive parameters (Fig. 3) (the solid line).

III. CONCLUSION

The obtained results show that the application of the identification algorithm based on the speed gradient method [4, 5] to the FitzHugh-Nagumo neuron model is an efficient way for modeling movement-related evoked potentials in the EEG signal. However, it should be noted that the algorithm does not take into account interference in the signal. It means that high-quality preprocessing is required for correct parameters estimation. This is a difficult circumstance when working with a real-time signal, so the possibility of perturbations accounting in the algorithm requires further research. Also, the research of the relationship between the obtained adaptive estimates and the simulated processes is of great interest. The existence of such dependence will allow the proposed approach to be used in the problem of recognizing a human's intention to perform a certain action, which is extremely important for the development of brain-computer interfaces.

ACKNOWLEDGMENT

REFERENCES

[1] Nicolas-Alonso L. F., Gomez-Gil J. Brain computer interfaces, a review. Sensors. 2012. Vol. 12. No. 2. Pp. 1211-1279.

[2] FitzHugh R. Impulses and physiological states in theoretical models of nerve membrane. Biophysical Journal. 1961. Vol. 1. P. 445–466.

[3] Nagumo J., Arimoto S., Yoshizawa S. An Active Pulse Transmission Line Simulating Nerve Axon. Proceedings of the IRE. 1962. Vol. 50. No. 10. Pp. 2061–2070.

[4] A. Fradkov, A. Shepeljavyi, and A. Rybalko. Identification of the FitzHugh-Nagumo neuron model based on the speed-gradient and filtering. 2022 Fourth International Conference Neurotechnologies and Neurointerfaces (CNN), IEEE. 2022. Pp. 29–31.

[5] Rybalko A., Fradkov A. Identification of two-neuron FitzHugh–Nagumo model based on the speed-gradient and filtering. Chaos: An Interdisciplinary Journal of Nonlinear Science. 2023. Vol. 33. No. 8. Pp. 083126.

[6] A. L. Fradkov, I. V. Miroshnik, and V. O. Nikiforov. Nonlinear and Adaptive Control of Complex Systems. Springer Netherlands, 2013.

[7] Lipkovich M. et al. Evoked Potentials Detection During Self-Initiated Movements Using Machine Learning Approach. 2023 Fifth International Conference Neurotechnologies and Neurointerfaces (CNN), IEEE. 2023. Pp. 47-50.

The Introduction of Optimal Propellant Maneuvers of the International Space Station Into Regular Operation

A.A. Prut'ko

Department of dynamics and G&NC software
S.P. Korolev Rocket and Space Corporation "Energia"
Korolev, Moscow region, Russian Federation
aleksey.prutko1@rsce.ru

Abstract—The paper presents the results of the introduction into regular operation of optimal propellant maneuvers of the International Space Station controlled by the Russian segment. It was necessary to develop a methodology for performing optimal maneuvers of the ISS in such a way that the specialists of the Mission Control Center (MCC) could independently plan and perform the required maneuvers without the need to include specialists in motion control and navigation systems, who developed both the onboard software and the optimal trajectories of the ISS maneuvers themselves. Since its introduction into regular operation, at least ten optimal maneuvers of the ISS have been performed.

Keywords—International Space Station, optimal maneuver, propellant consumption, spacecraft, orbital station, angular motion control

I. INTRODUCTION

From 2017 to 2019, specialists of RSC Energia conducted the space experiment MKS-Razvorot, the purpose of which was to develop and test alternative algorithms for performing ISS maneuvers [1-5]. Typical ISS maneuvers are the following: a 180° turn around the local vertical in the orbital coordinate system [6] and a reverse maneuver conducted to ensure light-to-dark conditions before and after the docking of Russian spacecraft. During the implementation of the space experiment, a software component for tracking a given trajectory was developed and introduced into the ISS onboard software. The space experiment was successfully completed, and the implementation of optimal propellant consumption maneuvers of the ISS by tracking a given trajectory was introduced into the ISS regular operation. During the experiment, propellant consumption for a typical maneuver was reducced to 5-10 kg instead of 50-60 kg. After the docking of the Multipurpose Laboratory module Nauka and the node module Prichal in 2021, the construction of the Russian segment was completed. In this regard, the mass inertia characteristics should not change significantly in the future. It was proposed to develop a list of optimal maneuvers trajectories for typical ISS reorientations, which MCC specialists could use during the long-term operation of the ISS. This paper presents the results of the work carried out.

II. DEVELOPMENT OF THE ISS OPTIMAL MANEUVERS TRAJECTORIES

Since the construction of the ISS Russian segment is currently considered completed, the private segment Axiom is under development; the change in mass inertia characteristics depends on the docked cargo and transport spacecraft: the Soyuz MS manned transport spacecraft, the Progress MS transport cargo spacecraft, Cygnus, SpaceX Dragon and Dragon Crew, Boeing Starliner, Dream Chaser, and HTV. Various possible combinations of docked spacecraft to various docking sites lead to hundreds of different ISS configurations. Since spacecraft dock to the ISS mainly in one plane – the plane of the orbit, the torque equilibrium attitude of the ISS changes practically only in pitch. Thus, all configurations were averaged into eight configurations with torque equilibrium attitudes about one and a half degrees apart. The list of ISS configurations is presented in Table 1. The list of torque equilibrium attitudes corresponding to the ISS configurations in the "acceleration" and "deceleration" positions is presented in Table 2.

For the presented configurations of the ISS and the corresponding torque equilibrium attitudes, motion control and navigation systems software developers obtained optimal propellant consumption maneuvers trajectories for typical turns of the ISS from the "acceleration" position to the "deceleration" position and back with a duration of 5500 seconds. Sixteen ISS reorientation trajectories have been obtained, which should be sufficient for the coming years of operation. All trajectories were worked out at the ground-based development complex [7] using all permissible jet configurations and were allowed to operate normally.

TABLE I. AVERAGED ISS CONFIGURATIONS FOR THE DEVELOPMENT OF OPTIMAL MANEUVERS TRAJECTORIES. J_{XX}, J_{YY}, J_{ZZ} – ISS DIAGONAL ELEMENTS OF THE INERTIA TENSOR, J_{XY}, J_{XZ}, J_{YZ} – OFF-DIAGONAL MOMENTS OF INERTIA WITH A NEGATIVE SIGN. THE DIMENSION – KG·M²

	J_{XX}	J_{YY}	J_{ZZ}
K1	142256416	191330160	81546568
K2	142046336	194799904	84795616
K3	141949632	198996912	88879448
K4	141505888	200807872	90237440
K5	139817120	199651968	87399136
K6	138726720	199555600	86213552
K7	138701456	200945728	87570016
K8	138723136	201136112	87782240
	J_{XY}	J_{XZ}	J_{YZ}
K1	10711161	4356336	443796
K2	9930085	4320057	441938
K3	9062321	4292516	446268
K4	7753397	4268089	456441
K5	6092651	4205908	520286
K6	4494287	4157578	556886
K7	3004080	4133860	549067
K8	1443159	4114245	530280

TABLE II.   THE TORQUE EQUILIBRIUM ATTITUDES FOR THE ISS CONFIGURATIONS FROM TABLE I REPRESENTED IN KRYLOV ANGLES IN THE YAW-PITCH-ROLL SEQUENCE IN DEGREES

	"Acceleration" position		
	Yaw	*Pitch*	*Roll*
K1	-176.0	11.47	-0.85
K2	-176.0	10.11	-0.85
K3	-176.0	8.7	-0.85
K4	-176.0	7.27	-0.85
K5	-176.0	5.76	-0.85
K6	-176.0	4.29	-0.85
K7	-176.0	2.9	-0.85
K8	-176.0	1.52	-0.85
	"Deceleration" position		
	Yaw	*Pitch*	*Roll*
K1	4.0	11.08	-0.85
K2	4.0	9.74	-0.85
K3	4.0	8.34	-0.85
K4	4.0	6.92	-0.85
K5	4.0	5.41	-0.85
K6	4.0	3.94	-0.85
K7	4.0	2.56	-0.85
K8	4.0	1.18	-0.85

Thus, the MCC specialists need to perform the following steps:

- select the required turn trajectory in advance, the one that is closest to the initial torque equilibrium attitude of the ISS from Table 2, and upload it to the onboard memory;

- start performing a maneuver by tracking a given trajectory and wait for the end of the maneuver;

- complete the ISS rotation to the standby attitude in the standard way along the shortest trajectory.

III. RESULTS OF PERFOMING ISS OPTIMAL MANEUVERS

At least ten maneuvers have been performed since the introduction of the ISS optimal propellant maneuvers into regular operation. Their results on propellant consumption are presented in Table 3. All the maneuvers were completed successfully. All propellant consumption results are less than eight kilograms per maneuver.

TABLE III.   THE RESULTS OF PERFORMING OPTIMAL PROPELLANT MANEUVERS OF THE ISS FOR 2022 AND 2023.

Date	Type	Direction	Consumption, kg
09.21.2022	K3	Forward	7.28
10.27.2022	K5	Forward	6.58
10.28.2022	K6	Reverse	4.03
02.19.2023	K5	Reverse	3.75
02.25.2023	K5	Forward	6.12
02.26.2023	K6	Reverse	4.38
05.24.2023	K4	Forward	4.94
05.25.2023	K4	Reverse	4.1
09.15.2023	K3	Forward	5.04
09.16.2023	K3	Reverse	4.09

IV. CONCLUSION

Sixteen trajectories have been developed for regular operation to perform optimal propellant maneuvers of the ISS. Of these, eight are from the "acceleration" position to the "deceleration" position and back. The trajectories were developed based on eight averaged ISS configurations, taking into account various combinations of docked visiting spacecraft. MCC specialists can use the proposed list of trajectories to perform maneuvers of the ISS in the coming years of operation. In 2022 and 2023, optimal maneuvers of the ISS were successfully carried out, which provided favorable conditions for docking and undocking spacecraft.

REFERENCES

[1] MKS-razvorot / JSC «TsNIIMash» — URL: https://tsniimash.ru/en/science/scientific-experiments-onboard-the-is-rs/cnts/experiments/?PAGEN_1=9 (Accesed 04/19/2024).

[2] Atroshenkov S.N., Platonov V.N., Gubarev F.V., Saratov A.A. Optimal'nyi po raskhodu topliva algoritm razvorota MKS s pomoshch'yu reaktivnykh dvigatelei s uchetom ogranichenii po nagruzkam na konstruktsiyu [Algorithm optimal in terms of the fuel consumption for the ISS rotation using jet engines with regard to the structural load limitations]. Vestnik MGTU im. N.E. Baumana. Ser. Priborostroenie, 2017, no. 4, pp. 118–138. DOI 10.18698/0236-3933-2017-4-118-138.

[3] Prutko A.A., Atroshenkov S.N., Bogachev A.V., Starchenko A.E. "Using Pseudospectral method to search for propellant-optimal trajectories of the International space station attitude control maneuvers," Kosmicheskaya tekhnika i tekhnologii [Space engineering and technology]. 2019. № 4(27). pp. 121–133.

[4] Atroshenkov S.N., Prutko A.A., Krylov A.N., Krylov N.A., Gubarev F.V., "Aerodynamic forces and torques simulation for the verificaion of International Space Station guidance, navigation and control system dynamic modes anf for optimal control synthesis," Space engineering and technology. 2017. № 4(19). pp. 72–88.

[5] Prutko A.A. Rezul'taty provedeniya manevrov krupnogabaritnoj orbital'noj stancii po traektoriyam, optimiziruyushchim raskhod topliva [The results of maneuvers of a large-sized orbital station along optimal propellant trajectories], Collection of materials of the All-Russian youth scientific and practical conference « Orbita molodyozhi-2019». — St. Petersburg: Baltic State Technical University "Voenmeh" D.F. Ustinov, 2019. — pp. 130—132.

[6] Borisenko N.Yu., Sumarokov a.V. Ob uskorennom postroenii orbital'noi orientatsii gruzovykh i transportnykh korablei serii «Soyuz MS» i «Progress MS» [About accelerated establishment of orbital attitude of logistics and crew transportation spacecraft of the Soyuz MS and Progress MS series]. Izvestiya RaN. Teoriya i sistemy upravleniya, 2017, no. 5, pp. 131–141. DOI: 10.7868/S0002338817050110.

[7] Mikrin E.A. Bortovye kompleksy upravleniya kosmicheskikh apparatov [Onboard spacecraft control systems]. Moscow, MGTU im. N.E. Baumana publ., 2014. 254 p.

Navigation Complex for UAV

D. B. Pazychev
Integral Ltd
Moscow, Russia
d.pazychev@integral-group.pro

K. S. Bakulev
NITU MISIS
Moscow, Russia
bakulev.ks.@misis.ru

Abstract— **The article discusses the navigation complex built on the basis of the Integral LLC products. We consider the composition of the complex and describe the algorithm of its operation; the results of the navigation system tests on a light-engine aircraft are also presented, giving an idea of its accuracy.**

Keywords: navigation system, air data system, magnetometer, navigation algorithm, UAV

I. INTRODUCTION

Modern unmanned aerial vehicles (UAVs) rely on navigation systems to determine their current position when performing flight tasks along a given route. These navigation systems are usually very compact, based on microelectromechanical systems (MEMS) sensors, and are often part of UAV flight controllers such as Pixhawk (Fig. 1), SpeedyBee, BetaFlight, and others.

Fig. 1. Pixhawk flight controller with Cube Orange navigation system

In addition, similar controllers, such as the Pixhawk, already include a three-component magnetic field sensor and it is possible to connect a total airflow pressure gauge with simple air tube (Fig. 2).

Fig. 2. Airspeed sensor for the Pixhawk control controller

It should be noted that, despite the abundance of measurement sensors in such controllers, the accuracy of the navigation problem solution in the event of loss of the satellite navigation system (SNS) signal, which is also included in the control controller, is extremely low. There are two reasons for this:

– The system sensors (gyroscopes, accelerometers, magnetic field sensors and pressure sensors) in such devices have extremely low accuracy and are only subject to simple calibration procedures. In most cases, the only procedure performed on them is to remove static sensor bias. The sensors still have large in-run bias instabilities that are not compensated by the software.

– The Extended Kalman Filter (EKF) in the controller's navigation algorithm is in most cases implemented in Attitude and Heading Reference System (AHRS) mode, i.e. it is only used to plot the UAV angular position..

The poor quality of the navigation solution provided by standard flight controllers, especially in SNC signal loss mode, makes it necessary to develop an additional navigation complex capable of solving this task.

This paper proposes an additional (external to the flight controller) navigation complex connected to the UAV controller via a CAN interface using the UAVCAN (droneCAN) communication protocol. Thus, when connecting the proposed complex to the UAV controller, the algorithms and software for UAV control remain unchanged and the navigation solution is replaced by a better one.

II. COMPLEX COMPOSITION

The navigation complex proposed in this paper (Fig. 3) consists of several devices connected by a single CAN interface:

- Strapdown Inertial Navigation System (SINS)

- Air data system (ADS)

- Digital magnetic heading sensor (DMHS).

Fig. 3. Navigation complex for UAV

The basis of the complex is AHRS NV-1 which forms a common high-frequency navigation solution of the whole

complex, including angular velocities, accelerations, angles, linear velocities and geographical coordinates of the UAV.

In order to eliminate the accumulation of errors of linear velocities of the complex, it includes the air data system ADS-1, which provides the whole complex with information on static and total pressures, air velocities, and barometric altitude of the UAV.

In order to eliminate the accumulation of errors in the azimuth channel of the AHRS NV-1, the complex includes a digital magnetic heading sensor DMHS-1, which makes it possible to measure the UAV's magnetic field components, magnetic heading and magnetic declination.

In addition, it is possible to use the data logger DR-1 in the complex to store all the data from all the devices on an SD card for further analysis.

The navigation system proposed in this article has a number of distinctive features:

– Flexibility in the placement of its components on the UAV: the navigation system is placed in a location with the least amount of vibration (vibration-isolated shocks can also be used), the ADS is placed as close as possible to the incoming airflow, and the magnetic heading sensor is placed in a location with the least amount of magnetic disturbance (wing or tail for an airborne UAV).

– It is possible to make some or all of the components of the set in an uncased version, which allows not only a significant reduction in the size of the set, but also in its weight (up to 150 g in the uncased version).

– Easy switching between individual components of the complex. Data exchange between the units and signal transmission to the UAV controller is via a single CAN interface using the standard UAVCAN version 0 (DroneCAN) protocol.

III. AHRS NV-1

It is well known that one of the peculiarities of using MEMS sensors as part of the navigation system on small UAVs is the strong influence of external vibration and temperature on their readings.

To eliminate vibration errors, shock-absorbing vibration mounts with a certain cut-off frequency are usually used to limit the penetration of high-frequency vibrations into the UAV navigation system.

Again, to eliminate sensor temperature errors, in most cases designers are limited to performing temperature compensation of errors according to the sensor temperature model given below (1):

$$\vec{\omega}_B^{Err} = \begin{bmatrix} \beta_{XX}(t) & \beta_{XY}(t) & \beta_{XZ}(t) \\ \beta_{YX}(t) & \beta_{YY}(t) & \beta_{YZ}(t) \\ \beta_{ZX}(t) & \beta_{ZY}(t) & \beta_{ZZ}(t) \end{bmatrix} \cdot \vec{\omega}_B^{Tr} + \begin{bmatrix} \beta_X(t) \\ \beta_Y(t) \\ \beta_Z(t) \end{bmatrix}$$

$$\vec{a}_B^{Err} = \begin{bmatrix} \alpha_{XX}(t) & \alpha_{XY}(t) & \alpha_{XZ}(t) \\ \alpha_{YX}(t) & \alpha_{YY}(t) & \alpha_{YZ}(t) \\ \alpha_{ZX}(t) & \alpha_{ZY}(t) & \alpha_{ZZ}(t) \end{bmatrix} \cdot \vec{a}_B^{Tr} + \begin{bmatrix} \alpha_X(t) \\ \alpha_Y(t) \\ \alpha_Z(t) \end{bmatrix} \tag{1}$$

where $\beta_{ii}(T_{Gi})$ и $\alpha_{ii}(T_{Ai})$ - temperature dependent scale factors errors of gyroscopes and accelerometers;

$\beta_{ii}(T_{Gi})$ и $\alpha_{ii}(T_{Ai})$ - temperature dependent non orthogonality errors of gyroscopes and accelerometers;

$\beta_i(T_{Gi})$ и $\alpha_i(T_{Ai})$ - temperature-dependent biases of gyroscopes and accelerometers.

Each of the coefficients of the sensor error model presented above (1) can, in turn, be represented as a high-order power function, the coefficients of which are to be determined during the calibration procedure (2):

$$f_j(T_i) = K_{j,N} \cdot T_i^N + K_{j,N-1} \cdot T_i^{N-1} + ... + K_{j,0} \tag{2}$$

Such a mathematical error model requires special technological equipment for the identification of its individual coefficients - the equipment should be able to change the orientation of the system (to change the effect of the Earth's free fall acceleration on the sensitivity axes of the gyroscopes and accelerometers), to rotate the system at different angular speeds (to identify the error coefficients of the gyroscopes) and, most importantly, to change the temperature of the navigation system in a wide range (from -40 °C to +70 °C).

To exclude such a complicated complex in the production of the system, another principle of compensation of sensors temperature errors–temperature stabilization– is used in the AHRS NV-1 (Fig. 4, Table 1).

Fig. 4. Enclosed and uncased version of the AHRS NV-1

The peculiarity of this stabilization circuit can be seen in the presence of two heaters at the same time (Fig. 5), which ensures the best maintenance of the sensor readings at the set temperature of +50 ^{0}C.

TABLE I. AHRS NV-1 TECHNICAL CHARACTERISTICS

Parameter	Value
Size	75 x 102 x 41 mm – enclosed, 39 x 65 x 16 mm – uncased
Weight	200 g - enclosed, 30 g - uncased
Voltage	= 12 … 36 V
Power	3 W during warm up, 350 mW stabilization
Interface	1 x RS-422, 1 x CAN
Frequency	50…400 Hz
Temperature	- 50 ^{0}C …+ 70 ^{0}C
Humidity	5% … 98 %
Gyroscopes range	± 250 ^{0} / sec (can be changed to ± 500 ^{0} sec, ± 1000 ^{0} / sec or ± 2000 ^{0} / sec)
Accelerometers range	± 2g (can be changed to ± 4g, ± 8g or ± 16g)
GNSS receiver	One frequency, one antenna, installed inside

The use of the above temperature stabilization of the sensors not only simplifies the system calibration procedure, but also increases the stability of the sensor readings (mainly zero offset and scaling factors) over the entire temperature range of the system application.

Fig. 5. AHRS NV-1 temperature stabilization system

IV. AIR DATA SYSTEM ADS-1

The air data system ADS-1 (Fig. 6, Table 2) is a device designed to measure the barometric height of an object and its speed relative to the air flow. The device consists of 2 pressure sensors: static and total, united in a single housing with a calculator of parameters inside. In addition, the device can be connected to the P-104M brake temperature sensor to improve its accuracy characteristics.

Fig. 6. Enclosed and uncased version of the ADS-1

A characteristic feature of the system should be the presence of a factory calibration of the system sensors according to the following mathematical model (3):

$$
\begin{aligned}
P_{St}^{Err} &= \rho_{SS}(t) \cdot P_{St} + \rho_{S}(t) \\
P_{Dn}^{Err} &= \rho_{DD}(t) \cdot P_{Dn} + \rho_{D}(t)
\end{aligned}
\tag{3}
$$

where $\rho_{SS}(t)$ and $\rho_{DD}(t)$ are errors in the scaling factors of the static and total pressure sensors respectively.

$\rho_{S}(t)$ and $\rho_{D}(t)$ - zero offset error of static and total pressure sensors respectively.

P_{St} and P_{Dn} are true values of static and total pressure respectively.

P_{St}^{Err} and P_{Dn}^{Err} - errors in static and total pressure sensor readings respectively.

TABLE II. ADS-1 TECHNICAL CHARACTERISTICS

Parameter	Value
Size	122 x 55 x 26 mm– enclosed, 70 x 40 x 18 mm– uncased
Weight	200 g - enclosed, 40 g - uncased
Voltage	= 12 … 30 V
Power	350 mW
Interface	1 x RS-422, 1 x CAN
Frequency	50…100 Hz
Temperature	- 50 ^{0}C … + 70 ^{0}C
Humidity	5% … 98 %
Speed range	3…750 km/h

By analogy with gyroscopic sensors, each of the coefficients in model (3) can be represented as a degree function (2) whose temperature-dependent coefficients must be identified.

Unlike the AHRS NV-1, which uses ultra-compact micro-mechanical angular velocity and acceleration sensors, the air data system ADS-1 uses very large pressure sensors. This circumstance does not allow to build a high quality and compact system of temperature stabilization of these sensors, which leads to the necessity of classical temperature compensation of errors of pressure sensors. Within the framework of the presented mathematical model of sensor errors (3), the company Integral Ltd has developed a technological bench for calibration of pressure sensors, which is a system of pressure adjusters, tubes and a heat-cold chamber, united in a common automated system with software control. With the help of this stand it is possible to identify the error model (3) even up to high order polynomials, which provide high accuracy of pressure measurements after calibration of sensors.

V. DIGITAL MAGNETIC HEADING SENSOR DMHS-1

The Digital Magnetic Heading Sensor DMHS-1 (Fig. 7) is a device designed to measure the magnetic heading of an object (the angle between the longitudinal axis of the object and the direction of the Earth's magnetic north, counted in the plane of the local horizon). In addition, the DMHS-1 measures the object's roll and pitch angles, as well as the three components of the magnetic field around its location. The instrument consists of several sensors: a multi-component angular velocity and acceleration sensor, and three magneto-induction magnetic field sensors arranged orthogonally to each other.

Fig. 7. Enclosed and uncased version of the DMHS-1

As in the case of the AHRS NV-1, this device also undergoes a factory calibration of the sensor readings. In addition to the mathematical error model (1) introduced for the identification of errors of gyroscopes and accelerometers,

this device introduces an additional mathematical model (4) for the identification of errors of magnetic field sensors:

$$\vec{M}_B^{Err} = \begin{bmatrix} \mu_{XX}(t) & \mu_{XY}(t) & \mu_{XZ}(t) \\ \mu_{YX}(t) & \mu_{YY}(t) & \mu_{YZ}(t) \\ \mu_{ZX}(t) & \mu_{ZY}(t) & \mu_{ZZ}(t) \end{bmatrix} \cdot \vec{M}_B^{Tr} + \begin{bmatrix} \mu_X(t) \\ \mu_Y(t) \\ \mu_Z(t) \end{bmatrix} \quad (4)$$

where $\mu_{ii}(t)$ - errors of scale factors of magnetic field sensors;

$\mu_{ij}(t)$ - errors of non-orthogonality of magnetic field sensor axes;

$\mu_i(t)$ - zero offset errors of magnetic field sensors.

$\vec{M}_B^{Tr}$ - vector of true magnetic field values.

$\vec{M}_B^{Err}$ - magnetic field measurement error vector.

TABLE III.   DMHS-1 TECHNICAL CHARACTERISTICS

Parameter	Value
Size	95 x 45 x 25 mm– enclosed 50 x 35 x 15 mm– uncased
Weight	200 g - enclosed, 25 g - uncased
Voltage	= 12 … 30 V
Power	300 мW
Interface	1 x RS-422, 1 x CAN
Frequency	20 Hz
Temperature	- 50 ^{0}C … + 70 ^{0}C
Humidity	5% … 98 %
Magnetic field range	± 800 mTesla

In most cases, magnetic field sensors are calibrated by rotating the device about three axes and recording data at the moment of rotation (Fig. 8).

Fig. 8.   Magnetic field sensor readings during rotation around three axes

The coefficients estimated by this method are characterized by a high level of error and do not allow their estimation as a function of temperature variation. For this reason, company Integral Ltd. has developed a technological stand for the calibration of magnetic field sensors, consisting of a magnetically shielded case and a system of 3 orthogonal inductors that allow the creation of a magnetic field of different strength and direction. Using this stand, it is possible to identify the error model (4) with high accuracy in order to ensure qualitative measurements of the magnetic field after calibration of the sensors.

VI. NAVIGATION COMPLEX ALGORITHM

The augmentation of inertial navigation systems with various external systems (such as global positioning systems, airspeed sensors, magnetometers, etc.) is in most cases carried out using Kalman-type filters. In this case, the classical mathematical model of inertial navigation system errors (5) is considered as a mathematical model of the filter, whose parameters should be identified in order to improve the accuracy of the SINS output signal:

$$\begin{cases} \delta\dot{E} = \delta V_E \\ \delta\dot{N} = \delta V_N \\ \delta\dot{V}_E = -g\Phi_N + a_N\Phi_{up} + B_E + w(t) \\ \delta\dot{V}_N = -g\Phi_E + a_E\Phi_{up} + B_N + w(t) \\ \dot{\Phi}_E = -\dfrac{\delta V_N}{R} + \omega_E^{dr} + w(t) \\ \dot{\Phi}_N = \dfrac{\delta V_E}{R} + \omega_N^{dr} + w(t) \\ \dot{\Phi}_{up} = \dfrac{\delta V_E}{R}tg\varphi + \omega_{up}^{dr} + w(t) \\ \omega_E^{dr} = \omega_{xB}^{dr} \cdot \cos H + \omega_{yB}^{dr} \cdot \sin H \\ \omega_N^{dr} = \omega_{yB}^{dr} \cdot \cos H - \omega_{xB}^{dr} \cdot \sin H \\ \dot{\omega}_{xB}^{dr} = -\beta\omega_{xB}^{dr} + A\sqrt{2\beta}w(t) \\ \dot{\omega}_{yB}^{dr} = -\beta\omega_{xB}^{dr} + A\sqrt{2\beta}w(t) \\ \dot{\omega}_{up}^{dr} = -\beta\omega_{up}^{dr} + A\sqrt{2\beta}w(t) \end{cases} \quad (5)$$

In addition to the error model of the inertial system (5) in the Kalman filter in the case we are considering, it is also necessary to introduce as a mathematical model the errors of external devices - the system of airborne devices and the magnetic heading sensor, which are themselves complex devices and have their own very complex and extensive error models. This situation will lead to the fact that the introduced Kalman filter will be quite cumbersome and will require a very large number of settings for its correct and qualitative work. Without detracting in any way from this approach to the construction of navigation system software, in this paper we propose an alternative variant that requires a smaller number of settings.

This method is based on the classic SINS algorithm, which calculates coordinates by double integration of accelerometer accelerations, which are recalculated into the navigation coordinate system using the cosine guide matrix updated in the Poisson equation (Fig. 9).

In order to eliminate accumulating SINS errors in the calculation of angles, velocities and coordinates, three control angular velocities are introduced into the algorithm through feedback coefficients K1, K2 and K3: coefficients K1 and K2 are introduced to eliminate SINS linear velocity errors, and coefficient K3 is introduced to reset the accumulating heading error.

The introduction of such feedbacks leads to a modification of the classical SINS error equations (5) to the following form (6):

Fig. 9. Navigation complex algorithm

$$
\begin{cases}
\delta\dot{V}_E^{SINS} = -g\Phi_N + a_N\Phi_{up} + B_E \\
\qquad - K_1(\delta V_E^{SINS} - \delta V_E^{ADS}) \\
\delta\dot{V}_N^{SINS} = g\Phi_E + a_E\Phi_{up} + B_N \\
\qquad - K_1(\delta V_N^{SINS} - \delta V_N^{ADS}) \\
\dot{\Phi}_E = -\dfrac{\delta V_N^{SINS}}{R} + \omega_E^{dr} - K_2(\delta V_N^{SINS} - \delta V_N^{ADS}) \\
\dot{\Phi}_N = \dfrac{\delta V_E^{SINS}}{R} + \omega_N^{dr} + K_2(\delta V_E^{SINS} - \delta V_E^{ADS}) \\
\dot{\Phi}_{up} = \dfrac{\delta V_E}{R} tg\varphi + \omega_{up}^{dr} - K_3 \cdot (\delta\Phi_{up} - \delta H_{MAGN}) \\
\delta\dot{S}_E^{SINS} = \delta V_E^{SINS} \\
\delta\dot{S}_N^{SINS} = \delta V_N^{SINS}
\end{cases}
\tag{6}
$$

where δV_E^{ADS} and δV_N^{ADS} - errors of air velocity projections on the east and north axis, respectively.

δH_{MAGN} - error of magnetic heading measurement by the DMHS-1 sensor.

Solution of the above system of equations in terms of velocity error and azimuthal channel gives the following equations (7):

$$
\begin{cases}
\delta\ddot{V}_E^{SINS} + K_1\delta\dot{V}_E^{SINS} + (v^2 + K_2 g)\delta V_E^{SINS} = \\
\qquad -g\omega_N^{dr} + \dot{B}_E + (v^2 + K_2 g)\delta V_E^{ADS} \\
\delta\ddot{V}_N^{SINS} + K_1\delta\dot{V}_N^{SINS} + (v^2 + K_2 g)\delta V_N^{SINS} = \\
\qquad -g\omega_E^{dr} + \dot{B}_N + (v^2 + K_2 g)\delta V_N^{ADS} \\
\dot{\Phi}_{up} + K_3 \cdot \delta\Phi_{up} = \dfrac{\delta V_E^{SINS}}{R} tg\varphi + \omega_{up}^{dr} + \delta H_{MAGN}
\end{cases}
\tag{7}
$$

The first two equations of the error system presented are second order oscillatory links, the third equation is a first order aperiodic link. Both these terms give damped transients

(with the correct choice of coefficients) which tend to the following steady state error values (8):

$$
\begin{cases}
\delta V_E^{SINS} = -g\omega_N^{dr} / K_2 g + \delta V_E^{ADS} \\
\delta V_N^{SINS} = -g\omega_E^{dr} / K_2 g + \delta V_N^{ADS} \\
\Phi_{up} = \omega_{up}^{dr} / K_3 + \delta H_{MAGN}
\end{cases}
\tag{8}
$$

As can be seen from the equations presented, the calculation errors of linear velocities and heading angles are determined by the accuracy of the external systems - the air signal system and the magnetic heading sensor.

In order to obtain the errors of the ADS velocities, it is possible to use the equations describing the airspeed transformations (9):

$$
\begin{cases}
V_E^{ADS} = V_{AIR} \cdot \cos(\vartheta) \cdot \sin(H^{SINS}) + W_E \\
V_N^{ADS} = V_{AIR} \cdot \cos(\vartheta) \cdot \cos(H^{SINS}) + W_N
\end{cases}
\tag{9}
$$

where V_{AIR} - air data system readings

W_E and W_N - values of the wind speed in projection on the east and north axes respectively. These components are taken into account because the airborne signal system measures the speed relative to the airflow, and to calculate the current coordinates of the object it is necessary to integrate the speeds relative to the Earth's surface. Thus, the use of SINS correction from the air data system implies the initial determination of wind speed projections using air and ground speed signals, i.e. the resolution of the velocity triangle (Fig. 10) in the presence of both GNSS and ADS signals.

Varying equations (9) from the airspeed projections themselves on the east and north axes, we can go to their errors (10):

$$
\begin{cases}
\delta V_E^{ADS} = \delta V_{AIR} \cdot \sin(H^{SINS}) + \\
\qquad V_{AIR} \cdot \cos(\delta H^{SINS}) + \delta W_E \\
\delta V_N^{ADS} = \delta V_{AIR} \cdot \cos(H^{SINS}) - \\
\qquad V_{AIR} \cdot \sin(\delta H^{SINS}) + \delta W_N
\end{cases}
\tag{10}
$$

where δV_{AIR} - residual error in airspeed measurement by the air signal system after calibration.

δW_E, δW_N - errors in the wind speed projections on the east and north axes. These errors are caused both by incorrect initial values at the stage of their estimation at the resolution of the velocity triangle (Fig. 10), and by changes in wind speed and direction during the integration process.

Fig. 10. Aircraft speeds in projection to the horizon plane

The total error in the determination of the coordinates is thus determined by three componentsи:

$$\delta S^{SINS} \sim \int_0^t \delta V_{AIR}\, dt + \int_o^t V_{AIR} \cdot \delta H^{SINS} dt + \int_o^t \delta W dt \qquad (12)$$

The first component is the integral of the airspeed error. This in turn is determined by both the design accuracy of the airspeed signal of the airspeed indicator system and the loss of accuracy during the pressure transfer from the air pressure instrument (API) to the ADS in the pipe system. The second component is the lateral deviation from the trajectory caused by integrating the speed with the heading error. The third component is the integral of the wind speed calculation error or its change during the integration process. Equation (12) is very useful in practice - on its basis it is possible to make a preliminary theoretical estimate of the accuracy of the whole complex.

Thus, the residual speed error of the air data system ADS-1 is in the order of 1 m/s. The error of magnetic heading measurement by the DMHS-1 instrument is in the range of 1°, but this instrument is located in the UAV and is influenced by the electronics of the UAV itself. Despite the additional procedure of writing off magnetic deviations as part of the object, practice shows that the residual error of magnetic heading measurement can be within 2°. Holding these values constant and substituting them into formula (12), at a UAV speed of 100 km/h, the following error in the calculation of complex coordinates can be expected during the first 10 minutes of loss of communication with the GNSS:

$$\delta S^{SINS} \approx 600s \cdot 1m/s + 600s \cdot \sin(2°) \cdot 27m/s$$
$$\approx 1100m \qquad (13)$$

At the same time, the range in this mode will be in the order of $600s \cdot 27m/s \approx 16\,km$. Thus, the error in determining geographic coordinates in this mode of motion will be about 7% of the path travelled by the UAV.

This value indicates the high accuracy of the proposed complex - the equivalent accuracy of geographic position determination using only the inertial navigation system is possible only with the use of more accurate fibre-optic gyroscopes (FOG) and quartz accelerometers, which will lead to a significant increase in the size of the system, its weight and, most importantly, its price.

VII. FIELD TESTS

Field tests of the navigation complex were carried out as part of the navigation unit VPNP-1M (Fig. 11) of the company Navigator on the aircraft AN-2 TVS 2MS (Fig. 12).

Fig. 11. External view of the navigation unit VPNP-1M

Fig. 12. AN-2 TVS 2MS airplane

This unit combines the AHRS NV-1 and the air data system ADS-1. In addition, it is possible to connect a digital magnetic heading sensor DMHS-1 to the unit in order to organize a heading-air mode of coordinate calculation according to the algorithm described in this article.

In total, three flights of the aircraft described above were carried out, during which the aircraft turns were performed with roll and pitch angles of about 30 degrees at a speed of about 150...200 km/h (Fig. 13, 14).

To simulate the loss of the GNSS signal, a special command was sent to the AHRS NV-1, causing the algorithm to switch to the Heading-Air mode of coordinate calculation. The GNSS was not physically switched off, which made it possible to compare the accuracy of the complex's coordinate calculation with the actual GNSS coordinates (Fig. 16).

Fig. 13. Attitude angles behavior during the flight

Fig. 14. Behaviour of airspeed (black) and ground speed (green) in one of the flights

Fig. 15. Trajectory of the aircraft in one of the flights

Fig. 16. Behaviour of the navigation complex error in the calculation of coordinates in comparison with SNS readings in one of the flights

VIII. TEST RESULTS

The results of the three flights are summarized in Table 4. As can be seen from the table, the navigation system shows an acceptable accuracy for aerobatics even when performing active aerobatics with large values of orientation angles.

TABLE IV.   NAVIGATION COMPLEX ACCURACY

Parameter	Test 1	Test 2	Test 3
Without GNSS time, sec	605	612	608
Coordinates error, m	1702	1814	1762
Distance travelled, m	23595	25784	24512
Error, %	7,2	7,0	7,1

As can be seen from the presented data, the obtained values of coordinate errors exceed the theoretical values calculated by formula (13). This is explained by the fact that the speed of the aircraft movement exceeded the calculated one by 1.5-2 times. If we judge the results by the error as a percentage of the travelled path, the obtained values are close to the theoretically calculated values.

IX. CONCLUSION

The navigation system presented in this paper is compact in size and weight and has the potential to be easily installed on board virtually any modern aircraft-based UAV. The navigation system combines several devices and provides the UAV's autopilot system with all the necessary information, including altitude-air parameters and magnetic heading angle.

The algorithm of the proposed navigation system uses the classical SINS algorithm with control signals proportional to the difference of the SINS, ADS and magnetic heading sensor errors. Changing the correction coefficients of the signals during the operation of the algorithm allows us not only to obtain optimal results in terms of accuracy, but also to calculate the parameters of the navigation algorithm with minimum delays in order to obtain the highest dynamic performance.

Field tests on a light aircraft have demonstrated the high dynamics and accuracy of this navigation system in the compact equipment segment.

REFERENCES

[1] Golovan A.A., Parusnikov N.A. Matematicheskie osnovy navigacionnyh sistem (Mathematical Foundations of Navigation Systems). Part 1.Max Press Publ., 2011. 136 s. (in Russian)

[2] Dmitriev, S.P., Inertsial'nye metody v inzhenernoi geodezii (Inertial Methods in Engineering Geodesy), St. Petersburg, CSRI Elektropribor, 1997. (in Russian)

[3] Salychev, O.S., MEMS-based inertial navigation: Expectations and reality, Moscow: Bauman MSTU Press, 2012, 207 p.

[4] Kuzovkov N. T., Karabanov S. V., Salychev O. S. Nepreryvnye i diskretnye sistemy upravleniya i metody identifikacii (Continuous and discrete control systems and identification methods), Moscow, Mashinostroenie, 1978 (in Russian).

[5] D. B. Pazychev and R. N. Sadekov, "MEMS Temperature Stabilization," 2020 27th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), St. Petersburg, Russia, 2020, pp. 1-4, doi: 10.23919/ICINS43215.2020.9133899

[6] D. B. Pazychev and R. N. Sadekov, "Simulation of INS Errors of Various Accuracy Classes," 2020 27th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), St. Petersburg, Russia, 2020, pp. 1-3, doi: 10.23919/ICINS43215.2020.9133869.

[7] D. B. Pazychev, K. S. Bakulev and R. N. Sadekov, "Low-Cost Navigation System for UAV," 2023 30th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), Saint Petersburg, Russian Federation, 2023, pp. 1-6, doi: 10.23919/ICINS51816.2023.10168469

[8] https://www.integral-group.pro/

Comparison of Gravity Anomaly Determination Methods using a Strapdown Scalar Gravimeter

E.V. Dranitsyna
Concern CSRI Elektropribor, JSC
ITMO University
St. Petersburg, Russia
ORCID 0000-0003-2591-5816

A.A. Krasnov
Concern CSRI Elektropribor, JSC
St. Petersburg, Russia
ORCID 0000-0003-0298-2418

A.V. Motorin
Concern CSRI Elektropribor, JSC
ITMO University
St. Petersburg, Russia
ORCID 0000-0002-2093-5079

O.A. Stepanov
Concern CSRI Elektropribor, JSC
ITMO University
St. Petersburg, Russia
ORCID 0000-0003-3640-3760

Abstract—**A comparison of direct and indirect methods for gravity anomaly determination on the sea using a strapdown gravimeter was carried out. The results of the bench test for direct and indirect processing methods are presented, which did not show significant advantages of any of the considered methods for marine applications. The variability of the accelerometer bias caused by the vertical accelerations has been revealed, which should be taken into account in the inertial measurement unit error model.**

Keywords—strapdown gravimeter, marine gravimetry, gravity anomaly, smoothing algorithms

I. INTRODUCTION

Strapdown gravimeters are increasingly used for the gravitational field surveys [1–18] due to their smaller weight and size, as well as lower energy consumption as compared to their platform counterparts [9, 15–16]. In this case, we are mainly talking about scalar gravimetry, i.e. determination of gravity anomaly (GA). It is known that strapdown gravimeters are widely use in aerogravimetry; currently, the work is also underway aimed for their marine applications [8, 11, 16]. When using strapdown technologies, it is common to distinguish two methods of data processing, which, following [13], will be referred to as direct and indirect methods. The direct method involves calculation of the orientation matrix by processing data from the inertial measurement unit (IMU) and global navigation satellite systems (GNSS) [4, 8, 10–11, 13, 15, 18] and subsequent use of the matrix to obtain the full value of the vertical specific force using the readings of the IMU accelerometers. It consists of two stages. At the first stage, the vertical specific force is formed; at the second stage, this vertical specific force is used to calculate the GA estimate in much the same way as is done when processing data in platform gravimetry. In contrast, the indirect method combines all available measurements generated by a strapdown inertial navigation system (SINS) built on the basis of IMU and GNSS, in a Kalman filter wherein the GA estimate is generated, which is included in the SINS error model along with IMU errors [7]. Note that a similar technique used to estimate horizontal components of the gravity disturbance vector is called an inertial-geodesic method [19].

The paper compares direct and indirect methods of GA determination for solving the problem of strapdown scalar marine gravimetry. Sections 2 and 3 present the problem statement and discuss the features of strapdown scalar gravimetry in comparison with platform gravimetry. The features of the two methods for GA determination are also analyzed here. Section 4 discusses the results obtained during the comparison of the methods based on simulation and bench test data. We do not use GNSS high-accuracy altitude measurements in this research, because, unlike aerogravimetry, in marine gravimetry, the problem of GA estimation at sea can be solved with the required accuracy without this information [9, 20, 21].

II. THE DIRECT METHOD OF GA DETERMINATION

Projections of the gravity disturbance vector on the axes of the local horizontal frame ENU $\delta\mathrm{g}^n = [\delta g^E \quad \delta g^N \quad \delta g^H]^T$ can be represented as [18]:

$$\delta\mathrm{g}^n = \ddot{r}^n - C_b^n f^b + \delta g_{eot}^n - \gamma^n - C_b^n \delta f^b, \quad (1)$$

where $\ddot{r}^n$ are projections of the carrier kinematic acceleration on the axes of the local horizontal coordinates ENU, usually obtained from GNSS; f^b – specific force measured by the IMU accelerometers in the projection on the carrier principal axes XYZ; C_b^n – orientation matrix characterizing the transition from the carrier principal axes XYZ to local horizontal coordinates ENU; $\delta\mathrm{g}_{eot}^n$ – Eotvos correction; γ^n – projection of the normal gravity vector on the ENU; δf^b – accelerometer errors. Equation (1) in vector form allows us to determine both the horizontal and vertical components of the gravity disturbance vector. In the scalar gravimetry problem considered here, only the vertical component GA ($\delta\mathrm{g}^H$) is determined.

In the direct method, the orientation matrix C_b^n is preliminarily determined, which is necessary to transform accelerometer readings to the local horizontal axes ENU. For this purpose, an SINS algorithm is implemented with correction from GNSS velocity and position measurements, based on the use of a recursive extended Kalman filter (EKF). After this, the problem is reduced to standard

processing of scalar GA measurements using smoothing algorithms used in platform marine or aerogravimetry [4, 8, 10–11, 13, 15, 18, 21]. When solving the problem of marine strapdown gravimetry, similar to platform gravimetry, GNSS high-accuracy altitude measurements are not required to succesfully estimate GA, as shown below. It is important to take into consideration the fact that specific force measurements f^b also contain errors of the IMU accelerometers δf^b. If the white noise component of the error can be successfully suppressed using filtering and smoothing algorithms, then the biases and scale factor errors of accelerometers are difficult to separate from the GA and, as a result, they must be carefully calibrated before and after the survey [5, 14, 16–18]. The direct method in this work is implemented using the SINS-GNSS filtering algorithm described below to calculate the orientation matrix C_b^n. The

GA estimate was calculated by a smoother that was used for processing of the platform gravimetry data [9, 21].

The possibility of using the experience with integrated algorithms for SINS and GNSS as well as filtering and smoothing algorithms for GA based on data from platform gravimeters makes this method widespread [1, 4, 6, 8–9, 11, 15].

III. THE INDIRECT METHOD OF GA DETERMINATION

The indirect method [4], also called the inertial navigation method [5], estimates GA when solving the problem of joint processing of data from the SINS algorithm and the GNSS. Integration of SINS and GNSS data, as in the previous method, is carried out using the recursive EKF [23, 24]. Figure 1 illustrates a block diagram of the indirect method for GA estimation.

Fig. 1. Block diagram of the indirect method for GA estimation.

The SINS algorithm based on angular velocity ω_{ib}^b and specific force f^b measurements calculates the orientation matrix C_b^n characterizing the transition from carrier principal frame XYZ (denoted by index b) to the local horizontal frame ENU (denoted by index n); attitude angles K, ψ, θ ; linear velocities $V_{eb}^{n\ INS}$ and geographic coordinates $p = \left[\varphi^{INS}, \lambda^{INS}, h^{INS} \right]^T$. Navigation parameters (geographic coordinates and linear velocity components), calculated by the SINS and received from GNSS, enter the EKF to generate differential velocity and position measurements. The EKF state vector includes attitude errors $\tilde{\alpha}, \tilde{\beta}, \tilde{\gamma}$, linear velocity errors $\delta \tilde{V}_{eb}^n$, coordinate errors $\delta \tilde{\varphi}, \delta \tilde{\lambda}, \delta \tilde{h}$, and accelerometer biases $\delta f_X, \delta f_Y, \delta f_Z$. The formation of Coriolis acceleration here corresponds to the calculation of the Eötvos correction: $\delta g_{eot}^n = (\Omega_{en}^n + 2\Omega_{ie}^n)V_{eb}^n$. Estimates of the system state

vector are fed back to the corresponding blocks of the SINS algorithm.

This is exactly the EKF used in the direct method.

When the indirect method is used, the components that describe the GA model are added to the state vector. In this work, we used the widespread Jordan model of GA [25].

To obtain a better GA estimate, also due to reduced phase delay, we need to use a smoothing algorithm. In this work, a smoother is implemented for this purpose by weighted averaging of state vector estimates obtained when solving the filtering problem in forward and reverse time [26, 27]. To obtain estimates in reverse time, it is necessary to invert the sign of measurements of the inertial sensors, linear velocity, Earth's angular rate, and gravity. Though the EKF structure (Fig. 1) is remains the same, the GA estimate will also have an inverted sign. The proposed scheme differs from the smoother in [13], however, in general, they are similar in accuracy [27].

IV. COMPARISON RESULTS

To compare direct and indirect methods, a modeling program for complex processing of IMU and GNSS data was developed using recurrent filtering algorithms based on the extended Kalman filter. At the first stage, comparison of methods was carried out through simulation. The results were similar for GA estimation in both cases without high-accuracy GNSS data on altitude when simulating the motion of a marine vessel. The second stage was a preliminary check of the operation of the direct and indirect methods on a vertical displacement stand. The IMU used in the experiment included a triaxial accelerometer and a triad of navigation-grade fiber-optic gyros. GNSS data were simulated, for which errors corresponding to code measurements were added to the reference coordinates. The coordinates of the vertical displacement stand were taken as reference latitude and longitude, and the altitude was recorded with a laser distance meter. The IMU was installed on the stand platform, and data recording started. The stand was kept motionless for 0.5 –1.5 hours. After that, harmonic vertical displacements started with a displacement amplitude of 2 m and various periods inherent in marine objects. The vertical pitching lasted for 1.5 – 2 hours. Then, the stand returned to its initial position and the IMU data continued to be recorded in a stationary position. At the same time, it is obvious that the GA will be constant in test. Therefore, to compare methods, we used an estimate of accelerometer bias projected on the vertical axis, which has a similar character in the time domain as the GA on a moving vessel, and is practically inseparable from the useful signal. The results of estimating the vertical component of the accelerometer bias by using the direct and indirect method for movements with a heave period of 18 s and 51 s are presented in Fig. 2.

Fig. 2. Estimates for the vertical component of the accelerometer bias by using the direct method (smoothing) and indirect method (filtering and smoothing) a) heave period 18 s; b) heave period 51 s.

Disturbances in the estimates obtained by the direct method during the beginning and end of the heaving are due to the choice of a stationary smoother. Replacing the stationary filter with a recursive Kalman filter eliminates the disturbances caused by the change of the mode, as shown in [21]. It can be seen that the estimates obtained by the indirect method in the filtering mode have a phase delay. When solving the smoothing problem by the indirect method, GA estimates do not have a phase delay, and the oscillatory component is smaller than in the filtering or when using the direct method, which is more consistent with the behavior of the accelerometer bias.

The tests on the stand have made it possible to identify the accelerometer bias component, which depends on the parameters of vertical motion and has a value greater than that of platform gravimeters with a quartz damped sensitive element [22]. The higher is the heave frequency, the greater is the displacement. The discovered dependence of the bias should be further taken into account in the error model of the accelerometers used in the IMU or eliminated technologically.

V. CONCLUSIONS

A comparison of direct and indirect methods in solving the problem of GA determination at sea using a strapdown scalar gravimeter has been carried out.

The direct method is implemented by joint processing of SINS and GNSS data using recursive filtering algorithms based on the recusive EKF, which allows obtaining orientation data with the required accuracy and vertical specific force signal. The GA estimate was calculated by processing the specific force signal using standard smoothing filters applied in marine platform gravimetry without involving high-accuracy GNSS altitude data.

The indirect method was implemented by joint processing of SINS and GNSS data using smoothing algorithms and was based on the same data from the GNSS as the direct method, i.e. without involving high-accuracy altitude data. In this case, GA were included in the estimated state vector.

Computer simulation and testing of direct and indirect processing methods were carried out based on the bench test data, which did not show significant advantages of any of the methods under consideration for marine applications. At the same time, based on the results of the bench tests, variability in the vertical accelerometer bias caused by vertical accelerations of the carrier was revealed, which is planned to be described and taken into account in the future in the IMU error model.

REFERENCES

[1] Jekeli, C., Garcia, R. GPS phase accelerations for moving-base vector gravimetry. In: Journal of Geodesy 71: 1997 C. 630–639. DOI: 10.1007/s001900050130.

[2] Wei, M., Schwarz, K.P. (1998): Flight test results from a strapdown airborne gravity system. In: Journal of Geodesy: 323–332. DOI: 10.1007/s001900050171.

[3] Jekeli, C., Kwon, J.H., Results of Airborne Vector (3-D) gravimetry, *Geoph. Res. Lett.*, 1999, 26(23):3533–3536

[4] Kwon, J.H., Jekeli, C., A new approach for airborne vector gravimetry using GPS/INS, J. Geod., 2001,74: 690–700, doi: 10.1007/s001900000130.

[5] Becker, D., Nielsen, J.E., Ayres-Sampaio, D. et al. Drift reduction in strapdown airborne gravimetry using a simple thermal correction, J. Geod., 2015, vol. 89, pp. 1133–1144, https://doi.org/10.1007/ s00190-015-0839-8.

[6] Jekeli, C., Theoretical Fundamentals of Airborne Gradiometry, Airborne Gravity for Geodesy Summer School, 23–27 May 2016.

[7] Becker, D., Advanced Calibration Methods for Strapdown Airborne Gravimetry, Technische Universität Darmstadt, Ph.D. Thesis, 2016.

[8] Cai, S., Tie, J., Zhang, K., Cao, J., and Wu, M., / Marine gravimetry using the strapdown gravimeter SGA-WZ, // Marine Geophysical Research, vol. 38, no. 4, pp. 325–340, 2017. doi:10.1007/s11001-017-9312-9.

[9] V.G. Peshekhonov, O.A. Stepanov, L.I. Avgustov et al., Methods and Technologies for Measuring the Earth's Gravity Field Parameters, Springer, 2022, https://doi.org/10.1007/978-3-031-11158-7.

[10] C. A. Lin, K. W. Chiang, C. Y. Kuo Integration of INS and GNSS for gravimetric application with uas // Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLII-1, 263–268, https://doi.org/10.5194/isprs-archives-XLII-1-263-2018, 2018.

[11] Wang W, Gao J, Li D, Zhang T, Luo X, Wang J. / Measurements and Accuracy Evaluation of a Strapdown Marine Gravimeter Based on Inertial Navigation. // Sensors. 2018; 18(11):3902. https://doi.org/10.3390/s18113902

[12] Jensen, T.E., Forsberg, R., Helicopter Test of a Strapdown Airborne Gravimetry System, Sensors, 2018, 18, 3121, doi:10.3390/s18093121.

[13] Johann F., Becker D., Becker M., Forsberg R. and Kadir M. The Direct Method in Strapdown Airborne Gravimetry - a Review. // ZFV 5, 2019. C. 323-333 DOI:10.12902/zfv-0263-2019

[14] Simav, M., Becker, D., Yildiz, H., Hoss, M., Impact of temperature stabilization on the strapdown airborne gravimetry: a case study in Central Turkey, Journal of Geodesy, 2020, 94, 41, https://doi.org/10.1007/ s00190-020-01369-5.

[15] Vyazmin V., Golovan A., Bolotin Y. New strapdown airborne gravimetry algorithms: testing with real flight data // 28th Saint Petersburg International Conference on Integrated Navigation Systems, ICINS 2021. 28. 2021. C. 9470826.

[16] Peshekhonov, V.G., Stepanov, O.A., Rozentsvein, V.G. et al. State-of-the-Art Strapdown Airborne Gravimeters: Analysis of the Development. Gyroscopy Navig. 13, 189–209 (2022). https://doi.org/10.1134/S2075108722040101

[17] Vyazmin, V.S., Golovan, A.A. & Govorov, A.D. Initial and Final Alignment of a Strapdown Airborne Gravimeter and Accelerometer Bias Determination. Gyroscopy Navig. 14, 48–55 (2023). https://doi.org/10.1134/S2075108723010091.

[18] A. A., Golovan, A.A., Vyazmin, V.S. Methodology of Airborne Gravimetry Surveying and Strapdown Gravimeter Data Processing. Gyroscopy Navig. 14, 36–47 (2023). https://doi.org/10.1134/S2075108723010029

[19] A.V. Motorin, D.A. Koshaev, O.A. Stepanov, O.M. Yashnilova. Identifikatsiya modeli i sglazhivaniye gorizontal'nykh komponent vozmushcheniya sily tyazhesti dlya resheniya zadachi vektornoy gravimetrii (Model identification and smoothing of horizontal components of gravity disturbance for solving a vector gravimetry problem) (in Russian)// Materials: XXXII konferentsii pamyati vydayushchegosya konstruktora giroskopicheskikh priborov N. N. Ostryakova., 2020. pp. 149-152.

[20] O. A. Stepanov, D. A. Koshaev, A. V. Motorin, A. A. Krasnov, and A. V. Sokolov, "Algorithms for Integrated Processing of Marine Gravimeter Data and GNSS Measurements," IFAC-PapersOnLine, vol. 53, no. 2, pp. 500–505, 2020.

[21] A.V. Sokolov, O.A. Stepanov, A.V. Motorin, A.A. Krasnov, Comparison of Wiener and Kalman approaches when solving the problem of processing the results of marine gravimetric surveys // Gyroscopy Navig. 2024, vol. 15, №1.

[22] Sokolov, A.V., Krasnov, A.A. & Zheleznyak, L.K. Improving the Accuracy of Marine Gravimeters. Gyroscopy Navig. 10, 155–160 (2019). https://doi.org/10.1134/S2075108719030088

[23] Vavilova N.B., Golovan A.A., Parusnikov N.A. Matematicheskiye osnovy inertsial'nykh navigatsionnykh sistem. (in Russian) M.: Izd-vo Moskovskogo universiteta, 2020.

[24] Yemel'yantsev G.I., Stepanov A.P. Integrirovannye inertsial'no-sputnikovyye sistemy oriyentatsii i navigatsii / pod obshchey red. akad. RAN V.G. Peshekhonova St. Petersburg, Russia: Concern CSRI Elektropribor, 2016.

[25] Jordan, S. K. (1972). Self-consistent Statistical Models for Gravity Anomaly and Undulation of the Geoid. Journal of Geophysical Research, 77/(20), 2156–2202.

[26] S. Sarkka, Bayesian Filtering and Smoothing. Cambridge University Press, 2013. doi: 10.1017/CBO9781139344203.

[27] O.A. Stepanov, Osnovy teorii otsenivaniya s prilozheniyami k zadacham obrabotki navigatsionnoy informatsii. Chast' 2. Vvedeniye v teoriyu filtratsii, 3rd ed. (in Russian), St. Petersburg, Russia: Concern CSRI Elektropribor, 2017.

Study of Accuracy Characteristics of Aircraft Navigation using Micro-Relief with Artificial and Natural Object Composition

Sazonova T.V.
JSC "Ramenskoye Design Company" Ramenskoye, Russia
tatsazonova@mail.ru

Shelagurova M.S.
JSC "Ramenskoye Design Company" Ramenskoye, Russia
marinashell@mail.ru

Korneva E.G.
JSC "Ramenskoye Design Company" Ramenskoye, Russia
ekorneva@rpkb.ru

Abstract—**The article is devoted to the description of the mathematical modeling results of a correlation-extreme navigation system with the use of real cartographic micro-relief data regarding the urban development area and the area with natural object composition.**

Keywords – micro-relief, a laser rangefinder, correlation-extreme navigation system.

I. INTRODUCTION

Currently, the urgent task is to provide autonomous navigation for aircraft (AC), for which correlation-extreme navigation systems (CENS) by physical fields of the Earth can be used. Practical realization has been implemented Масло помаслили….only for the correlation-extreme navigation systems that are applied in relation to relief of the earth's surface, however, these systems are not operable when aircraft is flying over uninformative terrain and their accuracy is estimated at a level of 100-200 m. Due to the development of production technologies for reception of micro-relief data with high resolution (terrain topography with consideration of heights regarding object composition of artificial and natural origin) it has become possible to use the data for high-precision correction of aircraft coordinates. A laser rangefinder is used as a micro-relief sensor, which provides a high-precision measurement of physical height. The remote sensing data of the Earth are used to generate maps with micro-relief. The results of aerial or aerospace photography processed by the stereophotogrammetry method [1] or the results of air-laser scanning can be applied.

II. THE PURPOSE OF THE STUDIES

The purpose of the studies was to assess the accuracy characteristics of CENS by micro-relief with artificial and natural object composition via mathematical modeling.

III. MATHEMATICAL MODELING OF CENS BY MICRO-RELIEF

During mathematical modeling we used real cartographic data of the urban development area of Kolomna and the area with natural object composition near Nizhny Tagil, which were prepared based on the cosmophotographic results processed by stereophotogrammetry. Information of micro-relief was prepared in Gauss-Kruger projection and stored in MTW format. Recording discrete is 1 m, mapping error in plan is 1-2 m, as per height - 2 m.

Micro-relief within the areas of Kolomna (Fig, 1) had a size of 5100x5900 sq. m

Micro-relief near Nizhniy Tagil (Fig, 2) had a size of 9650x10350 sq. m.

Evaluation of the informative value of the navigation field has shown that the root mean square deviation of micro-relief for these areas is the same - about 6 m, but correlation radius for area of Kolomna (38 m) is almost half as small as compared to area near Nizhny Tagil (65 m).

Fig. 1. Micro-relief within the areas of Kolomna

Fig. 2. Micro-relief near Nizhniy Tagil

The laser rangefinder (designation: LRF03C01S, supplier: "STC "LEMT" of the BelOMO") [2] was considered as a micro-relief sensor, which measured the

underlying surface profile with an error of 1 m at altitudes up to 3000 m in simple weather conditions (Fig. 3).

Fig. 3. Fig.3. The laser rangefinder LRF03C01S

At mathematical modeling, aircraft flight at 500 m altitude and at speed of 65 m/s was simulated within the conditions of inoperability of satellite navigation system. In this case, root mean square errors (RMSE) regarding determination of the navigation system parameters took the following values:

- RMSE for determination of orientation angles (heading, roll, pitch) - 1º;

- RMSE for determination of ground speed - 3 m/s;

- RMSE for determination of coordinates (a priori error) regarding 100 m and 300 m

The following parameters have been evaluated during the mathematical modelling:

- number of performed corrections Nc;

- average length of correction area in relation to micro-relief Lc;

- CENS accuracy by micro-relief σc (circular RMSE);

- average calculation time for one tact of correction tc.

The CENS micro-relief algorithm is based on the function that calculates a square of difference between the reference and the current value, while introducing the "attenuation" regarding weight of the functions calculated in the previous steps, which is very relevant at unsteady micro-relief. In the course of the algorithm's operation, selection regarding the gross sensor errors is carried out based on the knowledge of minimum and maximum reference micro-relief

values within the confidence square and errors of the laser rangefinder [3].

At implementation of mathematical modeling, maximum of the confidence square size was 900x900, the confidence square discretes (readout discretes, hereinafter - di) took the following values:

di = 2 m, 3 m, 4 m, 5 m, 6 m.

Table 1 summarizes the main results for CENS based on micro-relief for the areas of Kolomna (K) and near Nizhniy Tagil (NT).

T A B L E 1 . The CENS results based on micro-relief.

Reg	σa, m	di, m	Nc	Lc, m	σc, m	tc, mc
K	100	2	228	331	2,6	0,37
		3	225	270	2,8	0,25
		4	225	252	3,1	0,22
		5	148	259	3,8	0,17
		6	115	269	4,2	0,14
	300	3	134	299	4,6	1,48
		4	115	274	4,2	0,91
		5	86	274	4,2	0,62
		6	50	277	4,3	0,51
NT	100	2	292	582	3,5	0,24
		3	328	432	3,9	0,17
		4	422	358	4,1	0,16
		5	418	313	4,8	0,16
		6	420	274	5,5	0,16
	300	3	294	452	4,3	0,95
		4	295	382	4,6	0,74
		5	275	341	5,5	0,52
		6	249	297	5,0	0.48

IV. CONCLUSION

Thus, mathematical modeling has shown that CENS based on micro-relief can be implemented in real time with the characteristics that are comparable to differential mode accuracy of satellite navigation systems.

For practical implementation of the CENS by micro-relief it is necessary to solve the problem of organizing and promptly updating a micro-relief database created by processing high-resolution satellite images using stereophotogrammetry.

REFERENCES

[1] E.V. Kozin, A.G. Karmanov, N.A. Karmaonova. «Photogrammetry». University ITMO. Saint Petersburg. 2019.

[2] The laser rangefinder LRF03C01S Technical specification.

[3] Monograph: T.V. Sazonova, M.S. Shelagurova «Geoinformation in aircraft equipment set ». Moscow. 2018.

Gravity Matching Method Based on Optimized Particle Filtering

Tijing Cai
School of Instrument Science
and Engineering
Southeast University
Nanjing, China
caitij@seu.edu.cn

Shuaipeng Gao
School of Instrument Science
and Engineering
Southeast University
Nanjing, China
gsp0803@seu.edu.cn

Ying Liu
School of Instrument Science
and Engineering
Southeast University
Nanjing, China
liuyingseu@seu.edu.cn

Zhiqian Lu
School of Instrument Science
and Engineering
Southeast University
Nanjing, China
luzhi_qian1999@163.com

Abstract—**When employing conventional particle filtering method for gravity aided navigation, the phenomena of particle swarm impoverishment and inadequately precise importance distribution functions manifest. In this paper, a gravity matching method based on optimal importance differential particle filtering is proposed to solve this problem. In the particle update process, this method follows the principle of local linearization, introduces previous estimation and gravity anomaly measurement, and employs optimal importance sampling to obtain the particle distribution. This approach fundamentally reduces the degradation of the particle swarm and maintains its diversity. Moreover, this method utilizes the differencing of particle information at different time steps to mitigate the impact of zero drift in gravity meter, which improves the positioning accuracy of gravity aided navigation. Through post-processing simulation experiments with ocean test data, the results indicated that the proposed gravity matching method has better positioning accuracy and stability than the traditional particle filtering method, and has reduced the frequency of resampling phenomenon.**

Keywords—Bayesian Estimation, Particle Filtering, Optimal Importance Sampling

I. INTRODUCTION

The navigation of underwater vehicles often relies primarily on inertial navigation system (INS) due to the requirements for concealment, independence, and autonomy [1]. However, INS is generally unsuitable for prolonged and long-distance underwater navigation due to the accumulation of errors [2]. In recent years, gravity aided navigation that use gravity anomaly measurements and pre-stored gravity map to correct the INS position has received widespread attention from scholars both domestically and internationally [3]. In this case, the particle filtering (PF) method is a commonly used gravity aided navigation algorithm. PF method is a Bayesian filtering technique based on Monte Carlo method, which is advantageous in dealing with nonlinear and non-Gaussian state estimation problems [4].

The PF algorithm effectively addresses nonlinear filtering problems but exhibits inherent limitations, particularly

concerning particle depletion and the construction of importance distribution functions [5]. Particle depletion refers to the loss of diversity in the particle swarm, resulting in estimation deviations. To tackle this, the resampling technique is typically employed [6], with common methods including polynomial resampling [7], systematic resampling [8], and residual resampling [9]. Importance distribution functions relate to how particles and their weights are updated iteratively. Literature [10] introduced particle jitter in the state transition process and additional process noise in the particle update process, aiming to improve particle diversity. Additionally, literature [11] proposed a Gaussian sum PF without a resampling process, with playback simulation experiments indicating its effectiveness in approximating the Bayesian filtering model. However, these methods overlook particle distribution patterns and error propagation. Gordon [12], [13] proposed local linearization of the model and used optimal importance sampling to distribute particles.

This paper proposes an optimal importance and differential particle filter (ODPF) algorithm for gravity aided navigation. This algorithm introduces an optimal importance sampling method which can fundamentally reduce the number of particle degradations during the filtering process. Simultaneously, addressing the impact of measurement accuracy caused by the zero drift of gravity meter, the algorithm presents particle position differencing method, which corrects the current particles based on the estimation of the previous moment. By conducting post-processing simulation experiments from ocean test data, it is demonstrated that the ODPF method is superior to the typical PF method in the gravity aided navigation.

II. METHODOLOGY

A. Traditional Particle Filter Method

Due to the high number of integration operations involved in Bayesian estimation and the non-normal distribution of observations, obtaining integration values for posterior probabilities is challenging. Therefore, the Monte Carlo method is commonly employed to obtain estimates. Its recursive iterative versions are also known as particle

filtering or sequential Monte Carlo methods. As the sample size increases, the values obtained through sequential Monte Carlo methods converge towards optimal Bayesian estimation. To provide a detailed description of the algorithm, Let $\mathbf{x}_t$ denote the INS position at time t, y_t indicate the gravity measurement at time t, $\mathbf{u}_t$ represent the relative displacement of the INS position between adjacent instants, $\mathbf{v}_t$ be the process noise in the state update process of particle swarm between adjacent instants, $h(\mathbf{x}_t)$ means the nonlinear process that derives gravity map values of the position vector $\mathbf{x}_t$, and λ_t be the gravity measurement noise. Assuming that the $\mathbf{v}_t$ and λ_t are mutually independent and follow Gaussian distributions with known variances. The state equation and observation equation for the nonlinear stochastic system are respectively represented as:

$$
\begin{aligned}
\mathbf{x}_{t+1} &= \mathbf{x}_t + \mathbf{u}_t + \mathbf{v}_t \\
y_t &= h(\mathbf{x}_t) + \lambda_t
\end{aligned}
\tag{1}
$$

The gravity matching algorithm based on traditional PF method includes the following steps:

1. Initialization of particle swarm: a set of discrete points with Gaussian distribution is generated around the initial INS position. Each point represents a possible true position of the vehicle at current moment and has the same weight. The weighted average of the discrete particles is the position estimate at the initial time.

2. Particle update and measurement acquisition: following the state equation in Eq. (1), update the particles based on the $\mathbf{u}_t$. The y_t is obtained from the gravity meter.

3. Weight update: utilize the $h(\mathbf{x}_t)$ and the weights from the last instant to update the weights at this instant, and normalize the weights.

4. Position prediction: perform a weighted average of particle positions and weights to obtain the predicted position of the vehicle.

In the actual navigation process, constant iteration of steps 2-4 can continuously provide the trajectory of the vehicle.

B. Resampling

In the gravity matching algorithm based on the PF method, the weights assigned to individual particles can sometimes lead to a phenomenon known as particle impoverishment during the iterative process. Particle impoverishment occurs when certain particles have weights that become exceedingly small, nearly negligible, while a few particles retain significant weights. This situation arises when only a minority of particles significantly contribute to the final estimation, resulting in a loss of diversity within the entire particle swarm. Consequently, the filter's performance may be compromised, leading to inaccurate estimation results.

In such situations, the algorithm incorporates a resampling step to address the particle impoverishment issue. The specific approach is as follows: when the effective number of particles in the swarm meets certain conditions, indicating particle depletion, resampling is triggered.

$$
N_{eff} = \frac{1}{\sum_{i=1}^{N}(\omega_t^{(i)})^2} < N_{thres}
\tag{2}
$$

where N_{thres} represents the threshold value for effective particles, typically set as $2N/3$.

When resampling is required, the weights of all particles can be uniformly redistributed. This approach is effective in preventing significant estimation biases caused by particle depletion.

C. Resampling

Equation (1) highlights that once the initial position distribution of the particles is determined, subsequent updates during the iterative process do not alter the distribution patterns. However, the weight distribution of the particle swarm changes as measurements are updated. If there's a significant deviation in measurements at a specific moment, it will also affect the weight distribution of the particle swarm, potentially leading to substantial discrepancies between estimated and true positions, indicating algorithm failure.

Moreover, when the particle swarm experiences a significant offset due to INS divergence or gravity anomaly measurement noise, subsequent updates, being related to the previous state, may introduce erroneous offsets, rendering all subsequent estimations unreliable. In such cases, the filter becomes invalidated, regardless of whether resampling is performed.

During the particle state update process, the estimation results from the previous moment and the gravity anomaly measurements from the current moment are incorporated. Utilizing a local linearization method, a Gaussian distribution is employed to generate diverse particles. This approach effectively mitigates the impact of erroneous measurements and reduces the frequency of resampling.

The mean and variance of the particles at moment t are calculated as follows:

$$
\begin{aligned}
\boldsymbol{\Sigma}_t^{(i)} &= \mathbf{Q} - \mathbf{Q}\mathbf{h}_t^{(i)}(\mathbf{h}_t^{(i)^T}\mathbf{Q}\mathbf{h}_t^{(i)} + \lambda)^{-1}\mathbf{h}_t^{(i)^T}\mathbf{Q} \\
\boldsymbol{\mu}_t^{(i)} &= \boldsymbol{\Sigma}_t^{(i)}(\mathbf{Q}^{-1}(\mathbf{x}_{t-1}^{(i)} + \mathbf{u}_{t-1}) + \\
&\mathbf{h}_t^{(i)}(y_t - h(\mathbf{x}_{t-1}^{(i)}) + \mathbf{h}_t^{(i)^T}(\mathbf{x}_{t-1}^{(i)} + \mathbf{u}_{t-1}))\lambda^{-1})
\end{aligned}
\tag{3}
$$

where $\boldsymbol{\Sigma}_t^{(i)}$ represents the mean of the Gaussian distribution for the i particle at time t, $\boldsymbol{\mu}_t^{(i)}$ is the variance of the Gaussian distribution, $\mathbf{h}_t^{(i)} = \nabla h(\mathbf{x}_t^{(i)})$ means the gravity gradient value at location $\mathbf{x}_t^i$, and $\mathbf{Q}$ is the covariance matrix of the state noise.

To update the particle weights, the position of each particle from the previous moment is also considered. The updating process is as follows:

$$
\omega_t^{(i)} = \omega_{t-1}^{(i)} \frac{N(y_t - h(\mathbf{x}_t^{(i)}), \lambda)N(\mathbf{x}_t^{(i)} - (\mathbf{x}_{t-1}^{(i)} + \mathbf{u}_{t-1}), \mathbf{Q})}{N(\mathbf{x}_t^{(i)} - \boldsymbol{\mu}_t^{(i)}, \boldsymbol{\Sigma}_t^{(i)})}
\tag{4}
$$

where $N(a,b)$ represents the likelihood function, calculated as follows:

$$N(a,b) = \frac{1}{(2\pi)^{n/2}\sqrt{\det b}}\exp(-\frac{1}{2}a^T b^{-1}a) \qquad (5)$$

where n represents the dimension of matrix b.

D. Differential Particle Filter Method based on Optimal Importance Sampling

Based on the PF algorithm, this paper proposes the Optimal Importance and Differential Particle Filter (ODPF) algorithm for gravity-aided navigation. The detailed steps are as follows:

1. Initialization of particle swarm: use the Gaussian distribution to distribute the particles at the initial position $\mathbf{x}_0$ of the INS. Define the initial error covariance matrix as $\mathbf{P}_0$. The particle set is $\{\mathbf{x}_0^i\} \sim \mathcal{N}(\mathbf{x}_0, \mathbf{P}_0), i = 1, 2, ..., N$, with initial weights for each particle being $1/N$. The initial position estimate $\hat{\mathbf{x}}_0 = \sum_{i=1}^{N} w^{(i)}\mathbf{x}_0^{(i)}$.

2. Particle update and measurement acquisition: utilizing the relative displacement of INS between the last and current instants, the method outlined in Eq. (3) is employed to compute the mean and variance of particles. This process establishes potential positions for each particle. Following this, the gradient sequence on the gravity map is calculated based on the positions of individual particles.

3. Weight Update: update weights according to Eq. (4) and normalize them.

4. Position prediction: perform a weighted average of particle positions $\hat{\mathbf{x}}_t = \sum_{i=1}^{N} w^{(i)}\mathbf{x}_t^{(i)}$ based on their weights to obtain the predicted navigation position.

In the actual navigation process, constant iteration of steps 2-4 can continuously provide the trajectory of the vehicle.

The ODPF algorithm's procedure indicates that the particle distribution is no longer uniformly updated. Instead, each particle is individually updated with its own estimated mean and variance, ensuring diversity while incorporating information from the previous moment. Regarding weight updating, both the particle information from the previous moment and the directly updated particle values are utilized, effectively mitigating the impact of measurement errors.

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

To mitigate the impact of the zero drift of the gravity meter, the ODPF employs a differential method. Throughout the optimal importance sampling process, both the particle update and weight update stages utilize estimation information from the previous moment. The particle update and weight update processes are outlined as follows:

$$\omega_t^{(i)} = \frac{\omega_{t-1}^{(i)} N(y_t - h(\mathbf{x}_t^{(i)}) - (y_{t-k} - h(\hat{\mathbf{x}}_{t-k}))), 2\lambda)}{N(\mathbf{x}_t^{(i)} - \boldsymbol{\mu}_t^{(i)}, \boldsymbol{\Sigma}_t^{(i)})} \bullet$$

$$N(\mathbf{x}_t^{(i)} - (\mathbf{x}_{t-1}^{(i)} + \mathbf{u}_{t-1}), \mathbf{Q}) \qquad (6)$$

where $\mathbf{h}_t^{(i)\prime}$ represents the difference in gravity gradient values from the previous k instants.

III. EXPERIMENT

The simulation experiment data were obtained from marine gravity anomaly measurement tests. The parameters of the test equipment are as follows: The gyro zero bias stability of the INS is 0.003 °/h, with a random walk of 0.0005 °/h. The accelerometer bias is less than 5 mg, and the output frequency is 200 Hz. The marine gravity meter continuously measures gravity with an accuracy better than 1.5 mGal, and the gravity data output frequency is 1 Hz. The resolution of the gravity anomaly map is $1'\times1'$. GPS receivers record the real-time position of the carrier for precision evaluation of navigation.

It was experimentally determined that the gravity matching algorithm with 1000 particles effectively balanced real-time performance and reliability. The parameter settings for the gravity matching method based on traditional PF method and the ODPF method remained consistent. Trajectories generated by the PF and ODPF methods are depicted in Fig. 1, while Fig. 2 illustrates the error variation curve of navigation distance over time. The average position error for the PF method is 679.37 m, whereas for the ODPF method, it reduces to 354.97 m.

Fig. 1. Navigational trajectory.

Fig. 2 clearly illustrates that the positioning accuracy of the ODPF method surpasses that of the PF method. Particularly in the later stages of navigation, the matching results generated by the ODPF method closely approximate the real position compared to those of the PF method. The oscillations observed in the position error curve of the PF method in Fig. 2 can be attributed to the frequent resampling events during the matching process. In contrast, the ODPF algorithm experiences fewer resampling occurrences and maintains superior particle diversity throughout the navigation process.

Fig. 2. Positioning errors

Given that both the PF method and the ODPF method incorporate Gaussian distributions in their particle updating processes, inherent randomness exists within the algorithms themselves. To assess their reliability and stability, several experiments were conducted. After performing 10 repeated experiments under identical conditions, the results consistently demonstrated that the average position error of the ODPF algorithm remained below 500 m, indicating remarkable stability.

IV. CONCLUSION

In this paper, the traditional PF method is enhanced and evolved into the ODPF method. By incorporating optimal importance sampling, the ODPF method effectively addresses the issue of particle degradation. Additionally, the differential method is employed to mitigate the impact of gravity measurement drift. Experimental studies conducted using real trajectory data highlight the high accuracy and great stability of the ODPF method. These findings underscore its efficacy in leveraging gravity information to facilitate navigation for long-duration underwater vehicles.

REFERENCES

[1] B. Wang, T. Li, and X. Li, "Review of underwater inertial/gravity gradient matching navigation," Tactical Missile Technology, vol. 4, pp. 1–12, 2023, doi: 10.16358/j.issn.1009-1300.20230020.

[2] B. Ji, S. Bian, and J. Jin, Gravity Gradient Underwater Detection and Navigation. Beijing: Science Press, 2016.

[3] B. Wang, J. Zhu, Z. Deng, and M. Fu, "A Characteristic Parameter Matching Algorithm for Gravity-Aided Navigation of Underwater Vehicles," IEEE Transactions on industrial electronics, vol. 66, no. 2, pp. 1203–1212, Feb. 2019, doi: 10.1109/TIE.2018.2831171.

[4] F. Liu, X. Wei, D. Qian, and F. Li, "The application of optimization particle filter in gravity aided positioning," Journal of Harbin Institute of Technology, vol. 44, no. 12, pp. 145–148, 2012.

[5] M. Ouyang and Y. Yang, "Underwater gravity matching navigation simulation and experiment based on the improved particle filtering algorithm," Journal of Harbin Engineering University, vol. 43, no. 10, pp. 1514–1521, 2022.

[6] J. Li, X. Liu, C. Lu, and Y. Zuo, "Improvement of resampling algorithm of particle filter," Journal of Jilin University. Engineering and Technology Edition, vol. 45, no. 6, pp. 2069–2074, 2015.

[7] N. Gordon, D. Salmond, and A. Smith, "Novel-approach to nonlinear non-Gausian Bayesian state estimation," IEE Proceedings-F radar and signal processing, vol. 140, no. 2, pp. 107–113, Apr. 1993, doi: 10.1049/ip-f-2.1993.0015.

[8] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," IEEE Transactions on signal processing, vol. 50, no. 2, pp. 174–188, Feb. 2002, doi: 10.1109/78.978374.

[9] J. S. Liu and R. Chen, "Sequential Monte Carlo methods for dynamic systems," Journal of the american statistical association, vol. 93, no. 443, pp. 1032–1044, Sep. 1998, doi: 10.2307/2669847.

[10] R. Chen and C. Che, "Underwater Terrain Aided Navigation Method Based on Improved Particle Filter Algorithm," Equipment Enviromental Engineering, vol. 19, no. 6, pp. 91–96, 2022.

[11] Y. Han, P. Chen, and P. Shen, "Underwater Terrain-Aided Navigation Method Based on Gaussian Sum Particle Filtering for AUVs," Unmanned Systems Technology, vol. 3, no. 1, pp. 48–54, 2020, doi: 10.19942/j.issn.2096-5915.2020.01.006.

[12] N. Gordon, Sequential Monte Carlo Methods in Practice. New York: Springer Science+Business Media New York, 2001.

[13] V. G. Peshekhonov and O. A. Stepanov, Methods and Technologies for Measuring the Earth's Gravity Field Parameters. Springer, 2017.

Depth-Aided Vessel Navigation with the Use of a Neural Network

V. Deryabin

Navigation Department
Admiral Makarov State University of Maritime and Inland Shipping
Saint-Petersburg, Russia
deryabinvv@gumrf.ru

Abstract—**The use of a multilayered feedforward neural network for the task of vessel position fixing on the basis of preliminary measured depths is proposed. The results of the tests with the use of simulated and real data are in favor of the feasibility of the proposed method.**

Keywords—neural network, vessel, depth, navigation, map data

I. Introduction

The task of terrain-aided positioning can be solved by different ways. By the moment, two different approaches are widely used: TERCOM and SITAN. TERCOM [1] is based on the search of a similar to measured terrain height sequence throughout a general database generated on the basis of preliminary measurements. SITAN [2] uses the Kalman Filter. The measurement equation of the filter is derived on the basis of the linearization of a terrain surface nearby an approximate position which can be got, for example, with an inertial navigation system. TERCOM requires significant computational resources as the search is carried out in real time mode. SITAN does not perform the search procedure directly. Nevertheless, divergency problems remain relevant for it because the Kalman Filter is a linear estimation algorithm.

In order to improve computational performance of search-based methods, it is reasonable to use neural networks for terrain-aided navigation. Such ideas have already been proposed. For instance, in [3], Adaline predicts the coordinates and the flight bearing angle of an aerial vehicle on the basis of a height sequence given as an input to the network. LSTM is used to solve a similar task in [4].

The above and many other researches do not contain test results in the form of numerical indicators of the accuracy to be estimated in the operating mode of the neural network systems. This research is a development of [5]. It is intended to eliminate the above-mentioned drawbacks.

II. Network architecture

A. Input and output signals

A p-dimensional vector is given to the network input. Each element of the vector is a sequentially measured value of sea depth.

Vessel geodetic coordinates (latitude and longitude) are considered as the network output.

A general scheme of the system operation is presented in Fig. 1.

Fig. 1. A general scheme of the operation of the bathymetric navigation system

Moving alongside one of the reference meridians with longitude λ_*, the vessel takes spot soundings at its intersections with the reference parallels with latitudes φ_1, φ_2, ..., φ_p, ..., φ_{lim}, where φ_{lim} is a limited latitude value in the area for bathymetric navigation. It should be noted that latitudes φ_1, φ_2, ..., φ_{p-1} refer to data collection area (letter A in Fig. 1) whilst φ_p, φ_{p+1}, ..., φ_{lim} refer to system working area (letter B in Fig. 1).

B. Neural network configuration

A feedforward multi-layered and fully-connected neural network is used as the basic configuration. The hidden neurons have hyperbolic tangent activation functions, while the output neurons have linear ones. It is known that such a network can perform continuous multidimensional «input-output» mappings [6]. Figure 2 shows the architecture of the network.

Fig. 2. Neural network architecture (M – the number of hidden layers, N – the number of neurons)

III. Training set

There are two ways to get bathymetric data for the network training. The first way is to obtain training data as the result of preliminary measurements. The second way is to extract them from a database which was not intentionally

collected for the purposes of terrain-aided navigation. Let us consider these ways in more detail.

For the first case, it is possible to take soundings strictly at reference points. A hydrographic vessel visits the mentioned points in accordance with a predefined program, that is, it moves by a route with waypoints equal to the points.

For the second case, it is possible to use the data contained, for example, in vector electronic navigational charts (ENC). Spot soundings in these charts, in general, are known not for all nodes of the regular grid. Therefore, it is necessary to use interpolation methods to fix the problem. Besides ENCs, sounding data may be obtained on the basis of a digital terrain model, e.g., GEBCO, ETOPO, IBCAO, etc.

For all the above cases, sounding steps for latitude $\Delta\varphi$ and longitude $\Delta\lambda$ are parameters of the data preparation algorithm.

IV. TRAINING AND TESTING

A. Training

After forming the training set, the task of the neural network training arises in order to achieve the mapping capability.

Traditionally, the whole data set for training is divided into two subsets: training and validation sets. The training set is directly used in the optimization task, that is, it is used to correct the weights of a neural network. The validation set is used to mitigate overfitting. The later means that the neural network simply remembers certain data. For the certain case, such approach appears to be acceptable because the task is to map only chart data, and they are available in details enough. In connection with the above, the validation set is not used in this research.

During the training process, the mean squared error is minimized traditionally. At each epoch, the Adamax optimization method [7] is used and the distance between the true and predicted positions on the WGS-84 ellipsoid is calculated. A better network condition corresponds to the minimum value of the above distance. The size of the mini-batch is an important parameter of the Adamax method.

B. Testing

Testing is mainly conducted in order to check how adequately the neural network will predict vessel coordinates if the input signals are different from those used at the training stage. The differences result from different factors. Let us consider the main of them below.

Firstly, input samples can be distorted at the stage of the use of the model because of errors in soundings taken with a single-beam echosounder. Secondly, modifications of input samples may be interpreted as sea-level fluctuations which cannot be predicted in a certain case. Thirdly, the vessel may track in a recommended direction in a bathymetric system area only with finite accuracy. The last factor will be considered in more details below.

Underwater terrain may be featured differently in different directions. For instance, it changes significantly in the south-north direction, but in the west-east direction, practically, it does not. Therefore, within a bathymetric

navigation area, recommended track directions may be established. These directions are to be followed up as terrain profiles are known only for them. It is clear that a deviation from the recommended direction will result in the additional distortion of the input signal of the neural network.

V. SIMULATION

To verify the performance of the proposed method, computer simulation was conducted with the use of the Python programming language and the TensorFlow library, as well as some other libraries. The simulation was conducted as follows.

Within a sea-area with the size of approximately 50x50 nautical miles, a functional dependency of the depth d on coordinates φ, λ was modelled in the following way: $d = 10 + 30(\varphi - \varphi_S)^2 + 200(\lambda - \lambda_W)$, where the indexes «S» and «W» refer to the southern and western boundaries of the area correspondently. The number of depth values is $p = 51$. Step sizes for latitude and longitude are $\Delta\varphi \approx \Delta\lambda \approx 185$ m. It is assumed that the vessel goes strictly along the meridians from south to north.

Several variants of the neural network configuration are considered. They differ both the number of hidden layers M and neurons N. Each such neural network is trained in a mini-batch mode with 10 samples per the batch. The results of the training after 1000 epochs are performed in the table below. Each cell of the table contains a minimal value of the maximum positional error obtained during the training process.

On the basis of the data given in the table, the most precise configuration is chosen. For that configuration, the training process continues up to 10000 epochs. As the result, the positional error has been reduced to 912 m.

TABLE TRAINING RESULTS

N	M				
	1	2	3	4	5
10	1219	1102	1037	1058	1000
20	1139	1095	997	975	1024
30	1090	1168	1077	1042	1022
40	1213	1018	1054	966	1044
50	1223	1076	1164	1076	1020
100	1195	1158	1020	1207	1030

After the training, a test procedure of the neural network performance was conducted. During the test, it is assumed that the vessel moves alongside meridians that do not match precisely with those for which the sounding data are available and used for the creation of the training set. Firstly, some intermediate meridians are considered. Their longitudes are calculated as mean values of longitudes of the adjacent meridians. Secondly, 1000 random meridians are used. The longitude of the random meridian is taken with the use of a uniform random number generator from an interval of possible values within the sailing area. The neural network demonstrates 985 m in maximum positional error after the test has been performed.

ENC US3SC10M is also used to obtain the training data. With the *ogrinfo* command-line utility from the GDAL package, a depth array was extracted from a corresponding layer of the chart. Each component of the array contains

three values: the WGS-84 latitude, longitude of the point, and depth at it.

A sea-area of about 50x50 nautical miles is considered on the chart. The samples are obtained while moving alongside imaginary sounding tracks spaced with the distance of about 1 nautical mile. The number of depth values is $p = 21$. That is, the first bathymetric-based position fix is available after proceeding 20 miles in the direction from south to north. The value of depth at each point is considered as the value of depth in the nearest charted point. The training procedure has been conducted. The maximum positional error is 239 m for the samples in the training data.

VI. Conclusion

An underwater terrain-based navigation solution on the basis of a neural network is proposed. The network takes a sequence of depth values as its input and predicts the corresponding latitude and longitude. The network is preliminary trained on the basis of bathymetric data that are obtained for a certain sea-area in advance. Underwater terrain modelling and real navigational chart data used for training allows us to consider that the proposed method has acceptable navigational accuracy.

Further studies may be performed in the direction of the modernization of the algorithms for testing the bathymetric neural navigation system; navigational application of underwater terrain; and the use of information on sea level fluctuations due to tidal and/or wind factors. The assessment of potential accuracy of the system taking into account the most significant uncertainties is also considered as un important line of further research.

References

[1] J. P. Golden, "Terrain Contour Matching (TERCOM): a cruise missile guidance aid" Image processing for missile guidance, vol. 238, pp. 10–18, 1980.

[2] W. E. Longenbaker, "Terrain-aided navigation of an unpowered tactical missile using autopilot-grade sensors" Journal of Guidance, Control, and Dynamics, vol. 7, № 2, pp. 175–182, 1984.

[3] S. J. Chen, L. Luoh, "2-D TERCOM Guidance – Neural Net Approach" 2009 10th International Symposium on Pervasive Systems, Algorithms, and Networks, pp. 438–441, 2009.

[4] S. Lee, H. Bang, "Terrain Contour Matching with Recurrent Neural Networks" 2018 IEEE Aerospace Conference, Big Sky, MT, pp. 1–9, 2018.

[5] V. V. Deryabin, "Neural network-based method for determining vessel position by seabed relief" Vestnik Gosudarstvennogo universiteta morskogo i rechnogo flota imeni admirala S. O. Makarova, vol. 15, № 5, pp. 723–734, 2023 (*in Russian*).

[6] M. H. Stone, "The generalized Weierstrass approximation theorem" Mathematics Magazine, vol. 21, № 4, pp. 167–184, 1948.

[7] D. P. Kingma, J. L. Ba, "Adam: A method for stochastic optimization" 3rd International Conference on Learning Representations, preprint arXiv:1412.6980, 2014.

Correction Of A Strapdown Inertial Navigation System Using Information About Geolocated Landmarks

R.M. Antonov
JSC «Inertial technologies of "Technocomplex"»
Rameskoye, Russia
RAntonov@inertech.ru

A.N. Pushkin
JSC «Inertial technologies of "Technocomplex"»
Rameskoye, Russia
APushkin@inertech.ru

M.O. Kalinina
JSC «Inertial technologies of "Technocomplex"»
Rameskoye, Russia
MKalinina@inertech.ru

A.V. Nekrasov
JSC «Inertial technologies of "Technocomplex"»
Rameskoye, Russia
ANekrasov@inertech.ru

Abstract—**This paper describes a method for correction of a strapdown navigation system by determining spatial position and angular orientation from information received from a technical vision system. This method allows obtaining corrections with accuracy comparable to the accuracy of corrections from a satellite navigation system. To test the method, a model of the system which makes it possible to fix and process the information received from the technical vision system has been developed. The requirements for geolocated landmarks were defined to form corrections. Based on these requirements, a training set was assembled, and a neural network was trained to classify landmarks. The results of model tests confirmed the accuracy characteristics of the corrections and the suitability of using the system as a source of integrated information.**

Keywords—SINS correction, technical vision, autonomous navigation, neural networks, geolocated landmarks

I. Introduction

Primary means of determining spatial position and angular orientation are strapdown inertial navigation systems (SINS). SINS determine coordinates and orientation of vehicles with high accuracy, solving the problem of autonomous navigation. However, due to the integrating nature of the navigation algorithm, SINS incline to accumulate errors over time. This is caused by stochastic errors of gyroscopes and accelerometers, errors during initialization of algorithms at the stage of initial alignment, as well as errors of numerical integration methods. The accuracy characteristics of SINS are typically provided by means of various algorithms that compensate for errors using information from external sources. In this case, the main advantage of SINS - autonomy - is lost [1].

There are many sources of compensating information, such as Doppler navigation systems, magnetometer, and satellite systems [2]. The use of satellite-based differential correction systems is the most widespread. Information from the satellite navigation system, being integrated with inertial system readings, allows high-accuracy corrections to SINS. At the same time, the satellite signal has low resistance to interference and is subject to "spoofing" - signal substitution. In some areas the received signal can be quite weak and unstable. In the light of this, there are several scenarios where forming corrections based on the information from a satellite navigation system presents a problem, which makes it necessary to use autonomous information sources integrated into the onboard equipment complex.

A significant number of modern technically complex devices are equipped with technical vision systems (TVS) as part of their onboard equipment. Correlation-extreme navigation systems using visual information based on the calculation of the mutual or autocorrelation function of random processes characterizing the state of the object being located and determination of the coordinates of the main extremum of this function are widely used in aviation [3]. Typically, the developed methods assume that the technical vision system is oriented downward along the terrain's vertical [4, 5]. The use of information from the technical vision system directed at an acute angle with respect to the terrain horizon allows obtaining much more information about observed objects and solving problems inherent in classical correlation-extreme systems. Furthermore, the development of artificial intelligence systems has enabled processing the information obtained from such technical vision systems [6]. At the same time, there has been widespread study on application of artificial intelligence methods to improve the accuracy characteristics of navigation systems of different nature [7, 8].

Thus, the task arises to develop a navigation system that utilizes information from the technical vision system, oriented at an acute angle to the horizon. This system would enable autonomous navigation without reliance on external sources of corrective information.

II. Scheme of algorithm construction in the presence of corrective infrrmation from TVS

Technical vision systems provide image acquisition, processing, and analysis of the frame configuration with subsequent transfer of information to the control device. This allows creating a correction method based on video information analysis.

Image processing implies selecting unique features, classifying them according to a database of known geolocated features, determining their relative coordinates with respect to the location object, and calculating corrective measurements of coordinates and heading. Some objects of natural and anthropogenic origin serve as such features.

Solving the problem involves creating a catalog of geospatial data, with each object having known geographic coordinates. The task of landmark detection and classification is solved using machine learning methods, including artificial neural networks.

A camera angled to horizon offers much larger coverage compared to traditional way of capturing terrain images. This orientation allows interaction with both horizontal and vertical projection of objects. Application of this method allows to reduce the requirements for camera orientation and height of its location.

With two cameras pointing toward the horizon, it is possible to determine the distance to a landmark and the azimuth angle relative to it from the acquired images. Each camera processes its own two-dimensional image. Based on the known orientation of the cameras b and the distance L between them, the distance d to the landmark and the azimuth angle α relative to it are calculated (Fig. 1).

Fig. 1. Determination of the distance to a landmark using two cameras.

The algorithm operates by sequentially forming range and azimuth measurements based on the information from the TVS while moving along the trajectory. The information transmitted by the TVS is discrete, the transmission speed depends on the frame rate. The algorithm calculates corrections of geolocated landmarks based on the TVS data in the area.

The landmark base element is characterized by longitude λ_i and latitude φ_i. In the developed algorithm, a flat rectangular system is used as the base coordinate system. Since solving the TVS task takes place on a conditional plot of the Earth's surface no larger than 10 km by 10 km, the transition to a planar rectangular coordinate system (RCS) assumes that longitude is independent on latitude, and the relationship between linear displacements and coordinate changes remains constant. Transition to the rectangular coordinate system is carried out in accordance with GOST 32453-2017 with the GSK-2011 in Gauss-Krueger projection as the coordinate system. The error of coordinate transformation does not exceed 0.001 m [9].

Consider the situation when a frame contains information about several landmarks (Fig. 2).

Fig. 2. Frame containing information about multiple landmarks.

Let point O_1 with coordinates λ_1, φ_1 be the center of a circle with radius d_1, and let point O_2 with coordinates λ_2, φ_2 be the center of another circle with radius d_2. These two circles intersect at point O_3 with coordinates λ_3, φ_3 corresponding to the coordinates of the object. To find the intersection point, we need to solve the system of equations:

$$\begin{cases} (\lambda_3 - \lambda_1)^2 + (\varphi_3 - \varphi_1)^2 = d_1^2 \\ (\lambda_3 - \lambda_2)^2 + (\varphi_3 - \varphi_2)^2 = d_2^2 \end{cases} \quad (1)$$

The solution of the system of equations (1) will yield two equally probable points. The choice of which point contains the object's coordinates is based on current SINS readings.

Assume the coordinates of the obtained point are [0, 0] and draw a straight line perpendicular to the axis to the north (Fig. 3).

Fig. 3. Determination of the heading of an object by a single landmark.

According to (2), we determine the relative coordinates φ_1', λ_1', φ_2', λ_2'.

$$\begin{cases} \varphi_1' = \varphi_1 - \varphi_3 \\ \lambda_1' = \lambda_1 - \lambda_3 \\ \varphi_2' = \varphi_2 - \varphi_3 \\ \lambda_2' = \lambda_2 - \lambda_3 \end{cases} \quad (2)$$

According to (3), we determine the angles β and γ.

$$\begin{cases} \beta = \arcsin\left(\dfrac{\lambda_1'}{d_1}\right) \\ \gamma = \arcsin\left(\dfrac{\lambda_2'}{d_2}\right) \end{cases} \quad (3)$$

Then the heading can be determined according to (4):

$$\psi = \frac{\beta + \alpha_1 - \gamma - \alpha_2}{2} \quad (4)$$

When the frame contains only one landmark, the coordinates are determined sequentially at set time intervals.

During the first measurement, the distance to the landmark d_1 and azimuth angle α_1 are determined, with the angle ψ_{SINS} transmitted by SINS representing the object heading. At the second measurement, the object's coordinates are determined by integrating the path traveled since the first measurement, and calculating where it should have appeared following the heading ψ_{SINS}. The difference

between the result of integration and the actual coordinates of the object allows to determine the heading correction.

Let O be the center of a circle with radius d with coordinates x_O, y_O corresponding to the object's coordinates. Let d be the distance to the landmark. Let L represent the landmark with coordinates x_L, y_L. OL is the line connecting object and the landmark with length d. ON is the line connecting object and the direction to the north with length d (Fig. 4).

Fig. 4. Determination of the coordinates of an object by a single landmark.

The angle β between the direction to the north and the axis to the landmark is defined as:

$$\beta = \psi + \alpha \tag{5}$$

where α is the angle between the heading ψ and the axis to the landmark. If the landmark is located to the left of the heading axis, $\alpha > 0$, otherwise $\alpha < 0$. Then the y coordinate of the object is found according to (6):

$$y_O = 2d\sin\left(\frac{\beta}{2}\right)^2 - d + y_L \tag{6}$$

The coordinate x is determined according to (7). The result of (7) yields two equally probable points. The selection of which of these coordinates corresponds to the object's coordinate is made based of current SINS readings.

$$\begin{cases} x_O = x_L - 2d\sin\left(\frac{\beta}{2}\right)\sqrt{1 - \sin\left(\frac{\beta}{2}\right)^2} \\ x_O = x_L + 2d\sin\left(\frac{\beta}{2}\right)\sqrt{1 - \sin\left(\frac{\beta}{2}\right)^2} \end{cases} \tag{7}$$

III. LIBRARY OF GEOLOCATED LANDMARKS

Geolocated landmarks are prominent objects on the ground with a precise geographical reference, detectable by visual methods and unambiguously identifiable. They may include natural (water basins, landforms, forests, etc.) as well as anthropogenic objects (roadways, urban structures, power lines, clearings, agricultural facilities).

The solution of the problem involves creating a specialized library containing information about a set of geolocated landmarks. For each landmark in the library, a set of images is created to capture as many features as possible. Due to the horizontal orientation of the camera, each element

of the library can be described in much greater detail than with vertical orientation. Consequently, classification of landmarks by computer vision becomes possible.

IV. TESTING

Based on the results of the theoretical study, a model of a video navigation system has been developed. The model includes an NVIDIA Jetson Nano microcomputer, a stereo camera and a program module written in Python that implements the described algorithms.

A dataset containing distinguishable objects from Ramenskoye was collected to work out the navigation algorithm. Each object is described by a set of photos from different angles and its geographical coordinates. A neural network was trained on this dataset to perform landmark detection and classification. The classification accuracy achieved 0.96 using the Precision metric. Figure 5 shows an example of detection and classification of a landmark from the set, as well as determination of the distance to it.

Fig. 5. Landmark detection, classification and distance determination.

For the model test, a dataset was prepared including:

- video recordings containing images from a stereo camera mounted on a moving car. A number of frames contain landmarks, and corrections are generated using these frames;

- navigation solution from LINS-100RS navigation system installed on the vehicle, measurements from satellite navigation system, integrated navigation solution;

- map with geolocated landmarks corresponding to the vehicle movement.

Corrective measurements are formed based on information from the stereo camera with periodicity of 10-30 seconds. The accuracy of corrections is limited by map accuracy and is within 10-20 meters. Figure 6 shows the graph of geographic coordinates determination error in meters relative to the combined inertial and satellite readings. The accuracy of geographic coordinates determination was between 10-30 meters, with the greatest influence of the error in determining the distance to the landmark. The obtained results demonstrate the suitability of the developed system layout for solving autonomous navigation problem.

Fig. 6. Navigation algorithm test results.

V. CONCLUSION

This paper presents a method for creating corrections for SINS with an accuracy depending on the accuracy of the landmark distance determination, its geo-referencing and resolution of the vision system.

To test this algorithm, a system model realizing it was developed. Tests have shown the accuracy in determination of geographic coordinates in the region of 10-30 meters.

The developed method can be used as a source of corrective information in the absence of signals from external sources.

REFERENCES

[1] Dzhandzhgava G. I., Avgustov L. I., Babichenko A. V., Orekhov M. I., Sukhorukov S. Ya., Shkred V. K., Navigaciya letatel'nyh apparatov v okolozemnom prostranstve [Navigation of aircraft in near-Earth space]. Russia, Moscow: Nauchtekhlitizdat, 2015.

[2] Proletarskij A.V., Neusypin K.A. "Sposoby korrekcii navigacionnyh sistem i kompleksov letatel'nyh apparatov" [Methods of correction of navigation systems and aircraft complexes.]. Inzhenernyj zhurnal: nauka i innovacii, no.3 (3): 2012.

[3] Beloglazoe I.N., Tarasenko V.P. Korrelyacionno-ekstremal'nye sistemy [Correlation-extreme systems]. USSR, Moscow: Sov. radio, 1974.

[4] Tanchenko, A.P., Fedulin, A.M., Bikmaev, R.R. et al. UAV Navigation System Autonomous Correction Algorithm Based on Road and River Network Recognition. Gyroscopy Navig. 11, 293–299 (2020). https://doi.org/10.1134/S2075108720040100.

[5] Conte, G.; Doherty, P. Vision-Based Unmanned Aerial Vehicle Navigation Using Geo-Referenced Information. EURASIP J. Adv. Signal Process. 2009, 9, 387308.

[6] X. Wang, P. Cheng, X. Liu, and B. Uzochukwu. Fast and Accurate, Convolutional Neural Network Based Approach for Object Detection from UAV. arXiv preprint arXiv:1808.05756, 2018.

[7] QinYuan He, HuaPeng Yu & YuChen Fang Deep Learning-Based Inertial Navigation Technology for Autonomous Underwater Vehicle Long-Distance Navigation—A Review. Gyroscopy Navig. 14, 267–275 (2023). https://doi.org/10.1134/S2075108723030070.

[8] Golev I.M., Ugryumov R.B., Zaenceva T.I., Mandrykin A.V. Magnitometricheskaya sistema lokal'noj navigacii s nejrosetevoj obrabotkoj informacii [Magnetometer System of Local Navigation With Neural Network Information Processing]. Vozdushno-kosmicheskie sily. Teoriya i praktika, 2021, no.18. (In Russian)

[9] GOST 32453-2017. Global navigation satellite system. Coordinate systems. Methods of transformations for determinated points coordinate.

Method of Constructing Gravity Gradient Map Based on the Earth's Gravity Field Model

Li Da

Tianjin Navigation Instrument Research Institute
Tianjin, China
College of Intelligent Systems Science and Engineering, Harbin Engineer University
Harbin, China
18322617798@163.com

Gao Wei

Tianjin Navigation Instrument Research Institute
Tianjin, China
zlclbgw@163.com

Li Chengsuo

Tianjin Navigation Instrument Research Institute
Tianjin, China
chengsuol@163.com

Li Zhong

Tianjin Navigation Instrument Research Institute
Tianjin, China
createform@163.com

Zhao Lin

College of Intelligent Systems Science and Engineering, Harbin Engineer University
Harbin, China
zhaolin@hrbeu.edu.cn

Li Rui*

China Aero Geophysical Survey and Remote Sensing Center for Land and Resources
Beijing, China
li510231865@163.com

Abstract—To the demands for gravity gradient-assisted navigation and the construction of gravity gradient reference maps, this study introduced a method for developing a gravity gradient reference map based on the Earth's gravity field model. Initially, by defining the gravity gradient, the relationship between gravity potential and gravity gradient tensor in the geographic coordinate system was established. The study then derived the calculation formulas for all independent components of the gravity gradient based on the EGM2008 Earth gravity field spherical harmonic model. Subsequently, within a $5° \times 5°$ region, a gravity gradient reference map with a spatial resolution of $5'$ was meticulously constructed. The correctness of the construction was validated through Laplace constraint conditions. Furthermore, a comprehensive statistical analysis of the gravity gradient reference map in the selected region was conducted, offering substantial data support for the allocation of weight to matching quantities in gravity gradient-assisted navigation. The analysis revealed that among the six components of the gravity gradient tensor, the vertical component encapsulated the most diverse features of the Earth's gravity field. Its matching characteristics significantly surpassed the other five components, establishing it as the pivotal matching component in gravity gradient-assisted navigation. Consequently, in the realm of gravity gradient-assisted navigation research, assigning the highest weight to the vertical component was deemed crucial.

Keywords—gravity gradient reference map; Earth's gravity field model; EGM2008; vertical component

I. INTRODUCTION

Gravity gradient is the second-order spatial derivative of gravity potential. Compared to traditional gravity anomaly signals, gravity gradient signals possess higher spatial resolution and provide more dimensions of information about the Earth's gravity field. In the field of gravity-assisted navigation, gravity gradient-assisted navigation theoretically offers higher combined navigation accuracy and can increase the matching adaptation areas, thus enhancing the practical level of gravity gradient-assisted navigation compared to solely gravity anomaly-assisted navigation. A gravity gradient reference map is a crucial component of gravity gradient-assisted navigation. A high-resolution and high-precision gravity gradient reference map forms the basis for achieving high-precision gravity gradient-assisted navigation. Due to the scarcity of mature commercially available dynamic gravity gradient sensors worldwide, conducting extensive gravity gradient survey work is challenging. As a result, research on gravity gradient-assisted navigation currently remains limited to theoretical simulations and has not yet progressed to physical validation.

In response to the application demands for constructing gravity gradient reference maps in gravity gradient-assisted navigation, a method for developing a gravity gradient reference map based on the Earth's gravity field model was proposed in this study. The study initially established the relationship between perturbed gravity potential and perturbed gravity gradient tensor components in the geographic coordinate system based on the definition of gravity gradient. Calculation formulas for all independent components of the perturbed gravity gradient based on the EGM2008 Earth gravity field spherical harmonic model were derived. Subsequently, a gravity gradient reference map with a spatial resolution of $5'$ for all components of the perturbed gravity gradient within a $5°×5°$ region was constructed. The correctness of the constructed gravity gradient reference map was validated through Laplace constraint conditions. Lastly, a statistical analysis and matching adaptability analysis of the gravity gradient reference map for each component of the perturbed gravity gradient in the selected region were conducted, providing data support for the allocation of weight to matching quantities in gravity gradient-assisted navigation[1].

II. DEFINITION AND CALCULATION FORMULA OF GRAVITY GRADIENT

A. Definition of Gravity Gradient

Gravity gradient is the spatial second derivative of gravity potential, representing the spatial rate of change of the gravity vector. In the geographic coordinate system (the

East-North-Zenith system chosen in this study), the gravity vector can be decomposed into three components in the x, y, and z directions, each having a gradient along the axis parallel to the coordinate axes. Therefore, the gravity gradient tensor has a total of nine components and can be represented as:

$$\Gamma = \begin{bmatrix} \dfrac{\partial g_x}{\partial x} & \dfrac{\partial g_x}{\partial y} & \dfrac{\partial g_x}{\partial z} \\ \dfrac{\partial g_y}{\partial x} & \dfrac{\partial g_y}{\partial y} & \dfrac{\partial g_y}{\partial z} \\ \dfrac{\partial g_z}{\partial x} & \dfrac{\partial g_z}{\partial y} & \dfrac{\partial g_z}{\partial z} \end{bmatrix} = \begin{bmatrix} \Gamma_{xx} & \Gamma_{xy} & \Gamma_{xz} \\ \Gamma_{yx} & \Gamma_{yy} & \Gamma_{yz} \\ \Gamma_{zx} & \Gamma_{zy} & \Gamma_{zz} \end{bmatrix} \quad (1)$$

Where $\Gamma_{ij}(i, j = x, y, z)$ represents the corresponding directional component in the gravity gradient tensor matrix, indicating the rate of change of the gravity component gi in the j direction. Each component can be expressed as the sum of the corresponding normal gravity gradient component and the perturbed gravity component, as shown in the following expression:

$$\Gamma_{ij} = U_{ij} + T_{ij} \quad (2)$$

Where $U_{ij}(i, j = x, y, z)$ is the corresponding normal gravity gradient component in the direction, and $T_{ij}(i, j = x, y, z)$ is the corresponding perturbed gravity gradient component in the direction. The various components of the normal gravity gradient can be obtained through normal gravity calculations, while the components of the perturbed gravity gradient are derived through the expansion of the Earth's gravity field model.

B. Normal gravity gradient

Under the assumption of a rotating ellipsoid, the projection of the normal gravity vector g^g at the sea level onto the geographic coordinate system can be represented as:

$$g^g = \begin{bmatrix} 0 & -\beta_3 h \sin 2L & -g_{Lh} \end{bmatrix}^T \quad (3)$$

In the equation, h is the altitude, L is the geographic latitude, β_3 is a constant typically taken as $\beta_3 = 8.08 \times 10^{-9} \mathrm{s}^{-2}$, g_{Lh} is the value of normal gravity near the Earth's surface that varies with latitude and altitude. It is commonly calculated using the formula:

$$g_{Lh} = g_e(1 + \beta \sin^2 L - \beta_1 \sin^2 2L) - \beta_2 h \quad (4)$$

In the equation, g_e is the normal gravity value at the equator, typically taken as $g_e = 9.780325 \, m/s^2$, β, β_1, and β_2 are constants related to the shape of the Earth, typically taken as $\beta = 0.00530240$, $\beta_1 = -0.00000582$, and $\beta_2 = 3.08 \times 10^{-6} \mathrm{s}^{-2}$ respectively.

By spatially differentiating the gravity vector g^g, we can obtain that the values of U_{xx}, U_{xy}, U_{xz}, U_{yx} and U_{zx} in the normal gravity gradient tensor in the geographic coordinate

system are all zero. The formulas for the remaining components can be expressed as:

$$\begin{cases} U_{yy} = -\dfrac{2\beta_3 h}{R_M + h} \cos 2L \\ U_{yz} = -\beta_3 \sin 2L \\ U_{zy} = \dfrac{\beta \sin 2L + 2\beta_1 \sin 4L}{R_M + h} \\ U_{zz} = -\beta_2 \end{cases} \quad (5)$$

In fact, the normal gravity gradient is mainly determined by the corresponding parameters of the rotating ellipsoid, and its numerical variation is very slow when the parameters are fixed. Therefore, the perturbation gravity gradient is crucial for obtaining the reference map of oceanic gravity gradients.

C. Disturbance Gravity Gradient

The perturbed gravity potential subtracts the influence of centrifugal force from gravity, making the perturbed gravity field a conservative field that satisfies the constraint of the Laplace equation, given by:

$$\begin{cases} T_{ij} = T_{ji}(i, j = x, y, z; i \neq j) \\ T_{xx} + T_{yy} + T_{zz} = 0 \end{cases} \quad (6)$$

The perturbed gravity potential function T based on the spherical harmonic model can be expressed as[2]:

$$T(r, \theta, \lambda) = \dfrac{\mu}{r} \sum_{n=2}^{N} \left(\dfrac{R_e}{r}\right)^n \sum_{m=0}^{n} (\bar{C}_n^m \cos m\lambda \\ + \bar{S}_n^m \sin m\lambda)\kappa_m \bar{P}_{nm}(\cos \theta) \quad (7)$$

In the equation, N is the degree of the spherical harmonic model, which can go up to 2190, r is the distance from the target point to the center of the coordinates, λ is the longitude, θ is the residual latitude, μ=GM is the gravitational constant, R_e is the equatorial radius of the Earth's reference ellipsoid, $\bar{C}_n^m$ and $\bar{S}_n^m$ are the spherical harmonic coefficients, and $\bar{P}_{nm}(\cdot)$ are the normalized associated Legendre functions. By taking the first derivative of the perturbed gravity potential function, we can obtain the expression for the projection of the perturbed gravity vector in the geographic coordinate system[3].

$$\begin{cases} \dfrac{\partial T}{\partial x} = \dfrac{1}{r \sin \theta} \dfrac{\partial T}{\partial \lambda} \\ \dfrac{\partial T}{\partial y} = -\dfrac{1}{r} \dfrac{\partial T}{\partial \theta} \\ \dfrac{\partial T}{\partial z} = \dfrac{\partial T}{\partial r} \end{cases} \quad (8)$$

Taking the derivative of the above equation, we can obtain the expression for the projection of the perturbed gravity gradient in the geographic coordinate system.

$$\begin{cases} T_{xx} = \dfrac{\partial^2 T}{\partial x^2} = \dfrac{1}{r^2 \sin^2 \theta} \dfrac{\partial^2 T}{\partial \lambda^2} + \dfrac{1}{r^2 \tan \theta} \dfrac{\partial T}{\partial \theta} + \dfrac{1}{r} \dfrac{\partial T}{\partial r} \\[2mm] T_{xy} = \dfrac{\partial^2 T}{\partial x \partial y} = -\dfrac{1}{r^2 \sin \theta} \dfrac{\partial^2 T}{\partial \lambda \partial \theta} + \dfrac{\sin \theta}{r^2 \cos^2 \theta} \dfrac{\partial T}{\partial r} \\[2mm] T_{xz} = \dfrac{\partial^2 T}{\partial x \partial z} = \dfrac{1}{r \sin \theta} \dfrac{\partial^2 T}{\partial \lambda \partial r} - \dfrac{1}{r^2 \sin \theta} \dfrac{\partial T}{\partial r} \\[2mm] T_{yy} = \dfrac{\partial^2 T}{\partial y^2} = \dfrac{1}{r^2} \dfrac{\partial^2 T}{\partial \theta^2} + \dfrac{1}{r} \dfrac{\partial T}{\partial r} \\[2mm] T_{yz} = \dfrac{\partial^2 T}{\partial y \partial z} = -\dfrac{1}{r} \dfrac{\partial^2 T}{\partial \theta \partial r} + \dfrac{1}{r^2} \dfrac{\partial T}{\partial \theta} \\[2mm] T_{zz} = \dfrac{\partial^2 T}{\partial z^2} = \dfrac{\partial^2 T}{\partial r^2} \end{cases} \quad (9)$$

Substituting equation (7) into equation (9), the formula for calculating the perturbed gravity gradient can be obtained[4].

$$\begin{cases} T_{xx} = \dfrac{\mu}{r^3} \sum_{n=2}^{N} \left(\dfrac{R_e}{r}\right)^n \sum_{m=0}^{n} \begin{pmatrix} \bar{C}_n^m \cos m\lambda \\ +\bar{S}_n^m \sin m\lambda \end{pmatrix} \kappa_m \bar{P}_{nm}^{xx}(\cos \theta) \\[3mm] T_{xy} = \dfrac{\mu}{r^3} \sum_{n=2}^{N} \left(\dfrac{R_e}{r}\right)^n \sum_{m=0}^{n} \begin{pmatrix} \bar{C}_n^m \cos m\lambda \\ +\bar{S}_n^m \sin m\lambda \end{pmatrix} \kappa_m \bar{P}_{nm}^{xy}(\cos \theta) \\[3mm] T_{xz} = \dfrac{\mu}{r^3} \sum_{n=2}^{N} \left(\dfrac{R_e}{r}\right)^n \sum_{m=0}^{n} \begin{pmatrix} \bar{C}_n^m \cos m\lambda \\ +\bar{S}_n^m \sin m\lambda \end{pmatrix} \kappa_m \bar{P}_{nm}^{xz}(\cos \theta) \\[3mm] T_{yy} = \dfrac{\mu}{r^3} \sum_{n=2}^{N} \left(\dfrac{R_e}{r}\right)^n \sum_{m=0}^{n} \begin{pmatrix} \bar{C}_n^m \cos m\lambda \\ +\bar{S}_n^m \sin m\lambda \end{pmatrix} \kappa_m \bar{P}_{nm}^{yy}(\cos \theta) \\[3mm] T_{yz} = \dfrac{\mu}{r^3} \sum_{n=2}^{N} \left(\dfrac{R_e}{r}\right)^n \sum_{m=0}^{n} \begin{pmatrix} \bar{C}_n^m \cos m\lambda \\ +\bar{S}_n^m \sin m\lambda \end{pmatrix} \kappa_m \bar{P}_{nm}^{yz}(\cos \theta) \\[3mm] T_{zz} = \dfrac{\mu}{r^3} \sum_{n=2}^{N} \left(\dfrac{R_e}{r}\right)^n \sum_{m=0}^{n} \begin{pmatrix} \bar{C}_n^m \cos m\lambda \\ +\bar{S}_n^m \sin m\lambda \end{pmatrix} \kappa_m \bar{P}_{nm}^{zz}(\cos \theta) \end{cases} \quad (10)$$

In the equation, the Legendre polynomials for each component are:

$$\begin{cases} \bar{P}_{nm}^{xx}(\cos \theta) = \left[\dfrac{m \cos^2 \theta - m^2}{\sin^2 \theta} - (n+1)\right] \bar{P}_{nm}(\cos \theta) \\ \qquad\qquad - \cot \theta \bar{P}_{nm+1}(\cos \theta) \\[2mm] \bar{P}_{nm}^{xy}(\cos \theta) = \dfrac{-m(m-1)\cos \theta}{\sin^2 \theta} \bar{P}_{nm}(\cos \theta) \\ \qquad\qquad + \dfrac{m}{\sin \theta} \bar{P}_{nm+1}(\cos \theta) \\[2mm] \bar{P}_{nm}^{xz}(\cos \theta) = \dfrac{-m(n+2)}{\sin \theta} \bar{P}_{nm}(\cos \theta) \\[2mm] \bar{P}_{nm}^{yy}(\cos \theta) = \left[\dfrac{m^2 - m \cos^2 \theta}{\sin^2 \theta} - (n+1)^2\right] \bar{P}_{nm}(\cos \theta) \\ \qquad\qquad - \cot \theta \bar{P}_{nm+1}(\cos \theta) \\[2mm] \bar{P}_{nm}^{yz}(\cos \theta) = m(n+2)\cot \theta \bar{P}_{nm}(\cos \theta) \\ \qquad\qquad - (n+2)\cot \theta \bar{P}_{nm+1}(\cos \theta) \\[2mm] \bar{P}_{nm}^{zz}(\cos \theta) = (n+1)(n+2) \bar{P}_{nm}(\cos \theta) \end{cases}$$

This enables the calculation of the six components of the perturbed gravity gradient tensor, thereby completing the construction of the marine gravity gradient reference map.

III. CONSTRUCTION OF THE GRAVITY GRADIENT REFERENCE MAP

In this paper, the sea area with a longitude range of 112°E to 117°E and a latitude range of 10°N to 15°N was selected for the construction of the perturbed gravity gradient tensor reference map. The Earth's gravity field model was expanded to the 2190th order, with a spatial resolution of approximately 5 arc-minutes. The results are shown in Fig.1. to Fig.6.

Fig. 1. Calculation results of disturbance gravity gradient T_{xx} component

Fig. 2. Calculation results of disturbance gravity gradient T_{xy} component

Fig. 3. Calculation results of disturbance gravity gradient T_{xz} component

Fig. 4. Calculation results of disturbance gravity gradient T_{yy} component

Fig. 5. Calculation results of disturbance gravity gradient T_{yz} component

Fig. 6. Calculation results of disturbance gravity gradient T_{zz} component

Given that the perturbed gravity potential satisfies the Laplace equation, i.e., $T_{xx} + T_{yy} + T_{zz} = 0$. To validate the reliability and accuracy of the perturbed gravity gradient map constructed in this study, the sum of the co-axial components of the gravity gradient in the constructed sea area (i.e., T_{xx}, T_{yy}, and T_{zz}) was calculated, and the results are shown in Fig.7. Statistical analysis of the calculation results is presented in Table I.

Fig. 7. Calculation results of the sum of coaxial components of perturbation gravity gradient

<div align="center">

TABLE I. STATISTICAL ANALYSIS OF THE SUM OF COAXIAL COMPONENTS OF PERTURBATION GRAVITY GRADIENT

</div>

Number	Statistical subjects	Statistical value (E)
1	Maximum value	1.4×10^{-8}
2	minimum value	-1.3×10^{-8}
3	Peak to peak value	2.7×10^{-8}
4	average value	6.6×10^{-12}

From Fig.7 and Table I, it can be seen that the sum of the co-axial components of the perturbed gravity gradient fluctuates around zero with a very small range of variation. The difference between the maximum and minimum values is only 2.7×10^{-8}E. Therefore, the perturbed gravity gradient map calculated in this study satisfies the Laplace condition at any grid point in the sea area. Hence, the perturbed gravity gradient reference map obtained in this paper provides a reliable background for gravity gradient navigation simulation studies.

IV. ADAPTABILITY ANALYSIS OF GRAVITY GRADIENT MATCHING

The positioning accuracy of gravity gradient-assisted navigation is closely related to the characteristic distribution of the navigation area. In regions with significant gravity gradient variations, better positioning accuracy can usually be achieved. The characteristic parameters of gravity gradient distribution are typically used to describe the degree of gravity gradient variation or the richness of gravity gradient features in a region. Currently, research on gravity gradient-assisted navigation is not yet mature, so adaptability analysis for gravity gradients draws inspiration from characteristic parameters of gravity-assisted navigation, including standard deviation, roughness, information entropy, correlation coefficient, and average change. These parameters characterize different attributes such as the local gravity gradient distribution fluctuations, smoothness, and discreteness, reflecting the richness of information in the local gravity gradient distribution. Therefore, one or multiple characteristic parameters can serve as effective means for selecting navigation-adaptive zones. Assuming the size of the gridded region of gravity gradient distribution is m×n (where m is the number of rows in the data matrix and n is the number of columns), $T(i,j)$ represents the gravity gradient value at coordinate (i,j), the calculation formulas for each characteristic parameter are as follows.

A. Standard Deviation

The standard deviation reflects the dispersion of data samples, and the specific calculation formula is:

$$\begin{cases} \bar{T} = \dfrac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}T(i,j) \\ \sigma = \sqrt{\dfrac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\left[T(i,j)-\bar{T}\right]^2} \end{cases} \quad (11)$$

Where σ is the standard deviation of the gravity gradient distribution.

B. 3.2 Roughness

Roughness is used to describe the average smoothness of the trend surface of the gravity gradient distribution, capturing finer local fluctuations. The larger the value, the more pronounced the gravity gradient changes in the region. The specific calculation formula is:

$$\begin{cases} r_x = \dfrac{1}{(m-1)n}\sum_{i=1}^{m-1}\sum_{j=1}^{n}\left|T(i+1,j)-T(i,j)\right| \\ r_y = \dfrac{1}{m(n-1)}\sum_{i=1}^{m}\sum_{j=1}^{n-1}\left|T(i,j+1)-T(i,j)\right| \\ r = \sqrt{r_x^2 + r_y^2} \end{cases} \quad (12)$$

Where r_x and r_y are the roughness of the gravity gradient distribution in the x and y directions respectively, and r is the roughness of the gravity gradient distribution.

C. Information Entropy

Information entropy is defined as the probability of occurrence of discrete random events, where a higher probability implies greater uncertainty and thus a higher entropy value. Therefore, a lower value represents a more unique data value and a higher level of information richness. The specific calculation formula is:

$$H = -\sum_{i=1}^{m}\sum_{j=1}^{n}p_{ij}\log_2 p_{ij} \quad (13)$$

Where H is the information entropy of the gravity gradient distribution, and p_{ij} is the probability of $T(i,j)$ occurring in a local region of the gravity gradient distribution.

D. Correlation Coefficient

The correlation coefficient reflects the correlation between data. The closer its value is to 0, the poorer the correlation between the data, indicating more drastic changes in the gravity field in that data region. Conversely, the closer its value is to 1, the stronger the correlation between the data, indicating smoother changes in the gravity field in that data region. The specific calculation formula is:

$$\begin{cases} R_x = \dfrac{1}{(m-1)n\sigma^2}\sum_{i=1}^{m-1}\sum_{j=1}^{n}\left[T(i+1,j)-\bar{T}\right]\left[T(i,j)-\bar{T}\right] \\ R_y = \dfrac{1}{m(n-1)\sigma^2}\sum_{i=1}^{m}\sum_{j=1}^{n-1}\left[T(i,j+1)-\bar{T}\right]\left[T(i,j)-\bar{T}\right] \\ R = \dfrac{1}{2}(R_x + R_y) \end{cases} \quad (14)$$

Where R_x and R_y are the correlation coefficients of the gravity gradient distribution in the x and y directions respectively, and R is the correlation coefficient of the gravity gradient distribution.

E. Average Change

The average change describes the average difference in gravity gradient values between adjacent grid points in the gravity gradient data. A larger value indicates a more pronounced change in the gravity gradient distribution in the region. The specific calculation formula is:

$$A = \dfrac{1}{2}(r_x + r_y) \quad (15)$$

Where A is the average change in the gravity gradient distribution.

F. Feature Parameter Calculation

The characteristic parameter values of the gravity gradient different components distribution in the sea area were calculated according to the above formula, and the results are shown in Table II.

TABLE II. THE CHARACTERISTIC PARAMETER VALUES OF THE DISTRIBUTION OF DIFFERENT COMPONENTS OF GRAVITY GRADIENT

Gravity Gradient Component	T_{xx}	T_{xy}	T_{xz}	T_{yy}	T_{yz}	T_{zz}
Average Value(E)	-0.1	0.0	-0.4	-4.6	4.7	8.8
Peak to peak Value(E)	185.1	95.6	209.4	164.9	208.1	271.2
Standard Deviation(E)	13.6	8.3	16.5	13.8	16.5	23.3
Roughness(E)	2.9	1.9	3.6	2.9	3.6	5.1
Information Entropy(bit)	11.82	11.56	11.75	11.68	11.73	11.89
Correlation Coefficient	0.9733	0.9709	0.9724	0.9722	0.9713	0.9619
Average Change(E)	2.0	1.3	2.5	2.0	2.5	3.6

This indicates that several characteristic parameter values of the perturbed gravity gradient Γ_{zz} component were the best among the six components. This suggests that the vertical component contains the richest features of the Earth's gravity field, with matching features far superior to the other five components. It can be considered the most important matching component in gravity gradient-assisted navigation. Therefore, in the research of gravity gradient-assisted navigation, the Γ_{zz} component should be assigned the highest weight.

V. CONCLUSION

Through analysis, it was found that the perturbed gravity gradient is crucial for calculating the gravity gradient. The perturbed gravity gradient is the second derivative of the perturbed potential, thus deriving the formula for calculating the six components of the perturbed gravity gradient by deriving the second derivative of the perturbed potential in the local north-oriented coordinate system. Subsequently, a benchmark diagram of the full-tensor perturbed gravity gradient was constructed for the sea area ranging from longitude 112°E to 117°E and latitude 12°N to 17°N. The correctness of the constructed gravity gradient benchmark diagram was verified through Laplace's constraint. Statistical analyses such as standard deviation, roughness, information entropy, correlation coefficient, and average change were conducted on the benchmark diagram of the perturbed gravity gradient in the selected area. It was clearly identified that the Γ_{zz} component in the gravity gradient contains the richest information, providing data support for the weight allocation of the matching quantity in gravity gradient-assisted navigation.

REFERENCES

[1] Koneshov V N, Nepoklonov V B. Studying the Representation Accuracy of the Earth's Gravity Field in the Polar Regions Based on the Global Geopotential Models[J].Izvestiya Physics of the Solid Earth, 2018, 54(3):504-512.

[2] Pavlis N.K., Holmes S.A., Kenyon, S.C., et al., The development and evaluation of the Earth Gravitational Model 2008 (EGM2008)[J]. Geophys. Res., 2012, 117(B16):4406.

[3] Jia Y U, Li-Hua C, Yi-Kai F. Research on Method for Offshore Elevation Transmission Based on EGM2008 Model[J].Coastal Engineering, 2019.

[4] Pavlis N.K., Factor J.K., and Holmes S.A., Terrain-related gravimetric quantities computed for the next EGM, Proc. 1st Int. Symposium of the International Gravity Field Service (Istanbul, 2006), Harita Dergisi, 2007, 18:318–323.

Increasing the Navigation Safety of an Autonomous Underwater Vehicle during Observation in Ice Conditions

L.A. Martynova
Concern CSRI Elektropribor, JSC
Saint Petersburg, Russia
martynowa999@bk.ru

I.V. Pashkevich
Concern CSRI Elektropribor, JSC
Saint Petersburg, Russia
iv999@bk.ru

Abstract—**The problem of searching for a water opening in the ice conditions by an autonomous underwater vehicle for its observation for safety purposes is considered. Due to the time-accumulated error in the navigation data of the on-board inertial navigation system of the vehicle, observation by signals from satellite navigation systems is required, since the vehicle can enter navigation danger zones due to deviation from the given route. The difficulty of finding an ice hole lies in the constant drift of ice and the influence of the environment on its formation and condition. In this case, the time it takes for the vehicle to find a hole of suitable size affects its safety. The traditional use of an echo sounder is ineffective due to the narrow inspected zone. To search for an ice hole, it is proposed to use computer vision equipment located in the lower part of the vehicle body and intended for examining the sea bottom. The use of this equipment for examining the lower surface of the ice can significantly increase the width of the inspected zone. The authors present an algorithm for examining the ice using a wide range of computer vision equipment and give an estimate of the reduction in time for finding an ice hole of the suitable size. The example shows the advantage of the proposed solution, which accelerated the inspection of ice and the path traveled by the vehicle by more than 10 times, that, in turn, leads to significant vehicle energy saving. To transfer computer vision equipment from the lower part of the vehicle body to the upper part, the authors propose a method for controlled change of roll using four ballast tanks specially equipped on board the vehicle. An algorithm for controlled roll change is presented. Using the example of the Hugin vehicle (Norway), based on the results of simulation modeling, a controlled change in the vehicle roll and its transfer to the "overkeel" position is shown. The proposed solution makes it possible to examine, along with the bottom topography, the lower part of the ice surface and inclined surfaces of obstacles arising in front of the vehicle, which helps to increase its safety.**

Key words— autonomous underwater vehicle, ice conditions, roll change control, overkeel position, computer vision equipment.

I. Introduction

At present, the research of sea ice in the Arctic and Antarctic zones is relevant because in the Arctic zone there is a shelf with large reserves of hydrocarbons and the Northern Sea Route; in the Antarctic zone, changes in the state of a great ice swarm due to warming have a significant impact on the climate of the entire planet. Marine polar research using autonomous underwater vehicles (AUVs) allows us to better understand and assess the impact of processes occurring in the Arctic and Antarctic on climate change on the Earth and the marine ecosystem.

In the Arctic zone, the underwater observation by a single AUV [1] or a group [2] requires periodical surfacing of the vehicle to carry out observations by signals from satellite navigation systems (SNS), conduct a communication session with a coastal point or an escort ship in order to ensure its navigational safety. Observation of AUV by SNS signals in ice conditions is only possible either by "attaching" a special antenna to the lower edge of the ice, or by surfacing in an ice hole [3].

The observation is planned in the route mission of the AUV for the purpose of its safety: its deviation from the given route path can lead to the situation when the vehicle gets into navigation danger zones or fishing nets, collides with obstacles, etc. Timely observation helps to increase navigational safety of the vehicle.

In this regard, for the AUV it is necessary to survey the ice in order to find a suitable site for attaching an antenna or an ice hole during the predetermined period of time for observation [4-8]. The navigational safety of the AUV depends on how quickly a hole suitable for its surfacing is found.

Searching for an ice hole consists in examining the lower edge of the ice and estimating the size of the discovered hole for the safe surfacing of the AUV. The minimum size of the hole that can be considered safe for the AUV surfacing should be five to ten times its length, and even more in unfavorable conditions.

Usually [4], to survey the ice, the upper part of the AUV body is equipped with a single-beam or multi-beam echo sounder (MBE) or an under-ice sonar. That is why to estimate the size of the ice hole and determine the surfacing point, the AUV has to perform maneuvering on several tacks along and across the discovered ice hole. This leads to significant time and energy costs.

When examining the ice, a profilograph (PF) is also used to assess the thickness and structure of the ice.

The observation is difficult because it is impossible to predict the ice hole position in advance due to the continuous drift of ice and the influence of weather conditions on its formation and state. At the same time, ensuring the navigational safety of the AUV significantly depends on the time spent on detecting the ice hole and its size, since its location quickly changes when the ice moves. In this regard, the goal of the work was to reduce the time of finding the sites for observation – a site for attaching an antenna or an ice hole for the AUV to surface.

II. Reducing the time to seacrh for an ice hole

Currently, to survey the bottom, AUV are equipped with a wide range of computer vision equipment (CVE), using

various physical principles of operation and applied to search for objects on the seabed or in its mass [9]. Along with MBE and PF, side-scan sonars (SSS) and video cameras are used. The use of SSS for examining the lower surface of the ice allows us to significantly increase the width of the inspected zone (Fig. 1). In addition to the listed CVE, it is also proposed to use a thermal imager and a Doppler log directed vertically upward towards the lower surface of the ice.

Fig. 1.   Multual position of zones being observed by MBE and SSS

The operation of MBE and SSS is based on assessing the difference in the delays of high-frequency and low-frequency components of sonar signals reflected from the ice-air and water-ice interfaces. The MBE examines the surface directly above itself; the SSS examines the surface on the sides (Fig. 1). The combined use of MBE and SSS makes it possible to increase the areal productivity of surveying the lower ice surface by increasing the width of the surveyed zone from 160 m to 800 m as compared to the traditionally used echo sounder.

Optical equipment (video cameras) can detect an ice hole by recognizing images on a video picture: a bright spot on a darker background of pack ice during daylight hours. In the night time a thermal imager can be used to find an ice hole by the temperature difference between the water in the hole and the surrounding ice.

According to emission/reception of low-frequency acoustic signals the PF estimates the ice thickness and structure (packed/rotten ice) to determine the possibility of its safe punching by the AUV without damaging its body and outboard instruments and actuators protruding beyond the body.

In addition to the listed CVE, it is also proposed to use a Doppler log, directed vertically upward towards the lower surface of the ice, to estimate the speed of ice movement and take it into account when planning the point of safe surfacing of the AUV in the ice hole.

The additional use of a profilograph, video camera, and thermal imager will provide more accurate information about the age of ice, its structure and mechanical properties.

The use of various CVE will not only significantly reduce the time for examining the ice, but also predict changes in the state and position of the ice over time.

III.   THE ICE SOUNDING PROCEDURE

If observation is required, the ice sounding is carried out, first of all, by the most productive CVE, in terms of the sur-

veyed area, – MBE and SSS. They are capable of forming a hydroacoustic map of the ice relief at a distance of 200 m from it; the swath width when MBE and SSS are used simultaneously ranges from 160 m to 800 m (Fig. 1).

When a fragment of the ice surface, looking like a platform or an ice hole, is detected, the AUV approaches this fragment and turns on the PF, video camera and thermal imager. The video camera data specify the classification criteria of the ice hole (light spot, temperature difference between water and ice); the PF data are used in estimating the ice structure: if the ice is thin or loose, breaking through the ice will be safe for the AUV. The combined use of all the listed CVE also makes it possible to quickly estimate the size of the ice hole. The Doppler log will estimate the speed of ice movement and reasonably determine the AUV surfacing point for safe observation according to the algorithm given in [3].

To evaluate the benefit from the proposed solution, consider the method of finding an ice hole [3] (Fig. 2).

Fig. 2.   Trajectory of AUV motion when using a traditional single echo sounder for searching for an ice hole

From Fig. 2 it follows that with the use of an under-ice sonar the AUV had to travel 8 tacks from the search start point, that is, to examine a total zone with a width of $17 \times 8 = 136$ m. Using the solution proposed in this paper, the same result can be obtained in one AUV pass with a swath width of 200 m, provided by the combined use of MBE and SSS.

Thus, the gain along the distance traveled, taking into account transverse tacks and maneuvering when moving from tack to tack, can be no less than 10 times. Saving the distance traveled leads to saving time `for examination and saving the AUV energy resource spent on examination. This, in turn, makes it possible to improve the safety of AUV surfacing for observation and, hence, enhance the navigational safety of AUV.

Installation of CVE in the upper part of the body is possible by duplicating the CVE located in the lower part of the body. Another method is a controlled change of the AUV roll, wherein the CVE located in the lower part of the AUV body, after the AUV rolls, will be directed towards the surface being examined.

IV. CONTROLLED CHANGE OF AUV ROLL

The change in the AUV roll occurs due to origination of the moment M_x, shown in Fig. 3 in the O_{xyz} coordinate system with the center located at the point of the center of buoyancy (CB) of the AUV, to which the buoyancy force (Archimedes force) is applied.

Fig. 3. Position of rolling moment M_x.

To create a rolling moment M_x we assume that the AUV has cross-shaped or X-shaped stern hydroplanes; the AUV body is a figure of rotation about the longitudinal axis to ensure uniform distribution of the AUV body and equipment mass; devices inside the AUV body are arranged so that their mass is distributed uniformly in the AUV body.

Let's consider an AUV equipped with four ballast tanks 1,2,3,4 (Fig. 4), whose centers of gravity (CG) are located in a plane transverse to the centerline plane, and the ballast tanks themselves are located as follows: one at the bottom of the body, the second – on the starboard side, the third – at the top of the body, the fourth – on the port side; the ballast tanks are connected to each other by pipelines for distilling water from one tank to another using pumps in order to maintain constant buoyancy of the AUV. Some cargo is installed so that it can be moved vertically, for example, freely suspended in the center of the AUV.

Fig. 4. Position of ballast tanks (at the left – before rolling).

Let the initial value of the AUV roll angle before roll $\theta_0=0$; the preset value of the roll angle $\theta^*=180°$. At the initial moment of time, the current value of the roll angle agrees with the initial value.

The AUV roll algorithm is as follows.

1. Calculation of the mismatch $\Delta\theta$ between the current roll angle θ and the preset one θ^*.

2. Calculation of the required rolling moment Mx^*.

3. Determination of ballast tanks for water distillation in order to create a given rolling moment Mx^* in a preset direction.

4. Distillation of water between tanks, taking into account restrictions on the distilling rate.

5. Calculation of the rolling moment Mx, which occurs under the influence of: hydrodynamic forces, the force of initial (operational) excess buoyancy and the buoyancy force

resulting from compression of the AUV body [10], the buoyancy of ballast tanks.

6. Calculation of the coordinates of the new position of the AUV gravity center.

7. Calculation of the metacentric height of the AUV.

8. Calculation of the moment of inertia of the AUV.

9. Calculation of changes in the angular rate of the AUV body rotation.

10. Calculation of roll change rate.

11. Determination of the current roll angle θ.

12. Go to step 1.

The described algorithm for changing the AUV roll is repeated by the AUV control system [11,12] until the current value θ matches with the specified value θ^*.

Figure 5 shows the results of simulation modeling [13] of water distillation in tanks for a controlled change in AUV roll according to the proposed algorithm.

Fig. 5. Sequence of filling-drying the ballast tanks after water distillation.

When carrying out calculations, the "Hugin" (Norway) with the following characteristics was adopted as the AUV prototype [14]: weight 750 kg; length H = 4.50 m; diameter D = 0.75 m. The volume of each ballast tank was taken equal to 0.008 m³.

At the initial instant, the AUV is in the steady state; the filling of the tanks is shown in Fig. 4. The moment of inertia of the AUV was Ix=52.7 kg·m², the total underwater displacement was V=5.3 m³. As water is pumped from ballast tank 1 to ballast tank 2, and then from ballast tank 2 to ballast tank 3, a controlled change in the AUV roll occurs. After the AUV is transferred to the "overkeel" position, to give the AUV a steady state, the load from the AUV gravity center moves to the lowest position (Fig. 5).

Based on the results of simulation modeling, the dependences of the change in rolling moment Mx over time (Fig. 6 on the left) and the change in the roll angle over time (Fig. 6 on the right) were obtained.

The results presented in Fig. 6 show that the change in rolling moment Mx first increases smoothly, then reaches its maximum and then smoothly decreases. As a result, the roll gradually increases from 0° to +180°, and at $\theta = +180°$ the rolling moment and angular rate take zero values (Fig. 6 on the left). The function of the roll change from 0° to +180° is

close to the form of hysteresis (Fig. 6 on the right). The roll change time from 0° to +180° was 1.5 hours.

Fig. 6. Change of rolling moment (left) and roll angle (right) over time

The obtained simulation modeling results confirm the feasibility and reliability of the proposed method of AUV roll.

V. CONCLUSION

A method for examining the ice has been developed by using a wide range of different computer vision tools used to examine the bottom and search for objects on the bottom and in its mass. The procedure of using the computer vision equipment to examine the ice has been determined. The benefit from the proposed solution was assessed in comparison with the existing approach to searching for and inspecting the ice hole using a single- or multi-beam echo sounder. The resulting tenfold reduction in the inspection time and the distance traveled leads to an increase in the navigational safety of the autonomous underwater vehicle and saving its energy. To transfer the computer vision equipment from the AUV keel location to the top, a method for controlled change of the AUV roll to the specified roll angle is proposed, ensuring a stable position of the AUV.

REFERENCES

[1] L. A. Martynova, "Solution to the problem of underwater observation under intelligent noise conditions," Informatsionno-upravlyayushchie sistemy, no. 1, 2018, pp. 31–41.

[2] L. A. Martynova, and M. B. Rozengaus, "Determination of the effective behavior of an AUV group in the network-centric illumination system of the underwater environment," Informatsionno-upravlyayushchie sistemy, no. 3 (88), 2017, pp. 47–57.

[3] K. Z. Laptev, and A. V. Bagnitsky, "The task for long-range AUVs to search for polynyas when sailing in the Arctic regions," Izvestia of the Southern Federal University. Technical science. no. 1 (203), 2019. pp. 28-38. DOI 10.23683/2311-3103-2019.

[4] Yu. A. Bozhenov "The use of autonomous uninhabited underwater vehicles for the study of the Arctic and Antarctic," Fundamental and Applied Hydrophysics, vol. 4, no. 1, 2011, pp. 47–68.

[5] A. V. Bagnitsky, A. V. Inzartsev, and M. A. "Panin Algorithmic support for AUV operation under ice in the polar regions," Marine intelligent technologies. No. 3 part 1, 2023, pp. 98–109.

[6] A. V. Inzartsev, L. V. Kiselev, V. V. Kostenko, Yu. V. Matvienko, A. M. Pavin, and A.F. Shcherbatyuk, Underwater robotic systems: systems, technologies, applications, responsible editor L. V. Kiselev, FGBUN Institute for Problems of Marine Technologies, Far Eastern Branch of the Russian Academy of Sciences, Vladivostok, 2018.

[7] Pyatkin V. P., and Salov G. I. "Detection of polynyas," Collection of articles based on materials from the international scientific congress "Inter-expo Geo-Siberia", Novosibirsk, April 13–25 2015, pp.77–82.

[8] K. Z. Laptev, and A. V. Bagnitsky "Algorithms for choosing polynyas and places for glaciation of anpa when working in high latitudes," Eighth All-Russian Scientific and Technical Conference "Technical Problems of the Development of the World Ocean" TOPOMO-8. Conference materials, IPMT, September 30 – October 3, 2019, pp. 43–49.

[9] G. G. Bezruk, L. A.Martynova, and I. B. Saenko "Dynamic method for searching for anthropogenic objects in the seabed using autonomous uninhabited underwater vehicles," Proceedings of SPIIRAS, no. 3 (58), 2018, pp. 203–226.

[10] M. D. Ageev, L. V. Kiselev, and Yu. V. Matvienko and others, under general ed. M. D. Ageva, Autonomous underwater robots: systems and technologies Moskow, Nauka, 2005.

[11] L. A. Martynova, N. K. Kiselev, and A. A. Myslivy "Method for selecting the architecture of a multi-agent control system for an autonomous uninhabited underwater vehicle," Information and control systems. Informatsionno-upravlyayushchie sistemy, no.4 (107), 2020, pp. 31–41.

[12] L. A. Martynova, A. I. Mashoshin, I. V. Pashkevich, A. I. Sokolov "Algorithms implemented by the integrated AUV control system," News of the Southern Federal University, Technical science, no. 1 (162), 2015, pp. 50–58.

[13] L. A. Martynova, A. V. Grinenkov, A. O. Pronin, and Yu. V. Kulikovskikh "Simulation modeling of the functioning of a multi-agent control system for an autonomous underwater vehicle," In the collection: Simulation modeling. Theory and practice. Eighth All-Russian scientific and practical conference on simulation modeling and its application in science and industry, pp. 474–479, 2017.

[14] Kraft A. Nothing will happen: neither minesweepers, nor divers - just solid robots [Electronic resource]. URL: https://andrej-kraft.livejournal.com/142900.html.

Vehicle vector gravimetry method based on SINS/GNSS/LDV integrated system

Chengzhi Hou

College of Advanced Interdisciplinary Studies, National University of Defense Technology
Nanhu Laser Laboratory, National University of Defense Technology
Changsha, China
houchengzhi@alumni.nudt.edu.cn

Mailun Chen

College of Advanced Interdisciplinary Studies, National University of Defense Technology
Nanhu Laser Laboratory, National University of Defense Technology
Changsha, China
18874702536@163.com

Xu Zhu

College of Advanced Interdisciplinary Studies, National University of Defense Technology
Nanhu Laser Laboratory, National University of Defense Technology
Changsha, China
17333151763@163.com

Xudong Yu

College of Advanced Interdisciplinary Studies, National University of Defense Technology
Nanhu Laser Laboratory, National University of Defense Technology
Changsha, China
wind0909@163.com

Chunfeng Gao

College of Advanced Interdisciplinary Studies, National University of Defense Technology
Nanhu Laser Laboratory, National University of Defense Technology
Changsha, China
neil1989@126.com

Wenjian Zhou

College of Advanced Interdisciplinary Studies, National University of Defense Technology
Nanhu Laser Laboratory, National University of Defense Technology
Changsha, China tom7850@foxmail.com

Jiayi Cheng

College of Advanced Interdisciplinary Studies, National University of Defense Technology
Nanhu Laser Laboratory, National University of Defense Technology
Changsha, China
18841218702@163.com

Guo Wei*

College of Advanced Interdisciplinary Studies, National University of Defense Technology
Nanhu Laser Laboratory, National University of Defense Technology
Changsha, China
nudtweiguo@163.com

Abstract—**With the development of inertial devices, gravity disturbance has gradually become the main reason restricting the accuracy of high-precision inertial navigation. Because the measurement of gravity disturbance is greatly affected by the attitude, an aircraft with smooth flight is mainly used as the measurement carrier of the disturbance vector, that is, the airborne gravimetry. Due to the influence of the height and speed of the carrier, the resolution of airborne gravimetry is low, and the gravity results need to be extended to the ground, the accuracy will also be reduced, which is not conducive to the construction of the earth's gravity field information. In this paper, aiming at the problems encountered in traditional gravity measurement, combined with the Strapdown Inertial Navigation System (SINS) and Laser Doppler Velocimeter (LDV) developed by our team, a vehicle-borne gravity vector measurement method based on SINS/GNSS/LDV integrated system is proposed. By analyzing the gravity error model, the requirements of the equipment accuracy are obtained, and the integrated system is built for experiments. Without post smoothing, the internal coincidence accuracy of results of vehicle-borne gravity vector measurement are within 3.3mGal/1.7km. Therefore, the feasibility of the method is illustrated.**

Keywords—vehicle-borne gravity vector measurement, integrated navigation, internal coincidence accuracy

I. INTRODUCTION

In the inertial navigation solution, the earth's gravity field is usually based on the normal gravity field determined by satellite altimetry. This method ignores the difference between normal gravity and real gravity, that is, gravity disturbance. The gravity disturbance is called deflection of the vertical (DOV) in the direction, which is the deviation between the normal line of the reference ellipsoid and the vertical line of the geoid [1]. The horizontal component of the normal gravity field is zero, but the projection of real gravity in the horizontal direction of the reference ellipsoid is usually non-zero. Therefore, the horizontal component of the gravity disturbance is equivalent to the gravity vector.

Inertial Navigation System (INS) adopts the positioning method of dead reckoning. Due to the gravity disturbance, INS is not error-free even with non-drift gyros and perfect accelerometers [2,3]. In recent years, with the significant improvement of INS, especially in inertial sensors, horizontal gravity disturbance has become the most important error source in navigation solutions [4]. According to the different terrain and geology, the DOV can reach the maximum of 100" orders of magnitude, and the magnitude of the error is

significantly higher than the zero bias of modern accelerometers (about $10\mu g$, corresponding to the DOV of $2''$), which will directly affect the alignment process, and then affect the subsequent positioning [5]. In [6], Gao uses the Back Propagation Neural Network (BPNN) algorithm and combines the highest-order data of the existing spherical harmonic model EIGEN-6C4 to predict the gravity disturbance at the fine position, making up for the lack of resolution. He utilizes the results of gravity disturbance prediction to compensate. Compared with pure inertial solution, the position accuracy is improved by about 26% (1300m/6h). In [7], Tie gives a method to calculate the gravity disturbance by using the EGM2008 gravity field model and the fully normalized Legendre function formula. The calculated results are compared with DOV grid data of National Geospatial Intelligence Agency (NGA), and the average difference can be ignored. He uses the calculated disturbance value to compensate the ship-borne experiment, and the inertial navigation accuracy was increased by 1.5nmile/24h. It can be seen that the compensation of gravity disturbance plays an important role in improving the positioning accuracy.

In fact, there is still no mature gravity vector measurement technology, which is mainly limited by the accuracy of attitude measurement. Assuming that there is an attitude angle error of $1''$, the influence on the horizontal component is as high as 5 mGal [8]. Consequently, airborne gravity vector measurement with low dynamic attitude variation was primarily chosen [9,10]. At present, the disturbance accuracy of airborne gravimetry without post-processing is 7 mGal, and after post-processing, it is 2 mGal. [11]. Considering the closer to the earth's surface, the richer the gravity information of different frequency bands, and the more high-resolution disturbance information can be measured [12]. Therefore, vehicle-borne gravity vector measurement is an effective way to obtain high-precision and high-resolution earth gravity field database.

In this paper, according to the practical needs of gravity vector, a measurement system based on SINS/GNSS/LDV is given. By establishing the model of gravity vector, the measurement error caused by attitude, velocity and acceleration error are decoupled and analyzed. The SINS/GNSS/LDV integrated gravity vector measurement system is constructed, and the information calculation process is designed. Through relevant experiment, it is proved that the designed vehicle-borne scheme can achieve effective gravity vector measurement.

II. ERROR CHARACTERISTIC ANALYSIS OF VEHICLE-BORNE GRAVITY VECTOR MEASUREMENT

By transforming the velocity differential equation of strapdown inertial navigation, the gravity disturbance can be expressed as:

$$\delta \boldsymbol{g}^n = \dot{\boldsymbol{v}}_e^n - \boldsymbol{C}_b^n \boldsymbol{f}^b + (2\boldsymbol{\omega}_{ie}^n + \boldsymbol{\omega}_{en}^n) \times \boldsymbol{v}_e^n - \boldsymbol{\gamma}^n \qquad (1)$$

By expanding the vector of (1), the form of components can be obtained:

$$\delta g_E = \dot{v}_E - f_E + (2\omega_{ie} \cos L + \frac{v_E}{R_N + h}) \cdot v_U$$
$$- (2\omega_{ie} \sin L + \frac{v_E}{R_N + h} \tan L) \cdot v_N$$

$$\delta g_N = \dot{v}_N - f_N + (2\omega_{ie} \sin L + \frac{v_E}{R_N + h} \tan L) \cdot v_E + \frac{v_N}{R_M + h} \cdot v_U$$

$$\delta g_U = \dot{v}_U - f_U - \left[(2\omega_{ie} \cos L + \frac{v_E}{R_N + h}) \cdot v_E + \frac{v_N^2}{R_M + h} \right] - \gamma_U$$
$$(2)$$

In the expression, δg_E, δg_N, δg_U are the east, north and vertical components of the gravity disturbance vector at the location of the carrier; $\dot{v}_E$, $\dot{v}_N$, $\dot{v}_U$ are the east, north and vertical acceleration of the carrier motion; f_E, f_N, f_U are the east, north and vertical specific force information provided by IMU; ω_{ie} is the angular velocity of the earth's rotation; L is the latitude of the carrier position; v_E, v_N, v_U are the east, north and vertical speed of the carrier motion; R_M, R_N are the radius of the meridian circle and the radius of prime vertical circle of the carrier position; h is the height of the carrier position; γ_U is the normal gravity value calculated according to different gravity models.

The mathematical equation of gravity disturbance is differentiated and transformed, and the error model of gravity disturbance can be obtained as follows:

$$d\delta \boldsymbol{g}^n = \delta \dot{\boldsymbol{v}}_e^n - \boldsymbol{C}_b^n \delta \boldsymbol{f}^b - [\boldsymbol{f}^n \times]\boldsymbol{\psi} - [\boldsymbol{v}_e^n \times](2\delta \boldsymbol{\omega}_{ie}^n + \delta \boldsymbol{\omega}_{en}^n)$$
$$+ (2\boldsymbol{\omega}_{ie}^n + \boldsymbol{\omega}_{en}^n) \times \delta \boldsymbol{v}_e^n + (\dot{\boldsymbol{C}}_b^n \boldsymbol{f}^b + \boldsymbol{C}_b^n \dot{\boldsymbol{f}}^b)dT - \delta \boldsymbol{\gamma}^n$$
$$(3)$$

The error model includes: term of carrier acceleration measurement error $\delta \dot{\boldsymbol{v}}_e^n$; term of specific force measurement error $\boldsymbol{C}_b^n \delta \boldsymbol{f}^b$; term of attitude measurement error $[\boldsymbol{f}^n \times]\boldsymbol{\psi}$; correlation term of centrifugal acceleration and Coriolis acceleration calculation error caused by earth rotation $[\boldsymbol{v}_e^n \times](2\delta \boldsymbol{\omega}_{ie}^n + \delta \boldsymbol{\omega}_{en}^n)$; correlation term of carrier velocity measurement error $(2\boldsymbol{\omega}_{ie}^n + \boldsymbol{\omega}_{en}^n) \times \delta \boldsymbol{v}_e^n$; correlation term of time synchronization error $(\dot{\boldsymbol{C}}_b^n \boldsymbol{f}^b + \boldsymbol{C}_b^n \dot{\boldsymbol{f}}^b)dT$; calculation error of normal gravity $\delta \boldsymbol{\gamma}^n$.

Through simulation experiments, Cai found that the key error sources of gravity vector measurement are divided into two aspects: the attitude, velocity and acceleration measurement errors in the horizontal direction have a great influence on the horizontal component of the disturbance, and the velocity and acceleration measurement errors in the vertical direction have a great influence on the gravity anomaly [14]. The key error sources are analyzed below, and the requirements of the equipment accuracy of vehicle gravity vector measurement are obtained.

A. Gravity Error Caused by the Attitude

At present, relevant research has analyzed the influence of attitude error in airborne gravimetry, and it is considered that the aircraft is in a stable flight state without lateral acceleration. Different from the airborne gravimetry, there

will be a change in the running speed during the vehicle-borne gravimetry, which will lead to a change in the external acceleration and then affect the change of the specific force measurement value. The horizontal specific force in the carrier operation cannot be regarded as zero. Therefore, the attitude estimation for a vehicle-borne system needs to design simulation experiments. The relative term of attitude error in (3) is $-\left[\boldsymbol{f}^n \times\right]\boldsymbol{\psi}+\left(2\boldsymbol{\omega}_{ie}^n+\boldsymbol{\omega}_{en}^n\right)\times\left(\left[\boldsymbol{v}^n \times\right]\boldsymbol{\psi}\right)$, which is expanded, as shown in (4).

$$
\begin{aligned}
d\delta\boldsymbol{g}^n &= -\left[\boldsymbol{f}^n \times\right]\boldsymbol{\psi}+\left(2\boldsymbol{\omega}_{ie}^n+\boldsymbol{\omega}_{en}^n\right)\times\left(\left[\boldsymbol{v}^n \times\right]\boldsymbol{\psi}\right) \\
&= -\begin{bmatrix} \psi_U f_N - \psi_N f_U \\ \psi_E f_U - \psi_U f_E \\ \psi_N f_E - \psi_E f_N \end{bmatrix} + \begin{bmatrix} \lambda_E \\ \lambda_N \\ \lambda_U \end{bmatrix}
\end{aligned}
$$

$$
\lambda_E = \left(2\omega_{ie}\cos L + \frac{v_E^n}{R_N+h}\right)\times\left(\psi_N v_E^n - \psi_E v_N^n\right)
$$
$$
- \left(2\omega_{ie}\sin L + \frac{v_E^n \tan L}{R_N+h}\right)\times\left(\psi_E v_U^n - \psi_U v_E^n\right)
$$

$$
\lambda_N = \left(2\omega_{ie}\sin L + \frac{v_E^n \tan L}{R_N+h}\right)\times\left(\psi_U v_N^n - \psi_N v_U^n\right)
$$
$$
- \left(-\frac{v_N^n}{R_M+h}\right)\left(\psi_N v_E^n - \psi_E v_N^n\right)
$$

$$
\lambda_U = \left(-\frac{v_N^n}{R_M+h}\right)\left(\psi_E v_U^n - \psi_U v_E^n\right)
$$
$$
- \left(2\omega_{ie}\cos L + \frac{v_E^n}{R_N+h}\right)\left(\psi_U v_N^n - \psi_N v_U^n\right) \tag{4}
$$

The IMU sampling frequency of the integrated navigation system is very high. In a short time, the attitude error is generally within a few arc seconds. After conversion to radian, it is on the order of 10^{-5}, and reaches 10^{-11} after R_M+h and R_N+h attenuation. Thus, (4) can be simplified as:

$$
\begin{aligned}
d\delta\boldsymbol{g}^n &= -\left[\boldsymbol{f}^n \times\right]\boldsymbol{\psi}+\left(2\boldsymbol{\omega}_{ie}^n+\boldsymbol{\omega}_{en}^n\right)\times\left(\left[\boldsymbol{v}^n \times\right]\boldsymbol{\psi}\right) \\
&= -\begin{bmatrix} \psi_U f_N - \psi_N f_U \\ \psi_E f_U - \psi_U f_E \\ \psi_N f_E - \psi_E f_N \end{bmatrix}
\end{aligned} \tag{5}
$$

Unlike airborne gravimetry, it is difficult to maintain uniform motion during vehicle operation, so the horizontal direction specific force cannot be regarded as zero. Therefore, all terms of f_E, f_N will be retained in (5). Assuming that the vehicle is traveling at a speed of $20\ m/s$, $f_U \approx g$, the correspondence between the 1 mGal gravity measurement error and the attitude error can be obtained according to (5) as shown in Table 1.

It can be seen from the Table I, in order to keep the gravity vector measurement error better than 1 mGal, the horizontal attitude error should be kept in the order of seconds. Wang carried out time-frequency analysis of the misalignment angle error propagation equation and concluded that the attitude error is mainly caused by the zero-bias instability of gyro and accelerometer [14]. The bias

instability of accelerometer is mainly affected by temperature, and the bias characteristics will change slowly with time. Before the experiment, the IMU will be fully preheated by continuous power supply, and the integrated navigation system will be in a relatively stable temperature environment in the experiment. After the calibration of IMU bias error, scale factor error and installation error, the influence of accelerometer bias instability on attitude error can be ignored in short time experiment. Therefore, the simplified attitude error model is shown as follows [14]:

$$
\boldsymbol{\psi} = \frac{\boldsymbol{b}_g}{\omega_s}\sin\omega_s t \tag{6}
$$

TABLE I. 1 MGAL VEHICLE-BORNE GRAVITY MEASUREMENT ERROR CORRESPONDS TO ATTITUDE ERROR

Horizontal specific force	Vector measurement	
	Horizontal attitude angle error	*Heading angle error*
0.30g	0.50"	1.65"
	0.60"	1.99"
0.20g	0.50"	2.48"
	0.60"	2.98"
0.10g	0.50"	4.96"
	0.60"	5.96"
0.05g	0.50"	9.92"
	0.60"	11.92"
0.01g	0.50"	49.64"
	0.60"	59.64"

In the expression, gyro drift $\boldsymbol{b}_g$; Schuler angular frequency $\omega_s = \sqrt{g/R} \approx 1.24\times10^{-3}/s$. According to the pre-sequence experiment, the actual horizontal acceleration in the vehicle measurement experiment is about $0.3g$. When the horizontal attitude angle accuracy error is about $0.60"$ and the heading angle error is about $2.98"$, the corresponding gyro accuracy is:

$$
\boldsymbol{b}_g < \omega_s \cdot \boldsymbol{\psi} \approx 0.00074°/h \tag{7}
$$

B. Gravity Error Caused By the Velocity

The calculation error of gravity disturbance vector caused by velocity error is:

$$
d\delta\boldsymbol{g} = (2\boldsymbol{\omega}_{ie}^n+\boldsymbol{\omega}_{en}^n)\times\delta\boldsymbol{v}_e^n \tag{8}
$$

Expand it into (9):

$$
d\delta g_E = (2\omega_{ie}\cos L + \frac{v_E}{R_N+h})\cdot\delta v_U - (2\omega_{ie}\sin L + \frac{v_E\tan L}{(R_N+h)})\cdot\delta v_N
$$

$$
d\delta g_N = (2\omega_{ie}\sin L + \frac{v_E\tan L}{R_N+h})\cdot\delta v_E + \frac{v_N}{R_M+h}\cdot\delta v_U
$$

$$
d\delta g_U = -\left[(2\omega_{ie}\cos L + \frac{v_E}{R_N+h})\cdot\delta v_E + \frac{v_N}{R_M+h}\cdot\delta v_N\right]
$$
$$
\tag{9}
$$

It can be seen from (9) that for gravity measurement, in the north-south survey line, the speed projection of the vehicle is concentrated in the north direction, and the east speed is relatively small. All the terms containing the radius of the meridian circle and the unitary circle will be reduced due to the position of the radius in the denominator. The coupling term itself is on the order of μGal, and the re-multiplication speed error will be further reduced. Therefore, the (9) can be simplified to:

$$d\delta g_E = (2\omega_{ie}\cos L)\delta v_U - (2\omega_{ie}\sin L)\delta v_N$$
$$d\delta g_N = (2\omega_{ie}\sin L)\delta v_E \qquad (10)$$
$$d\delta g_U = -(2\omega_{ie}\cos L)\delta v_E$$

According to the simplified (10), it can be seen that the north and vertical components of gravity disturbance are mainly affected by the east velocity error, and the east component is mainly affected by the vertical and north velocity error. The simulation analysis sets the latitude as 28° and the earth rotation as $\omega_{ie} = 7.29\times10^{-5}\,rad/s$. The results are shown in Table Ⅱ and Table Ⅲ.

TABLE II.   THE INFLUENCE OF EASTWARD VELOCITY ERROR ON THE NORTH AND VERTICAL COMPONENTS (UNIT: MGAL)

Eastward velocity error (m/s)	Disturbance of north component	Disturbance of vertical component
0.01	0.0685	0.1283
0.05	0.3425	0.6415
0.10	0.6850	1.2831
0.20	1.3701	2.5662

TABLE III.   THE INFLUENCE OF NORTHWARD AND VERTICAL VELOCITY ERROR AND ON THE EAST COMPONENT (UNIT: MGAL)

North Error (m/s) \ Vertical error (m/s)	0.01	0.05	0.1	0.2
0.01	0.0598	0.5730	1.2145	2.4975
0.05	0.2142	0.2990	0.9405	2.2235
0.10	0.5567	0.0435	0.5980	1.8810
0.20	1.2417	0.7285	0.0870	1.1960

It can be seen from Table Ⅱ and Table Ⅲ that under the conditions of north-south survey line and latitude 28°, when the east velocity error is $0.1\,m/s$, it will cause about 1 mGal errors to the gravity anomaly. When the vertical velocity error is $0.1\,m/s$ and the north velocity error is $0.05\,m/s$, it will cause an error of about 1 mGal to the east component of the gravity disturbance.

C.   Gravity Error Caused By the Acceleration

The acceleration is usually obtained by the difference of the velocity signal measured by GNSS, and the influence of the error on the gravity disturbance is directly superimposed:

$$d\delta\boldsymbol{g}^n = \boldsymbol{C}_b^n\delta\ddot{\boldsymbol{v}}^b \qquad (11)$$

In order to ensure that the influence of acceleration on gravity disturbance is less than 1 mGal, the calculation error of acceleration in navigation system should be less than 1 mGal, at least. There is a coupling between the errors of velocity and acceleration, and it is necessary to analyze the influence of the two according to the concrete experiment.

III.   CALCULATION METHOD OF GRAVITY VECTOR MEASUREMENT BASED ON SINS/GNSS/LDV INTE-GRATED NAVIGATION SYSTEM

At present, the commonly used GNSS single point horizontal positioning accuracy is about $1.5\,m$, the differential horizontal positioning accuracy is $0.4\,m$, the theoretical velocity measurement accuracy is about $0.03\,m/s$ and the data output frequency is $5Hz$. However, the velocity information obtained by position difference will bring additional high-frequency error [15]. The measurement error will be greater when the signal is occluded. According to the previous experiments of this unit, after smooth filtering processing, the velocity measurement error is about $0.16\,m/s$ and the acceleration measurement error is about $0.082\,m/s^2$. The accuracy of the two-dimensional Laser Doppler velocimeter developed by our team is 0.05% (mileage ten thousand meters error in 5 meters). If the carrier is running at a speed of about $11.1\,m/s$ ($40\,km/h$), the estimated speed accuracy is about $5.6\times10^{-3}\,m/s$, and the external acceleration measurement accuracy is about 0.62 mGal. The data output frequency is $100\,Hz$. Its speed measurement accuracy, external acceleration measurement accuracy and data output frequency are better than GNSS and theoretically meet the requirements of gravity vector measurement for high-precision velocity measurement. In addition, the bias instability of the new high-precision laser gyro used in the SINS developed by our team is better than $0.0005°/h$, and the bias instability of the accelerometer used is better than $10\,\mu g$, which also meets the requirements of high-precision attitude measurement. The system equipped with high-precision equipment can accurately obtain the high-precision gravity vector directly output, which is more conducive to the subsequent processing of the data. Therefore, GNSS is used as position correction, LDV provides velocity information and is used to calculate acceleration, SINS provides attitude information, and constructs SINS/GNSS/LDV integrated system for vehicle gravity vector measurement. The schematic diagram of the vehicle-borne gravity vector measurement system is shown in Fig. 1.

Fig. 1.   Structure diagram of vehicle-brone gravimetry

For gravity measurement, the original data obtained from the experiment can be processed later. The process of calculating the gravity information mainly includes the

calibration of the installation error of the velocimeter, the calculation of the gravity disturbance, and the accuracy evaluation of the multi-line. The gravity data processing flow is shown in Fig. 2.

Fig. 2.   Data processing procedure

IV.   VEHICLE GRAVITY VECTOR MEASUREMENT EXPERIMENT

After selecting the equipment that meets the requirements, install it according to Fig. 3. The GNSS antenna and the LDV are installed in the upper and lower regions of the vertical line of the IMU respectively, and the remaining lever arm errors are compensated according to the need in subsequent data processing. The raw data are synchronized and stored in the host computer using UTC time after each sensor is turned on.

Fig. 3.   Installation diagram of experimental equipment

The zero-bias error, scale factor error and installation error of gyro and accelerometer are calibrated and warmed up in a relatively constant temperature environment before the experiment starts, and the initial attitude error can be ignored by reaching the starting point of the measurement line for a long period of static alignment.

Three repeated line gravity measurement experiments were carried out in Heimi Mountain, Changsha City. The gravity change of the mountain is obvious, which is convenient to eliminate the illusion of high internal coincidence accuracy caused by the small range of gravity change. The length of a single line is about 9 km, and the altitude change is about 328 m .The experimental trajectory is shown in Fig. 4.

Fig. 4.   The diagram of experimental trajectory

The gravity solution of the integrated navigation system SINS/GNSS and SINS/GNSS/LDV is carried out, and the results are shown in Fig. 5 to Fig. 10 and Table IV to Table IX below. Since the integrated navigation system performs relative gravity measurement and the correlation of random errors between different measurements is low, the internal coincidence accuracy is used for evaluation. The calculation method of internal coincidence accuracy is (12) and (13), and the value is usually used to describe the stability of the system and the repeatability of the measured data [16].

$$\varepsilon_j = \pm\sqrt{\frac{\sum_{i=1}^{n}\delta_{ij}^2}{n}},(j=1,2,...,m) \qquad (12)$$

$$\delta_{ij} = F_{ij} - F_i(i=1,2,...,n; j=1,2,...,m)$$

$$\varepsilon_{total} = \pm\sqrt{\frac{\sum_{j=1}^{m}\sum_{i=1}^{n}\delta_{ij}^2}{m \times n}} \qquad (13)$$

ε_j is the internal coincidence accuracy of the line j ; δ_{ij} is the difference between the measured value of the position point i on the repeated line j and the average of all measured value on the same position point; n is the number of points of repeated lines; m is the number of repeated lines, and ε_{total} is the total internal coincidence accuracy.

V. VEHICLE GRAVITY VECTOR SOLUTION RESULTS

Fig. 5. Measurement results of gravity anomaly based on SINS/GNSS

TABLE IV.   THE INTERNAL COINCIDENCE ACCURACY OF GRAVITY ANOMALY MEASURED BY SINS/GNSS (UNIT: MGAL)

Route	Max	Min	RMS (single)	RMS (total)
Line 1	4.47	-2.58	2.15	
Line 2	1.96	-2.75	1.41	1.62
Line 3	1.93	-2.27	1.14	

Fig. 6. Measurement results of the eastern component of gravity disturbance based on SINS/GNSS

TABLE V.   THE INTERNAL COINCIDENCE ACCURACY OF THE EAST COMPONENT OF GRAVITY DISTURBANCE MEASURED BY SINS/GNSS (UNIT: MGAL)

Route	Max	Min	RMS (single)	RMS (total)
Line 1	10.68	-7.27	6.48	
Line 2	6.62	-3.34	2.34	4.88
Line 3	5.69	-9.15	4.91	

Fig. 7. Measurement results of the northern component of gravity disturbance based on SINS/GNSS

TABLE VI.   THE INTERNAL COINCIDENCE ACCURACY OF THE NORTH COMPONENT OF GRAVITY DISTURBANCE MEASURED BY SINS/GNSS (UNIT: MGAL)

Route	Max	Min	RMS (single)	RMS (total)
Line 1	5.65	-10.22	4.20	
Line 2	2.76	-3.94	1.97	3.17
Line 3	7.57	-6.85	2.94	

After calculation, it can be obtained that the maximum value of the root mean square (RMS) of the internal coincidence accuracy in a single line of gravity anomaly solved by SINS/GNSS is 2.15 mGal, the minimum value is 1.14 mGal, and the total RMS is 1.62 mGal.

The maximum RMS value of the internal coincidence accuracy of the single line of the east component of the gravity disturbance solved by SINS/GNSS is 6.48 mGal, the minimum value is 2.34 mGal, and the total RMS is 4.88 mGal.

The maximum value of the RMS of the internal coincidence accuracy in the single line of the north component of gravity disturbance calculated by SINS/GNSS is 4.20 mGal, the minimum value is 1.97 mGal, and the total RMS is 3.17 mGal.

Fig. 8. Measurement results of gravity anomaly based on SINS/GNSS/LDV

TABLE VII. THE INTERNAL COINCIDENCE ACCURACY OF GRAVITY ANOMALY MEASURED BY SINS/GNSS/LDV (UNIT: MGAL)

Route	Max	Min	RMS (single)	RMS (total)
Line 1	3.02	-2.41	1.80	
Line 2	2.50	-1.91	1.22	1.35
Line 3	2.23	-1.22	0.88	

Fig. 9. Measurement results of the eastern component of gravity disturbance based on SINS/GNSS/LDV

TABLE VIII. THE INTERNAL COINCIDENCE ACCURACY OF THE EAST COMPONENT OF GRAVITY DISTURBANCE MEASURED BY SINS/GNSS/LDV (UNIT: MGAL)

Route	Max	Min	RMS (single)	RMS (total)
Line 1	6.58	-5.16	3.82	
Line 2	2.88	-2.48	1.38	3.24
Line 3	5.34	-6.09	3.86	

Fig. 10. Measurement results of the northern component of gravity disturbance based on SINS/GNSS/LDV

After calculation, it can be obtained that the maximum value of the root mean square (RMS) of the internal coincidence accuracy in a single line of gravity anomaly solved by SINS/GNSS/LDV is 1.80 mGal, the minimum value is 0.88 mGal, and the total RMS is 1.35 mGal.

TABLE IX. THE INTERNAL COINCIDENCE ACCURACY OF THE NORTH COMPONENT OF GRAVITY DISTURBANCE MEASURED BY SINS/GNSS/LDV (UNIT: MGAL)

Route	Max	Min	RMS (single)	RMS (total)
Line 1	3.96	-3.47	1.74	
Line 2	4.80	-4.18	2.70	1.98
Line 3	2.89	-1.86	1.18	

The maximum RMS value of the internal coincidence accuracy of the single line of the east component of the gravity disturbance solved by SINS/GNSS/LDV is 3.82 mGal, the minimum value is 1.38 mGal, and the total RMS is 3.24 mGal.

The maximum value of the RMS of the internal coincidence accuracy in the single line of the north component of gravity disturbance calculated by SINS/GNSS/LDV is 2.70 mGal, the minimum value is 1.18 mGal, and the total RMS is 1.98 mGal.

VI. CONCLUSIONS

By comparing the results of different integrated methods, it can be found that the gravity disturbance vector measured in this experiment has certain repeatability. The velocity and external acceleration measurement accuracy of LDV is higher than that of differential GNSS, so the internal coincidence accuracy of SINS/GNSS/LDV integrated system solution is also better. The accuracy of the inertial device used is very high, so the measurement results of the traditional method are also good. Compared with the SINS/GNSS solution, the gravity anomaly is increased by about 0.27 mGal (17%), the east component of the gravity disturbance is increased by about 1.64 mGal (34%), and the north component of the gravity disturbance is increased by about 1.19 mGal (38%).

Since the accelerometer is sensitive to temperature, the change of temperature will affect the measurement of specific force which will affect the calculation of gravity disturbance. In this experiment, the total internal coincidence accuracy of the three directions of gravity disturbance vector is better than 3.3 mGal without strict system temperature control and low frequency error compensation of data processing (the influence of attitude error and residual accelerometer bias on gravity disturbance mainly exist in the low frequency stage, which has been analyzed in [16,17]), which shows the feasibility of SINS/GNSS/LDV integrated vehicle gravity vector measurement.

REFERENCES

[1] L. Chang, F. Qin, and M. Wu, Gravity disturbance compensation for inertial navigation system, IEEE Transactions on Instrumentation and Measurement, vol. 68, no. 10, pp. 3751-3765, 2018.

[2] J. Wang, G. Yang, X. Li, and X. Zhou, Application of the spherical harmonic gravity model in high precision inertial navigation systems, Measurement Science and Technology, vol. 27, no. 9, pp. 095103, 2016.

[3] X. Zhou, G. Yang, J. Wang, and J. Li, An improved gravity compensation method for high-precision free-INS based on MEC--BP--AdaBoost, Measurement Science and Technology, vol. 27, no. 12, pp. 125007, 2016.

[4] C. Jekeli, Nga gravity support for inertial navigation, Bulletin Géodésique, vol.69, pp. 1-11, 1994.

[5] P. Kopcha, Airborne vector gravimetry using precise, position-aided inertial measurement units, Proceedings of the 60th Annual Meeting of The Institute of Navigation (2004), pp. 497-504, 2004.

[6] D. Gao, B. Hu, F. Qin, L. Chang, and X. Lyu, A real-time gravity compensation method for INS based on BPNN, IEEE Sensors Journal, vol. 21, no. 12, pp. 13584--13593, 2021.

[7] J. Tie, M. Wu, S. Cai, K. Zhang, and J. Lian, Gravity disturbance calculation method based on Earth Gravitational Model 2008, Journal of Chinese Inertial Technology, vol. 25, no. 5, pp. 624--629, 2017.

[8] D. Guo, Research on airborne vector gravity data processing and error compensation. Chang'an University, Xi'an, Shanxi, P.R.China, 2018.

[9] H. Tu, K. Liu, H. Sun, Q. Cui, Y. Yuan, S. Liu, J. He, L. Liu, Study on the CHZ-II Gravimeter and Its Calibrations along Forward and Reverse Overlapping Survey Lines, Micromachines, vol. 13, no. 12, pp. 2124, 2022.

[10] V. Vyazmin, A. Golovan, IMU and GNSS Postprocessing for High-Resolution Strapdown Airborne Gravimetry, Engineering Proceedings, vol. 54, no. 1, pp. 41, 2023.

[11] T. Cai, K. Fang, A New Method for Determining the Gravity Disturbance Vector in Strapdown Airborne Gravimetry, IEEE Transactions on Instrumentation and Measurement, vol. 71, pp. 1-7, 2022.

[12] R. Yu, S. Cai, M. Wu, J. Cao, and K. Zhang, An SINS/GNSS Ground Vehicle Gravimetry Test Based on SGA-WZ02, Sensors, vol. 15, pp. 23477-23495, 2015.

[13] S. Cai, The Research about Airborne Vector Gravimeter and Methods of Errors Separation. National University of Defense Technology, Changsha, Hunan, P.R.China, 2014.

[14] Z. Wang, Research on the airborne vector gravimetry based on data from SINS and GNSS. Wuhan University, Wuhan, Hubei, P.R.China, 2016.

[15] M. Senobari, A new method for accurate alignment and calibration of strapdown INS, Measurement, vol. 216, pp. 112943, 2023.

[16] M. Wang, J. Cao, S. Cai, M. Wu, K. Zhang and R. Yu, Improving the Strapdown Airborne Vector Gravimetry by a Backward Inertial Navigation Algorithm, Sensors, vol. 18, no. 12, pp. 4432, 2018.

[17] S. Cai, J. Cao, R. Yu, J. Cao, The mechanism and modification of low frequency error in strapdown airborne vector gravimetry, Journal of Chinese Inertial Technology, vol. 28, pp. 433, 2020.

Solid-State Wave Gyro in the Whole Angle Mode: Experimental Research of the Basic Parameters

S.E. Perelyaev

Ishlinsky Institute for
Problems in Mechanics
of the Russian Academy
of Sciences (IPMech RAS)
Moscow, Russia

B.P.Bodunov

JSC Research and Production
Enterprise MEDICON
Miass, Russia
e-mail: mdcn@medicon-miass.ru

S.B.Bodunov

JSC Research and Production
Enterprise MEDICON
Miass, Russia
e-mail: mdcn@medicon-miass.ru

Abstract – **The development of modern inertial systems for offline (autonomous) positioning and navigation substantially depends on the progress in the sphere of gyro design, based on the new physical principles. Among these a rather promising solid-state wave gyro (SSWG) could be distinguished in particular. The report demonstrates the experience in new-generation gyro design for modern air-space objects and represents the engineering solutions in design of the new-generation SSWG, based on a high-Q quartz hemispherical resonator, for low-dynamic and high-dynamic objects of air-space applications.**

Key words — hemispherical quartz resonator, solid-state wave gyro, inertial sensors test, tilt-rotation dinamic simulator.

1. INTRODUCTION

"SSWG-30" sensor basic accuracy parameters in known operational modes – force-to-rebalance (FTR or Close-Loop) and integrating gyro (WA or Open-Loop) were tested in 2023-2024 by means of a precise 3-axis tilt-rotation dynamic simulator with a built-in temperature chamber by JSC "Medicon" specialists.

High-precision FTR: the drift value of SSWG pilot model of new design (device decimal № D013) was tested. The tests were carried out at the ambient temperature of +40°C. No preliminary thermo-control or compensation of the SSWG was performed. The wave rotation angle was equal to 1.2 degrees. The bandwidth was equal to 1.5 Hz; sensor parameters sampling rate was 4Hz. The test lasted T=480 minutes (8 hours approximately) and was carried out in a day-time. SSWG systematic drift was estimated two hours after the temperature chamber was switched on. Test conditions for SSWG № D013 were as follows: initial temperature of the specimen was approximately 22°C; its basic parameters were estimated at 40°C, wave rotation angle 1.1 deg. Gyro warm-up, drift estimation, resonant frequency f_0=6550 Hz, temperature 42°C, and gyro control unit temperature 52°C. For the environmental conditions mentioned above SSWG systematic drift (FTR mode) was equal to *0.00327 deg/h*, standard root-mean-square deviation (MSD) of the angular velocity noise $\sigma\omega$ = 3.3 *deg/h*.

Night-time test with duration approximately T=15 h was performed. SSWG warm-up time was T=5 h before the measuring. Y-axis SSWG № D013 output was equal to 5.738 deg/h, noise root-mean-square value std(w0)=1.56 deg/h, SSWG systematic drift was *0,00073 deg/h,* angular velocity measuring range was 5 deg/sec. Noise and angle random walk (ARW) in FTR mode were: SIGMA(1s) σ=0,23 deg/h, ARW=σ/60=*0,00383 °/√h*.

WA (open-loop) mode: SSWG-30ig integrating gyro is a promising inertial sensor and used for highly-dynamic objects. It has a digital output and multichannel "push-pull" control system that operates according to algorithm designed by PhD V.F. Zhuravlev [1–3]. Digital part of electronic control unit comprises dedicated signal processor, ADC buffers, frequency synthesizer, and asynchronous serial interface controller. By means of special data handling algorithms the digital processor estimates wave angular position, oscillation amplitude and quadrature components and using this information calculates appropriate weight factors necessary for standing wave control. Data handling and control algorithms [2,3] are implemented simultaneously by a dedicated processor in FPGA of a new generation. Every control channel comprises a set of digital-analog converters (DAC); Every DAC gets following input digital data: control weight factors, and as bearings, sine/cosine signals and their derivatives from the pick-off electrodes [3,4], i.e. analog-digital converter (ADC) outputs. Control and data handling algorithm for every ADC has been implemented in modern FPGA of "Kintex-7" type with the technological topology level 28 nm.

2. WHOLE ANGLE OPERATIONAL MODE

WA (open-loop) mode: SSWG-30ig integrating gyro is a promising inertial sensor, which is used for highly-dynamic objects. It has a digital output and multichannel "push-pull" control system that operates according to algorithm [3,5] designed by Academician V.F. Zhuravlev. Digital part of electronic control unit comprises a dedicated signal processor, ADC buffers, frequency synthesizer, and an asynchronous serial interface controller. By means of special data handling algorithms, the digital processor estimates wave angular position, oscillation amplitude and quadrature components

and, using this information, it calculates appropriate weight factors necessary for standing wave control [4,5].

Data handling and control algorithms are implemented simultaneously by a dedicated processor in fast-acting FPGA of a new generation. Control and data handling algorithm [3,5] for every ADC has been implemented in FPGA Artix-7 [4,5].

3. DATA HANDLING RESULTS

3.1. *Bryan scale factor estimation*

Bryan scale factor [3,8] was calculated as a ratio of wave angular precession, specified over the entire data array (approximation factor at the first power of time), to the simulator absolute angular velocity [9].

The latter (absolute angular velocity) was calculated as a sum of relative angular velocity of simulator and vertical component of that of Earth rotation, taken as a well-known value for our latitude equal to 12,3423°/h.

The Bryan factor calculation results are shown below in Table 1.

TABLE I.

Simulator absolute angular velocity °/h (°/sec)	Simulator angular velocity MSD (°/h)	Wave precession angular velocity (°/h)	Wave precession angular velocity MSD (°/h)	Bryan factor
18012.207 (5)	$5 \cdot 10^{-5}$	4941.329	0.00072	0.2743323
7212.288 (2)	$2 \cdot 10^{-5}$	1978.571	0.00034	0.2743333
3612.314 (1)	$1 \cdot 10^{-5}$	990.979	0.00049	0.2743337
1812.329 (0.5)	$5 \cdot 10^{-6}$	497.453	0.0012	0.2744830
732.337 (0.2)	$3 \cdot 10^{-6}$	201.044	0.00096	0.2745247
372.3395 (0.1)	$0.7 \cdot 10^{-6}$	102.229	0.0019	0.2745579
192.341 (0.05)	$0.5 \cdot 10^{-7}$	52.814	0.00068	0.2745861
-167.656 (-0.05)	$0.9 \cdot 10^{-7}$	-45.965	0.0012	0.2741591
-347.655 (-0.1)	$0.7 \cdot 10^{-6}$	-95.408	0.0017	0.2744324
-707.652 (-0.2)	$3 \cdot 10^{-6}$	-194.259	0.0014	0.2745125
-1787.644 (-0.5)	$5 \cdot 10^{-6}$	-490.665	0.0022	0.2744757
-3587.630 (-1)	$1 \cdot 10^{-5}$	-984.279	0.0014	0.2743535
-7187.603 (-2)	$2 \cdot 10^{-5}$	-1971.738	0.001	0.2743248
-17987,523 (-5)	$5 \cdot 10^{-5}$	-4934.542	0.00043	0.2743314

a. Note: In the first column, the simulator relative angular velocity values are shown in brackets.

3.2 *Wave angular amplitude and angular velocity precession vs. angular velocity dependence estimation.*

Table 2 shows the results of harmonic components detection from angular detector data.

TABLE II.

Relative input velocity, °/s	First harmonic		Second harmonic	
	Frequency, rad/s	Amplitude, ang.min.	Frequency rad/s	Amplitude, ang.min.
5	0.096	0.97	0.01755	0.8
2	0.03875	0.37	-	-
1	0.01922	1.275	-	-
0.5	0.0096465	3.54	0.0118	0.62
0.2	0.003899	7.95	-	-
0.1	0.001979	19.77	0.004	0.62
0.05*	0.0010238	29.7	0.00207	1.1
-0.05*	0.0008915	55.09	0.00176	1.16
-0.1	0.00185	25.19	0.00365	0.61
-0.2	0.00376	11.59	-	-
-0.5	0.009503	6.41	0.0118	0.62
-1	0.0191	4.11	-	-
-2	0.03825	2.83	-	-
-5	0.09586	2.02	-	-

b. Note: At the input angular velocities larger than |0.05|°/s the rest of the components are omitted due to their small value in comparison to the first one.

c.

As it can be seen from the information above. the second and additional harmonics only become apparent at relatively small input velocities, and have one order less amplitude than that of the first one. Hence it is the first harmonic that has the primary effect on stability of the wave angular position relative to the resonator, and the further analysis will be performed for it. It is worthy of note that the full phase of this harmonic, including initial wave angular position and product of frequency and time, corresponds to actual wave angular position relative to the resonator. It gives an opportunity to register and compensate this harmonic.

The plot of amplitude of the wave angular position first harmonic vs input angular velocity is shown in Fig. 1.

Fig. 1.

Data analysis of Table 2 and diagram represented in Fig. 2 below show that the wave angular position first harmonic is apparently asymmetrical relative to the positive and negative input angular velocity values. Besides, its amplitude increases rapidly at the small values of input velocity. Let us consider how the wave angular velocity first harmonic amplitude depends on input angular velocity. There are two ways of the angular velocity harmonics detection out of the initial information: 1) to get them from the wave angle increment initial values and approximate them; 2) to estimate the angular velocity first harmonic amplitude by multiplying the frequency by wave angular position amplitude. Calculations have shown that the estimates coincide with an accuracy of 0.11°/h, which seems quite sufficient.

The plot of the wave angular velocity first harmonic amplitude vs input angular velocity is illustrated in Fig. 2. The

plot of its approximation is represented in the same plot as well in the following form: $A\omega = 10.53 - 6.097 \cdot \omega_{input}$.

Fig. 2.

Thus, the wave angular velocity first harmonic amplitude has a constant component that corresponds to the theory and a component that linearly depends on input angular velocity, the origin of which is not found yet. The approximation accuracy of the negative area of the input angular velocities is higher than that of the positive one. The wave angular position first harmonic amplitude asymmetry may be caused by amplitude dependence on the input velocity. Apparently, this fact complicates the error compensation considerably. The cause of this dependence will have to be further researched.

3.3 *SSWG accuracy parameters in WA mode.*
The device test results are shown in Table 3.

TABLE III.

Absolute angular velocity (°/h)	Estimated angular velocity(°/h)	Approximated angular velocity (°/h)	Difference between ppr. and estimated values
-72	-73.289	-72.549	0.74
-36	-36.621	-36.422	0.199
-21.6	-21.987	-21.971	0.016
-10.8	-10.986	-11.133	-0.148
-7.2	-7.288	-7.520	-0.232
-3.6	-3.594	-3.908	-0.314
0	0.104	-0.295	-0.398
3.6	3.639	3.318	-0.321
7.2	7.192	6.931	-0.262
10.8	10.864	10.543	-0.320
21.6	21.339	21.381	0.043
36	35.557	35.832	0.275
72	71.237	71.960	0.723

The approximation of estimated angular velocity was made as follows: $\omega_{est} = -0.295 + 1.00353 \cdot \omega_{inp}$ over the entire range of measure. $\omega_{est} = 0.0898 + 0.9874 \cdot \omega_{inp}$ over the positive area, $\omega_{est} = 0.056 + 1.0189 \cdot \omega_{inp}$ over the negative one.

The approximations were made for both positive and negative parts because the error form over entire range suggests the asymmetry of the measuring characteristic. The diagram of errors for these approximations is demonstrated in Fig. 3 below.

The implication of these results is that the device measuring characteristic in the range of $\pm 0.02^0$/s has a scale factor asymmetry equal to 0.0157^0/h relative to the mean value. Bias asymmetry (half-difference of the biases in negative and positive parts), is 0.0165^0/h.

Fig. 3.

3.4 *Insensitivity zone: presence/absence check-up.*
The results of two runs, carried out for the insensitivity zone detecting, are shown in Tables 4 and 5.

TABLE IV.

Simulator axis angle [°]	Estimated angular velocity [°/h]	Bias drift [°/h]	Calculated angular velocity increment [°/h]	Estimated angular velocity increment [°/h]	Angular velocity estimation error [°/h]
55.15	0.161972	0	0	-	-
55.3	0.106	-0.01289	-0.04	-0.04308	0.003082
55	0.148788	-0.02578	0.04	0.012597	0.027403
55.45	0.031181	-0.03867	-0.08	-0.09212	0.012119
54.85	0.161685	-0.05156	0.08	0.051276	0.028724
55.15	0.084628	-0.07734	-	-	-

TABLE V.

Simulator middle axis angle [°]	Estimated angular velocity [°/h]	Bias drift [°/h]	Estimated angular velocity increment [°/h]	Angular velocity estimation error [°/h]
55.15	-0.05626	0	-	-
55.3	-0.09013	0.002907	-0.03677	-0.00323
55	-0.04391	0.005814	0.006542	0.033458
55.45	-0.1368	0.008721	-0.08926	0.009256
54.85	0.025514	0.011628	0.070151	0.009849
55.6	-0.13311	0.014535	-0.09138	-0.02862
54.7	0.072464	0.017442	0.111287	0.008713
55.15	-0.03301	0.023256	-	-

SSWG bias drift is considered linear, time of device changing position is taken the same. The effect of insensitivity was not found upon the obtained data. Angular velocity estimation error does not exceed 0.033 °/h.

3.5 *Data handling results*
Analysis of pulse-wide modulated (PWM) signals for amplitude and quadrature control systems as a function of input angular velocity was performed. The technique is similar to that of the angular detector information analysis. PWM-signal harmonic set of the amplitude control system is different from that of the quadrature control one, though the first harmonics frequencies correspond to the same ones of the angular detector information. While the first and second harmonics are significant over the entire range of input

angular velocities for amplitude control system, the set of significant harmonics for quadrature control alters as a function of velocity. At the input angular velocity ±0.1°/s, the significant harmonics are at basic, doubled, quadruped and intermediate frequencies. As the input velocity increases, the harmonics with intermediate frequencies disappear, then quadruped harmonic disappears at the level of ±0.2°/s and higher, and doubled one disappears at more than ±1°/s. As the input angular velocity increases the full spectrum MSD of the quadrature control system PWM-signal increases from 0,04-0.07 to 0.38 relative units.

The initial information and associated discrepancies are shown below.

Fig. 4.

Fig. 5.

Fig. 6.

The residual discrepancy of the amplitude control PWM-signal is close to white noise. Its form is typical for input angular velocity from range in question.

Fig. 7.

Fig. 8.

Fig. 9.

Fig. 10.

Even after quadruped harmonic detection the discrepancy of quadrature control PWM-signal does not have a form of white noise.

Fig. 11.

Fig. 12.

The discrepancy form represented in Fig. 12 suggests signal modulation. It implies the noise additivity ratio upset, so further decrease of approximation MSD in the applied technique is impossible. There is a modulation at even less

input velocities. It apparently accounts for the form of finite discrepancy looks far from white noise. The plots of the first harmonics amplitude of both systems vs input angular velocity are illustrated in Figs. 13 and 14.

Fig. 13.

Fig. 14.

It can be seen from the diagrams that the first harmonic of the amplitude control PWM-signal has apparent asymmetry in both negative and positive range of input angular velocities. The increase in amplitude as the input velocity decreases is apparently explained by the fact that the product of the harmonic frequency and its amplitude (an analogue of the angular velocity amplitude) tends to a constant value. There is a correlation with the first harmonic amplitude of wave angular position peculiarities. The amplitude of the PWM first harmonic in quadrature control system decreases as the absolute input angular velocity increases. The fact can possibly be accounted for the influence of filters that limit the system bandwidth. The plots of MSD of the approximation for both systems vs input angular velocity are shown in Figs. 15 and 16.

Fig. 15.

Fig. 16.

It follows from the diagrams represented above that data approximation MSD of amplitude control system is less than that of the quadrature control. As the input angular velocity increases the ratio quadrature MSD/amplitude MSD reaches 12.3. It follows from the Fig. 15 that amplitude control system PWM data approximation accuracy at the input angular velocity faster than $|0.5|^0$/s is higher in case of positive effects than negative ones (MSD is less in case of positive effect than in case of negative one).

Conclusions and recommendations for the section
The job was performed basically to specify the device technical form for choosing its mode of operation, and particularly, design implementation proposition for WA mode and over the entire range of input angular velocities.

A special model was used in the experiment. It was designed for FTR operational mode with narrow bandwidth (less than 1 Hz) and measurement range 1.5^0/s and so the final conclusions can be made only after performing the same actions with a gyro, that satisfies all the technical requirements of the Customer.

Upon the test results and obtained data analysis the following conclusions can be made:

1. Insensitivity zone, i.e. angular velocities range that has no proportional connection between input and output information, with the background noise has not been detected;

2. The first harmonic amplitude of the standing wave precession angle depends on input angular velocity;

3. The first harmonic amplitude of the standing wave precession angular velocity is approximated by linear function of input angular velocity. The constant part of this function has an ample theoretical ground – resonator Q-mismatch effects to wave precession stability.

As this takes place, the initial formation of information regarding resonator oscillations is performed by the sequence of pulses with 15 B amplitude, given to *X* and *Y* SSWG resonator electrodes in turn with their further amplification. In so doing, the electrodes are used to form the initial data (as pick-offs) during one part of the oscillation period, and as control (forcers) electrodes during the other part. The drawings representing SSWG-FTR control unit operation for air-space application are shown in Figs. 17–19 [10].

The output signal of a quartz hemispherical resonator (HSR) lip oscillations that can be seen by means of digital oscilloscope is illustrated in Fig. 19, where *R* – in-phase (amplitude) component, *ρ* – quadrature component.

Fig. 17. SSWG-FTR functional diagram

Fig. 18. SSWG electrodes arrangement.   Fig. 19 SSWG-FTR output.

4.   AIR APPLICATION GYRO

Air application gyro is a typical inertial sensor – integrating SSWG – wave rotation angle is equal to integral of angular velocity projection onto its sensitivity axis. For this reason, the range of measured angular velocities actually is not limited that makes a device especially convenient for using in promising strap down inertial navigation systems (SINS). A new design of SSWG – SSWG-30ig, based on precise inertial sensor developed specially by JSC RPE MEDICON, has been implemented for highly dynamic objects. SSWG-30ig has a digital output and multi-channel push-pull control system, that operates according to algorithm [3,4], designed and implemented by specialists of the Systems Mechanic Laboratory (Ishlinsky Institute for Problems in Mechanics of the Russian Academy of Sciences) V.F. Zhuravlev and S.E. Perelyaev. Digital part of electronic control unit comprises signal processor, ADC buffers, frequencies synthesizer, and serial interface asynchronous microcontroller. Digital processor estimates angles of the wave position, amplitude and quadrature components according to special data handling algorithms and calculates appropriate factors necessary to control the standing wave. Data handling and control algorithms are implemented simultaneously by a dedicated processor in modern fast-acting FPGA.

Every control channel comprises a set of digital-analog converters (DAC). Every DAC gets following input digital data: control weight factors, and as bearings, sine/cosine signals and their derivatives from the analog-digital converters (ADC) outputs. Control and data pick-off algorithm for every ADC has been implemented in FPGA ARTIX-7 by means of dedicated "finite state machine" [4,5]. The ADC output data are converted by finite state machine into parallel 16-digit code that has READY bit. These data are transmitted to fast-acting real-time processor (RTP). According to synthesized algorithm RTP forms the direct digital control rules – DAC output signals (voltage). These signals are applied to gyro forcer electrodes embedded into the forcer/pickoff assembly. New digital electronic unit – gyro control unit (GCU-1) includes three signal processors for handling analogue output signals from three sensors (gyros). The structure of GCU-1 has been implemented in modern FPGA Artix-7 [5,6], that comprises embedded controllers, serial interface and principal 64-bit digital processor. A program that implements forming of data handling and control signals algorithm for each of three channels is performed simultaneously in each dedicated processor. This multitasking structure provides simultaneous data pick-off from all three gyros. Serial communication line controller performs data pickoff of every processor output units, converts the information into series of data blocks and sends it into standard communication line RS-422.

5.   SINE COMPONENTS ERROR RESEARCH

5.1. *Test purpose and conditions*: the research of bias error sine components for SSWG-30ig was performed in WA mode at high input angular velocities. The device under test was SSWG № E005. The test was aimed to detect sine harmonics set in SSWG output signal and to find out how their amplitudes and phases depend on input angular velocity. SSWG № E005 basic parameters during the lab test were as follows:

– Operational mode – WA (whole angle);
– Maximal bandwidth 25 Hz at angular velocity 400°/s and decreased in proportion to the input angular velocity;
– Input angular velocity: ±2, ±5, ±10, ±20, ±50, ±100, ±200, ±400 °/s;
– Sampling rate: 5, 5, 10, 25, 25, 25, 50, 50 Hz according to the input angular velocity;
–- Number of samples in each file: 1502, 1502, 1204, 1202, 1798, 956, 1904, 1329 according to the input angular velocity;
– Noise RMSV (root-mean-square value) at the bandwidth 25 Hz and sampling rate 50 Hz (for maximal angular velocity 400 °/s) is 433 °/h, harmonics amplitude estimation error is not higher than 11.9 °/h (considering the array of averaging 1329 samples).

Harmonics amplitude and phase calculation was performed by the Curve Fitting application of the Matlab 2020 software. The harmonics have the following form:

$$y = a_1 \cdot \sin(2\pi/180a+b_1) + a_2 \cdot \sin(4\pi/180a+b_2) + a_3 \cdot \sin(8\pi/180a+b_3) \qquad (1)$$

a – standing wave angular position, [°]; a_1, a_2, a_3 – harmonics' amplitudes, [°/h]; b_1, b_2, b_3 – harmonics origin phases, [rad].

5.2. *Gyro test results*
5.2.1 *Harmonic spectrum*

The output typical form (angular velocity deviation from the mean value of 400 °/c) is shown in Fig. 1. Output spectrum for 400 °/s is demonstrated in Fig. 2 (in linear scale) and Fig. 3 (in special semi-logarithmic scale).

Figure 1

Figure 2

Figure 3

Harmonics with frequencies higher than 5 Hz are caused by tilt-dynamic simulator. The first three harmonics belong to the device under test:

− a_1 at the frequency of 0.5859 Hz, period 180° by the resonator oscillation axis angular position. The harmonic is caused by the constant high voltage 200 V;

− a_2 at the frequency of 1.2205 Hz, period 90°, is caused by the HSR Q-mismatch and SSWG gap irregularity;

− a_3 at the frequency of 2.441 Hz, period 45°, is caused in SSWG by non-linearity of the electrode's capacity vibrations and electrodes arrangement error.

5.2.2 *Harmonics amplitudes dependence on angular velocity.*

The first harmonic amplitude a_1(Fig. 4) is all but invariant with angular velocity. The mean value is 151.3°/s, deviation is not higher than 3 °/s, so its value lies in the limits of estimation error.

Figure 4

The second harmonic amplitude a_2 (Fig.5,6) is a linearly function of input angular velocity.

Figure 5

Figure 6

The linearity upsets only at low angular velocities ω < 5°/s. At the rest velocities the deviation from linearity is not higher than 20 °/h.

The third harmonic amplitude a_3 (Figs. 5,6) linearly depends on angular velocity as well. The deviation from linearity takes place at low angular velocities of ω < 5°/s and extreme values of the range (up to 10 °/h). At the rest velocities the deviation is not higher than 4 °/h.

Figure 7

Figure 8

5.2.3 *Harmonics original phase dependence on angular velocity*

The plots of harmonics original phases vs angular velocity are shown in Figs. 9–11.

These diagrams demonstrate the presence of some constant value and a small linear dependence on input angular velocity. The original phase of the second b_2 and third b_3 harmonics strongly deviate from linearity at low angular velocities ω<20°/s.

Figure 9

Figure 10

Figure 11

5.3 Obtained information analysis
5.3.1 The first harmonics amplitudes

The first harmonic amplitude a_1 with period = 180° is invariant with input angular velocity within an accuracy of measuring. Harmonic is a bias error caused by the effect of high voltage 200B applied to SSWG quartz resonator and should be corrected by usual methods of adjustment. No extra correction is required. The second and third harmonics a_2 and a_3 at high angular velocities ($\omega > 5°/s$) are linear functions against angular velocity. The linearity upset at low angular velocities is accounted for the fact that signals with periods 90° and 45° comprise two components:

- the first one is dependent on input angular velocity and caused by the gap irregularity, electrodes arrangement error and non-linearity of the electrodes capacity vibrations;
- the other is independent of input angular velocity, and caused by the effects of SSWG HSR Quality/Frequency mismatch.

When harmonic a_2 or a_3 amplitude value diminishes to that of the components invariant with angular velocity they stop depending on input angular velocity that causes the linearity upset. Harmonics a_2 and a_3, invariant with angular velocity should be corrected by usual methods of adjustment.

To correct the angular velocity-dependent components, the correction system should generate a special correction signal of the following form:

$$\omega_{cor} = \omega K_{a2} \sin(2a + b_2) + \omega K_{a3}\sin(2a + b_3) \quad (2)$$

where ω – angular velocity, [°/h]; a – resonator oscillation axis angular position, [rad]; b_2, b_3 – original phases, [rad]; K_{a2}, K_{a3} – correction factors.

It will be sufficient to measure a_2 and a_3 harmonics amplitude only at two velocities for K_{a2}, K_{a3} estimation, for example: at -200 °/s and 200 °/s.

5.3.2 Harmonics original phases

The original phases comprise some constant value and a small addition linearly dependent on angular velocity:

$$b_2 = b_{02} + \omega K_{b2} ; \quad b_3 = b_{03} + \omega K_{b3} \quad (3)$$

where b_{02}, b_{03} – constant values; ω – angular velocity; K_{b2}, K_{b3} – correction factors.

Harmonic a_1 original phase is non-linear over the entire range of angular velocities. It is corrected during the process of adjustment, so it is not considered.

The original phases b_2, b_3 linearity upsets at low angular velocities for the same reason as for the corresponding amplitudes. Non-linearity of the original phases at low angular velocities is eliminated by bias error correction and usual methods of adjustment.

It will be sufficient to measure phases of a_2 and a_3 harmonics only at two velocities for K_{b2}, K_{b3} factors estimation, for example: at -200 °/s and 200 °/s.

6. CONCLUSIONS

1. While SSWG is being rotated, two additional harmonics of the resonator oscillation angular position with periods of 90° and 45° appear in its output spectrum. Their amplitude is in proportion to input angular velocity.

2. To correct these harmonics, it is necessary to generate appropriate correction signals according to formulas (2), (3) mentioned above.

3. It is sufficient to measure harmonics' amplitudes and phases at -200°/s and 200°/s angular velocities for the correction factors estimation. Measuring should be carried out for completely adjusted SSWG.

4. In this research measuring data and correction signal formulas for Normal Environmental Conditions (NEC) are demonstrated.

7. GYRO ACCURACY PARAMETERS

The inertial sensor SSWG-30i pilot model of new design was tested on a precise 3-axis dynamic simulator ACUTRONIC (Switzerland). Numerous experiments were carried out along with the customer; so, the basic characteristics of new generation device SSWG-30ig were confirmed and the dynamic range ±1500 °/s was obtained while the navigation grade accuracy demands for gyro operation (MSD≤0.015 °/h) were met. SSWG-30ig control runs were also performed on 2-axis tilt-rotation dynamic simulator "ACTIDYN" (France) and confirmed that

integrating gyro operates effectively in WA mode and has almost unlimited range of angular velocity/acceleration measuring. At the same time the device meets the navigation grade accuracy demands (MSD≤0.015 °/h) [10]. The lab test results of the new inertial sensor SSWG-30ig in JSC MEDICON at the room ambient temperature confirmed high in-run and run-to-run bias stability [11,12]. SSWG-№ D013 in-run standard Mean Square Deviation (MSD) did not exceed MSD=0.005°/h, and angular random walk (ARW) was less than $0.003°/\sqrt{h}$. As a result of the design and implementation of the new and completely digital control unit of SSWG-30ig for highly-dynamic objects and sampling rate f=1000 Hz, angular noise was reduced to:

$$\sigma_a = 0.11'' \text{ and } \varepsilon_a = 0.009''/\sqrt{Hz} \text{ .}$$

For the serious improvement of the new gyro a method for device autocalibration was developed. Applying this procedure to an industrial inertial sensor during its dynamic and environmental tests allows speaking only about positive results of the new technique for instrumental and technological errors compensation. To minimize turn-on time period and reduce warm-up duration of SSWG-30ig a quick starting algorithm of driving quartz resonator to its natural frequency has been implemented. At the same time an algorithm for gyro warm-up time minimization has been synthesized. It considerably reduced the time for driving the standing wave amplitude to its required value and provided quick quadrature suppression [13]. All this really helped to minimize SSWG-30ig turn-on time to required value, that does not exceed T=3s. On trials the maximal range of measured angular velocities was (±400°/s).

The value was obtained for industrial applications, specified by designer, under limited capabilities of SSWG-30ig control unit and that of the device programmable system, implemented in a microchip of industrial class, hence it was not conditioned by potential capabilities of the new integrating gyro proper [14]. Short warm-up time, necessary for a gyro to meet its required performance, indicates the ability to measure Earth rotation velocity and required accuracy for long-term (longer than 1h) offline positioning and navigation of SINS-SSWG without being corrected by global satellite systems [15].

FINAL CONCLUSIONS

A highly reliable small-sized integrating gyro SSWG-30ig with digital output and fully digital control designed by Russian specialists is a high-precision device for offline positioning. It is based on a pilot model of inertial sensor SSWG-30ig № D013 type of modern design. The radically new technique for quartz toothless HSR (d = 30mm) stem fixation has been implemented in the sensor. The new digital electronic control unit (multichannel system of direct digital control, pickoff and signal processing) has been used. The device can be competitive with fiber-optic gyro (FOG) and ring laser gyro (RLG) in terms of cost, accuracy and power consumption; but it is much more reliable in various kinds of missions and under drastic environmental conditions.

REFERENCES

[1] Zhuravlev V.F. Fundamental questions of the theory of new gyroscopic sensors of the "generalized Foucault pendulum" family. Problems of mechanics: Collection of articles. To the 90th anniversary of the birth of A.Yu. Ishlinsky / Edited by D.M. Klimov. – M.: Fizmazlit, 2003.-832 pages.

[2] Zhuravlev V.F. About the drift of a wave solid-state gyroscope on a rotating base at quadrature control in the fast and slow time modes. // Solid State Mechanics. RAS. № 3. 2003. pp.13-18.

[3] Klimov D.M., Zhuravlev V.F., Zhbanov Yu.K. Quartz hemispherical resonator (Solid-state wave gyroscope). M.: Kim L.A., 2017. 194 pp.

[4] Perelyaev S.E., Bodunov B.P., Bodunov S.B. Solid-state wave gyroscope - a new generation inertial sensor // Materials of the XXIV St. Petersburg International. conf. on integrated navigation systems. SPb.: JSC Concern Central Research Institute Elektropribor, 2017. pp. 287–290.

[5] Zhuravlev V.F., Perelyaev S.E., Bodunov B.P., Bodunov S.B. A new generation miniature solid-state wave gyroscope for strap-down inertial navigation systems of unmanned flying vehicle. // XXVI St. Petersburg International Conference on Integrated Navigation Systems. St. Petersburg: Concern Central Research Institute Elektropribor, 2019. – pp. 250-254.

[6] Perelyaev S.E., Zhuravlev V.F., Bodunov B.P., Bodunov S.B. Fundamental questions of the theory of new gyroscopic sensors of the "generalized Foucault pendulum" family and applied aspects of its implementation in the engineering practice of modern gyroscopy. // XXVII St. Petersburg International Conference on Integrated Navigation Systems. St. Petersburg: Concern Central Research Institute Elektropribor, 2020. pp. 262-272

[7] Zhuravlev V.F. On the formation of feedback in the Van der Pol spatial oscillator. // Applied mathematics and mechanics. Vol. 84. No. 2. 2020. pp. 151–157.

[8] Bryan G. H. On the beats in the vibrations of a revolving cylinder or bell // Proc. Camb. Phil. Soc., Vol VII, Nov. 24, pp. 101–111, 1890.

[9] Bodunov B.P., Lynch D.D, Voros A. Inexpensive hemispherical resonator for small-sized SWG navigation systems for civil purposes // Materials of the II St. Petersburg Int. conf. on Integrated Navigation Systems, 1995.- pp. 89-92.

[10] Bodunov B.P. Development of a solid-state wave gyroscope for low-dynamic and high-dynamic objects. Report at the XIII All-Russian Conference on Management Problems (VSPU-2019), Materials of the XIII All-Russian Conference on Management Problems. Russia, Moscow, Institute of Management Problems named after V.A. Trapeznikov RAS, June 17-20, 2019, pp. 1336-1341.

[11] Peshekhonov V.G. Prospects for gyroscopy. Report at the section "State and prospects for the development of modern navigation systems." Materials of the XIII All-Russian Conference on Management Problems. Russia, Moscow, Institute of Management Problems named after V.A. Trapeznikov RAS, June 17-20, 2019.

[12] Bodunov B.P., Bodunov S.B., Kotelnikov S.V. Patent No. 2362121 Russian Federation, IPC G-01C 19/56. Small-sized solid-state wave gyroscope. - No. 2007125894A; publ. 07/27/2009.

[13] Bodunov B.P., Bodunov S.B., Vladimirov V.A., Igonin A.N., Kostenok N.A. Solid-state wave gyroscope of dual-mode operation for space applications //Materials of the XX International. conf. on integrated navigation systems. St. Petersburg: JSC Concern Central Research Institute Elektropribor, 2013, pp. 173-178.

[14] Perelyaev S.E., Zhuravlev V.F., Bodunov B.P., Bodunov S.B. Theory of a two-dimensional Van der Pol oscillator. Technical applications in new solid-state wave gyroscopes. // Materials of the XXVIII St. Petersburg International Conference on Integrated Navigation Systems. St. Petersburg: Concern Central Research Institute Elektropribor, St. Petersburg, 2021, pp. 307-311.

[15] Solid-state wave gyroscope for aerospace applications of navigation accuracy class. // XXIX St. Petersburg International Conference on Integrated Navigation Systems. St. Petersburg: Concern Central Research Institute Elektropribor, 2022, pp. 185-188.

Dissipation of the Oscillation Energy in Adhesive Joints of Mechanical HRG Resonators

B. Lunin

Dept. of Chemistry,
Lomonosov Moscow State University
Moscow, Russia
luninboris@yandex.ru

M. Basarab

Dept. of Informatics and Control Systems
Bauman Moscow State Technical University
Moscow, Russia
basarab@bmstu.ru

Abstract — **Dissipation processes occurring in typical HRG adhesive joints are considered: gluing the resonator to the base and in adhesive joints of piezoelements to the surface of mechanical resonators. Part of the vibration energy of the resonator is dissipated due to bulk and thermoelastic internal friction in the adhesive layers, which deteriorates the characteristics of the devices. The influence of the configuration of adhesive joints and the properties of its material on the dissipation of vibration energy is analyzed.**

Keywords — adhesive, dissipation, internal friction

I. Introduction

Hemispherical resonator gyroscopes (HRG) are a large class of modern navigation devices used in various fields of technology. However, there are some technical problems that limit the widespread use of HRG. The most important of them is the high requirements for axial symmetry and high Q-factor of the resonator. Deviation of the resonator shape from axisymmetric leads to the appearance of a mass defect, which generates forces and moments acting on the center of mass of the resonator during oscillations and deteriorating the HRG characteristics [1-5]. For HRG resonators operating on the lower bending mode of vibration, the 1st and 3rd forms of the mass defect cause transverse (relative to the symmetry axis) forces, and the 2nd form causes a longitudinal force. In addition, the non-uniform distribution of mass along the height of the resonator generates torques. This leads to transverse and longitudinal vibrations of the resonator stem and to the dissipation of part of the vibration energy in the adhesive joint connecting the resonator stem to the base. If dissipation in this connection depends on the direction and magnitude of the vibration vector of the center of mass, then the Q-factor of the resonator will depend on the orientation of the wave pattern, and a systematic drift of the standing wave with speed will appear [6]

$$\frac{d\theta}{dt} = \frac{\omega \sin 4(\theta - \varphi_0)}{8}(\xi_1 - \xi_2),\qquad(1)$$

where θ – current angular position of the wave pattern; ξ_1 and ξ_2 – maximum and minimum internal friction in the fixed resonator; φ_0 – angle of orientation of the dissipative defect; ω – circular frequency of resonator oscillations.

The values ξ_1 and ξ_2 are determined both by the characteristics of the resonator itself and the configuration of the *resonator-base* connection, as well as internal friction in the cured adhesive.

In inexpensive HRGs with metal resonators, piezoelectric sensors and actuators glued to the surface of the resonator are usually used to measure vibration parameters and excite them [7]. Each of these piezoelectric elements introduces additional energy losses, which consist of losses in the piezoceramics itself and losses in the connecting adhesive layer. Considering that the energy losses introduced by eight glued piezoelements account for up to 50% of the total energy losses in the resonator, the influence of dissipation in the adhesive *piezoelement-resonator* joints on the accuracy characteristics of the HRG turns out to be high. The amplitude of the speed of systematic drift of the wave pattern (1), associated with dissipative processes in adhesive joints of HRG, can reach several tens of degrees per hour, and its stability depends on the constancy of internal friction in these joints. Essentially, it is the mass imbalance of the resonator and the dissipation of vibrational energy in adhesive joints that are the main cause of HRG errors.

The purpose of this work is to consider the main factors influencing the dissipative characteristics of the *resonator–base* and *piezoelectric element–resonator* adhesive joints, which are typical in most HRG designs described in the literature.

II. Dissipation of vibration energy at the place of mounting of the HRG resonator

Figure 1 shows a typical connection of the hemispherical resonator of the HRG with the base. Geometric deviations of the resonator shell of radius R can be arbitrarily distributed over the surface of the hemisphere. For a selected shell section, the wall thickness h is described by the Fourier series

$$h(\varphi) = h_0 + \sum_{k=1}^{\infty} h_k \cos k(\varphi - \varphi_k).\qquad(2)$$

Uneven resonator wall thickness leads to uneven mass distribution (per unit azimuthal angle φ)

$$m(\varphi) = m_0 + \sum_{k=1}^{\infty} m_k \cos k(\varphi - \varphi_k),\qquad(3)$$

where $m_k = \rho R^2 h_k$ and ρ – density of the resonator material.

For simplicity, Fig. 1 shows only the defect of the 1st form with the value of m_1, located at a small distance from the edge, determined by the angle α. Under vibration this mass defect causes a transverse force F, a moment of force associated with fastening the resonator outside the center of mass $M_{11} \approx F \cdot L$, and a moment of force associated with the vertical movement of the edge of the shell $M_{12} = F \cdot R$. The application of these forces and moments of force to the adhesive layer causes its deformation and energy dissipation, that is, it reduces the Q-factor of the mechanical resonator. The movement of the resonator along the X axis under the influence of force F is described by the equation

$$\ddot{x} + \xi_a \omega \dot{x} + \frac{SY}{d_a M} x = \frac{F}{M} \cos \omega t,\qquad(4)$$

where M – resonator mass; ξ_a – bulk internal friction in the adhesive layer; d_a – thickness of the adhesive layer; S – deformable area of the adhesive layer; Y – elasticity modulus of cured adhesive.

Fig. 1. Typical resonator-base connection in HRGs

In stationary mode, the amplitude of oscillations of the resonator stem is

$$x_0 = \frac{F}{M\sqrt{\left(\frac{SY}{d_aM} - \omega^2\right)^2 + \xi_a^2\omega^4}} . \tag{5}$$

In the vast majority of cases, the inequality $\xi_a^2\omega^4 << \left(\frac{SY}{d_aM} - \omega^2\right)$ is true, therefore $x_0 \approx \frac{Fd_a}{2SY}$.

Deformation energy of the adhesive layer is

$$W_a = \frac{F^2 d_a}{2SY} . \tag{6}$$

Part of the deformation energy of the adhesive layer (ΔW) is irreversibly dissipated in it; it depends on the bulk internal friction in the cured adhesive

$$\Delta W = 2\pi\xi_a W_a . \tag{7}$$

Additional internal friction introduced into the resonator is

$$\xi = \frac{\xi_a W_a}{W} , \tag{8}$$

where W– resonator oscillation energy.

For the hemispherical resonator according to the Rayleigh formula [6]

$$W = \frac{1,52961\pi R^2 h_0 \rho a^2 \omega^2}{8} , \tag{9}$$

where a – amplitude of resonator oscillations.

Assuming for the 1st form of mass defect $\alpha = \pi/2$ and $F = m_1 a\omega^2$, we obtain

$$\xi = \frac{4\xi_a d_a m_1^2 \omega^2}{1,52961SY\pi R^2 h_0 \rho} . \tag{10}$$

These losses are included as an additional term in the total internal friction in the resonator. If the resonator is balanced, that is $m_1 = 0$, then, according to (10), $\xi = 0$. In an unbalanced resonator, if the magnitude of the force

deforming the adhesive layer does not depend on the orientation of the standing wave (this occurs in the presence of only the 1st or only the 3rd form of mass defect), ξ=const. Then the internal friction along the resonator's own dissipative axes ξ_1 and ξ_2 will increase by the same amount ξ. The Q-factor of the resonator will decrease, but the difference $(\xi_1 - \xi_2)$, included as a factor in (1), will remain the same, that is, the speed of the systematic drift of the standing wave will not change. The presence of the 2nd form of a mass defect, as well as the simultaneous presence of its 1st and 3rd forms, lead to the dependence of the force F on the azimuthal orientation of the standing wave, while the value ξ is a function of the azimuthal angle φ, causing a change in the difference $(\xi_1 - \xi_2)$, and systematic drift speed.

Let us consider the influence of inhomogeneities of the adhesive layer on the dissipation of vibration energy. Let the hole in the base with a diameter of D and the resonator stem itself be perfectly round, but the stem is fixed asymmetrically, that is, the thickness of the adhesive layer on opposite sides of the leg is different, on one side $(d_a + \delta)$ and on the other $(d_a - \delta)$. When the variable force F is applied to the resonator stem, the connecting layer on both sides will be deformed by the same amount Δd but the nature of the deformation will be different: on one side the layer will be compressed, and on the other it will be stretched. In this case, the elastic forces arising on opposite sides of the stem due to different thicknesses of the connecting layer will not be the same, and the deformation energies of the adhesive layer on opposite sides of the stem will also be different. The total deformation energy of the adhesive layer is equal to

$$W_a = \frac{\Delta d}{d_a}SY\left(1 - \frac{\delta^2}{d_a^2}\right)^{-1} . \tag{11}$$

The last factor in (11) takes into account the asymmetry of the gluing. With the ratio $\delta / d_a = 0.1$, there will be an asymmetry of losses in fastening equal to approximately 1%, which will lead to corresponding change in the amplitude of the systematic drift velocity (1).

Now let the resonator stem be perfectly round, glued into the base symmetrically, but the hole itself in the base is oval, due to which the gap changes from d_a (along the X axis) to $(d_a + \delta)$ along the Y axis. With the same force F created by the mass resonator defect, the deformation energies of the adhesive layer along the X and Y directions will be different

$$\frac{\Delta W_y}{\Delta Wx} = \frac{d_a + \delta}{d_a} . \tag{12}$$

So, for example, if the clearance along the Y axis is 10% greater than the clearance along the X axis, then the energy dissipation in the Y direction is 10% greater compared to the losses in the X direction. This example shows that the ovality of the hole has a much greater effect on the HRG systematic drift than the asymmetry of the resonator stem fixation.

When gluing the resonator to the base, air bubbles may form in the layer. They also affect energy dissipation in the joint, since they change the deformation area of the material of the layer S. If there is an air bubble with area s in the adhesive layer, then according to (6), the deformation energy of the adhesive layer increases by $S / (S - s)$ times, leading to the change in the speed of the systematic drift.

Formulas (4), (5), (7), (8) include the value of bulk internal friction in the cured adhesive ξ_a which characterizes that part of the deformation energy of the connecting layer that is irreversibly converted into heat during periodic deformation of the material. This internal friction is determined by the structure of the cured adhesive, various types of defects, etc. Theoretical ideas about the nature of internal friction in solids make it possible to understand the inelastic behavior of cured adhesive, but the theoretical calculation of ξ_a is extremely difficult. Therefore, when determining losses in the adhesive layer, one should use experimental data obtained by dynamic mechanical thermal analysis in a fairly wide temperature range, during which the temperature dependences of the elastic modulus, internal friction and expansion coefficient of the cured adhesive are determined (for examples of such measurements, see, for example, [8]).

III. DISSIPATION OF VIBRATION ENERGY IN THE GLUED PIEZO ELEMENT-RESONATOR

Let us consider the dissipation of the oscillation energy of the resonator in the *piezoelement-resonator* gluing. In this case, the layer of cured adhesive between the piezoelectric element and the resonator is deformed along with the bending of the surface, and part of the deformation energy of the adhesive layer is irreversibly dissipated. During flexural deformation of the thin layer of cured adhesive, in addition to bulk internal friction in the adhesive structure, thermoelastic losses occur. The total internal friction in such layer is the sum of two quantities: bulk (ξ_a) and thermoelastic (ξ_{TE}) internal friction. The mechanism of thermoelastic losses was described by Zener [9]. During bending vibrations of the thin plate, the deformation of individual sections of its surface layer has different signs: on the convex side of the plate the material expands, and on the concave side it contracts, and vice versa. If the thermal coefficient of linear expansion of the material is different from zero, then during deformation the temperature of the material will locally change, and this change, depending on the sign of the deformation, can be either positive or negative. Thus, during bending vibrations of the thin plate, changes in local temperatures on its convex and concave sides at each moment of time have opposite signs, which is why a heat flow arises between these sides, equalizing these temperatures and converting the energy of mechanical vibrations into heat. Thermoelastic losses in the thin layer of cured adhesive can be estimated using the formula [9]

$$\xi_{TE} = \frac{\alpha^2 YT}{\rho C_p} \cdot \frac{\omega\gamma}{1 + \omega^2\gamma^2}, \text{ где } \gamma = \frac{\rho C_p d_a^2}{\pi^2\kappa}, \qquad (13)$$

where α - thermal coefficient of linear expansion of cured adhesive; C_p – specific heat; γ - characteristic time of heat transfer through the adhesive layer; κ - thermal conductivity; T – temperature.

Thermoelastic losses can be quite large for materials with large coefficients of thermal expansion, while their maximum occurs at a thickness corresponding to the condition $\omega\gamma=1$.

The deformation energy of the adhesive layer Wa between the piezoelectric element and the resonator can be calculated using the following formula, taking into account the dimensions of the layer and the elastic modulus of the material [10]

$$W_a = \frac{bx_0^2}{3L^3} nYd_a^3, \qquad (14)$$

where n – constant coefficient; b and L – width and length of the adhesive layer; x_0 – deformation amplitude.

Additional internal friction ξ introduced into the mechanical resonator by the adhesive layer connecting the piezoelement to the surface can be found using the formula

$$\xi = \left[\frac{bx_0^2 n}{3L^3 W}\right] \cdot \left[Yd_a^3\left(\xi_a + \xi_{TE}\right)\right], \qquad (15)$$

where W – total vibration energy of the resonator.

Let W=const. Since the first factor in (15) does not depend on the properties of the adhesive and its thickness, the losses introduced by the adhesive layer are proportional to $Yd_a^3(\xi_a + \xi_{TE})$. Fig. 2 shows the results of calculating these losses in the K-400 adhesive layer and, for comparison, in the POS solder layer in relative units. Normalization was carried out by the value of losses calculated for the cured epoxy adhesive with the layer thickness of 5 microns at room temperature, while the value of bulk losses in the adhesive was taken to be $\xi_a = 0.05$, and in the solder $\xi_a = 0.005$.

The graph shows that the energy dissipated in the layer of solder and adhesive differs by an order of magnitude. This is because the elastic modulus of the solder is significantly higher than the elastic modulus of the cured epoxy adhesive. With the same amplitude of deformation of the connecting layer, the energy spent on deformation of the solder turns out to be much greater and therefore the amount of energy dissipated in the solder is much higher than in the adhesive layer. This shows that the use of adhesives for installing piezoelectric elements on the surface of mechanical resonators is preferable.

Fig. 2. Losses (rel. units) introduced into the resonator by a layer of cured epoxy adhesive and solder: 1 – solder; 2 – adhesive

On the other hand, it is not always possible to create the thin adhesive layer due to the presence of relatively large filler particles in the adhesive, while the minimum thickness of the solder layer may well be a few microns. In this case, the use of soldering for attaching piezoelements is justified.

IV. Conclusion

Mass unbalance of the resonator leads to vibration of its stem and to energy dissipation in the adhesive layer connecting the resonator to the base. The insertion loss is proportional to the thickness of the adhesive layer and inversely proportional to the diameter of the stem and the thickness of the HRG base. Asymmetrical fixation of the stem in the base hole, ovality of the hole or stem, and the presence of bubbles in the adhesive layer entail an azimuthal dependence of losses and create additional systematic drift of the HRG. The bulk internal friction in the cured adhesive is a complex function of temperature, so the systematic drift of the standing wave in the HRG associated with these losses will also be temperature dependent.

Gluing piezoelectric elements to the surface of mechanical resonators also introduces additional internal friction into the resonator, which is caused by dissipation both in the piezoceramics and in the adhesive layer. It is caused by both bulk and thermoelastic internal friction, and at the same thickness, the adhesive layer introduces much less internal friction than the solder layer. The use of soldering for attaching piezoelements from the point of view of losses introduced into the resonator is justified only for small layer thicknesses (several microns).

References

[1] Yu.K. Zhbanov, N.V. Kalenova "Surface unbalance of a hemispherical resonator gyro", Mech. Solids, 2001, v.30, No.3, pp.7-12.

[2] Yu.K. Zhbanov, V.F. Zhuravlev "On the balancing of a hemispherical resonator gyro", Mech. Solids, 1998, v.33, No.4, pp.2-13.

[3] M.A. Basarab, V.A. Matveev, B.S. Lunin, S.V. Fetisov "Influence of nonuniform thickness of hemispherical resonator gyro shell on its unbalance parameters", Gyroscopy Navig., 2017, v.8, No.2, pp.97-103.

[4] Gireesh N. Sharma, Sundararajan, T. & Gautam, S.S. Effect of geometric imperfections on anchor loss and characterisation of a gyroscope resonator with high quality factor. Gyroscopy Navig. 11, 206–213 (2020). https://doi.org/10.1134/S2075108720030074.

[5] D.D. Gerrard, E.J Ng, C.H. Ahn, V.A. Hong, T.W. Kenny "Modeling the effect of anchor geometry on the quality factor of bulk mode resonators", 2015 Transducers - 2015 18th International Conference on Solid-State Sensors, Actuators and Microsystems (TRANSDUCERS), Anchorage, AK, USA, 2015, pp. 1997-2000.

[6] N.E. Egarmin, V.E. Yurin "Introduction to theory of vibratory gyroscopes", Moscow, Binom, 1993, 112 p.

[7] Raspopov, V.Y., Likhosherst, V.V. HRG with a Metal Resonator. Gyroscopy Navig. 14, 14–26 (2023).
https://doi.org/10.1134/S2075108723010066.

[8] L. Qiao, A.J. Eastel, C.J. Bolt, P.K. Coveny, R.A. Franich "Thermomechanical analysis and performance test of some EPI wood adhesives", Pigment and Resin Technology, 2000, v.29, No.4, pp.229-237.

[9] C. Zener "Internal Friction in Solids. II. General Theory of Thermoelastic Internal Friction", Phys. Rev. 1938, v.53, No.1, pp.90-99.

[10] V.I. Feodosiev "Strength of Materials", v.2, Bauman MSTU, 1999, 592 p.

Development of a Methodology for Analysis of Vibration Damping of a Hemispherical Resonator Gyroscope Cluster

D.I. Martynenko
JSC «Inertial technologies of "Technocomplex"»
Ramenskoye, Russia
DMartynenko@inertech.ru

S.V. Fetisov
JSC «Inertial technologies of "Technocomplex"»
Ramenskoye, Russia
SFetisov@inertech.ru

A.S. Malyugin
JSC «Inertial technologies of "Technocomplex"»
Ramenskoye, Russia
AMalyugin@inertech.ru

L.E. Kochegizova
JSC «Inertial technologies of "Technocomplex"»
Ramenskoye, Russia
LKochegizova@inertech.ru

Abstract—**The paper presents an approach to evaluate the characteristics of a hemispherical resonator gyroscope cluster damping system. The considered methodology enables to determine the mechanical characteristics of rubber-metal shock absorbers using bench equipment and software. Nonlinear characteristics, incompressibility parameters and rubber damping coefficient were identified by using a set of experimental data. According to the obtained information, the parameters of inertial measurement unit were calculated when exposed to external vibration. The conducted field tests made it possible to verify the calculation model. The finite element model simplified the approach to estimate the natural frequencies and displacements of the structure during design development phase to meet the requirements of the specifications.**

Keywords—HRG, hemispherical resonator gyroscope, vibration damping, inertial measurement unit

I. INTRODUCTION

Hemispherical resonator gyroscopes (HRG) are relevant autonomous angular motion sensors of aircraft. Critical features of inertial sensors include low magnitude drift rate, its long-term stability, ensuring reliability and long service life when subjected to external disturbances in operating conditions of the application object.

The effect of shock and vibration load on the equipment can cause numerous failures related to noise, stresses, resonances and damage accumulation in structures [1]. HRG, as a vibrating gyroscope, is subjected to the influence of external vibrations due to the imperfection of the sensing element, the resonator, as well as due to the presence of intrinsic resonances of the gyroscope's structural parts. Attainment of stable reproducible output characteristics is only possible if HRG-based inertial measurement unit (IMU) is protected from such mechanical disturbances as vibration and shock.

II. PROBLEM STATEMENT

A current problem of the HRG-based IMU development is to ensure vibration strength. The presence of the first, second, and third harmonics of the mass deviation across the resonator's circumferential angle under the effect of external vibration causes parasitic components that distort the useful signal [2]. Balancing the resonators against the first three harmonics of the mass defect reduces this effect. Shocks and vibrations from the base do not affect the dynamics of a fully balanced resonator operating at the second fundamental oscillation mode [3]. Fabricating a resonator in practice without instrumental assembly errors is impossible. These errors lead to connections between oscillation modes and vibration sensitivity. The decrease in accuracy of the wave pattern measurement leads to an additional drift rate.

Vibration isolation system is also necessary if there are vibrations of the structure at frequencies close to the operating frequency of the HRG resonator. The resonance effect at natural frequency depends on the magnitude of external influence. At low vibration levels the pick-off and drive systems may enter the instability zone. At high levels of vibration, significant stresses might occur in adhesive and welded joints due to inertial forces.

Thus, the vibration isolation system makes it possible to reduce the magnitude of HRG vibration-induced errors and the level of actioning effects on the cluster; stabilize the output characteristics and prevent the occurrence of resonances. The purpose of this work is to develop a methodology for analyzing vibration damping systems that requires a low information load of developers and has a low computational cost.

III. PROBLEM SOLUTION

Within the limits of the work performed, the task was set to ensure the IMU resistance to mechanical disturbances. Small dimensions of the IMU housing and a specified level of impacts justify for the choice of the protection method using shock absorbers.

The shock absorption is a system of elastic supports and the protected object, forming an oscillatory link [4]. The simplest shock absorbers are rubber-metal ones. The advantage of this type of shock absorbers is the ability to experience plastic deformations such as compression, tension, shear and torsion across a wide temperature range [5].

Currently, "digital twins" of the product are becoming increasingly widespread. The mathematical model should correspond to the real product and reliably reflect the behavior of the structure [6]. A finite element model is considered as a computational model of an IMU. The use of

verified models makes it possible to evaluate product characteristics at the development stage. Simulation software allows us to visualize the information of interest and simplify the setting up boundary conditions. The study results are reliable only if the mass, stiffness, and damping matrices are correctly constructed [7].

As part of the work, a commercially available shock absorber model was selected. The damping element is made of rubber compound IRP-1347. However, technical specifications for this compound do not reveal material nonlinear properties and damping characteristics. In the absence of energy dissipation parameters, calculations yield an infinitely large estimate of displacements in the natural frequency domain. The available studies on rubber characterization [8] require expensive equipment for tensile-compression of rubber samples. It should be noted that rubber features under dynamic loads may differ from the parameters determined in static experiments.

The method of determining the amplitude-frequency response (AFR) of a structure is devoid of these disadvantages. A loaded shock absorber 2 is fixed through a fixture 3 on the armature of vibration shaker table 4. Sensor 5, which measures vibration acceleration at the output of the oscillatory system, is installed on load 1. Sensor 6, measuring the vibration acceleration of the input effect, is placed on the base of the vibration shaker table (Fig. 1).

Fig. 1. Shock absorber AFR determination setup configuration

The automatic vibration control system maintained a constant sinusoidal vibration impact magnitude in sine sweep mode at a constant speed, ranging from 50 to 2000 Hz. The developed tooling provides both vertical and horizontal fastening of the loaded shock absorber to evaluate the AFR affected by axial and transverse vibration.

In the course of work, it was observed that the deformation magnitude depended non-linearly on the shock absorber load. Therefore, weights of three ratings (0.1, 0.2, and 0.3 kg) were selected based on the mass distribution magnitude of the proposed cluster.

The AFR of the shock absorber was obtained using finite element simulation software. The goal of the study was to establish correspondence between the experimental and calculated characteristics for each load. The elastomer behavior modeling is a non-trivial task. Rubber is classified as a hyperelastic material. This material model is described as nonlinearly elastic, isotropic, incompressible and independent of strain rate [9]. One of the simplest and most popular methods for describing hyperelastic materials in

nonlinear finite element analysis is the two-parameter Mooney-Rivlin model. Based on the characterization of rubber as an incompressible material, Poisson's ratio was assumed to be $\nu = 0.48$ m/m. According to study [10], for the considered rubber compound, the value of shear modulus at low deformations was chosen equal to $G = 0.97$ MPa. This value allows us to obtain Young's modulus E due to the relationship from the theory of strength of materials [11]:

$$G = \frac{E}{2 \cdot (1 + \nu)} \qquad (1)$$

where ν – Poisson's ratio, m/m; G – shear modulus, MPa; E – Young's modulus, MPa.

Based on the value of shear modulus, it is possible to determine a range of parameters C_{10} and C_{01} [12]:

$$G = 2 \cdot (C_{10} + C_{01}) \qquad (2)$$

where C_{10}, C_{01} are material constants of the Mooney-Rivlin model, MPa.

The value of the bulk modulus K for practically incompressible hyperelastic materials, according to [13] is determined in the range from 10^3 to 10^4 MPa. The defined values served as initial values for refinement of parameters for their selection according to the experimental AFR (Table I).

TABLE I. NATURAL FREQUENCIES OF THE SHOCK ABSORBER UNDER LOAD

Load mass, kg	Average value of natural frequency, Hz	Average value of maximum load displacement amplitude, mm
0.1	226	0.08815
0.2	153	0.21209
0.3	121	0.31798

The parameters C_{10}, C_{01}, K and damping coefficient η were selected as long as the model reaches a resonant frequency and maximum displacement amplitude equal to the experimental ones (Table II).

TABLE II. MATERIAL PARAMETERS OF SHOCK ABSORBER RUBBERCOMPOUND UNDER LOAD

Parameter	Load mass, kg		
	0.1	0.2	0.3
Natural frequency, Hz	226	153	121
K, MPa	1505	1260	1135
C_{10}, MPa	0.602	0.504	0.454
C_{01}, MPa	0.1505	0.1260	0.1135
η, m/m	0.295	0.215	0.235

Nonlinearity of material parameters related to load can be described by a polynomial in the form (3):

$$P(m) = a + b \cdot m + c \cdot m^2 \qquad (3)$$

where $P(m)$ is considered the parameter dependent on the load mass m, kg; a, b, c - coefficients of the polynomial.

The coefficients of material parameters are presented in Table III.

TABLE III.   POLYNOMIAL COEFFICIENTS OF MATERIAL PARAMETERS DEPENCE ON LOAD

Parameter	Coefficients of the polynomial		
	a	b	c
K, MPa	$1.87 \cdot 10^9$	$-4.25 \cdot 10^9$	$6.00 \cdot 10^9$
C_{10}, MPa	$7.58 \cdot 10^5$	$-1.70 \cdot 10^6$	$2.40 \cdot 10^6$
C_{01}, MPa	$1.87 \cdot 10^5$	$-4.25 \cdot 10^5$	$6.00 \cdot 10^5$
η, m/m	0.18	0.20	0.00

Figures 2–4 show both the experimental and calculated frequency response of the shock absorber under load.

Fig. 2. Comparison of model and experimental shock absorber AFR under 0.1 kg load

Fig. 3. Comparison of model and experimental shock absorber AFR under 0.2 kg load

Fig. 4. Comparison of model and experimental shock absorber AFR under 0.3 kg load

In this case the disadvantage of the considered Mooney-Rivlin model is the difference of steepness in the near-resonanceregion. The actual AFR is characterized by a sharper pattern of increase and decay of oscillations in the resonant region. In this task, it is necessary to estimate the maximum displacements of the cluster points in the region of significant resonances of the structure. Therefore, this model is applicable.

IV.   MODEL VERIFICATION

To protect the HRG IMU from external disturbances, a vibration damping system has been developed. It consists of four shock absorbers arranged in a "triangular pyramid" configuration. This layout offers good stability and meets the requirements for ease of installation and replacement. It does not require the development of additional fastening units and ensures strict requirements for compliance with tolerances of bases and angular connection dimensions. The IMU's center of mass coincides with the center of rigidity of the vibration damping system (Fig. 5).

Fig. 5. HRG IMU's model

According to the results of the design, a prototype of the HRG IMU was assembled. An experiment was conducted to determine the AFR of the cluster's points. A vibration acceleration sensor was fixed on the IMU's basement. The data were converted into vibration displacement during post-processing.

The mass-size evaluation model of the cluster and its components were built, and the properties of shock absorbers were specified, taking into account the load applied to them. The results are presented in Fig. 6. The estimation of the

output level of impact is given as the amplification factor of the current vibration.

Fig. 6. Comparison of experimental and simulation AFR of HRG IMU's affected by axial and transverse vibration

The calculations demonstrate sufficient validity of the results. Therefore, the model can be used to estimate the maximum displacements of the HRG cluster.

V. CONCLUSION

The paper presents an approach to evaluate the performance of the damping system, validated by the results of full-scale tests and simulation studies. Experiments conducted according to the developed methodology enabled the specification of the rubber characteristics. The use of a model for the vibration damping system gives the opportunity to significantly reduce the computational load, simplify tests, and evaluate the nature of reactions to the applied impacts. Verification of the finite element model provided its implementation at the stage of IMU's design development to estimate the maximum displacements and the damping system gain coefficient.

The shock absorber model meets the requirements for accurate of displacement estimation in regions of significant resonances. The disadvantage of this method is its inaccuracy of determining the gain for low-power resonances of the system. The Mooney-Rivlin model was adopted for its simplicity (2 related parameters) and the absence of the necessity for complex compression, tension and torsion studies. To achieve better results, it is imperative to use more complex models of hyperelastic or viscoelastic material, that involve selecting properties with a large number of material

constants, which can be a time-consuming task. A possible solution is to represent the additional properties of the shock absorber as a transfer element within the oscillation system.

REFERENCES

[1] Belyaev Y.K., Bogatyrev V.A., Bolotin V.V. *Nadezhnost' tekhnicheskikh sistem: Spravochnik* [Reliability of technical systems: Handbook]. Russia, Moscow: Radio i svyaz', 1985.

[2] Lunin B.S., Matveev V.A., Basarab M.A. *Volnovoi tverdotel'nyi giroskop. Teoriya i tekhnologiya* [Wave solid-state gyroscope. Theory and technology]. Russia, Moscow: Radiotekhnika, 2014.

[3] Merkur'ev I.V., Podalkov V.V. "Vliyanie vibratsii i udarov na dinamiku mikromekhanicheskogo vibratsionnogo giroskopa" [Effect of vibration and shock on the dynamics of a micromechanical vibrating gyroscope]. *Trends in Applied Mechanics and Mechatronics*, vol 1. Russia, Moscow: INFRA-M, 2015, pp. 109-119.

[4] Kasperovich A.V., Shashok Z.S., Vishnevskii K.V. *Tekhnologiya proizvodstva rezinotekhnicheskikh izdelii* [Production technology of rubber products]. Belarus, Minsk: BGTU, 2014.

[5] Surovtsev Y.A. *Amortizatsiya radioelektronnoi apparatury* [Shock absorber damping of radio and electronic equipment]. Russia, Moscow: Sovetskoe radio, 1974.

[6] "Tsifrovoi dvoinik – element, kotorogo tak ne khvatalo!" [A digital twin is an element that has been sorely missing!], *CAD/CAM/CAE Observer*, no. 6(114), p. 56, 2017. [Online]. Available: http://www.cadcamcae.lv/N114/56.pdf [Accessed: Apr. 2, 2024]

[7] Kazakova O.I., Smolin I.Y., Bezmozgii I.M. "Analiz amplitudno-zavisimykh dempfirovanii i vozmozhnosti ikh primeneniya pri raschete chislennymi metodami" [Analysis of amplitude-dependent damping and its application to numerical calculation]. *Vestnik Tomskogo gosudarstvennogo universiteta. Matematika i mekhanika*, no. 54: 2018, pp. 66-78.

[8] Ucar H., Basdogan I. *Dynamic characterization and modeling of rubber shock absorbers: A comprehensive case study.* Journal of Low Frequency Noise, Vibration and Active Control, vol. 37, issue 3, 2018, pp. 509-518.

[9] Sukonkina M.L, Gainov S.I. Obzor metodov i ustroistv vibrozashchity pribornykh platform [Review of methods and devices for vibration protection of sensor platforms]. *Trudy Nizhegorodskogo gosudarstvennogo universiteta im. R.E. Alekseeva*, no. 4 (101): 2013, pp. 311-319.

[10] Efremov S.K, Mikhailenko E.V.. Reshenie prostranstvennoi kraevoi zadachi nelineinoi teorii uprugosti dlya kombinirovannogo rezinometallicheskogo paketa tsilindricheskoi formy [Solution of the spatial edge of the nonlinear elasticity theory for a combined resinometallic cylindrical shape], *Nauchnye vedomosti BelGU. Ser. Matematika. Fizika*, vol. 50, no. 4, 2018, pp. 498-506.

[11] Belyaev N.M. *Soprotivlenie materialov* [Strength of materials]. Russia, Moscow: Nauka, 1965.

[12] Marchenko E.S., Kozulin A.A., Vetrova A.V., Baigonakova G.A. *Modelirovanie napryazhenno-deformirovannogo sostoyaniya pri otsenke mekhanicheskogo povedeniya konstruktsii iz nikelida titana.* [Modeling of stress-strain state in estimation of mechanical properties of titanium nickelide structures]. Russia, Tomsk: Izdatel'stvo Tomskogo gosudarstvennogo universiteta, 2021.

[13] Bonet J., Wood R.D. *Nonlinear Continuum Mechanics for Finite Element Analysis*. Cambridge, Cambridge Univ. Press, 1997.

Suppression of Anchor Loss in Hemispherical Resonator Based on Vibration Mode Optimization

Ning Wang
Space Contol and Inertial
Technology Research Center
Harbin, China
Ning.W@stu.hit.edu.cn

Guoxing Wang
Space Contol and Inertial
Technology Research Center
Harbin, China
ygx@.hit.edu.cn

Zhennan Wei
Space Contol and Inertial
Technology Research Center
Harbin, China
wzn@.hit.edu.cn

Yan Huo
Space Contol and Inertial
Technology Research Center
Harbin, China
yhuo@.hit.edu.cn

Lishan Yuan
Space Contol and Inertial
Technology Research Center
Harbin, China
lyuan@.hit.edu.cn

Yiwei Sun
Space Contol and Inertial
Technology Research Center
Harbin, China
syw@.hit.edu.cn

Abstract—This paper analyses the anchor loss of the hemispherical resonator. The mechanism of the hemispherical resonator anchor loss is revealed, and the optimization of the hemispherical resonator vibration mode is proposed. By adjusting the dimensional structure of the hemispherical resonator so that its swing modal node is exactly located at the connection position between the hemispherical shell and the anchor stem. This design effectively isolates the vibration of the hemispherical shell from the external anchoring structure. In this way, the vibration energy of the hemispherical resonator is not easily transmitted to the external environment, significantly reducing the energy loss of the hemispherical resonator due to anchoring. The finite element simulation results show that the design significantly suppressed the anchoring energy loss of the imperfect hemispherical resonator.

Keywords—Hemispherical resonator gyro(HRG), anchor loss, modal design, Quality factor(Q)

I. INTRODUCTION

The hemispherical resonator gyro (HRG) has been widely developed in the past few years due to its advantages of high accuracy, fast start-up speed, and long life[1-4]. The hemispherical resonator, as the core component of the HRG, the quality factor is a key factor restricting the development of the HRG's performance[5-6].

As the core parameter of the hemispherical resonator, the quality factor has been widely studied[7-11]. It has been shown that the quality factor of hemispherical resonators is mainly affected by thermoelastic damping[8], material damping, surface loss[11], and anchoring loss[7]. Fused silica with very low material damping and thermoelastic damping is an ideal material for the preparation of hemispherical resonators[12]. Surface loss can be suppressed by magnetorheological polishing and chemical etching[13]. Nevertheless, the anchor loss is related to the hemispherical resonator mechanical structure and anchor type. The anchor loss is the main mode of energy loss in the hemispherical resonator.

Normally the hemispherical resonator operates in the four-wave amplitude mode. For an ideal hemispherical resonator, the azimuthal angle of the standing wave is only related to the external input angular rate due to the perfect symmetry of its structure concerning the rotation axis. However, the machining process inevitably causes the hemispherical resonator to lose its perfect axial symmetry. The asymmetric mass distribution leads to vibration of the center of mass during the four-wave amplitude vibration of the hemispherical resonator, which in turn leads to the transfer of vibration energy to the support base, resulting in anchorage loss. Huo Yan et al. [6], [13] established a hemispherical shell dynamics model based on the elastic thin-shell theory and analyzed the effect of inhomogeneous mass on the hemispherical harmonic oscillator. The results show that the first four harmonics of the inhomogeneous mass are the main factors affecting the mechanical properties of the hemispherical resonator. The fourth harmonic of the inhomogeneous mass causes the frequency split of the hemispherical resonator, and the first three harmonics of the inhomogeneous mass significantly reduce the quality factor of the hemispherical resonator, which leads to an inhomogeneous distribution of the quality factor in the cyclic direction, and thus affects the accuracy of the hemispherical resonance gyroscope. Due to the complex structure of the hemispherical resonator, theoretical modeling makes it difficult to accurately describe its actual vibration characteristics. Ali Darvishian [7] has analyzed the effect of machining defects on the anchor loss of the hemispherical resonator using a finite element method (FEM) with a perfect-matched layer (PML). It is shown that any imperfection in the shell or misalignment between the shell and the stem increases the anchoring loss by several orders of magnitude. The above literature focuses on the generation mechanism of anchor loss. However, it does not reveal the impact mechanism of the vibration modes of the hemispherical resonator on the anchor loss. Revealing the anchor loss mechanism from the effect of modes can help guide the hemispherical resonator's structural design and reduce the anchor loss.

In this paper, the anchor loss of the hemispherical resonator is investigated. Firstly, the effect of anchor height on the hemispherical resonator's modal vibration shape and quality factor is analyzed using the finite element perfect-matched layer method. Subsequently, the energy dissipation mechanism is explored by analyzing the mode shapes, and it is proposed to avoid the energy transfer from the hemispherical shell to the outside through the design of the node position. Finally, the effectiveness of the above energy loss mechanism and node design is verified using finite element perfect-matched layer frequency response analysis.

II. FINITE ELEMENT SIMULATION OF THE HEMISPHERICAL RESONATOR

As shown in Fig. 1(a), the hemispherical resonator consists of a hemispherical shell and an anchor stem, usually in the shape of an umbrella. The hemispherical shell is excited as a four-wave amplitude standing wave, as shown in Fig. 1(b). When there is an external angle Ω input, the standing wave will proceed in the opposite direction φ, and the external input angle Ω can be obtained by measuring the azimuthal position of the standing wave. The ideal hemispherical resonator is an axisymmetric structure, and the four-wave amplitude vibration of the spherical shell hardly causes the vibration of the hemispherical resonator anchor rod. Thus, it has a shallow anchor loss. However, the actual processing of the hemispherical resonator will inevitably have dimensional errors, resulting in the hemispherical harmonic resonator mass circumferential inhomogeneous distribution. The inhomogeneous distributed mass will drive the anchor to vibrate causing anchor loss, which affects the quality factor and the homogeneity of the quality factor.

Fig. 1. (a) Schematic diagram of hemispherical resonator structure and anchor loss; (b) Schematic diagram of the standing wave progress of the hemispherical resonator.

Hemispherical resonator energy is dissipated from the substrate through elastic waves. By obtaining the energy transferred to the substrate, the quality factor of the hemispherical resonator determined by the anchor loss can be obtained. A perfect-matched layer (PML) is used, where the PML is added to the outside of the substrate to absorb the energy transferred from the substrate. The quality factor can be calculated by the following equation:

$$Q = \frac{|\mathrm{Re}(\omega)|}{|2\,\mathrm{Im}(\omega)|} \qquad (1)$$

where $|\mathrm{Re}(\omega)|$、$|\mathrm{Im}(\omega)|$ are the real and imaginary parts of the inherent frequency ω of the mode, respectively.

Fig. 2. (a) Schematic of the structure of the hemispherical resonator, substrate and perfectly matched layer (b) Diagram of the finite element meshing.

The simulation model shown in Fig. 2(a) was established using COMSOL software, which contains the hemispherical resonator, the substrate, and the PML. The outside of the PML is a fixed constraint, the different structures are bounded connections, and the other boundaries are free. The thickness of the PML is usually larger than the wavelength of

the solid elastic wave to ensure that the elastic wave is fully absorbed:

$$R_{PML} = \lambda + r_{PML} \qquad (2)$$

where R_{PML} is the diameter of the perfectly matched layer, λ is the wavelength of the solid elastic wave, and r_{PML} is the radius of the finite element substrate. Usually $r_{PML} = 20r$, r is the radius of the hemispherical resonator.

The wavelength is given by Eq:

$$\lambda = \frac{C}{f} \qquad (3)$$

where C is the wave speed of sound propagation in a solid, and f is the frequency of vibration.

For longitudinal waves in solid materials, C can be calculated from the following equation:

$$C = \sqrt{\frac{E(1-v)}{\rho(1+v)(1-2v)}} \qquad (4)$$

where E, v and ρ are Young's modulus, Poisson's ratio, and material density, respectively.

In addition to the size of the PML, the mesh delineation will also seriously affect the results of the calculation. Mapping and sweeping are chosen to delineate the mesh of the PML part to ensure the uniformity and denseness of the mesh delineation. A tetrahedral mesh is used for the hemispherical resonator and the substrate, and the mesh is encrypted in some areas. Mesh-independence verification was performed to ensure the accuracy of the calculations.

The structure's modal analysis was first carried out using the finite element PML method. The mode shapes, inherent frequencies, and quality factors of the main modes of the structure are analyzed. The dimensional parameters of the hemispherical resonator are shown in Table I.

TABLE I. DIMENSIONAL PARAMETERS OF HEMISPHERICAL RESONATORS

Dimensional parameters	Value(mm)
R_1	10.4
R_2	9.6
r_1	2
r_2	1.5
d	6.8
t	0.8
h	4

Fig. 3 illustrates the dimensions of the hemispherical resonator used for the simulation and the main mode shapes. The main modes of the hemispherical resonator are the four-wave amplitude vibration mode, the respiration mode, the first-order swing mode, and the second-order swing mode. The four-wave amplitude vibration mode is the working mode of the resonator. Due to the high symmetry of its structure and vibration mode, almost no energy is transferred out through the anchor rod during vibration in this mode, and this mode has a high-quality factor. However, due to the processing defects of the actual machined hemispherical resonator, the energy between modes is coupled to each other. The energy of the four-wave amplitude vibration mode of the hemispherical resonator is transferred to the

neighboring modes for dissipation. It can be seen from the mode shapes that the first- and second-order swing anchors of the resonator have large displacements that transfer energy to the fixed base and thus lead to rapid energy dissipation.

Fig.3. Dimensions and main mode shapes of the hemispherical resonator.

Fig. 4 illustrates the effect of anchor height h on each mode's frequency near the hemispherical resonator's operating frequency. The inherent frequency of the four-wave amplitude vibration of the hemispherical shell is 5409 Hz, which is not affected by the anchor height h. The other modes, on the other hand, show a significant decreasing trend with increasing anchor height h. The first-order swing mode frequency of the hemispherical resonator decreases at the fastest rate. In the modal design, the frequency of the non-working modes should be separated from the frequency of the working modes with a large interval, so as to reduce the energy transfer between the modes. When the anchor height h is at 2 mm and 4 mm, the respiration mode and the first-order swing mode are close to the four-wave amplitude vibration mode, and this interval should be avoided when designing the anchor height h. The second-order swing mode of the hemispherical resonator decreases from 7000 Hz to 5800 Hz with the increase of anchor height h, which is always larger than the operating frequency of the hemispherical resonator. However, as the anchor height h increases, it gradually approaches the operating mode of the hemispherical resonator, which is unfavorable.

Fig.4. Effect of anchor height h on the inherent frequencies of the hemispherical resonator.

Fig. 5 shows the effect of anchor height h on the quality factor of each mode of the hemispherical resonator. It can be seen that the quality factors of the hemispherical resonator four-wave amplitude mode and the breathing mode of the hemispherical resonator are significantly higher than those of the first-order and second-order swing modes of the anchor

stem, which are all higher than 1×10^9. The effect of the anchor height h on the quality factor does not change significantly. On the other hand, the quality factor of the swing modes of the hemispherical resonator is lower and is significantly affected by the anchor height h, with the quality factor ranging from 3000 to 80000. The first-order swing modes of the anchors are more sensitive to the anchor height h, and there exists a minimal value when the anchor height h is 4 mm. The quality factor of the second-order swing modes of the anchors has a slight increase with the increase of anchor height h.

Fig.5. Effect of anchor height h on the quality factors of the hemispherical resonators

III. OPTIMISED DESIGN OF HEMISPHERICAL RESONATOR MODES

Based on the above modal analysis, an equivalent model for the energy dissipation of the hemispherical resonator is proposed, as shown in Fig. 6(a). The modal mass of the hemispherical shell part of the hemispherical resonator can be denoted as M, and dM is the asymmetric inhomogeneous modal mass on the shell. The inhomogeneous mass dM transfers energy to the neighboring modes of the four-wave amplitude vibrational mode. Assume that the energy that dM has on it is dE. The energy dE that the inhomogeneous mass dM has is proportional to the total energy E that the hemispherical resonator has:

$$dE = E \times \frac{dM}{M} \tag{5}$$

The energy dE possessed on dM is partially transferred to the other modes of the hemispherical resonator, and then the energy gained by the other order mode of the hemispherical resonator is $\delta_i dE$:

$$\delta_i dE = K \cdot dE \tag{6}$$

where K is the ratio coefficient, which affects the energy transfer between the modes of the hemispherical resonator.

The unbalanced mass dM and the ratio coefficient K are the key factors affecting the vibration quality factor of the hemispherical resonator. In addition to reducing the unbalanced mass dM, the quality factor can be improved by reducing the ratio coefficient K through the optimization of the modal design. There are two design principles for reducing the K:

1) Avoid that the inherent frequencies of other modes are close to the resonant frequencies of the working modes;

2) Design the node of the modal vibration shape at the connection between the hemispherical shell and the anchor

stem to avoid the hemispherical shell energy dissipation through the anchor stem as much as possible.

Fig. 6. (a) Energy dissipation model of the hemispherical resonator; (b) Nodal design.

From the results of the modal analysis, it can be seen that the quality factor of the first-order swing mode and the second-order swing mode of the resonator is low, and its anchor energy dissipation is high. The resonant frequency and mode shape of this mode can be changed by adjusting the anchor height h or anchor diameter. As shown in Fig. 6(b), the simulation results show that different anchorage heights h affect the nodes of the second-order swing mode of the hemispherical resonator. When h is 4 mm, the second-order swing mode of the resonator is similar to the first-order oscillating mode and does not have nodes. When h is 10 mm, the node position is located at the connection between the hemispherical shell and the anchor stem. As the anchor height h increases, the node moves downward, and when h is 16 mm, the node is located below the connection between the spherical shell and the anchor stem.

IV. SIMULATION

In order to verify the accuracy and validity of the above-proposed node design, further finite element PML frequency response analysis was carried out. Based on the model in Fig. 2, a simple harmonic force F of 1 N is applied at the lip edge of the resonator. The swept frequency ranges from 5350 to 5374 Hz near the operating frequency with a frequency interval of 0.1 Hz.

Fig.7. Results of the harmonic response of the hemispherical resonator at different anchor heights.

Fig. 7 shows the effect of the excitation frequency and the anchor height h on the total energy possessed by the hemispherical resonator. It can be seen that the harmonic frequency of the resonator is at 5372 Hz, and the variation of the anchor height h hardly affects the operating frequency of

the resonator. However, the anchor height h significantly affects the total energy of the hemispherical resonator during the frequency response. The hemispherical resonator has the largest total energy of 6×10^{-5} J at an anchor height h of 10 mm, indicating that the hemispherical resonator has the highest quality factor and only a small amount of energy is dissipated from the anchor. Comparing the nodal design presented in Fig. 6(b), it is found that at an anchor height h of 10 mm, the second-order swing mode at this time is just at the anchor stem and hemispherical shell connection. In contrast, when the anchor height h is at a non-nodal position, the hemispherical resonator resonance energy decreases rapidly, highlighting the importance of the nodal position.

V. CONCLUSION

In this paper, the effect of anchor height on the quality factor of the hemispherical resonator is investigated. The individual inherent frequencies and modal shape of the hemispherical resonator are analyzed using the finite element method with a perfectly matched layer, and the vibrational energy loss mechanism of the imperfect hemispherical resonator is revealed. The concept of nodal design is proposed on the basis of the modal frequency design of the hemispherical resonator. By designing the node of the second-order swing mode of the hemispherical resonator at the connection between the hemispherical shell and the anchor stem, the modal coupling and energy transfer can be avoided, thus improving the quality factor of the hemispherical resonator. The finite element frequency response is used to verify the accuracy and validity of the design concept.

REFERENCES

[1] A. Jeanroy, A. Bouvet, and G. Remillieux, "HRG and marine applications," Gyroscopy and Navigation, 2014, vol. 5, no. 2, pp. 67–74.

[2] F. Delhaye, "SpaceNaute® the HRG Based Inertial Reference System of Ariane 6 European space launcher," Gyroscopy and Navigation, 2019, vol. 10, no. 1, pp. 1–6.

[3] A. Jeanroy, G. Grosset, J.-C. Goudon, and F. Delhaye, "HRG by Sagem from laboratory to mass production," IEEE International Symposium on Inertial Sensors and Systems, 2016.

[4] D. M. Rozelle, "The Hemispherical Resonator Gyro: From Wineglass to the Planets," AAS/AIAA 19th Space Flight Mechanics Meeting, in Advances in the Astronautical Sciences, vol. 134. Feb. 2009, pp. 1157–1178.

[5] N. Wang, G. Yi, Z. Wei, Y. Sun, and L. Yuan, "Non-contact laser-based test method for vibration performance parameters of hemispherical resonator," Advanced Fiber Laser Conference (AFL2023), P. Zhou, Ed., Shenzhen, China: SPIE, Mar. 2024, p. 384.

[6] Y. Huo, Z. Wang, S. Ren, Z. Wei, G. Yi, and C. Wang, "Standing wave drift of hemispherical resonator with quality factor non-uniformity under a ring electrode excitation," Chinese Journal of Aeronautics, 2022, vol. 35, no. 1, pp. 160–172.

[7] A. Darvishian, B. Shiari, J. Y. Cho, T. Nagourney, and K. Najafi, "Anchor Loss in Hemispherical Shell Resonators," Journal of Microelectromechanical Systems, 2017, vol. 26, no. 1, pp. 51–66.

[8]   A. Darvishian, T. Nagourney, J. Y. Cho, B. Shiari, and K. Najafi, "Thermoelastic Dissipation in Micromachined Birdbath Shell Resonators, Journal of Micro-electromechanical Systems," 2017, vol. 26, no. 4, pp. 758–772.

[9]   T. Nagourney, J. Y. Cho, A. Darvishian, B. Shiari, and K. Najafi, "Effect of metal annealing on the Q-factor of metal-coated fused silica micro shell resonators," IEEE International Symposium on Inertial Sensors and Systems (ISISS) Proceedings, 2015.

[10] Z. Wei, G. Yi, Y. Huo, B. Xi, and Y. Zhao, "High-precision synchronous test method of vibration performance parameters for fused quartz hemispherical resonator," Measurement, 2021, vol. 185.

[11] J. Chu et al., "Fundamental investigation of subsurface damage on the quality factor of hemispherical fused silica shell resonator," Sensors and Actuators A: Physical, 2022, vol. 335.

[12] B. Lunin and V. Lopatin, "Silica Glass for High-Q Mechanical Resonators," Inorganic Materials, 2020, vol. 56, no. 3, pp. 292–296.

[13] Y. Huo, Z. Wei, S. Ren, and G. Yi, "High Precision Mass Balancing Method for the Fourth Harmonic Mass Defect of Fused Quartz Hemispherical Resonator Based on Ion Beam Etching Process," IEEE Transactions on Industrial Electronics,2022, pp. 1–10.

Electrical Balancing of Wave Solid-State Gyroscope with Flat Electrodes

A.A. Maslov

Moscow Power
Engineering Institute,
Russia
Maslov954@ya.ru

D.A. Maslov

Moscow Power
Engineering Institute,
Russia

I.V. Merkuryev

Moscow Power
Engineering Institute,
Russia
MerkuryevIV@ya.ru

Abstract—**We consider the hemispherical resonator gyro with flat electrodes and electrostatic control sensors. The mathematical model of a gyroscope describing resonator oscillations during electrical balancing applicable to compensate frequency difference is constructed. The resonator dynamics is studied by taking into account the nonlinear effects caused by electrical balancing. It is shown that the resonator balancing by using eight electrodes does not completely compensate for the frequency difference. It is recommended to use sixteen electrodes for electrical balancing.**

Keywords— hemispherical resonator gyro, flat electrodes, nonlinear oscillations, drift, electrical balancing.

I. INTRODUCTION

One of the most promising gyroscopes for inertial navigation systems is the hemispherical resonator gyros (HRG), also referred to as wave solid-state gyros in Russian publications. The analysis of publications [1-3] shows that the scope of HRG research and development will expand in the near future, contributing to an increase in the accuracy of gyroscopes. Due to its characteristics and the possibility of increasing accuracy, HRG is becoming the most promising type of gyroscope for navigation systems in various fields of application [2,4,5].

The operating principle of the HRG is based on the inertia effect of elastic wave oscillations of a hemispherical resonator, which is created from quartz glass, which has an invariable isotropic pattern and quality factor. The metallized surface of the resonator, together with the electrodes, provides electrostatic sensors for control and information retrieval. HRG with a hemispherical resonator belongs to high and medium precision gyroscopes. Achieving high accuracy for HRG is possible if the hemispherical resonator has a high-quality factor (Q~107), as well as high-quality manufacturing and precise balancing, which significantly reduce the impact of the anisotropic properties of the resonator on the measurements obtained. The justification for the HRG includes methods for manufacturing a quartz hemispherical resonator and its cost, which is associated with the use of a number of precision technological operations [6]. Therefore, recently, the works aimed to improve the design of the HRG, associated with reducing the overall dimensions of the structure, simplifying the technology for the location of power and measuring electrodes have been carried out. It is known that the edge of a hemispherical resonator moves during oscillations of movement not only in reference to the axis of the resonator, but also parallel to it. This makes it possible to connect all the electrodes on a flat electrode assembly, stabilizing it parallel to the end edge of the

hemispherical processing of the resonator. In the new version of the HRG design, the metal coating is applied only to the end edge of the hemisphere, which significantly reduces internal friction in the resonator and simplifies its manufacturing technology. The metal coating of the edge of the hemisphere together with the base electrodes form electrostatic sensors for control and information readout. In this design, it is much easier to ensure the uniformity of the flat gap between the end surface of the resonator and the flat electrode assembly [6]. Using this design scheme, the French company Safran Electronics & Defense developed and launched the production of a high-precision HRG [4] with a quartz hemispherical resonator with a diameter of 20 mm.

It follows from publications that the most effective method to increase the HRG accuracy is to balance the resonator; however, this is one of the most labor-intensive and expensive operations of assembling and adjusting the HRG. The need for balancing arises due to small deviations of real resonator from the ideal axisymmetric shape and the appearance of an imbalance, leading to a splitting of the resonator's natural frequencies and to decreasing quality factor due to the dissipation of oscillations energy in the supports. The accuracy characteristics of the device largely depend on the quality of balancing [4,6]. Experimental data demonstrating the effectiveness of balancing hemispherical resonators by ion sputtering is presented in [6]. The HRG drift decreased by two orders of magnitude, and the quality factor increased by an order of magnitude. In this case, the residual frequency difference was 0.004 Hz. The French company Safran Electronics & Defense [4] has developed the procedure that balances the resonator and reduces frequency mismatch. Its duration is several hours. As a result, the frequency difference does not exceed 0.0005 Hz, the quality factor is 107, and the drift is less than 0.01 deg/h, regardless of the mechanical impact to which the object is subjected.

In the considered designs of HRG quartz resonators, the residual splitting of the natural frequency remaining after mechanical balancing was 0.004 Hz and 0.0005 Hz. To further correct residual imbalances without removing material in the gyroscope operating mode, electrostatic balancing based on applying a constant voltage of various magnitudes to individual electrodes is used [4,6,7]

Control of resonator oscillations, including balancing, is carried out by applying voltage to flat electrodes, which form a system of electrical capacitors with the resonator. A mathematical model of an HRG with electrodes on the spherical surface of a resonator, which describes linear electrical and mechanical oscillations in an interconnected form, including during electrical balancing, was obtained in [7]. A linear mathematical model of an HRG with flat electrodes is presented in [8].

Works [10-11] take into account the impact of nonlinear properties of electrostatic control sensors on the dynamics of HRG with a ring and cylindrical resonator. In [12-16] it is shown that the reference voltage causes a drift of the HRG when taking into account nonlinear terms in the dynamics equations. In this work, the goal is to study the dynamics of a hemispherical HRG resonator with flat electrodes during electrical balancing, taking into account the nonlinearity caused by electrical processes in the gyroscope control loop.

II. EQUATIONS OF SMALL FORCED RESONATOR OSCILLATIONS

We consider a thin elastic hemispherical resonator 1 (Fig. 1), the metallized edge 3 of which, together with flat electrodes 4 located on the base 2, forms sensors for control and information retrieval. The HRG resonator is made of fused quartz glass. It is assumed that the elastic properties of the resonator material are isotropic; instrumental manufacturing errors are not taken into account.

Let the hemispherical shell rotate around its axis of symmetry with an angular velocity Ω, which we will further consider to be small compared to the natural oscillation frequency ω, i.e. $\Omega / \omega \ll 1$.

Fig. 1. Wave solid-state gyroscope with a flat node of electrodes: 1 - resonator; 2 – gyroscope base; 3 – metallized resonator surface; 4 - electrode; 5 - the junction of the resonator with the base

As a model, we will use the equations of resonator dynamics HRG with flat electrodes [16]:

$$m\ddot{f} + cf - \frac{C_0}{2d^2} \sum_{i=1}^{n} \frac{U_i^2 \cos 2\theta_i}{(1 - f\cos 2\theta_i - g\sin 2\theta_i)^2} = 0,$$

$$m\ddot{g} + cg - \frac{C_0}{2d^2} \sum_{i=1}^{n} \frac{U_i^2 \sin 2\theta_i}{(1 - f\cos 2\theta_i - g\sin 2\theta_i)^2} = 0,$$

(1)

where $f = f_* / d$, $g = g_* / d$ are dimensionless generalized coordinates, f_* and g_* are generalized coordinates of the main second waveform, equal to the displacement in the direction of the HRG axis at two fixed points of the end surface of the resonator, spaced from each other at an angle of $\pi/4$; d is the gap between the undeformed resonator and the electrodes; m – the reduced mass of the resonator corresponding to the main waveform, equal to hemispherical resonator $m = 4.8\rho hR^2$; h – thickness of the resonator; R – radius of the hemisphere; $c = m\omega^2$ – reduced stiffness; $\omega = 2.62 \frac{h}{R^2} \sqrt{\frac{E}{3(1+\sigma)\rho}}$ – the frequency of free oscillations of a hemispherical resonator[16]; E – Young's modulus; σ – Poisson's ratio; ρ – density of the resonator material; $C_0 = \varepsilon_0 S / d$ – electrode capacitance with an

undeformed resonator; $\varepsilon_0 = 8.85 \cdot 10^{-12}$ F/m – electric constant S – electrode area; d – nominal gap between an undeformed resonator and electrodes; U_i – potential difference between resonator and control electrodes; $\theta_i = 2\pi(i-1)/n$ – the angle determining the position of the i th electrode ($i = 1,2, ...,n$). In equations (1) and further, the dot indicates the differentiation in time t.

Note that the equations of dynamics of HRG with spherical resonators [7,16] have the same form. They differ only in that the generalized coordinates of the main second mode of oscillations are equal to the radial displacements, and not to the displacements in the direction of the axis of symmetry of the HRG. In addition, in [7,16], when normalizing generalized coordinates, coefficients that depended on the geometric parameters of the spherical electrodes are used.

We rewrite the equations of the system, taking into account one of the main defects of the resonator – frequency difference:

$$m\ddot{f} + cf = fb\cos 4\alpha + gb\sin 4\alpha + \frac{F_1}{d},$$

$$m\ddot{g} + cg = fb\sin 4\alpha - gb\cos 4\alpha + \frac{F_2}{d},$$

(2)

where b is the modulus of frequency difference, α is the angle of orientation of the main axes of rigidity, F_1 and F_2 are control forces determined by the formulas:

$$F_1 = \frac{1}{2dC_0} \sum_{i=1}^{n} \frac{U_i^2}{(1 - f\cos 2\theta_i - g\sin 2\theta_i)^2} \cos 2\theta_i,$$

$$F_2 = \frac{1}{2dC_0} \sum_{i=1}^{n} \frac{U_i^2}{(1 - f\cos 2\theta_i - g\sin 2\theta_i)^2} \sin 2\theta_i.$$

(3)

III. ELECTRICAL BALANCING OF THE RESONATOR

The voltages U_i supplied to the electrodes are determined by a specific control law. Let's consider a gyroscope with sixteen electrodes. For electrical balancing, a voltage of the form [7] is applied:

$$U_i = U\cos 2(\theta_i - \alpha), \qquad (i = 1,2,\ldots,16) \quad (4)$$

Substituting (4) into (3), we find the control forces

$$F_1 = \mu(2f + 3f^3 + 3fg^2 + (f + 2f^3)\cos 4\alpha +$$
$$+(g + 3f^2g + g^3)\sin 4\alpha),$$

$$F_2 = \mu(2g + 3f^2g + 3g^3 - (g + 2g^3)\cos 4\alpha +$$
$$+(f + f^3 + 3fg^2)\sin 4\alpha),$$

(5)

where $\mu = 2U^2/(dC_0)$ is a coefficient characterizing electrostatic forces. In (5), terms containing f and g in powers higher than three are discarded.

Neglecting (5) nonlinear terms with respect to f and g, we obtain

$$F_1 = \mu(2f + f\cos 4\alpha + g\sin 4\alpha),$$
$$F_2 = \mu(2g + f\sin 4\alpha - g\cos 4\alpha).$$

(6)

Thus, forces F_1 and F_2 are formed, which at $\mu = -bd$ compensate for the difference in frequency in equations (2).

In this case, the terms *bf* and *bg* appear, causing a change in the oscillation frequencies of the resonator.

There are HRG designs that use not only sixteen control electrodes, but also eight. Therefore, consider a gyroscope with eight electrodes. In this case, $\theta_i = \pi(i-1)/4$ the angle that determines the position of the i th electrode ($i = 1, 2,\ldots, 8$).

Substituting a voltage of the form (4), ($i = 1, 2,\ldots, 8$) into (3) and linearizing the expressions for f and g, we will have

$$F_1 = \mu(f + f\cos 4\alpha),$$
$$F_2 = \mu(g - g\cos 4\alpha). \tag{7}$$

Forces F_1 and F_2, as follows from (7), can partially compensate for the difference in frequency in equations (2).

Let's consider the influence of nonlinearity that appears when electrically balancing a gyroscope with sixteen flat electrodes. Let us substitute force expressions (5) into equations (2), compensating for different frequencies (i.e. at $\mu = -bd$). Passing to dimensionless time $\tau = \omega t$, we get

$$\ddot{f} + f = \varepsilon(2f + 3(f^2 + g^2)f + 2f^3\cos 4\alpha +$$
$$+(3f^2 g + g^3)\sin 4\alpha),$$
$$\ddot{g} + g = \varepsilon(2g + 3(f^2 + g^2)g - 2g^3\cos 4\alpha + \tag{8}$$
$$+(f^3 + 3fg^2)\sin 4\alpha),$$

where $\varepsilon = -bd/(m\omega^2)$.

From equations (8) it follows that the reference voltage causes a decrease in the frequency of natural oscillations. These conclusions are consistent with the result obtained in [7]. We will study system (8) using the Krylov–Bogolyubov averaging method [18]. To do this, we bring it to a standard form by moving from variables $f, g, \dot{f}, \dot{g}$ to slowly changing variables r, k, θ, χ, called orbital elements in mechanics [7]

$$f = r\cos(\tau + \chi)\cos 2\theta - k\sin(\tau + \chi)\sin 2\theta,$$

$$g = r\cos(\tau + \chi)\sin 2\theta + k\sin(\tau + \chi)\cos 2\theta,$$

$$\frac{df}{d\tau} = -r\sin(\tau + \chi)\cos 2\theta - k\cos(\tau + \chi)\sin 2\theta, \tag{9}$$

$$\frac{dg}{d\tau} = -r\sin(\tau + \chi)\sin 2\theta + k\cos(\tau + \chi)\cos 2\theta.$$

Note that the Krylov–Bogolyubov averaging method is physically implemented in the electronic circuit of the gyroscope. High-frequency functions of time f, g measured using a capacitive system of electrodes are presented in the form of two normal vibration modes with amplitudes r and k. The orientation of the wave pattern of oscillations on a plane f, g is determined by the precession angle, the change in the phases of oscillations of two normal modes depends on the slow variable.

If the amplitudes of the main and quadrature oscillation waves ($r = k$), the precession angle is not determined, therefore this mode is not considered further.

The equations of motion in slow variables r, k, θ, χ, in the first approximation of the averaging method, have the form

$$\frac{dr}{d\tau} = -\frac{3}{8}k(r^2 + k^2)\,\varepsilon\sin 4(\theta - \alpha),$$

$$\frac{dk}{d\tau} = \frac{3}{8}r(r^2 + k^2)\,\varepsilon\sin 4(\theta - \alpha),$$

$$\frac{d\theta}{d\tau} = \frac{3}{8}\varepsilon kr + \frac{3}{8}\varepsilon kr\frac{r^2 + k^2}{r^2 - k^2}\cos 4(\theta - \alpha), \tag{10}$$

$$\frac{d\chi}{d\tau} = -\varepsilon - \frac{9}{8}\varepsilon(r^2 + k^2) - \frac{3}{4}\varepsilon\frac{r^4 + k^4}{r^2 - k^2}\cos 4(\theta - \alpha).$$

From the first two equations (10) it follows that the total energy of the resonator does not change E = (r2 + k2)/2 = const. The terms $\frac{3}{8}\varepsilon kr$ and $-\frac{9}{8}\varepsilon(r^2 + k^2)$ in the third and fourth equations of system (10) indicate a change in the drift and characteristic oscillation frequency of the resonator when the reference voltage is applied [17]. This can be explained by the fact that voltage is applied to the electrodes $U_i = U\cos 2(\theta_i - \alpha)$ ($i = 1, 2,\ldots, 16$), and the force acting on the resonator is proportional to the square of this voltage $U_i^2 = \frac{U^2}{2} + \frac{U^2}{2}\cos 4(\theta_i - \alpha)$. That is, the force has a constant component, independent of the angle θ, which causes a frequency drift and offset, the same as when a reference voltage is applied. The force component, which is dependent on the angle θ and is proportional to $\cos 4(\theta_i - \alpha)$, causes a drift $\frac{3}{8}\varepsilon kr\frac{r^2 + k^2}{r^2 - k^2}\cos 4(\theta - \alpha)$ and frequency offset $-\frac{3}{4}\varepsilon\frac{r^4 + k^4}{r^2 - k^2}\cos 4(\theta - \alpha)$ proportional to $\cos 4(\theta_i - \alpha)$. According to (10), the angular velocity of the HRG drift changes in direct proportion to the amplitudes of the main and quadrature waves of oscillations *r, k,* which is consistent with [19]. To eliminate the error caused by nonlinearity, in gyroscopes of this class the oscillation amplitude *r* is kept constant, and the quadrature *k* is sought to be reduced to zero [19].

IV. Conclusion

The study has shown that the voltage supply of this type can be used for electrical balancing of an HRG resonator with flat electrodes. At the same time, the average oscillation frequency of the resonator decreases. However, in addition to balancing, the applied voltage leads to an additional undesirable effect - gyroscope drift, which is directly proportional to the amplitudes of the main and quadrature oscillation waves. It is also shown that balancing the resonator with eight electrodes does not allow to completely compensate for the frequency difference. Therefore, sixteen electrodes are recommended for electrical balancing.

References

[1] Peshekhonov, V.G., The outlook for gyroscopy, Gyroscopy and Navigation, 2020, vol. 11, no. 3, pp. 193-197. https://doi.org/10.1134/S2075108720030062

[2] Perelyaev, S.E., Current state and scientific and technical forecast of the perspective application of foreign hemispherical resonator gyros (Analytical review of foreign materials), Novosti navigatsii, 2020, no. 3, pp. 14-28.

[3] Maslov, A.A., Maslov, D.A., Ninalalov, I.G., and Merkuryev, I.V. // Hemispherical Resonator Gyros. An Overview of Publications. Gyroscopy and Navigation, 2023, Vol. 14, No. 1, pp. 1–13. https://link.springer.com/article/10.1134/S2075108723010054

[4] Jeanroy, A., Bouvet, A., and Remillieux, G., HRG and marine applications, Gyroscopy and Navigation, 2014, vol. 5, no. 2, pp. 67-74. https://doi.org/10.1134/S2075108714020047

[5] Matveev, V.A., Lipatnikov, V.I., and Alekhin, A.V., Proektirovanie volnovogo tverdotel'nogo giroskopa (Designing a Wave Solid-State Gyro), Moscow: MSTU Publishing, 1997.

[6] Lunin, B.S., Matveev, V.A., and Basarab, M.A., Volnovoi tverdotel'nyi giroskop. Teoriya i tekhnologiya (Wave Solid-State Gyro. Theory and Technology), Moscow: Radiotekhnika, 2014.

[7] Zhuravlev, V.F. and Lynch D.D., Electrical model of the wave solid-state gyroscope, Izv. RAN MTT, 1995, no. 5, pp. 12–24.

[8] Klimov, D.M., Zhuravlev, V.F., and Zhbanov, Yu.K.,Kvartsevyi polusfericheskii rezonator (volnovoi tverdotel'nyi giroskop) (Quartz Hemispherical Resonator of a HRG), Moscow: Kim L.A., 2017.

[9] Wei, Z., Yi, G., Huo, Y.; Qi, Z., Xu, Z. The Synthesis Model of Flat-Electrode Hemispherical Resonator Gyro. Sensors 2019, 19, 1690. DOI 10.3390/s19071690

[10] Maslov, A.A., Maslov, D.A., and Merkuryev, I.V., Nonlinear effects in dynamics of cylindrical resonator of wave solid-state gyro with electrostatic control system, Gyroscopy and Navigation, 2015, vol. 6, no. 3, pp. 224-229. https://doi.org/10.1134/S2075108715030104

[11] Maslov, D.A. and Merkuryev, I.V., Impact of nonlinear properties of electrostatic control sensors on the dynamics of a cylindrical resonator of a wave solid-state gyroscope, Mechanics of Solids, 2021, vol. 56, no. 6, pp. 960-979. https://doi.org/10.3103/S002565442106011X

[12] Maslov, A.A., Maslov, D.A., Merkuryev, I.V., Podalkov, V.V. Methods to Eliminate Nonlinearity of Electrostatic Control Sensors of the Wave Solid-State Gyroscope // 25th Anniversary Saint Petersburg International Conference on Integrated Navigation Systems, Proceedings, 2018, pp. 288-290. DOI: https://doi.org/10.23919/ICINS.2018.8405895

[13] Denisov, R.A., Maslov, A.A., Maslov, D.A., Merkuryev, I.V., Podalkov V.V. How the reference voltage of electromagnetic control sensors affects the drift of wave solid-state gyroscopes // Gyroscopy Navig. 2016, no. 3, p. 231–238. https://doi.org/10.1134/S2075108716030032

[14] Maslov, A.A., Maslov, D.A., Merkuryev, I.V., and Mikhailov, D.V., Drift of wave solidstate gyro in the presence of reference voltage on control electrodes, Vestnik MEI, 2013, no. 2, pp. 11–14.

[15] .Maslov, A.A., Maslov, D.A., Merkuryev, I.V. The Effect of the Reference Voltage on the Drift of a Wave Solid-State Gyroscope with Flat Electrodes // 30th Saint Petersburg International Conference on Integrated Navigation Systems, Proceedings, 2023, pp. 331-334. https://doi.org/10.23919/ICINS51816.2023.10168316

[16] Maslov, A.A., Maslov, D.A., Merkuryev, I.V., Nonlinear effects in the dynamics of a wave solid-state gyroscope with flat electrodes // Giroskopiya i Navigatsiya, 2023, vol. 31, no. 4, pp. 67-77. EDN BPZTEA, DOI: 10.1134/S2075108724700044

[17] Strutt, J. W. (Baron Rayleigh). The theory of sound. Vol. I. London: Macmillan & Co. Publ., 1877. 470 p.

[18] Bogolyubov N.N., Mitropol'skij Yu.A. Asimptoticheskie metody v teorii nelinejnyh kolebanij (Asymptotic methods in the theory of nonlinear oscillations.). Moscow: Nauka, 1974.

[19] Zhuravlev, V.F. Foucault's controlled pendulum as a model of one class of free gyroscopes // Izv. RAN MTT, 1995, no. 5, pp. 27–35.

Simulation of the Frequency Mismatch Caused by the Inclined Errors of the Assembly of the Hemispherical Resonator Gyroscope

Y. Tao
National University of Defense Technology
Changsha, China
taoyunfengcn@foxmail.com

K. Deng
National University of Defense Technology
Changsha, China
dengkaixin19@163.com

Y. Pan
National University of Defense Technology
Changsha, China
panyao08@nudt.edu.cn email

W. Wu
National University of Defense Technology
Changsha, China
wuweihiter@163.com

K. Yang
National University of Defense Technology
Changsha, China
yky208@nudt.edu.cn

H. Luo
National University of Defense Technology
Changsha, China
luohui.luo@163.com

Abstract—**The assembly of the resonator and the flat electrode is the key process during the manufacturing of the hemispherical resonator gyroscope. However, the inclined error is inevitable during the assembly. Therefore, the electrostatic force produced by the bias voltage varies at different central angles, leading to an additional frequency mismatch. The theoretical model of the inclined assembly has been widely investigated. While the FEM analysis is rarely investigated. In this study, theoretical analysis and simulations have been conducted to investigate the frequency mismatch caused by the inclined error. The theoretical relationship between the gap and the inclined angle is deduced. A finite element model is established to calculate the frequency mismatch caused by the unevenness of the electrostatic force. The frequency mismatch under different inclined angles, bias voltages, and average gaps are investigated. The frequency mismatch reaches the order of 10^{-3} Hz when the inclined angle is 200 arc second and the average gap is 30 μm. The simulation results provide a guidance for choosing the average gap and controlling the inclined error.**

Keywords—hemispherical resonator; inclined assembly; frequency mismatch

I. INTRODUCTION

The hemispherical resonator gyroscope (HRG) is a kind of solid-state gyroscope with advantages of high accuracy, high reliability, and small size. Achieving high-precision gyroscope puts stringent requirement on the symmetry of the core elements. The assembly of the resonator and the flat electrode is a key process during the manufacturing of HRGs. Welding and gluing are usually employed to connect the resonator and the flat electrode, and the gap between the resonator lip and the flat electrode is about 30 μm -100 μm.

Although the flat-electrode structure proposed by Sagem is simpler and more compact than the sphere-electrode structure, assembly errors are inevitable [1]. The inclined error and the eccentric error are the primary errors of assembly [2, 3]. The inclined error directly affects the gap uniformity. Efforts have been made to identify and compensate for the assembly errors. Hang et al. established a mathematical model between the assembly pose error and the equivalent capacitance. The algorithm of identification was realized using the support vector regression [4]. Ruan et al. investigated the formation mechanism of the inclined

assembly. The inclined angle and the azimuth of the inclination were successfully obtained using the Nelder-Mead simplex algorithm. The bias instability reduced from 2.470°/h to 2.322°/h after compensation [5]. Ning et al. deduced the theoretical model of the inclined assembly. Based on the theoretical calculation, the frequency mismatch was at the order of 10^{-4} Hz when the inclined angle was 60 arc second for the 30mm-diameter resonator [6]. In addition to the theoretical analysis, researchers are also devoted to improving the assembly precision. Medicon developed an assembling machine that was made up of precision alignment module, gap measuring module, and acoustic parameter measuring module. The assembling machine adjusts the pitch and the location of the resonator according to the real-time gap. Yuan et al. also designed an assembly apparatus to control the gap between the resonator and the flat electrode. The average gap was 33 μm，and the maximum deviation was 3μm by using the assembly apparatus [7]. The nonuniformity of the gap is still quite large, even though the assembly is assisted by the precise apparatus. The nonuniformity of the gap leads to the nonuniformity of the electrostatic force, which results in the additional frequency mismatch. The resonators are well-balanced before assembly, their frequency mismatch are usually lower than 1 mHz. However, the frequency mismatch may increase to several mill hertz after assembly. Deng et al. proposed a method to separate the mass and gap errors. The frequency mismatch caused by the mass error remained unchanged when the bias voltage varies. While the frequency mismatch caused by the gap error will change at different voltages [8].

The theoretical model of the imperfect assembly has been intensively investigated. However, accurate experimental verification is hard to conduct, because it is difficult to control the inclined angle precisely during the actual assembly process. Moreover, the FEM analysis is rarely conducted to investigate the influence of the inclined assembly on the frequency mismatch. In this paper, theoretical analysis and simulations have been conducted to investigate the frequency mismatch caused by the inclined assembly of the resonator and the flat electrode. The actual gap and the actual electrostatic force are deduced based on the flat capacitors. The results show that the gap varies as a cosine function of the central angle. The unevenness of the gap is proportional to the

inner radius of the resonator and is inversely proportional to the average gap. The relationship between the electrostatic force and the central angle is also investigated. The frequency mismatch of the $n=2$ mode under different bias voltages, average gaps, and inclined angles are acquired. The influence caused by the inclined error is analyzed based on the simulation results.

II. INFLUENCE OF THE INCLINED ASSEMBLY

As shown in Fig. 1, the hemispherical resonator is assembled with an inclined angle of α. The line AB is defined by the lowest point and the highest point of the resonator lip. The projection of the line AB on the xoy surface is named as the inclined direction. The inclined direction coincides with the x-axis in this case. The surface of the flat electrode is coated with gold. The gold film is cut into eight or sixteen electrodes.

Fig. 1. The scheme of inclined assembly

The gap between the resonator lip and the flat electrode at different central angles can be expressed as

$$d = d_0 - r \cos \varphi \sin \alpha \qquad (1)$$

where r is the radius of the hemispherical resonator, α is the inclined angle, d_0 is the average gap, φ is the central angle.

The gap nonuniformity ΔG can be defined as $\Delta G = (r\sin\alpha/d_0)\times100\%$. Fig. 2 shows the gap distribution at different central angles where the radius of the hemispherical resonator is 10 mm and the average gap is 50 μm. When the inclined angle reaches 20 arc second, the maximum of the gap error is 0.97μm. Therefore, the inclined assembly has a significant influence to the gap between the resonator and the flat electrode.

Fig. 2. The gap distribution at different central angle.

The resonator lip and the electrodes form capacitors. Hundreds of bias voltage are applied to each capacitor. The electrostatic force of the unit area can expressed as

$$df = \frac{1}{2} \frac{\varepsilon_0 \varepsilon_r V^2}{\left(d_0 - r \cos \varphi \sin \alpha\right)^2} dA \qquad (2)$$

where ε_0 is vacuum dielectric constant, ε_r is relative dielectric constant, dA is the unit area, and V is the bias voltage.

The nonuniformity of the electrostatic force ΔK can be defined as

$$\Delta K = \frac{df_{max} - df_{min}}{df_{max} + df_{min}} \times 100\%$$

where $\qquad (3)$

$$df_{max} = \frac{1}{2} \frac{\varepsilon_0 \varepsilon_r V^2}{\left(d_0 - r \sin \alpha\right)^2} dA, df_{min} = \frac{1}{2} \frac{\varepsilon_0 \varepsilon_r V^2}{\left(d_0 + r \sin \alpha\right)^2} dA$$

According to Eq. (2), the electrostatic force varies from different central angles. The inclined angle leads to the nonuniformity of the gap. The resonator and the flat electrode form a capacitor whose capacitance is determined by the gap. The uniformity of the capacitance directly affects the electrostatic detection and the excitation. The variation of the gap further results in the nonuniformity of the electrostatic force. The electrostatic force produced by the bias voltage is time-independent. Thus, the force acts as a prestress to the resonator during the simulation.

III. SIMULATION RESULTS

The finite element simulation is used to investigate the frequency mismatch caused by the inclined errors. The pre-stressed analysis type is chosen in the Comsol. The electrostatic force is applied to the lip of the resonator as a boundary load. As shown in Fig 3, the boundary load coincides with the symmetry axis of the resonator. The distribution of the stress caused by the uneven electrostatic force is first acquired. Then, the eigenfrequency of the resonator of the $n=2$ mode is calculated. The additional stress changes the resonator's original stiffness, leading to the frequency mismatch.

Fig. 3.The scheme of the load on the resonator lip

The material and the structural parameters of the resonator are listed in Table 1.

TABLE I.   MATERIAL AND STRUCTURAL PARAMETERS
OF THE RESONATOR

Density	2203 kg/m^3
Young's modulus	71.7 GPa
Poisson's ratio	0.17
Radius of resonator	10 mm

The grid of the resonator is elaborately meshed. The frequency mismatch of the resonator of the $n=2$ mode is only 0.000801Hz. The influence of the bias voltages, the average gap, and the inclined angle on the frequency mismatch of the $n=2$ mode are investigated. Fig. 4 shows the frequency mismatch under bias voltage from 50V to 300V when the inclined angle α is 100 arc second and the average gap d_0 is 50μm. The frequency mismatch of the $n=2$ mode is proportional to the square of the bias voltage. The frequency mismatch increases from 0.000803 Hz to 0.000831 Hz when the bias voltage reaches 300 V. Higher bias voltage leads to a lager frequency mismatch.

Fig. 4. The frequency mismatch of the $n=2$ mode under different bias voltages.

The influence of the inclined angle on the frequency mismatch is shown in Fig. 5. The average gap maintains 50 μm. The frequency mismatch of the $n=2$ mode is 0.000857 Hz when the inclined angle α is 200 arc second and the bias voltage is 200 V. The frequency mismatch of the $n=2$ mode is 0.000926 Hz when the inclined angle α is 200 arc second and the bias voltage is 300 V. The frequency mismatch of the $n=2$ mode is proportional to the square of the inclined angle. Although the signal-to-noise ratio benefits from higher bias voltage, the frequency mismatch increases at a rapid rate when the bias voltage rises. The inclined error needs to be controlled during the assembly of the resonator and the flat electrode to reduce the frequency mismatch.

The frequency mismatch under different inclined angles when the average gap is 30μm, 50μm, and 70μm, respectively, is shown in Fig. 6. The bias voltage is 300V during the simulation. The frequency mismatch increases rapidly when the average gap decreases from 70μm to 30μm. The frequency mismatch is 0.000809Hz when the average gap is 70μm. The frequency mismatch is 0.001033Hz when the average gap is 30μm. As the increasing of the average gap, the gap nonuniformity gets smaller. For the average gap 30μm, 50μm, and 70μm, the gap nonuniformity are 16.1%, 9.7%, and 6.9%, respectively, when the inclined angle is 100 arc second.

Moreover, the electrostatic force and the nonuniformity of the electrostatic force decrease simultaneously. The maximum electrostatic force of the unit area is 629 N/m^2, and the nonuniformity of the electrostatic force is 31.5% when the average gap is 30 μm. The maximum electrostatic force of the unit area is 93.8 N/m^2, and the nonuniformity of the electrostatic force is 13.7% when the average gap is 70 μm.

Fig. 5. The frequency mismatch of the $n=2$ mode under different inclined angles.

Fig. 6. The frequency mismatch of the $n=2$ mode under different average gaps and inclined angles.

The simulation results indicate that the frequency mismatch of the $n=2$ mode is affected by the inclined errors of the assembly. Although the frequency mismatch caused by the inclined assembly is smaller than that caused by the geometric nonuniformity, it still has a detrimental effect on the high-precision gyroscopes.

IV. CONCLUSION

Simulation results illustrate that the frequency mismatch is proportional to the square of the inclined angle and the bias voltage. The smaller inclined angle, smaller bias voltage and larger average gap can reduce the frequency mismatch. Furthermore, the inclined assembly not only leads to the frequency mismatch but also leads to the gain errors. Therefore, the inclined angle should be carefully controlled by high-precision assembly. The simulation results provide guidance for choosing the average gap and controlling the inclined error.

REFERENCES

[1] A. Jeanroy, A. Bouvet, and G. Remillieux, "HRG and marine applications," Gyroscopy Navig., vol. 5, no. 2, pp. 67–74, 2014.

[2] R. Zhihu, D. Xukai, G. Yang, Q. Zhengcheng, and L. Hongsheng, "Analysis and Compensation of Bias Drift of Force-to-Rebalanced Micro-Hemispherical Resonator Gyroscope Caused by Assembly Eccentricity Error," J. Microelectromech. Syst., 2023.

[3] X. Zeyuan, Y. Guoxing, and W. Zhennan, "Influence analysis of the pole's eccentricity error on vibration characteristics of resonator," Journal of Chinese Inertial Technology, vol. 27, no. 1, pp. 101–107, 2019.

[4] Y. Hang et al., "Identification Method for Assembly Pose Errors of Flat Electrode Hemispherical Resonator Gyro," IEEE Sensors J., vol. 24, no. 4, pp. 4258–4266, 2024.

[5] R. Zhihu, D. Xukai, Q. Zhengcheng, J. Jia, and L. Hongsheng, "Modeling and Compensation of Assembly Inclination Error of Micro Hemispherical Resonator Gyroscope Under Force-to-Rebalance Mode," IEEE Sensors J., vol. 21, no. 13, pp. 14726–14738, 2021.

[6] N. Youhuan, Y. Guoxing, X. Boqi, X. Weinan, W. Zhennan, and W. Changhong, "The influence of assembly inclination error in hemispherical resonator on frequency splitting," Journal of Chinese Inertial Technology, vol. 29, no. 4, pp. 510–515, 2021.

[7] Y. Lishan, Z. Qingshuang, W. Changhong, R. Shunqing, W. Zhennan, and H. Yan, "High Precision Assembly and Welding Method for Flat-Electrode Hemispherical Resonator Gyro," IEEE Trans. Instrum. Meas., vol. 72, 2023.

[8] D. Kaixin et al., "Separation of mass and gap errors of hemispherical resonator gyroscopes by varying bias voltage," Measurement, vol. 228, 2024.

Estimation of the Resonant Frequency of Vibrations of a Solid-State Wave Gyro

S.E. Perelyaev

Ishlinsky Institute for Problems in Mechanics
of the Russian Academy of Sciences,
e-mail: ipm@ipmnet.ru

A.V. Alekhin

JSC Inertial technologies of technocomplex,
Ramenskoye, Russia,
e-mail: aalehin@itts.ru

Abstract – **The solution to the current problem of estimating the frequency of working vibrations of a hemispherical quartz resonator of a solid-state wave gyro (SWG) in the mode of switching information and control signals is considered. For a continuous information retrieval process, an algorithm for finding the frequency has been successfully developed, but in the conditions of switching information and control signals it is not operational. In the switching mode, the time interval for observing signals is limited to a few periods, which complicates the signal processing algorithm.**

Keywords - wave solid-state gyroscope, information-measurement channels of SWG, signal switching, vibration frequency.

I. INTRODUCTION

In this paper, an important problem of determining the vibration frequency of a wave solid-state gyroscope is considered. Knowing the vibration frequency makes it possible to control errors associated with temperature changes. The vibration frequency is not included in the parameters of the state vector of the device, therefore, the academic approach based on the solution of motion equations does not work in this case [1].

For this reason, a separate development is required, based on the analysis of the primary SWG signal from the standpoint of well-known digital methods for processing primary information and measurement signals.

II. SETTING THE TASK

Let the primary signal of the information-measuring system of the device have the form [2]:

$$z(t) = A\sin(\frac{2\pi t}{T} + \alpha) + kt + b \tag{1}$$

where A is the amplitude, T is the period, α is the signal phase; k, b are the coefficients of the linear trend, t is the time interval with a duration of N periods.

For gyroscopes with a quartz hemispherical resonator of the SWG class, there are usually several information channels, any of them can be considered, the signal frequency of all channels is the same. The presence of a linear trend is characteristic of the switching mode.

We consider period T to be known (in the first approximation). Over time, there is not an integer number of periods in the observation interval.

The reason for this is a shift in the frequency of vibrations. The observation interval is represented as $N(T + \delta)$, where δ is the error of the period change. The task is to determine parameter δ. The control of a mechanical standing wave in a quartz hemispherical resonator, which is a sensitive element of a wave solid-state gyroscope, is carried out by applying voltage to a system of electrodes forming,

together with a sprayed resonator, a system of electric capacitors. These voltages are controlled using signals taken from another electrode system - information (measuring) electrodes. The influence of control electrodes and information electrodes on the dynamics of a quartz hemispherical resonator is studied in sufficient detail in the works of academician V.F. Zhuravlev [1,3].

The dynamic model of the main working (second) waveform as an electromechanical system, which is a hemispherical resonator with control electrodes and information retrieval electrodes, is considered in detail in the well-known joint publication by D. Lynch and V.F. Zhuravlev [3]. The information electrodes allow measuring the generalized coordinates w1 and w2 of the main (second) form of elastic vibrations, equal to the radial displacement of the resonator at two fixed points separated from each other by an angle of 45° in angular measurement along the equator of the resonator. The dynamic equations of a hemispherical quartz resonator relate both to the case of information retrieval electrodes and to the case of control electrodes.

In known SWG designs, there are usually eight information removal electrodes (n=8), and the radial movements of the quartz resonator $w1$ and $w2$ represent real movements at the location of the electrodes with numbers $i = 1,2$.

In the case of control electrodes equal to $n= 16$ or n=8, the resistance R is small and it can be considered equal to zero. The distribution of voltage across the correction and control electrodes depends on the specific technical objectives of the control task. In this case, the amplitude of elastic vibrations (vibrations) of the resonator edge is maintained constant by the control electrodes. The amplitudes of the two measuring channels are proportional to the sine and cosine of the angle of orientation of the standing wave relative to the instrument body.

(1)

III. THE SOLUTION TO THE PROBLEM

We believe that we know the initial value of period T. We introduced a linear trend $kt + b$, since this error is characteristic of the switching mode, and greatly distorts the calculation of the frequency (period). We suggest subtracting the linear trend from the primary signal before starting the processing. Although the method for calculation of the linear trend is well known, we propose a simpler algorithm using the harmony of the primary signal.

Let us consider two convolutions of the signal we are studying:

$$S_1 = \int_0^{NT} z(x)\,dx$$

$$S_2 = \int_0^{NT} z(x)\sin\frac{4\pi x}{T}\,dx. \tag{2}$$

On the one hand, the left parts (2) are calculated numerically from experimental data, and the right parts after transformations are written as

$$S_1 = \frac{1}{2}T^2 N^2 k + TNb ,$$

$$S_2 = \frac{1}{4\pi}T^2 Nk .$$

(3)

From system (3), it is easy to find the linear trend parameters k and b. After subtracting the linear trend, the signal under study will be recorded as purely harmonic:

$$z(t) = A\sin(\frac{2\pi t}{T} + \alpha) .$$

(4)

We know the initial value of period T. Suppose that there was not an integer number of periods in the signal sampling interval (4), for example, $N(T + \delta)$ (N is an integer, δ – a small error associated with a change in the signal period). Consider the following convolutions of the signal (4):

$$R_s = \int_0^{N(T+\delta)} z(x)\sin(\frac{2\pi x}{T})dx ,$$

$$R_c = \int_0^{N(T+\delta)} z(x)\cos(\frac{2\pi x}{T})dx .$$

(5)

The left parts of formulas (5) are calculated digitally from experimental (measured) samples of the signal, the right parts are converted to the form (keeping the second degree of error δ):

$$\frac{R_s}{NA} = \frac{T}{2}\cos(\alpha) + \frac{\pi N}{T}\sin(\alpha)\delta^2 ,$$

$$\frac{R_c}{NA} = \frac{T}{2}\sin(\alpha) + \sin(\alpha)\delta + \frac{\pi N}{T}\cos(\alpha)\delta^2 .$$

(6)

Note that retaining only one first degree of δ in the system of equations (6) leads to a degenerate system, therefore, it is necessary to retain the second degree.

System (6) includes unknowns A, α, δ, so that it is very difficult to solve such a system analytically.

Therefore, we apply the method of successive approximations. Let $\delta = 0$ be a zero approximation; in the unperturbed system (6), we find the values of the unknowns.

From the system of equations (6), under the condition $\delta = 0$, we find the following variable parameters A_*, α_*:

$$\sin(\alpha_*) = \frac{2}{NTA_*}R_c , \quad \cos(\alpha_*) = \frac{2}{NTA_*}R_s$$

$$A_* = \frac{2}{NT}\sqrt{R_s^2 + R_c^2}$$

(7)

Substituting (7) into the perturbed system (6), we obtain the equation for the first approximation:

$$\frac{\pi N}{T}\delta^2 + \frac{1}{2}\sin(2\alpha_*)\delta + \frac{T}{4}\sin(2\alpha_*)$$
$$- \frac{1}{NA_*}(R_s\sin(\alpha_*) + R_c\cos(\alpha_*)) = 0$$

(8)

Let us write the solution of equation (8) as

$$\delta = \frac{T}{2\pi N}\left[\begin{array}{c} -\frac{1}{2}\sin(2\alpha_*) \\ +\sqrt{\frac{1}{4}\sin^2(2\alpha_*) - \pi N\sin(2\alpha_*) + \frac{4\pi}{TA_*}(R_s\sin(\alpha_*) + R_c\cos(\alpha_*))} \end{array} \right] .$$

(9)

The resulting formula (9) gives a value (in the first approximation) for changing the period of signal δ, thus completely solving the task.

The new value T_1 of the period can be found as $T_1 = T + \delta$. Value T_1 can be taken as a new initial value for processing the next sample of the signal. Recall that N is an integer number of periods in the initial sample of the signal.

IV. CONCLUSION

The problem of estimating the frequency of working vibrations of the WSG resonator in the mode of switching measurement and control signals has been completely solved. Note the importance of expressions (7) for calculating the signal phase; the calculation of the signal phase is necessary for the development of frequency tracking systems (PLL and similar). If necessary, the solution according to formula (9) can be inserted into system (6) and the following approximation can be found. However, the practice of processing experimental results shows that the first approximation according to formula (9) is quite sufficient for a reliable estimation of the signal frequency. Practical studies of the SWG have shown that the effect of temperature on the change in the natural drift velocity is 10^0/h, after the introduction of the resonant frequency model - no more than 0.05^0/h.

The results of the report can be used not only in SWG research, but also in any tasks related to estimating the frequency of a harmonic signal using digital methods in conditions of low-frequency interference and a limited amount of data.

REFERENCES

[1] 1. Klimov D.M., Zhuravlev V.F., Zhbanov Yu.K. Quartz hemispherical resonator (wave solid-state gyro). - M.: Kim L.A., 2017.

[2] 2. Matveev V.A., Lipatnikov V.I., Alekhin A.V. Development of a wave solid-state gyroscope. Bauman Moscow State Technical University, 1998, 167 p.

[3] 3. Zhuravlev V.F., Lynch D.D. Electric model of a wave solid-state gyroscope. Izv. RAS, MTT, No.5, 1995.

Investigation of the Accuracy Characteristics of Matrix Methods of Linear-Angular Measurements

Alexander Korolev
Laser Measurement and Navigation Systems Dept.
Saint Petersburg Electrotechnical University «LETI»
St Petersburg, Russia
al_korolev@mail.ru

Alexander Lukin
Department of Physics
Peter the Great Saint Petersburg Politechnical University
St Petersburg, Russia
alexander_lukin@inbox.ru

Pavel Ivanov
Laser Measurement and Navigation Systems Dept.
Saint Petersburg Electrotechnical University «LETI»
St Petersburg, Russia
ivanov_etu@mail.ru

Eugeny Bokhman
Laser Measurement and Navigation Systems Dept.
Saint Petersburg Electrotechnical university «LETI»
St Petersburg, Russia
edbokhman@mail.ru

Yuri Filatov
Laser Measurement and Navigation Systems Dept.
Saint Petersburg Electrotechnical University «LETI»
St Petersburg, Russia
yvfilatov@etu.ru

Abstract – The paper presents the results of the experimental studies and simulation of the accuracy in measuring the coordinates of the elements of the mark image. The possibilities of estimating distortions in the coordinates of image elements associated with lens distortion and mark distortions, as well as correcting these images, are shown. The obtained deviation of the coordinates of the elements of the mark image from the coordinates of the elements of the global matrix after correction of distortions is less than 0.1 microns. Experimental results of estimating the random error of measuring the values of the angle, coordinate and period of the mark are obtained. The estimates obtained coincide with the forecast and with the theoretical estimates.

Keywords – linear-angular measurement, 2D scale, digital camera, optical mark.

I. INTRODUCTION

The creation of modeling stands for testing and calibration of navigation equipment requires the use of precision linear-angular measurement methods. To implement these methods, it is necessary to measure linear movements at the level of fractions of micrometers and angular rotations at the level of fractions of an arc second. The most commonly used approach to performing linear and angular displacements is the use of encoders or interference systems on diffraction gratings. The use of these systems makes it possible to achieve very high accuracy both in measuring linear and angular displacements [1,2]. CMOS and CCD cameras are widely used in various fields of technology to solve problems of measuring linear displacement and/or rotation. The matrix measurement technology considered in this paper is based on the use of a measuring mark with tens of thousands of elements. Preliminary studies conducted earlier were devoted to solving the problems of restoring the periodic structure of the mark image and eliminating distortions associated with lens distortion and mark manufacturing errors. The predicted error of this technology is at the level of fractions of a nanometer on a linear scale and thousandths of an arc second on an angular scale. [3,4].

In this work, a mark was used (Fig.1) with a period of 50 microns and a diameter of elements of 30 microns. To carry out the research, a mock-up of a matrix measuring system with a heat sink and thermostating was used. As indicated in [3], the most important parameter of matrix meters that determines their resolution is the scale size, which, in turn, is limited by the size of the photodetector matrix of a digital camera. In these studies, a matrix angle meter (MAM) with a 1.3 Megapixel photodetector matrix was used.

Fig.1. Optical mark configuration

In the design of the MAM, a protective housing with a water cooling system was used, and thermal protection was made in the form of an additional casing covered with foil

insulation. Since the 1.3 Megapixel camera has a small heat dissipation, the proposed design made it possible to keep the temperature of the meter at 25 ° C.

The MAM used is characterized by the following parameters.

Digital CMOS Camera BR-1341LM-U [4]

1.3 Megapixels, matrix size 6.6x5.3 mm, pixel size 5.2 microns,

1280x1024 pixels

Lens: magnification G=1x.

Mark: chrome photomask on glass, the diameter of the elements is 30 microns, the period is 50 microns. The diameter of the scale (area of analysis) is 5 mm. The number of mark image elements in the analysis area is 10,000.

II. ANGLE MEASUREMENTS

Studies of the matrix angle meter were carried out on the rotating platform of the SG-03 goniometer [6] of INERTECH Ltd. The rotary platform is equipped with a LIR-3200 SKB IS sensor [7]. The sensor uses a holographic radial scale with a diameter of 120 mm. The scale step is 1 micron. The resolution of the sensor is 0.01 arc-sec.

When measuring an angle, as a rule, there is a combination of a deterministic cyclic error, which can be studied and corrected in detail, and a random error. It is the latter that limits the accuracy of angle measurement. The causes of cyclic error for MAM can be incomplete correction of lens distortion and mark manufacturing errors, the influence of uneven lighting, defects in brand manufacturing. The reasons for such errors for the LIR are primarily scale errors.

In the absence of a standard of the required accuracy, a convenient procedure for studying angle meters is the use of two angle meters mounted on the rotor of the rotary platform. During the rotation of the rotor, the differences in the angle values are measured on 2 scales. The mutual phase shift of the scales allows, if necessary, to identify systematic errors of both sensors.

Table 1 shows fragments of the synchronous measurement protocol for MAM and LIR sensors in the range from 0 to 360 deg. at a step of the rotation angle of the rotor of the rotary platform 10 deg.

The following designations are used in Table 1:

U1 - angle reading in degrees for the MAM matrix meter,

U0 is the angle reading in degrees for the LIR-3200 sensor of the rotary platform,

R01 is the difference U1-U0 in arc seconds,

dS is the RMS deviation of the measured coordinates of the mark image elements from the coordinates of the corresponding elements of the mathematical matrix in pixels.

The values of the measured angles in columns 2 and 3 are presented in degrees up to 7 decimal places, with $1\cdot10^{-7}$ degrees corresponding to 0.0004 arc-sec.

The RMS of the measurement results of the period between the images of the mark elements is $1\cdot10^{-5}$ pixels = 0.05 nm. In this case, B is the average value of the period between the elements of the mathematical matrix obtained in each experiment as a result of averaging. Such a small value of the RMS reflects the high efficiency of matrix technology in linear measurements, allowing you to calculate the exact magnification value of the lens for massaging the measurement results.

$$V = B*5.2/50 = 1.245086^x,$$

with a level of accuracy unattainable in any optical applications.

The ds value is an indicator of the quality of measurements of image coordinates of mark elements, including a high level of correction of optical image distortions, including image tilt, distortion, mark manufacturing errors and lighting quality. The value ds=0.017 pixels corresponds to 0.1 microns.

Figure 2 shows a graph of the normalized (minus the average value) difference in the measurement results R01. The maximum deviation of the difference is 1.2 arc-sec.

Fig. 2. The dependence of R01 on the angle of rotation

It is obvious that the cyclic component present in the functional in Fig. 3 is nothing more than the sum of the systematic errors of the two sensors.

In [8], the experience of using nonlinear programming methods to calculate the parameters of a cyclic function involving a number of harmonics was described using the Excel option "Solution search" while minimizing the sum of squares of the discrepancy.

Table 1.

№	U1	U0	R01	ds	B
	deg	deg	arc-sec	pixel	pixel
1	0.23741006	229.5189222	470586.55	0.014	9.40366
2	10.23976333	239.5218348	470584.54	0.016	9.40367
3	20.23906938	249.5210825	470584.75	0.016	9.40367
4	30.24244671	259.5244471	470584.79	0.015	9.40368
5	40.24271586	269.5247328	470584.73	0.016	9.40368
..					
32	310.2347521	179.5171854	470583.24	0.016	9.40366
33	320.2317531	189.514114	470583.50	0.016	9.40366
34	330.2327544	199.515056	470583.71	0.015	9.40367
35	340.2356453	209.5178908	470583.91	0.016	9.40367
36	350.2342324	219.5164	470584.19	0.016	9.40367

Limiting ourselves to four harmonics, we determine the coefficients a_n and b_n of the cyclic function using the "Solution Search" under the condition

$$S1_n = R01 - \sum_1^4 (a_n \sin(nx) + b_n \cos(nx)) \quad (1)$$
$$\sum_N (S1_n)^2 \to min$$

where n is the measurement number.

The cyclic function, defined as the sum of 4 harmonics, is shown in Fig. 3.

The subtraction of the cyclic function from the difference of counts according to formula (1) is identical to the fact that the systematic errors of both sensors were separated (measurement cycles with mutual phase shift between scales), and correctors are built into the sensor program, eliminating the influence of systematic errors on the measurement results.

Fig. 3. The sum of the four harmonics

The subtraction of the cyclic function from the difference of counts according to formula (1) is identical to the fact that the separation of the systematic errors of both sensors (measurement cycles with mutual phase shift between scales) was performed, and correctors are built into the sensor program, eliminating the influence of systematic errors on the measurement results.

Figure 4 shows a graph of the $S1_n$ discrepancy, which is the sum of the random errors of the MIU and LIR sensors. Its maximum values do not exceed ± 0.06 arc-sec. Since random errors, unlike systematic cyclic errors, cannot be divided by sensors, the following conclusion is obvious: the random error of the MIU matrix meter cannot exceed ± 0.05 arc-sec. This is an amazing level of accuracy of angle measurement for a device with a scale diameter of 5 mm.

Fig. 4. The sum of the random errors of the MIU and LIR sensors

The forecast of the angle measurement error for the configuration of the matrix angle meter with large sizes of the photodetector matrix may be less than ± 0.01 arc-sec [1]. Obviously, to confirm this level of accuracy, it is required to use at least an order of magnitude more accurate device as a reference. However, we do not have such a standard. To confirm these forecast values, the authors suggest using another measurement technology in the near future.

III. MEASUREMENT OF LINEAR DISPLACEMENTS

Metrological studies of the matrix displacement meter at the level of nanometer units are an even more difficult task. In this case, an interferometer could be used as a reference measuring instrument, but it is too expensive to create an interferometer with a resolution of several nanometers and with a measuring range of tens of mm for such an experiment.

Therefore, it was decided to use two identical matrix meters, using them to synchronously measure movements. In this case, two independent devices are compared. The diagram of the measuring stand is shown in Fig. 5.

Fig. 5. Diagram of a linear displacement measuring stand

Figure 5 shows 2 measuring channels with identical cameras (1,4), lenses (2,5), measuring marks (3,6). The marks are fixed on the site of a linear translator 7 with a stepper motor 8 for the software implementation of movements.

A linear miniature translator 8MT18-13 [9] with a displacement range of 13 mm and a resolution of 1 step - 1 μm was used.

The program provides simultaneous operation of two channels, forming two protocols of measurement results with each movement of the platform with marks. The program allows you to output "measured coordinates corresponding to any element of the mark" to the protocol. The definition in quotation marks needs to be clarified. In this case, the camera generates an image of the mark and calculates the coordinates of each element of the mark in the coordinate system of the photodetector matrix. Then, using the least squares method, the parameters of the mathematical matrix closest to the set

of coordinates of all analyzed elements in the mark image are searched. The indexes in the description of the elements of the mathematical matrix coincide with the indexes of the corresponding elements in the image of the mark. The value of dS is about 0.1 microns. The transition to a mathematical matrix is a way of averaging the coordinates of the entire set of analyzed elements. And the RMS of the average value is $1/\sqrt{N}$ of the RMS of a single measurement. At N = 10000, we get an increase in accuracy by 2 orders of magnitude. Thus, instead of the usual physically real stroke, the coordinates of which determine the movement, the matrix meter uses a virtual coordinate calculated over the entire set of elements, but tied to a specific element. Next, we will call these virtual coordinates the coordinates of the mark elements for shortening.

The measurement protocol displays the X and Y coordinates of the "central" (conditionally) element located in the middle of the height of the isosceles triangle connecting the reference points (Fig. 1). And if the starting point of the matrix is the point at the apex of the triangle with indices (0,0), then the "central" point has indices (0,18).

Due to the inability to provide complete measurement protocols, especially for 2 channels, we will give some basic data on the experiment: the number of measurements is 52, the range of movements is 320 microns, the average step of movement is 6.3 microns, the duration of the experiment is about 6 minutes. The main indicator determining the accuracy of measuring linear movements is the measurement of the coordinates of the brand elements. The average value of this value was 0.0002 pixels ≈ 1 nm

Based on the protocol data, the displacement value at each stage of the experiment is calculated using the formula

$$L_n = \sqrt{(X_n - X_1)^2 + (Y_n - Y_1)^2}$$

where the differences in parentheses are the differences of the measured coordinates along the X and Y axes at the current and initial points of the experiment.

The result is shown in Fig. 6.

Fig. 6. Random error in measuring linear displacements

This graph, as in the previous angle measurement experiment, represents the total random error of the used matrix length meters. The maximum deviation does not exceed ± 15 nm, and the total error is about 8 nm.

Obviously, the random measurement error of each of these meters cannot exceed the specified values

Thus, the real possibility of measuring displacements at the level of nm units by fairly simple means has been confirmed. Interestingly, this level of accuracy of length measurement is realized using a scale (measuring mark) for which the error of the coordinates of the elements is ± 1 μm.

It is quite obvious that with such a high level of accuracy, the practical use of matrix meters is possible only in the presence of highly efficient heat removal, thermal insulation and vibration protection systems.

IV. CONCLUSION

Studies of matrix meters made in the form of prototypes with a heat sink and thermal protection have been carried out. Based on experiments using cross-calibration, it was concluded that the random angle measurement errors for the matrix meters used do not exceed ± 0.06" and ±15 nm, respectively. Considering that cameras with small photodetector arrays and with a small field of analysis (N≈10000) were used in the experiments, the authors are convinced that it is possible to significantly improve the measurement accuracy in the presence of large-dimensional photodetector matrices with an analysis field that provides the ability to measure from hundreds of thousands to 1 million elements.

REFERENCES

[1] Bridges A., Yacoot A., Kissinger T., Humphreys D.A., Tatam R.P. "Correction of periodic displacement non-linearities by two-wavelength interferometry", Measurement Science and Technology, vol.32, №12, p. 125202, 2021,

[2] Changhai Zhao, Qiuhua Wan, Lihui Liang, "Compensation for Dynamic Subdivision Error When the Grating Displacement Sensor Code Disk is Stained", IEEE Sensors Journal, vol.23. №3. p.2403, 2023.

[3] Korolev, A.N., Lukin, A.Ya., Filatov, Yu.V., Venediktov, V.Yu., "Matrix technology of linear–angular measurements", Journal of Optical Technology, v. 89, №12, p. 733-739, 2022.

[4] Korolev, A.N., Lukin, A.Ya., Filatov, Yu.V., Venediktov, V.Yu., "Matrix measurement technology. The accuracy of measuring the coordinates of the elements and the control of the photomasks", Journal of Optical Technology, v. 91, №3, p. 115-123, 2024.

[5] http://www.es-exp.ru/

[6] http://inertech-ltd.com/

[7] https://skbis.ru/

[8] Korolev, A.N., Lukin, A.Ya., Bokhman, E.D., Venediktov, V.Yu., Ivashchenko E.M., "Research on methods of regression analysis and nonlinear programming for calibrating a goniometer", Journal of Optical Technology, v. 89, №6, c. 346-352, 2022.

[9] https://www.standa.lt/products/catalog/motorised_positioners?item=674&prod=miniature-motorized-linear-stage&print=1

A Technique for Calculating the Transient Temperature Profile in the Internal Cavity of a Strapdown Inertial Navigation System

A.V. Frolov
JSC «CNIIAG»
Moscow, Russia
frolov@frolov.moscow

Yu.V. Mikhaylov
JSC «CNIIAG»
Moscow, Russia
mikhaylov.yu@gmail.com

P.A. Shapovalov
JSC «CNIIAG»
Moscow, Russia
kinik.keni@yandex.ru

Abstract — **This paper describes a technique for calculating transient heat transfer in a closed cavity of a strapdown inertial navigation system. The presented technique for heat transfer calculation is based on an iterative solution of the heat conduction task in ANSYS Mechanical, during which the average air temperature is calculated at each time step of a given transient process. The technique adequacy and results accuracy were assessed. It was shown that the calculation results are in satisfactory agreement with experimental measurements and qualitatively describe the nature of the considered thermal transient processes, with the error less than 6%.**

Keywords—SINS, numerical simulation, ANSYS, heat transfer

I. INTRODUCTION

The trend towards increasing the accuracy of modern strapdown inertial navigation systems (SINS) [1, 2] requires advanced computational studies during the design and development of devices.

Significant causes for the occurrence and accumulation of errors in the output of SINS sensitive elements are inertial and thermal loads.

The computational analysis of inertial loads is quite straightforward, considering the availability of modern engineering simulation software packages, while taking into account the thermal expansion of the SINS parts due to internal heat generation of electronic components, heat exchange inside the device and external thermal loads requires careful consideration of thermal boundary conditions due to their strong influence on the final result [3, 4, 5].

One of the most difficult phenomena is the heat transfer in the closed cavities of the device, since for convective boundary conditions widely used for thermal analysis, both quantities are unknown: temperature and heat transfer coefficient; in the case of transient thermal processes, their change over time is also unknown.

The most accurate approach for simulation of heat transfer in the closed cavities is by means of computational fluid dynamics (CFD) and conjugate heat transfer. Application of such approach to SINS analysis is presented, for example, in [6], where it was shown that the heat transfer intensity is low, and the average heat transfer coefficient can be considered as constant. Thus, in order to solve considered task, it is necessary to estimate the temperature inside the device as a function of time, which is proposed to be done using the iterative method.

II. DESCRIPTION OF THE TECHNIQUE

With help of the developed technique, it is possible to perform a transient analysis of the SINS thermal state by the iterative method, solving heat conduction equations in ANSYS Mechanical. In this case, the average air temperature in the internal space of the SINS is calculated iteratively with assumption of fixed heat transfer coefficient.

Iterative recalculation of the air temperature in the internal space of the device is carried out on the basis of the Newton-Richmann law by the iterative method with the power law of relaxation aimed to stabilize the convergence. The technique allows you to calculate several isolated cavities simultaneously.

The system of equations is given below:

$$\begin{cases} T_{i+1}^t = \begin{cases} T_i^t + \beta_i \cdot \dfrac{Q_i^t}{A_j \cdot \alpha_j}, & Q_i^t < 0 \\[2mm] T_i^t - \beta_i \cdot \dfrac{Q_i^t}{A_j \cdot \alpha_j}, & Q_i^t > 0 \end{cases} \\[2mm] \beta_i(i) = \beta_0 \cdot \gamma^{i-1} \end{cases}$$

where T_i^t is the air temperature in the internal space of the device at time t of the transient thermal process at the i-th iteration of the calculation;

T_{i+1}^t – air temperature at $(i+1)$-th iteration of the calculation;

$\beta_i(i)$ – power law of relaxation, where it is assumed that $\beta_0 = 0.001$; $\gamma = 0,8$;

Q_i^t – heat flow through the surfaces on the j-th internal cavity at time t of the transient thermal process at the i-th iteration of the calculation;

A_j – surface area of the j-th internal cavity;

α_j – heat transfer coefficient from the surfaces of the j-th internal cavity.

The technique is implemented in the form of a Python script using the capabilities of the ANSYS Mechanical API (Application Programming Interface). Figure 1 represents the algorithm of the developed technique, showing the sequence of operations during the calculations.

Fig. 1. Algorithm of the calculation technique

Verification is carried out on the basis of a test problem solution – simulation of heat transfer for two contacting cubes with cavities and heat generation blocks inside them (see Fig. 2) in two different ways:

• using the developed technique;

• by solving the coupled problem of hydrodynamics and heat transfer inside these cavities in ANSYS Fluent and, at the same time, solving the problem of thermal conductivity in ANSYS Mechanical; iterative data transfer is carried out using the ANSYS System Coupling module.

Fig. 2. Calculation model for the test problem: a) solid model; b) finite element model; c) thermal BC; d) named sets of internal surfaces

Correct calculation of the SINS transient temperature field using the described technique implies development of a calculation model which contains a finite element mesh generated for considered parts, assigned material properties, contact pairs, as well as all significant thermal boundary conditions except the internal convection in cavities. The internal convection boundary conditions are generated by script during the solution.

The results of CFD calculations are shown in Figs. 3–5, which demonstrate good qualitative and quantitative agreement of temperature fields obtained by different methods.

In order to validate the developed technique, the calculation results were compared with experimental temperature measurements of the SINS presented in [6].

III. VERIFICATION AND VALIDATION OF THE TECHNIQUE

In order to assess the adequacy and accuracy of the developed technique, verification was carried out based on comparison of calculation results with the results obtained by means of CFD and conjugate heat transfer [7, 8], as well as validation based on comparison of calculation results with experimental temperature measurements.

a)

b)

Fig. 3. Calculation results using computational fluid dynamics: a) air circulation velocity vectors; b) air temperature in the calculation area

a)

b)

Fig. 4. Comparison of temperature fields: a) calculation using the developed technique; a) calculation using a verification model

The calculation model took into account the heat generation of the electronic components of the inertial measurement unit (the values were taken from manufacturer datasheets) and the boundary conditions of convective heat transfer with the environment, specified on the outer surfaces of the SINS. As in the case of the test task, the convective boundary conditions inside the casing of the inertial measurement unit were not specified, but were generated and iteratively recalculated by the script during the solution process.

The temperature profile inside the casing was calculated for the transient thermal process lasting 6000 s.

An experimental study of the SINS thermal state was carried out by means of thermocouples mounted on the surfaces of the accelerometer block parts. Temperature measurements were performed after turning on the device in the self-heating regime.

a)

b)

Calculation using the developed technique
Calculation using CFD

Fig. 5. Comparison of calculation results: a) in cavity 1; b) in cavity 2

Since for the purposes of this work the thermal state of the accelerometers and the bracket on which they are mounted is of most interest, it is sufficient to limit ourselves to considering the results of four measurement points: T1, T2, T3 - on the accelerometers and T4 - on the bracket (see Fig. 6).

Fig. 6. Thermocouples mounting on the accelerometer bracket

The thermocouples are fixed at the measurement points using heat-conducting double-sided tape. The values of the measured temperatures were displayed and recorded in the memory of the thermometers. All channels have an error of ± 1 °C.

Graphs comparing the results of calculations and experimental measurements at the mounting points of thermocouples for T1 - T4 are presented in Fig.7.

As can be seen from the graphs, the calculation results are in satisfactory agreement with the experimental measurements and qualitatively describe the nature of the considered transient thermal processes.

For a quantitative analysis of the results, the maximum errors were calculated during the entire estimated transient process for the corresponding thermocouples T1-T4: T1 – 3.7%, T2 – 3.0%, T3 – 5.8%, T4 – 4.7%.

Fig. 7. Comparison of calculation results and experimental measurements: a) at point T1; b) at point T2; c) at point T3; d) at point T4

The temperature distribution in the bracket and accelerometers at the end of the calculation is shown in Fig. 8.

Fig. 8. Temperature field in the accelerometer bracket

IV. Conclusion

Thus, the verification and validation of the developed technique has been successfully completed, from which it can be concluded that it is possible to use this technique in calculation of the transient thermal state of SINS.

It is also advisable to note the significant time saving for calculations using the described technique in comparison with the technique [6] with comparable accuracy in calculation of the parts average temperatures. However, in order to increase the accuracy of local temperatures calculation by the current technique, it requires some modifications in terms of taking into account the locality of the heat transfer coefficient and its change over time in a transient thermal process.

References

[1] V.V. Matveev and V.Ya. Raspopov, Osnovy postroeniya besplatformennyh inertsial'nyh navigatsionnyh sistem (Fundamentals of construction of strapdown inertial navigation systems), St. Petersburg, Concern CSRI Elektropribor, 2009.

[2] Peshekhonov, V.G. Gyroscopic navigation systems: Current status and prospects. Gyroscopy Navig. 2, 111–118 (2011).
https://doi.org/10.1134/S2075108711030096.

[3] V.E. Dzhashitov, V.M. Pankratov, K.V. Koffer et al., "Hierarchical thermal models of a strapdown inertial navigation system based on fiber-optic gyroscopes", Giroskopiya i navigatsiya, no. 1, 2013, pp. 49-63.

[4] Gromov D.S. Thermal protection and thermal stabilization of a fiber-optic gyroscope as part of a modern strapdown inertial navigation system / Scientific and Technical Bulletin of Information Technologies, Mechanics and Optics, 2014, No. 2 (90), pp. 137-142.

[5] A.V. Frolov, S.V. Smirnov, and E.A. Popov, "Study of heat influence on stability of SINS accelerometer set carrier system axes", 27th St. Petersburg International Conference on Integrated Navigation Systems, 2020.

[6] Frolov A.V., Mikhailov Yu.V., Smirnov S.V. Development and validation of a methodology for modeling the thermal and deformed state of parts of a strapdown inertial navigation system. Bulletin of MSTU im. N.E. Bauman. Ser. Instrumentation, 2022, No. 1 (138), p. 32–48.

[7] Menter F. CFD Best Practice Guidelines for CFD Code Validation for Reactor Safety Applications // Evaluation of Computational Fluid Dynamic Methods for Reactor Safety Analysis (ECORA), European Commission, 5th EURATOM framework programme. - Berlin: GRS, 2004. – 53 p.

[8] Sosnovwski M., Krzywanski J, Grabowska K., Gnatowska R. Polyhedral meshing in numerical analysis of conjugate heat transfer. EPJ Web Conferences 180, 02096, 2018.

Synthesis of Calibration Programs for a Triad of Accelerometers

S.V. Smirnov
AO «CNIIAG»
Moscow, Russia
1410@cniiag.ru

Yu.G. Egorov
AO «CNIIAG»
Moscow, Russia
1412@cniiag.ru

G.Yu. Kiryachenko
AO «CNIIAG»
Moscow, Russia
1412@cniiag.ru

G.S. Taranenko
AO «CNIIAG»
Moscow, Russia
grigoriy1702@list.ru

Abstract — The work synthesizes and evaluates the accuracy characteristics of calibration programs for a triad of accelerometers using orthogonal series and a genetic algorithm.

Keywords — calibration, a triad of accelerometers, genetic algorithm, instrumental errors

I. INTRODUCTION

The task of estimating the instrumental errors of a triad of accelerometers (TA) during the calibration process in the classical formulation involves obtaining estimates of: biases, relative errors of scale factors, direction cosines of the accelerometer input axes. There are a total of 15 parameters that are assumed constant during calibration. In the TA calibration algorithms, normalization conditions are used [1] (the sum of the squares of the direction cosines of the accelerometer input axes is equal to one).

To solve the problems of synthesizing programs for calibrating an orthogonal TA, a rational mathematical model of TA measurements was constructed. The mathematical model includes three linear estimation problems with identical observation matrices and independent state vectors of dimension 4 in each problem. The use of normalization conditions allows us to subsequently solve the problem of calibrating the orthogonal TA using the 12 obtained estimates with the calculation of random estimates of all 15 instrumental errors.

In the process of synthesizing TA calibration programs based on orthogonal series, restrictions are imposed on the number of measuring positions and kinematic parameters that specify the orientation of the TA, but using a genetic algorithm there are no such restrictions.

II. MATHEMATICAL MODEL OF MEASUREMENTS OF A TRIAD OF ACCELEROMETERS

The measurement equation for a triad of accelerometers during the calibration process on a test bench can be represented in the following vector-matrix form:

$$\bar{J}_a = \bar{\Delta} + g E_{\delta k} \cdot C_a \cdot \bar{\eta} + \bar{\omega}_a; \qquad (1)$$

where $\bar{J}_a = \begin{bmatrix} J_{aX} & J_{aY} & J_{aZ} \end{bmatrix}^T$ – measurements of the i-th accelerometer (i = X, Y, Z);

$\bar{\Delta} = \begin{bmatrix} \Delta_X & \Delta_Y & \Delta_Z \end{bmatrix}^T$ – bias of the i-th accelerometer (i = X, Y, Z);

g – gravitational acceleration at the test site;

$$E_{\delta k} = \begin{bmatrix} 1+\delta_{kX} & 0 & 0 \\ 0 & 1+\delta_{kY} & 0 \\ 0 & 0 & 1+\delta_{kZ} \end{bmatrix}$$ – matrix specified by the relative errors of the TA scale factors;

δ_{ki} – relative error of the scale factor of the i-th accelerometer (i = X, Y, Z);

$$C_a = \begin{bmatrix} a_X & b_X & c_X \\ a_Y & b_Y & c_Y \\ a_Z & b_Z & c_Z \end{bmatrix}$$ – transition matrix from the instrument trihedron associated with the faceplate of the test bench to the trihedron coinciding with the input axes of the TA;

a_i, b_i, c_i – direction cosines of the input axis of the i-th accelerometer (i = X, Y, Z);

$\bar{\eta} = \begin{bmatrix} \eta_X & \eta_Y & \eta_Z \end{bmatrix}^T$ – a vector consisting of vertical direction cosines in the axes of the instrument trihedron, specified by the test bench, for example, as follows:

$\eta_X = \sin\alpha_3 \cos\alpha_2$, $\eta_Y = \cos\alpha_2 \cos\alpha_3$, $\eta_Z = -\sin\alpha_2$

(α_2, α_3 – angles of rotation along the axes of the test bench);

$\bar{\omega}_a = \begin{bmatrix} \omega_{aX} & \omega_{aY} & \omega_{aZ} \end{bmatrix}$ – vector of TA measurement errors, where ω_{ai} – measurement error of the i-th accelerometer (i = X, Y, Z);

To build a rational mathematical model, we introduce the following vectors:

$$\bar{X} = \begin{bmatrix} \Delta X & X_1 & X_2 & X_3 \end{bmatrix}^T;$$
$$\bar{Y} = \begin{bmatrix} \Delta Y & Y_1 & Y_2 & Y_3 \end{bmatrix}^T;$$
$$\bar{Z} = \begin{bmatrix} \Delta Z & Z_1 & Z_2 & Z_3 \end{bmatrix}^T;$$
$$X_1 = (1+\delta k_x) a_X g;$$
$$X_2 = (1+\delta k_x) b_X g;$$
$$X_3 = (1+\delta k_x) c_X g;$$
$$Y_1 = (1+\delta k_Y) a_Y g;$$
$$Y_2 = (1+\delta k_Y) b_Y g;$$
$$Y_3 = (1+\delta k_Y) c_Y g;$$
$$Z_1 = (1+\delta k_Z) a_Z g;$$
$$Z_2 = (1+\delta k_Z) b_Z g;$$
$$Z_3 = (1+\delta k_Z) c_Z g; \qquad (2)$$

where $\overline{X}, \overline{Y}, \overline{Z}$ – vectors formed by the calibrated instrumental errors of the accelerometers X, Y and Z, respectively.

The following relations, defining the normalization condition, are valid:

$$\sum_{i=1}^{3} X_i^2 = \left(1 + \delta k_X\right)^2 g^2;$$

$$\sum_{i=1}^{3} Y_i^2 = \left(1 + \delta k_Y\right)^2 g^2; \quad (3)$$

$$\sum_{i=1}^{3} Z_i^2 = \left(1 + \delta k_Z\right)^2 g^2.$$

As a result, we have a rational mathematical model of TA measurements during the calibration process:

$$J_{aX} = \begin{bmatrix} 1 & \overline{\eta}^T \end{bmatrix} \overline{X} + \overline{\omega}_{aX};$$

$$J_{aY} = \begin{bmatrix} 1 & \overline{\eta}^T \end{bmatrix} \overline{Y} + \overline{\omega}_{aY}; \quad (4)$$

$$J_{aZ} = \begin{bmatrix} 1 & \overline{\eta}^T \end{bmatrix} \overline{Z} + \overline{\omega}_{aZ}.$$

The mathematical model (4) and relationships (3) allow for the synthesis of programs and algorithms for calibrating TA using orthogonal series and a genetic algorithm.

III. SYNTHESIS OF PROGRAMS FOR CALIBRATION OF AN ORTHOGONAL TRIAD OF ACCELEROMETERS

The task of synthesizing TA calibration programs involves determining the number of measuring positions (MP) in the calibration program - N and the values of kinematic parameters (stand angles α_2, α_3), specifying the orientation of the TA relative to the vertical in each MP.

The measurement matrices H in the problems of estimating intermediate vectors $\overline{X}, \overline{Y}, \overline{Z}$ with N MP in the TA calibration program will be the same:

$$H = \begin{bmatrix} 1 & \overline{\eta}^T(1) \\ & \dots \\ 1 & \overline{\eta}^T(i) \\ & \dots \\ 1 & \overline{\eta}^T(N) \end{bmatrix}; \quad (5)$$

As a result, optimal estimates of intermediate vectors $\overline{X}, \overline{Y}, \overline{Z}$ are found using the following algorithms with the same operators A:

$$\hat{\overline{X}} = A \cdot \overline{J}_{aX}, \quad \hat{\overline{Y}} = A \cdot \overline{J}_{aY}, \quad \hat{\overline{Z}} = A \cdot \overline{J}_{aZ};$$

$$A = \left(H^T H\right)^{-1} H^T; \quad (6)$$

where $\overline{J}_{aj} = \begin{bmatrix} J_{aj}(1) & \dots & J_{aj}(i) & \dots & J_{aj}(N) \end{bmatrix}^T$ – vector of measurements of the j-th accelerometer in the calibration program from N MP (j = X, Y, Z);

$J_{aj}(j)$ – measurement of the j-th accelerometer (j = X, Y, Z) in the i-th MP of the calibration program (i = 1..N).

The covariance matrices P of vector $\overline{X}, \overline{Y}, \overline{Z}$ estimation errors $\Delta\overline{X}, \Delta\overline{Y}, \Delta\overline{Z}$ will be the same:

$$P = \sigma_n^2 \left(H^T H\right)^{-1}; \quad (7)$$

where σ_n^2 – vector $\overline{X}, \overline{Y}, \overline{Z}$ measurement error variance.

The synthesis of TA calibration programs using orthogonal series was carried out for the case when the MPs are set by rotating alternately around all axes of the instrument trihedron. Each of the axes about which the rotation is carried out is nominally in the horizontal plane. The rotation angles are measured from the vertical position of each of the axes of the instrument trihedron. In this case, the measurement matrix H for estimating vectors $\overline{X}, \overline{Y}, \overline{Z}$ has the following form:

$$H = \begin{bmatrix} \overline{e}_1 & \overline{e}_s & \overline{e}_c & \overline{e}_0 \\ \overline{e}_1 & \overline{e}_c & \overline{e}_0 & \overline{e}_s \\ \overline{e}_1 & \overline{e}_0 & \overline{e}_s & \overline{e}_c \end{bmatrix};$$

$$\overline{e}_0 = \begin{bmatrix} 1 \dots 1 \dots 1 \end{bmatrix}^T; \quad \overline{e}_0 = \begin{bmatrix} 0 \dots 0 \dots 0 \end{bmatrix}^T;$$

$$\overline{e}_s = \begin{bmatrix} \sin(\alpha_1) \dots \sin(\alpha_i) \dots \sin(\alpha_n) \end{bmatrix}^T; \quad (8)$$

$$\overline{e}_c = \begin{bmatrix} \cos(\alpha_1) \dots \cos(\alpha_i) \dots \cos(\alpha_n) \end{bmatrix}^T;$$

$$\overline{e}_0, \overline{e}_1, \overline{e}_s, \overline{e}_c \in R^n;$$

where α_i – values of the rotation angles of the TA in the MP by rotation alternately around the axes of the instrument trihedron (i = 1..n);

n – number of MPs specified by rotation around each axis of the instrument trihedron;

The total number of MPs in the calibration program N in this case is a multiple of three:

$$N = 3n. \quad (9)$$

The rows of the matrix H represent the elements of the orthogonal series 1, $\sin\alpha$, $\cos\alpha$. Maximum suppression of the influence of measurement noise on the estimation of the coefficients of the orthogonal series is achieved by dividing the full revolution into equal intervals [2]. In this case, the covariance matrix of estimation errors will be diagonal.

Thus, the entire set of TA calibration programs synthesized on the basis of orthogonal series is defined as follows:

$$\alpha_i = \frac{2\pi}{n}(i-1);$$

$$\left(n = 2,3,4,5\dots\right), \left(N = 6,9,12\right) \quad (10)$$

Covariance matrices P of estimation errors have the form:

$$P = \sigma_n^2 \cdot \begin{bmatrix} \dfrac{1}{N} & 0 & 0 & 0 \\ 0 & \dfrac{3}{N} & 0 & 0 \\ 0 & 0 & \dfrac{3}{N} & 0 \\ 0 & 0 & 0 & \dfrac{3}{N} \end{bmatrix}; \quad (11)$$

where σ_n^2 – variance of vector $\bar{X}, \bar{Y}, \bar{Z}$ measurement errors.

From (4) it follows that the number of MPs in TA calibration programs can be four or more. To synthesize a complete set of calibration programs with a number of measuring positions N = 4, 5, 6, 7, etc. a genetic algorithm was applied [3]. Calibration programs are synthesized according to the criteria of the minimum trace of the covariance matrix and its determinant, which corresponds to the maximum suppression of the influence of measurement noise on the calibration results. The results of the synthesis of TA calibration programs for N = 4, 6 and 12 are given in Table 1.

TABLE I.   RESULTS OF SYNTHESIS OF TA CALIBRATION PROGRAMS USING A GENETIC ALGORITHM

№ MP	N = 4		N = 6		N = 12	
	α_2, °	α_3, °	α_2, °	α_3, °	α_2, °	α_3, °
1	4,85	185,56	28,64	64,79	156,46	284,51
2	32,26	302,25	204,17	168,98	185,99	295,60
3	159,96	252,79	287,17	245,99	256,29	467,09
4	253,97	521,92	188,44	667,28	34,32	-2,28
5			50,92	292,53	-106,47	283,95
6			165,16	35,02	6,87	247,73
7					23,20	-10,50
8					-141,80	36,99
9					348,09	52,44
10					52,89	190,21
11					356,67	-47,43
12					139,94	13,20

Covariance matrices of estimation errors $\Delta\bar{X}, \Delta\bar{Y}, \Delta\bar{Z}$ for synthesized calibration programs with N multiples of three and with a minimum number of MPs N = 4 correspond to the expression (11), that specifies the corresponding covariance matrices of calibration programs synthesized on the basis of orthogonal series.

Thus, calibration programs synthesized using a genetic algorithm are not inferior in accuracy to calibration programs synthesized based on orthogonal series.

IV. CONCLUSION

Optimal TA calibration programs have been synthesized based on orthogonal series, consisting of a multiple of three number of MPs (N = 6, 9, 12, etc.). with a diagonal error matrix for estimating calibrated parameters.

The use of a genetic algorithm has made it possible to synthesize a complete set of TA calibration programs with any number of measuring positions (N = 4, 5, 6, 7... etc.), which are not inferior in accuracy to TA calibration programs synthesized on the basis of orthogonal series.

The conducted studies have shown the high efficiency of both methods for synthesizing optimal calibration programs for accelerometer units.

REFERENCES

[1] Taranovskiy D.O. Bench Calibration of a Pendulum Float Accelerometer Unit of a Ship's Inertial Navigation System // Gyroscopy and Navigation. 2008. №4 (63). pp. 56 – 65.

[2] Egorov Yu. G., Dzuyev A.A., Popov E.A. Synthesis of Calibration Programs for an Accelerometer Unit of Strapdown INS: An Invariant Approach // Gyroscopy and Navigation. 2019. Vol. 27 №1 (104), pp. 61 – 71.

[3] Egorov Yu. G., Kiryachenko G. Yu., Popov E.A. Numerical Methods of Suboptimal Programs Synthesis for Accelerometer Unit // Gyroscopy and Navigation. 2021. Vol. 29 №2 (113) pp. 47 – 58.

Equations of Motion for a Protection System with a Nonlinear Dynamic Vibration Damper for a Strapdown Inertial Navigation System

Svetlana Topilskaya

Department of Design Inertial Systems

Mars Moscow Experimental Design Bureau

Moscow, Russian Federation

office@mokb-mars.ru

Abstract – **Based on the performed analysis of the advantages and disadvantages of known vibration protection methods, it was determined that steel elastic elements and nonlinear dynamic vibration dampers are the most suitable damping elements for inertial devices and systems. The paper presents a model of the equations of the damping system for a block of sensitive elements of a gyroscopic device, which allows determining the parameters of elastic and damping elements when designing gyroscopic devices, as well as modelling the mechanical motion of a damping object during operational impacts. The description of this motion can be used in a mathematical model of the accuracy parameters of an inertial device.**

Keywords: **damping system, mechanical effects, inertial devices, elastic elements, dynamic vibration damper.**

I. INTRODUCTION

Gyroscopes are the basis for the construction of orientation and navigation systems in aerospace industry. When creating high-precision inertial instruments that are a part of aircraft and spacecraft control systems, the vibration strength and vibration resistance of sensitive elements throughout the entire operating time (during the launch of a launch vehicle, orbital flight, etc.) is of great importance [1], [6].

The purpose of this work is to determine the most suitable option for the damping system of a gyroscopic device and create a calculation model of the proposed damping system, allowing for practical engineering calculations.

II. VIBRATION PROTECTION SYSTEM FOR AN INERTIAL DEVICE

The analysis of the advantages and disadvantages of known vibration protection methods has shown that widely used rubber dampers, which have a number of advantages (simple design, easy installation, manufacturability, high internal damping, a wide range of features and a wide range of materials), are not always acceptable due to their inherent disadvantages: rubber material features depend on temperature and changes in environment composition and background radiation, and degradation of material over time. Therefore, it is preferable to use special alloys with stable characteristics under inertial device operation conditions as a material for elastic damping elements [5]. However, the elastic properties of steel elastic elements determine the high quality of the damping system at resonance, which leads to unacceptable vibration effects for sensitive elements.

Known methods for suppressing resonant vibration ranges have disadvantages associated with an increase in vibration ranges in a frequency range remote from the resonant frequency of the damping system, i.e. in the high frequency range (critical frequency zone of gyroscopes) [3], [4]. As a result of the analysis of the advantages and disadvantages of known vibration protection methods, it was determined that one of the suitable vibration protection methods for inertial devices in the aerospace industry is the combined use of steel elastic elements and dynamic vibration dampers (DVD) with nonlinear elasticity and damping characteristics [5].

Fig. 1. Inertial instrument damping system.

Fig. 2. An elastic element and a nonlinear dynamic vibration damper, where K_1 and k are the elasticity coefficients; D_1 and d are damping coefficients; h is a gap; m_2 is the mass of blocks; t, a, L are spring parameters.

Figure 1 shows the proposed damping system with a nonlinear DVD, figure 2 shows its elements.

The peculiarity of the operation of the proposed damping system includes the manifestation of the damping properties of the damping system only in the presence of critical vibration ranges, which make it possible to include inertial, elastic, damping force into the operation of the damping system only at the natural damping system frequency [2]. Thus, the DVD nonlinear characteristics eliminate the influence of the damping device on the level of suppression of vibrations of the steel elastic element in the dangerous frequency zone of the sensitive element.

III. THE MODEL OF THE EQUATIONS OF MOTION OF THE DAMPING SYSTEM

Nonlinear equations of motion of the proposed vibration protection method elements are represented as

$$\begin{cases} M_1\ddot{x}_1 + D_1(\dot{x}_1 - \dot{x}_0) + K_1(x_1 - x_0) + \\ + \begin{cases} 0, \ for \ (x_2 - x_1) < \dfrac{h}{2} \\ d(\dot{x}_1 - \dot{x}_2) + k(x_1 - x_2), \ for \ (x_2 - x_1) \geqslant \dfrac{h}{2} \end{cases} = 0 \\ m_2\ddot{x}_2 + \\ + \begin{cases} 0, \ for \ (x_2 - x_1) < \dfrac{h}{2} \\ d(\dot{x}_2 - \dot{x}_1) + k(x_2 - x_1), \ for \ (x_2 - x_1) \geqslant \dfrac{h}{2} \end{cases} = 0 \end{cases} \quad (1)$$

Figure 3 shows the kinematic diagram of the damping system as described by the equations of motion (1), where M_1 and m_2 are the masses of the damping object and the DVD block; K_1 and k are coefficients of damper and DVD elasticity; D_1 and d are damping coefficients of the damper and DVD; h is a gap; x_0, x_1, x_2 are the movements of damper mounting points in the inertial device, the base of the damping object and the DVD blocks, respectively, in relation to the fixed coordinate system.

For the convenience of applied engineering calculations, a model of the equations of motion has been derived making it possible to model the dynamic behavior of the system for various parameters of the damping system.

Fig. 3.   Kinematic diagram of the damping system

Figure 4 presents a model of nonlinear equations of motion. In the nonlinear model, the presence of a gap h of the dynamic vibration damper is implemented using the Switch block. The k and d elastic and damping characteristics of the damper exert their influence on the vibration and dynamic characteristics of the damping system only if the condition $\Delta x_2 \geqslant h/2$ is met. In other cases, in the absence of critical vibration ranges, the elastic and damping characteristics of the damper do not affect the vibration amplitudes of the damping object. This model allows simulating the dynamic behavior of a damping object for various parameters M_1, K_1, D_1, m_2, K_2, D_2, k, d, h.

Fig. 4. Model of the equations of motion of the damping system

IV. SIMULATION DATA

Figure 5 presents the results of modelling the movements x_1 of the damping object (M_1) and x_2 of the DVD block (m_2) at resonant vibration ranges. The diagram shows that when the condition $\Delta x_2 \geq h/2$ is met, the DVD block m_2 starts executing vibrations in the gap h. These vibrations reduce the vibration range of the vibration protection object. Also, as a result of modelling, it was determined that in the frequency range remote from the resonance of the damping system, the elastic k and damping d characteristics of the DVD do not affect the vibration ranges of the damping object (Figure 6).

Fig. 5. Results of modeling the movements of the damping system at $\Delta x_2 \geq h/2$.

Fig. 6. Results of modeling the movements of the damping system at $\Delta x_2 < h/2$.

That is, due to the gap h in the DVD design, it is possible to create a positive influence of elastic and damping characteristics of the DVD on the vibration ranges of the damping object.

V. CONCLUSION

Based on the analysis of known vibration protection methods in various engineering fields, a damping system for an inertial device, consisting of a damping object, elastic elements, and nonlinear DVDs has been proposed. The main result of the work is the derivation of a model of the equations of motion for the proposed damping system, which allows for applied engineering calculations when designing inertial devices for aerospace industry.

REFERENCES

[1] S.P. Timoshenko, Vibration Problems in Engineering / S.P. Timoshenko; revised by D.H. Yangomper; translated from English by Ya.G. Panovko, from the 3rd American ed. – 2nd ed., stereotyped. Moscow: URSS, 2006, pp. 439.

[2] Ya.G. Panovko. Attached Dynamic Systems as Vibration Dampers / Ya.G. Panovko // Strength, Stability, Vibrations, Handbook. M.: Mechanical Engineering, 1968, pp. 331-346.

[3] Den-Hartog Jacob. Mechanical Vibrations/Den-Hartog Jacob//M.: Fizmatgiz, 1960, p. 574.

[4] S. Crandall. Random Vibrations/S. Crandall//M.: Mir, 1967, p. 356.

[5] A. Nashif. Vibration Damping/A. Nashif., D. Jones., J. Henderson; translated from English // M.: Mir, 1988, p. 448.

[6] V.S. Ilyinsky. Protection of Devices from Dynamic Influence/V. S. Ilyinsky // M.: Energy, 1970, p. 320.

Refinement of Hemispherical Resonator Gyroscope Actuation Algorithm

L.E. Kochegizova
JSC «Inertial technologies of "Technocomplex"»
Ramenskoye, Russia
larisakochegizova@gmail.com

S.V. Fetisov
JSC «Inertial technologies of "Technocomplex"»
Ramenskoye, Russia
SFetisov@inertech.ru

A.S. Malyugin
JSC «Inertial technologies of "Technocomplex"»
Ramenskoye, Russia
AMalyugin@inertech.ru

D.I. Martynenko
JSC «Inertial technologies of "Technocomplex"»
Ramenskoye, Russia
DMartynenko@inertech.ru

Abstract—**The paper presents the research of impact of different external influence methods on actuation of hemispherical resonator gyroscope using the digital model of this sensor. The conducted research allows formation of the requirements and approach to refinement of the actuation algorithm. The goal was to decrease functional availability time. The consistent patterns of oscillation amplitude rise were determined based on the number of modeling iterations. The analysis of the existing actuation algorithm was performed. The algorithm enabling the reduction of functional availability time was developed due to the results of the studies. The results were estimated using the digital model of a hemispherical resonator gyroscope.**

Keywords—HRG, hemispherical resonator gyroscope, actuation, algorithm, model

I. INTRODUCTION

Nowadays the central themes in developing navigational sensors and systems are miniaturization, improvement of accuracy and reduction of indication setting time. One of actively advancing prospects is engineering of hemispherical resonator gyroscopes (HRG) [1]. A key requirement for this type of gyroscope is providing high dynamic performance. One of the essential dynamic characteristic is functional availability time. It is defined by duration required to achieve the operating magnitude of hemispherical resonator oscillations.

II. PROBLEM STATEMENT

Optimization of actuation in hemispherical resonator gyroscope is a problem of current interest. Requirements for indication setting time applied to navigation systems using HRG as sensors are becoming increasingly demanding. Thus, it is necessary to ensure minimal availability time for each system element.

The main focus of current research is dedicated to accuracy characteristics. At the same time the subjects of actuation and availability are less common. There are two existing methods of actuating hemispherical resonator gyroscope [2]. The first one is called parametric excitation of thin-walled hemispherical shell and the second one is called position excitation of thin-walled hemispherical shell. HRG with position excitation operates in force-to-rebalance (FTB) mode as a rate sensor. In this case the stationary wave is held in specified circumferential position by control system during operational period. HRG with parametric excitation [3] operates in whole-angle (WA) mode as an integrating gyroscope. In this case the stationary wave is actuated in unspecified circumferential position and pivots proportionally to angular rate of the platform. The platform attitude information is measured by means of capacitive pick-offs. The useful signal is generated by fluctuations of the working gap.

Taking into consideration that HRG operates in whole-angle mode, actuation optimization comes down to minimizing resonant frequency search time and realizing the most effective influence on sensing element.

The purpose of this paper is to study impact of external influence methods on excitation and to refine actuation algorithm.

III. PROBLEM SOLVING

The research was conducted using a digital model of the hemispherical resonator gyroscope [4]. The digital model was verified by means of comparing the stationary wave behavior and technological parameters of the resonator to the theoretical ones [5]. The model consists of several mathematical modules that make it possible to simulate the operating sensor, considering the following features:

- Technological parameters of the resonator, including resonance frequency, quality factor, frequency splitting, quality factor variation.

- External angular rate.

- Control action.

Modeling parameters common for every study are shown in Table I.

TABLE I. MODELING PARAMETERS OF ACTUATION

Parameter	Value
Resonance frequency f_{res}, Hz	5000.0
Time of impact on resonator, s	5.0
Technological inaccuracies of resonator	Absent
Implemantation of external impact	By means of both control channnels
Amplitude of actuating impact, V	1.4
Operating amplitude, V	1.0

The sensing element model is assumed to be perfect because if there are technological inaccuracies then the energy is expanded not only to excitation of desired

oscillations but also to actuation of quadrature oscillations. This effect leads to complications when estimating the efficiency of certain influence method. The remaining parameters were assumed close to actual parameters of the sensor. The impact time was chosen to minimize time invested.

A. Investigation of the influence of the resonance frequency determination error

The actuation efficiency depends on the influence error of matching the excitation frequency to the resonance peak. The investigation of impact on actuation with error of resonance frequency determination was conducted using parameters from Table I. The quality factor of the resonator was defined at values of 10 million and 1 million. These values were chosen based on the experience of HRG development. The most common value of quality factor in real sensors ranges between 1 and 10 million.

Modeling results are presented in Table II.

TABLE II. MODELING RESULTS OF SENSOR ACTUATION WITH RESONANCE FREQUENCY DETERMINATION ERROR

Quality factor	Frequency of impact, Hz	Frequency error relative to resonance frequency, Hz	Amplitude of oscillations after impact, V
10 mil.	4999.9	0.1	1.16377
	4999.0	1.0	0.02196
	4990.0	10.0	0.00123
1 mil.	4999.9	0.1	1.12360
	4999.0	1.0	0.02155
	4990.0	10.0	0.00120

As shown by the data above the error of influence frequency of 0.1 Hz relative to resonance frequency of the sensing element ensures the actuation efficiency at 116% of operating amplitude. The error of influence frequency of 1.0 Hz relative to resonance frequency of the sensing element provides only 2.2% efficiency of actuation. Thus, it is necessary to determine resonance frequency with accuracy better than 1 Hz.

B. Investigation of the sweep sinusoidal influence on actuation

Widely known methods of excitation propose consequential sweep of frequencies through given range of values. When hitting resonance peak, the kinetic energy of oscillations starts accumulating due to the external force. The study of actuation efficiency dependent on the sweep range was carried out using the conditions shown in Table I. The frequency of sinusoidal impact varies linearly continually and symmetrically relative to the resonance frequency. Ranges of modeling frequencies were $f_{res} \pm 2$ Hz, $f_{res} \pm 5$ Hz, $f_{res} \pm 7$ Hz. The linear dependency of amplitude after influence ended on range of the sweep frequency was elucidated. Figure 1 shows the acquired dependency for resonators with quality factors of 10 million and 1 million.

In addition, the excitation of the sensing element subjected to sinusoidal influence with frequency asymmetrically distributed relative to resonance was studied. The sweep range width was assumed equal to 4 Hz because this interval was the most efficient due to results of the

previous study. The ratio between subdivisions of the sweep range is 3 to 1. The quality factor of the resonator is 10 million. Comparative graph in Fig. 2 shows the oscillations amplitude change when exposed to both types of excitation influence.

Fig. 1   Dependency of amplitude after influence ended on the width of sweep frequency range

Fig. 2   Comparison of amplitude change under sinusoidal influence with symmetrical and non-symmetrical sweep range

As the figure shows the efficiency of actuation decreases but not significantly, approximately by 0.02 V. Thus, it can be concluded that even if there is initially an error of resonance frequency determination, the performance of excitation decreases inconsequentially.

C. Investigation of the square wave influence with variable frequency on actuation

The actuating influence on resonator might be presented as square wave [6]. Moreover, square wave signal might be interpreted as alternating impact at certain frequency. Modeling was conducted under conditions from Table I. The frequency of square wave varies linearly and discretely, symmetrically relative to the resonance frequency. The value of discrete step is 1 Hz. The range of frequency variability was chosen $f_{res} \pm 2$ Hz because the wider range is less effective based on the results of investigation B.

The Table III shows comparison between square wave influence with variable frequency and sweep sinusoidal influence on the resonator.

TABLE III. MODELING RESULTS OF SQUARE WAVE INFLUENCE AND SINUSOIDAL INFLUENCE

Quality factor.	Amplitude under square wave influence, V	Amplitude under sinusoidal influence, V
10 mil.	0.4329	0.39592
1 mil.	0.4125	0.38189

The results indicate that the actuation efficiency of square wave exceeds 40% relative to the operating amplitude. It is accomplished by longer-term effect from square wave at the resonance frequency. To realize the significance of discrete frequency succession, the modeling was conducted when external influence frequency step hadn't hit the resonance frequency. Modeling parameters were the same as the previous study. The quality factor of the resonator was set to 10 million, frequency range was from 4997.5 Hz to 5001.5 Hz, the value of discrete step was the same. The graphs of amplitude change under the influence with and without resonance are shown in Fig. 3.

Fig. 3 Comparison of amplitude change under the influence of square wave with resonance frequency and without

The Fig. 3 shows the actuation efficiency of square wave without resonance decreases almost 2 times comparing with the case of square wave with resonance. Considering these results, it can be concluded that if there is an error of frequency range determination and choice of excessively large discrete step of frequencies, then the square wave influence is less effective than the sweep sinusoidal influence.

IV. DEVELOPMENT OF ACTUATION ALGORITHM

The most preferable way to actuate the gyroscope is to apply the external influence at its resonance frequency. However accurate determination of resonance frequency under service conditions is not a trivial problem because resonance frequency depends on temperature. The goal of this part of work is to develop an algorithm of actuation that ensures achievement of operating amplitude of sensing element oscillations equal to 1 V. The existing actuation algorithm of hemispherical resonator gyroscope employs the parametric excitation method. The algorithm consists of following stages:

1. Wide-range search of resonance frequency.
2. Refinement of the determined resonance frequency.

3. Sinusoidal influence at the refined resonance frequency.
4. Engagement of control loop for maintaining the oscillation amplitude once reached 60% of operating amplitude value (0.6 V).

Figure 4 shows the oscillation amplitude change when using the existing actuation algorithm.

Fig. 4 The oscillation amplitude change when using the existing actuation algorithm

The settling amplitude time is totaled to more than 60 seconds. Analyzing the stages of algorithm above, one may conclude that the developing algorithm should comply with following requirements:

1. To embody the search of resonance frequency.
2. To ensure the oscillation amplitude value after the influence of no less than 80% of operating.
3. To restrict the duration of external influence to 5 seconds.
4. To guarantee the total settling amplitude time of no more than 60 seconds.

Guided by the above requirements, the following actuation algorithm for HRG was developed:

1. Sweep sinusoidal influence on resonator within the space of 2 seconds. Whole range of frequencies is swept in the course of 1 second.
2. The assessment of sensor's resonance frequency is made upon received signals in the span of 1 second. The estimated frequency gets fixated.
3. The sinusoidal influence on resonator at frequency determined in item 2 during 2 seconds.
4. Engagement of control loop for maintaining the operating oscillation amplitude.

Figure 5 shows the amplitude change when using developed algorithm.

Thus, the total time of actuating influence on resonator is 5 seconds. Since the estimation of the resonance frequency is performed during actuation it gives the opportunity to assess the resonance frequency very precisely. Afterwards that the oscillations are excited at determined frequency which enables to achieve 80% of operating amplitude during short period of time, the remaining 20% is obtained by control loop action. The total actuation time is less than 30 seconds.

Fig. 5   The amplitude change when using the developed algorithm

V.   Conclusion

The paper presents the research of impact of different external influence methods on actuation of hemispherical resonator gyroscope using the digital model of this sensor. The conducted research has made it possible to define the requirements to refinement of actuation algorithm and determine the type of external influence necessary for minimal actuation time. The analysis of existing actuation algorithm has been performed, the disadvantages of the algorithm have been discovered and compensated in the developed algorithm. The results fully comply with the requirements, ensued from the conducted research.

References

[1]   Lunin B.S. *Fiziko-khimicheskie osnovy razrabotki polusfericheskikh rezonatorov volnovykh tverdotel'nykh giroskopov* [Physical-chemical fundumentals of wave solid-state gyroscope development], Russia, Moscow: Izdatel'stvo MAI, 2005.

[2]   Z. Wei, G. Yi, Y. Han, Z. Qi, Z. Xu, "The synthesis model of fla-electrode hemispherical resonator gyro", Sensors, vol. 19, issue 7, 2019, p.1690.

[3]   W. Zhao, H. Yang, F. Lui, Y. Su, L. Song, "The energy compensation of th HRG based on double-frequency parametric excitation of the discrete electrode", Sensors, vol. 20, 2020.

[4]   Kochegizova L.E., Fetisov S.V, Malyugin A.S., Martynenko D.I., Vakhlyarskii D.S. "Model' tverdotel'nogo volnovogo giroskopa v srede SIMULINK" [Model of hemispherical resonator gyroscope in SIMULINK] *Navigatsiya i upravlenie dvizheniem materialy XXV yubileinoi konferentsii molodykh uchenykh s mezhdunarodnym uchastiem,* 2023. - pp. 88-90.

[5]   Lunin B.S., Matveev V.A., Basarab M.A. *Volnovoi tverdotel'nyi giroskop. Teoriya i tekhnologiya* [Wave solid-state gyroscope. Theory and technology]. Russia, Moscow: Radiotekhnika, 2014.

[6]   X. Xu, X. Lui, Y. Zhang, "A quick start method for MEMS disk resonant gyroscope", Sensors (Basel), 2021 no.23: 7986.

Noise Suppression in the Output Signal of a YAG:Cr^{4+} Solid-State Laser Gyroscope with Cavity Length Control And Mode-Locking

Yu.Yu. Broslavets
MIPT
JSC «Lasex»
Dolgoprudny, Russia
laseruu@mail.ru

E.A. Polukeev
MIPT
JSC «Lasex»
Dolgoprudny, Russia
epolukeev2105@yandex.ru

D.S. Redichkina
MIPT
Dolgoprudny, Russia
dariaredichkina@gmail.com

A.A. Fomichev
MIPT
JSC «Lasex»
Dolgoprudny, Russia
laser@mail.mipt.ru

V.G. Semenov
MIPT
JSC «Lasex»
Dolgoprudny, Russia
valerii.semenov.g@gmail.com

A.R. Pokrovskaya
MIPT
Dolgoprudny, Russia
pokrovskaya3113@gmail.com

Abstract — **The present study investigates the noise in the output beat signal of a femtosecond solid-state laser gyroscope during mode-locking, taking into account cavity configuration, ring laser parameters, cavity length control system, and frequency bias. Kerr mode-locking requires strong focusing and high lasing stability, narrowing the linewidth of individual modes and thereby improving accuracy. We have determined the cavity configuration ensuring lasing stability under external disturbances with the first derivatives of these dependencies being zeroed. Cavity length control system included an acetylene cell and its influence on the amplitude and phase noise in the output signal was considered. Frequency biasing of a solid-state laser gyroscope via an electro-optical phase modulator was also investigated.**

Keywords — *solid-state laser gyroscope, ring laser, mode-locking, lock-in.*

I. INTRODUCTION

Nowadays almost all laser gyroscopes are based on He-Ne ring lasers [1] due to their numerous positive qualities. Nevertheless, achievements in the field of creating solid-state laser systems [2, 9, 11, 13], designing multilayered coatings, manufacturing crystal mirrors with record high reflection and the emergence of new gain media [2, 4, 5, 11], including those with inhomogeneous broadening, stimulate the search for new physical solutions for creating the next generation laser gyroscope. One of those promising solutions is a solid-state laser gyroscope based on a femtosecond ring laser with a YAG:Cr^{4+} gain medium, capable of broadband femtosecond lasing with weak competition of counterpropagating waves [11]. Operating in a multi-frequency regime with the oscillations of the modes being synchronized enables significant noise suppression due to the longitudinal modes, carrying information about rotation, being in-phase. A number of works by THALES proves the achievable accuracies of a YAG:Nd^{3+} solid-state laser gyroscope to reach the corresponding parameters of He-Ne gyroscopes [9, 12]. A solid-state Ti^{3+}:Al$_2$O$_3$ femtosecond ring laser with Kerr mode-locking created at the University of New Mexico has shown to overcome the fundamental obstacles to the successful operation of solid-state laser gyroscope [2].

It should be noted that previously created solid-state laser gyroscopes had no cavity length control systems [2, 9, 12] and the frequency bias was most often created by rotating the gyroscope or by applying a magnetic field to the gain medium (Faraday effect), which limited their accuracy. Systems for stabilizing the cavity length and creating and subtracting frequency bias play a significant role in the operation of a laser gyroscope and largely determine its random walk. Therefore, our work is aimed at creating these systems based on new physical principles for a solid-state laser gyroscope and determining their effect on its noise and accuracy parameters.

The work reflects the results obtained during the creation of laser gyroscope based on a broadband gain medium of YAG:Cr^{4+}. System for stabilizing the cavity length and frequency bias system via a lithium niobate electro-optical phase modulator are also presented. The influence of the used frequency bias and cavity length control system on the noise and stability of the beat frequency of counterpropagating waves during Kerr mode-locking was investigated. To provide the required phase modulation during frequency biasing we optimized the optical scheme, placing lithium niobate crystal at the Brewster angle and finding the optimal location of the modulator in the cavity. Optimal parameters of the PID controller of an acetylene-based absorbing cell in frequency stabilization systems have also been determined.

Numerical modeling included calculating the influence of the frequency stabilization system and frequency bias on the stability of lasing and fluctuations of the beat frequency of counterpropagating waves during mode-locking. The main parameters influencing the stability of the beat signal were determined. It is shown that the presence of amplitude modulation caused by the Kerr mechanism and the coupling between waves of the same direction leads to a lock-in region between waves of the same direction, which causes their phasing and thereby mode-locking and the formation of ultrashort pulses. The optimal value of such coupling was found.

II. NUMERICAL MODELING OF LASING DYNAMICS IN MULTIFREQUENCY REGIME DURING MODE-LOCKING

It is known that synchronization of longitudinal modes [10, 11, 14] leads to a strong narrowing of the linewidth of each mode, which increases maximum sensitivity and accuracy when employed in a laser gyroscope. Additionally, the use of femtosecond pulses [2, 11] reduces the coupling of

counterpropagating waves and corresponding nonlinearity of the output frequency response of the laser gyroscope. For that purpose we performed a study on multifrequency lasing in a YAG:Cr^{4+} solid-state laser gyroscopes in various operating regimes. The broad gain of YAG:Cr^{4+} medium allows for a lasing of a great number of longitudinal modes, which, when synchronized, can produce femtosecond pulses.

When modelling the lasing dynamics in multifrequency regime, we employed the wave equation for the field in the cavity, neglecting the transverse structure [8]:

$$\nabla \vec{E} - grad(div\vec{E}) = \frac{4\pi\sigma}{c^2}\frac{\partial \vec{E}}{\partial t} + \frac{\varepsilon}{c^2}\frac{\partial^2}{\partial t^2}\left(\vec{E} + \frac{4\pi\vec{P}}{\varepsilon}\right), \quad (1)$$

where σ - conductivity, reflecting the presence of volumetric losses in the medium (during mode-locking losses are modulated at the intermode frequency), ε - dielectric constant of the medium. In this case, the field in a solid-state ring laser was written as a superposition of the fields of counterpropagating waves, each of which was represented as an expansion in the longitudinal modes of the resonator:

$$\vec{E}(z,t) =$$
$$= \text{Re}\sum_j \left(e_{j1}\tilde{E}_{j1}(t)\exp\left(i\left(\omega_{j1}t + k_{j1}z\right)\right) + e_{j2}\tilde{E}_{j2}(t)\exp\left(i\left(\omega_{j2}t - k_{j2}z\right)\right)\right),$$
$$(2)$$

$$\vec{P}(z,t) =$$
$$= \text{Re}\sum_j \left(e_{j1}\tilde{P}_{j1}(t)\exp\left(i\left(\omega_{j1}t + k_{j1}z\right)\right) + e_{j2}\tilde{P}_{j2}(t)\exp\left(i\left(\omega_{j2}t - k_{j2}z\right)\right)\right).$$
$$(3)$$

The complex amplitudes of counterpropagating waves can be written in terms of amplitude and phase:

$$\tilde{E}_{j1}(t) = E_{j1}\exp\left(i\varphi_{j1}\right), \quad \tilde{E}_{j2}(t) = E_{j2}\exp\left(i\varphi_{j2}\right). \quad (4)$$

Solving the system of equations (5) was carried out under the assumption that the frequencies of longitudinal modes during lasing are significantly greater than their intermode frequency interval. The interaction of waves on the inversion gratings for various longitudinal modes was not taken into account. Synchronization of longitudinal modes was due to the appearance of side frequencies at the frequencies of neighboring modes during modulation of losses at the intermode frequency. The coupling of modes of one direction of propagation which is the result of modulation of losses is given by $\tilde{m}l_{jm}$ coefficients, determined by the modulation depth. The side frequencies play the role of a driving force capturing the frequency of the neighboring mode and thereby providing a constant phase shift between the frequencies of the longitudinal modes of each lasing direction. To a first approximation, we can assume that the frequencies of counterpropagating waves of each longitudinal mode are split by the same amount (due to rotation of the gyroscope or nonreciprocity). The shortened equations for complex amplitudes [8] of counterpropagating waves and population inversion describing multimode lasing in a solid-state ring laser have the form:

$$\frac{d\tilde{E}_{11}}{dt} = -\frac{\omega_{11}}{2Q_{11}}\tilde{E}_{11} + i\frac{\Omega}{2}\tilde{E}_{11} + \frac{i}{2}\tilde{m}_{11}\tilde{E}_{11} + \frac{i}{2}\tilde{m}l_{11}\tilde{E}_{21} +$$
$$+\Delta\omega r\tilde{E}_{11} + \frac{\sigma_1 l}{2T}(1-i\delta)\left(N_0\tilde{E}_{11} + N_-\tilde{E}_{11}\right)$$

$$\frac{d\tilde{E}_{12}}{dt} = -\frac{\omega_{12}}{2Q_{12}}\tilde{E}_{12} - i\frac{\Omega}{2}\tilde{E}_{12} + \frac{i}{2}\tilde{m}_{12}\tilde{E}_{12} + \frac{i}{2}\tilde{m}l_{12}\tilde{E}_{22} +$$
$$+\Delta\omega r\tilde{E}_{12} + \frac{\sigma_1 l}{2T}(1-i\delta)\left(N_0\tilde{E}_{12} + N_+\tilde{E}_{12}\right)$$

$$T_1\frac{dN_{1+}}{dt} = -N_{1+}\left(1 + a_1\left(|E_{11}|^2 + |E_{12}|^2\right)\right) - N_0 a_1 E_{11}^* E_{12}$$

$$T_1\frac{dN_0}{dt} = N_{th}(1+\eta) -$$
$$-N_0\left(1 + a_1\left(|E_{11}|^2 + |E_{12}|^2\right) + a_2\left(|E_{21}|^2 + |E_{22}|^2\right) + \ldots + a_n\left(|E_{n1}|^2 + |E_{n2}|^2\right)\right)$$

$$\frac{d\tilde{E}_{21}}{dt} = -\frac{\omega_{21}}{2Q_{21}}\tilde{E}_{21} + i\frac{\Omega}{2}\tilde{E}_{21} + \frac{i}{2}\tilde{m}_{21}\tilde{E}_{21} + \frac{i}{2}\tilde{m}l_{21}\left(\tilde{E}_{11} + \tilde{E}_{31}\right) +$$
$$+\Delta\omega r\tilde{E}_{21} + \frac{\sigma_2 l}{2T}(1-i\delta)\left(N_0\tilde{E}_{21} + N_-\tilde{E}_{21}\right)$$

$$\frac{d\tilde{E}_{22}}{dt} = -\frac{\omega_{22}}{2Q_{22}}\tilde{E}_{22} - i\frac{\Omega}{2}\tilde{E}_{22} + \frac{i}{2}\tilde{m}_{22}\tilde{E}_{22} + \frac{i}{2}\tilde{m}l_{22}\left(\tilde{E}_{12} + \tilde{E}_{32}\right) +$$
$$+\Delta\omega r\tilde{E}_{22} + \frac{\sigma_2 l}{2T}(1-i\delta)\left(N_0\tilde{E}_{22} + N_+\tilde{E}_{22}\right)$$

$$T_1\frac{dN_{2+}}{dt} = -N_{2+}\left(1 + a_2\left(|E_{21}|^2 + |E_{22}|^2\right)\right) - N_0 a_2 E_{21}^* E_{22}$$

$$\ldots$$

$$\frac{d\tilde{E}_{n1}}{dt} = -\frac{\omega_{n1}}{2Q_{n1}}\tilde{E}_{n1} + i\frac{\Omega}{2}\tilde{E}_{n1} + \frac{i}{2}\tilde{m}_{n1}\tilde{E}_{n1} + \frac{i}{2}\tilde{m}l_{n1}\tilde{E}_{(n-1)1} +$$
$$+\Delta\omega r\tilde{E}_{n1} + \frac{\sigma_n l}{2T}(1-i\delta)\left(N_0\tilde{E}_{n1} + N_-\tilde{E}_{n1}\right)$$

$$\frac{d\tilde{E}_{n2}}{dt} = -\frac{\omega_{n2}}{2Q_{n2}}\tilde{E}_{n2} - i\frac{\Omega}{2}\tilde{E}_{n2} + \frac{i}{2}\tilde{m}_{n2}\tilde{E}_{n2} + \frac{i}{2}\tilde{m}l_{n2}\tilde{E}_{(n-1)2} +$$
$$+\Delta\omega r\tilde{E}_{n2} + \frac{\sigma_n l}{2T}(1-i\delta)\left(N_0\tilde{E}_{n2} + N_+\tilde{E}_{n2}\right)$$

$$T_1\frac{dN_{n+}}{dt} = -N_{n+}\left(1 + a_n\left(|E_{n1}|^2 + |E_{n2}|^2\right)\right) - N_0 a_n E_{n1}^* E_{n2}$$
$$(5)$$

Here j − is a number of a longitudinal mode, $\tilde{m}_{jm} = m_{jm}\exp(\pm i\Theta_{jm})$ - complex coupling coefficients of counterpropagating waves, $\tilde{m}l_{jm} = ml_{jm}\exp(\pm i\Psi_{jm})$ - complex coupling coefficients of longitudinal modes, T_1 - population inversion relaxation time, $\Delta\omega r$ − coefficient, defining random fluctuations, N_{th} − threshold population inversion, L − cavity length, l − crystal length, $\delta_j = (\omega_j - \omega_0)/\Delta\omega_g$ - relative detuning of the lasing frequency of mode j from the center of the gain line, $\sigma_j = \sigma/(1 + \delta_j^2)$ - transition cross section, different for different longitudinal modes, $a_j = (\sigma_j cT_1)/(4h\omega)$ - saturation parameter, $\Omega = \omega_{j1} - \omega_{j2}$ - difference in mode frequencies for counterpropagating waves, Q_j - resonator quality factor, $k_j = 2\pi n_j/L$ - wave number, n_j - number of wavelengths along the cavity length, $N_0 = \frac{1}{l}\int_0^l N dz$, $N_{j\pm} = \frac{1}{l}\int_0^l N\exp(\pm i2kz)dz$ - spatial harmonics of population inversion.

The frequency difference of counterpropagating waves for each longitudinal mode j can be obtained through the field phases changing with time:

$$\Delta\omega_j = \omega_{j1} - \omega_{j2} = \frac{d\varphi_{j1}(t)}{dt} - \frac{d\varphi_{j2}(t)}{dt}. \qquad (6)$$

During numerical modeling cases j=1..4 were considered (lasing from one to four longitudinal modes in each direction). Figures 1–3 show the obtained dependences of the frequency difference of a pair of counterpropagating waves on nonreciprocity Ω, similar in meaning to the output characteristic of a gas laser gyroscope.

Fig. 1. Output frequency response for a ring laser with two longitudinal modes. At large nonreciprocity a second lock-in zone is observed.

Fig. 2. Output frequency response with two longitudinal modes when they are in phase.

Fig. 3. Output frequency response with one longitudinal mode.

The calculation shows that the presence of synchronized longitudinal modes leads to a decrease in the lock-in zone (Figs. 2, 3). At the same time, the nonlinearity of the characteristic is increased near the lock-in zone (Fig. 1), and the second lock-in zone, which appears at large nonreciprocity values, is shifted to the region of their smaller values. A significant narrowing of the linewidth of longitudinal modes during mode-locking in broadband media with hundreds of thousands of longitudinal modes

[10, 11, 14] and a reduced lock-in zone promote the use of broadband media in laser gyroscopes.

Figure 4 shows the results of numerical simulation of the lasing of counterpropagating waves in the beat mode, the nonstationary population inversion grating that appears in the gain medium is presented. The beat frequency corresponds to the magnitude of frequency nonreciprocity. The traveling wave mode is actually a unidirectional lasing. The main reason for switching to this mode is the uniform broadening of the gain. In this case the inversion grid does not appear (Fig. 5). This operating regime poses a big problem for the use of solid-state ring lasers in gyroscopy since during it, measuring the rotation is impossible, but it can be useful for other purposes such as spectroscopy to create a source of monochromatic radiation. To suppress it one can introduce losses for waves proportional to their intensity.

During lock-in the frequencies of the counterpropagating waves are the same, which leads to the formation of a stationary inversion grating (Fig. 6).

Fig. 4. Population inversion grating in beat mode. The beat frequency is determined by frequency nonreciprocity.

Fig. 5. The inversion grid does not appear. The traveling wave regime is actually a unidirectional lasing.

Fig. 6. Population inversion grating during lock-in of counterpropagating waves.

III. MODELLING OF AN ASTIGMATIC RING CAVITY FOR KERR MODE-LOCKING

Significant suppression of noise in mode-locked laser spectrum requires a certain configuration of an astigmatic cavity. This ensures both strong focusing of light in the gain medium for the formation of a Kerr lens and sets first derivatives of the disturbance values of an astigmatic Gaussian beam to zero (displacement of mirrors, fluctuations of parameters of the gain medium). The laser cavity also ensures the maximum pulse shortening when passing through the formed Kerr lens and the diaphragm. The oblique incidence of a Gaussian beam onto the gain medium at the Brewster angle ensures low losses and lasing of linear polarization and leads to an astigmatic beam with its waists displaced along the optical axis. To ensure the indicated characteristics of the ring laser cavity we performed the modelling according to the well-known "ABCD" rule for complex parameters describing the beam in the plane of incidence and in the perpendicular plane.

The cavity was built according to a z-shaped ring design, providing strong focusing of laser radiation in the gain medium and a wide beam in the dispersion arm (Fig. 7). Spherical mirrors M3 and M4 provide focusing of radiation in the gain medium to enable Kerr mode-locking. An active element of 2 cm long with its edges cut at the Brewster angle is installed between the focusing mirrors, which leads to beam astigmatism. Volumetric Kerr and thermal lenses are formed in the gain medium, and to ensure the greatest change in the beam size due to the Kerr lens the medium is located at a certain distance from the middle between the focusing mirrors. This increases the depth of radiation modulation and thereby the stability of Kerr mode-locking, which reduces noise by narrowing the linewidth of individual modes. As mode-locking depends on the noise of amplitude modulation, it is important to choose the optimal location of the gain medium and the distance between the focusing mirrors to ensure the parameters of Kerr amplitude modulation.

For each element of the astigmatic cavity (Fig. 7) two corresponding beam matrices were compiled for p and s planes to describe the passage of the beam. Then two corresponding p and s matrices were calculated to describe the cavity round trip and to find the waist sizes for the p and s planes with a symmetrical change in the distance between the focusing mirrors. In the calculation the influence of the prism was considered small, since the transverse size and curvature of the wavefront in its place are large.

Fig. 7. Optical scheme of an astigmatic ring cavity for Kerr mode-locking.

For the sagittal plane s, for a section in the middle of the gain medium, the ray matrix for cavity round trip can be written as:

MMMLs=M2s*MYAG2orts*MYAG1ort*M10a*M5s*M8* M12*M9*M5s*M11a*MYAG3ort*MYAG2orts,   (7)

where the beam matrices for the cavity elements are represented below as:

$$M2s = \begin{pmatrix} 1 & 0 \\ -1/f_s & 1 \end{pmatrix}$$ - thermal or Kerr lens,

$$MYAG2orts = \begin{pmatrix} 1 & Lyag/ng/2 \\ 0 & 1 \end{pmatrix}$$ - gain medium,

$$MYAG1ort = \begin{pmatrix} ng & 0 \\ 0 & nv \end{pmatrix}$$ - interface,

$$M10a = \begin{pmatrix} 1 & Lo/2 + lo - loo \\ 0 & 1 \end{pmatrix}, \; M11a = \begin{pmatrix} 1 & Lo/2 - lo - loo \\ 0 & 1 \end{pmatrix}$$ - area between the spherical mirror and the gain medium,

$$M5s = \begin{pmatrix} 1 & 0 \\ -2\cos(\alpha)/R_m & 1 \end{pmatrix}$$ - spherical mirror,

$$M8 = \begin{pmatrix} 1 & a_1 \\ 0 & 1 \end{pmatrix}, \; M12 = \begin{pmatrix} 1 & a_3 \\ 0 & 1 \end{pmatrix}, \; M9 = \begin{pmatrix} 1 & a_2 \\ 0 & 1 \end{pmatrix}$$ - free space matrices, $$MYAG3ort = \begin{pmatrix} nv & 0 \\ 0 & ng \end{pmatrix}$$ - interface.

For the plane of incidence p for the section in the middle of the gain medium the ray matrix of the cavity round trip can be written as:

MMMLp=M2p*MYAG2pp*MYAG1p*M10a*M5p*M8* M12*M9*M5p*M11a*MYAG3p*MYAG2pp,   (8)

$$M2p = \begin{pmatrix} 1 & 0 \\ -1/f_p & 1 \end{pmatrix}$$ - thermal or Kerr lens,

$$MYAG2pp = \begin{pmatrix} 1 & Lyag/ng/2 \\ 0 & 1 \end{pmatrix}$$ - gain medium,

$$MYAG1p = \begin{pmatrix} A1 & 0 \\ 0 & D1 \end{pmatrix}$$ - interface,

$$M10a = \begin{pmatrix} 1 & Lo/2 + lo - loo \\ 0 & 1 \end{pmatrix}, \; M11a = \begin{pmatrix} 1 & Lo/2 - lo - loo \\ 0 & 1 \end{pmatrix}$$ - area between the spherical mirror and the gain medium,

$$M5p = \begin{pmatrix} 1 & 0 \\ -2/(R_m\cos(\alpha)) & 1 \end{pmatrix}$$ - spherical mirror,

$$M8 = \begin{pmatrix} 1 & a_1 \\ 0 & 1 \end{pmatrix}, \; M12 = \begin{pmatrix} 1 & a_3 \\ 0 & 1 \end{pmatrix}, \; M9 = \begin{pmatrix} 1 & a_2 \\ 0 & 1 \end{pmatrix}$$ free space matrices, $$MYAG3p = \begin{pmatrix} D1 & 0 \\ 0 & A1 \end{pmatrix}$$ - interface.

For interface matrices the coefficients were found through the angles of incidence and refraction in the gain medium:

A1=ng*cos(φ_{B2})/cos(φ_{B1}),   D1=nv*cos(φ_{B1})/cos(φ_{B2}),

where the angles of incidence (Brewster angle) and refraction are:

φ_{B1}=arctg(ng/nv), φ_{B2}=arcsin(nv/ng*sin(φ_{B1})).

From the resulting matrices describing the cavity round trip for the section in the middle of the gain medium one can find A, B, C, D elements as:

$$MMMLp,s = \begin{pmatrix} A_{p,s} & B_{p,s} \\ C_{p,s} & D_{p,s} \end{pmatrix}.$$

The beam radius and the radius of curvature of the wave front were found by solving the quadratic equation of complex parameter transformation:

$$\omega_{p,s} = \sqrt{\frac{2 \cdot \lambda \cdot B_{p,s}}{\pi \sqrt{4 - (A_{p,s} + D_{p,s})^2}}} \quad (9.a), \quad R_{p,s} = \frac{2B_{p,s}}{D_{p,s} - A_{p,s}} \quad (9.b)$$

Fig. 8. Stability parameters for p (*blue*) and s (*green*) planes when changing the distance Lo between the spherical mirrors (stable for |(A+D)/2|<1).

Fig. 9. Beam sizes in the gain medium for the plane of incidence on the mirror ω_p (*top*) and in the perpendicular plane ω_s (*bottom*) when changing the distance Lo between the spherical mirrors.

In general, the position of the waists between the focusing mirrors for the planes p and s is determined by the oblique incidence of the Gaussian beam on the gain medium at the Brewster angle and on the spherical mirrors.

As a result of performed calculations, the optimal angle of incidence on mirrors was found to be 15°, which ensures the greatest overlap of the beam stability regions for s and p planes and zeroes the first derivatives of the dependence of the waist sizes on the displacement of the mirrors. The optimal distance between the focusing mirrors was found to be ~0.118 m with the first derivatives of the beam size on the mirror displacement approaching zero simultaneously for p and s planes. This ensures the stability of Kerr mode-locking

and lasing power, as changing the beam size in the gain medium changes a number of lasing parameters, in particular the number of active centers used for amplification. The cavity was also optimized according to the angles of incidence on the spherical mirrors and the distance between them to ensure the required waist sizes in the gain medium for an astigmatic beam. The dimensions of the beam near the output mirror largely determine the Kerr amplitude modulation, since this region contain a diaphragm that ensures duration reduction of generated pulses.

Fig. 10. Beam size on the output mirror for the plane of incidence on the mirror ω_p (*blue*) and in the perpendicular plane ω_s (*green*) when changing the distance Lo between the spherical mirrors.

IV. Experimental Setup

In this work we have created a laser gyroscope based on YAG:Cr^{4+} gain medium (Fig. 11) with broad spectrum (1.3-1.6 μm), which allows to adjust the wavelength to the region of optimal lasing and to generate femtosecond pulses during mode-locking. As a result of optimization and tuning of the parameters of YAG:Cr^{4+} laser cavity, stable bidirectional lasing was obtained during mode-locking. The resulting output power in each direction reached 100 mW. The setup was based on a wavelength-tunable YAG:Cr^{4+} laser pumped by a fiber ytterbium laser or a mode-locked YAG:Nd^{3+} laser. To create non-reciprocity for counterpropagating waves we used an electro-optical phase modulator. To adjust the wavelength and create negative group delay dispersion we also used an IR quartz prism with low absorption in the operating wavelength range. The lasing linewidth and the wavelength were adjusted by changing the width and position of the diaphragm D and by rotating mirrors M3 and M4 (Fig. 11).

A YAG:Cr^{4+} crystal 2 cm long and 6 mm in diameter with edges oriented at the Brewster angle was used as gain medium. The concentration of Cr^{4+} ions was 5×10^{17} *cm*$^{-3}$. Characteristic crystal parameters [5, 11] are: upper state lifetime 3.6 μs, gain cross section at wavelength of 1.42 μm - $(7-8) \times 10^{-19}$ *cm*^{2}. The center of the absorption line is at wavelength of 1 μm and corresponds to the transition $^3B_1\left(^3A_2\right) \rightarrow ^3A_2\left(^3T_1\right)$. Absorption coefficient at wavelength 1.06 μm – around 2.5 *cm*$^{-1}$. Thus, the pump absorption over the length of the crystal is 90%. The laser operates between two wide bands $^3T_2$ and $^3A_2$ consisting of a set of energy vibrational sublevels. Thanks to this, the laser operates according to a four-level scheme.

Fig. 11. Experimental setup of a YAG:Cr^{4+} laser gyroscope.

During lasing the active element was placed in a precision thermostat and cooled to 10-12 degrees Celsius, which significantly reduces multiphonon relaxation and increases the lasing efficiency (Fig. 12). Stronger cooling is limited by condensate in the open cavity but can be achieved through sealing the optical scheme.

Experimental studies have shown that the competition of counterpropagating waves in a YAG:Cr^{4+} ring laser is significantly weakened (Fig. 12). This is possible due to such features of the gain medium as the ability to generate femtosecond pulses and peculiarities of interaction of Cr^{4+} active centres with polarized radiation associated with the orientation and placement of ions in the garnet crystal.

Fig. 12. Lasing power when changing the crystal temperature during minimal pump power.

The Kerr mechanism was used for mode-locking with Kerr lens forming in gain medium due to strong focusing and resulting in expansion of the beam at the leading and trailing edges of the pulse. Then, when a pulse with such transverse spatial distribution passes through the limiting diaphragm the leading and trailing edges are reduced, having greater losses due to the larger transverse size. This leads to the formation of shorter pulses.

On the other hand, amplitude modulation of radiation due to the Kerr lens and diaphragm leads to the appearance of side lasing components at the frequencies of neighboring modes, since the modulation frequency corresponds to the period of the pulse traversal of the cavity. This leads to the lock-in of the frequencies of neighboring modes by the side components, which makes the modes synchronized, and the spectrum becomes completely equidistant with phase difference between the modes being constant. Due to the modes being phased, the number of locked modes significantly reduces linewidth of each mode [14]. But this is only possible when loss modulation frequency is very stable and its linewidth is significantly narrower than the mode linewidth. Therefore, the stability of Kerr amplitude modulation is crucial for obtaining narrow longitudinal modes with low noise.

Fig. 13. Mechanism of Kerr loss modulation.

The Kerr lens in the gain medium is formed due to the Gaussian distribution of the transverse structure of the fundamental mode and requires very strong focusing of ultrashort pulses of high intensity. That required a more detailed consideration of the ring cavity configuration with fulfillment of a number of contradictory conditions as a large response to a change in the Kerr lens is required and at the same time high stability of the mode is desirable. After appropriate adjustments of the ring laser during mode-locking a pulse duration of about 400 fs was obtained. A decrease in intensity fluctuations, a narrowing of individual mode linewidth and reduced fluctuations of counterpropagating wave beat frequency were also observed.

Another important factor determining the stability of Kerr mode-locking and the lasing of femtosecond pulses is the amount of intracavity dispersion, which ensures the soliton mode of pulse propagation. We have performed calculations (Fig. 14) of the intracavity group delay dispersion for a YAG:Cr^{4+} gain medium of 2 cm long and a section in a quartz prism of 7.1 mm long, depending on the lasing wavelength.

Fig. 14. Dependence of the group delay dispersion for a YAG:Cr^{4+} gain medium of 2 cm long and a section of a quartz prism of 7.1 mm long on the lasing wavelength. Zero dispersion is observed at the wavelength of 1.46 μm.

It was found that a certain length of the medium section in a quartz prism together with gain medium provides zero disper-

sion at a lasing wavelength of 1.46 μm. This allows using only one dispersive prism, thereby reducing losses and distortion of the Gaussian beam in the laser cavity.

V. Noise suppression during mode-locking

Mode-locking leads to a significant noise suppression for each longitudinal mode. Physically, this is due to the phasing of oscillations in the modes and the uncorrelated nature of the noise in the modes themselves [14, 15]. To analyze and simulate the influence of lasing parameters in a ring laser during mode-locking on the mode linewidth and noise in the beat signal we used the following approach and system of equations [14]:

$$
\begin{aligned}
\frac{d\varphi_1}{dt} &= S_1 + \nabla_\nu - \frac{\alpha E_2}{E_1}\sin(\varphi_2 - \varphi_1) + \frac{F_1}{E_1} + \omega_{El} \\
\frac{d\varphi_2}{dt} &= S_2 + \frac{\alpha}{E_2}\left(E_3\sin(\varphi_2 - \varphi_3) + E_1\sin(\varphi_2 - \varphi_1)\right) + \frac{F_2}{E_2} + \omega_{El} \\
\frac{d\varphi_3}{dt} &= S_3 - \nabla_\nu - \frac{\alpha E_2}{E_3}\sin(\varphi_2 - \varphi_3) + \frac{F_3}{E_3} + \omega_{El}
\end{aligned}
\tag{10}
$$

Here φ_1, φ_2, φ_3 – phases of the central and two neighboring modes, E_1, E_2, E_3 – mode amplitudes, F_k – noise sources (uncorrelated for different modes), $\nabla_\nu = \omega_m - \Omega$, ω_m – modulation frequency, $\alpha = \omega_m \rho_m/(4\pi)$ - coupling coefficient between modes due to loss modulation: $\rho = \rho_0 - \rho_m.\cos(\omega_m t)$, Ω - frequency difference between adjacent modes, S_1, S_2, S_3 – contribution made by polarization of the gain medium (assumed to depend only on the amplitudes of the modes), and loss modulating signal is monochromatic, ω_{El} – frequency shift caused by phase electro-optical modulator. By solving the system of equations one can obtain an expression [14] for the spectral density of frequency fluctuations in the approximation of equal quality factors of modes and for low frequencies ($\omega \ll \alpha$):

$$
g_{\nu k}(0) = \frac{2}{2 + 1/x^2}\frac{D}{E_1^2}, \quad \text{where } x = E_1/E_2 = E_3/E_2.
$$

The mode linewidth can be written as: $\Delta\omega_k = \frac{1}{2}g_{\nu k}(0)$, and for $x \approx 1$ during mode-locking:

$$
\Delta\omega_k \approx \frac{1}{3}\Delta\omega_{k0}, \tag{11}
$$

where $\Delta\omega_{k0} = D/E_1^2$ - natural linewidth during free lasing. It is evident, that a reduction in the mode linewidth is proportional to the number of synchronized modes.

For a multi-frequency lasing regime during mode-locking one can obtain a system of equations similar to (12) in the approximation of amplitudes of all modes being equal ($E_k = E_0$) [14]:

$$
\begin{aligned}
\frac{d\varphi_1}{dt} &= -\alpha\sin(\varphi_2 - \varphi_1) + \frac{F_1}{E_0} + \omega_{El}, \\
\frac{d\varphi_n}{dt} &= -\alpha\sin(\varphi_{n+1} - \varphi_n) - \alpha\sin(\varphi_n - \varphi_{n-1}) + \frac{F_n}{E_0} + \omega_{El}, \\
\frac{d\varphi_N}{dt} &= -\alpha\sin(\varphi_{N+1} - \varphi_N) + \frac{F_N}{E_0} + \omega_{El}, \quad (n = 2, 3, \dots N - 1)
\end{aligned}
\tag{12}
$$

By solving the system for the case of uncorrelated noise in different modes for low frequencies and $\omega \ll \alpha$, the spectral density of frequency fluctuations can be written as:

$$
g_{\nu k}(0) = \frac{2}{N^2 E_0^2}\sum_{i=1}^{N} D_i
$$

Then for the mode linewidth: $\Delta\omega_k = \frac{1}{N^2}\sum_{i=1}^{N}\Delta\omega_i^0$ and if the natural mode linewidths during free lasing are close in value, one can obtain a reduction in the mode linewidth by N times:

$$
\Delta\omega_k = \frac{1}{N}\Delta\omega^0 \tag{13}
$$

Using systems of equations (10, 12) we simulated the dynamics of establishing the spectral parameters of the modes for different types of disturbances F_k introduced during lasing. In particular, we considered the influence of a frequency bias based on an electro-optical phase modulator and a perimeter control system based on an absorbing cell. It has been found that the variable phase shift introduced by the phase modulator should be significantly lower than the coupling between modes α.

VI. Alternating frequency bias based on electro-optical phase modulator

The coupling of counterpropagating waves through scattering on the resonator elements and on the inversion gratings in a solid-state laser gyroscope makes it important to choose the right mechanism for a frequency bias. The nonlinearity of the output frequency response, caused by the coupling of waves, leads to additional noise in the output beat signal of counterpropagating waves despite the significant suppression of this coupling due to the pulses passing the scattering elements one by one.

Using an alternating frequency bias based on phase modulator leads to a periodic phase shift of the lasing wave in a time shorter than the cavity round trip time. This leads to the necessity to take into account these mode-locking disturbances, as if the rate of phase change exceeds the magnitude of wave coupling, then the mode-locking is suppressed.

To check that we carried out a study of a frequency bias system based on phase delay of pulses in an electro-optical modulator. To model noise arising from periodic phase changes we used equations (10, 12). Periodically changing the optical length of the cavity using a modulator so that for a pulse propagating clockwise, the optical length increases, and for a counterclockwise pulse decreases, leads to a change in the beat frequency of counterpropagating waves. The signal is recorded when the pulses coincide at the output optics. To create an alternating frequency bias an electro–optical modulator inside the cavity is controlled by this signal from a high-speed photodetector that registers the output pulses obtained during mode-locking. The signal then goes to an amplifier and an adjustable delay line. After a broadband high-frequency power amplifier with a 50 Ohm load, connected in parallel to the electrodes of an electro-optical modulator, the signal creates a variable field strength on the LiNbO$_3$ optical crystal with the ability to adjust the field amplitude. Applying a control voltage to the modulator during the passage of a pulse leads to a change in the refractive index and to a shift in the frequency of the generated radiation relative to the counterpropagating wave according to (14, 15).

$$\Delta\varphi(U) = \frac{\pi n_e^3 r_{33} lU}{\lambda D}, \quad r_{ij} = \begin{pmatrix} 0 & -r_{22} & r_{13} \\ 0 & r_{22} & r_{13} \\ 0 & 0 & r_{33} \\ 0 & r_{51} & 0 \\ r_{51} & 0 & 0 \\ -r_{22} & 0 & 0 \end{pmatrix}, \quad (14)$$

where n_e – extraordinary wave refractive index, r_{ii} – electro-optical coefficient tensor, $r_{33}=30.8*10^{-12}$ m/V – electro-optical coefficient for a selected direction of the crystal of LiNbO$_3$), U – voltage applied to crystal, l – crystal length, λ – wavelength, D – crystal thickness. The LiNbO$_3$ crystal is oriented so that the radiation passes it at the Brewster angle. The change in lasing frequency when changing the phase in a ring laser (Fig. 16) can be written as:

$$\Delta\nu(U) = \frac{c}{2\pi L}\Delta\varphi(U). \quad (15)$$

Thus, a frequency difference arises between the counter-propagating beams and the bias is formed (Fig. 15). To record pulses, we employed a high-speed InGaAs PIN-photodiode with an amplifier (Fig. 16).

Fig. 15. Dependence of the frequency bias on the delay of the signal on the photodetector (U$_0$ ~ 25 V).

Fig. 16. Optical scheme of a YAG:Cr^{4+} ring laser gyroscope with mode-locking via an electro-optical phase modulator.

$$\Delta\nu = \frac{c}{2\pi L}\Delta\varphi(U(\tau)), \text{ where } U(\tau) = \left| U_0 \sin\left(\frac{2\pi c}{L}\tau - \varphi_0\right) \right|. \quad (16)$$

To form a control pulse, the signal from the photodetector was delayed so that the modulator voltage reached its maximum value at the moment of pulse of one direction of propagation passing it. The maximum opposite polarity is for a pulse of the opposite direction of propagation.

To obtain an alternating frequency bias, it is necessary to change the phase of the control signal, so that when a pulse passes in one direction, the phase shift increases, and for the pulse of opposite direction decreases (Fig. 16), and then vice versa. For that the counterpropagating pulses must pass through the modulator in antiphase.

VII. STABILIZING FREQUENCY OF MODE-LOCKED RING LASER BY AN ABSORBING CELL

The lasing of femtosecond pulses during mode-locking occurs during multifrequency lasing, when the modes are separated in frequency by in interval of ~ 283 MHz, and with gain bandwidth of 76 THz, which is enough for ~2.7*10^5 modes. Of these, 1.4*10^4 modes fall into the acetylene absorption line (4 THz) (Fig. 17). As some of the modes overlap with the acetylene absorption lines, the lasing frequency can be stabilized along these lines. Optimizing the mode frequency spacing can increase the absorption value, making the laser frequency stabilization system more accurate. That is important, as the high stability of the cavity length control system is a major factor affecting the accuracy a laser gyroscope.

Fig. 17. Acetylene absorption spectrum in the region of 1.5 μm.

A frequency stabilization system was simulated based on an acetylene absorbing cell with two bands of discrete absorption lines in the wavelength region of 1.5 μm (Fig. 16 - 18) for a mode-locked laser. Simulation has shown that by optimizing the intermode frequency interval one can obtain narrow absorption lines of ~ 0.5% with a width of ~ 250 MHz, ensuring frequency stabilization during multifrequency lasing. As the change in absorption is small, we used radiation modulation in front of the absorbing cell and a synchronous amplifier, which allows to obtain the necessary signal for the piezomotor control system. As the absorption lines of acetylene lie in the range of 1510-1540 nm, we had to tune the wavelength of YAG:Cr^{4+} laser to this range.

To stabilize the cavity length of the ring laser we employed an LIHV162 electronic unit and an acetylene cell made by VitaWave. The length of an acetylene cell $^{12}C_2H_2$ - 12 cm, diameter is 8 mm. At a pressure of 20 Torr, the width of the resonance line P9 ~ 5 pm (~666.2 MHz). Temperature dependence <0.01 pm/°C (1.33 MHz/°C). The electronic unit allows to link the frequency of the tunable laser to the resonance formed by the absorbing medium. The unit itself consists of a triangle signal generator, a sine wave generator, a phase shifter, an analog multiplier, a PID controller with adjustable gain bands and a power supply. Additionally, there is an optional high voltage output for piezomotor control. The operating principle of the frequency stabilization is

based on the fact that the value of the product of two sinusoidal signals of the same frequency is determined by the phase difference of these signals: $2*\sin(\omega t+\alpha)*\sin(\omega t) = \cos(\alpha) - \cos(2\omega t+\alpha)$, with a high frequency component being filtered. At the extremum the phase of the detected modulation signal changes sign, which can be registered. After filtering the high-frequency component the remaining signal of the dispersive form allows to bind the laser frequency to the reference resonance source. That allows to narrow the absorption line by more than two times, improving the accuracy of cavity length stabilization. The temperature drift of the acetylene absorption line will lead to zero drift of the gyroscope due to the dispersion of the gain medium and due to anomalous dispersion at the energy level. However, as the cavity is tuned to a group velocity dispersion close to zero, the effect of the frequency shift on the zero drift is insignificant.

Fig. 18. Dependence of the transmittance α of an acetylene cell on the laser frequency tuning $\Delta\nu$ during multifrequency lasing. The linewidth of the absorption maximum is ~250 MHz (line near $\Delta\nu$~0).

To stabilize the frequency during mode-locking an alternating frequency bias via an electro-optical modulator was used, with a synchronous detector extracting signal from a photodetector installed after an absorbing cell with acetylene (Fig. 16).

VIII. CONCLUSION

As a result of calculations and numerical modeling of lasing dynamics of a multifrequency ring laser based on equations for complex amplitudes and for phases of locked modes, it was determined that mode-locking narrows spectral linewidth of each mode in direct ratio to the number of modes being locked, which also significantly reduces noise and fluctuations of each mode. It was found that the variable phase shift introduced by the phase modulator should be significantly lower than the coupling coefficient between modes due to modulation. Simulation of lasing dynamics has shown that the inversion gratings for a YAG:Cr^{4+} ring laser are significantly reduced compared to a YAG:Nd^{3+} laser, which substantially weakens the coupling of counterpropagating waves. After modeling the formation of an astigmatic Gaussian beam, a cavity configuration was determined that ensures lasing stability under external disturbances with the first derivatives of these dependences being zeroed.

The parameters of the cavity length stabilization system were found, ensuring stability of the lasing frequency at the level of 0.01 pm/°C with intracavity group velocity being close to zero, which allows for insignificant shifts in frequencies of counterpropagating waves and, accordingly, to the lowered zero drift of the gyroscope. Additionally, the influence of the cavity length stabilization system on the amplitude and phase noise in the output signal was studied.

The influence of the frequency bias via an electro-optical phase modulator on the noise associated with disturbances in the phase of the lasing modes has been determined.

Overall, the studies have shown the possibility of stable operation of a YAG:Cr^{4+} mode-locked solid-state laser gyroscope with a cavity length stabilization system and an electro-optical frequency bias.

REFERENCES

[1] Aronowitz, F., Laser gyros, In: Lasers Applications. Russian edition, M: «Mir», 1974. p. 182-269.

[2] M.J. Bohn, The Ti:Sapphire Ring Laser Gyroscope, dissertation, University of New Mexico, 1998.

[3] Sorokin, E. Solid-State Materials for Few-Cycle Pulse Generation and Amplification. In: Kärtner, F.X. (eds) Few-Cycle Laser Pulse Generation and Its Applications. Topics in Applied Physics, vol 95. Springer, Berlin, Heidelberg.. https://doi.org/10.1007/978-3-540-39849-3_1

[4] H. Eilers et al., "Performance of a Cr:YAG laser", in IEEE Journal of Quantum Electronics, vol. 29, no. 9, pp. 2508-2512, Sept. 1993.

[5] A.G. Okhrimchuk and A.V. Shestakov, "Absorption saturation mechanism for YAG:Cr4+crystals", Phys. Rev. B 61, 988, 2000.

[6] G.M. Zverev and A.V. Shestakov, "Tunable Near-Infrared Oxide Crystal Lasers," in Advanced Solid State Lasers, M. Shand and H. Jenssen, eds., Vol. 5 of OSA Proceedings Series (Optical Society of America, 1989), paper BB5.

[7] I.T. Sorokina, S. Naumov, E. Sorokin, E. Wintner, and A.V. Shestakov, Directly diode-pumped tunable continuous-wave room-temperature Cr4+:YAG laser, Opt. Lett. 24, 1578-1580 (1999).

[8] Nikolai V. Kravtsov and E. G. Lariontsev, Self-modulation oscillations and relaxation processes in solid-state ring lasers, Quantum Electronics, 1994, Volume 24, Issue 10, Pages 841–856. DOI: https://doi.org/10.1070/QE1994v024n10ABEH000192

[9] Schwartz, S., Lauprêtre, T., Goldfarb, F., Bretenaker, F., Ghosh, R., Carusotto, I., Feugnet, G., Gutty, F., Pocholle, J.-P., The solid-state ring laser gyro: current and future trends, Proc. SPIE 8636, Advances in Slow and Fast Light VI, 86360M (March 6, 2013).

[10] Malakhov, A.N. and Mal'tsev, A.A., Width of the spectral lines of an N-mode system which has an equidistant spectrum of natural frequencies and operates in a self-synchronous mode, Radiofizika, 1973. Vol. 16, No. 8, p. 1211-1220.

[11] Y.Y. Broslavets, E.A. Polukeev, V.P. Surovtseva, A.A. Fomichev and V. G. Semenov, "Multifrequency YAG:Cr4+ Solid-State Laser Gyroscope, Perimeter Control and Dither Creation System, Operating Regimes", 29th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), Saint Petersburg, Russian Federation, 2022, pp. 1-4, doi: 10.23919/ICINS51784.2022.9815385.

[12] Schwartz, S., Gutty, F., Feugnet, G., Loil, É., and Pocholle, J.-P., Solid-state ring laser gyro behaving like its helium–neon counterpart at low rotation rates, OPTICS LETTERS. 2009 .Vol. 34, № 24, p. 3884-3886.

[13] Kravtsov, N.V., L'vov, B.V., Petrun'kin, V.Yu., Samusev, K.B., Shelayev, A.N.; Shokalo, V.I., Longitudinal mode locking in a solid state ring laser with semiconductor pumping. AVMEB, no. 4, 1987, 104-106.

[14] Ya.Z. Virnik, A.S. Kovalev and E.G. Lariontsev, Natural frequency fluctuations in a laser having locked modes, Radiophysics and Quantum Electronics, Volume 13, pages 1363-1367, (1970).

[15] V. Arzt, H. Haken, H. Risken, H. Sauermann, Ch. Schmid, and W. Weidlich "Quantum Theory of Noise in Gas and Solid State Lasers with an Inhomogeneously Broadened Line" Zeitschrift fur Physik 197, 207-227 (1966).

Study of Noise Characteristics of Signals from Navigation System Devices Using Wavelet-like Generalized Allan Estimators

M. Basarab

Dept. of Informatics and Control Systems
Bauman Moscow State Technical University
Moscow, Russia
basarab@bmstu.ru

B. Lunin

Dept. of Chemistry,
Lomonosov Moscow State University
Moscow, Russia
luninboris@yandex.ru

Abstract— **Novel approaches to the construction of new classes of two-point estimators, similar to the Allan variance, are proposed. One approach is based on the wavelet-like properties of infinitely differentiable atomic functions as a generalization of compactly supported B-splines. This variance is designed to analyze the phase and frequency instability of signals that occur in systems that require high-quality time and frequency sources. Unlike the usual Allan variance and its well-known generalizations and modifications, the proposed variance is suitable for analyzing signals in the presence of arbitrary power-law fluctuations due to the unique temporal and spectral properties of the atomic function. The second approach is based on windowed sine transformation and allows one to organize fast procedures for assessing frequency instability. Theoretical analysis of model signals with known power-law spectra, as well as a computational experiment on real data, demonstrate the advantages of new mathematical tools for studying frequency instability.**

Keywords— *Allan variance, power spectral density, windowed Fourier transform, wavelet, atomic function, wave solid-state gyroscope.*

I. INTRODUCTION

Allan variance (AVAR), or its square root (Allan deviance, ADEV), proposed by D. Allan in the 1960s [1, 2], is currently widely used as a measure of the stability of the output data of various navigation devices characterized by oscillatory behavior [3]. Despite its widespread use, the Allan variance has theoretical limitations on the types of power-law noise that can be detected, the power spectrum of which must be proportional to the power of the frequency with an exponent in a narrow range from -2 to 0. For exponents less than -2, the Allan variance is not defined and requires the introduction of a lower frequency cutoff limit. This problem for noise indicators in the ranges from -4 to -2, in particular, can be solved using the so-called Hadamard variance (HVAR) [4-6]. At the same time, to analyze noise with positive power spectra, it is practically necessary to introduce an upper frequency cutoff limit for both AVAR and HVAR. Thus, these types of variances are not very effective in the case of fast noisy processes. A number of works mentioned the possible presence of noises in signals, the nature of which is different from traditional types: WPM (white phase modulation), FPM (flicker phase modulation), WFM (white frequency modulation), FFM (flicker frequency modulation), RWFM (Brownian noise frequency modulation). These include, for example, sinusoidal noise, Markov noise, etc. [7]. In this regard, in addition to AVAR and HVAR, some other improved estimation methods were introduced: modified Allan variance (MVAR, Λ-counter

[8]), Ω-counter (parabolic variance, PVAR) [8-11], etc. A close connection was established between Allan variance and its generalizations with wavelet transform based and wavelet variance [12, 13]. At the same time, in general, the problem associated with limiting the frequency band of the studied types of noise has not been eliminated, which is especially important when studying complex ensembles of noise signals from sensitive elements of inertial navigation systems devices: microelectromechanical gyroscopes and accelerometers, solid-state wave gyroscopes (hemispherical resonator gyroscopes, HRG), fiber-optical gyroscopes, etc.

The aim of this work is to consider new universal estimates of frequency instability that allow solving the above problems. One of the possible approaches is based on the wavelet-like properties of finite infinitely differentiable atomic functions [14], which are solutions to special classes of functional differential equations and are used in problems of digital signal processing of various physical nature. Families of such functions and variances based on them can serve as the basis for constructing banks of high-order differentiating filters, which makes it possible to detect noise with power spectral density with indicators that vary over a wide range, including fractional indicators. The paper examines in more detail the variance of upVAR [15] based on the most well-known atomic function up(x).

In addition, using the connection between wavelet analysis and windowed Fourier transform [16], a fundamentally new type of windowed cosine variance (CVAR) was derived based on the windowed segment of the sine function. This type of variance has broad prospects for optimizing calculations through the use of discrete transformations [17] and the use of fast Fourier analysis algorithms.

As an example, the work demonstrates the applicability and effectiveness of the new family of Allan variances for analyzing the inhomogeneity (roughness) profiles of the surface of the HRG resonator sensitive element before and after various types of technological processing (grinding, balancing by chemical etching) [18]. The generalized Allan variances based on an infinitely differentiable atomic function and the sine function demonstrate the greater sensitivity in comparison with the classical Allan variance, which allows for fine analysis of inhomogeneous structures.

Computational aspects of estimating the proposed variances are discussed. A conclusion is drawn about the advisability of using new classes of estimators when studying processes in sensitive elements of inertial navigation devices characterized by complex nonlinear

dynamics, as well as when analyzing spatial structures with roughness of fractal geometry.

II. CLASSIC ALLAN VARIANCE

Let us consider a general approach to the analysis of frequency fluctuations of continuous signals based on an analog filtering apparatus [4, 6]. Let $y(t)$ be a continuous time signal with zero mean. Its autocorrelation function has the form

$$R_y(\tau) = \langle y(t)y(t+\tau) \rangle = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} y(t)y(t+\tau)\, dt.$$

The two-way power spectral density (PSD) of the signal, $S_y^{\mathrm{II}}(f)$, and its autocorrelation function are related by the Fourier transform:

$$S_y^{\mathrm{II}}(f) = \int_{-\infty}^{\infty} R_y(\tau)e^{-i2\pi f\tau}\, d\tau,$$

$$R_y(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_y^{\mathrm{II}}(f)e^{i2\pi f\tau}\, df.$$

In measurement practice, one-sided PSD is more often used:

$$S_y(f) = \begin{cases} 2S_y^{\mathrm{II}}(f), & f \geq 0, \\ 0, & f < 0. \end{cases}$$

We will study model noise signals of a power-law spectrum, the PSD of which can be written in the form

$$S_y(f) = A_\alpha f^\alpha, \tag{1}$$

where A_α is the amplitude and α is noise type indicator.

Due to the fact that the signal true variance is not defined for integer values $\alpha < 0$, different estimators of frequency instability are used for different types of noise.

To describe the signals determined by the PSD components at low frequencies, it is necessary to study the variance of the signal $y(t)$ averaged over the interval τ using a moving average filter with an impulse response

$$g(t,\tau) = g_\Pi(t,\tau) = \frac{1}{\tau}\Pi_\tau\left(t - \frac{\tau}{2}\right), \tag{2}$$

where

$$\Pi_\tau(t) = \begin{cases} 1, & |t| \leq \tau/2, \\ 0 & |t| > \tau/2. \end{cases}$$

Thus, the averaged signal $z(t,\tau)$ is determined by convolution

$$z(t,\tau) = g(t,\tau)*y(t) = \frac{1}{\tau} \int_{t-\tau}^{t} y(t)\, dt. \tag{3}$$

The frequency response of the moving average filter has the form

$$G(f) = G_\Pi(f) = \mathrm{sinc}(\pi\tau f)e^{i\pi\tau f}, \tag{4}$$

where $\mathrm{sinc}(x) \equiv x^{-1}\sin x$.

Power spectral density of the averaged profile function is

$$S_z(f) = |G(f)|^2 S_y(f). \tag{5}$$

True variance of the average signal,

$$\sigma_z^2 = \int_0^\infty \mathrm{sinc}^2(\pi\tau f)S_y(f)\, df, \tag{6}$$

is not defined for the power law (1) at $\alpha < 0$, therefore, in series with G, it is necessary to introduce an additional differentiating filter D. The variance of the resulting output signal $u(t,\tau)$ has a spectrum

$$\sigma_u^2 = A_\alpha \int_0^\infty \mathrm{sinc}^2(\pi\tau f)|D(f)|^2\, f^\alpha\, df. \tag{7}$$

For the convergence of integral (7), the amplitude-frequency characteristic of the new filter must have the form $D(f) = f^\beta$ with an exponent $\beta > -\alpha/2$ as f asymptotically tends to zero. Such filters can be obtained using several continuous derivatives of the signal $z(t,\tau)$ in time, since the derivative in the time domain corresponds to the PSD multiplied by f^2. PSD of signal $z^{(n)}(t,\tau)$ equals

$$S_u^{(n)}(f) = (2\pi f)^{2n} S_z(f) =$$
$$= (2\pi f)^{2n}\mathrm{sinc}^2(\pi\tau f)S_y(f). \tag{8}$$

The variance of the resulting signal is determined as

$$\sigma_u^2(\tau;n) = \int_0^\infty (2\pi f)^{2n}\mathrm{sinc}^2(\pi\tau f)S_y(f)\, df. \tag{9}$$

It is impossible to practically implement (9), since the corresponding frequency response must be linear. Therefore, difference filter D should be used that approximate the frequency response near zero frequency. The simplest family of n-th order difference filters has an impulse response

$$d^{(n)}(\tau;n) = \sum_{k=0}^{n} (-1)^k \binom{n}{k}\delta(t - k\tau), \tag{10}$$

where $\delta(t)$ is Dirac delta function.

The Fourier transform (10) gives the impulse response of the filter

$$D^{(n)}(f) = (2i)^n \sin^n(\pi\tau f)e^{-i\pi n\tau f}. \tag{11}$$

Filter output is equal to

$$u^{(n)}(t,\tau) = \frac{1}{\mu_n}\left[d^{(n)}(t,\tau)*g(t,\tau)\right]*y(t), \tag{12}$$

where μ_n is the normalization factor used to keep the white noise variance produced by the filter unchanged, or to ensure the response to frequency drift remains unchanged.

Variance of $u^{(n)}(t,\tau)$ can be written as

$$\sigma_u^2(\tau;n) = \frac{1}{\mu_n^2} \int_0^\infty \left|D^{(n)}(f)G(f)\right| S_y(f)\, df =$$
$$= \frac{2^{2n}}{\mu_n^2} \int_0^\infty \sin^{2n}(2\pi f)\mathrm{sinc}^2(\pi\tau f)S_y(f)\, df. \tag{13}$$

This variance converges at $\alpha > -(n+2)$. To ensure convergence of $\sigma_u^2(\tau;n)$ for positive values of α, the upper limit of integration in (13) must be limited by the cutoff frequency.

Combining a low pass filter $G(f)$ with a high pass filter $D^{(n)}(f)$ gives us a bandpass filter. By varying the filter

parameter τ, we obtain a bank of different bandpass filters, similarly to multiple-scale wavelet analysis.

Most known stability variances are generated by equation (13). In this case, the Allan variance (AVAR) corresponds to the case $n = 1$:

$$\sigma_u^2(\tau;1) = 2\int_0^\infty \sin^2(2\pi f)\operatorname{sinc}^2(\pi\tau f)S_y(f)\,\mathrm{d}f. \qquad (14)$$

AVAR is defined for a power spectrum with $-2 \le \alpha \le 0$. For $\alpha > 0$ it is necessary to take into account a high cutoff frequency f_{high}, and for a power law with $\alpha = -3$ or $\alpha = -4$ a low cutoff frequency f_{low} must be chosen. Allan variance is also called two-sample variance in contrast to the one-sample true variance. The impulse response of the Allan variance filter is a Haar wavelet basis function [12].

Other types of Allan variance were also introduced, such as the modified Allan variance (MVAR), the parabolic variance (PVAR), etc. Having a more complex form compared to AVAR and HVAR, they still have similar restrictions on the power parameter due to the fact that their impulse characteristics represent a limited number of times continuously differentiable finite polynomial or piecewise polynomial functions. Thus, the most general approach to constructing two-point estimators of various orders is based on the following scheme [8]:

$$\mathrm{xxx}\,\sigma_y^2(\tau) = \left\langle\left[\int_{-\infty}^\infty y(t')w_{\mathrm{xxx}\,\sigma}(t'-t)\,\mathrm{d}t'\right]^2\right\rangle;$$
$$w_{\mathrm{xxx}\,\sigma}(t) = \frac{1}{\sqrt{2}}\left[w_{\mathrm{xxx}}(t+\tau/2) - w_{\mathrm{xxx}}(t-\tau/2)\right]. \qquad (15)$$

Here xxx denotes the type of variance depending on the function $w_{\mathrm{xxx}}(t)$. Some types of weight functions $w_{\mathrm{xxx}}(t)$ and their spectra $W_{\mathrm{xxx}}(f)$ are given in Table I. The corresponding spectra of functions $w_{\mathrm{xxx}\,\sigma}(t)$ have the form

$$W_{\mathrm{xxx}\,\sigma}(f) = -i\sqrt{2}\sin(\pi f\tau)W_{\mathrm{xxx}}(f). \qquad (16)$$

TABLE I. WEIGHT FUNCTIONS OF COMMON TYPES OF ALLAN VARIANCES

Type	Weight function $w_{\mathrm{xxx}}(t)$	Fourier transform $W_{\mathrm{xxx}}(f)$
AVAR	$w_\Pi(t) = \dfrac{1}{\tau},$ $\lvert t\rvert \le \tau/2$	$W_\Pi(f) = \operatorname{sinc}(\pi f\tau)$
MVAR	$w_\Lambda(t) = 1 + \dfrac{\lvert t\rvert}{\tau^2},$ $\lvert t\rvert \le \tau$	$W_\Lambda(f) = \operatorname{sinc}^2(\pi f\tau)$
PVAR	$w_\Omega(t) = \dfrac{3}{\tau}\left(\dfrac{1}{2} - \dfrac{2t^2}{\tau^2}\right),$ $\lvert t\rvert \le \tau/2$	$W_\Omega(f) = 3\dfrac{\operatorname{sinc}(\pi f\tau) - \cos(\pi f\tau)}{(\pi f\tau)^2}$

The choice of different functions $w_{\mathrm{xxx}}(t)$ does not ensure sufficient smoothness of function (15) at zero, as well as at the edges of the support. This negatively affects the spectral characteristics, resolution and sensitivity of generalized Allan filters. This paper proposes a new approach to directly constructing function (15) based on functions of high and even infinite smoothness. In the next section, the variance based on periodic segments of the sine function of multiple frequency (CVAR), which in its properties is close to Ω-variance (parabolic variance, PVAR) [9-11], will be described for the first time. Next, a variance will be presented based on the most well-known basic function up(t) from a wide family of infinitely differentiable compactly supported atomic functions [15]. A new type of wavelet-like variance (upVAR) will allow us to analyze noise signals with arbitrary indicators α.

III. VARIANCE CVAR

It is easy to see that while the impulse response of the AVAR filter is discontinuous, the first and second derivatives of the MVAR and PVAR are discontinuous, respectively. A further increase in the degree of smoothness is possible by using piecewise polynomial functions of higher degrees as impulse responses (in particular, derivatives of the corresponding Schoenberg B-splines [19]), however, such an increase in the degree of polynomials can lead to a loss of stability and large errors in calculations, which is critical when processing noisy data.

An alternative way is to use non-polynomial generating functions, in particular harmonic ones. Let

$$w_{\mathrm{xxx}}(t) = w_{\cos}(t) = \frac{a}{2}\cos\left(\frac{\pi}{\tau}t\right) \quad (\lvert t\rvert \le \tau/2). \qquad (17)$$

It is easy to see that

$$w_{\mathrm{xxx}\,\sigma}(t) = w_{\cos\sigma}(t) = a_\tau\sin\left(\frac{\pi}{\tau}t\right) \quad (\lvert t\rvert \le \tau). \qquad (18)$$

Thus, the impulse function represents one period of a sinusoid cut out by a rectangular window of width 2τ. Since the generating function is actually the half-period of the cosine function, we will call this type of variance CVAR (cos-variance).

Coefficient

$$a_\tau = \frac{\pi}{2\sqrt{2}\tau} \qquad (19)$$

is found from the condition of ensuring a standard response to frequency drift:

$$\left(2a_\tau D_y\int_0^\tau tw_{\cos\sigma}(t)\,\mathrm{d}t\right)^2 = \frac{D_y^2\tau^2}{2}. \qquad (20)$$

The Fourier transform (16) has the form

$$W_{\cos\sigma}(f) = -i2\pi\tau a_\tau\frac{\sin 2\pi f\tau}{\pi^2 - (2\pi f\tau)^2}. \qquad (21)$$

In digital signal processing, the choice of averaging interval τ is important. If we limit ourselves to variance without overlap, then for a discrete signal $y_k = y(t_k)$, $k = 0,1,...,N-1$, with length $N = 2^p$ it seems natural to choose binary averaging intervals

$$\tau_l = \frac{1}{2^l}, \quad l = 1, 2, ..., p. \qquad (22)$$

With this choice of parameters, the expression for finding the corresponding variance is calculated as

$$\sigma_y^2(\tau_{l-1}) = \frac{\pi^2}{2^{1-2l}} \sum_{j=0}^{2^l-1} \left\{ \sum_{k=j2^{p-l}}^{(j+1)2^{p-l}-1} y_k \sin\left[\frac{2\pi}{2^{p-l}}\left(k + \frac{1}{2}\right)\right]\right\}^2. \qquad (23)$$

and turns out to be related to the coefficients of the discrete windowed sine Fourier transform [17].

IV. VARIANCE upVAR

The impulse response CVAR (18) is infinitely differentiable at the origin. At the same time, a discontinuity of the first kind (non-differentiability) is formed at both ends of the interval, which also leads to limitations in terms of the spectral resolution of various types of noise. To avoid this, you can use artificial techniques by selecting more complex combinations of trigonometric functions (for example, continuity of the first derivative can be achieved by taking its square, $\cos^2 t$, instead of the weight function $\cos t$). Next, we give an example of a function of a special form that ensures continuity of both the impulse response of the filter and all its derivatives over the entire averaging interval.

The atomic function $up(t)$ is a finite solution to the following functional differential equation [14]:

$$\frac{1}{2} up'(t) = up(2t+1) - up(2t-1). \qquad (24)$$

Among the important properties of the function $up(t)$, it should be noted that it is even, infinitely differentiable, $\operatorname{supp} up(t) = (-1, 1)$, $up(0) = 1$, and

$$\int_{-1}^{1} up(t)\,dt = 1.$$

Both the first and all higher derivatives of $up(t)$ are similar to the original function and can be calculated recursively using (24). All derivatives of the function $up(t)$ at the ends of the support and at the origin are equal to zero (ideally flat curve at points -1, 0, 1).

Graphs $up(t)$ and its first derivative are shown in Fig. 1.

Despite its infinite smoothness, like other atomic functions, the function $up(t)$ is not analytic, i.e. its Taylor series expansion does not converge to the function itself. However, to find values of $up(t)$ at arbitrary points of the interval $[0,1]$, we can use the Fourier series, as well as other rapidly convergent series and expansions. In addition, the function $up(t)$ can be accurately calculated at binary rational points $k2^{-n}$ through its moments [14, 15], which is important for discrete implementation.

The Fourier transform of $up(t)$ is a rapidly converging infinite product

$$\widehat{up}(f) = \prod_{k=1}^{\infty} \operatorname{sinc} \frac{f}{2^k}. \qquad (25)$$

Fig. 1. Atomic function $up(t)$ (solid line) and its first derivative (dashed line)

It is also necessary to know the value

$$\chi = 2\int_0^1 up^2(t)\,dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} \widehat{up}^2(f)\,df, \qquad (26)$$

which can be approximately evaluated using the Fourier expansion $up(t)$:

$$\chi = \frac{1}{2} + \sum_{j=1}^{\infty}\prod_{k=1}^{\infty} \operatorname{sinc}^2 \frac{\pi(2j-1)}{2^k} \approx 0.809. \qquad (27)$$

Two-sided moments of any even order of $\widehat{up}^2(f)$ can also be calculated (two-sided odd moments are zero):

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} f^{2n}\,\widehat{up}^2(f)\,df = 2^{n(n+1)}\chi. \qquad (28)$$

Integral (28) converges for any n due to the fact that the spectrum $\widehat{up}(f)$ decays faster than any power law.

The calculation of odd moments of $\widehat{up}^2(f)$ can be realized numerically using the expansion [15]:

$$\frac{1}{2\pi}\int_0^{\infty} f^{2n-1}\,\widehat{up}^2(f)\,df = \frac{1}{2\pi}\sum_{k=1}^{\infty} k^{2n-1}\,\widehat{up}^2(k). \qquad (29)$$

Let us introduce the impulse response of the filter based on the derivative of the atomic function $up(t)$:

$$w_{up\sigma}(t, \tau) = -\frac{1}{\tau\sqrt{2}} up'\left(\frac{t}{\tau}\right). \qquad (30)$$

The corresponding frequency response has the form

$$W_{up\sigma}(f, \tau) = -i\sqrt{2}\pi\tau f\,\widehat{up}(2\pi\tau f). \qquad (31)$$

Obviously,

$$\left|W_{up\sigma}(f, \tau)\right|^2 = 2(\pi\tau f)^2 \sin^2(\pi\tau f)\prod_{k=1}^{\infty} \operatorname{sinc}^2 \frac{\pi\tau f}{2^k}. \qquad (32)$$

In particular, as f tends to zero, we have

$$\left|W_{up\sigma}(f, \tau)\right| \approx \sqrt{2}\pi\tau\,|f|, \qquad (33)$$

which corresponds to the FM-noise dependence with the characteristic f^{-2}.

Using the PSD $S_y(f)$, we can calculate the new upVAR variance as

$$\sigma_{\text{upVAR}}^2(\tau) = \int_0^\infty \left|W_{\text{up}}(f)\right|^2 S_y(f)\, d f. \tag{34}$$

Replacing $S_y(f)$ with any polynomial power dependence of even order, $S_y(f) = A_{2\alpha} f^{2\alpha}$, we obtain the following analytical expression:

$$\sigma_{\text{upVAR}}^2(\tau) = \frac{2^{\alpha(\alpha+1)}\chi A_{2\alpha}}{\pi^{2\alpha}\tau^{2\alpha+1}}. \tag{35}$$

For the case of polynomial power laws of odd order, one can use expansion (29), keeping a small number of first terms.

Discretization can be done on a binary rational grid, similarly to CVAR.

Table II provides a summary of some known two-point estimators (AVAR, MVAR, PVAR) and new ones (CVAR, upVAR). Here the following trigonometric integrals are used for CVAR:

$$\text{Si}(t) = \int_0^t \frac{\sin x}{x}\, d x, \quad \text{Cin}(t) = \int_0^t \frac{1-\cos x}{x}\, d x.$$

All variances are normalized from the condition of ensuring the same response $D_y^2\tau^2/2$ to frequency drift $\dot{y} = D_y$, similar to (20).

TABLE II. EXPRESSIONS FOR RESPONSE OF SOME KNOWN AND NEW TYPES OF VARIANCES FOR COMMON TYPES OF NOISE (τ_0 IS THE NYQUIST FREQUENCY, $\gamma \approx 0.577$ IS THE EULER–MASCHERONI CONSTANT)

Noise	$S_y(f)$	Variance				
		$\sigma_{\text{AVAR}}^2(\tau)$	$\sigma_{\text{MVAR}}^2(\tau)$	$\sigma_{\text{PVAR}}^2(\tau)$	$\sigma_{\text{CVAR}}^2(\tau)$	$\sigma_{\text{upVAR}}^2(\tau)$
WPM	$A_2 f^2$	$\dfrac{3A_2}{8\pi^2\tau_0\tau^2}$	$\dfrac{3A_2}{8\pi^2\tau^3}$	$\dfrac{3A_2}{2\pi^2\tau^3}$	$\dfrac{\pi^2 A_2}{64\tau^3}$	$\dfrac{4\chi A_2}{\pi^2\tau^3}$
FPM	$A_1 f^1$	$\dfrac{\left(3\gamma - \ln 2 + 3\ln\frac{\pi\tau}{\tau_0}\right)A_1}{4\pi^2\tau^2}$	$\dfrac{(24\ln 2 - 9\ln 3)A_1}{8\pi^2\tau^2}$	$\dfrac{3(\ln 16 - 1)A_1}{2\pi^2\tau^2}$	$\dfrac{\pi A_1}{16\tau^2}\text{Si}(2\pi)$	$\dfrac{1.383A_1}{\pi\tau^2}$
WFM	$A_0 f^0$	$\dfrac{A_0}{2\tau}$	$\dfrac{A_0}{4\tau}$	$\dfrac{3A_0}{5\tau}$	$\dfrac{\pi^2 A_0}{16\tau}$	$\dfrac{\chi A_0}{\tau}$
FFM	$A_{-1} f^{-1}$	$2A_{-1}\ln 2$	$\dfrac{(27\ln 3 - 32\ln 2)A_{-1}}{8}$	$\dfrac{2(7-\ln 16)A_{-1}}{5}$	$\dfrac{\left[\pi\,\text{Si}(2\pi)+\text{Cin}(2\pi)\right]A_{-1}}{4}$	$0.643\pi A_{-1}$
RWFM	$A_{-2} f^{-2}$	$\dfrac{2\pi^2 A_{-2}\tau}{3}$	$\dfrac{11\pi^2 A_{-2}\tau}{20}$	$\dfrac{26\pi^2 A_{-2}\tau}{35}$	$\dfrac{3\pi^2 A_{-2}\tau}{4}$	$\chi\pi^2 A_{-2}\tau$

V. NUMERICAL EXPERIMENT

In [18], the main theoretical prerequisites for the application of the Allan variance method in the analysis of the roughness of surface profiles were presented. The classical continuous Allan variance is described in terms of time-continuous signals with zero mean. In this case, a formal transition is made from the time signal to the spatial function of the height of the surface profile and, accordingly, from the time coordinate to the spatial coordinate. The term "signal" is formally used to refer to the height of a profile. As an example, let us consider the analysis of the roughness profile of a section of the surface of a quartz HRG resonator before and after several cycles of chemical etching (Fig. 2).

Fig. 2. Height of surface profiles (µm) before (thin line) and after treatment (bold line) along the length (µm)

A qualitative change in the nature of the surface relief is clearly visible: the concentration of small-scale surface

defects has decreased, and large-scale ones have increased. The average surface roughness (Ra) increased from 0.5 µm to 1.5 µm. However, since this parameter does not describe all changes in the surface during chemical treatment, a quantitative analysis was carried out using the classical Allan variance.

Skipping the quantitative side of the issue, we present the results of assessing profiles using three types of variances: AVAR, CVAR, upVAR (Fig. 3). The digitized profile length was 1024 samples, the averaging step τ was chosen to be binary.

It is clear from the graphs that the behavior of the classical Allan variance curve is practically independent of scale and corresponds to the same noise component (its own for situations before and after chemical treatment). The CVAR variance plots already show slight differences in the slope angle at small scales. But still, these differences become most noticeable when using the upVAR variance.

CONCLUSION

Despite the fact that the variances proposed in this work (CVAR, upVAR) refer to the so-called non-overlapping two-point estimators, it is not difficult to obtain higher-level generalizations (3-point estimators, etc.) with varying degrees

of overlap. In this case, in the case of CVAR, it is necessary to take more than 2 half-waves of the harmonic function, and for upVAR, derivatives of the original function higher than the 1-st order. The results of numerical experiments show the greater sensitivity of upVAR to the analysis of noisy structures of time series and roughness of surface profiles. At the same time, it should be noted that the computational costs are significantly higher than those of traditional variances such as AVAR, MVAR or PVAR. If such a drawback is not critical for the analysis of surface profile data, then for processing time series data from gyroscopic sensors in real time, more optimal schemes for calculating atomic functions, similar to fast "divide and conquer" algorithms in spectral and wavelet analysis (FFT etc.), should be developed. The CVAR variance, in turn, being comparable to the indicated classical variances, has great capabilities in terms of speed and implementation of fast calculation schemes.

Fig. 3. Different types of variances for profiles before treatment (upper curve) and after treatment (lower curve)

REFERENCES

[1]   D. W. Allan, "Statistics of atomic frequency standards," Proc. of the IEEE, vol. 54, no. 2, pp. 222-231, 1966.

[2]   IEEE 1554-2005. IEEE recommended practice for inertial sensor test equipment, instrumentation, data acquisition and analysis. IEEE, 2005.

[3]   D. W. Allan, "Historicity, strengths, and weaknesses of Allan variances and their general applications," Gyroscopy Navig., 2016, vol. 7, no. 1, pp. 1-17. doi: 10.1134/S2075108716010028

[4]   J. Rutman, "Characterization of phase and frequency instabilities in precision frequency sources: Fifteen years of progress," Proc. IEEE, vol. 66 (9), pp. 1048-1075, 1978, doi: 10.1109/PROC.1978.11080.

[5]   W. Riley, Handbook of frequency stability analysis, NIST special publication 1065, Washington, 2008.

[6]   A. Makdissi, F. Vernotte, and E. De Clerq, "Stability variances: A filter approach," IEEE Trans. on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 57 (5), pp. 1011-1028, 2010, doi: 10.1109/tuffc.2010.1513.

[7]   B. S. Lunin, V. A. Matveev, and M. A. Basarab, Solid-state wave gyroscope. Theory and technology. Moscow, Radiotekhnika, 2014. (in Russian)

[8]   S. T. Dawkins, J. J. McFerran, and A. N. Luiten, "Considerations on the measurement of the stability of oscillators with frequency counters," IEEE Trans. on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 54 (5), pp. 918-925, 2007, doi: 10.1109/tuffc.2007.337.

[9]   E. Benkler, C. Lisdat, and U. Sterr, "On the relation between uncertainties of weighted frequency averages and the various types of Allan deviations," arXiv:1504.00466v3, April 2015.

[10] E. Rubiola, M. Lenczner, P.-Y. Bourgeois, and F. Vernotte, "The Ω counter, a frequency counter based on the linear regression," arXiv:1506.05009v1, June 2015.

[11] F. Vernotte, M. Lenczner, P.-Y. Bourgeois, and E. Rubiola, "The parabolic variance (PVAR), a wavelet variance based on the least-square fit," IEEE Trans. on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 63 (4), pp. 611-623, 2016, doi: 10.1109/tuffc.2015.2499325.

[12] D. B. Percival, "A wavelet perspective on the Allan variance," IEEE Trans. on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 63 (4), pp. 538-554, 2016, doi: 10.1109/TUFFC.2015.2495012.

[13] D. B. Percival, "On estimation of the wavelet variance," Biometrika, 82, pp. 619–631, 1995, doi: 10.1093/biomet/82.3.619.

[14] V. A. Rvachev, "Compactly supported solutions of functional-differential equations and their applications," Russian Math. Surveys, vol. 45 (1), pp. 87-120, 1990.

[15] M. A. Basarab, "The new wavelet-like Allan variance based on the atomic function," 2021 Photonics & Electromagnetics Research Symposium (PIERS), Hangzhou, China, 2021, pp. 2870-2877, doi: 10.1109/PIERS53385.2021.9694896.

[16] C. K. Chui. An Introduction to Wavelets. San Diego, CA: Academic Press, 1992.

[17] B. G. Sherlock and Y. P. Kakad, "Windowed discrete cosine and sine transforms for shifting data," Signal Processing, 81 (7), pp. 1465-1478, 2001, doi:10.1016/S0165-1684(01)00033-0.

[18] B. S. Lunin, M. A. Basarab, N. S. Konnova, and I. S. Stroganov, "Introduction of the Allan variance method in studying the physical and chemical processes on the solid body surface," Herald of the Bauman MSTU, Series Natural Sciences, 2023, vol. 108 (3), pp. 20-36, doi: 10.18698/1812-3368-2023-3-20-36.

[19] M. Unser, A. Aldroubi, and M. Eden, "Polynomial spline signal approximations: Filter design and asymptotic equivalence with Shannon's sampling theorem," IEEE Trans. on Information Theory, vol. 38 (1), pp. 95-103, 1992, doi: 10.1109/18.108253.

Four-Frequency Zeeman Laser Gyroscope: System for Separating Counterpropagating Waves Beat Signals for Orthogonal Polarizations and Control of the Cavity Perimeter using the Co-Directed Wave Beat Signal

Yu.Yu. Broslavets
MIPT
JSC «Lasex»
Dolgoprudny, Russia
laseruu@mail.ru

E.A. Polukeev
MIPT
JSC «Lasex»
Dolgoprudny, Russia
epolukeev2105@yandex.ru

D.S. Redichkina
MIPT
Dolgoprudny, Russia
dariaredichkina@gmail.com

A.A. Fomichev
MIPT
JSC «Lasex»
Dolgoprudny, Russia
laser@mail.mipt.ru

V.G. Semenov
MIPT
JSC «Lasex»
Dolgoprudny, Russia
valerii.semenov.g@gmail.com

A.R. Pokrovskaya
MIPT
Dolgoprudny, Russia
pokrovskaya3113@gmail.com

Abstract— **The paper presents a system for separating beat signals of counterpropagating waves of orthogonal circular polarizations for a Zeeman four-frequency laser gyroscope that is based on mathematical processing of beat signals after a non-polarizing beam combining optics. It was found that a small channel crosstalk requires a significantly higher frequency of the bias compared to fully optical signal separation for the same angular velocities. To stabilize the cavity length of a four-frequency laser gyroscope and mathematically separate beat signals, a system based on maximizing the beat frequency of co-directed waves of orthogonal polarization is found to be best suited.**

Keywords — laser gyroscope, four-frequency laser gyroscope, frequency lock-in, nonplanar cavity, Zeeman frequency dither.

I. INTRODUCTION

Zeeman He-Ne laser gyroscopes are widely used due to their high mechanical strength and workability under high accelerations and impacts [1-5]. At the same time, they have high magnetic sensitivity. To subtract the influence of external magnetic field and increase sensitivity to rotation a new generation of Zeeman laser gyroscopes operating in four-frequency regime is being developed. This has become possible due to the latest technological advances such as: new designs of multilayered dielectric coatings, superpolished mirror substrates, FPGA-based compact electronic components with real-time data processing, new error-compensating computer algorithms. Four-frequency lasing requires separation of the beat signals of counterpropagating waves of orthogonal polarization, which is complicated due to circularly polarized light becoming highly elliptical when passing through the output mirror at oblique incidence. In this paper we present two types of four-frequency Zeeman laser gyroscopes based on laser cavities in the form of an isohedral tetrahedron and a regular tetrahedron with accordingly different angles of incidence on

the output mirror and different ellipticity of the output radiation. To separate the beat signals of counterpropagating waves of orthogonal polarizations we utilized both fully optical and mathematical method, with latter one implemented both in the form of real-time electronic signal processing and algorithmic post-processing. Maintaining the cavity length was carried out by a novel method based on measuring the beat signal of co-directed waves and tuning the beat frequency to a maximum, which leads to stabilization along the neon spectral line. Through this method the suppression of the influence of external magnetic fields by a factor of ~1000 was achieved.

Numerical modeling of the beat signal of co-directed waves of orthogonal polarizations changing through varying cavity length, applied magnetic field and other ring laser parameters was carried out on a model based on the previous works [1, 5]. However, the beat frequency of co-directed waves has not been previously considered as a signal source for the perimeter control system. As a result of modelling, the optimal cavity bend angle and other parameters were determined that provide the greatest steepness of the adjustment characteristic for the cavity length stabilization.

The first type of used four-frequency Zeeman laser gyroscope [2] had a cavity in the form of a regular tetrahedron with a length of 28 cm. An angle of incidence on the output mirror being 30º leads to high ellipticity of the output beam. Fully optical splitting of beat signals of orthogonal polarizations is still possible through using a compensating dielectric coating and a polarizing beamsplitter. This, however, makes the output optics very complex and requires high precision manufacturing and assembling. Even minor errors in dielectric coatings and geometry of optical elements can lead to signal crosstalk, with the source of the noise being hard to pinpoint. Therefore, we carried out research aimed at separating beat signals of counterpropagating waves of orthogonal polarization through processing in analog electronic circuits

without polarization-based output optics. To separate the beat signals of waves of orthogonal polarization we utilized the signals from two photodetectors with a phase shift of 90 degrees. The separation of signals is possible due to interference fringes for waves of orthogonal polarization moving in opposite directions. After further phase shifting and summation in analog electronic circuits the beat signal of counterpropagating waves can be successfully isolated for each polarization. The chosen cavity geometry leads to a non-equidistant frequency spectrum for eigenmodes of left- and right-handed circular polarization with the gyroscope being able to operate with larger (~606.5 MHz) or smaller (~462.5 Mhz) wave frequency splitting.

The second type of used four-frequency laser gyroscope had a cavity in the form of a isohedral tetrahedron with a length of 20 cm. To ensure four-frequency lasing a ~53/47% mixture of neon isotopes ^{20}Ne and ^{22}Ne was used in the gain medium This gyroscope offers plane of incidence rotation by 22.5° on each mirror and has a high frequency splitting for orthogonal eigenmodes (~756.365 MHz), which unlike the first type of gyroscope allows stable operation both in four-frequency and two-frequency regimes with appropriate adjustment of cavity length. The operating point corresponding to the lasing frequencies of modes of left and right circular polarization relative to the gain line was initially set based on the equality of lasing powers of waves of left and right circular polarization. Then fine tuning of the cavity length took place with lasing frequencies of orthogonal modes being set so that their frequency biases were equal (Fig. 5). This makes the sensitivity to the effect of a magnetic field on the frequency splitting of orthogonally

polarized radiation identical in magnitude and different in sign. For the two-frequency lasing the cavity length was adjusted to the maximum frequency of the bias. As the angle of incidence on the output mirror for this gyroscope is ~43.866°, the output radiation has even higher ellipticity, which further complicates fully optical beat signals splitting and encourages mathematical signal processing in electronic units. Still, a conventional non-polarizing optical mixer was used to obtain the mixed beat signals of counterpropagating waves.

II. Separating Beat Signals of Counterpropagating Waves of Orthgonal Polarization of a Zeeman Laser Gyroscopes by an Electronic Method

Operating a Zeeman laser gyroscope with a two-isotopic gain medium with an appropriate choice of the angle of rotation of the beam during cavity round-trip allows for simultaneous lasing of two orthogonal polarizations. That requires separating the beat signals of counterpropagating waves of orthogonal polarizations. Beat signals come from two photodetectors with a phase shift of $\pi/2$ between each other, and signal separation of orthogonal polarizations is performed according to the mathematical description (1-17) and circuitry on Fig. 1. In addition to in-phase cos(x) signals, signals shifted by $\pi/2$ – sin(x) are also obtained. To isolate the signals, amplifiers, phase-shifting electronic circuits, an adder and subtractor are used (Fig. 1). Signal separation (Fig. 1) was performed with a constant frequency bias up to ~30 kHz. We also considered increasing the frequency bias to one and half megahertz and working with an alternating bias.

Fig. 1. Block diagram of separation of beat signals of counterpropagating waves of a Zeeman laser gyroscope according to orthogonal polarizations.

The beat signals of counterpropagating waves of orthogonal polarization can be written as:

$$Sa = A_a\cos(\omega_a t - \pi/2), \quad Sb = A_b\cos(\omega_b t + \pi/2) \quad (1)$$

- the original signals of each of the orthogonal polarizations. Each of the photodetectors records the sum of the beat signals of counterpropagating waves of two polarizations:

$$Sab1 = Sa + Sb, \quad Sab2 = Sa(+\pi/2) + Sb(-\pi/2) \quad (2)$$

At the output of the first photodetector:

$$Sab1 = A_a\cos(\omega_a t) + A_b\cos(\omega_b t) \quad (3)$$

At the output of the second photodetector:

$$Sab2 = A_a\cos(\omega_a t + \pi/2) + A_b\cos(\omega_b t - \pi/2) =$$
$$= A_a\sin(\omega_a t) - A_b\sin(\omega_b t) \quad (4)$$

After a phase shift of $+\pi/2$ by electronics of the second photodetector:

$$Sab2_{\pi/2} = A_a\cos(\omega_a t + \pi) + A_b\cos(\omega_b t) =$$
$$= -A_a\cos(\omega_a t) + A_b\cos(\omega_b t) \quad (5)$$

By summing the signal of the first photodetector and phase-shifted signal of the second photodetector:

$$Sab1 + Sab2_{\pi/2} =$$
$$= (A_a\cos(\omega_a t) + A_b\cos(\omega_b t)) + (-A_a\cos(\omega_a t) + A_b\cos(\omega_b t)) =$$
$$= 2A_b\cos(\omega_b t) \quad (6)$$

One can obtain the original signal of one polarization:

$$Bb = Sab1 + Sab2_{\pi/2} = 2A_b\cos(\omega_b t). \quad (7)$$

By subtracting from the signal of the first photodetector the phase-shifted signal of the second photodetector:

$$
\begin{aligned}
&\mathrm{Sab1} - \mathrm{Sab2}_{\pi/2} = \\
&= (\mathrm{A}_a\cos(\omega_a t) + \mathrm{A}_b\cos(\omega_b t)) - (-\mathrm{A}_a\cos(\omega_a t) + \mathrm{A}_b\cos(\omega_b t)) = \\
&\qquad\qquad = 2\mathrm{A}_a\cos(\omega_a t)
\end{aligned} \tag{8}
$$

One can obtain the original signal of orthogonal polarization:

$$
\mathrm{Aa} = \mathrm{Sab1} - \mathrm{Sab2}\pi/2 = 2\mathrm{A}_a\cos(\omega_a t) . \tag{9}
$$

For the signal of the first photodetector (3):

$$
\mathrm{Sab1} = \mathrm{A}_a\cos(\omega_a t) + \mathrm{A}_b\cos(\omega_b t) . \tag{10}
$$

After a phase shift of $+\pi/2$ by electronics of the first photodetector:

$$
\begin{aligned}
\mathrm{Sab1} &= \\
&= \mathrm{A}_a\cos(\omega_a t + \pi/2) + \mathrm{A}_b\cos(\omega_b t + \pi/2) = \\
&= -\mathrm{A}_a\sin(\omega_a t) - \mathrm{A}_b\sin(\omega_b t)
\end{aligned} \tag{11}
$$

By summing the phase-shifted signal of the first photodetector Sab1 (11) and the signal of second photodetector Sab2 (4):

$$
\begin{aligned}
&\mathrm{Sab1}_{\pi/2} + \mathrm{Sab2} = \\
&= (-\mathrm{A}_a\sin(\omega_a t) - \mathrm{A}_b\sin(\omega_b t)) + (\mathrm{A}_a\sin(\omega_a t) - \mathrm{A}_b\sin(\omega_b t)) = \\
&\qquad\qquad = -2\mathrm{A}_b\sin(\omega_b t)
\end{aligned} \tag{12}
$$

By subtracting from the phase-shifted signal of the first photodetector Sab1 (11) the signal of the photodetector Sab2 (4):

$$
\begin{aligned}
&\mathrm{Sab1}_{\pi/2} - \mathrm{Sab2} = \\
&= (-\mathrm{A}_a\sin(\omega_a t) - \mathrm{A}_b\sin(\omega_b t)) - (\mathrm{A}_a\sin(\omega_a t) - \mathrm{A}_b\sin(\omega_b t)) = \\
&\qquad\qquad = -2\mathrm{A}_a\sin(\omega_a t)
\end{aligned} \tag{13}
$$

Thus, we can obtain two source signals separated by polarization state and two source signals with $\pi/2$ phase shifts: (7), (9), (12), (13).

Source signals:

$$
\mathrm{Sab1} + \mathrm{Sab2}_{\pi/2} = 2\mathrm{A}_b\cos(\omega_b t) , \tag{14}
$$

$$
\mathrm{Sab1} - \mathrm{Sab2}_{\pi/2} = 2\mathrm{A}_a\cos(\omega_a t) . \tag{15}
$$

Source signals with $\pi/2$ phase shifts:

$$
\mathrm{Sab1}_{\pi/2} + \mathrm{Sab2} = -2\mathrm{A}_b\sin(\omega_b t) , \tag{16}
$$

$$
\mathrm{Sab1}_{\pi/2} - \mathrm{Sab2} = -2\mathrm{A}_a\sin(\omega_a t) . \tag{17}
$$

The obtained signals phase-shifted by $\pi/2$ allow to determine the direction of rotation if the polarization is known in advance, or vice versa, knowing the direction of rotation, the polarization can be determined.

III. Cavity Length Stabilization by the Beat Signals of Co-directed Waves

One of the main methods for stabilizing the cavity length in a four-frequency Zeeman laser gyroscope is equalizing the frequency biases, which simultaneously gives the same magnetic sensitivity. The sum of frequencies of Zeeman beats of beams of orthogonal polarizations [1, 5] can be written as: $f_1 + f_2 = \mathrm{K}\xi$, where $\mathrm{K} = \Delta\nu\eta\mu\delta\,K(\mu,\delta)$, ξ – frequency detuning, $\Delta\nu$ – resonator bandwidth, η – gain exceeding losses, μ – Zeeman splitting of sublevels, δ – intermode frequency

interval, $K(\mu,\delta)$ – magnetic detuning sensitivity function. When controlling the cavity length using signals of frequencies of modes of orthogonal polarization, due to the dependence of their frequency on the change in the refractive index according to the dispersion curve [4], and without taking into account Zeeman splitting, one can write for mode frequencies: $f(\xi) = mc/Ln(\xi)$, where $n(\xi) = n_0 - 2s\xi e^{-\xi^2}$, $m = L/\lambda$, $\xi = (f - f_0)/\Delta F$, where L – cavity length, ΔF – gain bandwidth, λ - wavelength, $n(\xi)$ – change in the refractive index due to anomalous dispersion of the gain medium, f – 'cold' resonator frequency. The frequency difference $f = f1 - f2$ (beat frequency) between the modes of left and right circular polarization, each of which $f1$ and $f2$ changes according to the given expression is presented in Fig. 2. This dependence has a maximum shown in Fig. 2 corresponding to the minimum in Fig. 4. It reflects the state of modes of orthogonal polarization being symmetrical relative to the center of the gain curve.

Fig. 2. Dependence of lasing frequency change of a mode of one circular polarization f1, of orthogonal polarization f2 and of their frequency difference on the cavity length detuning (frequency detuning) f=f1 – f2.

IV. Experimental Setup

An experimental setup (Fig. 3) included four-frequency Zeeman laser gyroscopes with two different bend angles, gyroscope control and power supply units, radio-electronic circuits of cavity length control (CLC), computers with an ADC and DAC board and measuring equipment such as oscilloscopes, multimeters, radio spectrum analyzer. To stabilize the cavity length (Fig. 3) we used two methods, one based on measuring the frequency of an alternating bias for each of the orthogonally polarized modes and equalizing those biases through cavity length tuning (Fig. 5). The values of the biases for orthogonal polarizations were obtained either through a polarizing beam-combining optics or by non-polarizing optics and mathematical signal processing in an analog electronic unit. The other method for stabilizing the cavity length does not require beat signals of counterpropagating waves and beam-combining optics. It is based on measuring the beat frequency of co-directed waves of orthogonal polarization (~462.5 MHz, ~606.5 MHz or ~756.365 MHz) and tunes the cavity length according to the maximum beat frequency. A high-speed avalanche photodetector with an amplifier is installed behind one of the cavity mirrors and its signal is lowered to ~10.7 MHz through heterodyning with further processing and calculating the feedback signal for the piezomotors (Fig. 3). Due to the

frequency shift in the radio mixer, the maximum frequency can be converted into a minimum (Figs. 2, 4).

Fig. 3. Experimental setup.

a.

b.

Fig. 4. Dependences of the beat signal of orthogonal polarizations on the cavity frequency detuning for two laser gyroscopes (a – cavity length of 20 cm, two piezo motors, b – cavity length of 28 cm, one piezo motor).

Fig. 5. Dependence of the dither of the beats of orthogonal polarizations Asin and Bsin on lasing frequency tuning by piezo motors. The signal separation was performed by an electronic system.

Fig. 6. Allan variance of a Zeeman four-frequency laser gyroscope in case of electronically separated beat signals of orthogonal polarizations.

Using signal separation via mathematical processing in an analog electronic unit, we have obtained data for the gyroscope and built the Allan variance (Fig. 6), showing that under certain conditions mathematical processing outperforms fully optical signal splitting.

V. CONCLUSION

Overall, the research resulted in creating a system for separating beat signals of counterpropagating waves of orthogonal circular polarizations based on mathematical processing of combined beat signals after a non-polarizing beam combining optics. Insignificant channel crosstalk can be achieved at bias frequency that is significantly higher than with optical separation at the same angular velocities. To stabilize the cavity length of a four-frequency laser gyroscope and mathematically separate beat signals a system based on maximizing the beat frequency of co-directed waves of orthogonal polarization is found to be best suited.

REFERENCES

[1] Nazarenko, M.M., Saveliev, I.I., Skulachenko, S.S., Khromykh, A.M., Yudin, I.I. Investigation of Zeeman beats in a two-mode traveling-wave laser, Soviet Journal of Quantum Electronics (1977), 7(8): 984.

[2] Yu. Yu. Broslavets et al., "New Four-Frequency Zeeman Laser Gyroscope with a Nonplanar Symmetric Cavity, its Parameters and Operation", 2021 28th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), 2021, pp. 1-8, doi: 10.23919/ICINS43216.2021.9470859.

[3] Azarova, V.V., Golyaev, Yu.D., Dmitriev, V.G. Ring gas lasers with magneto-optical control for laser gyroscopy // Quantum Electronics, Volume 30, Issue 2, 2000, Pages 96-104. doi: 10.1070/QE2000v030n02ABEH001668.

[4] Kudravtsev, A.S., Savchenko, N.A., Saveliev, I.I., "Studying Nonreciprocal Device Based on Permanent Magnet for Zeeman Laser Gyroscope", Trudy MAI, 2020, no 111. DOI: 10.34759/trd-2020-111-18.

[5] Nazarenko, M.M., Saveliev, I.I., Skulachenko, S.S., Khromykh, A.M., Yudin, I.I. Interaction of modes characterized by orthogonal circular polarizations in a ring Zeeman laser, Soviet Journal of Quantum Electronics (1979), 9(8): 1000.

Measurement Method, Statistics and Sources of the Lock-In Zone in a Zeeman Laser Gyroscope

I.N. Khokhlov

POLYUS Research Institute
of M.F. Stelmakh
Joint Stock Company POLYUS
Research Institute of M.F. Stelmakh
Moscow, Russia
https://orcid.org/0000-0002-2676-4880

Y.D. Golyaev

POLYUS Research Institute
of M.F. Stelmakh
Joint Stock Company POLYUS
Research Institute of M.F. Stelmakh
Moscow, Russia
bereg@niipolyus.ru

E.A. Petrukhin

POLYUS Research Institute
of M.F. Stelmakh
Joint Stock Company POLYUS
Research Institute of M.F. Stelmakh
Moscow, Russia
petruhin53@mail.ru

Abstract— **A method for measuring the size of the lock-in zone of a laser angular velocity sensor is proposed. The method is based on the use of measurement data of the nonlinearity of the scale factor, and allows you to control the size of the capture zone at all stages of production.**

Keywords—laser gyroscope, lock-in zone, frequency stand, scale factor, output characteristic.

I. INTRODUCTION

One of the main error sources in laser gyroscopes (LG) is the backscattering of light on the inhomogeneities of the mirrors of the ring resonator. This phenomenon leads to synchronization of the frequencies of counterpropagating waves at low angular velocities [1].

The use of an alternating frequency stand makes it possible to significantly reduce the nonlinear distortion of the output characteristic and not reduce the sensitivity of the LG at low rotational speeds [2, 3].

The Zeeman LG [4, 5] mainly uses a rectangular stand. The main distinguishing feature of such a stand is significant nonlinear distortions of the output characteristic of the LG at a rotation speed close to the amplitude of the stand. The nature of these distortions is determined by the value of the lock-in threshold (Ω_L).

When the rotation speed coincides with the amplitude of the stand ($\Omega \approx \Omega_0$), in one of the half-periods of the stand, the LG is in a state of synchronization of the frequencies of counterpropagating waves. Dynamic zones of parametric resonance, discovered relatively recently, are also present near the amplitude of the stand [6]. The zone of significant nonlinear distortions of the output characteristic of the Zeeman LG is approximately 2°/sec.

II. THE THEORETICAL PART

In [7], it was shown that the nonlinear distortions (ΔK) of the output characteristic of an LG with a rectangular stand are adequately described by the ratio (1) obtained in the mode of weak coupling of counterpropagating waves [8] (when $|\Omega - \Omega_0| >> \Omega_L$):

$$\Delta K(\Omega) = 0.5(S_+^2 / A + S_-^2 \cdot B / C / D)10^6 \text{ [ppm]} \quad (1)$$

Where:

- $A = \Omega_0^2 - \Omega^2$ [°/sec];
- $B = \Omega_g^2 + \Omega^2 - \Omega_0^2$ [°/sec];
- $C = (\Omega - \Omega_0)^2 + \Omega_g^2$ [°/sec];
- $D = (\Omega + \Omega_0)^2 + \Omega_g^2$ [°/sec].

In the ratio (1), the value (Ω_g) is called the strength of the limit cycle and characterizes the inertia of the ring laser active medium and depends on the bandwidth of the optical resonator and the excess of the pump level (α) over the loss level (δ). The parameters (S_+ and S_-) are the following combinations of the coupling coefficients counterpropagating waves due to backscattering:

$$S_\pm = K(r_{cw}^2 + r_{ccw}^2 \pm 2\, r_{cw} r_{ccw} \cos(\varphi_{cw} + \varphi_{ccw}))^{1/2} \text{ [°/sec]} \quad (2)$$

Where:

K – scale factor LG;

$r_{cw, ccw}$ – parts of the field of self-oscillation scattered in the opposite direction;

$\varphi_{cw, ccw}$ – phase shifts that occur during backscattering from the mirrors of an ring resonator when waves propagate clockwise (CW) and counterclockwise (CCW).

The parameters (S_+ и S_-) represent the dissipative and conservative components of backscattering. They differ in the magnitude of the phase shift. The value of the total phase shift ($\varphi = \varphi_{cw} + \varphi_{ccw}$) is: 2π for dissipative and π for conservative components. The value of the LG lock-in threshold is determined by the dissipative component (S_+).

It was shown in [9] that the ratio of the average values of the parameters ($<S_->$ / $<S_+>$) for an array of LG sensors assembled from mirrors of approximately the same quality ranges from 3 to 7. In this case, the average value of the parameter (S_-) is determined by the integral scattering coefficient of the mirrors. It is also noted that there is no correlation between the components of backscattering.

In Fig. 1, the ratio (1) is represented graphically, at S_+ = 0.1; S_- = 0.8; Ω_0 = 30; Ω_g = 65; expressed in dimension [°/sec].

As can be seen from Fig. 1, a sharp surge in the nonlinearity of the output characteristic is observed at an angular velocity ($\Omega = 30$°/sec), equal to the amplitude of the frequency stand (Ω_0). Its value is characterized by the parameter (S_+). The value of the positive component of the output characteristic is mainly determined by the parameter (S_-). The output characteristic zone is indicated in gray, which is not described by the ratio (1). In the output characteristics of real devices, the width of this zone is about 2°/sec.

Fig. 1. A typical view of the output characteristics of a laser gyroscope.

In Fig. 2, two types of output characteristics are presented. In the first case, the parameter (S_-) is zero, and in the second case, the parameter (S_+) is zero (both cases are excluded in real devices). Thus, the first term of the ratio (1) is graphically shown at number 1, and the second at number 2. Their summation will give the output characteristic of the real device shown in Fig. 1.

Fig. 2. A special view of the output characteristic of a laser gyroscope; type 1: $S_- = 0$; type 2: $S_+ = 0$.

As can be seen from Fig. 2, the relative deviation of the scale factor (ΔK) at angular velocity ($\Omega = 90°$/sec) is determined mainly by the value of the parameter (S_-).

This paper presents an analysis of the statistics of the values of the lock-in threshold (K-5 type sensors) obtained by a method based on measurements of the nonlinearity of the scale factor.

III. THE EXPERIMENTAL PART

Measurements of the LG scale factor were carried out on a single-axis test bench with a vertical axis of rotation. Next, the nonlinear corrections ($\Delta K(\Omega)$) of the scale factor for each value of the angular velocity were calculated using the formula:

$$\Delta K(\Omega) = (K(\Omega) \ / \ K(180) - 1)10^6 \ \text{[ppm]} \qquad (3)$$

Where:

$K(180)$ – is the scale factor measured at an angular velocity of 180°/sec. At such an angular velocity, the output characteristic has practically no distortion, which makes it possible to accept $K(180)$ as a reference value;

$K(\Omega)$ – the scale factor measured at a certain angular velocity.

The angular velocities were selected in such a way as to obtain a detailed dependence of the nonlinearity of the scale factor on the angular velocity not exceeding the amplitude of the frequency stand (Fig. 1).

Approximation of the obtained dependence using formula (1) made it possible to calculate the dissipative (S_+) and conservative (S_-) components of backscattering, which is discussed in detail in [7, 10-12].

Thus, arrays of experimental data were obtained containing parameters (S_+), (S_-) and nonlinear corrections of the scale factor ($\Delta K(\Omega)$) at various angular velocities.

Measurements of the scale factor of K-5 type sensors and gyroscopes based on them are carried out at three angular velocities: 18, 90 and 180°/sec. As shown earlier (Fig. 2), at an angular velocity of 90°/sec, the nonlinear correction of the scale factor ($\Delta K(90)$) is determined by the parameter (S_-). Statistical processing of experimental data arrays made it possible to establish the following functional relationship between these two parameters:

$$S_- = 0.0215\Delta K(90) + 0.353 \ \text{[°/sec]} \qquad (4)$$

The nonlinear correction at an angular velocity of 18°/sec ($\Delta K(18)$) is determined by the values of the parameters (S_+ and S_-), as shown in Fig. 2. Thus, replacing the parameter (S_-) with formula (4) in formula (1) and converting it allows us to calculate the parameter (S_+) or the value of the lock-in threshold (Ω_L) through the following ratio:

$$S_+ = \Omega_L = (A[2\Delta K(18) - (0.0215\Delta K(90) + 0.353)^2 \cdot \\ \cdot B \ / \ C \ / \ D])^{1/2} \ \text{[°/sec]} \qquad (5)$$

IV. DISCUSSION OF THE RESULTS

The number of sensors tested was 175. Taking into account that the tests are carried out on four adjacent modes of each sensor, the total number of measured values of the lock-in threshold (Ω_L) was 700 pieces.

The measured values of the lock-in threshold (Ω_L) were converted from the dimension [°/sec] to [Hz]. A histogram of their distribution in 5 Hz increments is shown in Fig. 3.

Fig. 3. Histogram of the distribution of the values of the lock-in threshold.

The average value of the lock-in threshold was 177 Hz. The Rayleigh distribution density is also presented here, which describes the resulting field of an array of coherent oscillators with a random phase value [13]. The probability density of this distribution is described by the following ratio:

$$f(\Omega_L) = \Omega_L / \sigma^2 exp(-\Omega_L^2 / 2 / \sigma^2) \qquad (6)$$

Where:

Ω_L – random parameter (lock-in zone threshold);

σ – a scale parameter approximately equal to the average value of the lock-in zone threshold.

A comparison of the histogram of the measured values of the lock-in threshold with the Rayleigh distribution density revealed their noticeable difference. Two reasons for this difference can be pointed out.

The first is the contribution to the backscattering of sources that do not obey the Rayleigh distribution. These include backscattering from LG photodetectors and the presence of large (submicron-sized) defects in the working area of mirrors. This difference can also be explained by the backscattering of the aperture of the selective diaphragm.

The second, more significant (in our opinion) reason is the limitation that arises when using the weak coupling mode to describe nonlinear corrections of the LG scale factor with a rectangular stand.

As shown in [7], the strength of the limit cycle (Ω_g) in the ratio (1) is significantly lower than the value calculated based on the model of cubic nonlinearity of the polarizability of the active medium.

To resolve this contradiction, instead of dependence ($\Omega_g \propto \alpha / \delta - 1$), dependence ($\Omega_g \propto 1 - \delta / \alpha$) was introduced into the initial ratio (1). With such a replacement, the ratio (1) more or less adequately describes the output characteristic of LG.

V. Conclusions

This paper describes a method for measuring the lock-in threshold value of LG with a rectangular stand. The method is based on the analysis of nonlinear scale factor corrections. The results of statistical data processing are presented.

The use of this method in the technological cycle of testing laser sensors (and gyroscopes based on them) will significantly increase awareness of the state of technological processes in production, and somewhat reduce the time and complexity of testing.

It is noted that for the correct use of the LG frequency response model obtained in the weak coupling mode, it is necessary to change the type of functional dependence of the strength of the limit cycle.

It is also relevant to use this method at the testing stage of inertial navigation units in order to clarify the causes of possible instrument failures due to exceeding the permissible nonlinearity of the scale factor.

References

[1] Aronowitz F., in Laser Applications. Ed. by M. Ross (New York: Academic, 1971) pp 133–200.

[2] Aronowitz, F., Optical Gyros and their Applications, RTO AGARDograph, 1999, 339, p.3.1.

[3] Klimontovich Yu.L. (Ed.) Volnovye i fluktatsionnye protsessy v lazerakh (Wave and Fluctuation Processes in Lasers) (Moscow: Nauka, 1974).

[4] Azarova V.V., Golyaev Yu.D., Dmitriev V.G. Quantum Electron., 30, 96 (2010) [Kvantovaya Elektron., 30, 96 (2010)].

[5] Azarova V.V., Golyaev Yu.D., Saveliev I.I. Quantum Electron., 45, 171 (2015) [Kvantovaya Elektron., 45, 171 (2015)].

[6] Gorshkov V.N., Grushin M.E., Lariontsev E.G., Saveliev I.I., Khokhlov N.I. Quantum Electron., 46, 1061 (2016) [Kvantovaya Elektron., 46, 1061 (2016)].

[7] Beketov, S.E., Bessonov, A.S., Petrukhin, E.A., Khokhlov, I.N., Khokhlov, N.I. Quantum Electron., 49, 1059 (2019) [Kvantovaya Elektron., 49, 1059 (2019)].

[8] Khoshev I.M. Cand. Diss. (Moscow: Lomonosov Moscow State University, 1978).

[9] Petrukhin E.A., Khokhlov I.N., Khokhlov N.I. Quantum Electron., 51, 351 (2021) [Kvantovaya Elektron., 51, 351 (2021)].

[10] Khokhlov I.N., Sinelnikov A.O. Comparative analysis of methods for measuring lock-in zone in zeeman laser angular velocity sensors // Physical education in universities, 2021, vol. 27, no.S4, pp. 76-80. DOI: 10.54965/16093143_2021_27_S4_55

[11] Khokhlov I., Sinelnikov A. A Method for Measuring the Lock-In Zone in Laser Gyro Sensors, 2021 28th St. Petersburg International Conference on Integrated Navigation Systems (ICINS), Saint Petersburg, Russia, 2021, pp. 1-3, doi: 10.23919/ICINS43216.2021.9470854.

[12] Khokhlov I.N., Sinelnikov A.O., Fetisova N.E. Scale Factor Correction Model for Zeeman Laser Gyroscopes, 2022 29th St. Petersburg International Conference on Integrated Navigation Systems (ICINS), Saint Petersburg, Russian Federation, 2022, pp. 1-4. doi: 10.23919/ICINS51784.2022.9815382

[13] Rayleigh, Lord. Phylos. Mag., 10, 73 (1880).

Simulation of Initial Error Calibration of Uniaxial Rotating Laser Inertial Navigation System Assisted by Total Station

Shaojun Du
Beijing Aerospace Times Laser
navigation Technology Co., Ltd
Beijing Aerospace Institute of Control
devices
Beijing,China
dushaojunjg@sina.com

Lili
Beijing Aerospace Times Laser
navigation Technology Co., Ltd
Beijing Aerospace Institute of Control
devices
Beijing,China
liliii1683@sina.com

Lin Zhang
Beijing Aerospace Times Laser
navigation Technology Co., Ltd
Beijing Aerospace Institute of Control
devices
Beijing,China
zhanglinmail126@126.com

Jiangang Guo
Beijing Aerospace Times Laser
navigation Technology Co., Ltd
Beijing Aerospace Institute of Control
devices
Beijing,China
jiangang_tiger@126.com

Lei Wang
Beijing Aerospace Times Laser
navigation Technology Co., Ltd
Beijing Aerospace Institute of Control
devices
Beijing,China
wangleinuaa@163.com

Shitao Huang
Beijing Aerospace Times Laser
navigation Technology Co., Ltd
Beijing Aerospace Institute of Control
devices
Beijing,China
huangshitao11@163.com

Abstract—**For the extremely slow and extremely low dynamic error estimation situation of shield tunneling machine, this paper presents the scheme of shield machine guidance system with uniaxial rotating laser inertial navigation system as the core. By rotating the IMU itself, the observability of the installation error between inertial navigation system and shield tunneling machine is greatly improved. At the same time, this paper proposes the Kalman filter estimation scheme for the error estimation assisted by the total station system, and the system state space model is also derived. The effects of the IMU rotation, the total station pitch and the rod arm between the inertial navigation system and the observation points on the estimation observability and accuracy are analyzed in detail, the optimal scheme "uniaxial rotational excitation + reducing oblique pitch + increasing rod arm" of error estimation is proposed, which has a strong guiding value for practical engineering.**

Keywords—Uniaxial rotation, observability, oblique pitch, rod arm, estimation accuracy.

I. Introduction

In recent years, applying inertial navigation technology to shield machine guidance system has become a research hotspot. In order to obtain the high-precision attitude information of the shield machine, the most important research direction is to suppress the initial alignment error and the navigation attitude drift. For the convenience of maintenance, the inertial navigation system is generally installed on the shield machine through the bracket. In general, the frame coordinate system of the bracket and the coordinate system of the shield machine do not coincide, and there is an installation error bias Angle, so accurate measurement and compensation are needed to improve the measurement accuracy of the attitude of the shield machine. There are generally two ways to measure the installation deflection Angle between inertial navigation system and shield machine: manual direct measurement method and optimal estimation method.

Manual direct measurement method is the main method to determine the initial attitude of shield machine at present, which mainly includes artificial methods[1] such as horizontal ruler method, side wall method, partial middle method, three-point method, cutting plane method or fitting circle method.

Assuming that the coordinate frame of shield machine is defined as b' system, that of inertial navigation system is defined as b system, and that of navigation system is defined as n system, the attitude matrix $C_{b'}^{n}$ is obtained through direct measurement, and C_b^n is obtained through inertial navigation measurement, thus obtaining the installation bias Angle $C_b^{b'}$ between inertial navigation system and shield machine :

$$C_b^{b'} = (C_{b'}^n)^{-1} C_b^n$$

Manual direct measurement method has many disadvantages, for example, the horizontal ruler method is only applicable to the active articulated type shield machine, and requires less shield tail parts and enough space to place the horizontal ruler, the applicable environment is limited by the construction of shield mechanism. Although the side wall method has wider range than that of the horizontal ruler method, the operation process is complicated. The operation of the three points is simple and fast, but in the operation process, this method should be strictly controlled at the same mileage, and the deformation of the shield machine shell should be small. The fitting circle method is not affected by the construction of shield mechanism, and is widely used in the attitude measurement of shield machine and forming tunnel detection. However, when the shield machine is used for a long time, the shell deformation is large or the measurement data is less, the measurement effect will change significantly. Therefore, more scientific methods of initial attitude determination need to be sought.

Since laser target is the mainstream method of attitude measurement of shield machine, the optimal estimation method of the installation Angle between the inertial navigation system and shield machine has not been discussed in the literature, and the estimation measurement of the installation Angle between inertial navigation system and star sensor, odometer and DVL are performed on literature review.

Liqiang Luo[2] uses speed and position information provided by GPS as observation, the odometer scale coefficient and installation error of inertial navigation system are quickly and accurately estimated in about 50 seconds by

using recursive least squares. Xingpeng Da[3]bases on the expanded Kalman filter algorithm, an iterative calibration method of odometer system error and lidar installation error is proposed, which can make the calibration of the above error in real time while positioning. Xin Wang[4] designs a sensor installation error ground rapid calibration method, which uses the inertial navigation system attitude, position and star sensor attitude to construct the observation measurement, by two axial maneuver of the carrier, the triaxial installation error of the star sensor can be quickly estimated by Kalman filter. Jinwu Tong[5] designs a USBL installation error estimation algorithm based on two-vector reconstruction,with which the position calculation of the target responder and USBL installation error angle estimation can be completed in a circle. The operation process is simple and easy to implement, and has important engineering application value. Wang Li[6] presents a fast-high-precision installation error estimation method on Levenberg-Marquart for INS/CNS, which transfers the observation vector at different time to the same frame at the same time, constructs the observation vector error of multi-stars observed and least squares objective function of the navigation star vector , and uses the adaptive-step L-M iteration algorithm to solve the variation of the system installation matrix in real time. Jingke An[7] designs the dynamic identification method based on Kalman filter, which can quickly calibrate the installation error angle of the star sensor online, and the calibration value of the installation error angle can reach more than 85% of the actual error value, and effectively improves the accuracy of the INS/CNS integration navigation system. Zhaofa Zhou[8] puts forward a calibration method of odometer scale factor and installation error based on short-time SINS motion speed information,with which the initial calibration of odometer parameters can be realized within 1 minute, without the assumption that the error value of relevant parameters is small, and the influence of rod-arm effect on the calibration effect can be ignored.

The above installation error estimation methods are all conducted in the process of carrier movement or maneuver, so the parameter estimation is highly observable. For the very slow and very low dynamic use of shield machine, if the inertial navigation system has no motion parts, the observability of the installation error is very weak, even not observable. Therefore, this paper presents the scheme of shield machine guidance system with uniaxial rotating laser inertial navigation system as the core.By rotating the IMU itself, the observability of the installation error between inertial navigation system and shield tunneling machine is greatly improved. At the same time, this paper proposes the Kalman filter estimation scheme for the error estimation assisted by the total station system, and the system state space model is also derived. The effects of the IMU rotation, the total station pitch and the rod arm between the inertial navigation system and the observation points on the estimation observability and accuracy are analyzed in detail, the optimal scheme "uniaxial rotational excitation + reducing oblique pitch + increasing rod arm" of error estimation is proposed, which has a strong guiding value for practical engineering .

II. COORDINATE SYSTEM DEFINITION AND ITS TRANSFORMATION

The main coordinate system definitions and their interrelationships are shown in Figure 1.

Fig. 1. Equipment installation relationship and the definition of the main coordinate systems

Fig. 2. total station measurement coordinate system (s system)

2.1 Definition of the coordinate system

2.1.1 Total station measurement coordinate system

This is called s system expressed as $O_sX_sY_sZ_s$, O_s is in the center of the total station, X_s axis points towards the measurement target, Y_s axis points to the left, Z_s axis is determined according to the "right-hand rule", as shown in Figure 2. When $O_sX_sY_sZ_s$ is in the horizontal plane, the Angle between O_sX_s and true north is called the initial azimuth angle of the total station, and O_sX_s is positive clockwise. The s system represented in Figure 1 is measured for INS.

2.1.2 INS coordinate system

This is called b system expressed as $O_bX_bY_bZ_b$, Solid connection with the shield machine, O_b is in the inertial navigation prism center, X_b axis points to the right of INS, Y_b axis points to the anterior side of INS, Z_b axis points to the upper side of INS.

2.1.3 Characteristic point coordinate system of shield machine

This is called b' system expressed as $O_{b'}X_{b'}Y_{b'}Z_{b'}$, Solid connection with the shield machine, $O_{b'}$ is at the shield machine feature point, $X_{b'}$ axis points to the right side of the

shield machine, $Y_{b'}$ axis points to the front side of the shield machine, $Z_{b'}$ axis points to the upper side of the shield machine, as shown in Figure 1.

2.1.4 Navigation coordinate system

This is called n system expressed as $O_nX_nY_nZ_n$, O_n coincides with O_b, the local horizontal geographic coordinate system (east-north-sky), in which position is expressed by latitude, longitude and altitude, or by X、Y or Z in ground rectangular coordinate system .

2.2 Conversion between coordinate systems

2.2.1 The relationship between the total station measurement coordinate system (s system) and the navigation coordinate system (n system)

Suppose that the initial azimuth angle of the initial total station is ϕ_0, firstly rotates ϕ around the n system O_nZ_n axis to $O_nX_{n1}Y_{n1}Z_n$; then rotates Θ around O_nY_{n1} to $O_nX_{n2}Y_{n1}Z_{n1}$; finally turns $180°$ around O_nX_{n2} , coincides with $O_sX_sY_sZ_s$, the attitude matrix C_n^s between the two coordinate systems is

$$C_n^s = T_x(\pi)T_y(\Theta)T_z(\phi_0 + \phi)$$
$$= \begin{bmatrix} cos(\phi_0 + \phi)cos\Theta & sin(\phi_0 + \phi)cos\Theta & -sin\Theta \\ sin(\phi_0 + \phi) & -cos(\phi_0 + \phi) & 0 \\ -sin\Theta cos(\phi_0 + \phi) & sin\Theta sin(\phi_0 + \phi) & -cos\Theta \end{bmatrix}$$

2.2.2 The relationship between the characteristic point coordinate system of shield machine (b' system) and coordinate system of INS (b system)

l^b represents the vector of the origin $O_{b'}$ of the characteristic point coordinate system of the shield machine in INS coordinate system (b system). $C_b^{b'}$ represents the rotation relationship between the two coordinate systems:

$$C_b^{b'} = I - \alpha \times = \begin{bmatrix} 1 & \alpha_z & -\alpha_y \\ -\alpha_z & 1 & \alpha_x \\ \alpha_y & -\alpha_x & 1 \end{bmatrix}$$

Where, $\alpha = [\alpha_x \quad \alpha_y \quad \alpha_z]^T$ represents the angular vector of the b' system from the b system, $\alpha \times$ represents the antisymmetry matrix of angle α.

2.2.3 The relationship between INS coordinate system (b system) and the navigation coordinate system (n system)

The navigation coordinate system coincides with INS coordinate system through three rotations: firstly rotates $-\Psi$ around the navigation coordinate system O_nZ_n axis , to reach the $OX_{n1}Y_{n1}Z_n$; then rotates θ around O_nY_{n1} axis, gets to $O_nX_{n2}Y_{n1}Z_{n1}$; Finally rotates γ around O_nX_{n2} axis, and gets to $O_nX_{n2}Y_{n2}Z_{n2}$, Coincides with $O_bX_bY_bZ_b$, the attitude matrix between the two coordinate systems is:

$$C_n^b = T_y(\gamma)T_x(\theta)T_z(-\Psi)$$
$$= \begin{bmatrix} cos\gamma cos\Psi + sin\gamma sin\Psi sin\theta & -cos\gamma sin\Psi + sin\gamma cos\Psi sin\theta & -sin\gamma cos\theta \\ sin\Psi cos\theta & cos\Psi cos\theta & sin\theta \\ sin\gamma cos\Psi - cos\gamma sin\Psi sin\theta & -sin\gamma sin\Psi - cos\gamma cos\Psi sin\theta & cos\gamma cos\theta \end{bmatrix}$$

2.2.4 The relationship between navigation coordinate system (n system) and platform coordinate system (n' system)

The matrix between n system and n′ system is:

$$C_n^{n'} = I - \emptyset \times = \begin{bmatrix} 1 & \emptyset_z & -\emptyset_y \\ -\emptyset_z & 1 & \emptyset_x \\ \emptyset_y & \emptyset_x & 1 \end{bmatrix},$$

Where, $\emptyset = [\emptyset_x \quad \emptyset_y \quad \emptyset_z]^T$ represents the angular vector of the n′ system from the n system, $\emptyset^n \times$ represents the antisymmetry matrix of angle $\emptyset^n$.

2.2.5 The relationship between the longitude-latitude-altitude coordinate system and navigation rectangular coordinate system

The origin location of the navigation rectangular coordinate system is P_s^n (0,0,0), whose geographical location is P_s^n (λ_0, L_0, h_0) , where any point location is assumed P (x, y, z) , it's converted to P (λ, L, h) defined in longitude-latitude-altitude coordinate system as follows:

$$P (\lambda, L, h) = AP (x, y, z) + P_s^n (\lambda_0, L_0, h_0)$$

$Where, A = diag \left(\frac{1}{R_N} \quad \frac{1}{R_M cosL_0} \quad 1 \right). R_M$、$R_N$ is the radius of curvature of the earth meridian and prime vertical circle, respectively.

IV. ANALYSIS OF THE SYSTEM STATE-SPACE MODEL

3.1 System state equation

The installation error angle between the inertial navigation system and the shield machine is constant, so

$$\dot{\alpha} = 0$$

When the traditional 15-dimensional state equation of the conventional inertial navigation system is expanded to 18 dimensions, then the equation of state of the new inertial navigation system is:

$$\dot{X} = FX + GW \qquad (1)$$

Where, X is the system state :

$$X = [\emptyset \quad \delta v \quad \delta p \quad \varepsilon_b \quad \nabla_a \quad \alpha]^T$$

$\emptyset = [\emptyset_e \quad \emptyset_n \quad \emptyset_u]^T$ Is the platform error Angle;

$\delta v = [\delta v_e \quad \delta v_n \quad \delta v_u]^T$ is speed error;

$\delta p = [\delta \lambda \quad \delta L \quad \delta h]^T$ is position error;

$\varepsilon_b = [\varepsilon_{bx} \quad \varepsilon_{by} \quad \varepsilon_{bz}]^T$ is the gyro constant value drift;

$\nabla_a = [\nabla_{ax} \quad \nabla_{ay} \quad \nabla_{az}]^T$ is the accelerator constant value drift;

$\alpha = [\alpha_x \quad \alpha_y \quad \alpha_z]^T$ is the installation error angle between the inertial navigation system and the shield machine.

F is the system state matrix:

$$F = \begin{bmatrix} F_{\emptyset\emptyset} & F_{\emptyset v} & F_{\emptyset p} & -C_b^n & 0_{3\times3} & 0_{3\times3} \\ F_{v\emptyset} & F_{vv} & F_{vp} & 0_{3\times3} & C_b^n & 0_{3\times3} \\ 0_{3\times3} & F_{pv} & F_{pp} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{9\times3} & 0_{9\times3} & 0_{9\times3} & 0_{9\times3} & 0_{9\times3} & 0_{9\times3} \end{bmatrix}$$

where,

$$F_{\emptyset v} = \begin{bmatrix} 0 & -\frac{1}{R_M+h} & 0 \\ \frac{1}{R_N+h} & 0 & 0 \\ \frac{tanL}{R_N+h} & 0 & 0 \end{bmatrix} F_{pv} = \begin{bmatrix} \frac{secL}{R_N+h} & 0 & 0 \\ 0 & \frac{1}{R_M+h} & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$F_{pp} = \begin{bmatrix} 0 & \frac{v_e tanLsecL}{R_N + h} & -\frac{v_e secL}{(R_N + h)^2} \\ 0 & 0 & -\frac{v_n}{(R_M + h)^2} \\ 0 & 0 & 0 \end{bmatrix}$$

$F_{\emptyset\emptyset}$

$$= \begin{bmatrix} 0 & \omega_{ie}sinL + \frac{v_e}{R_N + h}tanL & -(\omega_{ie}cosL + \frac{v_e}{R_N + h}) \\ -(\omega_{ie}sinL + \frac{v_e}{R_N + h}tanL) & 0 & -\frac{v_n}{R_M + h} \\ \omega_{ie}cosL + \frac{v_e}{R_N + h} & \frac{v_n}{R_M + h} & 0 \end{bmatrix}$$

$$F_{\emptyset p} = \begin{bmatrix} 0 & 0 & \dfrac{v_n}{(R_M+h)^2} \\ 0 & -\omega_{ie}sinL & -\dfrac{v_e}{(R_N+h)^2} \\ 0 & \omega_{ie}cosL + \dfrac{v_e}{R_N+h}sec^2L & -\dfrac{v_e tanL}{(R_N+h)^2} \end{bmatrix}$$

$$F_{v\emptyset} = \begin{bmatrix} 0 & -f_u & f_n \\ f_u & 0 & -f_e \\ -f_n & f_e & 0 \end{bmatrix}$$

$$F_{vv} = \begin{bmatrix} \dfrac{v_n tanL - v_u}{R_N+h} & 2\omega_{ie}sinL + \dfrac{v_e}{R_N+h}tanL & -(2\omega_{ie}cosL + \dfrac{v_e}{R_N+h}) \\ -2(\omega_{ie}sinL + \dfrac{v_e}{R_M+h}tanL) & -\dfrac{v_u}{R_M+h} & -\dfrac{v_n}{R_M+h} \\ 2(\omega_{ie}cosL + \dfrac{v_e}{R_M+h}) & \dfrac{2v_n}{R_M+h} & 0 \end{bmatrix}$$

$$F_{vp} = \begin{bmatrix} 0 & 2\omega_{ie}(v_u sinL + v_n cosL) + \dfrac{v_e v_n}{R_N+h}sec^2L & \dfrac{v_e v_u - v_e v_n tanL}{(R_N+h)^2} \\ 0 & -(2\omega_{ie}cosL + \dfrac{v_e^2}{R_N+h}sec^2L) & \dfrac{v_n v_u + v_e^2 tanL}{(R_N+h)^2} \\ 0 & -2v_e\omega_{ie}sinL & \dfrac{v_e^2+v_n^2}{(R_N+h)^2} \end{bmatrix}$$

W is the system noise:

$$W = [\omega_{gx} \quad \omega_{gy} \quad \omega_{gz} \quad \omega_{ax} \quad \omega_{ay} \quad \omega_{az}]^T$$

Where, $\omega_{gx} \quad \omega_{gy} \quad \omega_{gz}$ is the gyro drift noise, $\omega_{ax} \quad \omega_{ay} \quad \omega_{az}$ is the accelerator drift noise.

G is the system noise matrix:

$$G = \begin{bmatrix} C_b^n & 0_{3\times3} \\ 0_{3\times3} & C_b^n \\ 0_{12\times3} & 0_{12\times3} \end{bmatrix}$$

Where, C_b^n is the INS attitude matrix.

3.2 Measurement equation

As shown in Figure 1, S system is the measurement coordinate system of the total station, b system is the INS coordinate system, and b 'system is the characteristic point coordinate system of shield machine. l^b is defined in b system, indicating the vector of $O_{b'}O_b$ of the characteristic point coordinate system of shield machine in the INS coordinate system (b system). $\rho^{b'}$ is defined in the b' system, as the vector of the origin $O_sO_{b'}$ of the total station coordinate system (s system) in the b' system. ρ^s represents slant-range for the origin O_b of the INS coordinate system (b system) measured by the total station in s system.

Assuming that the total station has completed the north determination, the initial Angle between its O_sX_s axis and true north is ϕ_0, and the center position of the measuring station $P_{Q0}^n(L_0, \lambda_0, h_0)$ has been nearly determined. The initial position of INS $P_{INS0}^n(L_0, \lambda_0, h_0)$ is determined by measuring the slant-range ρ^s, elevation angle Θ_1 and azimuth angle ϕ_1 between the total station and the inertial navigation system as follows:

$$P_{INS0}^n = P_{Q0}^n + A[C_n^s(\Theta_1, \phi_1)]^{-1}\rho^{s_1}$$

where,

$$A = diag\left(\frac{1}{R_N} \quad \frac{1}{R_M cos(P_{Q0}^n(L_0))} \quad 1\right)$$

$$C_n^s(\Theta_1, \phi_1)$$
$$= \begin{bmatrix} cos(\phi_0+\phi_1)cos\Theta_1 & sin(\phi_0+\phi_1)cos\Theta_1 & -sin\Theta_1 \\ sin(\phi_0+\phi_1) & -cos(\phi_0+\phi_1) & 0 \\ -sin\Theta_1 cos(\phi_0+\phi_1) & sin\Theta_1 sin(\phi_0+\phi_1) & -cos\Theta_1 \end{bmatrix}$$

After INS is initialized with P_{INS0}^n, the initial alignment begins, and then goes into the pure inertial navigation state,

thus the real-time position $P_{INS}^n(L, \lambda, h)$ of INS can be obtained.

Through the real-time position $P_{INS}^n(L, \lambda, h)$ of INS, the rod arm l^b, the slant-range $\rho^{b'}$, and the installation error matrix $C_b^{b'}$ between the b system and the b' system, the position $\hat{P}_m^n$ of the total station can be determined as follows:

$$\hat{P}_m^n = P_{INS}^n + AC_{n'}^n C_b^{n'}(C_{b'}^b \rho^{b'} + l^b) \qquad (2)$$

where, $C_b^{n'}$ is the inertial navigation attitude matrix, $C_{b'}^b$ is the installation error matrix between INS and the shield machine, and $C_{n'}^n$ the platform error angle matrix.
In formula (2),

$$C_{n'}^n = I + \emptyset \times \qquad (3)$$
$$C_{b'}^b = I + \alpha \times \qquad (4)$$

where, $\emptyset = [\emptyset^e \quad \emptyset^n \quad \emptyset^u]^T, \alpha = [\alpha_x \quad \alpha_y \quad \alpha_z]^T$. $\emptyset$ is the platform error angle vector, and α is the installation error angle vector between INS and the shield machine.

Formula (2) is substituted by formula （3） and （4）, and simplified, then linear equation of $\hat{P}_m^n$ by ignoring the second order and above infinitesimal is obtained as follows:

$$\hat{P}_m^n = P_{INS}^n + A(I + \emptyset^n \times)C_b^{n'}[(I + \alpha \times)\rho^{b'} + l^b]$$
$$= P_{INS}^n + AC_b^{n'}(\rho^{b'} + l^b)$$
$$- \begin{bmatrix} A[(C_b^{n'}(\rho^{b'}+l^b))\times] & AC_b^{n'}(\rho^{b'}\times) \end{bmatrix}\begin{bmatrix} \emptyset^n \\ \alpha \end{bmatrix}$$

If the exact position P_m^n of the total station is got, the position measurement error v is:

$$v = \hat{P}_m^n - P_m^n$$
$$= -\begin{bmatrix} A[(C_b^{n'}(\rho^{b'}+l^b))\times] & AC_b^{n'}(\rho^{b'}\times) \end{bmatrix}\begin{bmatrix} \emptyset^n \\ \alpha \end{bmatrix} + P_{INS}^n$$
$$+ AC_b^{n'}(\rho^{b'}+l^b) - P_m^n \qquad (5)$$

Assume:

$$h = \begin{bmatrix} A[(C_b^{n'}(\rho^{b'}+l^b))\times] & AC_b^{n'}(\rho^{b'}\times) \end{bmatrix}$$
$$X = [\emptyset^n \quad \alpha]^T$$
$$z = P_{INS}^n + AC_b^{n'}(\rho^{b'}+l^b) - P_m^n$$

Then formula （5） can be simplified as follows:

$$z = hX + v \qquad (6)$$

Where, X is 6×1 dimensional column vector, z is 3×1 dimensional observation column vector, h is 3×6 dimensional matrix, and v is 3×1 dimensional observation error column vector.

According state equation (1),

$$X = [\emptyset \quad \delta v \quad \delta p \quad \varepsilon_b \quad \nabla_a \quad \alpha]^T$$

The measurement matrix h' in equation (6) can be defined as follows:

$$h' = \begin{bmatrix} A[(C_b^{n'}(\rho^{b'}+l^b))\times] & 0_{3\times12} & AC_b^{n'}(\rho^{b'}\times) \end{bmatrix}$$

So the new measurement equation is:

$$z = h'X + v \qquad (7)$$

3.3 State space model

Combining equation (1) and (7), the state space model of the INS/total station integration systems is:

$$\begin{cases} \dot{X} = FX + GW \\ z = h'X + v \end{cases} \qquad (8)$$

Where, $X \in R^{18\times1}, F \in R^{18\times18}, G \in R^{18\times6}, W \in R^{6\times1}$, $z \in R^{3\times1}, h' \in R^{3\times18}, v \in R^{3\times1}$. F is the system state matrix, G is the system noise matrix, W is the system noise state column vector, z is the measurement column vector, h' is

the measurement matrix, and v is the measurement noise column vector.

The statistical characteristics of the state vector X are:
$$\hat{X}(t_0) = E(X(0)) = m_{x_0}$$
$$P(t_0) = E(X(0)X^T(0)) = var(X(0)) = P_{x_0}$$

The statistical characteristics of the system noise W and the measurement noise v are:
$$E[W(t)] = 0，\ E[W(t)W^T(\tau)] = cov(W(t), W(\tau))$$
$$= Q(t)\delta(t - \tau)$$
$$E[V(t)] = 0，\ E[V(t)V^T(\tau)] = cov(V(t), V(\tau))$$
$$= R(t)\delta(t - \tau)$$
$$E[W(t)V^T(\tau)] = cov(W(t), V(\tau)) = 0$$

$Q(t)$ is the system noise variance intensity matrix, 6×6 dimensional non-negative definite matrix; $R(t)$ is the measurement noise variance intensity matrix, 3×3 dimensional positive definite matrix; $\delta(t - \tau)$ is the Dirac shock function as follows,
$$\delta(t - \tau) = \begin{cases} 1，\ t = \tau \\ 0，\ t \neq \tau \end{cases}$$
$$Q(t) = diag[\sigma_{gx}^2 \quad \sigma_{gy}^2 \quad \sigma_{gz}^2 \quad \sigma_{ax}^2 \quad \sigma_{ay}^2 \quad \sigma_{az}^2]$$

Where, $\sigma_{gx}^2 \quad \sigma_{gy}^2 \quad \sigma_{gz}^2$ are gyroscopes' drift noise variance and $\sigma_{ax}^2 \quad \sigma_{ay}^2 \quad \sigma_{az}^2$ are accelerators' drift noise variance.

Discretize the state-space model (8) to obtain:
$$\begin{cases} X_k = \phi_{k,k-1}X_{k-1} + \Gamma_{k-1}W_{k-1} \\ z_k = H_kX_k + V_k \end{cases}$$
（9）

Where,
$$\phi_{k,k-1} = \sum_{n=0}^{\infty} \frac{[F(t_{k-1})T_s]^n}{n!}$$
$$\Gamma_{k-1} = \left\{\sum_{n=0}^{\infty} \frac{[F(t_{k-1})T_s]^{n-1}}{n!}\right\} G(t_{k-1})T_s$$
$$Q_k = Q(t)/T_s$$
$$R_k = R(t)/T_s$$

Q_k is the scattered noise covariance matrix of the system, and R_k is the measurement noise covariance matrix, T_s is the filtering period.

3.4 Observability analysis[9]

Without considering the system noise, the discrete system (9) is simplified to:
$$\begin{cases} X_k = \phi_{k,k-1}X_{k-1} \\ z_k = H_kX_k + V_k \end{cases}$$
（10）

Where,
$$E(V_k) = 0$$
$$E(V_kV_k^T) = R_k$$
$$E(X_0) = \bar{X}_0$$
$$E[(X_0 - \bar{X}_0)(X_0 - \bar{X}_0)^T] = P_0$$

Assume:
$$\phi_{k,0} = \phi_{k,k-1}\phi_{k-1,k-2}\cdots\phi_{2,1}\phi_{1,0} = \prod_{i=1}^{i} \phi_{i,i-1}$$
$$L_{0,k} = \sum_{i=0}^{k} \phi_{k,0}^T H_k^T R_k^{-1} H_k \phi_{k,0}$$

Defining the diagonal of the matrix $(I + P_0L_{0,k})^{-1}$ is the observable index corresponding to each state variable, and the observable index of the i subcomponent in the time t_k

state variables is defined as the i diagonal element of matric $(I + P_0L_{0,k})^{-1}$, that is:
$$\sigma_k^i = \left\{\text{diag}(I + P_0L_{0,k})^{-1}\right\}_i \quad （11）$$

According to definition of the above observability, the closer $|\sigma_k^i|$ is to 0, the faster the estimation error is decreasing, the more bigger the observability of the corresponding state parameter is. On the contrary, the closer $|\sigma_k^i|$ is to 1, the slower the estimation error is decreasing, the more smaller the observability of the corresponding state parameter is.

3.5 Optimal estimation method of state variables

To the discrete system (11), the standard Kalman filter algorithm is:
$$\hat{X}_{k,k-1} = \phi_{k,k-1}\hat{X}_{k-1}$$
$$P_{k,k-1} = \phi_{k,k-1}P_{k-1}\phi_{k,k-1}^T + \Gamma_{k-1}Q_{k-1}\Gamma_{k-1}^T$$
$$K_k = P_{k,k-1}H_k^T[H_kP_{k,k-1}H_k^T + R_k]^{-1}$$
$$\hat{X}_k = \hat{X}_{k,k-1} + K_k[Z_k - H_k\hat{X}_{k,k-1}]$$
$$P_k = [I - K_kH_k]P_{k,k-1}$$

Where, $\hat{X}_k$ is the current optimal estimated state variable, $\hat{X}_{k,k-1}$ is one-step prediction of the state variable Estimation, P_k is the current optimal estimated covariance matrix, $P_{k,k-1}$ is the one step prediction covariance matrix, K_k is the optimal gain matrix, Q_k is the system noise covariance matrix, and R_k is the measurement noise covariance matrix.

$$\phi_{k,k-1} \approx I + F(t_{k-1})T_s + F^2(t_{k-1})\frac{T_s^2}{2}$$
$$\Gamma_{k-1}Q_{k-1}\Gamma_{k-1}^T \triangleq Q_{k-1}$$
$$= Q(t_{k-1})T_s$$
$$+ [F(t_{k-1})Q(t_{k-1})$$
$$+ (F(t_{k-1})Q(t_{k-1}))^T]\frac{T_s^2}{2!}$$

V. SYSTEM SIMULATION

4.1 Block diagram of the system simulation principle

The schematic block diagram of the system simulation is shown in figure 3.

Fig. 3. Schematic block diagram of the system simulation

In the block diagram, A and A′ are position transformation matrics from navigation rectangular coordinate system to the longitude-latitude-altitude coordinate system, which are expressed as follows:
$$A = A(P_b^n), A' = A(P_b^{n'})$$

For example with the latter:
$$A' = diag\left(\frac{1}{R_N} \quad \frac{1}{R_M cos[P_b^{n'}(L)]} \quad 1\right)$$

The initial position $P_{b'}^n$ and the arm of the total station l^b have been determined in advance, we can obtain:

$$\rho^{b'} = (AC_b^n)^{-1}(P_{b'}^n - P_b^n) - l^b$$

Thus measurement z and measurement matrix h' can be got:

$$z = P_b^{n'} + AC_b^{n'}((AC_b^n)^{-1}(P_{b'}^n - P_b^n) + l^b) - P_{b'}^n$$

$$h' = \left[A'\left[\left(C_b^{n'}(AC_b^n)^{-1}(P_{b'}^n - P_b^n)\right)\times\right] \quad 0_{3\times12} \quad A'C_b^{n'}(\rho^{b'}\times)\right]$$

4.2 System Simulation

The simulation conditions are set up as follows:

Initial position: $[40° \quad 116° \quad 50m]^T$;

Gyro zero bias: $0.01°/h(1\sigma)$, random walk: $0.001°/\sqrt{h}$;

Accelerator zero bias: $30\mu g(1\sigma)$, random walk: $1\mu g/\sqrt{h}$;

Initial value of the status variable:

$$\varnothing = \begin{bmatrix}-1' & 2' & 3'\end{bmatrix}^T$$

$$\delta v = [0.1m/s \quad 0.1m/s \quad 0.1m/s]^T$$

$$\delta p = [0.5m \quad 0.5m \quad 0.5m]^T$$

$$\varepsilon_b = [0.01°/h \quad 0.01°/h \quad 0.01°/h]^T$$

$$\nabla_a = [30\mu g \quad 30\mu g \quad 30\mu g]^T$$

$$\alpha = \begin{bmatrix}-11' & 9' & 20'\end{bmatrix}^T$$

Initial variance of state variable,:

$$P_0 = \left[\text{diag}(\varnothing \quad \delta v \quad \delta p \quad \varepsilon_b \quad \nabla_a \quad \alpha)\right]^2$$

System noise variance:

$$Q = \text{diag}[\text{diag}(0.001°/\sqrt{h} \quad 0.001°/\sqrt{h} \quad 0.001°/\sqrt{h} \quad 1\mu g/\sqrt{h} \quad 1\mu g/\sqrt{h} \quad 1\mu g/\sqrt{h} \quad 0_{12\times1})]^2$$

With the total station measurement error 1mm, the noise variance matric is determined as follow:

$$R = [\text{diag}(1mm \quad 1mm \quad 1mm)]^2$$

INS is installed at the lower right of the rear vertical plane of the shield machine; the characteristic point on the shield machine is selected on the upper left of INS，and the arm in INS coordinate system is $l^b = [4m \quad 0m \quad 6m]^T$; the total station is set 50m behind the shield machine，that's $\rho^{b'} = [0 \quad -50m \quad 0]^T$; the Kalman filter period is 5 seconds; the simulation time is 350 seconds.

4.2.1 Simulation of motion state effect on the error estimation

The following three states are set for comparison simulation:

1）Stationary state;

2）INS uniaxial azimuthal rotation $360°$, and the rotation angle rate $1.2°/s$;

3）INS uniaxial azimuthal rotation $180°$, rotation angle rate $1.2°/s$; then rolling rotation $90°$, rotation angle rate $0.6°/s$.

Figure 4-6 are the observability curve, initial alignment error estimation curve and installation angle estimation curve respectively. Table 1-2 are the results of observability and error estimation, estimation accuracy statistics at the time of filtering 300s,in which,p1 stands for stationary state,p2 stands for azimuth rotation,and p3 stands for "azimuth + rolling "rotation.

(1) stationary state　　(2) azimuth rotation　　(3) azimuth/rolling rotation

Fig. 4.　Observability curve

(1) stationary state　　(2) azimuth rotation　　(3) azimuth/rolling rotaion

Fig. 5.　Initial alignment error estimation curve

(1) stationary state　　(2) azimuth rotation　　(3) azimuth/rolling rotaion

Fig. 6.　Installation error angle estimation curve

TABLE 1-1 INITIAL ALIGNMENT ERROR OBSERVABILITY (300S)

item	$\varnothing_E$	$\varnothing_N$	$\varnothing_U$
P1	3.8619e-01	1.0274e-02	3.5204e-02
P2	1.6719e-03	2.2428e-03	2.0137e-04
P3	2.8405e-04	2.0395e-03	5.8146e-05

TABLE 1-2 INSTALLATION ANGLE OBSERVABILITY (300S)

item	α_x	α_y	α_z
P1	7.0026e-03	5.4746e-05	3.9609e-03
P2	2.7693e-05	1.4978e-05	2.1357e-05
P3	4.3054e-06	3.1380e-05	5.0380e-06

TABLE 2-1 INITIAL ALIGNMENT ERROR ESTIMATION RESULTS AND ACCURACY(300 SECONDS, UNIT: MIN ('))

item		$\varnothing_E$	$\varnothing_N$	$\varnothing_U$
Set value		-1	2	3
estimation value	P1	-3.9957e-01	1.7780e+00	3.6609e+00
	P2	-8.3981e-01	1.8762e+00	2.2529e+00
	P3	-1.5869e+00	1.8144e+00	3.0245e+00
estimation accuracy	P1	6.0043e-01	-2.2201e-01	6.6093e-01
	P2	1.6019e-01	-1.2377e-01	-7.4709e-01
	P3	-5.8686e-01	-1.8558e-01	2.4494e-02

TABLE 2-2 INSTALLATION ANGLE ESTIMATION RESULTS AND ACCURACY(300 SECONDS, UNIT: MIN ('))

item		α_x	α_y	α_z
Set value		-11	9	20
estimation value	P1	-1.1551e+01	9.1067e+00	1.9174e+01
	P2	-1.1069e+01	8.9981e+00	1.9111e+01
	P3	-1.0911e+01	8.7929e+00	1.9639e+01

estimation accuracy	P1	-5.5120e-01	1.0673e-01	-8.2554e-01
	P2	-6.9073e-02	-1.9391e-03	-8.8851e-01
	P3	8.8522e-02	-2.0708e-01	-3.6139e-01

The following conclusions can be drawn from Figure 1-3 and Table 1~2: IMU maneuver will improve the observability and parameters' estimation accuracy in the following two aspects:

1）observability impact analysis

Compared with stationary state, the parameter estimation observability under uniaxial azimuthal rotation decay rapidly, especially $\varnothing_U$ and $\varnothing_E$,α_x and α_z , decay by nearly 1 / 100. Compared with and uniaxial azimuth rotation, the parameter estimation observability under the azimuth and rolling sequential rotation is further improved, and the observability of $\varnothing_U$,$\varnothing_E$,α_x and α_z decay by nearly 1 / 10.

2）Analysis of parameter estimation accuracy impact

Compared with the static state, $\varnothing_E$ 、 $\varnothing_N$ 、 α_x and α_y estimation accuracy under uniaxial rotation are all improved,while that of $\varnothing_U$ and α_z remains unchanged. Compared with single axial azimuth rotation,$\varnothing_U$ and α_z estimation accuracy under the azimuth and rolling sequential rotation is significantly improved, from nearly 1' to within 30".

4.2.2 Simulation of the influence of total station position on installation error angle estimation

In the case of uniaxial rotation, the following two conditions:

1) The total station is set at 10m behind the shield machine, $\rho^{b'} = [0 \quad -10m \quad 0]^T$.

2) The total station is set at 200m behind the shield machine, $\rho^{b'} = [0 \quad -200m \quad 0]^T$.

Figure 7-9 are the observability curve, initial alignment error estimation curve and installation angle estimation curve respectively. Table 3-4 are the results of observability and error estimation, estimation accuracy statistics at the time of filtering 300s,in which,p1 stands for $\rho^{b'} = [0 \quad -10m \quad 0]^T$,p2 stands for $\rho^{b'} = [0 \quad -200m \quad 0]^T$.

（1）$\rho^{b'} = [0 \quad -10m \quad 0]^T$ （2）$\rho^{b'} = [0 \quad -200m \quad 0]^T$
Fig. 7.   Observability curve

（1）$\rho^{b'} = [0 \quad -10m \quad 0]^T$ （2）$\rho^{b'} = [0 \quad -200m \quad 0]^T$
Fig. 8.   Initial alignment error estimation curve for INS

（1）$\rho^{b'} = [0 \quad -10m \quad 0]^T$ （2）$\rho^{b'} = [0 \quad -200m \quad 0]^T$
Fig. 9.   Installation error angle estimation curve

TABLE 3-1 INITIAL ALIGNMENT ERROR OBSERVABILITY (300S)

item	$\varnothing_E$	$\varnothing_N$	$\varnothing_U$
P1	2.3020e-03	1.2968e-03	2.0380e-04
P2	1.5833e-03	5.0169e-03	1.9009e-04

TABLE 3-2 INSTALLATION ANGLE OBSERVABILITY (300S)

item	α_x	α_y	α_z
P1	2.7617e-05	1.9379e-05	2.1292e-05
P2	2.7401e-05	5.1171e-06	2.0997e-05

TABLE 4-1 INITIAL ALIGNMENT ERROR ESTIMATION RESULTS AND ACCURACY(300 SECONDS, UNIT: MIN ('))

item		$\varnothing_E$	$\varnothing_N$	$\varnothing_U$
Set value		-1	2	3
estimation value	P1	-8.1966e-01	2.0344e+00	1.7120e+00
	P2	-1.0023e+00	1.7847e+00	2.1651e+00
estimation accuracy	P1	1.8034e-01	3.4449e-02	-1.288e+00
	P2	-2.2937e-03	-2.1527e-01	-8.3491e-01

TABLE 4-2 INSTALLATION ANGLE ESTIMATION RESULTS AND ACCURACY(300 SECONDS, UNIT: MIN ('))

item		α_x	α_y	α_z
Set value		-11	9	20
estimation value	P1	-1.1029e+01	8.9952e+00	1.9684e+01
	P2	-1.1075e+01	9.0523e+00	1.9127e+01
estimation accuracy	P1	-2.9174e-02	-4.8488e-03	-3.1616e-01
	P2	-7.5364e-02	5.2289e-02	-8.7301e-01

The following conclusions can be drawn from Figure 7-9 and Table 3~4: The distance between the total station and the shield machine has a significant impact on the convergence speed and accuracy of the installation error Angle estimation, which is embodied in the following two aspects:

1) Influence analysis on observability

The closer the total station is to the shield machine, the faster the convergence of the parameter estimation observability is, but there is no significant difference from the convergent stability value.

2) Influence Analysis on estimation accuracy

The closer the total station is to the shield machine, the estimated stable value accuracy of the installation error Angle α_x、 α_y and α_z is improved in some degree, reaching nearly 30" at 10 meters' distance, while the stable value

accuracy of attitude, heading error($\varnothing_E \varnothing_N$ and $\varnothing_U$) does not change significantly.

4.2.3 Simulation of the influence of the rod arm on the parameter estimation

In the case of uniaxial azimuthal rotation and $\rho^{b'} = [0 \quad -50m \quad 0]^T$,according to the following two conditions:
1) $l^b = [0.4m \quad 0m \quad 0.6m]^T$;
2) $l^b = [4m \quad 0m \quad 6m]^T$.

Figure 10-12 are the observability curve, initial alignment error estimation curve and installation angle estimation curve respectively. Table 5 are the results of observability and error estimation, estimation accuracy statistics at the time of filtering 300s,in which,p1 stands for $l^b = [0.4m \quad 0m \quad 0.6m]^T$,p2 stands for $l^b = [4m \quad 0m \quad 6m]^T$.

（1）$l^b = [0.4m \quad 0m \quad 0.6m]^T$ （2）$l^b = [4m \quad 0m \quad 6m]^T$

Fig. 10. Observability curve

（1）$l^b = [0.4m \quad 0m \quad 0.6m]^T$ （2）$l^b = [4m \quad 0m \quad 6m]^T$

Fig. 11. Initial alignment error estimation curve

（1）$l^b = [0.4m \quad 0m \quad 0.6m]^T$ （2）$l^b = [4m \quad 0m \quad 6m]^T$

Fig. 12. Installation error angle estimation curve

TABLE 5-1 INITIAL ALIGNMENT ERROR OBSERVABILITY (300S)

item	$\varnothing_E$	$\varnothing_N$	$\varnothing_U$
P1	1.8041e-03	2.9027e-03	5.0610e-03
P2	1.6826e-03	2.5214e-03	2.0481e-04

TABLE 5-2 INSTALLATION ANGLE OBSERVABILITY (300S)

item	α_x	α_y	α_z
P1	2.8983e-05	2.9831e-05	5.5584e-04
P2	2.8889e-05	1.5103e-05	2.1899e-05

TABLE 5-3 INITIAL ALIGNMENT ERROR ESTIMATION RESULTS AND ACCURACY(300 SECONDS, UNIT: MIN ('))

item		$\varnothing_E$	$\varnothing_N$	$\varnothing_U$
Set value		-1	2	3
estimation value	P1	-1.0131e+00	1.7659e+00	7.8325e+00
	P2	-8.4034e-01	1.8751e+00	2.2672e+00
estimation accuracy	P1	-1.3098e-02	-2.3408e-01	4.8325e+00
	P2	1.5966e-01	-1.2493e-01	-7.3279e-01

TABLE 5-4 INSTALLATION ANGLE ESTIMATION RESULTS AND ACCURACY(300 SECONDS, UNIT: MIN ('))

item		α_x	α_y	α_z
Set value		-11	9	20
estimation value	P1	-1.1001e+01	8.9910e+00	1.3509e+01
	P2	-1.1064e+01	8.9988e+00	1.9094e+01
estimation accuracy	P1	-1.1421e-03	-9.0423e-03	-6.4908e+00
	P2	-6.4077e-02	-1.2467e-03	-9.0613e-01

The following conclusions can be drawn from Figure 10-12 and Table 5~6: the rod arm l^b has a significant influence on the convergence accuracy of the initial heading error and that of heading installation angle, which is reflected in the following two aspects:

1) Influence analysis on observability

The rod arm l^b has no effect on the observability of the parameter estimation.

2) Analysis of parameter estimation accuracy

The larger the arm l^b is, the higher the convergence accuracy of the initial heading error $\varnothing_U$ and the heading installation error angle α_z is.When $\rho^{b'} = [0 \quad -50m \quad 0]^T, l^b = [4m \quad 0m \quad 6m]^T$,the convergence accuracy of $\varnothing_U$ and α_z is better than 1 ', while the stable value accuracy of the horizontal attitude error ($\varnothing_E$ and $\varnothing_N$)and that of horizontal installation angle (α_x and α_y) don't change significantly.

4.2.4 Simulation Summary

According to 4.2.1~4.2.3, the effects of inertial navigation IMU maneuver, variation of $\rho^{b'}$ and l^b on parameter estimation observability and convergence accuracy are shown in Table 6,in which,p1 stands for observability ,p2 stands for convergence accuracy.

TABLE 6-1 INITIAL ALIGNMENT ERROR ESTIMATION SUMMARY

item		$\varnothing_E$	$\varnothing_N$	$\varnothing_U$
Zimuth rotation	P1	raised	raised	raised
	P2	raised	raised	no change
Zimuth + rolling rotation	P1	raised	raised	raised
	P2	raised	raised	greatly raised
decrease$\rho^{b'}$	P1	no change	no change	no change
	P2	no change	no change	no change
Increase l^b	P1	no change	no change	no change
	P2	no change	no change	raised

TABLE 6-2 INSTALLATION ANGLE ESTIMATION SUMMARY

item		α_x	α_y	α_z
Zimuth rotation	P1	raised	raised	raised
	P2	raised	raised	no change
Zimuth + rolling rotation	P1	raised	raised	raised
	P2	raised	raised	greatly raised
decrease $\rho^{b'}$	P1	no change	no change	no change
	P2	raised	raised	raised
Increase l^b	P1	no change	no change	no change
	P2	no change	no change	raised

a. Note: (1) "azimuth rotation" and "azimuth + rolling rotation" are compared with static; (2)"decrease $\rho^{b'}$" and "increasel^b" are compared with "azimuth rotation".

VI. ENGINEERING EXAMPLE ANALYSIS

It can be seen from Table 6 that the best combination to achieve the high-precision estimation compensation of the initial error of INS is "zimuth rotation + rolling rotation + decrease $\rho^{b'}$ +increase l^b ", but only combination of "decrease $\rho^{b'}$ +increase l^b " for the low-cost uniaxial rotation INS. Because "increasing l^b " improves the estimation accuracy of the Zimuth error angle only ("raised" in Table 6 is different from "greatly raised"), how to improve the estimation accuracy of the initial azimuth error angle of INS becomes the key problem. Since the uniaxial rotation INS adopts two position alignment mode, the initial alignment heading accuracy can be within 40″, and the horizontal Angle error can reach the second level, thus, when initial error of INS is calibrated and compensated, it's not necessary to consider the high precise calibration of INS initial alignment error, while the installation angle calibration should only be considered. To the shield machine of 10m diameter , If INS is installed at the 45° edge of the upper right corner of the rear vertical plane of the shield machine, and the characteristic point of the shield machine is selected in the upper left corner of symmetric position, then $l^b = [7.07m \quad 0m \quad 3.535m]^T$, and the total station is set as near as possible from the shield machine,such as $\rho^{b'} = [0 \quad -10m \quad 0]^T$,the installation angle estimation curve and estimation accuracy results are shown in Figure 13 and Table 7.

TABLE 7. ACTUAL ENGINEERING ESTIMATION CASES (300 SECONDS, UNIT: MIN (′))

item	α_x	α_y	α_z
set value	-11	9	20
estimation value	-1.1036e+01	8.9973e+00	1.9798e+01
bias	-3.6231e-02	-2.7211e-03	-2.0164e-01

As can be seen from Figure 13, the three installation angles quickly converged within 180 seconds without fluctuation and overshoot. As can be seen from Table 7, the estimation accuracy of the three installation angle α is better than 15″, so the initial alignment error $\varnothing$ overlaying installation angle α is generally better than 1′ as a whole, which fully meets the requirements of the actual engineering accuracy of 1 mrad.

Fig. 13. Actual Engineering estimation curve(α)

VII. CONCLUSION

For the extremely slow and extremely low dynamic installation error estimation situation of shield tunneling machine, this paper presents the scheme of shield machine guidance system with uniaxial rotating laser inertial navigation as the core. By rotating the IMU itself, the observability of the installation error parameter estimation is greatly improved; At the same time, this paper proposes the Kalman filter estimation scheme for the inertial navigation installation error assisted by the total station system, the system state space model of uniaxial rotating laser inertial navigation assisted by total station is derived, the effects of the IMU rotation, the total station pitch and the rod arm between the inertial navigation system and the observation points on the estimation observability and accuracy of the installation error are analyzed in detail, the optimal scheme of installation error estimation of "uniaxial rotational excitation + reducing oblique pitch + increasing rod arm" is proposed, which has a strong guiding value for practical engineering.

REFERENCES

[1] Li Hailiang. Measurement method and applicability analysis of the initial attitude of the shield tunneling machine [J], LOW CARBON WORLD, March,2020,pp:179-180.

[2] Luo Strong. The integrated navigation system calibration of inertial navigation and odometer based on the least squares method[J], MISSILES AND SPACE VEHICLES, No.1,2014,Sum No.330,pp:30-33.

[3] Da Xingpeng.Mobile robot /odometer system error and lidar installation error online calibration[J], ROBOT, vol.39, No.2, Mar. 2017, pp:206-213.

[4] Wang Xin.Star sensor installation error calibration technology research[J], Navigation, Positioning & Timing, Vol. 6, No.3, May 2019, pp:126-130.

[5] Tong Jinwu. Analysis on the impact of ultra-short baseline installation error on positioning accuracy and its Calibration Technology research [J], Navigation, Positioning & Timing, Vol.7, No.2, Mar. 2020,pp:19-27.

[6] Wang Li.An online rapid estimation method of all-day star tracker[J], Acta Aeronautica et Astronautica Sinica, Vol.41,No.8,August 2020,pp:624117-2~8.

[7] Anjing Ke. Online calibration method for installation error of space spacecraft[J], Navigation, Positioning & Timing, Vol.9, No.2, March 2022,pp:83-90.

[8] Zhou Zhafa. Rapid Calibration Method for Milometer Parameters based on SINS Speed Information [J], Journal of Southwest Jiaotong University,Network first paper, https://kns.cnki.net/kcms/detail/51. 1277.U.20220608.1507.004.html, June 9,2022.

[9] Shi Zhiyong. Online calibration technology of rotating strapdown inertial navigation system[M]. Beijing: Beijing Institute of Technology Press, August 2021,first edition.

Magnetic Compass with a New Correction System: Field Test Results

D. G. Gryazin

Concern CSRI Elektropribor, JSC, ITMO University,
St. Petersburg, Russia
volnagdg@yandex.ru

T. V. Paderina

Concern CSRI Elektropribor, JSC,
St. Petersburg, Russia
paderinata@gmail.com

The paper studies the performance of a magnetic compass with a new correction system based on a complementary filter. The correction system is intended to reduce the effect of dynamic errors occurring under the ship rolling and pitching motion and conditioned by the influence of ship magnetic forces (heeling deviation) and translational accelerations if the compass is installed at some distance from the ship oscillation center.

The results from bench and field tests of the magnetic compass (in the Lake Ladoga) are presented, which confirm the effectiveness of the correction system.

Keywords—magnetic compass, correction system, heeling deviation, translational accelerations, MEMS gyroscope

I. INTRODUCTION

Under the ship rolling and pitching motion, the errors in magnetic compass occur, conditioned by two factors. First, the redistributed ship magnetic forces affect the magnetic system of the compass card, which results in heeling deviation [1], which can be eliminated only for a certain latitude. Second, tangential and centripetal (translational) accelerations conditioned by the compass installation at some distance from the ship oscillation center influence the compass card, being a short-period pendulum [2].

These errors are most pronounced in high latitudes, which additionally complicates the heading indication, because the card directional magnetic moment caused by the horizontal component of the Earth's magnetic field is initially small. Considering the growing number of ships, including fishing ones, which work in the northern seas, and therefore the increasing use of the Northern Sea Route, suppressing the dynamic errors of magnetic compass becomes a key problem.

The paper presented at the previous International Conference on Integrated Navigation Systems (May 2023) was devoted to the development of a new dynamic error compensation method for a magnetic compass in roll and pitch conditions. It was proposed to use a MEMS gyroscope in the compass design and to develop a complementary filter integrating the data from the magnetic heading sensor (MHS) generating the signal from the card and the MEMS gyro in order to improve the performance.

II. MAGNETIC COMPASS CORRECTION SYSTEM BASED ON A COMPLEMENTARY FILTER

The block diagram of the compass correction system based on a complementary filter is presented in Fig. 1 [3].

Fig. 1. Block diagram of the magnetic compass correction system based on a complementary filter.

The principle of the complementary filter consists in compensating the dynamic errors due to the translational accelerations and ship magnetism during the roll/pitch motion with the MEMS gyroscope data, and to compensate the errors due to the gyro drift low-frequency components with the MHS data. Thus, the complementary filter functions as a high-pass filter (HPF) for the MEMS gyroscope and as a low-pass filter (LPF) for the MHS [4]. The complementary filter is described with an equation [5]:

$$\hat{\alpha} = W_{HPF}(p) \cdot \alpha_G + W_{LPF}(p) \cdot \alpha_{MHS}, \qquad (1)$$

where $W_{HPF}(p)$ is the HPF transfer function, $W_{LPF}(p)$ is the LPF transfer function, α_G, α_{MHS} are the total heading and yaw angles from the MEMS gyroscope and MHS, respectively, $\hat{\alpha}$ is the sum of heading and yaw angles estimated by the filter.

In the simplest case, the relevant links with the transfer functions

$$W_{HPF}(p) = \frac{Tp}{Tp+1}, W_{LPF}(p) = \frac{1}{Tp+1}, \qquad (2)$$

can be used as HPF and LPF, where T is the HPF and LPF time constant.

The difference equation for estimating the current sum of heading and yaw angles $\hat{\alpha}$ can be presented as follows:

$$\hat{\alpha}_t = \frac{T}{T + \Delta t} \left[\hat{\alpha}_{t-1} + \omega_G \Delta t \right] + \frac{\Delta t}{T + \Delta t} \alpha_{MHS}, \quad (3)$$

where $\hat{\alpha}_{t-1}$ is the estimate of the sum of heading and yaw angles at the previous time;

ω_G is the angular rate from the MEMS gyroscope;

α_{MHS} is the heading and yaw angle from MHS;

Δt is the sampling interval.

To provide invariance conditions, the factors K_1 and K_2 in Fig. 1 should be selected as follows: $T = \frac{K_2}{K_1}$, if

$K_1 = 1$, K_2 should be selected equal to the time constant $K_2 = T$.

Therefore, the total heading and yaw angle $K_M + \varphi$ (K_M is the magnetic heading, φ is the yaw angle) in Fig. 1 should be generated according to the following algorithm:

1. The MHS output comes to the link with the gain factor K_1, and the MEMS gyro output comes to the link with the gain factor K_2.

2. MHS signal $K_1 \cdot \alpha_{MHS}$ ($\alpha_{MHS} = K_M + \phi$) is summed with the MEMS gyro signal $K_2 p \alpha_G$ ($\alpha_G = K_M + \phi$).

3. The obtained signal sum is applied to the input of the aperiodic link $\dfrac{1}{Tp+1}$. The signal from the aperiodic link $\hat{\alpha} = K_M + \varphi$ being the estimated sum of the heading and yaw angles free of the compass error due to roll/pitch motion is applied to the remote display.

The MEMS gyroscope applied in the compass correction system has the following characteristics: measured angular rate – 75 – 900 deg/s (adjustable); bias instability – max 24 deg/h, ARW – max 0.28 deg/√h, bandwidth up to 160 Hz. Considering the gyroscope performance, sampling interval $\Delta t = 0.01$ s, and based on the simulation modeling results, the filter time constant T was selected to be 10 s.

II. RESULTS FROM BENCH AND FIELD TESTS OF A MAGNETIC COMPASS WITH A NEW CORRECTION SYSTEM

The capability of the magnetic compass correction system based on a complementary filter with the time constant $T = 10$ s of reducing the compass dynamic errors under the ship motion was experimentally estimated on a specialized test bench (Fig. 2) [6]. The test bench consists of a DC drive, a platform for the compass bowl, and two constant magnets positioned at some distance from the drive rotation axis. The compass bowl is installed at a distance l relative to the same axis of rotation.

Fig. 2.   Yaw test bench.

When the compass bowl is rotated by a certain angle, additional magnetic field is induced to the compass card, causing it to deviate from the physical yawing angle by an additional angle. The test bench design allows applying an extra magnetic moment to the card magnetic system, which provides the card additional rotation. This additional magnetic moment simulates the effect of ship magnetic forces due to motion. The compass magnetic system then should compensate the effect of this additional error.

The compass was tested at 60 deg N (horizontal component of the magnetic field strength is (15±1) µT). Yawing motions within ±10º with different periods from 4 to 30 s were set. The constant magnets positioned near the bowl simulated the effect of heeling deviation at extreme positions – plus and minus 10 deg, which resulted in additional card rotation within a doubled range ±20º. The compass correction system should correct the compass readings to the yawing angles commanded by the test bench, i.e., angles within ±10º. As a measure of the efficiency of correction system, we used the efficiency factor:

$$K_{ef} = \left| DK^M_{nocorr} \right| / \left| DK^M_{corr} \right| \qquad (4)$$

where DK^M_{nocorr}, DK^M_{corr} are the compass errors being the limiting differences between the noncorrected/corrected compass outputs and true yawing angles.

The dependency of K_{ef} on the disturbance periods (in bench tests, the periods of commanded harmonic oscillations of yawing angles) is shown in Fig. 3. The values of K_{ef} range from 3.7 to 9.8 as the yawing periods vary from 4 to 30 s, with the maximum K_{ef} observed at 18 s.

Fig. 3.   Efficiency factor vs. the disturbance periods.

The field tests of the compass with a correction system were conducted on board Azimuth ship with 100 tons water displacement in the Lake Ladoga. The compass was installed on the upper bridge. Additionally, a strapdown attitude determination system with GNSS receiver was used to determine the ship oscillation angles, and a gyrocompass was used as a source of reference heading. The additional vertical magnets were removed from the binnacle during the measurements to investigate the effect of noncompensated heeling deviation on the correction system (in standard mode, the magnets compensate the heeling deviation).

As an illustration, Figure 4 presents the errors of the compass with the running and disabled correction system obtained by forming the differences between the corresponding output signals of the compass and the

readings of the gyrocompass when the ship is moving on a constant course.

Fig. 4.   Compass errors with the enabled and disabled correction system (heading $K = 0°$).

The field tests have demonstrated that under rolling with amplitude up to 17° and pitching with amplitude up to 4.5° and periods close to 4 s, application of the correction system in the compass measurement scheme reduces the error, with the efficiency factor K_{ef} depending on the oscillation periods being from 2.4 to 3.5. This agrees with the bench test results with the extra magnets simula   g deviation, where K_{ef} for the yawing periods about 4 s was

$$K_{ef} = 3.7.$$

III. CONCLUSIONS

The paper presents the results from bench and field tests of the magnetic compass with a new correction system. The correction system is intended to reduce the effect of dynamic errors occurring under the ship rolling and pitching motion and conditioned by redistributed ship magnetic forces (heeling deviation) and translational accelerations if the compass is installed at some distance from the ship oscillation center.

The bench tests results have confirmed the efficiency of the correction system based on complementary filter, which reduces the compass dynamic errors 3.5 to 9.8 times depending on the prevailing frequencies (periods) of ship oscillations.

The field tests conducted in the Lake Ladoga with maximum 4 s oscillation periods have shown that the dynamic errors are reduced 2.4 to 3.5 times, which agrees with the bench test results.

The findings can be applied in developing the modern magnetic compasses to provide higher heading indication accuracy. Moreover, the correction system enhances the stability of the current heading and yawing when using the magnetic compass in auto pilot control loop.

REFERENCES

[1] V.P. Kozhukho, V.V. Voronov, and V.V. Grigoryev, Deviatsiya magnitnogo kompasa (Magnetic Compass Deviation), Moscow: Transport, 1981.

[2] N.Yu. Rybaltovskii, Magnitno-kompasnoe delo (Magnetic Compass Technology), Leningrad, 1954.

[3] Yu.V. Matveev, D.G. Gryazin, and T.V. Paderina, A method of magnetic heading measurement using the correction system, Patent RU 2804444, registered on 29.09. 2023.

[4] A.Yu. Ivoilov, "Application of MEMS sensors in developing a two-wheeled robot automatic stabilization system", Sbornik nauchnykh trudov NGTU, 2017, no. 3 (89), pp. 32–51.

[5] O.A. Stepanov and M. Mansur, "Integrated processing algorithms in the correction of navigation system data under nonlinear measurements", Izvestiya Tulskogo gosudarsvennogo universiteta. Tekhnicheskie nauki, 2016, no. 6, pp. 89–102.

[6] D.G. Gryazin, A method of determining the dynamic error of a magnetic compass with a correction system when exposed to pitching and apparatus for its implementation, Patent RU 2783479, registered on 14.11.2022.

Quaternion Algorithm for Strapdown Gyro-Vertical

V. Smirnov	A. Prokhortsov	V. Zarubin
FSBEI HE	FSBEI HE	JSC "VNII "Signal"
Tula State University	Tula State University	Kovrov, Russia
Tula, Russia	Tula, Russia	Zarubin@gmail.com
veld071@rambler.ru	ProxAV@rambler.ru	

Abstract — **Quaternion equations of strapdown gyrovertical are considered. In addition to the integral correction, which does not provide damping of oscillations with the Shuler period, an algorithm for bringing to the horizon and damping oscillations is proposed, which is practically invariant to linear accelerations.**

Key worlds —gyro-vertical, quaternion, strapdown inertial navigation system

I. INTRODUCTION

Strapdown inertial navigation systems (SINS) are widely used for navigation on the Earth's surface. The most commonly used algorithm for determining orientation in such SINS is the integration of kinematic equations, also called Poisson equations [1, 2]. At the same time, in SINS of medium and high accuracy it is necessary to compensate for the angular velocities due to the daily rotation of the Earth and the fact that movement along the Earth's surface occurs along an arc, the radius of which is determined by the size of the Earth. The disadvantage of this algorithm is the rather rapid increase in time of errors in determining the angular position. To eliminate these errors, a correction based on the signals of accelerometers included in the SINS, is introduced into the algorithm for integrating angular velocities. Depending on the correction algorithm, it is possible to eliminate the increase in errors in roll and pitch (gyrovertical), heading (gyrocompass) or in all three angles (heading gyrovertical). These algorithms are widely known, but their justification is based on the mathematical apparatus of Euler-Krylov angles [3,4], which presents some inconveniences during their practical implementation in modern navigation systems using the quaternion apparatus.

II. MATHEMATICAL MODEL FOR DETERMINING GYROVERTICAL ORIENTATION PARAMETERS

Let's consider a coordinate system $OENH$, the center of which is connected to the center of mass of the moving object, the axis ON is directed along the meridian to the north, the OH axis is directed along the local vertical, and the OE axis complements the system to the right. Let's associate with the inertial measurement unit a coordinate system $OXYZ$ rotated relative of the $ONHE$ to the Euler-Krylov angles (Figure 1).

It is known [1,2,5] that the equation for calculating the orientation quaternion from angular velocities has the form:

$$\dot{\Lambda} = \frac{1}{2}\Lambda \circ \overline{\omega}_{\text{c}} , \qquad (1)$$

where $\overline{\omega}_{\text{c}}$ – vector of absolute angular velocity in a coupled coordinate system; Λ – quaternion conversion from coupled to base coordinate system; $\circ$ – quaternion multiplication operator.

Fig. 1. Coordinate systems

When moving along the Earth's surface, the absolute angular velocity vector $\overline{\omega}_{\overline{6}}$ includes the angular velocity vector around the Earth and the Earth's rotation velocity vector:

$$\overline{\omega}_{\overline{6}} = \overline{\Omega}_{g\overline{6}} + \overline{\Omega}_{3} + \overline{\Omega}_{V} , \qquad (2)$$

where $\overline{\Omega}_{g\overline{6}} = [\Omega_{E_g} \quad \Omega_{N_g} \quad \Omega_{H_g}]^{T}$ – vector of the angular velocity of the inertial measurement unit relative to the surface of the earth's ellipsoid in the coordinate system $OENH$; $\overline{\Omega}_{3} = [0 \quad \Omega_{3}\cos\varphi \quad \Omega_{3}\sin\varphi]^{T}$ – vector of the angular velocity of the Earth's rotation in the coordinate system $OENH$; $\overline{\Omega}_{V} = \left[-\dfrac{V_N}{R_\varphi + H} \quad \dfrac{V_E}{R_\lambda + H} \quad \dfrac{V_E\text{tg}\varphi}{R_\lambda + H}\right]^{T}$ – vector of angular velocity due to the flight around the earth's ellipsoid; V_E , V_N – projections of the linear velocity of the inertial measurement unit to the eastern and northern axes of the accompanying geographic trihedron, respectively; R_φ , R_λ – radii of curvature of normal sections of a common earth ellipsoid; φ – longitude; H – altitude.

To eliminate the increase in errors in orientation angles, correction angular velocities ω_N^k , ω_E^k , ω_H^k , are added to the angular velocities $\overline{\omega}_{\overline{6}}$, calculated from accelerometer signals and ensuring that the virtual accompanying geographic trihedron $ONHE$ is brought to the horizon and azimuth.

Corrective angular velocities ω_N^k , ω_E^k , ω_H^k are calculated from accelerometer signals in accordance with the formulas:

$$\overline{\Omega}_k = K \cdot \overline{a}_{\overline{6}} + \int_0^t K_I \cdot \overline{a}_{\overline{6}}\,dt \qquad (3)$$

where $\overline{\Omega}_k = [\omega_E^k \quad \omega_N^k \quad \omega_H^k]^T$ – vector of correction angular velocities in the coordinate system $ONHE$; K – matrix of transmission coefficients of the proportional correction component; K_I – matrix of transmission coefficients of the integral correction component; $\overline{a}_б = [a_E \quad a_N \quad a_H]$ – apparent acceleration vector in the coordinate system $OENH$.

Translating (1) into the base coordinate system and substituting (3) and (2), we may find:

$$\dot{\Lambda} = \frac{1}{2}\left(\overline{\Omega}_{gб} + \overline{\Omega}_3 + \overline{\Omega}_V - K \cdot \overline{a}_б - \int_0^t K_I \cdot \overline{a}_б dt\right) \circ \Lambda \qquad (4)$$

Let's represent the vector $\overline{\Omega}_V$ in the form $\overline{\Omega}_V = K_V V_б^{\text{ист}}$, where

$$K_V = \begin{bmatrix} 0 & -\dfrac{1}{R_\varphi + H} & 0 \\ \dfrac{1}{R_\lambda + H} & 0 & 0 \\ \dfrac{\text{tg}\varphi}{R_\lambda + H} & 0 & 0 \end{bmatrix}, \quad \overline{V}_б^{\text{ист}} = \begin{bmatrix} V_E \\ V_N \\ V_H \end{bmatrix}.$$

Considering that the apparent acceleration vector is $\overline{a}_б = \overline{a}_б^{\text{ист}} + \overline{a}_б^{\text{вр}}$, where $\overline{a}_б^{\text{ист}} = \dot{V}_б^{\text{ист}}$ – acceleration vector relative to the base coordinate system; $\overline{a}_б^{\text{вр}} = K_{Va}\overline{V}_б^{\text{ист}} + K_{\Omega a}\overline{V}_б^{\text{ист}} + \overline{g}$ – "harmful" component of the apparent acceleration vector;

$$K_{\Omega a} = \Omega_3 \begin{bmatrix} 0 & -2\sin(\varphi) & 2\cos(\varphi) \\ 2\sin(\varphi) & 0 & 0 \\ -2\cos(\varphi) & 0 & 0 \end{bmatrix};$$

$$K_{Va} = \begin{bmatrix} 0 & -\dfrac{V_E\text{tg}(\varphi)}{R_\lambda + H} & \dfrac{V_E}{R_\lambda + H} \\ \dfrac{V_E\text{tg}(\varphi)}{R_\lambda + H} & 0 & \dfrac{V_N}{R_\varphi + H} \\ -\dfrac{V_E}{R_\lambda + H} & -\dfrac{V_N}{R_\varphi + H} & 0 \end{bmatrix}; \quad \overline{g} = \begin{bmatrix} 0 \\ 0 \\ g \end{bmatrix}$$ – gravity

vector, let us transform equation (4) to the form:

$$\dot{\Lambda} = \frac{1}{2}\left(\begin{array}{c} \overline{\Omega}_{gб} - K \cdot \overline{a}_б + \overline{\Omega}_3 + (K_V - K_I)\int_0^t \overline{a}_б dt - \\ - K_V \int_0^t (K_{Va}\overline{V}_б^{\text{ист}} + K_{\Omega a}\overline{V}_б^{\text{ист}} + \overline{g})dt \end{array}\right) \circ \Lambda \qquad (5)$$

When $K = 0$, $K_V = K_I$ equation (4) will take the form

$$\dot{\Lambda} = \frac{1}{2}\left(\overline{\Omega}_{gб} + \overline{\Omega}_3 - K_V \int_0^t (K_{Va}V_{\text{ист}} + K_{\Omega a}V_{\text{ист}})dt\right) \circ \Lambda \qquad (6)$$

The form of equation (6) confirms the well-known result [1-4,7,8] that with integral correction adjusted to the Schuler period, linear accelerations will not affect the orientation parameters.

Let us obtain the error equation, taking into account that in land navigation the orientation relative to the global coordinate system is of interest, the quaternion of which can be calculated using the equation:

$$\dot{\Lambda} = \frac{1}{2}\overline{\Omega}_{gб} \circ \Lambda. \qquad (7)$$

In this case, subtracting (6) from (5) and neglecting terms of higher orders of smallness, the orientation error equation can be written in the form:

$$\Delta\dot{\Lambda} + \frac{1}{2}K_V \int_0^t (\overline{g} \circ \Lambda \circ \Delta\tilde{\Lambda} + \Delta\Lambda \circ \tilde{\Lambda} \circ \overline{g})dt \circ \Lambda =$$

$$= \frac{1}{2}\left(\overline{\Omega}_3 - K_V \int_0^t (K_{Va}V_{\text{ист}} + K_{\Omega a}V_{\text{ист}})dt\right) \circ \Lambda + \qquad (8)$$

$$+ \frac{1}{2}\overline{\Omega}_{gб} \circ \Delta\Lambda + \frac{1}{2}(\overline{\Omega}_{gб} \circ \Lambda \circ \Delta\tilde{\Lambda} + \Delta\Lambda \circ \tilde{\Lambda} \circ \overline{\Omega}_{gб}) \circ \Lambda$$

Equation (11) confirms the well-known result [1-4,7,8], that in the case of integral correction at the $\Delta\Lambda \neq 0$, the gravitational acceleration vector $\overline{g}$ creates a component that increases in time, tending to reduce the magnitude of the orientation error.

It is well known that integral correction set to the Schuler period does not provide damping for oscillations with the Schuler period. To dampen oscillations and accelerate the bringing of the virtual platform to the horizon, either external sources of navigation information are used (usually speed from Doppler or other meters) or they use not only integral, but also proportional correction based on accelerometer signals [10,11,12], which violates invariance to linear accelerations.

One of the options for correction using accelerometers was proposed in [13], where the orientation quaternion is determined from the acceleration vector as a solution to the minimization problem:

$$\min_{\Lambda_a}\left(\tilde{\Lambda}_a \circ \overline{g} \circ \Lambda_a - \overline{a}_c\right) \qquad (9)$$

where Λ_a – orientation quaternion determined by accelerations; $\overline{a}_c$ – measured acceleration vector in a coupled coordinate system.

In this case, to solve this problem numerically, it is proposed to use the gradient descent method. However, this minimization problem has an exact solution [8]:

$$\Lambda_a = \begin{cases} \dfrac{1 + \overline{g} \cdot \overline{a}_c + \overline{g} \times \overline{a}_c}{\left|1 + \overline{g} \cdot \overline{a}_c + \overline{g} \times \overline{a}_c\right|} & \text{при } \left|1 + \overline{g} \cdot \overline{a}_c + \overline{g} \times \overline{a}_c\right| \neq 0; \\ \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} & \text{при } \left|1 + \overline{g} \cdot \overline{a}_c + \overline{g} \times \overline{a}_c\right| = 0. \end{cases} \qquad (10)$$

To obtain the resulting orientation quaternion, the quaternion determined by formula (10) is summed with the quaternion determined by the angular velocity sensors. In this case, the weighting coefficients are calculated depending on how much the measured acceleration vector differs from the acceleration of gravity:

$$\Lambda_\Sigma = (1 - \gamma)\Lambda_a + \Lambda \cdot \gamma, \qquad (11)$$

where $\gamma = \dfrac{g}{g + \left\|\Lambda_\Sigma \circ \overline{a}_c \circ \tilde{\Lambda}_\Sigma - \overline{g}\right\|}$.

To confirm the performance of the obtained algorithms for correcting orientation parameters based on accelerometer signals, experimental studies were carried out. The AIST-350 unit produced by Aisens LLC, built on the basis of micromechanical gyroscopes and accelerometers, the

parameters of which can be found in [14], was used as a block of inertial sensitive elements.

To check the invariance of the proposed algorithm to linear accelerations, the block was installed on a centrifuge where rotation was set at a constant speed. As can be seen from the graphs of orientation parameters, there are no roll and pitch deviations (Fig. 2, Fig. 4). Also, there are practically no heading deviations (Fig. 3, Fig. 5).

Note, that the first ten seconds are spent to preparing for the initial alignment and no data is transmitted. The next 30 seconds the initial alignment is carried out. In the Fig. 4 and Fig. 5, work in the setting-up mode and the initial alignment mode are not shown.

The results of experimental studies presented in the work confirmed the performance of the proposed gyrovertical algorithms.

Fig. 2.   Roll and pitch angles during rotation with acceleration 0.2g

Fig. 3.   Heading angle during rotation with accleration 0.2g

Fig. 4.   Roll and pitch angles during rotation with acceleration 1g

Fig. 5.   Heading angle during rotation with accleration 1g

III. Conclusion

The generalized equations of operation of a strapdown gyro in quaternions obtained in this work can be useful for research and construction of new algorithms for navigation systems. The proposed method for accelerated alignment to the horizon has also shown its effectiveness.

References

[1]   Anuchin O.N., Yemel'yantsev G.I. Integrirovannyye sistemy oriyentatsii i navigatsii dlya morskikh podvizhnykh ob"yektov (INS for marine venichles) SPb.: «Elektropribor», 2003 – 390 pp. (in Russian)

[2]   Yemel'yantsev G.I., Stepanov A.P. Integrirovannyye inertsial'no-sputnikovyye sistemy oriyentatsii i navigatsii (Integrated inertial-satellite attitude control and navigation systems). SPb.: Concern «Elektropribor», 2016. – 394 pp. (in Russian)

[3]   Plotnikov P.K. Elements of the theory of operation of one type of strapdown inertial orientation systems // Giroskopiya I navigatsiya. 1999. No. 3. pp. 23–35. (in Russian)

[4]   Plotnikov P.K. O sopostavlenii uravneniy dvizheniya i svoystv korrektiruyemogo girokompasa s algoritmami funktsionirovaniya i svoystvami yego besplatformenno-komp'yuternogo analoga // Vestnik SGTU. 2011. No. 4 (60). Vypusk 2, pp. 151 – 156.

[5]   Branets V.N., Shmyglevskiy I.P. Primeneniye kvaternionov v zadachakh oriyentatsii tverdogo tela. – M.: Nauka, 1973. – 320 pp.

[6]   Oriyentatsiya i navigatsiya podvizhnykh ob"yektov: sovremennyye informatsionnyye tekhnologii. / Pod red. B.S.Aleshina, K.K. Veremeyenko, A.I. Chernomorskogo – M.: FIZMATLIT. 2006.– 424 pp.

[7]   Salychev O. Applied Inertial Navigation: Problems and Solutions / O. Salychev. – M.: BMSYU Press, 2004. – 304 pp.

[8]   Paul G. Savage. Strapdown analytics. Part 1 // Strapdown associates, Inc Maple Plain, Minnesota, 2000. – 817 pp.

[9]   Chelnokov YU.M. Opredeleniye mestopolozheniya i oriyentatsii podvizhnykh ob"yektov po pokazaniyam chuvstvitel'nykh elementov BINS posredstvom resheniya na bortovom vychislitele kvaternionnykh uravneniy dvizheniya giroskopicheskikh sistem / YU.N. Chelnokov // Izv. RAN. Mekhanika tverdogo tela. 1991. No. 4. pp. 3-12.

[10]   Pat. RU 2647205 C2 Adaptive strap down inertial attitude-and-heading reference system // 14.03.2018 Bull. no. 8

[11]   Shvedov A.P., Ivanov Yu.V., and Raspopov V.Ya. Utility model patent 96235 Russian Federation. Strapdown Inertial Vertical Gyroscope. Tula State University, 2010.

[12]   Matveyev V.V. Analiz komplementarnykh fil'trov pri postroyenii beskardannoy girovertikali // Izvestiya TulGU. Tekhnicheskiye nauki. 2019. Vyp. 8, pp.153-164.

[13]   Madgwick S., Harrison A., & Vaidyanathan R. Estimation of IMU and MARG orientation using a gradient descent algorithm. Proceedings of the IEEE Int. Conf. Rehabil. Robot. 2011. P. 179 – 185.

[14]   Bogdanov M.B. Inertial measurement unit AIST-350t: results of mechanical investigation tests // Gyroscopy and Navigation. 2014. T. 5. No 4. pp. 238-244.

Estimation of Possibilities of Building a Block of Sensing Elements on the Basis of an Array of Micromechanical Inertial Sensors

D.V. Furtas
JSC «Inertial technologies of "Technocomplex"»
Rameskoye, Russia
DFurtas@inertech.ru

A.V. Nekrasov
JSC «Inertial technologies of "Technocomplex"»
Rameskoye, Russia
ANekrasov@inertech.ru

I.H. Shaimardanov
JSC «Inertial technologies of "Technocomplex"»
Rameskoye, Russia
IShaimardanov@inertech.ru

E.V. Babaev
JSC «Inertial technologies of "Technocomplex"»
Rameskoye, Russia
EBabaev@inertech.ru

A.A. Dzuev
JSC «Inertial technologies of "Technocomplex"»
Rameskoye, Russia
ADzuev@inertech.ru

A.N. Kostornoi
JSC «Inertial technologies of "Technocomplex"»
Rameskoye, Russia
AKostornoi@inertech.ru

Abstract-The use of a cluster of micromechanical sensors allows reducing noise characteristics, improving the accuracy of measurement information, accuracy of the resulting navigation solution, and increasing the level of reliability and failure control.

The paper is devoted to the problems of building navigation systems based on cluster of IMU, analyzing the accuracy and noise characteristics of the navigation system depending on the number of sensors, analyzing the main factors limiting the accuracy of modern micromechanical systems, and searching for ways to improve the accuracy of micromechanical navigation systems.

Keywords-Micromechanical systems, Allan variance, ARW, micromechanical IMU.

I. INTRODUCTION

A feature of modern micromechanical gyroscopes (MMG) is a large drift from launch to launch and a much smaller drift in launch [1]. In contrast to the error of gyroscopes, micromechanical accelerometers (MMAs) have comparable drift values both in launch and from launch to launch [2].

In recent years, works have appeared that consider the possibility of building an inertial measurement unit (IMU) based on a cluster of cheap and coarse sensors which can be configured to create a navigation system. This makes it possible, depending on the cost and accuracy requirements, to configure the composition of the navigation system by changing the number of sensors within the IMU. The use of a cluster of sensors allows reducing noise characteristics, increasing the accuracy of measurement information, accuracy of the resulting navigation solution, as well as the level of reliability and failure control.

The paper is devoted to the problems of building navigation systems based on cluster of IMU, analyzing the accuracy and noise characteristics of the navigation system depending on the number of sensors, analyzing the main factors limiting the accuracy of modern micromechanical systems, and searching for ways to improve the accuracy of micromechanical navigation systems.

II. BASIC APPROACHES TO BUILDING A CLUSTER SOLUTION

Based both on the error distribution density and the assumption of independence of errors of measurement modules, combining several inertial modules provides noise reduction of the sensor array solution. Also due to the additional control of measurements its reliability increases [3].

Consumer-grade MEMS can be used for the construction of IMU (the cost is a few dollars per measurement module). Assuming the probability theory and independence of errors of measuring modules, the error of the final measurement can be reduced by √N times on N inertial sensors [4]. The noise characteristics of the array of measurement modules should also be reduced.

Figure 1 shows the theoretical dependence of the measurement unit accuracy improvement as the number of used MEMS measurement modules increases. Averaging the output data of measurement modules leads to noise reduction in accordance with the following expression:

$$ARW_N = \frac{ARW_1}{\sqrt{N}} \qquad (1)$$

The graph of Angle Random Walk (ARW) dependence on the number of measuring modules (N) is shown in Figure 1.

Fig.1. Graph of Angle Random Walk (ARW) dependence on the number of measuring modules (N) at $ARW_1 = 0.3°/\sqrt{h}$

Experimental data confirm the theoretical approaches of constructing these solutions. However, the real improvement in the accuracy of the measurement unit is less than the theoretically expected one:

$$ARW_N^{\text{эксп}} \leq \frac{ARW_1}{\sqrt{N}} \qquad (2)$$

Due to the large number of inertial modules in IMU, one or more modules can fail. For example, the modules can become noisy or the sensor signal may jump at a certain temperature. This can severely affect the accuracy performance of the IMU. Therefore, it is important to apply filtering of inertial module readings when building the solution.

The main variants of filtering are median filter and 3-sigma filtering. The advantage of the first variant is low labor intensity; however, its potential accuracy is limited by the constant number of filtered channels. The advantage of the second one is a variable number of filtered channels, but this method is more labor-intensive and fails at high frequency of solution construction.

III. INVESTIGATION OF CLUSTER IMU ERRORS WITH INCREASING NUMBER OF INERTIAL MODULES USING ALLAN VARIANCE

In order to evaluate the possible improvement of accuracy characteristics and noise reduction, a number of IMUs based on 32, 64, 128, 256 inertial modules was manufactured. Using technological software, data recording was performed on the fixed base of each sample. Allan variance plotting was performed to evaluate the achievable accuracy and noise performance. Also, data recording of one inertial module was performed to compare the results.

To evaluate the Allan variance, IMU recordings of at least 650 seconds duration with a frequency of 500 Hz were used. The IMU was mounted in a horizontal position on a tilt-and-turn table. Comparative noise characteristics of the IMU obtained during the experiments for different numbers of inertial modules are given in Table 1.

According to the studies, the character of experimentally obtained dependence of ARW, Velocity Random Walk (VRW) and zero instabilities of accelerometers and gyroscopes coincide with the theoretically justified dependence. Thus, as the number of inertial modules in the IMU increases, the ARW of gyroscopes decreases by a factor

of 4 for 32 modules and up to 12 times for 256 modules, the VRW of accelerometers decreases to 6.5 times for 32 modules and up to 15.5 times for 256 modules. Meanwhile, according to ARW theory, the VRW should decrease by a factor of 5.6 and 16 for 32 and 256 modules, respectively.

Also, increasing the number of measurement modules reduces the accelerometer zero errors to 4.65 and 10.3 times for 32 and 256 modules and the gyroscope zero instability to 1.85 and 5.5 times for 32 and 256 modules, respectively.

TABLE I. IMU CHARACTERISTICS DEPENDING ON THE NUMBER OF INERTIAL MODULES

N	VRW Acc, m/(s²√h)	ARW Gyr, °/√h	Bias Acc, m/s²	Bias Gyr, °/h
1	3.49E-4	7.86	1.76E-4	2.49
32	5.377E-5	2.048	3.78E-5	1.34
64	3.03E-5	1.366	2.57E-5	0.75
128	2.78E-5	0.88	1.78E-5	0.67
256	2.24E-5	0.659	1.71E-5	0.45

IV. CONCLUSION

The main approaches to the formation of the cluster of IMU are analyzed and the accuracy and noise characteristics of the Allan variance-based IMU using 1, 32, 64, 128, and 256 inertial modules are shown. Analysis of the experimental data has confirmed the theoretically predicted errors of IMU based on the cluster solution.

The use of a cluster of 256 inertial modules allows us to significantly increase the autonomous operation time of the micromechanical navigation system and makes it possible, in a number of applications, to use a cluster of sensors as a cheaper analog with smaller mass-size characteristics of navigation systems based on laser, fiber-optic and solid-state-wave gyroscopes..

REFERENCES

[1] Matveev V.A.: A gyroscope is simple, MSTU N.E. Bauman, 2012.

[2] Gurtov V.A., Belyaev M.A. Micromechanical systems. - M.: PetrSU, 2016.

[3] Lukianov D.P., Raspopov V.Y., Filatov Y.V. Applied gyroscope theory, «Elektropribor», Saint-Petersburg, 2015.

[4] J.-O. Nilsson and I. Skog, "Inertial sensor arrays—A literature review", in Proc. Eur. Navigat. Conf. (ENC), May/Jun. 2016

Three-component Angular Velocity Sensor Based on a Modulation Micromechanical Gyroscope

Vladimir Bogolyubov
Kazan National Research Technical University named after A.N. Tupolev
Kazan, Russia
bvm200@yandex.ru

Oleg Tsyganov
Kazan National Research Technical University named after A.N. Tupolev Kazan, Russia
modelistrocky@mail.ru

Lyalya Bakhtieva
Kazan Federal University
Kazan, Russia
lbakhtie@yandex.ru

Abstract— **Mathematical model of the hybrid type gyroscope has built based on the modulation MMG. In order to eliminate the "zero displacement", the model uses the principle of modulation of primary information in the mechanical contour and its readout in a rotating coordinate system, which is well developed in the theory of the rotor vibrating gyroscopes. This approach makes it possible without interfering with the mechanical circuit of the device, to provide it with the ability to measure three components of the angular velocity of rotation of the base, i.e. create an absolute angular velocity sensor. Based on numerical experiments carried out in high-level mathematical programming packages, two algorithms for determining the third component of angular velocity, which coincides in direction with the vector of the gyroscope angular momentum, are formulated.**

Keywords— gyroscope, angular velocity sensor, beat oscillations

I. INTRODUCTION

Micromechanical gyroscopes (MMG), widely used in modern aircraft control systems [1-5], are mostly single-component angular velocity sensors. When solving problems of navigation (orientation) of moving objects, it is necessary to install three such sensors on board, the measuring axes of which must be oriented with sufficiently high accuracy relative to the axes of the object, which requires appropriate design solutions and time costs. In addition, the operation of such an instrument arrangement in conditions of temperature fields reaching a difference of 100^0C or more leads to errors in the orientation of the measuring axes of the devices, not to mention solving the issue of mutual influence of gyroscopes and their electromagnetic compatibility. Therefore, developers of primary information sensors strive to combine a comprehensive set of metrological capabilities in one device at the stage of their creation [6-8].

We also note that the small size of the sensitive element (SE) of the device creates the problem of ultra-high sensitivity of the system for readout of its vibrations, which limits the accuracy of determining the object location. In addition, the MMG sensitivity to inaccuracy of manufacturing the SE and the lack of the possibility of fine-tuning it leads to a significant "zero displacement", compensation for which is one of the main problems of the considered class of devices [9-10].

When developing modern primary information sensors, it is desirable to use new methods for improving their characteristics, based on circuit solutions without interfering with the mechanical circuit of the device. Below we will show that, taking into account the developments of instrument-making enterprises, the solution to the above problems can be implemented on the basis of hybrid MMGs [11]. In particular,

the two-step modulation MMG is characterized by significantly higher accuracy, has a large arsenal of metrological capabilities, and, in addition, it does not require high level of technological support for production.

This work is based on the results of works [12-14]. The authors obtained a mathematical model of a hybrid type gyroscope based on a two-step modulation MMG. To eliminate the "zero displacement", the principle of modulating primary information in a mechanical circuit and obtaining it in a rotating coordinate system, well developed in rotary-vibrating gyroscopes, is used.

This model makes it possible, without interfering with the mechanical design of the device, to provide it with the ability to measure three components of the angular velocity of rotation, i.e. to create an absolute angular velocity sensor. After simulating in the Maple and Matlab+Simulink packages [15-17] it was possible to determine the conditions for implementing the device's ability to measure the third component of the angular velocity, which coincides in the direction with the vector of the gyroscope angular momentum.

II. FORMULATION OF THE PROBLEM

Figure 1 shows the kinematic diagram of a modulation MMG with an RR-type (rotary-rotary) sensitive element made by planar technology. Here, 1 is the base of the device, 2 – the rotor electric drive, 3 – the rotor-sensitive element, 4 – torsion bar suspension, 5 – the torque sensor, 6 – the elements fastening the rotor to the drive shaft, and 7 is the angle sensor.

Fig. 1. Kinematic diagram of the modulation MMG

Note that in the above device, the elasticity of the rotor suspension along the OY_p axis is significantly greater than the elasticity of the rotor suspension along the OX_p axis,

which provides two degrees of freedom to the rotor (rotation with the electric drive shaft and angular oscillations around OX_B axis).

Figure 2 shows the coordinate system $OX_{06}Y_{06}Z_{06}$ rigidly connected to the device, its origin coincides with the center of mass of the object.

Fig. 2. Coordinate systems

The projections of angular velocity $\overline{\dot{\boldsymbol{\Phi}}}$ of rotational motion relative to the origin of coordinates will be denoted by $\dot{\boldsymbol{\Phi}}_x, \dot{\boldsymbol{\Phi}}_y, \dot{\boldsymbol{\Phi}}_z$. The coordinate system $OX_\text{B}Y_\text{B}Z_\text{B}$ associated with the drive motor shaft rotates with angular velocity Ω around the OZ_{06} axis. The coordinate system $OX_\text{p}Y_\text{p}Z_\text{p}$, associated with the SE, along with rotation together with the shaft, rotates around the torsion bar suspension at an angle θ with angular velocity $\dot{\boldsymbol{\theta}}$.

Let us determine the motion parameters of the gyroscope in the specified coordinate systems. Projections of the angular velocities of movement of the base on the axis of the system $OX_\text{B}Y_\text{B}Z_\text{B}$ after turning the shaft through an angle Ωt are determined by the expressions

$$\omega'_x = \dot{\Phi}_x \cos \Omega t + \dot{\Phi}_y \sin \Omega t,$$
$$\omega'_y = \dot{\Phi}_y \cos \Omega t + \dot{\Phi}_x \sin \Omega t,$$
$$\omega'_z = \Omega + \dot{\Phi}_z.$$

Projections of angular velocities on the axis of the system $OX_\text{p}Y_\text{p}Z_\text{p}$ after rotation through an angle θ can be written in the form

$$\omega_{px} = \omega'_x \cos \theta - \omega'_z \sin \theta,$$
$$\omega_{py} = \omega'_y + \dot{\theta},$$
$$\omega_{pz} = \omega'_z \cos \theta + \omega'_x \sin \theta.$$

To derive the differential equations of motion of the gyroscopic system, we will use the Lagrange method [18], choosing the rotation angle $\theta(t)$, which determines the SE position, as the generalized coordinate, and the damping moment and the elastic moment of the torsion bar as the generalized forces. After linearizing the equations using the Jacobian matrix, we obtain a mathematical model of modulation MMG for the case of movement of the device base with a constant angular velocity, which can be represented in the form

$$\ddot{\theta} + 2a\dot{\theta} + \omega_0{}^2\left(1 + \frac{2\kappa\dot{\Phi}_z}{\omega_0}\right) =$$
$$= (1+\kappa)\Omega\sqrt{\dot{\Phi}_x^2 + \dot{\Phi}_y^2}\,\sin(\Omega t + \psi),$$

$\dot{\boldsymbol{\Phi}}_x, \dot{\boldsymbol{\Phi}}_y, \dot{\boldsymbol{\Phi}}_z$ are projections of the angular velocity of the base movement, a is the damping coefficient of SE oscillations, $\omega_0 = \sqrt{\frac{k+(C-A)\Omega^2}{B}}$ is the frequency of natural oscillations of the SE, A, B, C – respectively equatorial and polar moments of inertia of the SE, $\Omega = \sqrt{\frac{k}{A+B-C}}$ is the resonant value of the angular velocity of rotation of the SE, $\kappa = \frac{C-A}{B} \approx 1$ is the coefficient of centrifugal pendulum rigidity, $\psi = arctg(\dot{\Phi}_x/\dot{\Phi}_y)$.

The initial conditions for the model can be taking in the form

$$\theta(0) = \dot{\theta}(0) = 0.$$

The peculiarity of this mathematical model is the presence in the free term of the equation of the addendum $\frac{2\kappa\dot{\Phi}_z}{\omega_0}$, caused by the projection of the angular velocity of rotation of the base $\dot{\Phi}_z$ onto the SE axis of rotation (usually this term is not taken into account in the equation, since it is of the second order of smallness). In order to eliminate the interference associated with the influence of angular vibration of the base with double the rotor speed on the accuracy of the device, the ratio of its moments of inertia is selecting so that the coefficient κ is close one [19]. The presence of the specified term in the equation is equivalent to a change in the value ω_0 − the frequency of natural oscillations of the SE, which leads to detuning of the device and, consequently, a change in the amplitude and phase of oscillations of the gyroscope rotor. Based on the magnitude of these changes, we can judge the value of the parameter $\dot{\Phi}_z$. that caused them. This is the basis for one of the methods proposed below for determining the third component of angular velocity.

In addition, from the vibration's theory we know [20] that if a weakly damped oscillatory system is acted upon by a force function with a frequency slightly different from the frequency of its free (natural) oscillations, then the observed response of the system to this effect leads to the occurrence of beat oscillations. Moreover, the frequency of beat oscillations turns out to be equal to the difference in the frequencies of natural oscillations of the system and the force function. Consequently, by measuring the magnitude of the phase shift of the natural oscillations of the system relative to the resonant value or the magnitude of the beat frequency, one can judge the proximity of the frequency of the force function to the frequency of free oscillations of this system.

The solution to the problem can be represented in the following form [20]

$$\theta(t) = K_g((1-e^{-at})\dot{\Phi}_X - 2e^{-at}\dot{\Phi}_Y \sin \dot{\Phi}_z t \sin \Omega t)$$
$$-(1-e^{-at})\dot{\Phi}_Y + 2e^{-at}\dot{\Phi}_X \sin \dot{\Phi}_z t \cos \Omega t, \qquad K_g = \frac{1+\kappa}{2a}.$$

The formula describes the MMG motion in the case of rotation of the base of the device with an absolute angular velocity. It consists of undamped forced oscillations caused

by the angular velocities $\dot{\Phi}_x, \dot{\Phi}_y$, and damped beat oscillations created by the third component $\dot{\Phi}_z$.

The type of beat oscillations in weakly damped and damped gyroscopes is presenting in Figures 3a and 3b.

Taking into account the high-quality properties of modulation MMG, beat oscillations have a significant duration, which ensures their observability. This circumstance allows us to measure the third component of the angular velocity $\dot{\Phi}_z$ with a weakly damped MMG.

Fig. 3. Character of oscillations of weakly damped beats (a) and damped (b) MMG

After demodulation, the output voltages of the demodulators for the corresponding measuring channels can be represented in the following form

$$U_X(t) = \frac{K_g}{2}((1 - e^{-at})\dot{\Phi}_X - 2e^{-at}\dot{\Phi}_Y \sin \dot{\Phi}_z t,$$

$$U_Y(t) = \frac{K_g}{2}((1 - e^{-at})\dot{\Phi}_Y + 2e^{-at}\dot{\Phi}_X \sin \dot{\Phi}_z t.$$

The resulting expressions and Figures 3 lead to the conclusion: knowing the beat frequency, one can determine the value of the angular velocity $\dot{\Phi}_z$ directed along the axis of rotation of the gyroscope SE.

Thus, information about the third component of the angular velocity $\dot{\Phi}_z$ can be obtained in two ways: by measuring the phase of oscillation of the MMG rotor, as well as by measuring the oscillation frequency of the beats. Algorithms for implementing these methods will be described below.

III. MEASUREMENT METHODS OF THE THIRD COMPONENT ANGULAR VELOCITY

A. *Determination of the third component by the magnitude of the change in the phase of oscillations of the SE*

The first method was implemented by solving the equation of gyroscope motion in the Maple environment.

From the frequency characteristics of the SE shown in Figure 4, it is clear that a change in the phase of its oscillations up or down relative to the resonant value equal to -90° is determined by the magnitude of the detuning caused by the angular velocity $\dot{\Phi}_z$ and the nature of the phase change determined by its direction.

$$\Delta\Phi = \Phi_{res} - \Phi_c,$$

$\Phi_{res} = \frac{\pi}{2}$ is the resonant value of phase, Φ_c is the current value of phase caused by parameter $\dot{\Phi}_z$.

Isolation of a signal proportional to the magnitude of the phase shift is carried out by phase detection, where the signal of the reference voltage generators is used as a reference signal. Note that the modulation MMG is a high-Q oscillatory

system is tuned to the resonance, and its frequency response has a significant steepness, which makes it possible to create a channel for measuring the angular velocity $\dot{\Phi}_z$ with high resolution.

Fig. 4. MMG frequency characteristics

When implementing the method under consideration it is necessary to keep in mind that two components of the angular velocity $\dot{\Phi}_X$ and $\dot{\Phi}_Y$ are measured in the sensitivity plane of the device. Therefore, depending on their ratio and sign, an additional phase shift corresponding to the polar angle ψ is entered into the information channel for measuring the phase shift created by the angular velocity $\dot{\Phi}_z$. Therefore, to eliminate this phase shift, it is necessary to use an additional channel that identifies the current value of the polar angle, for example, as shown in [21], and then, by performing an algebraic summation of the output signals of both channels, obtain the required information about the magnitude and direction of the angular velocity $\dot{\Phi}_z$.

B. *Determination of the third component by the frequency of beat oscillations*

The second method was implemented by modeling in the Matlab+Simulink environment. The block diagram of the model, based on the equation of gyroscope motion, is shown in Figure 5.

Fig. 5. Block diagram of the MMG simulation model

The simulation model diagram consists of four blocks:
- modulation MMG – mask MMG block,
- demodulator with low-pass filters – mask DM Filtr block,
- scheme for identifying the sign $\dot{\Phi}_z$ – block mask +/- PhiZ,
- frequency demodulator with frequency meter – mask Freq block. DM.

To expand the range of the measured angular velocity $\dot{\Phi}_z$, it is necessary to ensure a sufficient level of amplitude of beat oscillations, which decreases with detuning. To do this, it is advisable to expand the frequency bandwidth of the device. In [19], it is shown that to expand the bandwidth of the SE frequency response, it is advisable to ensure, by choosing the ratio of its moments of inertia, the value of the coefficient of centrifugal pendulum stiffness of the suspension $\kappa \approx 1$, thereby bringing the shape of the rotor closer to a parallelepiped.

Figure 6 shows the frequency characteristics of the gyroscope at various values κ.

Fig. 6. Frequency characteristics at different values of κ

From Figure 6 we can see that as κ increases, the bandwidth of the amplitude-frequency characteristic of the device expands and thereby ensures the required level of amplitude of beat oscillations under conditions of an expanded range of measured angular velocity $\dot{\Phi}_z$, however, the steepness of the phase-frequency characteristic decreases. Therefore, the value of the SE centrifugal pendulum stiffness should be chosen in accordance with the application area of the device.

IV. CONCLUSION

The mathematical model constructed by the authors, as well as the algorithms proposed in this work for determining the third component of angular velocity make it possible:
- to implement an absolute angular velocity sensor based on modulation MMG;
- to ensure the preservation of the resonant operating mode of the device by introducing an auto-tuning loop using the measured angular velocity $\dot{\Phi}_z$;
- to eliminate the so-called "zero displacement" characteristic of conventional MMGs based on the principle of modulation of primary information in a mechanical circuit and its acquisition in a rotating coordinate system;
- to provide a measurement range of absolute angular velocity along the third axis of sensitivity that significantly exceeds the measurement range along the other two axes, which allows the device to be used in rotating objects;
- to ensure the stable spatial orientation of a moving object with one three-component device under conditions of its operation in a wide range of temperature fields, reaching a difference of 100^0C or more.

All of the above capabilities may be implemented without interfering with the mechanical circuit of the device.

REFERENCES

[1] V.V. Matveev, and V.Ya. Raspopov, "Devices and systems for orientation, stabilization and navigation on MEMS sensors", XXIV St. Petersburg International Conference on Integrated Navigation Systems, 2017, 24, pp. 264-266.

[2] V. G. Peshekhonov, Y. A. Nekrasov Y. A., P. Pfluger, C. Kergueris, H. Haddara, and A. Elayed, "The results of an RR-type micromechanical gyroscope", IEEE Aerospace and electronic systems magazine, 2011, vol 1, pp. 14-21.

[3] D.P. Lukyanov, V.Ya. Raspopov, and Yu.V. Filatov, Applied theory of gyroscopes: textbook for universities, 2015, 315 p.

[4] A.S. Kovalev, M.I. Evstifeev, and D.P. Eliseev, "Analysis of the electromechanical model of the RR-type micromechanical gyroscope", Abstracts of reports. Proceedings of the XXVIII conference in memory of the outstanding designer of gyroscopic devices N.N. Ostryakov, St. Petersburg, 2012, 13 p.

[5] D.P. Eliseev, "Classification of methods for increasing the resistance of micromechanical gyroscopes to inertial influences", Navigation and motion control. Materials of the XVII Conference of Young Scientists, St. Petersburg: State Scientific Center of the Russian Federation JSC Concern Central Research Institute Elektropribor, 2015, pp. 569−577.

[6] Ts. Liang, Yu.A. Litvinenko, and O.A. Stepanov, "Method for processing measurements from two blocks of micromechanical gyroscopes when solving the orientation problem", Gyroscopy and navigation, 2018, vol. 2 (101). pp. 29-42.

[7] Yu.A. Ponomarev, D.V. Mayorov, A.V. Polynkov, V.E Chulkov., S.F.Konovalov, and A.G. Andreev, "A triad of accelerometers with a common magnetic system based on quartz compensation accelerometers", In the book: XLVI Academic Readings on Cosmonautics, dedicated to the memory of Academician S.P. Korolev and other outstanding domestic scientists - pioneers of space exploration. Korolyovsky readings, 2022, pp. 375-377.

[8] I.E. Lysenko, D.V. Naumenko, S.A. Sinyutin, and O.A. Ezhova, "Integral micromechanical gyroscope", RF Patent 2778622 C1, 2022.

[9] S.F. Konovalov, Yu.A. Ponomarev, and D.V. Mayorov, "Magnetic compensation of the zero signal in a hybrid two-coordinate R-R-R-type mems gyroscope", Bulletin of the Moscow State Technical University. N.E. Bauman. Instrumentation series, 2013, vol. 4 (93), pp. 122-131.

[10] S.F. Konovalov, V.P. Podchezertsev, and Yu.A. Ponomarev, "Modeling of zero signal compensation processes in a MEMS gyroscope", 2012, http://technomag.edu.ru/doc/476265.html.

[11] S.F. Konovalov, Yu.A. Ponomarev, and D.V. Mayorov, "Hybrid microelectromechanical gyroscopes and accelerometers", Scientific publication of MSTU named N.E. Bauman, 2011, vol. 10, pp. 1-23.

[12] L. Bakhtieva, and V. Bogolyubov, "Modulation of Damping in the Rotor Vibratory Gyroscopes", Russian Aeronautics, 2019, vol. 61, pp. 599-604.

[13] V. Bogolyubov, and L. Bakhtieva, "Astatic Gyrocompass Based on a Hybrid Micromechanical Gyroscope", IEEE East-West Design and Test Symposium (EWDTS), 2021, pp. 1-5.

[14] L. Bakhtieva, V. Bogolyubov, and M. Tumakov, "Simulation of a multifunctional micromechanical gyroscope", Program Systems: Theory and Applications, 2022, vol. 13, 3(54), pp. 17-27.

[15] V.N. Govorukhin, and V.G. Tsibulin, "Computer in mathematical research (Matlab, Maple, LaTex)", https://studfile.net/preview/878182

[16] M.N.Kirsanov, Mathematics and programming in Maple, Moscow, 2020, 164 p.

[17] A.I. Gerasimov, V.V. Regeda, and O.N. Regeda, Modeling in the MATLAB-Simulink environment: method. instructions for laboratory work, Penza: PSU Publishing House, 2017, 104 p.

[18] O.I. Kanakov, and M.I. Motova, Lagrange and Hamilton methods in the study of oscillatory systems: Educational and methodological manual, Nizhny Novgorod: Nizhny Novgorod State University, 2016, 39 p.

[19] B.N. Sukhanov, On increasing the accuracy of a two-dimensional single-channel angular velocity meter, In the book. Applied gyroscopy. Ed. V.A. Pavlova, L.: Publishing house Leningr. Univ., 1974, 206 p.

Usage of Differential Optical Measurements for Mutual Binding of Navigation Devices by Angular Position

S.Yu. Perepelkina

Scientific and Production Association
of Automatics named after academician
N.A. Semikhatov, JSC,
Yekaterinburg, Russia
avt@npoa.ru

A.A. Fedotov

Scientific and Production Association
of Automatics named after academician
N.A. Semikhatov, JSC,
Yekaterinburg, Russia
avt@npoa.ru

Abstract—**A method for non-contact mutual angular binding of navigation devices as part of a distributed control system for a rocket and space technology is considered. The implementation of the method involves retrofitting each navigation device with additional optical equipment - a laser module and an optical camera. The method is based on a simple and fast tracking procedure of the mutual drift of laser marks during operation process of the control object. Experimental results on the mutual angular binding of two navigation devices simulators using low-power laser modules and an electronic endoscope are presented.**

Key words—mutual angular binding, strapdown inertial navigation system, differential optical measurements, cross-type laser module, small deformations of the control object

I. Introduction

The problem of mutual angular binding of various instruments and devices has a multitude (including patented) of technical solutions in the field of geodesy, construction, mechanical engineering and, of course, rocket and space technology [1 − 8]. In the patent roll, the author's certificate [9] should be noted, which is closest to the concept of mutual binding by angular position, using modern optical equipment, outlined in this article.

Assuring high-precision mutual angular binding of navigation devices is a high-demand problem when various gyroscopic equipment and devices are used. This problem is especially relevant for strapdown inertial navigation systems (SINS)-based control systems. There is no possibility of fully solving the tare problem for SINS due to its stiffening as part of a rocket and space technology object. For the same reason, the initial binding of the SINS by angular position is carried out into great expenses. In such conditions, a more accurate navigation system, for example, a platform INS, can be used in order to improve the basic technical characteristics of SINS [10].

The method of mutual binding, based on vector matching of SINS and platform INS measurement information, is described in the technical literature [11]. This method can be used to adjust the SINS orientation relative to the vertical axis. In this case, a special measuring section is required with a turn of the longitudinal axis of the control object by 90 angular degrees. This method makes it possible to reduce SINS errors, but it has application constraint as a high-precision mutual binding of the SINS and the platform INS as part of the control object.

A stable mutual angular position of the devices can be ensured mechanically by placing on one bracket or one mounting plate. However, in practice, this approach is not always justified due to deformations of the body (structure) of the control object or its components when operating conditions change. This, in particular, can be related to external mechanical or climatic effect. In addition, navigation devices can use their own shock absorption systems with individual characteristics, which are determined by the design features of the sensitive elements of the measuring units and other design solutions [12]. In such conditions, it is necessary to control and algorithmically take into account changes in the mutual angular position of navigation devices. First, this is in demand for distributed control systems with spatially separated measuring units.

This article considers a non-contact method for solving the problem of angular binding of navigation devices using counter-directed cross-type laser modules. A simple and fast procedure for mutual angular binding of navigation devices is used by tracking the mutual drift of laser marks during the control object operation. Some of presented experimental estimates show the possibility of practical use of differential optical measurements and the method as a whole. Some characteristics of angular binding estimate range were obtained.

II. Mutual binding by angular position with the use of laser modules

Technically, to obtain differential optical measurements, two navigation devices must be spaced apart from each other in a certain direction (hereinafter referred to as the angular binding axis). An area is additionally marked (or determined additionally) on bodies of both navigation devices, directed towards each other, nominally orthogonal to the angular binding axis. Figure 1 shows a standard view of a marked area, where larger blue markings are applied to a fine green grid.

A laser module is rigidly mounted to each navigation device, nominally co-directed with the angular binding axis and looking onto the marked area of the adjacent device. Two optical cameras located away from the angular binding axis synchronously record images of the corresponding marked areas before and after switching on the laser modules. The marking used in this case should ensure the required binding accuracy of the laser mark screen coordinates to it. Mutual deviations of laser marks from the angular binding axis are intended to control mutual

orientation of navigation devices among themselves and to adjust their mutual binding by angular position.

Fig. 1. Design of marked area

Let us describe an approximate procedure for mutual angular binding of two navigation devices NP1 and NP2 using cross-type laser modules. It is considered that the devices belong to the same control object and are located at a distance (d ± 0,3%) relative to each other. In this case, the mutual angular deviations of the seats do not exceed ~1 angular degree. The necessary sequence of actions is functionally divided into three main stages.

At the first stage, the initial mutual binding of NP1 and NP2 is carried out (certification of the initial position). This stage includes the alignment of optical cameras, connection of laser modules and initial alignment of laser marks. The alignment of each optical camera is carried out using a video image of the corresponding marked area and consists of a standard procedure for determining correction coefficient (including scaling factors, axial misalignment, change in perspective, etc.) [13, 14] for correct recalculation of the screen coordinates of the optical camera into the reference binding coordinates of the marked area. The initial alignment of laser marks characterizes the initial relative position of the devices. Each laser mark is initially fixed in screen coordinates, and then is recalculated into the plane of the marked area. The resulting initial positions of the laser mark centers denote as $(X1, Y1)$ and $(X2, Y2)$ for the second and first marked areas, respectively. In this case, to calculate the center of each mark, let us take the intersection of curves that approximate the vertical and horizontal lines of a cross-type laser mark. The obtained coordinates, as well as the alignment correction coefficient of the optical cameras, are stored in the calculator's memory.

At the second stage, the change in the mutual angular position of the NP1 and NP2 devices compared to the initial one is estimated. New positions of laser marks $(X1n, Y1n)$ and $(X2n, Y2n)$ are calculated according to the same scheme as at the first stage. Note that at this stage the process of optical camera alignment can also be partially repeated to adjust/verify the corresponding correction coefficients.

The third stage consists of coprocessing of $(X1, Y1), (X1n, Y1n)$ and $(X2, Y2), (X2n, Y2n)$ pairs. The result of coprocessing is flat angles α and β, defining the change in the relative position of NP1 and NP2 in the longitudinal (see Fig. 2) and transverse planes:

$$\alpha \approx \alpha_1 + \alpha_2 \approx Arctg\left(\frac{\Delta Y1 + \Delta Y2}{d}\right),$$

$$\beta \approx Arctg\left(\frac{\Delta X1 + \Delta X2}{d}\right),$$

where $\Delta Y1 = Y1 - Y1n$, $\Delta Y2 = Y2 - Y2n$ – laser marks decentration in the longitudinal plane relative to the original (certificate) value for NP1 and NP2, respectively,

$\Delta X1 = X1 - X1n$, $\Delta X2 = X2 - X2n$ – laser marks decentration in the transverse plane relative to the original (certificate) value for NP1 and NP2, respectively.

Fig. 2. Change in the mutual angular position of NP1 and NP2 in the longitudinal plane

Assuming that the angles α and β are small (up to ~3 angular degrees), it is possible to calculate the spatial angle of change in the mutual angular position of NP1 and NP2 as

$$\gamma \approx \sqrt{\alpha^2 + \beta^2}.$$

III. DESCRIPTION OF THE EXPERIMENT RESULTS

An empirical research of the considered method of mutually binding of two navigation devices by angular position was carried out at an experimental operating position. This position is equipped with the following standard equipment: marked area 100×100 mm (2 pcs.), red laser module 14,5×48 mm with focusing 650 nm 5 mW (cross) with power supply and bracket (2 pcs.), electronic endoscope Orbit OT-SME12 (8 mm, 1280×720 5m) (2 pcs.), laptop (1 pc.). The marked areas are spaced 2 m apart from each other. The laser modules are located in close proximity to the angular binding axis to reduce methodological errors. Endoscopes are installed in front of the corresponding marked areas at an angle of up to 30 degrees so as not to obscure the laser mark.

To select and subsequently bind laser marks, images of marked areas were taken before and after switching on the laser modules. In this case, two modes of the endoscope built-in lighting and a mode without lighting were used. Note that maintaining a light balance turned out to be important from the viewpoint of finding procedure of the laser mark center. Examples of images processed in different lighting modes are shown in Fig. 3.

To initially identify laser marks lines, several directions were considered along the sides of the marked area. Such directions are nominally orthogonal to the lines of the laser marks. The cross-position of the laser marks was determined using approximation of the laser mark lines and subsequent conversion into the coordinates of the marked area. During the initial processing of images, color filtering and

normalization of images were performed according to the data before and after switching on the laser module. A corresponding example of color filtering and normalization is shown in Fig. 4.

Fig. 3. Examples of images processed in different lighting modes of the endoscope

Fig. 4. Separation of the laser module light line before and after primary processing

The resulting scattering of light lines and laser mark crosshairs makes it possible to evaluate the characteristics of the calculated angles α and β. In the longitudinal and transverse directions, the resulting scattering of angles does not exceed ~2 angular min.

Let us make a special emphasis on the use of a cross-type laser module. It allows to additionally obtain an estimate of the twist angle, which characterizes the turn relative to the angular binding axis. Despite the fact that the determination accuracy of twist angle is lower (according to preliminary estimates at the level of ~30 angular min), it is important for diagnostic purposes.

IV. CONCLUSION

The experimental implementation of the known scheme of mutual angular binding of two navigation devices with the use of modern optical equipment is considered. This scheme allows to track changes in mutual angular orientation by calculating angles in the longitudinal and transverse directions. The resulting accuracy may be of practical interest. In the option of using a cross-type laser module, the twist angle relative to the axis of the angular binding can be additionally calculated, although with less accuracy. Nevertheless, it is possible to consider a three-dimensional estimate of the relative angular orientation for these conditions.

Despite the need for stability of the laser module beam under external influences, the use of such a scheme can provide benefits in a number of operational characteristics. Thus, several different types of devices can be combined into one navigation system, which will allow, in particular, to naturally assign tasks solved during pre-flight preparation and in flight. When implementing such a scheme, there is no need for high-precision installation of equipment as part of the control object, and the requirements for the permissible deformation level of the control object during operation are significantly weakened.

Thus, the presented research can be considered as a prospective line of development of distributed control systems for rocket and space technology objects.

REFERENCES

[1] Sazonnikova, N.A., Nonin, A.S., Tkachenko, A.S., and Voblikov, D.N., Laser Measurement System for Monitoring Construction Elements Angular Position, Polzunovskij vestnik, 2015, no. 10, pp. 53–57. (in Russian).

[2] Stroilov, N.A., Nikitin, A.V., Kurkina, A.N., Bessonov, R.V., Belinskaya, E.V., and Voronkov, S.V., Ground Intercalibration Methods of Star Sensor Coordinate Frames, Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa, 2017, vol. 14, no. 4, pp. 52–66. (in Russian).

[3] Raissouni, F.Z., De-La-Llana-Calvo, Á, Lázaro-Galilea, J.L., Gardel-Vicente, A., Cherkaoui, A., and Bravo-Muñoz, I., Vehicular Visible Light Positioning System Based on a PSD Detector, Sensors, 2024, 24, 2320. https://doi.org/10.3390/s24072320

[4] Xu, X., Dai, Z., and Tan, Y., Noncontact measurement method of linear and angular displacement based on dual-beam feedback interferometric system, 2022. https://doi.org/10.48550/arXiv:2204.03488

[5] Zhang, P., Liu, J., Yang, H., and Yu, L., Position Measurement of Laser Center by using 2-D PSD and Fixed-axis Rotating Device, IEEE Access, 2019, vol. 7, pp. 140319–140327, doi: 10.1109/access.2019.2943910

[6] Liang, X., Lin, J., Yang, L., Wu, T., Liu, Y., and Zhu, J., Simultaneous Measurement of Absolute Distance and Angle Based on Dispersive Interferometry, IEEE Photonics Technology Letters, 2020, vol. 32, no. 8, pp. 449–452.

[7] Brunner, D., Meier, R., Weber, J., Measuring Device for Measuring the 3D Movement Between Two Objects, Measuring Device for a Maritime Observation and Defence Platform, and Platform, WO 2012/123405 A3, 2012.

[8] Bokhman, E.D., Ivanov, P.A., Larichev, R.A., Nikolaev, M.S., Pavlov, P.A., Filatov, Yu.V., Berkovich, S.B., and Kotov, N.I., Automated Angle Measurement System for Directions Transmission, Gyroscopy and Navigation. 2019, vol. 10, no. 3, pp. 180–185.

[9] Alekseev, V.E., and Orlovskij, V.P., Device for Measuring the Mutual Angular Movement of Two Objects, SU № 1472762 author's certificate, 1987. (in Russian).

[10] Vodicheva, L.V., Parysheva, Yu.V., and Savel'eva, Ya.S., Comparative Analysis of Algorithms for Initial Alignment of a Strapdown INS with the Help of a Gimballed INS, 30th Anniversary St. Peterburg International Conference on Integrated Navigation Systems, St. Peterburg, CSRI Electropribor, 2023.

[11] Vdovin, M.V., Gokhfeld, V.D., and Gurjeva, V.K., Vector Matching Method of SINS with High-Precision INS, Rocket and Space Technology: Scientific and Technical Collection, Series XI Control Systems of Missile Complex, Ekaterinburg, 2011, issue 1, pp. 71–78. (in Russian)

[12] Podchezertsev, V.P., and Topilskaya S.V., Choosing Dampling Parameters for the Inertial Orientation System, Herald of the Bauman Moscow State Technical University, Series Instrument Engineering, 2021, no. 3 (136), pp. 113–128. (in Russian).

[13] Remondino, F., and Fraser, C., Digital Camera Calibration Methods: Considerations and Comparisons, IAPRS, 2006, no. 36 (5), pp. 266–272.

[14] Rongfu Tang, Mathematical Methods for Camera Self-Calibration in Photogrammetry and Computer Vision, Dr.-Ing., University of Stuttgart Munchen, 2013.

Gyrocompassing Technique for a Strapdown Inertial Measurement Unit with Two Angular Rate Sensors

L. Vodicheva

JSC Academician N.A. Semikhatov Scientific
and Production Association of Automatics
Yekaterinburg, Russia
avt@npoa.ru

Yu. Parysheva

JSC Academician N.A. Semikhatov Scientific
and Production Association of Automatics
Yekaterinburg, Russia

V. Yakimov

JSC Academician N.A. Semikhatov Scientific
and Production Association of Automatics
Yekaterinburg, Russia

D. Kabanova

JSC Academician N.A. Semikhatov Scientific
and Production Association of Automatics
Yekaterinburg, Russia

Abstract—**A technique for estimating elevation and azimuth angles of a measurement device— optical, in particular—with the help of a strapdown inertial measurement unit (SIMU) including a triad of accelerometers and two angular rate sensors is considered. The algorithms, the results of the theoretical error analysis and the results of bench testing of a prototype of such an SIMU based on the micromechanical inertial measurement unit MPU-9250 and two fibre-optic gyroscopes (FOG) FOGS170B are described.**

Key words—strapdown inertial measurement unit, elevation angle, azimuth, error analysis.

I. INTRODUCTION

A conventional strapdown gyroscopic device for determining elevation and azimuth angles includes a triad of accelerometers and a triad of angular rate sensors (ARS) measuring correspondingly vectors of the normal reaction force to gravity (hereinafter called 'gravity') and the earth rotation rate. Accelerometers measuring the gravity with sufficient accuracy are relatively low-cost devices whereas only high precision ARS, the cost of each of which can be much higher than the cost of the rest of an SIMU, can provide the required accuracy of the earth rotation rate. Reducing the number of ARS of a SIMU allows for a notable decrease in its cost.

Obviously, an azimuth angle cannot be unambiguously determined in a single orientation with only one ARS, even if its sensitive axis is ideally levelled. Introduction of a second ARS into a SIMU allows solving the problem.

The problem of azimuth determination with the help of two ARS, i.e. under incomplete measurements, has been considered in a number of papers, in particular in the fundamental work [1], with regard to gyroscopic inclinometers. However, every practical application has its own peculiarities and every developer has his own approaches to algorithms and error analysis, thus the results presented in the paper appear to be original.

II. STATEMENT OF THE PROBLEM AND ERROR MODEL

The problem is solved with the use of the following inertial sensors:

– a triad of accelerometers included in an inertial measurement unit MPU-9250;

– two fibre-optic gyroscopes (FOG) FOGS170B.

To describe the statement of the problem, we use the following coordinate frames:

$ONHE$ is the geographical coordinate frame (GCF), (North, Height, and East);

$OXYZ$ is the right-handed orthogonal coordinate frame connected with a device body (BCF), the OX is directed along the longitudinal axis of a device, OY is oriented in the plane formed by the OX and the vertical under zero slope angle, OZ complements the frame to a right-handed one.

$OX_IY_IZ_I$ is the right-handed orthogonal SIMU instrumental coordinate frame (ICF) geometrically attached to SIMU accelerometers sensitive axes (SA).

Coordinate transformation from the GCF to the BCF is defined as three consecutive rotations:1) about the axis OH by an azimuth angle A clockwise, 2) about the axis OE by an elevation angle θ counterclockwise, 3) about the axis OX by a slope angle ψ counterclockwise. The transformation matrix is as follows:

$$M_{BCF \Leftarrow GCF} = \qquad (1)$$

$$\begin{bmatrix} \cos\theta\cos A & \sin\theta & \cos\theta\sin A \\ -\cos A\sin\theta\cos\psi - \sin A\sin\psi & \cos\theta\cos\psi & -\sin A\sin\theta\cos\psi + \cos A\sin\psi \\ \cos A\sin\theta\sin\psi - \sin A\cos\psi & -\cos\theta\sin\psi & \sin A\sin\theta\sin\psi + \cos A\cos\psi \end{bmatrix}$$

The specified angle ranges are

1) $A \in [0; 360°)$; 2) $\theta \in [45°; 80°]$; 3) $\psi \in [-5°; +5°]$.

Misalignments between the ICF and BCF axes are defined by three small angles of consecutive rotation $\Delta\alpha_x, \Delta\alpha_y, \Delta\alpha_z$.

The ICF axes coincide nominally with the accelerometers sensitive axes. Accelerometers error model is typical and includes biases, scale factor errors, and SA misalignments.

The FOG sensitive axes coincide nominally with the BCF axes OY and OZ. SA misalignments are parametrized by nonorthogonalities $\Delta\xi_y, \Delta\zeta_y, \Delta\xi_z, \Delta\eta_z$ relative to the BCF axes:

$$\vec{l}_y^{BCF} = \left[-\sin\Delta\xi_y, \sqrt{1-\sin^2\Delta\xi_y - \sin^2\Delta\zeta_y}, -\sin\Delta\zeta_y \right] \tag{2}$$

$$\vec{l}_z^{BCF} = \left[-\sin\Delta\xi_z, -\sin\Delta\eta_z, \sqrt{1-\sin^2\Delta\xi_z - \sin^2\Delta\eta_z} \right],$$

where $\vec{l}_y^{BCF}$, $\vec{l}_z^{BCF}$ are the unit vectors of the FOG sensitive axes projected on the BCF axes.

FOG error model includes biases and scale factor errors.

The task is to determine elevation and azimuth angles and to estimate errors of this determination.

With the help of a three-axis sensor of the specific force vector, determination of an elevation angle does not pose a problem. SIMU accelerometers outputs are as follows:

$$\begin{bmatrix} \hat{g}_x \\ \hat{g}_y \\ \hat{g}_z \end{bmatrix} = M_{BCF \Leftarrow GCF} \cdot \begin{bmatrix} 0 \\ g \\ 0 \end{bmatrix} = g \begin{bmatrix} \sin\theta \\ \cos\theta\cos\psi \\ -\cos\theta\sin\psi \end{bmatrix} \tag{3}$$

The elevation and slope angles are uniquely determined from the expression (3) without methodical errors. The expression for the elevation angle estimation error induced by the accelerometers errors under zero slope angle, lack of correlation of error sources and the same variances of the similar errors takes the form:

$$D[\delta\theta]_{акс} = \frac{D_{\tau_A}}{g^2} + 2\sin^2\theta \cdot \cos^2\theta \cdot D_{\rho_A} + \sin^4\theta \cdot D_{\beta_A}, \tag{4}$$

where D_{τ_A}, D_{ρ_A}, D_{β_A} are correspondingly the variances of biases, scale factor errors and misalignments of accelerometers sensitive axes.

Misalignments between the ICF and BCF axes reduced to a horizontal axis are included in the estimation error with the unit coefficient, i.e.

$$D[\delta\theta]_{\Delta\alpha} = D_{\Delta\alpha} \tag{5}$$

To determine the azimuth angle, consider two methods. In the first one (direct method) an azimuth is determined directly from sensor outputs [1] (first levelling and then gyrocompassing are implemented). The second one is based on the help of a virtual FOG, which allows using the TRIAD method having the minimum instrumental errors if there are all three projections of the earth rotation rate [2].

III. DIRECT METHOD FOR AZIMUTH ANGLE DETERMINATION

Earth rotation rate projections on the BCF axes are as follows:

$$\begin{bmatrix} \Omega_x \\ \Omega_y \\ \Omega_z \end{bmatrix} = \begin{bmatrix} \Omega_N \cos A\cos\theta + \Omega_H \sin\theta \\ -\Omega_N \cos A\sin\theta\cos\psi - \Omega_N \sin A\sin\psi + \Omega_H \cos\theta\cos\psi \\ \Omega_N \cos A\sin\theta\sin\psi - \Omega_N \sin A\cos\psi - \Omega_H \cos\theta\sin\psi \end{bmatrix} \tag{6}$$

where $\Omega_N = \Omega\cdot\cos\varphi$, $\Omega_H = \Omega\cdot\sin\varphi$, Ω is the magnitude of the earth rotation rate, φ is the latitude of a location.

From (6) obtain

$$\cos A = \frac{\Omega_H \cdot \cos\theta + \left(-\Omega_y \cdot \cos\psi + \Omega_z \cdot \sin\psi\right)}{\Omega_N \cdot \sin\theta} \tag{7}$$

$$\sin A = -\frac{\left(\Omega_y \cdot \sin\psi + \Omega_z \cdot \cos\psi\right)}{\Omega_N}$$

The azimuth angle is determined uniquely from the expressions (7) under $\theta \neq 0$ and with the help of elevation and slope angles determined before from accelerometers outputs. From (6) and (7) analytical expressions for azimuth angle estimation error can be obtained. Up to the first-order terms, the error caused by FOG additive errors takes the form:

$$\delta A_{FOG}^{(1)} = \left(-\cos A \cdot \sin\psi + \frac{\sin A \cdot \cos\psi}{\sin\theta}\right) \cdot \frac{\delta\Omega_y}{\Omega_N} - \left(\cos A \cdot \cos\psi + \frac{\sin A \cdot \sin\psi}{\sin\theta}\right) \cdot \frac{\delta\Omega_z}{\Omega_N} \tag{8}$$

The error caused by accelerometers additive errors including misalignment between their SA and a device body is as follows (under zero slope angle):

$$\delta A_{accel}^{(1)} = \left(\tan\varphi \cdot \sin A \cdot \cos\theta + \frac{\sin A \cdot \cos A \cdot \cos^2\theta}{\sin\theta}\right) \cdot \frac{\delta g_x}{g} - \left(\tan\varphi \cdot \sin A \cdot \sin\theta + \sin A \cdot \cos A \cdot \cos\theta\right) \cdot \frac{\delta g_y}{g} + \left(\tan\varphi \cdot \cos A + \frac{\cos^2 A \cdot \cos\theta}{\sin\theta} - \frac{1}{\sin\theta \cdot \cos\theta}\right) \cdot \frac{\delta g_z}{g} \tag{9}$$

From (2), (6) and (8), the expression for the azimuth angle estimation error as a function of misalignments between FOG SA and a device body can be obtained.

It is obvious that the method does not work under elevation angles close to zero. It is clear as well, that if $\theta \to \frac{\pi}{2}$, the method also does not work since the azimuth angle becomes indeterminable in this case.

IV. USE OF A VIRTUAL FOG

The value of the third projection of the earth rotation rate can be unambiguously obtained from the vector magnitude and two measured projections as follows:

$$\left|\hat{\Omega}_x\right| = \sqrt{\Omega^2 - \hat{\Omega}_y^2 - \hat{\Omega}_z^2} \tag{10}$$

The sign of the third projection is the same as the sign of the expression $\Omega_N \cos A\cos\theta + \Omega_H \sin\theta$ including the estimated azimuth angle. The sign of the third projection (as

well as the projection itself) can be obtained from the scalar product of the vectors of gravity and earth rotation rate:

$$\hat{\Omega}_x = \frac{g \cdot \Omega_H - \hat{g}_y \cdot \hat{\Omega}_y - \hat{g}_z \cdot \hat{\Omega}_z}{\hat{g}_x} \quad (11)$$

Provided that $\cos A > 0$ or $\theta + \varphi > \dfrac{\pi}{2}$, the sign of the third projection is necessarily positive.

If there are all three projections of angular rate and specific force, the direction cosine matrix between GCF and BCF $\hat{M}_{BCF \Leftarrow GCF}$ can be obtained with the help of the TRIAD method [2]. Then using its components m_{11}, m_{13} (expression 1), the azimuth angle can be uniquely determined by

$$\hat{A} = \text{arctg}\, \frac{\hat{g} \cdot \left(\hat{\Omega}_y \cdot \hat{g}_z - \hat{\Omega}_z \cdot \hat{g}_y \right)}{\hat{\Omega}_x \cdot \hat{g}^2 - \hat{g}_x \cdot \left(\hat{\vec{\Omega}}, \hat{\vec{g}} \right)}, \text{ where} \quad (12)$$

$$\hat{g} = \sqrt{\hat{g}_x^2 + \hat{g}_y^2 + \hat{g}_z^2}\,, \quad \left(\hat{\vec{\Omega}}, \hat{\vec{g}} \right) = \hat{\Omega}_x \cdot \hat{g}_x + \hat{\Omega}_y \cdot \hat{g}_y + \hat{\Omega}_z \cdot \hat{g}_z \quad (13)$$

Obtain errors of the method by varying expression (12). The error induced by FOG errors has more complicated form than the error of the first method

$$\delta A_{FOG}^{(2)} = \frac{\tan \varphi \cdot (\sin A \cdot \cos \psi - \cos A \cdot \sin \theta \cdot \sin \psi) - \cos \theta \cdot \sin \psi}{\cos A \cdot \cos \theta + \tan \varphi \cdot \sin \theta} \cdot \frac{\delta \Omega_y}{\Omega_N} - $$
$$- \frac{\tan \varphi \cdot (\sin A \cdot \sin \psi + \cos A \cdot \sin \theta \cdot \cos \psi) + \cos \theta \cdot \cos \psi}{\cos A \cdot \cos \theta + \tan \varphi \cdot \sin \theta} \cdot \frac{\delta \Omega_z}{\Omega_N} \quad (14)$$

The error caused by accelerometers errors has simpler form and is equal to the error of physical gyrocompassing [2]. Under $\psi = 0$, it can be written as:

$$\delta A_{accel}^{(2)} = \tan \varphi \cdot \sin A \cdot \cos \theta \cdot \frac{\delta g_x}{g} - \tan \varphi \cdot \sin A \cdot \sin \theta \cdot \frac{\delta g_y}{g} + $$
$$+ \left(\tan \varphi \cdot \cos A - \frac{\sin \theta}{\cos \theta} \right) \cdot \frac{\delta g_z}{g} \quad (15)$$

The expression for the azimuth estimation error as a function of misalignment between FOF sensitive axes and the body frame can be obtained from expressions (2), (6), (14).

This method also does not work if $\theta \to \dfrac{\pi}{2}$. The second singularity point (surface) for this method is determined by the expression $\cos A \cdot \cos \theta + \tan \varphi \cdot \sin \theta = 0$ meaning that the axis OX is parallel to the equatorial plane.

V. Error Analysis

Verification of the algorithms and theoretical estimation of azimuth determination errors were implemented with the help of mathematical simulation using two approaches: the method of weight functions and the method of statistical simulation.

In the method of weight functions, error sources considered as random variables are assumed to be uncorrelated and contribution of each source to azimuth estimation error is calculated separately. Total variance is determined as a sum

of component variances. This method allows verifying the results of calculation with the help of analytical expressions. The errors calculated by the simulating program coincide with the numerical values of the errors calculated by means of analytical expressions with the accuracy of three significant decimal digits.

When using the device, accelerometers will have in-run ellipsoidal calibration [3], which is why their errors were set at the following level (3σ): biases $5 \cdot 10^{-3}$ m/s², scale factor errors $5 \cdot 10^{-2}$%, and misalignments 1′. FOG biases were set according to Datasheet 0.06°/h, misalignments between accelerometers SA and body axes 0.5°, and misalignments between FOG SA and body axes 1°. Azimuth estimation errors for both methods are shown as function of azimuth angle in Fig. 1, as function of elevation angle in Fig. 2.

As it can be seen from Fig. 1, the errors of the second method are somewhat less than those of the first method. If $A \in [90°; 270°]$, it is the first method that has an advantage of accuracy. Figure 2 shows that if $\theta \to 0$, the first method stops working whereas the errors of the second method do not increase.

Fig. 1. Azimuth estimation errors as functions of the azimuth angle under $\varphi = 56.8°$, $\theta = 45°$, $\psi = 0°$

Fig. 2. Azimuth estimation errors as functions of the elevation angle under $\varphi = 56.8°$, $A = 45°$, $\psi = 0°$

To estimate errors, the method of statistical simulation was used as well. Latitude and attitude angles were simulated as uniformly distributed random variables, instrumental errors as normally distributed random variables with zero mathematical expectations. Several series of runs with 15 000 in each one were carried out. Numerical characteris-

tics (3σ) of azimuth determination errors induced by FOG biases with the level (3σ) of 0.6°/h were estimated. Corresponding series of runs were intended 1) for assigned ranges of attitude angles, 2) for all elevation angles (except for close to zero), 3) for singularity points of the first method (elevation angles close to zero), 4) for singularity points of the second method ($\Omega_x = \Omega_N \cdot \cos A \cdot \cos\theta + \Omega_H \cdot \sin\theta = 0$). The results of estimation are shown in Table I.

TABLE I

Azimuth determination errors estimated with the help of statistical simulation (3σ), °

Series of runs No	Simulated intervals	First method	Second method
1	$\varphi:[10°,80°]$; $A:[3°,355°]$; $\theta:[45°,80°]$; $\psi:[-5°,+5°]$	0.52	0.53
2	$\varphi:[0°,80°]$; $A:[0°,90°]$; $\theta:[0°,87°]$; $\psi=0°$	9.31	0.96
3	$\varphi:[0°,70°]$; $A:[0°,90°]$; $\theta:[0°,3°]$; $\psi=0°$	47.5	2.17
4	$\Omega_N \cdot \cos A \cdot \cos\theta + \Omega_H \cdot \sin\theta = 0$, $\varphi:[0°,70°]$, $\psi=0°$	0.365	0.358

As it can be seen from Table I, the methods are close in accuracy under the assigned range of elevation angle. When expanding the range of elevation angle, it is the second method that has the advantage of accuracy. It is interesting that both methods have the same accuracy for singularity points of the second method.

Experimental estimation of azimuth determination errors was carried out using the equipment of JSC "Scientific and Production Association of Electromechanics" (Miass). Misalignments between device axes and the reference azimuth direction were about several arc min. Before experiments, the device was calibrated, which is worth dedicating a special paper. The results of error estimation are shown in Table II. As it can be seen from the Table, the second method is more accurate according to testing results.

TABLE II

Azimuth determination errors estimated in testing, °

Reference azimuth	Before calibration		After calibration	
	1 method	2 method	1 method	2 method
140°	−6.7	−5.3	−1.17	−0.48
230°	−2.6	−1.9	1.86	0.84
320°	3.1	2.2	1.58	0.20

VI. CONCLUSION

The results of the conducted study have confirmed a possibility of gyrocompassing with the help of an SIMU including a triad of accelerometers and two ARS. The methods considered in the paper are close in accuracy but have their own singularity points, which is to be taken into consideration when choosing a method. By and large, these methods can be used simultaneously to control device operation and improve azimuth estimation accuracy.

REFERENCES

[1] Paderina, T.V. Development of a Strapdown Inclinometer with an Angular Rate Sensor for Arbitrary Oriented Wells, Cand. Sci. Dissertation, St. Petersburg, 2005.

[2] Vodicheva, L., Alievskaya, E., and Parysheva, Yu., Initial Alignment of a Strapdown Inertial Navigation System on a Stationary Base: Methods and Their Errors, 20th St. Petersburg International Conference on Integrated Navigation Systems, St. Petersburg, CSRI Elektropribor, 2013, pp. 93–96.

[3] DT0059 Design tip. ST life augmented. Ellipsoid or sphere fitting for sensor calibration. October 2018. www.st.com

The Influence of Power Supply Switching Frequency Noise on the Fiber Optic Gyroscope Bias Stability

Gang Wang
No.16 Institute The 9th Academy
China Aerospace Science and
Technology Corporation
Xi'an, China
wanggang040727@126.com

Bo Xie
No.16 Institute The 9th Academy
China Aerospace Science and
Technology Corporation
Xi'an, China
xiebo0001@163.com

ShitaoJi
No.16 Institute The 9th Academy
China Aerospace Science and
Technology Corporation
Xi'an, China
Jishitao167773298@1263.com

Honggang Chen
No.16 Institute The 9th Academy
China Aerospace Science and
Technology Corporation
Xi'an, China
cnki1985@163.com

Wei Hong
No.16 Institute The 9th Academy
China Aerospace Science and
Technology Corporation
Xi'an, China
hongwei735@163.com

Hanrui Yang
School of Automation Engineering
Northeast Electric power
University
Jilin, China
Yanghanrui1208@163.com

Abstract—**Fiber optic inertial measurement units (IMU) primarily utilize high-precision fiber optic gyroscopes. However, the fiber optic gyroscopes power supply can introduce switching frequency noise through the data processing circuit power supply, which is then amplified by the long-distance transmission cable within the inertia group. This noise, present in the P-Intrinsic-N Field Effect Transistor （PIN-FET） of the fiber optic gyroscopes data processing circuit, manifests as random voltage spikes on the component side. Consequently, when using analog to digital (AD) sampling to solve the rotation phase in fiber optic gyroscopes, random phases are introduced, leading to a reduction in bias-stability. To address this issue, we analyzed the power supply channel in the inertial group, considering the signal path and characteristics. By utilizing actual measured power supply switching frequency noise data, we established a mathematical model to evaluate the impact of this noise on the gyroscopes bias. Through simulation analysis, we discovered that when subjected to a power switching frequency amplitude of 120mVp-p, the bias-stability of the fiber optic gyroscope decreased from 0.0012°/h (100s, 1δ) to 0.0026°/h (100s, 1δ). Further analysis revealed that for optimal performance of a high-precision fiber optic gyroscope, the power supply switching frequency should not exceed 70mVp-p. To meet the specified requirements, we implemented a filter to regulate the power input of the fiber optic gyroscope. As a result of this filtering process, the amplitude of the fiber optic gyroscope's response to the power switching frequency does not exceed 60mVp-p. During the gyroscopes bias-stability test conducted under these conditions, it was observed that the bias-stability of the fiber optic gyroscopes increased by more than 20%. This indicates that the implemented filter is an effective measure in reducing power switching frequency interference. To ensure optimal performance, it is recommended that the power supply switching frequency should not exceed 70mVp-p.**

Keywords—High-precision fiber optic gyroscope, Power switching frequency noise, filter, Bias-stability

I. INTRODUCTION

As a new all-solid-state gyroscope, fiber optic gyroscope has broad application prospects in many military fields such as inertial guidance, satellite positioning, and inertial navigation self-alignment due to its advantages of high reliability, fast startup, lightweight, and low cost[1]. At present, domestic high-precision fiber optic gyroscopes are generally on the order of 0.001°/h. As high-precision fiber optic gyroscopes develop towards higher precision and are gradually used in large quantities in inertial groups, high-precision fiber optic gyroscopes have increasingly higher power requirements. Performance has become one of the main factors affecting the engineering application of high-precision fiber optic gyroscopes.

Research on fiber optic gyroscope power supply systems at home and abroad mainly focuses on the suppression of power supply noise and the distributed power supply of the system. A dedicated power filter network was developed for the noisy fiber optic gyroscope power supply system. At the same time, the crosstalk between the gyroscopes modulation and demodulation systems was reduced through the distributed power supply design of the gyroscopes system [2]; for medium and low precision fiber optic gyroscopes, Research The influence of the duty cycle changing characteristics of the switching power supply on the accuracy of the load gyroscopes when the switching power supply is working [3] is focused on the causes of the spike noise of the switching power supply and the mechanism of its impact on the performance of the fiber optic gyroscope. On this basis, it is proposed that slew rate control is an applicable Low-noise power supply for fiber optic gyroscope system power supply [4]. However, the above studies did not provide a clear solution to the mechanism and measures for increasing the bias error of fiber optic gyroscope due to the switching frequency characteristics of the switching power supply and the crosstalk introduced by the power supply transmission cable.

To ensure the zero-bias performance of high-precision fiber optic gyroscopes, the power supply for fiber optic gyroscopes is required to be miniaturized, have a large input

voltage range, low noise, low ripple, and stable performance. At present, domestic switching power supplies have zero requirements for high-precision fiber optic gyroscopes. The biased influence needs to be further improved. This paper combines the high-precision fiber optic gyroscopes power supply path and its signal characteristics to analyze the impact mechanism of switching power supply noise on the fiber optic gyroscopes signal acquisition circuit, and on this basis, uses a designed filter to complete the fiber optic gyroscopes switching power supply filtering process. A certain type of three-self strap-down inertial unit is used to build a test platform to verify the zero-bias performance of the fiber optic gyroscope before and after filtering the switching power supply.

II. HIGH-PRECISION FIBER OPTIC GYROSCOPE POWER SUPPLY PATH AND ITS SIGNAL CHARACTERISTICS

The power supply channel for high-precision fiber optic gyroscopes and its signal characteristics are crucial for their performance in inertial units. Typically, the power supply channel for fiber optic gyroscopes involves a DC 26~30V input, which is then converted to a DC ±5V output using a DC/DC power module. This ±5V power is then distributed through the inertial unit to power various components such as the photoelectric conversion P-Intrinsic-N Field Effect Transistor(PIN-FET)components, preamplifier, AD sampling circuit, FPGA circuit, DA output circuit, and Amplified Spontaneous Emission(ASE) light source. Due to space constraints within the inertial unit, the power cable may intersect with the rotating motor and other power supply cables, potentially causing electromagnetic coupling. The quality of the power supply at the entrance of the fiber optic gyroscope is critical as it directly impacts the stability of the aforementioned devices. The power supply path structure of the fiber optic gyroscope in the inertia group is illustrated in Fig. 1.

Fig. 1. Schematic diagram of the fiber optic gyroscope power supply path structure.

The fiber optic gyroscope power supply path is outlined in the schematic diagram, with the analog part of the circuit featuring analog-to-digital conversion (AD), digital-to-analog conversion (DA), amplifier circuits, and ASE light source drive circuits. The digital section consists of the FPGA digital signal processing unit and communication module. The switching power supply provides power to all components of the system, but power supply noise and cable noise can introduce interference in the circuit module. In a digital closed-loop fiber optic gyroscope system, noise in the sampling signal channel can impact gyroscopes output zero-bias performance and stability.

III. POWER SWITCHING FREQUENCY INTERFERENCE MODEL

A. Power supply noise sources

Power supply noise in fiber optic gyroscope circuit typically originates from switching power supplies, which can manifest as power supply switching frequency noise and peak burr noise due to the switching frequency. Fig. 2 displays typical noise test outcomes resulting from the measured switching frequency at the PIN-FET terminal. The noise generated by the switching frequency of the power supply being tested is 100mVp-p, with a switching rate of 352 kHz.

Fig. 2. Typical switching power supply noise test results at the gyroscopes PIN-FET

The measurement results in Fig. 2 show that the test power supply switching frequency noise of PIN-FET is approximately 200mVp-p. This noise is amplified nearly 2 times through the inertia power supply cable when compared to the power supply switching frequency.

B. Mathematical model of power supply switching frequency interference.

The output signal of the fiber optic gyroscopes detector is a square wave signal, while the interference from the switching frequency of the fiber optic gyroscopes power supply can be considered a sinusoidal signal. By applying the demodulation principle of the fiber optic gyroscope, formula (1) can be derived.

$$\Delta\varphi(t) = (V_2 + \sin(f_0 t_1)) - (V_1 + \sin(f_0 t_2))$$
$$= (V_2 - V_1) + (\sin(f_0 t_1)) - \sin(f_0 t_2)) \tag{1}$$

Where the first term, V_1 and V_2, are the two levels of the square wave signal respectively, and the second term is the zero-bias of the fiber optic gyroscope caused by the power supply switching frequency noise.

To further examine the impact of power supply switching frequency noise on the bias of fiber optic gyroscope, let $\Delta\varphi(t) = \varphi_s(t) - \varphi_F(t)$ and conduct a Fourier transform on the detector signal free from power supply interference. This will be as in (2):

$$U(t) = \frac{4V_0 \sin(\Delta\varphi(t))}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{2n-1} \cos\left[(2n-1)f_0 t\right] \tag{2}$$

The switching frequency interference signal introduced by the power supply, as in:

$$n_p = n_0(t) + n_w(t) \tag{3}$$

Where: $n_0(t)$ is the white noise, and $n_w(t)$ is the noise of the switching power supply.

The fiber optic gyroscope demodulation process, as in:

$$R(t) = \lim_{t \to \infty} \frac{1}{2T} \int_{-T}^{T} s_1(t) s_2(t-\tau) dt$$

$$= \lim_{t \to \infty} \left(\frac{1}{2T} \int_{-T}^{T} i_1(t) s_2(t-\tau) dt \right)$$

$$+ \lim_{t \to \infty} \left(\frac{1}{2T} \int_{-T}^{T} n_0(t) s_2(t-\tau) dt \right) \quad (4)$$

$$+ \lim_{t \to \infty} \left(\frac{1}{2T} \int_{-T}^{T} n_w(t) s_2(t-\tau) dt \right)$$

Where:

$$s_1(t) = U(t) + n_p(t)$$

$$= \frac{4V_0 \sin(\Delta\varphi(t))}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{2n-1} \cos[(2n-1)f_0 t] \quad (5)$$

$$+ n_0(t) + A\sin(f_w t)$$

$$s_2(t) = S \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{2n-1} \cos[(2n-1)f_0 t] \quad (6)$$

In (4): the first term is the actual demodulation output value of the fiber optic gyroscope. Since $n_0(t)$ is white noise, the second term is a small quantity. The third term is the fiber optic gyroscope output error caused by the power supply switching frequency noise switching frequency noise, which is set to R_e. Its essence is that the interference of the detector signal caused by the power supply switching frequency is multiplied by the demodulation signal, and it is expanded, that is:

$$R_e = n_w(t) s_2(t)$$

$$= A\sin(f_w t) s \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{2n-1} \cos[(2n-1)f_0 t] \quad (7)$$

$$= \frac{AS}{2} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{2n-1} \left\{ \begin{array}{l} \sin[f_w + (2n-1)f_0]t + \\ \sin[f_w - (2n-1)f_0]t \end{array} \right\}$$

IV. EFFECT OF POWER SUPPLY SWITCHING FREQUENCY ON GYROSCOPES BIAS PERFORMANCE

A. Simulation of the influence of power supply switching frequency on gyroscopes bias-stability.

Simulation analysis based on power supply switching frequency interference model. The power supply switching frequency is 352KHz, while the fiber optic gyroscopes modulation frequency is 68KHz. The fiber optic gyroscopes is not affected by the power supply switching frequency (amplitude: 120mVp-p), but it is interfered with by the switching frequency. The comparison results of the output are illustrated in Fig 3.

The noise band of the zero bias of the gyroscopes output widens noticeably when the gyroscopes data is affected by the power switching frequency, as shown in Figure 3. According to the simulation results, the gyroscopes bias-stability (100s smoothing) is 0.0012°/h when not interfered

with by the power switching frequency. However, this stability decreases to 0.0026°/h when the power switching frequency interference is present, causing the gyroscopes bias-stability to more than double.

Fig. 3. Simulation of switching power supply noise interfering with gyroscopes data

B. Analysis of the mechanism of the influence of power supply switching frequency on gyroscopes bias-stability.

According to the high-precision fiber optic gyroscope power supply path and its signal characteristics, the fiber optic gyroscope is extremely susceptible to interference from the switching frequency of the inertial group's power supply to the gyroscopes when used in fiber optic inertial measurement unit, resulting in each gyroscopes signal acquisition input terminal (PIN-FET) The signal superposition during the transit time interferes with the trip point. A certain type of gyroscope is affected by power supply noise in the inertial group.

The power supply switching frequency has a significant impact on the input PIN-FET signal of the gyroscopes closed-loop demodulation, amplifying the amplitude of the sampled original signal level by 2 to 3 times, and the above influence shows random distribution. After the above errors are solved by gyroscopes sampling, they will bring random phase errors, which will eventually lead to the deterioration of the bias- stability of the fiber optic gyroscope.

Through simulation analysis, it can be seen that the power switching frequency has a significant impact on the bias-stability of the gyroscope. To give full play to the accuracy of the high-precision fiber optic gyroscope, the power switching frequency of the high-precision fiber optic gyroscope signal input end (PIN-FET end) needs to be controlled within 70mVp-p. To achieve the index requirements without changing the high-precision fiber optic gyroscope circuit and optical path structure, it is necessary to add appropriate filter suppression to the fiber optic gyroscope power input end.

V. POWER FILTER DESIGN

A. Power filter principle

The noise generated by the switching power supply can be amplified as it travels through the power lines of different electrical devices. This noise interference can be categorized into common mode interference and differential mode interference. Common mode interference refers to an unwanted potential difference between any current-carrying conductor and a reference ground, while differential mode interference is an unwanted potential difference between any two current-carrying conductors. In applications where the power line supplying power to the gyroscope in an inertial unit is long, exceeding 1m, such as in a three-axis strap-down inertial unit, the noise from the switching power supply can disrupt the coupling of the motor and its control frequency through the lengthy cable. The interference path of the fiber optic gyroscope is illustrated in Figure 4, based on the classification and route of interference.

Fig. 4. Common mode (CM) noise and differential mode (DM) noise.

Simulation results indicate that switching power supply noise can impact the bias-stability of fiber optic gyroscopes, as well as the coupling of power cables with multiple external noise sources. To address this issue, a filter is incorporated into the power input of the fiber optic gyroscope to eliminate external interference frequencies such as the power switching frequency and the power cable coupling motor control frequency. The insertion loss curve of the designed filter shows an attenuation of -25dB to -30dB for the 352kHz frequency and its high-frequency harmonics, effectively suppressing interference from the power switching frequency and power cable coupled motor control frequency.

B. Filter verification.

A high-precision fiber optic gyroscope was utilized to construct a test platform for evaluating the filtering efficacy of the designed filter. The outcomes are detailed in Table 1, while the test findings regarding the switching power supply noise from the PIN-FET terminal are illustrated in Figure 5.

Fig.5 demonstrates that the interference at the PIN-FET end of the fiber optic gyroscope decreased from 100mVp-p in Fig.2 to 60mVp-p after filtering the high-precision power supply inlet signal. Subsequently, the bias- stability of the fiber optic gyroscope was evaluated, with the test outcomes outlined in Table 1.

Table 1 illustrates that incorporating a filter gyroscopes at the output end of the power module has enhanced bias-stability, particularly evident in the significant improvements in the 10s and 100s data.

TABLE I. COMPARISON OF BIAS-STABILITY TEST RESULTS OF A CERTAIN TYPE OF FIBER OPTIC GYROSCOPE POWER SUPPLY WITH AND WITHOUT FILTER

Fiber optic gyroscope status		Bias-stability(°/h，1δ)			
		0.1s	1s	10s	100s
X-Axis	Unfiltered	0.052	0.014	0.0052	0.0030
	Filter	0.050	0.014	0.0045	0.0018
Y-Axis	Unfiltered	0.378	0.052	0.0073	0.0019
	Filter	0.328	0.048	0.0048	0.0016
Z-Axis	Unfiltered	0.315	0.041	0.0068	0.0024
	Filter	0.262	0.037	0.0051	0.0020

Fig. 5. Test results of power supply interference at the PIN-FET end of the gyroscope after filtering

VI. CONCLUSION

This study delves into the impact of switching frequency noise from the switching power supply on high-precision fiber optic gyroscopes systems. By analyzing the signal characteristics of the fiber optic gyroscopes power supply channel and the propagation path of this noise, a mathematical model is established to simulate its effect on the zero-bias performance of the gyroscopes. The findings reveal that interference from the external 100mVp-p power supply significantly reduces the bias-stability of the fiber optic gyroscope. Implementing a filter to mitigate this interference results in a substantial improvement in the bias-stability of the gyroscope by over 20%.

References

[1] Wang Wei High-precision fiber optic gyroscope technology, Chinese Journal of Inertial Technology, 2003, vol.16, no. 3, pp, 105-112.

[2] Keith G W, Mead D, Smith D E U.S. Patent 8542364B1, 2013

[3] Zhang Jingpei, Zhang Chunxi, Song Ningfang Research on the influence of switching power supply noise on the accuracy of fiber optic gyroscope, Chinese Journal of Inertial Technology, 2003, no. 1 pp, 46-49.

[4] Chen Xian, Yang Jianhua, Zhou Lan Application of a low-noise switching power supply in fiber optic gyro systems Optoelectronic Engineering, 2018, vol.45, no.1, pp, 34-42.

[5] Yao Linjie, Wang Yaobin, Jiang Wei Selection and design of EMI filter core for switching power supply, Mechanical and Electrical Engineering Technology, 2023, vol.52, no.5, pp, 253-257.

[6] Wang Na, Zhao Huanyu Ripple and noise analysis of switching power supply module[. Railway Communication Signals, 2021, vol.57, no.4, pp, 18-21.

[7] Tao C B. Fog noise analysis and suppression technology research, PhD thesis, Harbin Engineering University.

[8] Blank M, Gluck T, Kugi A Digital slew rate and s-shape control for smart power switches to reduce EMI generation, ITransactions on Power Electronics, 2015, vol.30, no.9, pp, 5170–5180.

[9] Li R J Study on inhibition of DC/DC EMI converter based on FPGA, Popular Science & Technology, 2013, vol.15, no.10, pp, 72–74.

Research on Input Axis Misalignment Angle Error Suppression Technology for FOGs Using the Fiber Coil with Skeleton

Longgang Li

XI'an Aerospace Precision Electromechanical Institute China Aerospace Science and Technology Corporation
Xi'an, China
598619798@qq.com

Bin Ren

XI'an Aerospace Precision Electromechanical Institute China Aerospace Science and Technology Corporation
Xi'an, China

Jiawei Li

XI'an Aerospace Precision Electromechanical Institute China Aerospace Science and Technology Corporation
Xi'an, China

ruifeng Xu

XI'an Aerospace Precision Electromechanical Institute China Aerospace Science and Technology Corporation
Xi'an, China

Yu Liu

XI'an Aerospace Precision Electromechanical Institute China Aerospace Science and Technology Corporation
Xi'an, China

Fengjun Li

XI'an Aerospace Precision Electromechanical Institute China Aerospace Science and Technology Corporation
Xi'an, China

Yihua Wang

XI'an Aerospace Precision Electromechanical Institute China Aerospace Science and Technology Corporation
Xi'an, China

Abstract—**We use the skeleton with special surface treatment and three kinds of adhesives (A, B and C) with different parameters to make fiber coils. By measuring fiber coil loss, fiber coil polarization crosstalk and input axis misalignment angle of FOGs, it is determined that C adhesive can effectively reduce the input axis misalignment angle error. Furthermore, according to the stress analysis curve of fiber coils with C adhesive and different coiling tensile forces (6.5g, 8.5g and 10.5g), the internal stress of fiber coil with the coiling tensile force of 8.5g is minimum. In the actual production process, the qualification rate of the input axis misalignment angle of FOGs has increased from 53.2% to 92.8%, which improves by 39.6%.**

Keywords—fiber optic gyroscope, fiber coil, error suppression, adhesive, input axis misalignment angle

I. INTRODUCTION

The fiber optic gyroscope (FOG) is a kind of all solid-state inertial instrument, which is different from traditional mechanical inertial instrument. It has the characteristics of simple structure, low cost, small volume, and light weight, and is widely used in the fields of aerospace, aviation and weapon systems [1]. The FOG inertial navigation system has high requirements for the change input axis misalignment angle caused by the temperature, which is also one of the important factors affecting the performance of the FOG inertial navigation system. The error of input axis misalignment angle will cause zero drift error and scale factor error, which will affect the FOGs precision [2].

Fiber coils with skeleton are mainly composed of quartz (fiber core and cladding), polymer materials (fiber coating and adhesive), and the skeleton. The general preparation process is using the quadrupole symmetrical winding method to wind the optical fiber on the skeleton, brushing adhesive while winding or winding the optical fiber after passing through the adhesive groove, and then curing and shaping. The FOG is composed of fiber coils with skeleton, circuit boards, and structural components, etc. This article mainly focuses on the research of the method to suppress input axis misalignment angle error of FOGs with the skeleton.

II. MECHANISM OF INPUT AXIS MISALIGNMENT ANGLE FOR FOGS

The equivalent plane normal of the fiber coil composed of multiple-turn fibers is the input axis (IA) of the fiber optic gyroscope. The gyroscope output is maximized when the input axis is parallel to the carrier's rotational angular velocity direction. Ideally, the gyroscope input axis should be parallel to the normal of the installation surface, namely the input reference axis (IRA). However, the sensing unit fiber coil of the FOG is composed of multiple-turn coils, and its input axis is the result of the vector sum of the plane of each turn of the optical fiber coil. Due to inherent issues of quadrupole symmetrical winding method and manufacturing process defects, IA generally does not coincide with IRA. The angle between IA and IRA is usually defined as the input axis misalignment angle γ[3-4],which is shown in Fig. 1.

The main reasons for the misalignment angle error of FOGs used the fiber coil with skeleton include errors of the gyroscope structure, errors of the fiber coil with skeleton, and installation errors of the fiber coil with skeleton [5], etc. The misalignment angle error can be divided into constant

error and temperature sensitive error. The constant error is mainly caused by factors such as the accuracy of the gyroscope platform structure, the flatness of the gyroscope installation surface, and installation errors of the fiber coil with skeleton. And it does not change as the environment changes. The temperature sensitive error refers to the variation of the misalignment angle of the FOG as the temperature changes, and the main factor causing temperature sensitive errors is the error of the fiber coil with skeleton. In this paper, by surface treatment on the skeleton, choosing the appropriate adhesive of fiber coil and choosing the appropriate coiling tensile force, the error of the fiber coil with skeleton is suppressed. Moreover, it will reduce input axis misalignment angle error of FOGs.

Fig. 1. Schematic diagram of input axis misalignment angle

III. SPECIAL SURFACE TREATMENT ON THE SKELETON

Three kinds of adhesives (A, B and C) are chosen to make fiber coils. After high-temperature aging and temperature cycling, the volume of adhesives shrinks to a certain extent. The volume shrinkage rates are 1.28% for A adhesive, 1.07% for B adhesive, and 0.85% for C adhesive. This is because the cured adhesive undergoes physical aging during temperature cycling, with a very small amount of solvent and small molecules volatilizing from the adhesive, and the molecular chain segments gradually occupying the free voids left by the evaporation of solvents and small molecules through thermal motion. This results in a reduction in the free volume of the adhesive and a slowing down of molecular chain segment movement. After a certain temperature cycling, the group reactions or filling of free voids end, and the entire cured adhesive enters a relatively stable state where the volume no longer changes. To solve the issues of volume shrinkage of adhesives and uneven brushing adhesives during winding, and to reduce the input axis misalignment angle error of FOGs, we adopt special surface treatment on the skeleton.

There are two special surface treatment process sequences between sulfuric acid anodizing and sandblasting for the skeleton (2A12 aluminum alloy). One method is sandblast followed by anodizing, while the other method is to anodize followed by sandblasting. Both sandblasting and sulfuric acid anodizing have an impact on the surface roughness of the skeleton. What's more, they can increase the surface roughness, which will be beneficial for adhesive bonding. In the adhesive strength tests for these two processes, the skeleton treated with anodizing before sandblasting exhibits higher adhesive strength compared to the skeleton treated with sandblasting before anodizing. This is because sulfuric acid anodization is a chemical method for treating metal surfaces. If sandblasting is conducted before

anodizing, it will destroy the chemical groups on the metal surface polished by sandblasting, hindering the chemical reaction or chemical bonding between the metal surface and the adhesive. Therefore, anodizing before sandblasting is conducive to increasing the adhesive strength between the adhesive and the metal.

We choose five skeletons that have undergone anodization followed by sandblasting, and clean the skeletons with organic solvents and then dry them to ensure the cleanliness of the skeleton surface. The cross-sectional diagram and physical image of the skeleton are shown in Figure 2. Use a coordinate measuring machine to measure the flatness and perpendicularity of the skeleton. The test data in Table I shows that the flatness of the installation surface, the flatness of the parallel surface, and the perpendicularity between the bottom and side surfaces of the skeleton, which meet the design requirements. And the surface treatment of the skeleton does not affect the flatness and perpendicularity.

a)   b)

Fig. 2. The skeleton: a) cross-sectional diagram, b) physical image

TABLE I. SKELETON FLATNESS AND PERPENDICULARITY

Skeleton number, #	Installation surface flatness(A-side), mm	Parallel surface flatness, mm	Perpendicularity of bottom and side surfaces, mm
1	0.0046	0.0032	0.0033
2	0.0049	0.0048	0.0039
3	0.0041	0.004	0.0048
4	0.0044	0.0037	0.0038
5	0.0049	0.0046	0.0047

Applying anodizing followed by sandblasting to the surface of the skeleton can effectively enhance the bonding strength between the fiber coil and the skeleton, which resists the shrinkage of the fiber coil. So, it can avoid to exist weak gaps between the fiber coil and the skeleton and reduce misalignment angle error.

IV. EXPERIMENT

A. Impact of different adhesives on the misalignment angle of FOGs

Three kinds of adhesives (A, B and C) with different parameters have been chosen, which are viscosity, glass transition temperature, curing shrinkage rate, thermal expansion coefficient, elastic modulus. They are shown in Table II.

The viscosity of the three adhesives is between 3500 mPa•s and 5000 mPa•s, which is suitable for use as adhesives of fiber coils. The glass transition temperature should avoid the temperature range of -40°C to 60°C in the FOG. The thermal expansion coefficients of three adhesives are close, which means that different adhesives have similar effects on the fiber coils. The curing shrinkage rate and

elastic modulus of A adhesive and C adhesive are smaller than those of B adhesive.The stresses generated during the curing of adhesive include shrinkage stress and adhesive stress, where the shrinkage stress mainly depends on the volume shrinkage of adhesives, and adhesive stress mainly depends on the elastic modulus of adhesives[6]. Choosing the adhesive,which has a smaller elastic modulus and curing shrinkage rate, as well as a suitable thermal expansion coefficient, can reduce the force between the fiber coil and the skeleton. Therefore, it will reduce the deformation of the fiber coil and effectively suppress the input axis misalignment angle error of FOGs.

TABLE II.   PARAMETERS OF DIFFERENT KINDS OF ADHESIVES

Parameters	A adhesive	B adhesive	C adhesive
Viscosity, (mPa • s, 25℃)	3502	4556	3721
Glass transition temperature, ℃	-50	-54	-63.5
Curing shrinkage rate, %	1.22	1.58	0.93
Thermal expansion coefficient, (10⁻⁶/K, 25℃)	226	223	248
Elastic modulus, (MPa, 25℃)	1.78	2.1	1

Use three kinds of adhesives (A, B, and C) to make the fiber coils (under the same conditions: the coiling tensile force of 6.5g), which contain 12 coils of each adhesive and totally 36 coil samples. According to the data in Fig. 3, the average loss of the three kinds of coils is between 0.16 dB and 0.18 dB at room temperature, and there is no obvious difference. By measuring fiber coil polarization crosstalk across the full temperature range from -45°C to 75°C, it is determined that fiber coils with A and C adhesives have the same performance (the average of the maximum polarization crosstalk is about -19.55dB). However, the performance of the fiber coil with B adhesive is slightly inferior (the average of the maximum polarization crosstalk is about -16.46dB), as shown in Fig. 3. The possible reason may be that the curing shrinkage rate and elastic modulus of A and C adhesives are smaller than those of B adhesive. During the temperature change process, the stress on the optical fiber is smaller, resulting in smaller the maximum polarization crosstalk of fiber coils across the full temperature range.

Fig. 3.   Fiber coils with different adhesives: a) the loss, b) the maximum polarization crosstalk

The input axis misalignment angle of the gyroscope with different adhesives is shown in Table III. And the specification requirement for the input axis misalignment angle of this type gyro is less than 5′. Before temperature cycling, the qualification rate of the input axis misalignment angle of FOGs is 58.3% with A adhesive; 75% with B adhesive; and 91.6% with C adhesive. After temperature cycling, the qualification rate of the misalignment angle of FOGs with each adhesive has not changed. And the absolute value of the maximum variation of misalignment angle for FOGs is 0.93′ with A adhesive; 1.11′ with B adhesive; and 0.78′ with C adhesive.

TABLE III.   THE INPUT AXIS MISALIGNMENT ANGLE OF THE GYROSCOPE WITH DIFFERENT ADHESIVES

Different adhesives	Input axis misalignment angle(′)		Variation for misalignment angle (′)
	Before temperature cycling	After temperature cycling	
A adhesive	3.912	3.257	-0.655
	0.890	1.432	0.542
	3.480	4.411	0.931
	-0.384	0.484	0.868
	-8.185	-8.441	-0.256
	-11.283	-10.887	0.396
	9.114	9.687	0.573
	3.880	4.524	0.644
	0.198	-0.128	-0.326
	-5.178	-5.633	-0.455
	-2.984	-2.323	0.661
	7.866	8.264	0.398
B adhesive	-10.079	-9.337	0.742
	-3.592	-3.806	-0.214
	-2.973	-4.080	-1.107
	-2.287	-2.556	-0.269
	-0.401	-1.138	-0.737
	4.225	4.989	0.764
	6.125	6.792	0.667
	2.450	2.004	-0.446
	-1.220	-0.697	0.523
	3.412	2.899	-0.513
	8.784	8.028	-0.756
	-2.010	-1.355	0.655
C adhesive	2.655	2.069	-0.586
	-3.179	-2.445	0.734
	1.811	1.339	-0.472
	-0.832	-0.486	0.346
	1.733	2.284	0.551
	9.126	8.636	-0.490
	-0.216	-0.856	-0.640
	-2.411	-1.630	0.781
	4.416	4.826	0.410
	0.594	0.682	0.088
	1.75	1.254	-0.496
	1.251	0.884	-0.367

B. Impact of different coiling tensile forces on misalignment angle of FOGs

Use C adhesive to make the fiber coils with different coiling tensile forces (6.5g, 8.5g and 10.5g). By measuring fiber coil loss and polarization crosstalk, it is determined that fiber coils with coiling tensile forces of 6.5g and 8.5g have the same performance (the average loss of fiber coils with 6.5g tension is 0.16 dB, and the average of the maximum polarization crosstalk over the full temperature range is -19.68 dB; the average loss of fiber coils with 8.5g tension is 0.19 dB, and the average of the maximum polarization crosstalk over the full temperature range is -20 dB). However, the performance of the fiber coil with the coiling tensile force of 10.5g is slightly inferior (the average loss of fiber coils with 10.5g tension is 0.35 dB, and the average of the maximum polarization crosstalk over the full temperature range is -17.43 dB), as shown in Figure 4.

a)   b)

Fig. 4. Fiber coils with different coiling tensile forces: a) the loss, b) the maximum polarization crosstalk

Before temperature cycling, the average misalignment angle of FOGs with the coiling tensile force of 8.5g is smallest among them, which is 1.02′ (the average misalignment angle of FOGs with the coiling tensile force of 6.5g is 1.90′, the average misalignment angle of FOGs with the coiling tensile force of 10.5g is 3.11′), as shown in Figure 5. According to the stress analysis curve of fiber coils with skeleton, the internal stress of fiber coil with the coiling tensile force of 8.5g is minimum, as shown in Figure 6.

Fig. 5. The misalignment angle of FOGs with different coiling tensile forces

Fig. 6. Stress analysis curve of fiber coils with different coiling tensile forces at different temperature

After temperature cycling and random vibration, there is no abnormality in misalignment angle of FOGs with the coiling tensile force of 8.5g. Finally, the reliability of this state has been fully proved.

V. CONCLUSION

In conclusion, by special sandblasting treatment on the skeleton, choosing the C adhesive of fiber coils and choosing the coiling tensile force of 8.5g, the input axis misalignment angle error of FOGs is effectively suppressed. Moreover, the qualification rate of the input axis misalignment angle of FOGs has increased from 53.2% to 92.8%, which improves by 39.6%.

REFERENCES

[1] Barhour, N., Schmidt, G., "Inertial sensor technology trends", IEEE Sensors Journal, no. 1(4), 2001, pp, 332 - 339.

[2] Yang, M.Y., Song, Z.Q., Zhang, X.L., and Chen, Y.Z., "Temperature characteristics of input axis misalignment angle of fiber optic gyro", Journal of National University of Defense Technology, 2014,36(3): 49-50.

[3] GJB 2426A-2004, Methods for fiber optic gyroscope test, Beijing: the Commission of Science, Technology and Industryfor National Defense，2004．

[4] Wang, H., Yang, D.W., "Research on the stability of the input axis of fiber optic gyroscope sensitivity coil", Modern Electronic Technology, 2012, 35(20): 137-139.

[5] Ji, S.T., Qin, Y.Y., Lan, Y., and Liu, Z.R.,"Research on temperature compensation for input axis misalignment angle of fiber optic gyroscope", Measurement and Control Technology, 2010, 29(10): 6-8.

[6] Zhang, S.Y., Zhang, G.C., Ma, L., and Yang, Y., "Error analysis and improved ways for fiber gyroscope output in shock vibration and half cardinals movement experiments", Saint Petersburg international conference on integrated navigation systems, 2016, pp, 579 - 583.

Research on the Technology for Improving Scale Factor Stability of High Precision Fiber Optic Gyroscope

Pei Zhang
XI'an Aerospace Precision
Electromechanical Institute
China Aerospace Science and
Technology Corporation
Xi'an, China
1184526321@qq.com

Yichen Wu
XI'an Aerospace Precision
Electromechanical Institute
China Aerospace Science and
Technology Corporation
Xi'an, China

Wei Ding
Beijing Institute of Control
and Electronic Technology
China Aerospace Science and
Industry Corporation Limited
Beijing, China

Honggang Chen
XI'an Aerospace Precision
Electromechanical Institute
China Aerospace Science and
Technology Corporation
Xi'an, China

Wei Hong
XI'an Aerospace Precision
Electromechanical Institute
China Aerospace Science and
Technology Corporation
Xi'an, China

Bo Huang
XI'an Aerospace Precision
Electromechanical Institute
China Aerospace Science and
Technology Corporation
Xi'an, China

Longgang Li
XI'an Aerospace Precision
Electromechanical Institute
China Aerospace Science and
Technology Corporation
Xi'an, China

Wei Jiang
XI'an Aerospace Precision
Electromechanical Institute
China Aerospace Science and
Technology Corporation
Xi'an, China

Yu Bai
XI'an Aerospace Precision
Electromechanical Institute
China Aerospace Science and
Technology Corporation
Xi'an, China

Abstract—**Firstly, the mathematical model of scale factor of the fiber optic gyroscope (FOG) is studied in this paper, and the main cause of scale factor changes in variable temperature environment is identified as variations in the length and diameter of fiber coil. To solve this problem, a method based on finite element simulation analysis is proposed to calculate the equivalent thermal expansion coefficient of fiber coil, and the finite element simulation analysis model for fiber coil is established on this basis. Then, the simulation analysis platform is established for the variation in scale factor of FOG caused by deformation of fiber coil in variable temperature environment. The influence of material properties of each component of the fiber coil encapsulation structure on the scale factor is quantitatively analyzed. The simulation results show that the deformation of fiber coil under variable temperature conditions can be effectively suppressed by reducing the thermal expansion coefficient of the winding adhesive and skeleton, as well as the elastic modulus of the winding adhesive. Thereby the stability of scale factor is improved. Finally, the material properties of the fiber coil encapsulation structure are optimized and experimentally verified in variable temperature (-40~+70) °C environment. The experimental results show that the stability of scale factor of the FOG is improved from 54.6ppm to 23.2ppm, which is of great significance for further realizing the engineering of high precision FOG.**

Keywords—fiber optic gyroscope (FOG), scale factor, fiber coil, finite element simulation, optimal design

I. INTRODUCTION

With the continuous improvement of the level of FOG devices and related technologies, as well as the increasingly mature manufacturing process, FOG is constantly developing towards high precision. As the main performance indicator of FOG, scale factor is very sensitive to temperature disturbances. The stability of scale factor will decrease due to the change of ambient temperature, which is one of the main factors limiting the performance and application of high precision FOG at present, and is also the main shortcoming of FOG compared with laser gyroscope [1-3].

At present, the stability of scale factor is improved by optimizing light source, temperature compensation for the scale factor and optimization of the software algorithm. W.S. Feng proposed a method to improve the temperature performance of the scale factor based on the wavelength control of light source[4]. H.Y. Chen analyzed the influence of the average wavelength of the wide spectrum light source on the scale factor, established the scale factor system model of the open-loop FOG and the closed-loop FOG, and verified the correctness of the system model[5]. Y.P. Zhang studied a scale factor temperature compensation scheme that combines software and hardware, which can reduce the temperature error of the scale factor[6]. Y.J. Wu analyzed the relationship between temperature and scale factor based on closed-loop FOG, established the corresponding formulas, and provided temperature coefficients for corresponding parameters[7]. Other studies have shown that by optimizing the software algorithm of FOG and using the second closed-loop control to compensate for the scale factor of FOG, the stability of the scale factor of FOG can also be effectively improved[8].

In this paper, the mathematical model of the scale factor of FOG is studied and analyzed, and the main reason of the scale factor change in the variable temperature environment

is determined. The influence of material physical property parameters of each component of fiber coil encapsulation structure on the scale factor of fiber optic gyroscope is quantitatively analyzed, and a scheme to improve the stability of the scale factor of FOG is proposed and verified.

II. THEORY

The scale factor of FOG is the proportional factor between the digital output of the gyroscope and the input angular velocity, which can be expressed as:

$$K_{SF} = \frac{2\pi LD}{\lambda c} \cdot \frac{1}{K_m K_{DA}} \tag{1}$$

where

L length of fiber coil,

D diameter of fiber coil,

λ average wavelength of light source,

C speed of light in vacuum,

K_m Y-waveguide modulation coefficient,

K_{DA} D/A conversion and amplifier gain.

It can be seen from equation (1) that the scale factor of FOG is related to the length and diameter of fiber coil, the average wavelength of light source, and the product of Y-waveguide modulation coefficient and D/A conversion and amplifier gain. At present, the amplified spontaneous emission (ASE) light source is usually used in high precision FOG, and the average wavelength of ASE light source can be controlled by temperature control and constant current measures in full temperature environment. Moreover, the Y-waveguide modulation coefficient, D/A conversion and amplifier gain are tracked and measured in real time by optimizing the gyro software algorithm, at the same time, the feedback channel proportion coefficient is adjusted to ensure that the feedback channel gain of the gyroscope remains unchanged in variable temperature environment.

However, as the core component of FOG, fiber coil is composed of fiber and winding adhesive, which is fixed in the Skeleton by bonding adhesive. The temperature changes will cause the thermal expansion and cold contraction deformation of each component of fiber coil encapsulation structure, thereby changing the length and diameter of fiber coil. The variation in scale factor of FOG caused by the product of the length and diameter of fiber coil over the full temperature range is more than 1000ppm, which is the main factor affecting the stability of scale factor. Therefore, in order to improve the stability of the scale factor of FOG, it is necessary to effectively suppress the product variation of the length and diameter of fiber coil.

III. FINITE ELEMENT SIMULATION MODEL OF FIBER COIL

A. 3D Model of Fiber Coil

The fiber coil used in high precision FOG is skeleton free fiber coil. It's a hollow cylindrical structure with rectangular cross section is shown in Fig. 1.

The fiber consists of a cladding layer and a coating layer. The diameter of fiber cladding is d1, the diameter of fiber coating is d2, and the equivalent fiber diameter after coating with winding adhesive is d. The rectangular cross section of the fiber coil wound with 8-pole symmetrical method is shown in Fig. 2.

Fig. 1. 3D model of fiber coil

Fig. 2. Rectangular cross section of fiber coil wound with 8-pole symmetrical method

B. Analysis of Physical Parameters of Fiber coil

In order to facilitate the model calculation, it is necessary to consider the fiber coil as a whole. Based on ANSYS finite element simulation and analysis software, a cube model with rectangular section of fiber coil (width W, height H) as the main view and length L is established. The minimum temperature and maximum temperature under the condition of temperature variation are set as T_1°C and T_2°C respectively, and the average elongation ΔW in the width direction of the cube model is calculated. The thermal expansion coefficient of fiber coil can be expressed as:

$$\alpha = \frac{\Delta W}{W \times (T_2 - T_1)} \tag{2}$$

For a certain type of fiber coil wound by the 8-pole symmetrical method, the physical parameters of fiber coil are calculated based on the wrapping parameters, the above calculation method of thermal expansion coefficient, and the volume proportion of each component, as shown in Table I.

TABLE I.   PHYSICAL PARAMETERS OF FIBER COIL

Physical Parameters	Cladding	Coating	Winding Adhesive	Fiber coil
Height(mm)	-	-	-	13.5
Inner diameter (mm)	-	-	-	65
Number of layers	-	-	-	64
Number of turns	-	-	-	98
Outer diameter (mm)	0.08	0.135	-	-
Thermal expansion coefficient (×10⁻⁶/°C)	0.54	12	65	12.76
Volume proportion (%)	30.7	56.7	12.6	100
Density (g·cm⁻³)	2.65	1.2	1	1.62
Elastic modulus (GPa)	70	15	1	30.12
Poisson's ratio	0.168	0.38	0.42	0.32

IV. SIMULATION ANALYSIS OF SCALE FACTOR VARIATION CAUSED BY FIBER COIL DEFORMATION

A. Encapsulation Model of Fiber Coil

The skeleton free fiber coil is fixed in the skeleton by bonding adhesive, and the cover plate is welded to the skeleton to complete the encapsulation of the fiber coil. The encapsulation model of fiber coil is shown in Fig. 3.

Fig. 3. Encapsulation model of fiber coil

B. Finite Element Simulation Analysis of Fiber Coil Encapsulation Structure

The finite element simulation analysis of fiber coil encapsulation structure is carried out by ANSYS software. The environmental temperature ranged from -45°C to +70°C, and the physical parameters of fiber coil encapsulation structure are shown in Table II.

TABLE II.   PHYSICAL PARAMETERS OF FIBER COIL ENCAPSULATION STRUCTURE

Physical parameters	Fiber coil	Bonding adhesive	Skeleton	Cover plate
Density (g·cm⁻³)	1.62	1.1	2.77	2.77
Thermal expansion coefficient ($\times 10^{-6}$/℃)	12.76	60	23	23
Elastic modulus (GPa)	30.12	1.2	71	71
Poisson's ratio	0.32	0.38	0.33	0.33

The simulation result of fiber coil deformation is shown Fig. 4.

Fig. 4. Simulation result of fiber coil deformation

C. Analysis of Scale Factor Variation

The equivalent fiber diameter can be expressed as:

$$d = \frac{H}{n+0.5} \tag{3}$$

The length of fiber coil can be expressed as:

$$L = mn\pi \left[D_{inner} + d + \frac{\sqrt{3}(m-1)}{2}d \right] \tag{4}$$

Where

D_{inner} inner diameter of fiber coil.

The diameter of fiber coil can be expressed as:

$$D = D_{inner} + d + \frac{\sqrt{3}(m-1)}{2}d \tag{5}$$

The variations of inner diameter and outer diameter of coil are shown in equation 6 and 7 respectively.

$$\Delta D_{inner} = \frac{1}{g} \sum_{i=1}^{g} \sqrt{X_{inner-i}^2 + Y_{inner-i}^2} \tag{6}$$

$$\Delta D_{outer} = \frac{1}{k} \sum_{j=1}^{k} \sqrt{X_{outer-j}^2 + Y_{outer-j}^2} \tag{7}$$

Where

$i = 1,2,\cdots g$. g is the number of nodes on the inner cylindrical surface of fiber coil,

$j = 1,2,\cdots k$. k is the number of nodes on the outer cylindrical surface of fiber coil,

$X_{inner-i}$ deformation of node i on the inner cylindrical surface of fiber coil in the X direction,

$Y_{inner-i}$ deformation of node i on the inner cylindrical surface of fiber coil in the Y direction,

$X_{outer-j}$ deformation of node j on the outer cylindrical surface of fiber coil in the X direction,

$Y_{outer-j}$ deformation of node j on the outer cylindrical surface of fiber coil in the Y direction.

The variation of equivalent fiber diameter can be expressed as:

$$\Delta d = \frac{\Delta D_{outer} - \Delta D_{inner}}{\sqrt{3}(m-1)+2} \tag{8}$$

The scale factor variation caused by fiber coil deformation can be expressed as:

$$\Delta K_{SF} = \frac{L \times \Delta D + D \times \Delta L + \Delta L \times \Delta D}{LD} \tag{9}$$

Where

ΔL variation in the length of fiber coil,
ΔD variation in the diameter of fiber coil.

V. ANALYSIS OF THE INFLUENCE OF MATERIAL PROPERTY PARAMETERS ON SCALE FACTOR

The thermal expansion coefficient α (from 0.5 times to 2 times) and elastic modulus E (from 0.5 times to 2 times) of each component of fiber coil encapsulation structure are adjusted under the condition of keeping the other physical property parameters unchanged, and the scale factor change of FOG caused by the fiber coil deformation at varying temperature (-45~+70)°C is calculated. The simulation results are shown in Table III.

TABLE III. INFLUENCE OF PHYSICAL PROPERTY PARAMETERS ON SCALE FACTOR

Physical property parameters			LD product of fiber coil (mm²)	Variation of scale factor (ppm)	Relative variation (ppm)
α	Winding adhesive	0.5 times	104186073	1340.2	536.9
		2 times	104241937	1877.1	
	Bonding adhesive	0.5 times	104204557	1517.9	3.8
		2 times	104204958	1521.7	
	Skeleton	0.5 times	104185467	1334.4	555.2
		2 times	104243230	1889.6	
	Cover plate	0.5 times	104203802	1510.6	25.8
		2 times	104206480	1536.4	
E	Winding adhesive	0.5 times	104195573	1431.5	185.8
		2 times	104214904	1617.3	
	Bonding adhesive	0.5 times	104204924	1521.4	-3.9
		2 times	104204512	1517.5	
	Skeleton	0.5 times	104201570	1489.2	56.2
		2 times	104207419	1545.4	
	Cover plate	0.5 times	104204225	1514.7	8.7
		2 times	104205130	1523.4	

It can be seen from the analysis results that by reducing the thermal expansion coefficient of the winding adhesive and skeleton, as well as the elastic modulus of the winding adhesive, the deformation of fiber coil under variable temperature environments can be effectively reduced, and the stability of scale factor of FOG under variable temperature environments can be improved.

VI. EXPERIMENT

According to the application environment of FOG, the adhesive with low expansion coefficient ($4.2 \times 10^{-5}/°C$) and low elastic modulus (0.6GPa) is used as the winding adhesive, and the soft magnetic alloy with low expansion coefficient ($1.05 \times 10^{-5}/°C$) is used as the skeleton and cover material. The stability of scale factor of FOG with temperature compensation is improved from 54.6ppm to 23.2ppm, and the experimental results are shown in Fig. 5.

Fig. 5. The experimental results

VII. CONCLUSION

The mathematical model of scale factor of FOG is studied in this paper, and the main cause of scale factor changes in variable temperature environment is identified as variations in the length and diameter of fiber coil. The simulation analysis platform is established for the variation in scale factor of FOG caused by the deformation of fiber coil in variable temperature environment on this basis. The influence of material properties of each component of the fiber coil encapsulation structure on the scale factor is quantitatively analyzed. The scale factor stability of high precision FOG is improved effectively by optimizing the physical property parameters of the fiber coil encapsulation structure.

REFERENCES

[1] W. Wang, Technology of interference fiber optical gyroscope, Beijing, China Aerospace Press, 2010.

[2] Vittorio P, Antonello C, Lorenzo V, et al, Gyroscope technology and applications: A review in the industrial perspective, Sensors, 2017, vol. 17, no. 2284, pp. 1-22.

[3] Y.L. Na, L.K Li, Y.J Wu, et al, Measurement error analysis of FOG's scale factor, Journal of Chinese Inertial Technology, 2012, vol. 20, no. 4, pp. 472-477.

[4] W.S. Feng, W. Wei, H.C. Yu, et al, Temperature Performance Improvement Method for Fiber Optic Gyroscope Based on Wavelength Control, Navigation and Control, 2019, vol. 18, no. 4, pp. 46-51.

[5] H.Y. Chen, Y. Zheng, X.X. Wang, et al, Influence of Mean Wavelength on Scale Factor of Fiber Optic Gyroscope, Chinese Journal of Lasers, 2019, vol. 46, no. 3, pp. 0306002-1-0306002-7.

[6] Y.P. Zhang, Z.J. Pan, Z.W. Wei, et al, Hardware implementation of temperature compensation for FOG's scale-factor. Journal of Chinese Inertial Technology, 2013, vol. 21, no. 5, pp. 660-662.

[7] Y.J. Wu, X.L. Huang, Study on Temperature Stability of Scale-Factor in Closed-Loop Fiber Optical Gyroscope, Transactions of Beijing Institute of Technology, 2007, vol. 27, no. 7, pp. 618-620.

[8] S.T. Chen, F. Sun, X.Y. Li, Compensation for phase drift in an integrated optical phase modulator. Journal of Harbin Engineering University, 2008, vol. 29, no. 1, pp. 45-49.

Research On The Technology For Suppressing Magnetic Field Error Of High Precision Fiber Optic Gyroscope

Yichen Wu
XI'an Aerospace Precision
Electromechanical Institute
China Aerospace Science and
Technology Corporation
Xi'an, CHINA
18829043344@163.com

Pei Zhang
XI'an Aerospace Precision
Electromechanical Institute
China Aerospace Science and
Technology Corporation
Xi'an, CHINA

Wei Ding
Beijing Institute of Control
and Electronic Technology
China Aerospace Science and
Industry Corporation Limited
Beijing, CHINA

Wei Hong
XI'an Aerospace Precision
Electromechanical Institute
China Aerospace Science and
Technology Corporation
Xi'an, CHINA

Zewei Pan
XI'an Aerospace Precision
Electromechanical Institute
China Aerospace Science and
Technology Corporation
Xi'an, CHINA

Bo Huang
XI'an Aerospace Precision
Electromechanical Institute
China Aerospace Science and
Technology Corporation
Xi'an, CHINA

Yunjiao Li
XI'an Aerospace Precision
Electromechanical Institute
China Aerospace Science and
Technology Corporation
Xi'an, CHINA

Hui Cao
XI'an Aerospace Precision
Electromechanical Institute
China Aerospace Science and
Technology Corporation
Xi'an, CHINA

Yu Bai
XI'an Aerospace Precision
Electromechanical Institute
China Aerospace Science and
Technology Corporation
Xi'an, CHINA

Abstract—This article studies the generation mechanism of magnetic field error in fiber optic gyroscope (FOG) and proposes a method to improve the gyro's resistance to magnetic field interference through magnetic shielding technology. Based on the ANSYS finite element analysis software, a simulation analysis platform is built to establish a simulation model of Helmholtz coils to simulate the magnetic field. By changing the current passing through the coils to generate a uniform magnetic field of different intensities, it is used to simulate and analyze the fiber coil under different magnetic field intensities. The simulation analysis results determine that the optimal solution is to use iron-nickel soft magnetic alloy and multi-layer magnetic shielding structure, which can increase the FOG's magnetic shielding effectiveness to over 40dB, meaning that external magnetic fields are attenuated to within 1% of the original field strength after passing through the magnetic shielding structure. Furthermore, an experimental platform is set up for validation specifically for the multi-layer magnetic shielding structure. The experimental results show a decrease in magnetic sensitivity coefficient by 2 orders of magnitude, indicating the significant importance of the double-layer magnetic shielding structure in improving the FOG's resistance to magnetic field interference.

Keywords—fiber optic gyroscope (FOG), magnetic field error, finite element analysis, iron-nickel soft magnetic alloy, multi-layer magnetic shielding structure

I. INTRODUCTION

Interferometric fiber optic gyroscope(IFOG), due to its advantages such as no moving parts, simple process, wide coverage of accuracy, large dynamic range, fast start-up, and long lifespan, are widely used in the defense fields of aerospace, aviation, and maritime. It serves as the strategic and tactical inertial navigation instrument, providing essential support for high-precision navigation of weapon systems and equipment[1,2].

Fiber optic gyroscope (FOG) is an inertial device that uses a fiber coil as a sensitive component and utilizes the Sagnac effect to measure the angular velocity of light paths. The output angular rate signal is proportional to the non-reciprocal phase shift of the two counter-propagating beams inside the fiber optic coil[3].When non-reciprocal phase shift occurs in the FOG under non-rotating conditions, it can lead to errors in the gyroscope, resulting in a decrease in its accuracy. Magnetic field interference is one of the main error sources for FOGs. Changes in the external magnetic field can cause zero bias errors in the FOG. Depending on the direction of the magnetic field, the mechanisms causing errors differ and can be classified as radial magnetic field error and axial magnetic field error. Radial magnetic field error is caused by the Faraday effect and fiber twisting, while axial errors are due to non-reciprocal phase shift effects and fiber bending[4].

Currently, methods to suppress the magnetic field error in fiber optic gyroscopes mainly include using polarization-maintaining fibers, adding depolarizers, software compensation, and adding magnetic shielding structures outside the fiber coil. Among these methods, using polarization-maintaining fibers and employing magnetic shielding structures are the most commonly used suppression measures. Regarding magnetic shielding, the magnetic shielding effectiveness is an important performance indicator used to measure the quality of shielding structures. For FOGs, the single-layer magnetic shielding structure can only weaken the magnetic field around the fiber coil by tens of times, which still has a significant impact on the accuracy of the FOG. Therefore, optimization design of the magnetic shielding structure is necessary.

This article, based on ANSYS finite element analysis software, establishes a simulation analysis platform, comprehensively designs an optimal solution for a multi-layer magnetic shielding considering various influencing factors, and sets up an experimental platform to verify its

effectiveness. This approach plays a guiding role in comprehensive analysis of the impact of magnetic field interference on FOGs and provides important basis for future designs of magnetic shielding structures for FOGs.

II. THE MECHANISM OF MAGNETIC FIELD ERRORS IN FIBER OPTIC GYROSCOPE

The magnetic fields that affect FOGs are divided into two categories: one is the environmental magnetic field generated by the Earth's magnetic field, and the other is the electromagnetic field generated by the device near the gyroscope. This article mainly discusses the impact of the environmental magnetic field generated by the Earth's magnetic field on FOGs. The environmental magnetic field can be decomposed into radial magnetic field and axial magnetic field. The error caused by the magnetic field on fiber optic gyroscopes can be regarded as the sum of the effects of these two orthogonal magnetic fields. The following introduces the mechanisms by which these two types of magnetic fields affect the accuracy of FOGs[5].

Radial magnetic-induced error refers to the error in gyroscope accuracy caused by the magnetic field perpendicular to the sensitive axis of the gyroscope obtained by decomposing the Earth's magnetic field. This error mainly originates from the Faraday effect and the twisting of the fiber during drawing and winding processes. The Faraday effect occurs when light propagates through a medium, and if there is a parallel magnetic field along the direction of light propagation, the polarization plane will be rotated during the propagation of this light[6]. The relevant formula is shown as follows：

$$\psi = VBL \tag{1}$$

where

V — the Verde constant

B — the intensity of the environmental magnetic field

L — the propagation length of the fiber

The Faraday effect is shown in Fig. 1:

Fig. 1. Faraday effect in optical fibers

The radial magnetic error of the FOG is related to two factors: the Faraday effect and fiber twist.

Axial magnetic induced error refers to the error caused by the magnetic field component parallel to the sensitive axis of the gyro obtained by resolving the Earth's magnetic field, when the axial magnetic field acts on the fiber coil. In the process of winding the fiber optic coil, a small helix angle a will be generated relative to the sensitive axis of the fiber in each turn due to the helical angle of the fiber coil. The quasi-TM mode (q-TM mode) in the polarization-maintaining fiber will experience a non-reciprocal phase shift effect due to the bending of the fiber[7]. Decompose the axial magnetic field B into components parallel to the direction of light propagation and perpendicular components. The helix angle a can be expressed as

$$\alpha = \tan^{-1}\left(\frac{D}{2\pi MR}\right) \tag{2}$$

where

D the width of the fiber coil;

M the number of turns;

R the radius of the fiber coil.

The spiral angle is shown in Fig. 2:

Fig. 2. Spiral angle during fiber coil wrapping process

The axial magnetic error of the FOG is related to two factors: non-reciprocity phase shift and fiber bending.

III. MAGNETIC SHIELDING STRUCTURE DESIGN

Magnetic shielding technology is a type of electromagnetic shielding. Unlike electromagnetic shielding materials that absorb or reflect electromagnetic waves, magnetic shielding diverts the magnetic field. Typically, soft magnetic alloys are used for this purpose. The technology involves adding a structure composed of high magnetic permeability soft magnetic alloy outside the fiber coil to divert and guide the magnetic field passing through the fiber coil, concentrating the magnetic field lines within the shielding structure and preventing them from entering the protected area of the shielding structure, thereby achieving shielding effect. The diversion principle is similar to electrical resistance diversion, so circuit methods can be used to equivalently calculate the shielding effectiveness of the magnetic shielding structure[8]. Shielding effectiveness is an important parameter used to quantify the shielding effect. For an ideal cylindrical shell, its shielding effectiveness is:

$$SE = 20\lg\frac{H_1}{H_2} \tag{3}$$

where:

H_1 the magnetic field strength of the internal structure before shielding；

H_2 the magnetic field strength of the internal structure after shielding.

The schematic diagram of magnetic shielding is shown in Fig. 3.

a） b）

Fig. 3. Magnetic Shielding Principle: a)Structure Diagram of Magnetic Shielding Principle; b) Equivalent Magnetic Circuit

In response to the characteristics of magnetic shielding, a shielding structure design is applied to the fiber coil. The single-layer magnetic shielding consists of the fiber coil skeleton, fiber coil cover plate, and lining panel in terms of structure. Laser welding is used for fixed connection between the fiber coil skeleton and the fiber coil cover plate, enclosing the fiber coil in the protected area , as shown in Fig. 4.

Fig. 4.   Single-layer magnetic shielding structure

In terms of structure, the double-layer magnetic shielding consists of the coil skeleton, fiber coil cover plate, lining panel, upper magnetic shielding shell and lower magnetic shielding shell. Laser welding is used for fixed connection between the inner fiber coil skeleton and the fiber coil cover plate. The outer upper shielding shell is fixed on the lining panel with screws, while the outer lower shielding shell is fixed on the pillar with screws. Additionally, the outer upper and lower shielding shells are fixed together through laser welding, as shown in Fig. 5：

Fig. 5.   Double-layer magnetic shielding structure

IV.   SIMULATION OF MAGNETIC SHIELDING STRUCTURE

Due to the complexity of the structure, the magnetic shielding effectiveness cannot be accurately obtained through theoretical calculations. It is necessary to build a simulation analysis platform using simulation software to conduct simulation analysis. In this study, a Helmholtz coil with a diameter of 1m and a spacing of 0.25m was simulated to analyze the magnetic shielding structure. By adjusting the orientation of the fiber coils, the simulation can mimic the axial and radial magnetic induction intensities of the fiber coils in a uniform magnetic field, and calculate the magnetic shielding effectiveness. The simulation analysis platform is established as shown in Fig. 6.

For the radial uniform magnetic field, the magnetic induction intensity of the fiber coil is shown in Fig.7. Without a magnetic shielding structure, the magnetic induction intensity of the fiber coil is 0.0016T. When using a single-layer magnetic shielding structure, the magnetic field strength decreases to 7.3×10^{-5}T. The external magnetic field attenuates to 4.6% of the original magnetic field strength after passing through the single-layer magnetic shielding structure, and the magnetic shielding effectiveness is

26.82dB. When using a double-layer magnetic shielding structure, the magnetic field strength decreases to 9.09×10^{-6}T. The external magnetic field attenuates to 0.57% of the original magnetic field strength after passing through the double-layer magnetic shielding structure, and the magnetic shielding effectiveness is 44.91dB. Compared to single-layer magnetic shielding, the double-layer magnetic shielding shows an increase in magnetic shielding effectiveness of 18.09dB, and the magnetic induction intensity decreases by an order of magnitude.

Fig. 6.   Simulation Analysis Platform: a) Radial Magnetic Field Simulation Analysis Platform;b) Axial Magnetic Field Simulation Analysis Platform

Fig. 7.   Simulation results of radial magnetic field：a) Magnetic induction intensity without magnetic shielding，b) Magnetic induction intensity during single-layer magnetic shielding，c) Magnetic induction intensity during double-layer magnetic shielding

For the axial uniform magnetic field, the magnetic induction intensity of the optical coil is shown in Fig.8.

Fig. 8.   Simulation results of radial magnetic field：a) Magnetic induction intensity without magnetic shielding，b) Magnetic induction intensity during single-layer magnetic shielding，c) Magnetic induction intensity during double-layer magnetic shielding.

Without a magnetic shielding structure, the magnetic induction intensity of the fiber coil is 0.0016T. When using a single-layer magnetic shielding structure, the magnetic field strength decreases to 4.79×10^{-5}T. The external magnetic field attenuates to 3% of the original magnetic field strength after passing through the single-layer magnetic shielding structure, and the magnetic shielding effectiveness is 30.48dB. When using a double-layer magnetic shielding structure, the magnetic field strength decreases to 5.96×10^{-6}T. The external magnetic field attenuates to 0.37% of the original magnetic field strength after passing through the double-layer

magnetic shielding structure, and the magnetic shielding effectiveness is 48.58dB. Compared to the single-layer magnetic shielding, the magnetic shielding effectiveness of the double-layer magnetic shielding is increased by 18.1dB, and the magnetic induction intensity decreases by an order of magnitude.

V. VERIFICATION OF MAGNETIC SHIELDING STRUCTURE TEST

Aiming at the effect of double-layer magnetic shielding structure on improving the accuracy of FOG, an experimental platform is built. By adjusting the current size and generating different magnetic field strengths, the magnetic sensitivity coefficient of FOG is tested under different magnetic field strengths. The formula for the magnetic sensitivity coefficient β is as follows:

$$\beta = \frac{P_0 - P_i}{H_0 - H_i} \qquad (4)$$

where

P_0 the bias value of the gyroscope in the geomagnetic environment

P_i the bias value of the gyroscope under different magnetic fields

H_0 the strength of the geomagnetic field

H_i the set magnetic field strength

By averaging the measured magnetic sensitivity coefficients, the FOG's magnetic sensitivity coefficient in the magnetic field is obtained. The radial magnetic sensitivity coefficient is shown in Table 1, and the axial magnetic sensitivity coefficient is shown in Table 2.

TABLE I.   RADIAL MAGNETIC SENSITIVITY COEFFICIENT

Structural form	*No magnetic shielding*	*single-layer magnetic shielding*	*double-layer magnetic shielding*
Magnetic sensitivity coefficient (° /h)/Gs)	0.013	0.00077	0.00043

TABLE II.   AXIAL MAGNETIC SENSITIVITY COEFFICIENT

Structural form	*No magnetic shielding*	*single-layer magnetic shielding*	*double-layer magnetic shielding*
Magnetic sensitivity coefficient (° /h)/Gs)	0.2	0.00076	0.00031

Based on the test results, it is found that FOGs with magnetic shielding structures have a significantly lower magnetic sensitivity coefficient compared to those without magnetic shielding structures. The magnetic sensitivity coefficient of FOGs with double-layer magnetic shielding structures decreases by more than 40% compared to single-layer structures, leading to a significant improvement in FOG accuracy.

VI. CONCLUSION

This article starts with suppressing the magnetic field error of FOG, it optimizes the design of the FOG structure and proposes a double-layer magnetic shielding structure for FOG. Through theoretical analysis, simulation analysis, and experimental verification, a thorough analysis of the magnetic shielding structure is conducted. It is concluded that when the FOG is subjected to an axial or radial magnetic field, the double-layer magnetic shielding structure can effectively reduce the intensity of external magnetic field induction. The magnetic shielding effectiveness can reach above 40dB, meaning that the attenuation of external magnetic fields is less than 1% of the unshielded state. The magnetic sensitivity coefficient of FOGs decreases by two orders of magnitude compared to the unshielded state. The double-layer magnetic shielding structure provides reliability assurance for FOGs operating in external magnetic field environments, which is of great significance in enhancing the anti-magnetic field interference ability of FOG.

REFERENCES

[1] Nayak J, Fiber-optic gyroscopes: from design to production, Applied Optics, 2011, vol.50, no.25, pp. 152-161.

[2] G.C.Zhang, Fiber Optic Gyroscope Principles and Technology, National Defense Industry Publishing, Beijing, 2008.

[3] W.Wang, Interferometric Fiber Optic Gyroscope Technology, China Aerospace Publishing, Beijing, 2010.

[4] X.Y.Li, P.Liu, X.X.Guang, et al, Temperature Dependence of Transverse Magnetic Error in a Polarization Maintaining Fiber Optic Gyroscope, Acta Optica Sinica, 2018, vol.38, no.1, pp. 0106006.

[5] Y.Z.Chen, X.X.Wang, Y.Y.Gao, et al, Research on the influence mechanism of earth's magnetic field on zero bias of high precision FOG, Electronic Measurement Technology, 2016, vol.39, no.1, pp. 147-149.

[6] Kumar P, Purohit G, Nayak J, et al. Effect of magnetic field on the performance of Fiber Optic Gyroscope. Journal of Optics, 2012, vol.41, no.4, pp. 231-234.

[7] D.W.Zhang, Z.F.Wang, L.Chen, et al. Nonreciprocal in single-mode fiber coil induced by orthogonal magnetic field. Journal of Optics, 2013, vol.15, no.5, pp. 055706.

[8] P.Liu, Research on Mechanism and Suppression Measures of Magnetic Field Error of Fiber Optic Gyroscope, PhD thesis, Harbin Engineering University, 2018, pp. 77-78.

Improvement of Precision of High-accuracy Fiber Optic Gyroscope Based on the Application of "Three-self motion" Inertial Measurement Unit

Yun-jiao Li
XI'an Aerospace Precision Electromechanical Institut
China Aerospace Science and Technology Corporation
Xi'an, China
liyj870309@126.com

Shao-feng Lou
XI'an Aerospace Precision Electromechanical Institut
China Aerospace Science and Technology Corporation
Xi'an, China
loushaofeng@163.com

Yong-liang Zhao
XI'an Aerospace Precision Electromechanical Institut
China Aerospace Science and Technology Corporation
Xi'an, China
312348152@qq.com

Wei Hong
XI'an Aerospace Precision Electromechanical Institut
China Aerospace Science and Technology Corporation
Xi'an, China
hongwei735@163.com

Yi-chen Wu
XI'an Aerospace Precision Electromechanical Institut
China Aerospace Science and Technology Corporation
Xi'an, China
18829043344@163.com

Gang Wang
XI'an Aerospace Precision Electromechanical Institut
China Aerospace Science and Technology Corporation
Xi'an, China
wanggang040727@126.com

Bo Huang
XI'an Aerospace Precision Electromechanical Institut
China Aerospace Science and Technology Corporation
Xi'an, China
bohuangmail@126.com

Abstract—**Based on the composition characteristics of the "three self motion" inertial measurement unit, a method is proposed to significantly improve the cost-effectiveness of the "three self motion " inertial measurement unit by improving system technology rather than overly demanding inertial devices. Based on a detailed analysis of its internal signal interference and path, the PWM wave interference signal is first introduced according to the closed-loop fiber optic gyroscope simulation model established in Matlab. Subsequently, the correctness of the simulation results was verified through experiments, and measures to suppress interference were proposed based on the experimental results, verifying the feasibility of the method.**

Keywords—"three self motion" inertial measurement unit, fiber optic gyroscope, interference

I. INTRODUCTION

With the increasing demand for inertial navigation systems in weapon systems, the accuracy of the "three self motion" inertial measurement unit (self calibration, self alignment, and self detection functions) is also increasing. Fiber optic gyroscopes, as all solid-state inertial instruments, do not cause additional disturbance to the installation carrier during operation, and have inherent advantages when applied to high-precision "three self motion " inertial combinations. On the basis of realizing the functions of the

strapdown inertial unit, the "Three Self Motion " inertial measurement unit adds a self rotating locking mechanism and high-speed digital processing circuit to receive control instructions from the flight control computer or measurement and control equipment, achieving online self calibration, self alignment, and self detection functions of the strapdown inertial unit. The traditional approach improves the accuracy of inertial instruments (gyroscopes, accelerometers) to achieve higher performance in the "three self motion " inertial measurement unit.

Given the composition characteristics of the "three self motion " inertial measurement unit, namely the existence of a self rotating locking mechanism, the inherent PWM wave signal of the "three self motion " inertial measurement unit always exists during the hot standby process throughout its entire life cycle (storage period, launch preparation stage, and navigation tracking stage), which interferes with the electrical signal in the fiber optic gyroscope and leads to a decrease in the accuracy of the inertial instrument, thereby limiting its performance in the "three self motion " inertial measurement unit. Effectively understanding the internal interference paths of the "three self motion " inertial measurement unit and suppressing them, significantly improving the cost-effectiveness of the "three self motion" inertial measurement unit through system technology improvement rather than excessive demands on inertial devices, may be of great significance for improving the

accuracy of the "three self motion " inertial measurement unit products.[1-3]

II. "THREE SELF MOTION" INERTIAL MEASUREMENT UNIT COMPOSITION CHARACTERISTICS

The "three self motion " inertial measurement unit is a highly integrated product of optics, mechanics, and electronics. It is generally composed of a self rotating locking mechanism, fiber optic gyroscope, accelerometer, signal processing, and self detection circuit. The typical composition diagram of the "three self motion " inertial measurement unit is shown in the figure 1:

Fig. 1. "Three self motion" Inertial Measurement Unit Composition Block Diagram

During the initial self calibration stage, the "three self motion " inertial measurement unit flips multiple positions through a rotating locking mechanism. When flipped to a specified position, the inner and outer frames are locked. The PWM output duty cycle changes from a maximum to a minimum or from a minimum to a maximum, until it reaches the locked state. The PWM wave outputs a fixed duty cycle signal; During the navigation tracking phase, the rotary locking mechanism is in a locked state, and the PWM wave outputs a fixed duty cycle signal. During the hot standby period of the "three self motion" inertial measurement unit throughout its entire lifecycle, its inherent PWM wave signal always exists.[5-6]

Fig. 2. Interference Path

III. ANALYSIS OF DIGITAL CLOSED-LOOP CONTROL FIBER OPTIC GYROSCOPE

The digital closed-loop control fiber optic gyroscope system consists of two parts: a non reciprocal optical path and a digital demodulation circuit.

Fig. 3. FOG Composition Block Diagram

When the fiber optic gyroscope has an angular velocity signal input, under the modulation of the square wave signal, the signal at the interference end will output an error signal of the same frequency as the modulation signal. The PIN-FET photodetector can convert this error optical signal into a corresponding voltage signal. The voltage signal is filtered and amplified, and finally converted into a digital signal through an A/D converter. The digital signals of adjacent half cycles are subtracted to give an error signal, and the error signal is integrated to generate a step increment for feedback control. Based on the step increment, a digital step wave signal is generated. Finally, the D/A converter is used to convert the superimposed value of the modulation signal and the stepped wave signal into an analog voltage signal, which is amplified by an amplifier and ultimately acts on the phase modulator to complete a basic feedback control task. Among them, the square wave bias mentioned is the introduction of a bias working phase shift to ensure the sensitivity of the system; The stepped wave signal is mainly used for feedback control.

The main function of the PIN-FET photodetector is to achieve photoelectric conversion, converting error light signals into corresponding voltage signals, and is the core sensor for measuring the intensity of interference light signals. The implementation of digital closed-loop control involves the A/D converter performing multi-point sampling and summing up the sampled data within each cycle. If there is noise interference on the output signal of PIN-FET photodetector, there is a risk of the mean value of the multi-point sampling and summing up data changing, which leads to the addition of additional errors in the error signal obtained by subtracting the digital signals of the two half cycles, resulting in the degradation of the gyroscope performance indicators. For ease of understanding, the following simplification has been made:

- Gyroscopic transit time（τ）10μ s.

- PWM wave period (t) 125 μ s. The duty cycle is 50%.

- The sampling frequency of the A/D converter is 20MHz, and there are 100 effective sampling points in a flat area.

- The modulation mode is square wave two state modulation.

Fig. 4.   output signal general view of PIN-FET photodetector

$$\nabla V_k = \sum_{i=1}^{n} v_i + \sum_{j=1}^{4} v_j$$

In the formula, ∇V_k for the th crossing time （τ） The error introduced by the cycle and Error accumulative total.

m for numbers of PWM wave high-level sampling points.

n for numbers of PWM wave low-level sampling points. If there are 100 valid sampling points, then m+n=100.

According to the principle of fiber optic gyroscope modulation and demodulation, it is known that the first and second cycles do not introduce error signals when the interference level of the PWM wave introduced is the same; When the jumping time of PWM wave occurs at the transition time （τ） The introduction of PWM waves in the sixth and twelfth cycles shown in the figure leads to additional error signals.

A closed-loop fiber optic gyroscope simulation model was built using MATLAB as shown in figure 5, and the required simulation parameters were set according to the table 1. The introduction of PWM wave interference is implemented in the PIN-FET module.

Fig. 5.   Closed-loop fiber optic gyroscope simulation model

Table 1 Related simulation parameters

Parameter	Value
Average wavelength of light source	1550 nm
conversion efficiency of PIN-FET Photodetector	0.95 A/W
Bias voltage of PIN-FET Photodetector	-1.6 V
cross resistance of PIN-FET Photodetector	40 kΩ
Input optical power of PIN-FET Photodetector	0.45 mW
Fiber optic coil length	2200 m
Equivalent diameter of fiber optic coil	100 mm
half wave voltage of Y-waveguide Integrated Optical Device	3.58 V
input range of A/D converter	2 Vp-p
input angular velocity	8.4°/h

Set PWM amplitude values of 0V (without interference),, and 1V respectively, and simulate the output of the gyroscope. The simulation graph is shown in the figure 6:

a）PWM wave amplitude was 0V

b）PWM wave amplitude was 1V

Fig. 6.   Simulated waveform

From the simulation curve observation, after introducing PWM waves, the output signal noise shows an amplification trend，zoom in to 1.67 times.

Based on the simulation results, a high-precision "three self motion" inertial measurement unit product of a certain type in the laboratory was selected, and the output data of the gyroscope using motor locking (PWM interference of 12.5kHz) and mechanical locking (no motor PWM interference) were compared, as shown in the figure 7:

Fig. 7.   Experimental data of FOG output

Comparison of the numerical result and the experimental data shows the correctness of theoretical analysis.

Due to the limitation of the structure of the "three self motion" inertial measurement unit, PWM wave interference always exists. To reduce the impact of interference during use, it is possible to consider adding shielding measures to the circuit board to weaken the interference of PWM waves on the circuit. Based on this idea, a circuit board shielding cover was designed and corresponding experiments were conducted. The experiments have shown that adding shielding measures can reduce the output noise of the gyroscope to a certain extent.

IV. CONCLUSION

The paper adopts a system design concept. Without adding hardware circuits and software architecture, by analyzing the internal signal interference composition of the system and taking interference suppression measures, the performance of the fiber optic gyroscope can be improved to a certain extent, thereby alleviating the excessive demand for inertial devices and significantly increasing the price ratio of the "three self motion " inertial measurement unit. This has certain guiding significance for engineering practice.

REFERENCES

[1] ShupeD.ThermoopticallyInducedNonreciprocityintheFiberOpticInterferometers.ApplOpt, 1980,19: 654- 655.

[2] Wang Wei, Yu haicheng, Feng wenshuai, et al High Precision Fiber Optic Gyroscope Technology[D]. Beijing, National Defense Industry Press, 2021.

[3] Li jianning Research on Control Technology of High Speed Hollow Cup Motor[D]. University,2023.

[4] V. N. Logozinskii. Magnetically Induced Non-Faraday Nonreciprocity in a Fiber-Optic Gyroscope[J]. Journal of Communications Technology and Electronics, 2006, Vol. 51, No. 7, pp. 836–840.

[5] Wu Junwei, Miao Lingjuan, Wu Yanji, et al. Research on Fast Start of Fiber Optic Gyroscope Driven by SLD Light Source Converter [J]. Navigation, Positioning and Timing, 2019,6 (4): 88-93.

[6] Li Lijing, Yang Hui, Ma Yingjian, et al. Research on high-precision temperature controlled pulse width modulation methods and spectral issues [J]. Infrared and Laser Engineering, 2013,42 (6), 1588-1592.

Mathematical Model of Polarization Error in Fiber Optic Gyroscope

Wei Jiang
No.16 Institute The 9th Academy
China Aerospace Science and
Technology Corporation
Xi'an, China
380992975@qq.com

Honggang Chen
No.16 Institute The 9th Academy
China Aerospace Science and
Technology Corporation
Xi'an, China

Zewei Pan
No.16 Institute The 9th Academy
China Aerospace Science and
Technology Corporation
Xi'an, China

Binjie Lin
No.16 Institute The 9th Academy
China Aerospace Science and
Technology Corporation
Xi'an, China

Wei Hong
No.16 Institute The 9th Academy
China Aerospace Science and
Technology Corporation
Xi'an, China

Hanrui Yang
School of Automation Engineering
Northeast Electric Power
University
Jilin, China

Abstract—**Traditional model of polarization error is established by multiplying the Jones matrices of optical devices directly, whose calculation is large and results are complex. In this paper, an effective method is proposed to solve this problem, by calculating the interference intensity of all coupled wave-trains based on the coherence function of the light source, any parameters don't need to be simplified. Take the practical parameters into the mathematical model; the precision of the fiber optic gyroscope is increased through utilizing a designed broadband laser with 17nm bandwidth instead of the traditional 10nm bandwidth laser, the experimental results and analysis results of this model are in agreement, which shows that the mathematical model is reasonable.**

Keywords—fiber optic gyroscope, polarization error, coupled wave-trains, coherence function

I. INTRODUCTION

The fiber optic gyroscope (FOG) has been widely used in various fields across the world, and a number of brilliant techniques have been specially designed and utilized to improve the FOG performances [1]. Since the polarization error limits the accuracy of FOG, the study on polarization error arose from optical devices and fused points is necessary [2]. Usually, the physical model of polarization error is established by multiplying the Jones matrices of optical devices directly, whose calculations are large and result is complex, some parameters such as the shape of light source spectrum, the effect of the fiber coil can't be reflected in the model [3]. In this paper, an effective method is proposed to solve this problem, which covers almost all the optical parameters and the result doesn't need to be simplified.

II. THEORY

Take the FOG using low-polarization and polarization-maintaining hybrid light path for example, and its configuration is illustrated as Fig.1. Ideally the light input into the multifunction integrated optic circuit (MIOC) is linearly polarized along the transmission axis of the input polarizer. In practice this ideal condition is not met, and the input field has two polarization states [4]. Polarization crosstalk is the main source of optical system polarization error. A crosstalk point separates one optical signal into two orthogonal polarization states, the two orthogonal signals propagate independently. And the two orthogonal signals are transformed into four independent orthogonal coupled waves by the next crosstalk point, and so on. The difference between interference intensity of those coupled wave-trains and the interference intensity of the principle wave-train is the polarization error.

Fig. 1. Illustration of FOG using hybrid light path

III. MODEL ANALYSIS

A. The Coherence Function

The FOG generally uses a broadband light source, and the temporal coherence should be considered when the light signals interfere. The interference power is [5]:

$$I = \frac{I_0}{2}\left[1 + \gamma(\tau)cos\left(2\pi\overline{f}\tau\right)\right] \qquad (1)$$

Where $\gamma(\tau)$ is the coherence function of the light source.

Since the spectrum of the light source measured by spectrometer can't be expressed by formula directly, it is necessary to discretize the spectrum. The Fourier transform of the segmented spectrum is superimposed to obtain a continuous coherence function, and a more accurate coherence function can be obtained [6]. According to the Wiener-Khinchin theorem, the relationship between the coherence function and the power spectral density (PSD) is:

$$\Gamma(\tau) = \int_{-\infty}^{+\infty} a^2(f)e^{i2\pi f\tau}df \qquad (2)$$

According to the spectrum measured by the spectrometer, two adjacent data points are approximated as a wavelength segment, and the continuous Fourier transform is calculated in segments. After Fourier transform and superposition of

each discrete point, the coherence function of the whole spectrum is obtained.

B. The Cross-Coupling in Fiber Coil

The process of winding, curing and bonding will introduce a large number of polarization crosstalk points in the polarization maintaining fiber coil. Fig.2 is the position distribution of polarization crosstalk points in a polarization maintaining fiber coil tested by PXA-1000, the abscissa in the figure indicates the distribution position of the crosstalk points, and the ordinate indicates the coupling strength of the crosstalk points. Each polarization crosstalk point is equivalent to a splice point in turn, and the length of the polarization maintaining fiber between each two polarization crosstalk points is equivalent to the length of the polarization maintaining fiber between each two splice points.

Fig. 2.   Distributions of crosstalk points in a PM fiber coil

The high crosstalk points in Fig.2 are selected as the polarization crosstalk points of the coil in Fig.1, and the parameters are shown in TABLE I.

TABLE I.   DISTRIBUTIONS OF CROSSTALK POINTS IN THE COIL

No. in Fig.1	Fiber Length(m)	Crosstalk(dB)	Equivalent Angle (rad)
a_4	86.3	-40.5	0.0094
a_5	380	-51.5	0.0027
a_6	640	-52.5	0.0024
a_7	773.3	-55	0.0018

C. The Calculation of Coupled Waves-trains

The main purpose of this part is to use MATLAB to automatically generate the amplitude and phase expressions of all coupled waves. The main influences of the amplitude of the coupled optical signal are the equivalent coupling angle and the amplitude suppression ratio of the MIOC. Taking Fig.1 as an example, the distribution of coupling points of clockwise and counterclockwise optical signals is as follows:

$$a_{cw} = [a_1 \ a_2 \ a_3 \ a_4 \ a_5 \ a_6 \ a_7 \ a_8 \ a_9] \tag{3}$$

$$a_{ccw} = [a_1 \ a_9 \ a_8 \ a_7 \ a_6 \ a_5 \ a_4 \ a_3 \ a_2] \tag{4}$$

Where a_i is the equivalent coupling angle along the optical path.

For other more complex optical paths, such as dual-path FOG[3], it is only necessary to adjust the elements in (3) and (4) according to the actual clockwise and counterclockwise optical path. One optical signal will generate two mutually perpendicular components through the coupling point a_i, and there are 512 combinations of the nine crosstalk points. For example, the amplitude of all coupled optical signals propagating in the clockwise direction is represented by a one-dimensional array c_cw[]. The program flow is shown in Fig.3, and the amplitude of the optical signal in the counterclockwise direction is solved similarly.

Fig. 3.   Flow chart of amplitude calculation

For example, (5) is the amplitude of the signal incident from the fast axis of the fiber segment l_1, and (6) is the amplitude of the signal incident from the slow axis of the fiber segment l_1, where d is the polarization degree of the light source, and ε is the amplitude suppression ratio of the MIOC.

$$cx_{cw}(8) = \varepsilon\sqrt{\frac{1+d}{2}}cos(a_4)cos(a_5)cos(a_6)cos(a_7)\cdot$$
$$cos(a_8)cos(a_9)sin(a_1)sin(a_2)sin(a_3) \tag{5}$$

$$cy_{cw}(20) = \sqrt{\frac{1-d}{2}}cos(a_3)cos(a_4)cos(a_6)cos(a_7)\cdot$$
$$cos(a_8)cos(a_9)sin(a_1)sin(a_2)sin(a_5) \tag{6}$$

According to the specific expressions of $cx_{cw}[]$ and $cy_{cw}[]$, the corresponding optical path distribution can be automatically obtained. The path distributions of clockwise and counterclockwise optical signals are:

$$l_{cw} = \begin{bmatrix} l_1 & l_{23} & l_3 & l_4 & l_5 & l_6 & l_7 & l_8 & l_9 & l_{24} \end{bmatrix} \quad (7)$$

$$l_{ccw} = \begin{bmatrix} l_1 & l_{24} & l_9 & l_8 & l_7 & l_6 & l_5 & l_4 & l_3 & l_{23} \end{bmatrix} \quad (8)$$

Equations (7) and (8) can be adjusted according to different optical path structures. The clockwise and fast axis incident light signal is still taken as an example, and the corresponding phase array $gx_{cw}[]$ is obtained according to its amplitude array $cx_{cw}[]$. The program flow is shown in Fig.4.

Fig. 4.   Flow chart of optical paths calculation

Equations (9) and (10) are the optical path corresponding to cx_{cw} (8) and cy_{cw} (20), respectively.

$$gx_{cw}(8) = l_1 n_x + l_3 n_x + l_4 n_y + l_5 n_y + l_6 n_y + \\ l_7 n_y + l_8 n_y + l_9 n_y + l_{23} n_y + l_{24} n_y \quad (9)$$

$$gy_{cw}(20) = l_6 n_x + l_7 n_x + l_8 n_y + l_9 n_y + l_{23} n_x + \\ l_{24} n_x + l_1 n_y + l_3 n_y + l_4 n_y + l_5 n_y \quad (10)$$

D. The Calculation of Phase Error

The expression of the kth column coupled wave signal in the clockwise direction is:

$$E_k = cx_{cw}(k) exp \left[i \left(\varpi_0 t - \beta_0 * gx_{cw}(k) + \varphi_s + \varphi_b \right) \right] \quad (11)$$

where

φ_s the Sagnac-Laue phase ;

φ_b the phase bias;

$\beta_0 = 2\pi/\lambda$ propagation constant of light in vacuum.

The expression of the mth column coupled wave signal in the counterclockwise direction is:

$$E'_m = cx_{ccw}(m) exp \left[i \left(\varpi_0 t - \beta_0 * gx_{ccw}(m) \right) \right] \quad (12)$$

The optical signal transmitted along the fast axis of the gyro optical path is the main optical signal. The interference between them is the main interference, and the intensity is:

$$I_{main} = cx_{cw}(1)^2 + cx_{ccw}(1)^2 + \\ 2 * cx_{cw}(1) cx_{ccw}(1) cos \left(\varphi_s + \varphi_b \right) \quad (13)$$

All clockwise wave trains and counterclockwise interference at the MIOC, and the interference intensity value is:

$$I_{sum} = \sum_{k=1}^{2^n} \sum_{m=1}^{2^n} \left\{ cx_{cw}(k)^2 + cx_{ccw}(m)^2 + 2\gamma(\Delta l_{km}) cx_{cw}(k) * \\ cx_{ccw}(m) cos \left(\beta_0 gx_{ccw}(m) - \beta_0 gx_{cw}(k) + \varphi_s + \varphi_b \right) \right\} \quad (14)$$

where

$\gamma(\Delta l_{km})$ the coherence function of the light source;

$\Delta l_{km} = gx_{ccw}(m) - gx_{cw}(k)$, the optical length difference.

The intensity values after demodulation are ΔI_{main} and ΔI_{sum}. According to the principle of FOG demodulation, the intensity value after demodulation is proportional to the sensitive phase shift of the FOG optical path. The main optical signal of the FOG should be sensitive to the Sagnac-Laue phase. The phase error of the FOG optical path can be obtained as [3]:

$$\frac{\Delta I_{main}}{\Delta I_{sum}} = \frac{\varphi_{main}}{\varphi_{sum}} = \frac{\varphi_s}{\varphi_{sum}} \quad (15)$$

$$\varphi_e = \varphi_{sum} - \varphi_{main} = \varphi_s \frac{\Delta I_{sum}}{\Delta I_{main}} - \varphi_s \quad (16)$$

According to the formula of FOG rotation speed, the phase error value of FOG can be converted into bias error.

IV. MODELING AND EXPERIMENT RESULTS

To verify the accuracy of the modeling, the FOG system is established as Fig.1. Three different kinds of ASE are used separately, whose spectrums are shown in Fig.5. The 3dB bandwidth of light source A (blue line) is 10.6 nm, the 3dB bandwidth of light source B (black line) is 6.5 nm, the 3dB bandwidth of light source C (red line) is 17 nm. Their coherence functions calculated using (2) are plotted in Fig.6. It can be seen from the diagram that the wider the spectral width, the faster the coherence of the light source decreases.

The three light sources are connected to the same fiber coil for static test at room temperature. The test time is more than 4 hours. The test results are shown in TABLE II. The diameter of the fiber coil is 90.5 mm, the length is 1600 m, and the angle of the fusion points are considered as 0.5 degree. The model analyses are consistent with the experimental results. The broadband light source can improve the gyro accuracy, which shows that the model is reasonable.

TABLE II.   SIMULATION COMPARED WITH TESTING RESULTS

Light Source	Calculated Phase Error	Tested Bias Stability (100s, °/h)
A	3.23E-08	0.0034
B	4.27 E-08	0.0041
C	1.69 E-08	0.0031

Fig. 5.   Spectrum of the ASE

Fig. 6.   Autocorrelation function of the ASE

V. CONCLUSION

Based on the interferences among all the coupled waves-trains, combined with the coherence function of the light source, the polarization error model of the FOG optical path is established. The model can be flexibly changed according to different optical path forms and has a wide range of applications. The model takes into account the optical parameters of each optical device, which has reference significance for the selection of device parameters and optical path design, and can truly reflect the polarization error of the optical path of the FOG.

REFERENCES

[1] Zheng,Y., Xu,H.J., Song,J.M., Li,L.J., Zhang,C.X., Excess Relative-Intensity-Noise Reduction in a Fiber Optic Gyroscope Using a Faraday Rotator Mirror. Journal of Lightwave Technology, 2020, Vol.38, no.24, pp.6939-6947.

[2] Li,Z.Z., Li,Z.H., Yao,S., Li,G.H., Research on Influence of Polarization Crosstalk on the Zero Drift and Random Walk of Fiber Optic Gyroscope. Acta Optica Sinica, 2014, Vol.34, no.12, pp.1-11.

[3] Xu,H.J., Zhang,W.Y., Xu,X.B., Song,N.F., Hu,Y.K., Polarization Bias Error Model and Simulation of Fiber-Optic Gyroscope with Double Optical Length. Acta Optica Sinica, 2014,Vol.34, no.10, pp.1-5.

[4] Jacob,N., Michel,J.F., Noise and Bias Error Due to Polarization Coupling in a Fiber Optic Gyroscope. Journal of Lightwave Technology, 2015, Vol.33, no.13,pp.2839-2847.

[5] Hervé C. Lefevre, The Fiber-Optic Gyroscope,Boston, Artech House, 2022, third edition.

[6] Zhao,C.J., Yang,Y.H., Zhang,C.X., Simulation Research for Coherence characteristics of Broad-Band Sources. Chinese Journal of Spectroscopy Laboratory, 2005, Vol.22, no.2,pp.268-272.

Evaluating Temperature Stability Of Power Balance In Fiber-Optic Gyroscope With Compensation Of Relative Intensity Noise Of A Light Source

Maxim Belousov
Scientific and Technical Center
Perm Scientific-Industrial Instrument Making Company
Perm
BelousovMA@pnppk.ru

Anton Krivosheev
Scientific and Technical Center
Perm Scientific-Industrial Instrument Making Company
Perm
KrivosheevAI@pnppl.ru

The paper considers the results of tests of fiber-optic gyroscope with compensation of relative intensity noise of light source. The performance of fiber-optic gyroscope layout in the extended temperature range is evaluated. Decrease of efficiency of the proposed method of compensation in the extended range of ambient temperatures is shown. The analysis of possible ways of solving the problem of increasing the efficiency of the proposed method of compensation of relative intensity noise of light source in the extended range of operating temperatures is carried out.

Key words – fiber-optic gyroscope, light source, relative intensity noise, temperature stability

I. INTRODUCTION

Today, fiber optic gyroscopes (FOG) are increasingly used as part of strapdown inertial navigation systems (SINS) [1-4]. In this connection, a high level of requirements is imposed on the FOG accuracy.

There are several indicators that characterize the accuracy of a gyroscope, such as angle random walk (ARW), bias instability (BI), scale factor (SF), and others. In particular, ARW is an important parameter related to the noise component of FOG readings, which needs to be reduced in order to decrease the availability time and improve the accuracy of the SINS orientation angles.

For this purpose, earlier, specialists of PJSC «PNPPK» carried out work to reduce the ARW level by compensating for excess noise of the light source intensity (Relative intensity noise – RIN) [5]. The compensation method described in [6] was used as a basis. The RIN compensated FOG circuitry was modernized to improve its noise reduction performance [7]. According to the test results of the proposed method, it was noted that the efficiency of RIN compensation depends on the balance of optical powers between the measurement arm and the RIN arm [6,8]. This report summarises the results of a power balance stability study over a temperature range of 0°C to +60°C.

II. INITIAL EXPERIMENTAL CONDITIONS

Figure 1 shows the optical circuit of a RIN compensated fiber optic gyroscope for which the temperature stability of the power balance was investigated.

When setting up the experiment, the optical components of the FOG (except for the light source and electronic processing board) were placed in a temperature chamber. The outputs of both channels were connected to a two-channel signal analyser ERM-202. Power balance equalisation was performed using an attenuator at the beginning of the experiment under normal conditions. Figure 2 shows the temperature cycle set for this experiment,

Fig. 1. Scheme of a fiber-optic gyroscope with light source intensity compensation

III. EXPERIMENT RESULTS

Figure 2a shows the change of channel powers with temperature change. The observed change of powers can be caused by the change of the division factor of the 90/10 balance splitter and the change of the optical loss level in the fiber coil.

Figure 2b shows the calculated variation of the ARW value of the FOG readings taking into account the variation of the FOG arm powers and temperature. At the elevated temperature of 60 °C, where the maximum power imbalance is observed ($P_{MEAS}/P_{RIN} = 0.85$), the ARW value increased by a factor of 1.5. This is due to the reduction of the RIN compensation level. The total ARW level is defined, according to the source [9], as the square root of the sum of the squares of the remaining ARW components:

- Photodetector shot noise;

- Thermal noise of the photodetector preamplifier;

- Photodetector dark current noise;

- Residual value of RIN.

Graphically, the dependence of residual ARW components on temperature is shown in Figure 3. As can be seen from the evaluation results, the efficiency of the RIN compensation scheme decreases when the FOG operates in the extended temperature range due to the change in the power balance. The residual value of ARW_{RIN} is expressed according to eq. 1.

$$ARW_{RIN_{residual}} = ARW_{RIN} \cdot \left| \frac{P_{meas}}{P_{RIN}} - 1 \right|, \tag{1}$$

где ARW_{RIN} − value of the ARW component of light source RIN, P_{meas} − measuring arm power, P_{RIN} − RIN arm power.

(a)

(b)

Fig. 2. Graphical results of the experiment on measuring the temperature dependence of (a) output powers of the arms of the VOG circuit with RIN compensation and (b) calculated value of ARW

Fig. 3. Dependence of ARW components on temperature

To make the RIN compensation effective over the entire temperature range, temperature compensation of the power balance in the FOG with a closed-loop feedback circuit can be implemented.

IV. CAPACITY BALANCING

Adjustment of the power balance is carried out in two stages. At the first stage, according to the optical loss characteristics of the fiber coil, a balanced optical splitter is

selected (e.g. 90/10 or 95/5). In the second step, the balance is fine-tuned. As an example in this work, using an optical attenuator. As mentioned in the source [6], there is an alternative way to fine-tune the power balance by changing the phase modulation depth (φ_m) in the algorithm of the FOG board. The correction $\Delta\varphi(T)$ to the operating depth of phase modulation from temperature is determined by the eq. 2.

$$\Delta\varphi(T) = \arccos\left[\frac{2 \cdot P_{RIN}(T)}{P_0(T)} - 1\right] - \arccos\left[\frac{2 \cdot P_{meas}(T)}{P_0(T)} - 1\right], \quad (2)$$

где $P_0(T)$ − optical power of the measuring arm without auxiliary phase modulation.

Schematically, the principle of changing the working depth of phase modulation of the FOG feedback signal to adjust the power level of the P_{MEAS} measurement arm to the P_{RIN} level is shown in Figure 4. In the feedback signal generated by the FOG processing board, the component $\Delta\varphi(T)$ obtained on the basis of the previously described temperature dependence is added.

Fig. 4. Scheme of power balance adjustment in FOG with RIN compensation with phase modulation depth

REFERENCES

[1] S. Sanders, A. Taranta, S. Mosor, M. Alden, L. Hendry, R. DeMaio, N. Giere, and J. Sewell. "Fiber optic gyros in a high-performance, highreliability inertial reference unit for commercial satellites". Proc. SPIE, 2012, vol.8421, OFS2012 22nd International Conference on Optical Fiber Sensors.

[2] U. Probst, M. Deck, S. Voigt. LFK-150: Development of a high-accuracy Marine Inertial Reference System. Proc. DGON Inertial Sensors and Systems - Symposium Gyro Technology 2017, Karlsruhe, Germany, 19-20 September 2017.

[3] Y. Paturel, V. Rumoroso, A. Chapelon, and J. Honthaas. "MARINS, the first FOG navigation system for submarines". Symposium Gyro Technology, 2006.

[4] Yu.N. Korkishko, V.A. Fedorov, V.E. Prilutskiy, V.G. Ponomarev, I.V. Fedorov, S.M. Kostritskii, I.V. Morev, D.V. Obuhovich, S.V. Prilutskiy, A.I.Zuev, V.K.Varnakov. High-precision inertial measurement unit IMU-5000. 5th IEEE International Symposium on Inertial Sensors and Systems, Lake Como, Italy, 26-29 March 2018.

[5] M. Belousov, A. Krivosheev. Compensation of Excess Intensity Noise of a Light Source in a Fiber-Optic Gyroscope. 30th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS) 2023.

[6] Guattari F., Chouvin S., Moluçon C., Lefèvre. «A Simple Optical Technique to Compensate for Excess RIN in a Fiber-Optic Gyroscpe». IXBlue France - Karlsruhe 2014.

[7] V.S. Ermakov, A.V. Subbotin, A.I. Krivosheev, M.A. Belousov. Optical scheme of FOG to reduce light source noise. Patent RU 203287 U1, 30.03.2021.

[8] Xudong Hu, Wei Hong, Zerun Zang, Zijun Pan, Xiong Yang, Kaixu Liang, Mianzhi Zhang, Xinliang Jia, Hanrui Yang. Theoretical Analysis and a New Optical Scheme for RIN Subtraction. 30th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), 2023.

[9] Jin Jing , Wang XueQin, Song NingFang, Zhang ChunXi. Effect of 60Co-gamma radiation on the random walk error of interferometric fiber optic gyroscopes. Science China Technological Sciences 53(11), November 2010.

Principles of Software Development for Inspecting Parameters of Inertial Measurement Units with Nonorthogonally Oriented Measuring Axes

D. M. Kalikhman

Yuri Gagarin State Technical University of Saratov, Russia, Saratov;
Branch of Academician Pilyugin Center – Production Association Korpus,
Russia, Saratov.
lidkalihman@yandex.ru

V.A. Turkin

Branch of Academician Pilyugin Center – Production Association Korpus,
Russia, Saratov, tyrkinba@mail.ru

A.A. Akmaev

Branch of Academician Pilyugin Center – Production Association Korpus,
Russia, Saratov
aleks.akmaev@yandex.ru

V.V. Skorobogatov,

Branch of Academician Pilyugin Center – Production Association Korpus,
Russia, Saratov
vvskorobogatov@yandex.ru

Abstract — **The report considers the principles of development of specialized software for inspection of inertial measurement units with nonorthogonally oriented measuring axes when applying the complex parameter technique at the manufacturer enterprise and during the incoming inspection at consumer enterprise prior to the installation of units into spacecraft control systems. The effectiveness of the proposed methodology and software is confirmed by many years of Soyuz and Progress spacecraft flight testing.**

Key words — methodology, complex parameter, angular rate sensors IMU, accelerometers IMU, gyroscope, accelerometer, flight tests, software.

I. Introduction

Inertial measurement units (IMUs) for strapdown inertial navigation systems (SINSs) containing angular rate sensors (ARSs) and accelerometers (sensors of apparent acceleration) as inertial sensing elements (ISEs), the sensitivity axes (SAs) of which are oriented nonorthogonally, i.e. located at certain angles to the axes of aircraft reference system (RS), have become widespread in control systems (CSs) of rocket and space technology, which was discussed in fundamental scientific papers on the subject [3-10, 12].

Specialists at Production Association Korpus have proposed a methodology for inspecting IMUs using a complex parameter technique, which was reported at International Conferences on Integrated Navigation Systems and was studied in the papers [13-16], where it was also noted that the proposed inspection technique has been successfully applied since 2002 to the present day and was tested on more than 86 Soyuz and Progress spacecraft. At the same time, in the same works it is noted that it is impossible to apply the proposed methodology without appropriate specialized software [1, 2, 11, 15, 16]. Since 2002, however, the software has been modified or completely reworked multiple times due to the advancement of computer technology and testing equipment. As noted in [13–16], the methodology can be applied not only for units of ARSs and accelerometers used in Soyuz and Progress SINSs, but also in any other control systems, where the application of IMUs with nonorthogonal orientation of measuring axes is necessary. And if the production of ISEs for Soyuz and Progress is naturally small-scale, then in case of application of IMUs with nonorthogonal orientation of sensing axes in other systems, where large-scale production is required, a steep increase in the number of workplaces in production divisions at manufacturing enterprises is expected to maintain quality inspection with a sufficiently large number of produced devices.

The report discusses the formation of new approaches in the development of software for inspection of IMUs with nonorthogonal orientation of measuring axes using the Client-Server architecture with the use of microservices.

II. Research

Prior to the introduction of microservices, specialized software was primarily based on a monolithic architecture where each application was a separate entity. This approach was valid until applications became too complex. To change a small portion of code in a monolithic system, the entire system had to be rebuilt, tested, and a new version of the application deployed.

With the advent of microservices, software is broken down into small elements that can be developed and deployed independently. Microservice architecture is an approach to building software as a set of decentralized services. These services are related only loosely, independently deployed and developed, and are easily maintained.

The structural diagram of the software client part is shown in Fig. 1.

The Client is a baseline specialized software package (compatibility tools for a specific testing equipment kit, means of visualization of inspections and means of communication with the Server). The Server represents a set of services responsible for receiving requests and sending responses; it also contains a storage for mathematical logic built on the microservice principle (a single calculated parameter is

associated with a single microservice), as well as a storage for the data of conducted tests. The structural diagram of the software server part is shown in Fig. 2.

Fig.1 Structural diagram of the software client part.

Fig. 2. Structural diagram of the software server part.

The major advantage of using such an architecture (Fig. 2) is the ease of maintenance and independence of individual software packages. The number of clients is limited only by network bandwidth and server hardware capacity. Clients can be distinctively different: automated workstation clients, a control and monitoring client (Fig. 3), a client that requests statistics data (Fig. 4) on all devices of the same type (e.g. yearly report), etc.

Fig. 3. "Control client and server" structural diagram.

The control and monitoring client is designed for server administration (single or several), namely to perform commands: stop, start, restart, maintenance shut down, gather statistics on the operation of each applications server and on the number of connected clients.

As seen from the diagram, the control client accesses the mathematical logic. This allows for "live" addition and modification of mathematics, i.e. without interrupting the applications server operation.

Sometimes there are situations when it is necessary to collect statistics on a particular calculated parameter for all products for a certain period of time. A separate client is needed for such purposes, the scheme of interaction with the server of which is shown in Fig. 4.

Fig. 4. "Statistics collection client and server" structural diagram.

As seen from the above, the number of solved client tasks and their nature can vary widely. Algorithmic language, in which the client is realized, can be chosen freely, based on the requirements to the architecture of the workstation (operating system, bit capacity, etc.), the most important thing is that the formation of requests and decoding of responses must use a predetermined byte-protocol.

It is reasonable to use "packet" data transmission for encoding and decoding of transmitted requests and responses. The scheme of the transmitted data packet is shown in Fig. 5.

Fig. 5. Transmitted data packet scheme.

The packet start and end marker are used to synchronize reception, in other words, the receiving side, upon reading the start marker, initiates data reception until the packet end marker is encountered.

Application of such an architecture in production will provide: **scalability** (easy commissioning of new workstations), **flexibility** (ability to replace or modify testing equipment, change testing sequence order and introduce new inspection procedures for the manufactured devices in the shortest possible time), **control efficiency** (introduction of microservices ensures the ease of software maintenance, cross-platform capabilities, and protection against possible data loss).

III. CONCLUSION

The proposed approach is especially effective for the inspection of IMUs for SINSs with nonorthogonal orientation of measuring axes, considering the complexity of ground testing algorithms, but, in principle, this approach can be applied in combination with other types of gyroscopic devices, inspection of which is also required within the framework of ground tests. In the future, further development plans for the proposed approach include application of artificial intelligence algorithms.

REFERENCES

[1] Akmaev, A.A., Turkin. V.A., Program for Inspecting Parameters of Inertial Measurement Unit of Accelerometers, certificate of state registration of computer program RU no. 2018663845, 2018.

[2] Akmaev, A.A., Turkin. V.A., Program for Inspecting Parameters of Inertial Measurement Unit of Angular Rate Sensors, certificate of state registration of computer program RU no. 2017662856, 2017.

[3] Branets, V.N., Shmyglevskii, I.P., Vvedenie v teoriyu besplatformennykh inertsial'nykh navigatsionnykh sistem (Introduction to the Theory of Strapdown Inetrial Navigation Systems), Moscow, Nauka, 1992.

[4] Branets, V.N., Sevast'yanov, N.N., Fedulov, R.V., Lektsii po teorii sistem orientatsii, upravleniya dvizheniem i navigatsii (Lectures on the Theory of Orientation, Motion Control and Navigation Systems), Tomsk, Tomsk University, 2013.

[5] Vodicheva, L.V., Bel'skiy, L.N. et al., Inertial Measurement Units of Advanced Rocket and Space Technology Instruments: Ensuring Fault Tolerance, Vestnik Samarskogo universiteta: Aerokosmicheskaya tekhnika, tekhnologii i mashinostroenie, 2018, vol. 17, no. 1, pp. 28-44.

[6] Vodicheva, L.V., Improvement of Reliability and Accuracy of Strapdown Inertial Measurement Units at Redundant Number of Measurements, Giroskopiya i navigatsiya, 1997, no. 1, pp. 55 – 67.

[7] Epifanov A.D. Izbytochnye sistemy upravleniya letatelnymi apparatami (Redundant Aircraft Control Systems), Moscow, Mashinostroenie, 1978.

[8] Izmailov, E.A, Modern Trends in the Development of Technologies of Inertial Sensing Elements and Aircraft Systems, Trudy FGUP NPTsAP, 2010, no. 1, pp. 27–35.

[9] Kalikhman, D.M., Nakhov, S.F. et al., Opyt proektirovaniya i izgotovleniya blokov izmeritelei lineinogo uskoreniya na kvartsevykh mayatnikovykh akselerometrakh s analogovoi i tsifrovoi sistemami upravleniya (Experience in Design and Manufacture of Units of Quartz Pendulum Accelerometers with Analog and Digital Control Systems), Saratov, Yuri Gagarin Saratov State Technical University, 2021.

[10] Kalikhman, D. M., Samitov, R.M. et al., Results of Flight Tests of the Six-Axis Inertial Measurement Unit Based on a Quartz Pendulum Accelerometer with Digital Feedback as Part of the Soyuz MS-14 Spacecraft, 27th St. Petersburg International Conference on Integrated Navigation Systems, CSRI Elektropribor, 2020, pp. 184-187.

[11] Kalikhman, D.M., Turkin, V.A. et al., Principles of Development of Digital Noise Resistant Controllers of Angular Rate and Apparent Acceleration Measurement Channels in Modern SINS and Testing Software, 20th St. Petersburg International Conference on Integrated Navigation Systems, CSRI Elektropribor, 2013, pp. 285–291.

[12] Negri, C., Lignon, C. et al., A New Generation of IRS with Innovative Architecture Based on HRG for Satellite Launch Vehicles, Gyroscopy and Navigation, 2016, vol. 7, no. 3, pp. 223–230.

[13] Turkin, V.A., Results of Development of the Methodology for Inspecting Inertial Measurement Units with Nonorthogonal Orientation of Measuring Axes, Izvestiya Tul'skogo gosudarstvennogo universiteta: Tekhnicheskie nauki, Tula, TulGU, 2023, no. 2, pp. 257 -264.

[14] Turkin, V.A., Application of Mathematical Modeling in the Development of Design Methodology for Inertial Measurement Units with Nonorthogonal Orientation of Measuring Axes, Izvestiya Tul'skogo gosudarstvennogo universiteta: Tekhnicheskie nauki, Tula, TulGU, 2023, no. 4, pp. 326-334.

[15] Kalikhman, D.M., Turkin, V.A. Techniques for Inspecting IMUs with Nonorthogonal Orientation of Their Sensitive Axes. Gyroscopy Navig. 14, 305–319 (2023). https://doi.org/10.1134/S2075108724700032.

[16] Turkin, V.A., Kalikhman, D.M. et al., Techniques for Inspecting IMUs with Nonorthogonal Orientation of Their Sensitive Axes, 30th St. Petersburg International Conference on Integrated Navigation Systems, CSRI Elektropribor, 2023, pp. 244-248.

Investigation of the Effect of Structural Nonrigidity of a Precision Test Bench with Inertial Sensing Elements for Testing Angular Rate Sensors on its Accuracy Characteristics

D. M. Kalikhman,
Yuri Gagarin State Technical University of Saratov,
Russia, Saratov;
Branch of Academician Pilyugin Center – Production
Association Korpus,
Russia, Saratov.
lidkalihman@yandex.ru

N. V. Tarakanov,
Yuri Gagarin State Technical University of Saratov,
Russia, Saratov;
Branch of Academician Pilyugin Center – Production
Association Korpus,
Russia, Saratov.
nikolaytarakanov87@gmail.com

E. A. Deputatova,
Yuri Gagarin State Technical University of Saratov,
Russia, Saratov;
Branch of Academician Pilyugin Center – Production
Association Korpus,
Russia, Saratov.
deputatova@bk.ru

S. V. Pchelintseva
Russian State Agrarian University - Moscow Timiryazev
Agricultural Academy,
Russia, Moscow.
pchelintseva@inbox.ru

Abstract — **The paper investigates the effect of structural nonrigidity of overall design and of the mounting of tested gyroscopic devices on the platform of a precision rotary test bench with inertial sensing elements on its accuracy characteristics. The results of mathematical modeling of the bench operation under disturbance effects with the introduction of a shock absorber pad for mounting platform are given.**

Key words — precision rotary test bench, structural nonrigidity, error, accuracy characteristics, shock absorber node.

I. INTRODUCTION

Technologies of strapdown inertial navigation systems (SINSs) are becoming widely used in all areas of navigation development: aviation, space, marine systems, and pedestrian navigation [1, 2]. In their turn, inertial sensing elements (ISEs) – primary SINS sensors – are also advancing, and angular rate sensors based on new physical principles are becoming increasingly widespread as well, while electromechanical angular rate sensors are found in SINS of various purposes [3–7]. In any case, precision test benches are needed to accurately test for the scale factor error in angular rate sensors (ARSs). The authors discussed the principles of designing precision test benches with inertial sensing elements at international conferences and in a number of publications [8–14].

A design concept for precision test benches with inertial sensing elements for inspection of gyroscopic devices was proposed in [12, 13]. It was shown that the concept can be used to generate a state primary standard for setting and storing angular rate values, as well as to create hierarchically subordinate standards. Angular rate sensors and linear accelerometers of different accuracy certifications and principles of operation can be used as inertial sensing elements, along with high-precision optical angular encoders, and auxiliary sensors, i.e. Hall effect sensors. Precision test benches can feature both analog and digital control systems for both primary and ISE feedback loops [11, 15]. However, the more precise a bench is, the more accurate are its ISEs, and the influence of a variety of errors inevitably found in a bench design affecting its accuracy characteristics rises significantly. The reason for that is all ISEs (as part of a bench design) are included in the primary feedback system of the bench automatic control system, and since the elements are highly sensitive, they will measure any distorted signals that appear due to structural errors. Distorted signals will be sent to the feedback system of the bench, and consequently, affect its accuracy characteristics in a negative way. Such structural errors include nonrigidity of attachment of the main axis of platform rotation to the bench body and nonrigidity of mounting of tested devices on bench platform. Additionally, when setting harmonic oscillations of the bench platform around the axis of rotation, vibration errors may occur, also affecting the evaluation of the accuracy characteristics of devices under testing [5–7, 16–21]. This paper investigates the effect of structural nonrigidity on the accuracy characteristics of a precision test bench for the inspection of gyroscopic devices.

II. Main elements in the bench structure affecting its accuracy characteristics

Figure 1 shows the functional and kinematic diagram of the investigated bench, where the following designations are introduced: ω_α – angular rate; TD – test device; PA – power amplifier; AE – angular encoder; FOG – fiber-optic gyroscope; LG – laser gyroscope; LAS – linear acceleration sensors; τ – tangential acceleration; c – centripetal acceleration; n – number of LASs; DCS – digital control system; DCU – data conversion unit.

Fig. 1. Functional and kinematic diagram of the precision rotary test bench

The considered bench contains a platform for mounting of tested devices. Its motor control system is digital and built on the difference principle using various inertial sensing elements. An optical angle encoder is used as the angle sensor of the bench. The output signals of all bench sensors are transmitted to the control and measurement systems by a non-contact method of information transfer [13].

Figure 2 shows the main elements in the bench structure affecting its accuracy characteristics, namely:

I – inertial sensing elements (ISEs) – angular rate sensors (ARSs) and linear acceleration sensors (LASs);

II – external shaft of rotation on supports, which is rigidly connected with the platform;

III – tested device (TD) – an angular rate sensor (ARS) of any type and operating principle;

IV – mounting system for fixing a TD on the platform.

Two rotorless gyroscopes – laser and fiber-optic with the best characteristics in terms of stability of the scale factor and zero signal [2], respectively, as well as precision sensors of tangential and centripetal acceleration, i.e. quartz pendulum accelerometers [22–24], are taken as inertial sensing elements in consideration of exclusion of their mutual influence.

To determine the instability of the bench angular rate of rotation, let us write down the probability density functions of the corresponding errors of indicated sensors, in particular, for ΔU_{fog}, ΔU_{lg} – for the instability of the zero signals of FOG, LG and for $M\varphi$ – for the change in the pulling force acting around the sensitivity axis of the quartz plate of the accelerometer, leading to a temporal change of its zero signal. For FOG and LG as ISEs with zero signal instability values of 0.005 °/$h = 2{,}4\cdot10^{-8}$ rad/s and 0.01 °/$h = 5\cdot10^{-8}$ rad/s, respectively, the root mean square deviation (RMSD) will be $1.2\cdot10^{-8}$ rad/s and $2.5\cdot10^{-8}$ rad/s, then the distribution laws will be as follows

$$\phi(\Delta U_{fog}) = \frac{1}{1,2\cdot10^{-8}\sqrt{2\pi}}\cdot e^{-\frac{(\Delta U fog)^2}{2\cdot(1,2\cdot10^{-8})^2}}, \qquad (1)$$

$$\phi(\Delta U_{lg}) = \frac{1}{2,5\cdot10^{-8}\sqrt{2\pi}}\cdot e^{-\frac{(\Delta U lg)^2}{2\cdot(2,5\cdot10^{-8})^2}}. \qquad (2)$$

Fig. 2. Main elements in the bench structure affecting its accuracy characteristics

The mathematical model of the error M_φ can be represented by a random function with zero mathematical expectation and the RMSD corresponding to the random drift component in startup, which for the accelerometer in consideration [24] is $0.5\cdot10^{-6}$ g, i.e. the RMSD is $7.5\cdot10^{-7}$ g, then the distribution law will be as follows

$$\phi(M_\varphi) = \frac{1}{7,5\cdot10^{-7}\sqrt{2\pi}}\cdot e^{-\frac{(M_\varphi)^2}{2\cdot(7,5\cdot10^{-7})^2}}. \qquad (3)$$

Let us investigate the effect of the tested device and the mounting system. If rotorless gyroscopes are tested, then vibration isolation between the platform with the device to be tested and the bench ISEs is not required. If a tested device features a gyromotor that excites vibrations that may be transmitted to the bench sensing elements, causing undesirable decrease in their accuracy, then a shock absorber node must be used.

A shock absorber node containing an elastic shock pad is introduced into the platform structure (Fig. 3) in accordance with

the vibration isolation rules [17] in order to effectively isolate the object from the vibration source.

Fig. 3. Structural diagram of the shock absorber node of the platform

The shock absorber must have intrinsic frequency that is 3 to 5 times lower than the lower limit of the excitable frequency spectrum. Working gyromotors of the TD and ISEs excite frequency spectra, the main (carrier) frequency of which is equal to the circular rotation frequency of the corresponding gyromotor, while all other frequencies of the spectrum are multiples of the carrier frequency. Selection of the shock absorber rigidity is based on the effective shock absorption condition [17]:

$$\bar{C}_{12} \le n_{min}^2 \cdot M_\Sigma / 1600,$$ (4)

where n_{min} is the smallest value of the TD and ISE gyromotors rotor rates of rotation (in *rpm*); M_Σ is the total mass of the elements loaded on the platform, such as the TD, platform cap, and shock absorber node.

If we consider a float angular rate sensor with a random component of zero signal in startup of 0.06 °/*h* [8] as the tested device, the probability density function of this error will adhere to the law:

$$\phi(M_\beta) = \frac{1}{5,5 \cdot 10^{-6} \sqrt{2\pi}} \cdot e^{-\frac{(M_\beta - 1,1 \cdot 10^{-6})^2}{2 \cdot (5,5 \cdot 10^{-6})^2}},$$ (5)

where M_β is the instability of the pulling force along the precession axis of the ARS gyro node, which determines the random drift component in startup, independent on overload.

Let us consider the rigidity of the external shaft of rotation for the cases of application of ball bearing and aerostatic supports. This factor is included in the expression for the instability of the moment of resistance along the bench axis of rotation in the form of the moment M_{frict}, determined by the Coulomb friction model:

$$M_\alpha = M_{frict} + \Delta M_{motor} - H \cdot \omega_{Earth.} \cdot \cos\varphi \cdot \sin\omega_0 t,$$ (6)

where H is the kinetic moment of the TD; ω_0 is the actual angular rate of bench rotation.

The same expression accommodates the component associated with the error of angular rate reproduction by the motor ΔM_{motor}. Additionally, external conditions may affect the

accuracy characteristics of the bench, but since the considered bench is ground-based equipment installed in a specialized facility on a "decoupled" (vibrationally isolated) platform, the greatest influence may be caused by the horizontal component of the Earth rotation rate $\omega_{Earth.} \cdot \cos\varphi$, which can be compensated by the bench digital control system.

The probability density functions in this case will take the form:

- if ball bearing supports are used:

$$\phi(M_\alpha) = \frac{1}{13,8\sqrt{2\pi}} \cdot e^{-\frac{(M_\alpha - 10,35)^2}{2 \cdot 13,8^2}};$$ (7)

- if aerostatic supports are used:

$$\phi(M_\alpha) = \frac{1}{0,01\sqrt{2\pi}} \cdot e^{-\frac{(M_\alpha - 0,3)^2}{2 \cdot 0,01^2}}.$$ (8)

The accuracy characteristics of the bench will also be affected by the instability of the input voltage source ΔU_{in}. This error is a step function with an amplitude of 15.5 μV and occurs in the amount of 2 to 3 spikes per 10 minutes [8]. The distribution law of the error ΔU_{in} will take the form:

$$\phi(\Delta U_{in}) = \frac{1}{3,5 \cdot 10^{-6}\sqrt{2\pi}} \cdot e^{-\frac{(\Delta Uin - 5 \cdot 10^{-6})^2}{2 \cdot (3,5 \cdot 10^{-6})^2}}.$$ (9)

III. RESULTS

Figs. 4 and 5 show the results of the investigation of structural nonrigidity of a precision test bench under simultaneous action of all the considered errors with FOG as ISE (identical results with LG as ISE) in the case of application of ball bearing supports and in the case of application of aerostatic supports and a shock absorber node for TD mounting, respectively.

Therefore, the instability of the bench rotation rate in the first case (Fig. 4) is $5 \cdot 10^{-4}$ °/s, and in the second case (Fig. 5) is $2.5 \cdot 10^{-4}$ °/s.

Fig. 4. Instability of the bench angular rate under simultaneous action of errors M_α, ΔU_{in}, M_β, ΔU_{fog} and M_φ (in the case of ball bearing supports application).

Fig. 5. Instability of the bench angular rate under simultaneous action of errors M_α, ΔU_{in}, M_β, ΔU_{fog} and M_φ (in the case of aerostatic supports and shock absorber node application).

REFERENCES

[1] Izmailov, E.A., Modern Trends in the Development of Technologies of Inertial Sensing Elements and Aircraft Systems, Trudy FGUP NPTsAP. Sistemy i pribory upravleniya, 2010, no. 3, pp. 27–35.

[2] Rivkin, B.S., Analytical Review of the State of Research and Development in the Field of Navigation Abroad, St. Petersburg, CSRI Elektropribor, 2017–2023, nos. 1–8.

[3] Reshetnikov, V.I., Doronin, V.P. et al., Gyroscopic Sensing Elements for Orbital Spacecraft Orientation and Stabilization Control Systems, 8th St. Petersburg International Conference on Integrated Navigation Systems, St. Petersburg, CSRI Elektropribor, 2001, pp. 17–30.

[4] Reshetnikov, V.I., Volyntsev, A.A. et al., Experience in Creating High-Precision Float Gyro Devices Used in Angular Orientation and Stabilization Systems of Spacecraft and Stations, 10th St. Petersburg International Conference on Integrated Navigation Systems, St. Petersburg, CSRI Elektropribor, 2003, pp. 226–234.

[5] Andreichenko, K.P., Dinamika poplavkovykh giroskopov i akselerometrov, Moscow, Mashinostroenie, 1987.

[6] Pel'por, D.S., Giroskopicheskie sistemy: v trekh tomakh (Gyroscopic Systems: In Three Volumes), Moscow, Vyshaya shkola, 1986.

[7] Pel'por, D.S., Mikhalev, I.A., Bauman, V.A. et al., Giroskopichekie pribory i sistemy (Gyroscopic Instrumentation and Systems), Moscow, Vyshaya shkola, 1988.

[8] Kalikhman, D.M., Pretsizionnye upravlyamye stendy dlya dinamicheskikh ispytanii giroskopicheskikh priborov (Precision Test Benches for Dynamic Testing of Gyroscopic Devices), Peshekhonov, V.G., Ed., St. Petersburg, CSRI Elektropribor, 2008.

[9] Deputatova, E.A., Kalikhman, D.M. et al., Digital Motion Stabilization in Precision Controlled Platforms with Inertial Sensing Elements. Part 1: Application of a Float Angular Rate Sensor, Izvestiya RAN: Teoriya i sistemy upravleniya, 2011, no. 1, pp. 120–132.

[10] Deputatova, E.A., Kalikhman, D.M. et al., Digital Motion Stabilization in Precision Controlled Platforms with Inertial Sensing Elements. Part 2: Application of Float Angular Rate Sensors and Pendulum Accelerometers, Izvestiya RAN: Teoriya i sistemy upravleniya, 2011, no. 2, pp. 131–146.

[11] Deputatova, E.A., Kalikhman, D.M. et al., New Generation of Precision Rotary Test Benches with Inertial Sensing Elements and Digital Control, Izvestiya RAN: Teoriya i sistemy upravleniya, 2014, no. 2, pp. 130–146.

[12] Kalikhman, D.M., Deputatova, E.A. et al., Development of a Design Concept for a Class of Precision Mechatronic Test Benches Using Inertial Sensing Elements Combined with High-Precision Angle Sensors, 29th St. Petersburg International Conference on Integrated Navigation Systems, St. Petersburg, CSRI Elektropribor, 2022, pp. 212–217.

[13] Kalikhman, D.M., Deputatova, E.A., Pchelintseva, S.V. et al. Design Philosophy for a Class Of Precision Rate Tables with Inertial Sensors in the Feedback Loop. Gyroscopy Navig. 13, 155–169 (2022). https://doi.org/10.1134/S2075108722030051.

[14] Kalikhman, D.M., Deputatova, E.A. et al., Development of Control Algorithms for a Precision Test Bench with Inertial Sensing Elements and Digital Control System and Analysis of Its Errors, Izvestiya Tul'skogo gosudarstvennogo universiteta: Tekhnicheskie nauki, Tula, TulGU, 2023, no. 9, pp. 66–80.

[15] Deputatova, E.A., Wide-Range Rotary Test Benches with Digital Control Systems and Navigation Instruments as Sensing Elements for Inspecting Sngular Rate Sensors, Cand. Sci. Dissertation, Yuri Gagarin State Technical University of Saratov, Saratov, 2012.

[16] Blekhman, I.I., Sinkhronizatsiya dinamicheskikh sistem (Synchronization of Dynamic Systems), Moscow, Nauka, 1971.

[17] Il'inskii, V.S., Zashchita apparatov ot dinamicheskikh vozdeistviy (Protection of Apparatus from Dynamic Effects), Moscow, Energiya, 1970.

[18] Wrigley, W., Hollister, W., Denhard, W., Teoriya, proektirovanie i ispytaniya giroskopov (Gyroscopic Theory, Design, and Instrumentation), Moscow, Mir, 1972.

[19] Ishlinksii, A.Yu., Mekhanika otnositel'nogo dvizheniya i sily inertsii (Mechanics of Relative Motion and Forces of Inertia), Moscow, Nauka, 1981.

[20] Kalikhman, D.M., Deputatova, E.A., Gorbachev , V.O., Investigation of the Effect of Structural Nonrigidity of a Precision Test Bench Design and of Tested Device Mounting on its Accuracy Characteristics, Izvestiya Tul'skogo gosudarstvennogo universiteta: Tekhnicheskie nauki, Tula, TulGU, 2023, no. 9, pp.507–519.

[21] Kalikhman, D.M., Deputatova, E.A., et al., All Purpose Precision Test Bench with Inertial Sensing Elements and Aerostatic Suspension for Angular Rate Sensors, 30th Anniversary St. Petersburg International Conference on Integrated Navigation Systems, St. Petersburg, CSRI Elektropribor, 2023, pp. 304–309.

[22] Kalikhman, D.M., Skorobogatov, V.V. et al., Quartz Pendulum Accelerometer with Digital Feedback as an Automatic Control System: Results of Synthesis of Controllers and Software. Prospects for Further Development, Navigatsiya i upravlenie letatelnymi apparatami, 2019, no. 3 (26), pp. 2–37.

[23] Grebennikov, V.I., Deputatova, E.A. et al., Pendulum Accelerometer with Digital Control and New Functional Capabilities, Izvestiya RAN. Teoriya i sistemy upravleniya, 2021, no. 2, pp.71–93.

[24] Nakhov, S.F., Kalikhman, D.M. et al., Experience in Design and Manufacture of Units of Quartz Pendulum Accelerometers with Analog and Digital Control Systems, Yuri Gagarin State Technical University of Saratov, Saratov, 2021.

Improving Accuracy Characteristics and Functionality in a Rotary Table for Inspection of Navigation Devices and Their Element

A.V. Polushkin

Branch of Academician Pilyugin Center – Production Association Korpus,
Russia, Saratov.
korpus_otd231@npcap.ru

I.V. Slistin

Branch of Academician Pilyugin Center – Production Association Korpus,
Russia, Saratov.

N.A. Kaldymov

Branch of Academician Pilyugin Center – Production Association Korpus,
Russia, Saratov.

A.A. Ivanov

Branch of Academician Pilyugin Center – Production Association Korpus,
Russia, Saratov.

I.A. Nazarov

Branch of Academician Pilyugin Center – Production Association Korpus,
Russia, Saratov.

V.F. Vasil'ev

Branch of Academician Pilyugin Center – Production Association Korpus,
Russia, Saratov.

A.K. Gerte

Branch of Academician Pilyugin Center – Production Association Korpus,
Russia, Saratov.

A.V. Pugovkin

Branch of Academician Pilyugin Center – Production Association Korpus,
Russia, Saratov.

S.O. Gorelov

Branch of Academician Pilyugin Center – Production Association Korpus,
Russia, Saratov.

Abstract — The paper considers the experience in creating a precision rotary table, the ways to improve its accuracy characteristics and provide advanced functionality. An analytical model of plane angle setting and measurement is shown. The results of expansion of functionality of a rotary table used for inspection of navigation devices and their elements are presented: characteristics of setting of angular rates and measurements, introduction of time gating for synchronization with the inspected device and its test equipment, and implementation of measurement of time intervals for devices with phase coding of output data. Two samples of the new rotary table (labeled SPP-19.001) were tested at D.I. Mendeleev Institute for Metrology to ascertain the accuracy class. The tests have confirmed high accuracy and advanced functionality.

Key words—angle measurement, positioning, angle-setting rotary table, control algorithms, test equipment, navigation instruments, angular rate measurement, precision instrumentation, synchronization, time gating, time domain measurements, mathematical modeling, automatic settings, systematic errors, error model, algorithmic compensation.

I. INTRODUCTION

Since the 1980s, Academician Pilyugin Center has been engaged in the development and production of precision angle-setting rotary tables for inspection of in-house manufactured navigation devices. The control systems of launch vehicles and spacecraft, and consequently all the included instrumentation, are constantly advancing, thus the associated test and inspection equipment must be improved accordingly. For the rotary tables produced by Academician Pilyugin Center, the accuracy level of setting and measurement of angular positions of less than 0.5", which is the limit for working measuring instruments, has already been reached. Recently, high-precision instruments that require angle setting accuracy level of tenths of an arc second, as well as non-standard test methods, have emerged. Such devices include precision gyroscopes (drift of less than 0.1 °/h) and accelerometers (sensitivity threshold of 10^{-6} to $10^{-7}g$), tiltmeters, inclinometers, star coordinate determination units (star sensors), gyrocompasses, precision angle sensors (induction, optical, holographic), etc. [1, 2, 3]. Analysis of the National Register of Measuring Equipment of the Russian Federation shows that rotary tables of such a high accuracy class are not mass-produced. In 2014, two models of precision rotary tables developed at Academician Pilyugin Center were entered into the Register: the angle-setting complex "UK" with error of less than ±0.5" (register no. 58244-14 [4]) and the digital rotary table "SPC-383" with error of less than ±0.35" (register no. 56468-14 [5]) [6]. These rotary tables are successfully operated at the manufacturer's enterprise production sites and were supplied to the leading scientific centers of the Russian Federation. A decade of operation and maintenance experience [7] has revealed design flaws and the lack of certain functionality that could allow for a wider range of devices eligible for inspection. As a result, the developers came up with technical solutions that made it possible to create a more accurate and overall superior rotary table model, the SPP-19.001.

II. MODEL OF ROTARY TABLE ERRORS. WAYS TO IMPROVE ACCURACY

The precision rotary table SPP-19.001 uses a high-resolution (~0.01″) angle sensor (AS), the resolution of which is 5 times higher than that of the previous models. The error models of SPC and UK rotary tables are discussed in [8, 9]. The SPC-383 error model can be applied to the SPP-19.001 with some adjustments and different weight coefficients. Let us consider ways to reduce individual errors in the new rotary table.

For the SPP-19.001, the layout and kinematics of the table, the designs of the spindle and the drive have been significantly revised, the body rigidity has been increased, while friction and free play have been reduced. Novel hardware and software solutions were introduced to reduce random errors in plane angle measurement and to improve smoothness of spindle rotation. To eliminate systematic errors, calibration and algorithmic compensation were implemented.

Fig. 1 shows the graphical representation of the SPP-19.001 rotary table error model.

Fig. 1. Rotary table error model.

The analytical error model of the precision rotary table SPP-19.001 can be expressed as follows:

$$\Delta\alpha(\alpha,t) = \Delta\alpha_{AS}(\alpha) + \Delta\alpha_{I}(\alpha,t) + \Delta\alpha_{MI}(\alpha) + \Delta\alpha_{E}(\alpha,t);$$

$$\Delta\alpha_{I}(\alpha,t) = \Delta\alpha_{SI}(\alpha,t) + \Delta\alpha_{OT}(\alpha,t) + \Delta\alpha_{HV}(\alpha,t) + \\ + \Delta\alpha_{DC}(\alpha,t);$$

$$\Delta\alpha_{MI}(\alpha) = \Delta\alpha_{C}^{rand}(\alpha,t) + \Delta\alpha_{PP}^{rand}(\alpha,t) + \Delta\alpha_{AC}^{rand}(\alpha,t) + \\ + \Delta\alpha_{C}^{syst}(\alpha) + \Delta\alpha_{PP}^{syst}(\alpha) + \Delta\alpha_{AC}^{syst}(\alpha);$$

$$\Delta\alpha_{E}(\alpha,t) = \Delta\alpha_{T}(\alpha,t) + \Delta\alpha_{TI}(\alpha,t) + \Delta\alpha_{VI}(\alpha) + \Delta\alpha_{EM}(\alpha,t).$$

where:

$\Delta\alpha$ is the SPP-19.001 resulting error;

$\Delta\alpha_{AS}$ is the optical angle sensor (AS) error;

$\Delta\alpha_{I}$ is the angle sensor installation error;

$\Delta\alpha_{SI}$ is the error caused by spatial instability of the spindle;

$\Delta\alpha_{OT}$ is the error caused by the increase of operating temperature due to the heat emitted by units and assemblies of the rotary table;

$\Delta\alpha_{HV}$ is the error dependent on the position of the spindle (horizontal or vertical) [10]

$\Delta\alpha_{DC}$ is the error caused by linear and angular dimensions changes of parts;

$\Delta\alpha_{MI}$ is the instrumental error of measuring instruments;

$\Delta\alpha_{C}^{rand}$ is the random component of calibration error;

$\Delta\alpha_{PP}^{rand}$ is the random component of polygonal prism error;

$\Delta\alpha_{AC}^{rand}$ is the random component of autocollimator error;

$\Delta\alpha_{C}^{syst}$ is the systematic component of calibration error;

$\Delta\alpha_{PP}^{syst}$ is the systematic component of polygonal prism error;

$\Delta\alpha_{AC}^{syst}$ is the systematic component of autocollimator error;

$\Delta\alpha_{E}$ is the error caused by the influence of external factors;

$\Delta\alpha_{T}$ is the thermal error;

$\Delta\alpha_{TI}$ is the temporal instability;

$\Delta\alpha_{VI}$ is the error caused by oscillations, vibration or impact;

$\Delta\alpha_{EM}$ is the error caused by the influence of electromagnetic fields.

The composition of errors of the optical angle sensor is similar to the composition of errors of the AS, as discussed in detail in [11]. The errors that depend on the rotary table design and operating conditions can be controlled and minimized. The errors of measuring instruments used to test the accuracy of angle measurement are of particular importance.

The level of errors characteristic of the SPP-19.001 for plane angle setting and measurement is ±0.25″, which is close to the level of state reference standards. When calibrating and inspecting rotary table parameters during the manufacture process, metrological equipment with the total value of instrumental measurement error of less than 0.08″ (the upper limit for a polygonal prism, which corresponds to the plane angle secondary reference standard) is required. Such highly precise instruments are not readily available at regular instrument-making enterprises. Therefore, in the course of research and development, an original method was found, and a device was developed that, using the high stability of the polygonal quartz prism angles, provided increased reliability of input corrections of the SPP-19.001 calibration curve and allowed for automating the process of verifying the rotary table angle measurement accuracy (the method is patented [12]), while a patent is pending for the associated device. In 2022-2023, this method and device were used to inspect and research the Automated Angle Measuring Complex (AAMC) with a 24-faced quartz prism [13]. The results of comparison of experimental data with the results of rotary table inspection at D.I. Mendeleev Institute for Metrology with the primary reference standard for plane angle measurement in 2014 are shown in Table 1.

For the rotary table design, the rigidity of the body was strengthened and an original arrangement of power units was introduced, which reduced the angular deformations of the body and other errors caused by elastic pliability. Additionally, an original method for controlling the body deformations described in the patent [14] was applied.

Table 1.

Corrections of adjacent angles of a 24-faced prism found at the AAMC workstation, ...″						Certificate of verification no. №2511-11/338, ...″
Face nos.	Meas. 1 06.06.22	Meas. 2 13.06.22	Meas. 3 15.02.23	Meas. 4 20.02.23	Avg.	
1-2	-1.20	-1.20	-1.18	-1.19	-1.19	-1.18
2-3	3.15	3.13	3.08	3.08	3.11	3.15
3-4	1.48	1.48	1.50	1.52	1.5	1.52
4-5	-2.28	-2.26	-2.31	-2.33	-2.3	-2.29
5-6	0.34	0.32	0.35	0.39	0.35	0.36
6-7	1.93	1.92	1.96	1.97	1.95	1.89
7-8	-5.18	-5.19	-5.18	-5.2	-5.19	-5.16
8-9	1.86	1.87	1.86	1.86	1.86	1.85
9-10	-1.41	-1.42	-1.38	-1.39	-1.4	-1.44
10-11	1.30	1.30	1.24	1.27	1.28	1.35
11-12	2.33	2.34	2.34	2.37	2.35	2.33
12-13	0.00	-0.02	0.03	0.01	0.01	0.01
13-14	-0.52	-0.51	-0.46	-0.49	-0.5	-0.51
14-15	-1.18	-1.18	-1.22	-1.23	-1.2	-1.15
15-16	0.75	0.75	0.71	0.71	0.73	0.71
16-17	-0.36	-0.36	-0.35	-0.35	-0.36	-0.45
17-18	-2.68	-2.69	-2.71	-2.7	-2.7	-2.70
18-19	3.72	3.73	3.63	3.63	3.68	3.79
19-20	-2.53	-2.53	-2.48	-2.48	-2.51	-2.60
20-21	3.80	3.79	3.84	3.8	3.81	3.84
21-22	-1.94	-1.93	-1.90	-1.92	-1.92	-1.99
22-23	0.76	0.78	0.72	0.72	0.75	0.83
23-24	0.20	0.21	0.23	0.23	0.22	0.18
24-1	-2.31	-2.31	-2.30	-2.28	-2.3	-2.33

Systematic errors of the rotary table are periodic and can be represented as a harmonic series. Fourier analysis of the inspection results using a 24-faced reference prism and an autocollimator was applied to create the compensation curve. The result of Fourier series expansion of the obtained curves for vertical and horizontal positions of the rotary table axis are shown in Table 2.

Table 2.

Vertical axis			Horizontal axis		
Harmo-nic no.	Ampli-tude, ″ ...	Phase, ...°	Harmo-nic no.	Ampli-tude, ″ ...	Phase, ...°
0	-0.413	-	0	-0.557	-
1	0.011	18.605	1	0.099	46.871
2	0.386	16.750	2	0.468	-4.740
3	0.015	-9.340	3	0.053	5.752
4	0.131	112.416	4	0.050	-142.344
5	0.025	-149.992	5	0.030	-150.395
6	0.155	46.145	6	0.038	144.045
7	0.024	49.006	7	0.039	88.715
8	0.068	51.779	8	0.040	-76.333
9	0.017	-67.423	9	0.026	-46.098
10	0.075	176.570	10	0.039	-47.544
11	0.024	78.488	11	0.032	51.654
12	0.008	-180.000	12	0.008	-180.000

The algorithmic compensation curves for the corresponding position of the rotary table axis are determined by means of substituting the coefficients given in Table 2 into the formula:

$$s(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (A_k \cos(k\omega_1 t + \varphi_k))$$

Long-term observations during the production of precision rotary tables such as SPC-383 and SPP-19.001 have shown that the amplitudes of higher harmonics contribute less to the resulting error curve. Therefore, in order to reduce labor intensity, it is optimal and sufficient to perform tests with a 24-faced prism and to use only the first 12 harmonics in the analysis.

The graph of the angle measurement error of the SPP-19.001 rotary table following the introduction of algorithmic compensation (plotted over 10 reversed rotations) is shown in Fig. 2.

Fig.2 Graphs of rotary table errors after compensation.

The first two samples of rotary tables SPP-19.001 (work nos. H201001, H201002) were sent to D.I. Mendeleev Institute for Metrology to ascertain the type of measuring instrument. The results of testing at the leading metrological institute of the Russian Federation conducted in accordance with the metrologically approved methodology [15] confirmed the claimed accuracy class. Graphs of the plane angle measurement error test results are shown in Fig. 3.

Fig.3 Graphs of errors of two rotary tables obtained at D.I.Mendeleev Institute for Metrology.

Graphs in Fig. 4 show observations of the temporal instability of the angle measurement error of two sample SPP-19.001 rotary tables.

Fig. 4 Graphs of errors of two rotary tables obtained for one year of observation.

A series of tests for two SPP-19.001 rotary tables, conducted over a period of 11 months, confirmed a sufficiently high convergence of results. It should be taken into account that the testing at the manufacturer's enterprise and at D.I. Mendeleev Institute for Metrology was carried out using different measuring instruments, before and after transportation tests (Saratov – St. Petersburg, St. Petersburg – Saratov, 1580 km each way and back). A verification interval of 2 years was assigned for the SPP-19.001, according to the results of documentation analysis and actual tests conducted.

III. EXPANSION OF FUNCTIONAL CAPABILITIES

SPP-19.001 was created for inspecting a wide range of devices produced by Academician Pilyugin Center, as well as in consideration of the possibility of supplying the rotary table to other enterprises of rocket, space, aviation and other industries. The manufacturer of the SPP-19.001 produces more than 10 varieties of high-precision induction angle sensors and sine-cosine rotary transformers, which are important elements in gyrostabilized platforms, gyrocompasses and other rocket and space equipment. In these types of angle sensors, the angle of rotation of the rotor relative to the stator is proportional to the time interval between the Start and Stop pulses generated by service electronics. Considering this, it was important to supplement the rotary table with a metrologically validated time interval measurement function. The results of rotary table certification by this parameter are summarized in Table 3.

Therefore, the SPP-19.001 rotary table can be implemented in the composition of an automated workplace for verification of parameters of induction angle sensors, making it possible to exclude expensive standard time interval meters, and to significantly reduce the inspection time. The

SPP-19.001 can also be used to inspect other devices with time proportional output signal.

Table 3.

Frequency, Hz	Frequency meter		BPU-17.001 (channel 1)		BPU-17.001 (channel 2)	
	Period, μs	Interval, μs	Period 1, μs	Interval 1, μs	Period 2, μs	Interval 2, μs
0.1	9999999.938	4999999.9717	9999999.94	4999999.97	9999999.94	4999999.98
1	999999.993	499999.9942	999999.99	499999.98	999999.99	499999.99
10	99999.9994	49999.9967	100000.00	49999.99	99999.99	50000.00
100	9999.99996	4999.9972	10000.00	5000.00	10000.00	5000.00
1000	999.99996	499.9972	1000.00	500.00	999.99	500.00
5000	199.9999992	99.9972	200.00	100.00	199.99	100.00
10000	99.9999996	49.9972	100.00	50.00	100.00	50.00

During the manufacturing process, when inspecting navigation devices and their elements, not only static tests on precise angular positioning, but also dynamic tests are conducted, which imply setting of angular rates and synchronization of data from the test equipment and from the device under inspection. This feature is fully realized in the SPP-19.001. The range of reproducible angular rates from 0.005 to 10°/s was selected in consideration of the parameters of instruments used in spacecraft and upper (booster) rocket stages, i.e. precision angular rate sensors, integrating gyroscopes, etc. The claimed high accuracy of angular rate setting is confirmed by verification conducted at D.I. Mendeleev Institute for Metrology (see Table 4).

Table 4.

ω_{set}, °/c	δ_{change}, %		Δ_{dev}, %		δ_{instab}, %	
	+	-	+	-	+	-
10.0	0.00073	0.00090	0.00158	0.00093	0.00276	0.00135
5.0	0.00090	0.00090	0.00172	0.00358	0.00017	0.00275
1.0	0.00090	0.00089	0.00610	0.00583	0.00307	0.00199
0.5	0.00088	0.00088	0.00990	0.00933	0.00048	0.00111
Tolerance	0.005		0.01		0.01	

ω_{set}, °/c	δ_{change}, %		Δ_{dev}, %		δ_{instab}, %	
	+	-	+	-	+	-
0.1	0.00091	0.00091	0.02072	0.01272	0.00904	0.02469
0.05	0.00091	0.00092	0.02198	0.00933	0.02735	0.03233
Tolerance	0.005		0.05		0.05	

ω_{set}, °/c	δ_{change}, %		Δ_{dev}, %		δ_{instab}, %	
	+	-	+	-	+	-
0.01	0.00092	0.00092	0.01314	0.01530	0.01687	0.02905
0.005	0.00090	0.00091	0.02137	0.03458	0.02505	0.01860
Tolerance	0.005		0.1		0.1	

The SPP-19.001 has another feature to simplify its integration with newly developed or already implemented test equipment sets – a synchronization and time gating function that allows the user to assign a "leader" and a "follower" during the inspection process. If SPP-19.001 is assigned as the leader, it will form time gating pulses tied to events (when a specified angle or a series of angular positions is reached, at regular time intervals, etc.), while the follower will record the values of the output signal from the device under inspection as guided by the received gate pulses. If a test equipment set of the inspected device is assigned as the leader, the rotary table receives gating pulses from it and records proper values of angle, time, etc. by the edge of the gate pulse. The results of testing of time gating when reaching a specified angular position, conducted at D.I. Mendeleev Institute for Metrology, are shown in Table 5.

The control algorithms and software of SPP-19.001 have been significantly redesigned to expand functionality. New drive control algorithms allow for softer and quieter rotation, acceleration to the required speed is faster, and more accurate positioning is ensured [16]. Clear division of functions between the two controller units and the single-board

computer is introduced. The controller unit for drive control and angle sensor data processing was specially developed at Production Association Korpus for this task, and it is installed in the mechanical part of the rotary table. Command reception and data exchange with the single-board computer built into the table control unit is performed via a high-speed serial interface with a unique exchange protocol. The second controller unit is located in the control unit, and is deigned to receive and process the Start and Stop pulses. Its data is also transmitted to the single-board computer via the unique exchange protocol interface.

Table 5.

Gate no.	Gate step, …°	Angular rate, °/s	Angular position	Time, s
1			000° 00′ 00.00″	0
2			359° 59′ 59.98″	36.0011
3	360	10	360° 00′ 00.12″	36.0013
4			359° 59′ 58.18″	36.0014
5			359° 59′ 58.92″	36.0013
6			359° 59′ 57.91″	36.0014
1			000° 00′ 00.00″	0
2			015° 00′ 00.30″	150.6656
3	15	0.1	015° 00′ 00.04″	150.6876
4			014° 59′ 59.90″	150.7113
5			015° 00′ 00.44″	150.7007
6			014° 59′ 59.97″	150.6483
1			000° 00′ 00.00″	0
2			005° 00′ 00.11″	916.7797
3	5	0.005	005° 00′ 00.06″	916.2312
4			005° 00′ 00.04″	915.9768
5			004° 59′ 59.93″	917.0102
6			004° 59′ 59.88″	916.4354

The SPP-19.001 uses built-in, original in-house developed software, designed to control the process of setting spindle rotation parameters and measurements. All software is certified with the "High" level of protection in accordance with metrological regulations and recommendations. A possibility of unauthorized influence on software and measurement data is effectively eliminated. The end user is also able to connect external devices via standard communication interfaces using dynamic libraries, as well as to remotely control table operation modes via standard Client and Server procedure, using a specially designed exchange protocol. The manufacturer provides full technical support when the SPP-19.001 is used as part of end user's custom test equipment set.

IV. CONCLUSIONS

As a result of the research and development initiative, a more advanced rotary table has been created, which can be used in the development of automated workplaces for inspection of parameters of navigation devices and their elements, as well as in other areas of science and technology, in which precise setting and measurement of plane angles and angular rates are necessary. Visual appearance of the precision rotary table SPP-19.001 is shown in Fig. 5, and technical characteristics are given in Table 6.

Table 6.

Characteristic	Value
Angular rate measurement range, °/s	0.005 up to 10
Tolerance limits of relative error of angular rate measurement for 360° angle (δ_{change}), %	±0.005
Angular rate setting range, °/s	0.005 up to 10
Relative deviation of angular rate from the set value (Δ_{dev}), %, not more than: - for 360° angular displacement - for 15° angular displacement - for 5° angular displacement	±0.01 ±0.05 ±0.1
Relative instability of angular rate (δ_{instab}), %, not more than: - for 360° angular displacement - for 15° angular displacement - for 5° angular displacement	±0.01 ±0.05 ±0.1
Measuring range of rotation angles	±360°
Tolerance limits of absolute error of rotation angle measurements	±0.25″
Maximum deviation from the set rotation angle - in automatic mode - in manual mode	±0.5″ ±0.1″
Measurement range of interval between Start and Stop signals and period of these signal sequencing, s	$1 \cdot 10^{-5}$ up to 10
Tolerance limits of absolute error of measurement of interval between Start and Stop signals and period of these signals sequencing, μs	±0.01
Measurement range of synchronization pulses sequencing period, s	$1 \cdot 10^{-5}$ up to 10
Weight of device under test, kg, not more than - with vertical position of the rotary axis - with horizontal position of the rotary axis	50 35
Dimensions, mm, not more than: - rotary table LxWxH - control unit LxWxH	320x330x320 320x210x325
Weight, kg, not more than: rotary table/control unit	85/10
Verification frequency, month	24
National Register of Measuring Equipment no.[17]	no. 91991-24

Fig. 5 The appearance of SPP-19.001.

Considering the experience of using the previously developed rotary tables UK and SPC-383, which have proven history of operation reliability, the improved model of the rotary table is expected to find wide application at industrial enterprises and the leading scientific centers of the Russian Federation.

REFERENCES

1. Polushkin, A.V., Kaldymov, N.A., Dmitriev, N.V., Nakhov, S.F., Bessonov, R.V., Brysin, N.N., Polyanskii, I.V., Methods and Tools for Evaluating the Technical Characteristics of Star Sensors in Ground Tests, 23rd St. Petersburg International Conference on Integrated Navigation Systems, St. Petersburg, CSRI Elektropribor, 2016, pp. 401-406.

2. Konovalov, S.F., Navigatsionnye akselerometry: uchebnoe posobie (Navigation Accelerometers: a Tutorial), Moscow, Bauman Moscow State Technical University, 2023.

3. Peshekhonov, V.G. The Outlook for Gyroscopy. Gyroscopy Navig. 11, 193–197 (2020). https://doi.org/10.1134/S2075108720030062.

4. Russian Federation National Register of Measuring Equipment no. 58244-14, Angle Setting Complexes "UK", https://grmetr.ru/gosreestr/371242-58244-14-uk

5. Russian Federation National Register of Measuring Equipment no. 56458-14, Digital Rotary Tables "SPC-383", https://grmetr.ru/gosreestr/369164-56458-14-spts-383

6. Polushkin, A.V., Kaldymov, N.A., Slistin, I.V., Borchaninov, D.G., Nakhov, S.F., Polushkin, A.A., Yankovskii, A.A., Selection of the Optimal Method of Algorithmic Compensation of Systematic Errors in the Precision Angular Positioning Rotary Table, Materialy VI Mezhdunarodnoi nauchno-tekhnicheskoi konferentsii "Izmereniya i ispytaniya v sudostroenii i smezhnykh otraslyakh - Sudometrika-2016" (Proceedings of the 6th International Scientific and Technical Conference "Measurement and Testing in Shipbuilding and Allied Industries - Sudometrika-2016"), 2016, pp. 16-27.

7. Polushkin, A.V., Borchaninov, D.G., Slistin, I.V., Dmitriev, N.V., Nakhov, S.F., Using Scripting Mechanisms to Quickly Adapt Automated Workstations to User Tasks, Vserossiiskaya nauchno-prakticheskaya konferentsiya s mezhdunarodnym uchastiem "Novye tekhnologii, materialy i oborudovanie rossiiskoi aviakosmicheskoi otrasli - AKTO-2016" (All-Russian Scientific and Practical Conference With International Participation "New Technologies, Materials and Equipment of the Russian Aerospace Industry - AKTO-2016"), 2016, pp. 182-195.

8. Kaldymov, N.A., Polushkin, A.V., Nakhov, S.F., Ermakov, R.V., Slistin, I.V., Shatskov, S.N., Study of Accuracy Characteristics of the Precision Full-Angle Rotary Table for Inspecting Elements of Inertial Devices. Algorithms of Automatic Drive Control Improving Accuracy of Angular Positioning, Systemy i kompleksy avtomaticheskogo upravleniya letatelnykh apparatov, Moscow, FGUP "NPTsAP", 2013, pp. 169 -178.

9. Slistin, I.V., Aspects of Designing a Multipurpose Angle-Setting Rotary Table for Instrumentation Industry, XVI konferentsiya molodykh uchenykh "Navigatsiya i upravlenie dvizheniem" (16th Conference of Young Scientists "Navigation and Motion Control"), St. Petersburg, CSRI Elektropribor, pp. 357-364.

10. Polushkin, A.V., Kaldymov, N.A., Dmitriev, N.V., Nakhov, S.F., Lenskii, Yu.V., Yumaguzin, R.M., Plotnikov, P.K., Yankovskii, A.A., Investigation of the Influence of Spatial Instability of the Rotation Axis of the Rotary Table Platform During Measurements on the Results of Inspections of Various Types of Navigation Devices and Their Elements, 23rd St. Petersburg International Conference on Integrated Navigation Systems, St. Petersburg, CSRI Elektropribor, 2016, pp. 198-210.

11. Ermakov, R.V., Lvov, A.A., Analysis of Errors of Angle Measuring Test Bench Based on Optical Non-Contact Angle Sensor, Sbornik trudov IV Mezhdunarodnoi nauchnoi konferentsii "Problemy upravleniya, obrabotki i peredachi informatsii" (Proceedings of the 4th International Scientific Conference "Problems of Control, Information Processing and Transmission"), vol. 2, 2015, pp. 116-123.

12. Kaldymov, N.A., Slistin, I.V., Polushkin, A.V., Method to Inspect Errors of Rotary Table Circular Scale, Patent RU no. 2790074 C1, 2023.

13. Polushkin, A.V., Ermakov, R.V., Kaldymov, N.A., Nakhov, S.F., Pukhov, D.B., Yankovskii, A.A., Development and Practical Application of Automatic Angle Measuring Complex for Rotary Table Inspection, 21st St. Petersburg International Conference on Integrated Navigation Systems, St. Petersburg, CSRI Elektropribor, 2014, pp. 171-176.

14. Kaldymov, N.A., Slistin, I.V., Polushkin, A.V., Method of Determining the Error of the Angle Measuring Rotary Table, Patent RU no. 2779239 C1, 2022.

15. Metodika proverki MP 253-0114-2023 "Stoly povorotnye pretsizionnye SPP-19.001" (Verification Procedure PP 253-0114-2023 "Precision Rotary Tables SPP-19.001"), https://grmetr.ru/gosreestr/1416098-91991-24-spp-19-001

16. Polushkin, A.V., Ermakov, R.V., Kaldymov, N.A., Nakhov, S.F., Plotnikov, P.K., Lisitskii L.A., Improving the Positioning Accuracy of the Precision Rotary Table Platform by Introducing Friction Control Algorithms, 21st St. Petersburg International Conference on Integrated Navigation Systems, St. Petersburg, CSRI Elektropribor, 2014, pp. 177-182.

17. Russian Federation National Register of Measuring Equipment no. 91991-24, Precision Rotary Tables "SPP-19.001", https://grmetr.ru/gosreestr/1416098-91991-24-spp-19-001

The Goniometric Stand for Angle Encoder Calibration. Research Methods and Results

P.A. Pavlov

Department of Laser Measurement and Navigation Systes
Saint Petersburg Electrotechnical University LETI
Saint-Petersburg, Russia
ORCID 0000-0002-6254-3145

E.M. Ivashchenko

Department of Laser Measurement and Navigation Systes
Saint Petersburg Electrotechnical University LETI
Saint-Petersburg, Russia
ORCID 0000-0003-3645-4402

Abstract — **When creating gyroscopic devices and their testing tools, digital angle encoders that have passed metrological control are widely used. According to the state verification schedule, angle measuring devices operating in dynamic mode – dynamic goniometers - are used for measuring a flat angle. The paper considers a dynamic goniometer, provides an analysis of its instrumental error, which allows increasing the accuracy of angle encoder calibration.**

Keywords — dynamic goniometer, angle encoder, angle measurements

I. INTRODUCTION

The Euramet guidelines [1] state that the primary means for calibrating angle encoders is a reference angle measuring table with a measurement limit of 360° and the ability to be positioned at any given angular position.

Typically, such a table contains a high-precision optical angle encoder with more than 36,000 marks per revolution with an electronic interpolator and several optical reading heads, which allows you to obtain an angular resolution of less than 10^{-2}". It is noted that the advantage of using such encoder is the ability to calibrate angle encoder with a large number of discrete steps and work in dynamic mode.

The dynamic goniometer (DG) [2] includes the reference encoder mounted on a spindle rotating in bearings, an engine and its control system, a signal registration unit and a computer. The dynamic goniometer calibrates the encoder mounted on the DG spindle, with its joint continuous rotation with the reference encoder. The DG allows you to calibrate various types of angle encoders in dynamic mode with high accuracy and minimal span time.

Increasing the accuracy of a gyroscopic device is closely related to increasing the accuracy of its components, in particular, the reference angle encoder and increasing the accuracy of their calibration.

For this purpose, it is necessary to improve the methods and techniques for improving DG accuracy. The identification of the DG systematic error allows you to eliminate it, reducing it to a random value. Therefore, it is an important task to identify and eliminate sources of random error.

II. METHOD

A. Determination of the DG systematic error

Cross-calibration methods are used to determine the systematic error of the angle measuring table [3-5]. The systematic errors of the two angle encoders involved in the

measurements are determined by these methods without the use of additional tools. The cross-calibration method consists in performing serial turns of the tested angle encoder relative to another angle encoder at angles determined by the discreteness of the encoder, taking into account that the systematic error in angular displacement by 2π is zero.

For prisms with a small number of faces, measurements take a limited time, and the calculation process does not cause significant problems. However, if the discreteness of the tested optical angle encoder is tens of arc seconds, then the use of the method becomes technically difficult due to the large number of turns at small angles.

For DG, the cross-calibration method can be modified [6]. In this case, a ring laser is used as a tested encoder. The ring laser is characterized by a high uniformity of the angular scale created by the electromagnetic wave during its rotation. The unevenness of its angular scale is mainly located on the first two harmonics [2, 7]. This makes it possible to identify the systematic error of the DG with a small number of turns of the ring laser. The measurements showed that with 3 angular displacements of the ring laser with a step of 120 °, the methodological error does not exceed 10^{-2}", which is comparable to the random error of the measurements.

B. Method for determining the random error

The DG random processes largely determine its potential accuracy. At the same time, random processes can be stationary and non-stationary. In the case of stationary processes, their contribution to the random error can be reduced by multiple measurements.

In the case of non-stationary processes, the situation is more complicated. The reason for the non-stationary random processes in the DG are the processes in the ball bearings in which the DG spindle is fixed.

There are certain methods for analyzing non-stationary processes, for example, Allan and Hadamard variances, and Wavelet analysis. The Allan variance was proposed as a characteristic of time and frequency standards back in the 1960s [8]. However, there are no theoretical restrictions on the use of the Allan variance for the analysis of other types of measured quantities. The Hadamard variance is insensitive to linear drift, so its application in conjunction with the Allan variance is useful to refine the noise model [9]. Wavelets [10] were introduced in the 1980s for the analysis of seismic and acoustic signals, and are now widely used in the analysis of non-stationary signals. Unlike Fourier-spectra, in which local signal features (discontinuities, steps) manifest themselves in the form of barely noticeable bursts, the wavelet allows you to identify these features, as well as determine the temporary change in the frequency composition of the signal [11].

III. EXPERIMENTAL STUDIES

Experimental studies of a dynamic goniometer with a angular-contact ball bearing and an optical angle encoder with a holographic scale and two measuring heads have been carried out. The resolution of the encoder's scale is 4", the unevenness is less than 0.5". The measurements were carried out during 25 revolutions of the shaft of the dynamic goniometer at rotational speeds 0.2 π, 3 π and 5.6 π rad/s from the dynamic range of rotation speeds.

To measure the DG random error, the method proposed in [7] was used. The method includes: taking data at equal angular intervals simultaneously from two diametrically positioned measuring heads during several revolutions of the spindle; averaging over revolutions; forming an array of random variables as a deviation from the average in each revolution.

The data was read out at angular intervals equal to 0.05". The array of random variables was analyzed using Allan variances, in accordance with the recommendations of [13] and Hadamard variances. Since the data capture clock was set at equal angular intervals, the Allan variance was built depending on the angle of rotation of the DG spindle.

Figure 1 shows Allan variances at different rotational speeds.

Fig. 1. Allan variances depending on the angle of rotation at rotation speeds: 1 – 0,2 π rad/s; 2 – 3 π rad/s; 3 – 5,6 π rad/s

It can be seen from Fig. 1 that the nature of the curves at different speeds of rotation differs slightly from each other and depends on the angle of rotation. At small averaging angles (up to 1°), white noise prevails (slope $-1/2$), then sinusoidal noise prevails.

Note that the Hadamard variance at different rotational speeds has a similar character as the Allan variance. That is, at small averaging angles, white noise prevails, then sinusoidal noise prevails.

Figure 2 shows the variances of Allan and Hadamard for a rotation speed of 3 π rad/sec. Having a similar dependence on the angle of rotation, the values of the variances differ. Taking into account that the Hadamard variance, according to [9] unlike the Allan variance, is not sensitive to linear drift, it can be assumed that the difference between curves 1 and 2 (Fig. 2) is due to the fact that linear drift in the

frequency of sinusoidal noise is present in the array of random variables.

Fig. 2 Variances at a rotation speed of 3 π rad/s: 1 – Allan; 2 – Hadamard

Thus, random processes in dynamic goniometers with a ball bearing are non-stationary, white noise, sinusoidal oscillations and linear drift are appeared.

From comparing the nature of random processes taking place in a DG with an air bearing and a ball bearing [14] it follows that sinusoidal oscillations are caused by processes taking place in the bearing.

To study the change in the frequency composition of the signal over time, we use Wavelet analysis. The wavelet spectrum, which represents the values of the wavelet coefficients in the time-frequency domain, for low frequencies for a speed of 3 π rad/s is shown in Fig. 3. The frequency domain is selected in such a way as to focus on the most significant impact associated with the rotational speed of the ball bearing separator used in the DG.

Fig. 3 The wavelet spectrum of an array of random variables characterizing a random error at a rotational speed of 3 π rad/s

Figure 3 shows that frequencies in the range from 0.54 to 1.5 Hz are manifested unevenly throughout the measurement time, both the amplitude changes and the effect of frequency blurring is observed. Presumably, this is the reason for the detection of linear frequency drift when using Allan and Hadamard variations.

IV. CONCLUSION

Methods to determine and study the instrumental error of the DG have been developed. The DG design features allow you to determine the quality of its bearing assembly without using additional equipment. Using the Allan and Hadamard variance methods, it was determined that in the arrays of random variables characterizing the random error, in addition to white noise, there is sinusoidal noise, as well as linear frequency drift. The wavelet analysis confirmed the frequency change over time. These processes do not depend on the speed of rotation, therefore, in the future, an urgent task is to explore the possibilities of filtration, in order to further increase the accuracy of measurements with a dynamic goniometer.

The developed methods made it possible to create a number of angle measuring units operating in dynamic mode in the range of angular velocities $(0.1 \div 20)\pi$ rad/s with an error of $0.2" \div 1"$, which are listed in the register of measuring instruments.

REFERENCES

[1]   Emilio Prieto, M del Mar Perez Tanfer Yandayan, Joanna Przybylska, Andreas Just, Ralf Geckeler, Guidlines on the Calibration of Angular Encoders. 2018.

[2]   M. N. Burnashev, P. A. Pavlov, и Y. V. Filatov, «Development of precision laser goniometer systems», Quantum Electron., T. 43, Vol. 2, cc. 130–138, Feb. 2013, doi: 10.1070/QE2013v043n02ABEH015045.

[3]   Sim P.J., «Modern Techiques in Metrology» Singapore: World Scientific, pp. 102–121, 1984.

[4]   R. Probst, R. Wittekopf, M. Krause et al., «The new PTB angle comparator» Measurement Science and Technology no. 9, pp. 1059–1066, 1998.

[5]   W.T. Estler, «Uncertainty analysis for angle calibration using circle closure», Journal Res. Natl. inst. Stand. Technol., no. 103, pp. 141–145, 1998.

[6]   P.A. Pavlov, «Aspects of the cross-calibration method in laser goniometry», Measurement Techniques, vol. 58, No. 9, 2015, doi: 10.1007/s11018-015-0827-7.

[7]   E.P. Krivtsov, Yu.V. Filatov, « Investigation of the error in reproducing a unit of a plane angle with a ring laser », Measurement techniques, №12, pp. 11–12, 1989.

[8]   Allan, D.W., «Statistics of atomic frequency standards», Proc. IEEE, т. 54, № 2, pp. 221–230, 1966.

[9]   W.J. Riley, Handbook of Frequency Stability Analysis. Washington: U.S. Goverment Printing Office, 2008.

[10]   Mary Beth Ruskai, Gregory Beylkin, Ronald Coifman, Ingrid Daubechies, and Stephane Mallat, Wavelets and their Applications. Boston: Jones and Barlett Publisher, 1992.

[11]   A.N. Yakovlev, Introduction to wavelet-transform, NGTU Publ. 2003.

[12]   E.A. Barinova C.V. Gordeev, E.M. Ivashchenko, P.A. Pavlov, «Method and results of optical encoders random error investigation»,. Metrologia, 7, pp. 17–25, 2011.

[13]   «IEEE Std 647-2006.», IEEE Standard Specification Format Guide and Test Procedure for single-axis laser gyros..

[14]   X. Zhang, C. Yan, Y. Liu, P. Yan, Y. Wang, и L. Wu, «Dynamic Modeling and Analysis of Rolling Bearing with Compound Fault on Raceway and Rolling Element», Shock and Vibration, 2020, pp. 1–16, 2020, doi: 10.1155/2020/8861899.

Optimal Deadband Estimation Algorithm for Fiber Optic Gyroscopes

A.E. Morozov

Perm Scientific-Industrial Instrument Making Company
Perm, Russia
Andrew_Morozov_@mail.ru

The main objective of this work is to develop a method for determining the deadband of fiber optic gyroscopes during the production stage for controlling a target parameter that affects the quality of the final product. The proposed algorithm can be used both at the adjustment stage to select the feedback coefficient, which reduces the size of the deadband, and at the stage of control tests in the operating temperature range to assess the quality of the manufactured sensor. The study of the algorithm has shown the limiting accuracy of the results obtained, which can be increased by changing the test method. The proposed algorithm is based on the least squares method, from which all the advantages are inherited: the absence of preliminary signal filtering and robustness to outliers. The work demonstrates a way to optimize calculations and automate the adjustment process.

Keywords — deadband, dead zone, fiber optic gyroscopes, least square method, automation.

I. INTRODUCTION

Fiber optic gyroscopes (FOG) are navigation-grade angular velocity sensors that require precise rotation measurement. In closed-loop interferometric fiber optic gyroscopes, a phenomenon known as the deadband (DB) is observed. Much research [1,2,3] has focused on developing methods to eliminate or compensate for this error in the FOG output signal. However, in mass production, measuring the deadband's magnitude remains necessary to ensure the absence of errors in specific units.

To minimize the impact of the deadband, adjusting the FOG's feedback parameters is employed with measuring the current deadband, which requires some form of quantitative assessment. This necessitates automated test processing and the development of a reliable algorithm for calculating the target deadband parameter.

Deadband observation tests involve slowly and uniformly changing the direction of the sensitivity axis near the zero projection of the Earth's angular velocity. This paper proposes a novel algorithm based on the least squares method (LSM), that enables efficient calculation of the deadband's magnitude. LSM offers advantages such as robustness to outliers and the ability to handle data without preliminary filtering.

II. ALGORITHM DESCRIPTION

The proposed method for calculating the deadband of a fiber optic gyroscope is based on the least squares method. However, instead of determining the coefficients of a linear regression, the task is interpreted as finding a state vector. To achieve this, it is necessary to define a model for the behavior of the deadband error, with the estimation set to zero within the expected deadband region (1):

$$\begin{cases} \hat{\omega}(t) = \varepsilon t + \omega_0, \ for\ t < t_1 \\ \hat{\omega}(t) = 0, \ for\ t_1 \leq t \leq t_2 \\ \hat{\omega}(t) = \varepsilon t + \omega_0, \ for\ t > t_2 \end{cases} \quad (1)$$

where $\hat{\omega}(t)$ – evaluation of the FOG signal in the low angular velocity zone [°/ч], [°/ч], ε – rate of change of Earth's rotation projection onto the sensitivity axis [°/ч²]; ω_0 – Earth's rotation projection onto the sensitivity axis at the initial position [°/ч]; t – time elapsed since the start of the test [ч]; t_0 – time of true zero angular velocity projection, denoted as ω_0/ε [ч]; t_1 – time of entering the FOG deadband [ч]; t_2 – time of exiting the FOG deadband [ч].

While the actual behavior of the FOG signal within the deadband might exhibit slight variations [4], this representation serves as a sufficiently accurate approximation for estimating the deadband magnitude (Figure 1). Since the deadband, by definition, pertains only to the region of small angular velocities, trigonometric functions in the model description can be neglected based on the small-angle approximation. This allows us to transition to a linear law for describing the test with a zero deadband.

The coefficients of the linear regression are predetermined from the testing methodology: $\varepsilon = U_{Up/N} \cdot \sin(\Omega t) \approx U_{Up/N} \cdot \Omega t$ and $\omega_0 = U_{Up/N} \cdot \sin(\alpha) \approx U_{Up/N} \cdot \alpha$, where Ω is the specified rotation rate of the test equipment [rad/s], α is the initial positioning angle [rad], and the projections of the Earth's angular velocity $U_{Up}=U \cdot \sin(\varphi)$ and $U_N=U \cdot \cos(\varphi)$ are chosen based on the direction of the rotation projection on the test equipment (U is the Earth's angular velocity, φ is the latitude of the test location). In practice, these coefficients can be easily computed numerically before starting the algorithm. However, incorporating this a priori information can enhance the accuracy of the estimated parameters.

Fig. 1. . Visualisation of the algorithm result

Due to the presence of discontinuities in the function within the FOG signal's deadband, the fundamental properties of linearity are violated. Consequently, novel approaches are required to find the optimal solution for this piecewise-smooth function. One such method is multi-criteria optimization [5], which allows for retaining the optimality criterion based on the least squares method $(\omega - \hat{\omega})^2 \to \min$ or selecting an analogous criterion to reduce computation time. To further optimize calculations, sliding data processing methods can be employed, transitioning to an asymptotic algorithm complexity of $O(N^2)$. An example of the main algorithm loop, following the initialization of all initial conditions and the initial FOG signal estimation $\hat{\omega}(t)$ with zero unknown variables, is presented as a flowchart in Figure 2.

Fig. 2. Block diagram of the main cycle of the algorithm

Where SD - storage area of the minimum standard deviation value, DB - storage area of the optimal value of the deadband with the minimum standard deviation, N - last element of the data array.

Depending on the specific objectives, the target parameter can be defined as either the entire deadband region $(t_2 - t_1)$ or the maximum deviation from the zero signal $\max(t_0 - t_1; t_2 - t_0)$. When evaluating the quality of an angular velocity sensor to meet the requirements of an inertial navigation system, it is essential to consider the maximum deviation from the zero signal. This is because the accuracy of gyrocompassing depends on the magnitude of this parameter.

III. ALGORITHM ACCURACY ANALYSIS

The primary errors in calculating the deadband using the proposed algorithm can be categorized as methodological (a single realization with different pre-averaging of the signal yields different results) and instrumental (different realizations with the same pre-averaging of the signal produce different results). For visualizing the deadband manifestation, it is convenient to use signal averaging with a certain window for smoothing the noise component. However, any filtering distorts the entry/exit boundaries of the deadband and introduces methodological errors in

subsequent data analysis. The proposed method does not require preliminary signal smoothing and can be applied directly to the raw data. Figure 3 illustrates the dependence of the distortion in the deadband estimation on the pre-averaging time of the signal with several different noise characteristics, using simulated data.

Fig. 3. Dependence of the error in determining the deadband value on averaging

Instrumental errors pertain to the accuracy of the measured signal. For FOGs, the standard deviation (SD) is determined based on its noise characteristics, primarily angle random walk, which follows a normal distribution: $\sigma = ARW/\sqrt{\Delta t}$ (where Δt is the sensor sampling period). The accuracy of the deadband calculation is inversely proportional to the square root of the time spent within the deadband region: $\sigma_{DB} \sim \sigma/\sqrt{t}$. Typically, it is not feasible to modify the sensor's measurement accuracy to obtain a more adequate deadband estimation. Therefore, achieving the required accuracy is ensured by increasing the time spent within the deadband region. This can be accomplished by reducing the rate of change of the Earth's angular velocity projection. However, in certain situations, the minimum rotation speed of the test equipment might be insufficient. In such cases, one can employ rotations of the gyroscope's sensitive axis with discrete steps and hold each position for the required amount of time. However, it is important to note that the estimated deadband will also be discrete, depending on the chosen step size. It is permissible to utilize combined continuous and discrete rotations, as the calculation algorithm remains unchanged for either method. An alternative approach is to repeat the test k times and average the obtained results, which is equivalent to increasing the time spent within the deadband region by a factor of k times.

Fig. 4. Accuracy of determining the deadband size under various conditions

The smaller the actual deadband magnitude, the more challenging it becomes to obtain a reliable estimation amidst measurement noise (Figure 4). The dispersion of the calculated deadband estimation not only varies with the size of the actual deadband but also transitions from a normal distribution of a random variable to a Laplace distribution (Figure 5). When selecting a testing methodology, it is crucial to ensure that the calculation accuracy at zero deadband overlaps with the required sensor accuracy or the required gyrocompassing accuracy in the case of an inertial navigation system.

Fig. 5. Distribution of the calculated estimate of the deadband for different given (true) values of the deadband.

IV. APPLIED USE OF THE ALGORITHM

To reduce the deadband level during FOG production, feedback coefficient adjustment is employed. This operation involves conducting a series of tests to verify the FOG's deadband with a predefined range of coefficients and selecting the optimal one with the least deadband magnitude. After each test, there is a need for quantitative assessment of the measured value, which can be obtained using the proposed algorithm. Having a numerical deadband estimation at each step makes finding the optimal coefficient with minimal deadband magnitude a trivial task (Figure VI).

Fig. 6. Dependence of the deadband on the feedback coefficient

The availability of an algorithm for calculating the deadband estimation and searching for the optimal feedback coefficient with the required accuracy allows for automation of the adjustment process. By connecting the workstation to the laboratory stand and the fiber-optic gyroscope, it is possible to control the stand and obtain immediate results from the sensor. To save time, a fast test cycle can be performed to preliminarily identify the vicinity of the optimal feedback coefficient, followed by a slow cycle for precise determination of the optimal feedback coefficient.

V. CONCLUSION

A high-quality and optimized algorithm for calculating the deadband of fiber-optic gyroscopes (FOGs) has been proposed. This algorithm allows for the determination of target parameters from test results. Methodological errors in deadband estimation calculations have been reduced to zero, and accuracy is limited only by the sensor's instrumental errors – specifically, noise characteristics. Recommendations have been provided for testing methods to achieve the required accuracy levels. The algorithm is suitable for any chosen methodology: continuous, discrete, or combined. Utilizing the proposed algorithm in production enables automation of test calculations and reduces risks associated with human error. Additionally, having an estimate of the target deadband parameter allows for the automation of the FOG adjustment process.

REFERENCES

[1] Xiao-Feng, Z. Research on Dead Zone Error Reduction Technique in Closedloop Fiber Optic Gyroscope / Z. Xiao-Feng, Z. Gui-cai // Piezoelectrics & Acoustooptics. — 2009. — No .31 (2)

[2] Ning-Fang, S. Method For Eliminating The Deadband Of FOG / S. Ning-Fang, W. X. Xiao, W. Zhan-Jun // Journal of chinese inertial technology. — 2006. — No. 14 (4) . — Pp.. 177-179.

[3] Kurbatov, A.M., Kurbatov, R.A. & Goryachkin, A.M. Fiber-Optic Gyroscope Accuracy Improvement by Suppressing the Parasitic Effects in Integrated Optic Phase Modulators. Gyroscopy Navig. 10, 256–267 (2019). https://doi.org/10.1134/S2075108719040114

[4] Alejnik A. S. «Research and improvement of the accuracy parameters of fiber-optic gyroscopes», Dissertation for the degree of Candidate of Technical Sciences, ITMO university, Saint Petersburg – 2012. Pp. 69-98.

[5] Verzhbickij V.M. «Fundamentals of numerical methods: textbook for universities», Grad. Sch., 2002.

Recursive and Nonrecursive Algorithms Applied to Navigation Data Processing: Differences and Interrelation with Factor Graph Optimization Algorithms

O.A. Stepanov
Concern CSRI Elektropribor, JSC
ITMO University
St. Petersburg, Russia
soalax@mail.ru

A.V. Motorin
Concern CSRI Elektropribor, JSC
ITMO University
St. Petersburg, Russia
motorin.a@mail.ru

V.P. Zolotarevich
Concern CSRI Elektropribor, JSC
ITMO University
St. Petersburg, Russia
zolotarevich@yandex.ru

A.M. Isaev
Concern CSRI Elektropribor, JSC
ITMO University
St. Petersburg, Russia
itmo_student@mail.ru

Yu.A. Litvinenko.
Concern CSRI Elektropribor, JSC
ITMO University
St. Petersburg, Russia
ya_litvinenko@mail.ru

Abstract—**Within the Bayesian approach, a nonlinear estimation problem is considered and the solution algorithms based on linearization and using recursive and nonrecursive measurement processing schemes are discussed. A recursive iterative batch linearized smoother is proposed, which has the advantages of the non-recurrent processing scheme and provides accuracy close to potential. The relationship of synthesized algorithm with factor graph optimization algorithms is discussed.**

Keywords—nonlinear estimation, Bayesian approach, factor graph, optimization, recursive algorithms, nonrecursive algorithms, Kalman filter

I. INTRODUCTION

Algorithms based on Bayesian estimation theory have been widely used in navigation information processing. Significant progress has been made in constructing these algorithms, especially those applied to linear estimation problems solved using recursive Kalman-type algorithms which are essentially universal. As for nonlinear estimation, there is no such a universal algorithm for them, so the development of efficient algorithms, including those for navigation applications, is still a relevant task. Recently, a lot of literature devoted to the navigation information processing is focused on nonrecursive algorithms based on the methods of factor graph optimization, known as factor graph optimization algorithms (FGOA) [1–14]. The factor graph optimization methods are being actively implemented in the algorithms for integrated navigation information processing, especially in robotic applications [1, 3, 5, 7–14], and in the algorithms for simultaneous localization and mapping (SLAM) [2, 4, 5, 10–12].

The idea of FGOA consists in reducing the estimation problem to minimization of the functional, which is a quadratic form of a high-dimensional vector with a sparse matrix, i.e. a matrix with a small number of non-zero elements. The advantages of FGOA in comparison with recursive Kalman-type filtering methods (primarily the extended Kalman filter (EKF)) include higher accuracy [1, 3, 7, 9], robustness to possible measurement gaps and non-Gaussian measurement errors [3, 9, 13], and lower computational complexity of algorithms thanks to the sparse form of matrices [2, 7]. It should be noted that in most works on the FGOA application, little attention is paid to the real reasons for the mentioned advantages of the obtained algorithms in terms of accuracy and robustness compared to the estimation algorithms synthesized within the Bayesian approach; the authors usually limit themselves to comparison with the EKF. However, such a comparison would be useful, since both EKF and FGOA are actually approximate solutions to the optimal estimation problem formulated within the Bayesian approach. In other words, both algorithms are suboptimal from the standpoint of the Bayesian approach, and both are based on linearization. At the same time, it should be borne in mind that the main

distinguishing feature of the estimation algorithms implemented using the software for FGOA application is associated with a nonrecursive (batch) scheme of measurements processing. In this case, in contrast to the EKF, the estimate is found on assumption that the entire set of measurements (a batch of measurements) accumulated by the current time is used simultaneously at the current step.

Taking into account the above, it is advisable to discuss the features of factor graph optimization-based algorithms for solving the estimation problems, basing on the estimation problem statement within the Bayesian approach and taking into account the large experience gained in the synthesis of algorithms and research of their efficiency, without limiting ourselves to comparison only with the EKF, which is the purpose of this presentation.

II. BAYESIAN STATEMENT OF NONLINEAR ESTIMATION PROBLEM, AND RECURSIVE AND NONRECURSIVE SCHEMES FOR ITS SOLUTION

A. Problem Statement

We consider the problem of nonlinear estimation of a random *n*-dimensional vector described by means of a shaping filter

$$x_k = f_k(x_{k-1}) + \Gamma_k w_k + u_k \qquad (1)$$

according to *m*-dimensional measurements in the following form:

$$y_k = h_k(x_k) + v_k. \qquad (2)$$

In these relations, it is assumed that *k* is the discrete time index; $u_k = (u_{1k},\dots,u_{nk})^T$ is the *n*-dimensional vector of known input signals (if double subscripts are used, the first subscript hereinafter corresponds to the number of component, and the other one to the time); x_0 is the *n*-dimensional random Gaussian vector with a given probability density function (PDF) $p(x_0) = N(x_0; \bar{x}_0, P_0)$; w_k is the n_w-dimensional zero-mean Gaussian white noise which is independent of x_0 and has a known covariance matrix Q_k with dimension $n_w \times n_w$; Γ_k is matrix with dimension $n \times n_w$; v_k is *m*-dimensional zero-mean discrete Gaussian white noise which is independent of x_0 and w_k and has a covariance matrix R_k. The specificity of the nonlinear estimation problem under study lies in the fact that $f_k(\bullet)$, $h_k(\bullet)$ are known nonlinear *n*- and *m*-dimensional vector functions which describe the behavior of the state vector and the model of measurements.

The solution to the estimation problem within the Bayesian approach essentially consists in obtaining some optimal estimate of the state vector $\hat{x}_k^{opt}(Y_k)$ according to the measurements $Y_k = (y_1^T,\dots,y_k^T)^T$ and, if possible, a respective conditional covariance matrix of estimation errors $P_k^{opt}(Y_k)$, which characterizes its current (corresponding to a particular set of measurements Y_k) accuracy.

In the theory of estimation, the problems are usually divided into three groups: filtering, prediction, and smoothing problems [16, 17]. In this work, we will consider the algorithms used in filtering and smoothing problems. The specific feature of filtering is that the estimate at some *j*-th time point is obtained using only the measurements with indices $i \leq j$, while in smoothing the entire available set of measurements $i = \overline{1.k}$ is used.

Estimate $\hat{x}_k^{opt}(Y_k)$, which is RMS-optimal, is known to be defined as follows [16, 17]:

$$\hat{x}_k^{opt}(Y_k) = \int x_k\, p(x_k / Y_k)\, dx_k, \qquad (3)$$

where $p(x_k / Y_k)$ is PDF that is posterior (conditional) in relation to measurements.

It should be noted that within the Bayesian approach, along with the RMS-optimal estimate (3), the estimate (4) corresponding to the maximum of the a posteriori PDF is often found:

$$\hat{x}_k(Y_k) = \arg\max_{x_k} p(x_k / Y_k). \qquad (4)$$

It is clear that it is necessary to know the a posteriori PDF $p(x_k / Y_k)$, in order to find the estimates of the form (3), (4).

B. Recursive Algorithms

As a rule, when discussing the proposed algorithms, it is often assumed by default that a recursive processing scheme will be used, i.e., algorithms that a recursive in relation to measurements are constructed. Such algorithms are synthesized widely using the recursive relationships for a posteriori PDF, which have the following forms [16, 17]:

$$p(x_k / Y_k) = \frac{p(y_k / x_k)\, p(x_k / Y_{k-1})}{\int p(y_k / x_k)\, p(x_k / Y_{k-1})\, dx_k}, \qquad (5)$$

$$p(x_k / Y_{k-1}) = \int p(x_k / x_{k-1})\, p(x_{k-1} / Y_{k-1})\, dx_{k-1}, \qquad (6)$$

where $p(y_k/x_k)$ is the likelihood function; $p(x_k/Y_{k-1})$ is the PDF of prediction; $p(x_k/x_{k-1})$ is transition PDF.

Taking into account (1), (2) and the above assumptions, we can write:

$$p(y_k / x_k) = N(y_k; h_k(x_k), R_k), \qquad (7)$$

$$p(x_k / x_{k-1}) = N\left(x_k; f_k(x_{k-1}) + u_k, \Gamma_k Q_k (\Gamma_k)^T\right). \qquad (8)$$

To construct efficient suboptimal algorithms, it is necessary to develop procedures with low computational load, which provide an estimate $\hat{x}_k^{sub}(Y_k)$ slightly differing from nonlinear optimal estimate (3) in terms of accuracy, and also provide a computational accuracy characteristic in the form of a covariance matrix of estimation errors, which is consistent to real. As was mentioned in the introduction, the algorithms meeting the latter requirement are called consistent [18].

Another popular group of algorithms which is of interest for the purposes of this work is intended for the above-mentioned smoothing problems. For these, recursive relationships for PDF have also been obtained to facilitate the construction of appropriate algorithms, and for the linear problems, computationally efficient algorithms based on the Kalman filter (KF) and subsequent processing of the respective estimates in reverse time have been developed [16, 19].

C. Nonrecursive Algorithms

Along with recursive algorithms, the desired estimate of the state vector at an arbitrary time point can be obtained by finding it as part of the estimate for vector $X_k = (x_1^T, ... x_k^T)^T$ which includes the values of x_i for all time points $i = \overline{1.k}$, by the vector measurement $Y_k = (y_1^T, ... y_k^T)^T$. To obtain this estimate, it is necessary to solve the problem formulated as follows. Constant vector X_k with a priori PDF $p(X_k)$ is to be estimated by $m \times k$ measurements

$$Y_k = H(X_k) + V_k, \tag{9}$$

where $H(X_k) = (h_1^T(x_1), ... h_k^T(x_k))^T$, $V_k = (v_1^T, ... v_k^T)^T$ is the composite vector of measurement errors with a block-diagonal covariance matrix $\mathbf{R}_k$ with matrices R_i, $i = \overline{1.k}$ on the main diagonal. To obtain the desired estimate, a posteriori PDF $p(X_k / Y_k)$ should be known, which is defined as

$$p(X_k / Y_k) = cp(X_k)p(Y_k / X_k), \tag{10}$$

where c is normalizing factor.

Since the estimated sequence is of Markov type, it is easy to verify that $p(X_k)p(Y_k / X_k)$ can be represented in the form

$$p(X_k)p(Y_k / X_k) = p(x_0)\prod_{i=1}^{k} p(x_i / x_{i-1})\prod_{i=1}^{k} p(y_i / x_i). \tag{11}$$

Taking into account the Gaussian form of the vector of initial conditions and the vectors of generating and measurement noise, we can write the equations

$$p(X_k) = c_1 \exp\left\{\frac{1}{2}J_1(X_k)\right\}, \tag{12}$$

$$p(Y_k / X_k) = c_2 \exp\left\{\frac{1}{2}J_2(X_k)\right\}, \tag{13}$$

in which

$$J_1(X_k) = -\left(\begin{matrix} x_0^{\mathrm{T}}P_0^{-1}x_0 + \sum_{i=1}^{k}(x_i - f_i(x_{i-1}) - u_i)^{\mathrm{T}} \times \\ \times\left(\Gamma_i^{\mathrm{T}}Q_i\Gamma_i\right)^{-1}(x_i - f_i(x_{i-1}) - u_i) \end{matrix}\right), \tag{14}$$

$$J_2(X_k) = -\left(\sum_{i=1}^{k}(y_i - h_i(x_i))^{\mathrm{T}}R_i^{-1}(y_i - h_i(x_i))\right), \tag{15}$$

where c_1, c_2 are normalizing factors.

Thus,

$$p(X_k)p(Y_k / X_k) = c\exp\left\{\frac{1}{2}J(X_k)\right\}, \tag{16}$$

where

$$J(X_k) = -\left(\begin{matrix} x_0^{\mathrm{T}}P_0^{-1}x_0 + \sum_{i=1}^{k}(x_i - f_i(x_{i-1}) - u_i)^{\mathrm{T}} \times \\ \times\left(\Gamma_i^{\mathrm{T}}Q_i\Gamma_i\right)^{-1}(x_i - f_i(x_{i-1}) - u_i) + \\ +\sum_{i=1}^{i}(y_i - h_i(x_i))^{\mathrm{T}}R_i^{-1}(y_i - h_i(x_i)) \end{matrix}\right). \tag{17}$$

When implementing the nonrecursive scheme, the estimate of the entire composite vector X_k is defined as

$$\hat{X}_k^{opt}(Y_k) = \int X_k p(X_k / Y_k) dX_k. \tag{18}$$

If an estimate only at the time point $j \le k$ is required, we can replace (3) in the nonrecursive scheme with the following:

$$\hat{x}_{j/k}^{opt}(Y_k) = \int x_j p(X_k / Y_k) dX_k, \tag{19}$$

where $\hat{x}_{j/k}^{opt}$ is the estimate obtained by k measurements.

Usually, when $j = k$, the notation $\hat{x}_{j/j}^{opt} = \hat{x}_j^{opt}$ is used. It obviously follows from this that when the estimate of the composite vector (18) is found, in fact, the estimate corresponding to the solution of smoothing problem is found for all its components $j < k$. There are two special cases of nonlinear problems. One of them relates to the estimation of Gaussian Markov sequences when the forming filter equations are linear and only the equations of measurements (2) are nonlinear. In this case, there are no problems with finding PDF $p(X_k)$, since it will be Gaussian. In the other special case, the measurements are linear, and equations (1) are nonlinear. In this situation, there are no problems with calculating $p(Y_k / X_k)$ and the main challenge consists in defining $p(X_k)$. Moreover, even finding the Gaussian approximation of this density is a nontrivial task [23].

It is easy to see that for solving the linear Gaussian problem in which in (1) and (2) $f_k(x_{k-1}) = \Phi_k x_{k-1}$, $h_k(x_k) = H_k x_k$, where Φ_k and H_k are known matrices with dimensions $n \times n$ and $m \times n$, the following relationships will be true for the optimal estimate of the composite vector X_k and the covariance matrix corresponding to it:

$$\hat{X}_k^{opt}\left(Y_k\right) = \overline{X} + \mathbf{P}_k \mathbf{H}_k^T \mathbf{R}_k^{-1}\left(Y_k - \mathbf{H}_k \overline{X}\right), \qquad (20)$$

$$\mathbf{P}_k\left(Y_k\right) = (\mathbf{P}_X^{-1} + \mathbf{H}_k^T \mathbf{R}_k^{-1}\mathbf{H}_k)^{-1}, \qquad (21)$$

where $\mathbf{H}_k$ is the block-diagonal matrix with matrices H_i, $i = \overline{1.k}$ on the main diagonal, and $\mathbf{P}_X$ is the covariance matrix corresponding to the composite vector $X_k = (x_0^T, \ldots x_k^T)^T$ [17, p. 337].

It is well known that the linear Gaussian problem, according to the recursive scheme, can be solved using known relations for a discrete recursive Kalman filter in the filtering mode. If it is necessary to solve the smoothing problem, then computationally efficient smoothing algorithms should be used. The essence of these algorithms lies in the fact that after the filtering problem has been solved using the KF, the obtained estimates are processed in reverse time, using the relations similar to the KF [16, 19, 25]. In the context of this work, a distinctive feature of the recursive processing scheme is that only the matrices with dimension n are subject to conversion when implementing the estimation algorithms, while in the nonrecursive scheme, the matrices with dimension $n \times k$ are converted.

III. Suboptimal Recirsive and Nonrecursive Bayesian Algorithms Based on Linearization

It is important to emphasize that in the case of fixed number of measurements used, it does not matter for an estimate that is optimal in one sense or another, whether the processing scheme used for its calculation is recursive or nonrecursive. The situation is different when using suboptimal algorithms, especially those based on a linearized representation of nonlinear functions. It is such algorithms that are considered further, i.e. it is assumed that

$$f_i(x_{i-1}) \approx f_i(x_i^{\pi_1}) + \Phi_i(x_{i-1} - x_i^{\pi_1}),$$
$$h_i(x_i) \approx h_i(x_i^{\pi_2}) + H_i(x_i - x_i^{\pi_2}), \qquad (22)$$

where Φ_i, H_i are Jacobi matrices of functions $f_i(x_{i-1})$, $h_i(x_i)$, calculated at the linearization points $x_i^{\pi_1}$, $x_i^{\pi_2}$.

The simplest recursive algorithm based on linearization is the so-called linearized KF (Alg. 1) and the corresponding linearized Kalman smoother (KS). Their distinctive feature is that a priori mathematical expectations, i.e. $x_i^{\pi_1} = \overline{x}_{i-1}, x_i^{\pi_2} = \overline{x}_i$, are chosen as the linearization points. The algorithm number indicated in parentheses (Alg. 1) is used to designate it in the simulation section. The same numbering is used for other algorithms.

Algorithm Alg. 2 which is a EKF and the corresponding extended KS has been widely used in solving applied problems. Their distinctive feature is that $x_i^{\pi_1} = \hat{x}_{i-1}, x_i^{\pi_2} = \hat{x}_{i/i-1}$, i.e., the estimate from the previous step and its predicted values generated at the current step are selected as linearization points.

Another class of algorithms includes filters with local iterations, or iterative EKF (Alg. 3), in which the same measurement is processed repeatedly (iteratively) at the current step and at the same time, during the next iteration,

the linearization points are adjusted in accordance with the results of processing at the previous iteration.

Note that in this paper we study the algorithms intended for the problems where the a posteriori PDF of the estimated state vector changes in the process of measurements accumulation from multi-extreme density to single-extreme one, often close to the Gaussian form. It is clear that the application of algorithms using the Gaussian approximation of a posteriori PDF at each step and involving an estimate from the previous step for calculating the linearization points will be ineffective in such problems, since at the initial stage, when there are little measurements, form of a posteriori PDF differs significantly from Gaussian, and this leads to significant linearization errors.

At the same time, as was stated in [21, p. 263], it is logical to expect that starting from the moment when the a posteriori PDF becomes single-extreme, algorithms such as recursive batch linearized smoother (Alg. 4) processing a set (batch) of measurements that have been accumulated by the current time will be efficient. Obviously, the greatest effect will be obtained using an iterative batch linearized filter (Alg. 5). The specific feature of this algorithm is that simultaneous processing of all measurements accumulated by the current time point is carried out repeatedly (iteratively) during linearization at fixed points, and at each iteration the linearization points are adjusted simultaneously in accordance with the results of the previous iteration. Iteration is understood as the re-processing of all measurements accumulated by the current moment. At the same time, it is clear that the use of such an algorithm will require a high computational load. However, when implementing this algorithm, in order to reduce the amount of calculations, it is proposed to use recursive iterative batch linearized smoother (Alg. 6). There are two stages at each iteration of this algorithm. The first one is filtering for all time points, using a recursive linearized KF, with memorization of the obtained estimates $\hat{x}_i$ and covariance matrices $\hat{P}_i$, $i = \overline{1.k}$. The second stage is calculation of estimates in the smoothing mode and their corresponding covariance matrices in reverse time, using the expressions similar to the KF [16, 19, 25]. Obviously, the resulting estimates will completely coincide with estimates (19) for Alg. 5, since they also correspond to the solution of the smoothing problem if the linearization points have been selected identically. These estimates are then used as new linearization points at the next iteration, and all measurements accumulated by the current time point are processed again.

IV. Example

Let us consider an example. We assume that it is required to estimate an exponentially correlated sequence x_k described by means of a linear forming filter:

$$x_k = \Phi x_{k-1} + \Gamma w_k + u, \qquad (23)$$

where $\Phi = e^{-\alpha\Delta t}$, $\Gamma = \dfrac{\sqrt{2\sigma^2\alpha}}{\alpha}\left(1 - e^{-\alpha\Delta t}\right)$, σ^2 is the process variance, α is the inverse of the correlation interval τ_k, and

u is a known time-constant input. Nonlinear measurements (2) have the form:

$$y_k = h_1 + h_2 x_k + h_3 x_k^2 + h_4 x_k^3 + v_k, \qquad (24)$$

where h_1, h_2, h_3, h_4 are known values. We carry out comparative simulation using the statistical test method according to the methodology described in [26]. The algorithms are compared to an algorithm aimed at calculating the RMS-optimal estimates and implemented with the use of a particle filter (OPT) with the number of particles $N = 20{,}000$ [16, 21, 30].

Simulation was performed with the following parameters: $\sigma = 1.5$, $\alpha = 0.1$, $r = 0.1$, $h_1 = 0.0875$, $h_2 = -0.1825$, $h_3 = 0.01$, $h_4 = 0.01$, $u = 1$, $L = 10{,}000$, where

L is the number of realizations in the statistical test method, the sampling interval $\Delta t = 1\,s$, the time of simulation $T = 20\,s$. The number of iterations of repeated processing of measurements in the iterative algorithms was assumed to be equal to 3. Solid blue line in Fig. 1 (left) shows the nonlinear function (24); the black dotted lines indicate the a priori region $\pm 3\sigma$, the red line indicates the value of one realization of the estimated sequence at time point $k = 1$, and the purple line indicates the measurement corresponding to this realization value. The graphs in the middle and on the right of Fig. 1 depict a posteriori PDF $p(x_k / Y_k)$ constructed for this realization, using sequential Monte Carlo methods at time points $k = 1$ and $k = 10$.

Fig. 1.   Nonlinear function of measurements and a posteriori PDF for $k = 1$ and $k = 10$.

In the simulation, unconditional and averaged estimated variances were calculated for every μ-th algorithm as follows:

$$G_k^\mu \approx \frac{1}{L} \sum_{j=1}^{L} (x_k^{(j)} - \hat{x}_k^\mu (Y_k^{(j)}))^2, \quad \tilde{G}_k^\mu \approx \frac{1}{L} \sum_{j=1}^{L} P_k^\mu (Y_k^{(j)}). \quad (25)$$

All curves in Fig. 2 were obtained at time k using measurements with indices $i \le k$. According to the simulation results shown in Fig. 2, Alg. 2 and Alg. 3 do not work, which is fully explained by the evolution pattern of the a posteriori PDF shown in Fig. 1, which at the initial time point is multi-extreme, and at $k = 10$ it is single-extreme. The accuracy and the level of consistency of Alg. 1 and Alg. 4, as well as Alg. 5 and Alg. 6 coincide. At the same time, the results of Alg. 5 and Alg. 6 at $k \ge 10$ are close to the potential accuracy corresponding to the OPT.

The analysis of computational costs showed that Alg. 1 is ≈ 4.5 times simpler than Alg. 4, and Alg. 6 is ≈ 6.5 tomes simpler than Alg. 5.

Note that the obtained results fully confirm the conclusions made in the previous section: the recursive Alg. 6 coincides with the nonrecursive Alg. 5 in filtering and smoothing problems, but its computational complexity is much lower. It is very important that both of them provide accuracy close to potential, starting from some time point. Similar results confirming the efficiency of Alg. 5 and Alg. 6 were also obtained in trajectory tracking [15].

Fig. 2.   Unconditional and averaged estimated values of RMS errors.

V.   FACTOR GRAPH OPTIMIZATION ALGORITHMS. DIFFERENCE AND INTERRELATION WITH BAYESIAN ESTIMATION ALGORITHMS

Taking into account the linearized representation of functions (22), criterion (17) can be represented in a quadratic form:

$$J(X) = -(AX + b)^T \mathbf{S}(AX + b), \qquad (27)$$

where

$$A = \begin{bmatrix} E & 0 & 0 & ... & 0 & 0 \\ -\Phi_1 & E & 0 & ... & 0 & 0 \\ ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & E & 0 \\ 0 & 0 & 0 & ... & -\Phi_k & E \\ 0 & -H_1 & 0 & ... & 0 & 0 \\ 0 & 0 & -H_2 & ... & 0 & 0 \\ ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 0 & -H_k \end{bmatrix}; \quad b = \begin{bmatrix} x_0 \\ -\tilde{u}_1 \\ ... \\ -\tilde{u}_k \\ \tilde{y}_1 \\ \tilde{y}_2 \\ ... \\ \tilde{y}_k \end{bmatrix};$$

$$\mathbf{S} = \begin{bmatrix} P_0^{-1} & 0 & ... & 0 & 0 & 0 & ... & 0 \\ 0 & (\Gamma_1^\tau Q_1 \Gamma_1)^{-1} & ... & 0 & 0 & 0 & ... & 0 \\ ... & ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & ... & (\Gamma_k^\tau Q_k \Gamma_k)^{-1} & 0 & 0 & ... & 0 \\ 0 & 0 & ... & 0 & R_1^{-1} & 0 & ... & 0 \\ 0 & 0 & ... & 0 & 0 & R_2^{-1} & ... & 0 \\ 0 & 0 & ... & 0 & 0 & 0 & ... & 0 \\ 0 & 0 & ... & 0 & 0 & 0 & ... & R_k^{-1} \end{bmatrix},$$

and $\tilde{u}_i = u_i + f_i(x_i^{\pi_1}) - \Phi_i x_i^{\pi_1}$, $\tilde{y}_i = y_i - h_i(x_i^{\pi_2}) + H_i x_i^{\pi_2}$. The estimation problem can thus be reduced to an optimization problem, i.e., finding the value of X that provides the minimum of the quadratic form (27). Its solution, in turn, is reduced to solving a set of linear algebraic equations [27]:

$$A^T \mathbf{S} A X = A^T \mathbf{S} b, \tag{28}$$

which can be easily verified by taking the derivative of criterion (27) with respect to X and equating it to zero. Note that in many cases it is difficult to obtain a solution to set (28) by direct conversion of matrix $A^T \mathbf{S} A$ due to its large dimension. At the same time, it should be borne in mind that this matrix, like matrices A and $\mathbf{S}$, is sparse. Moreover, matrix $A^T \mathbf{S} A$ is tridiagonal, i.e. it has non-zero blocks only on the principal and lateral diagonals. Reducing the estimation problem to the problem of the quadratic form optimization for a high-dimensional vector with sparse form of the mentioned matrices is, in fact, the reason for using the FGOA.

Note that historically, the need to minimize the quadratic criteria in the form of (27) with sparse matrices occurred in various problems not related to the theory of estimation. In particular, this need may occur when solving partial differential equations, including the calculations of elasticity, hydrodynamics, and acoustics [27].

It follows from the above that the problems solved with FGOA are similar to those solved within the Bayesian approach using suboptimal nonrecursive algorithms based on linearization. Their common distinguishing features are: a minimizable criterion represented by a quadratic form for a high-dimensional vector, and sparse type of matrices making up this form. It is these features that explain the popularity of these algorithms in solving the estimation problems. The nonlinear nature of the criterion is taken into account in the FGOA and in the algorithms synthesized within the Bayesian approach in basically the same way,

based on the iterative scheme for processing a batch of measurements, which makes it possible to refine the linearization points. In the terminology of the FGOA, in this case, this is called re-linearization [13]. In the linearized problems, much depends on the choice of the linearization point, but in general, the FGOA will correspond to a recursive iterative batch linearized smoother (Alg. 6) which combines the advantages of nonrecursive refinement of the linearization point, and recursive measurement processing. Hence, the results achieved using the FGOA and nonrecursive algorithms will be the same in terms of accuracy. The advantage of the FGOA computing characteristic over the nonrecursive algorithms is that they take into account the sparseness of the matrices. At the same time, following [28], it can be shown that the FGOA in this sense do not provide any gain compared to the algorithms developed within the estimation theory for solving the linear filtering and smoothing problems. This means that the proposed recursive iterative batch linearized smoother synthesized within the Bayesian approach fully takes into account the sparseness of the matrices.

Thus, it follows from the above that for the class of problems under consideration, the FGOA do not have any advantages either in terms of accuracy or in terms of computational load compared to algorithms synthesized within the Bayesian approach. Nevertheless, the software implementing the FGOA (such as, for example, the open source library GTSAM – Georgia Tech Smoothing and Mapping [29]) proves to be useful in developing onboard algorithms. In addition, this software allows the developers who do not have any significant experience in creating the estimation algorithms within the Bayesian approach to successfully synthesize such algorithms for their applied problems related to integrated information processing [10, 11]. Often, when using the FGOA, usual formulation of the problem in the form of (1), (2) is not given. For example, if there are m sets of measurements of type (9) given in the form

$$Y_k^\mu = H^\mu(X_k) + V_k^\mu, \; \mu = \overline{1, m},$$

then, assuming that the measurement errors are independent of each other and factoring the posterior PDF, one can write the following representation:

$$p(X_k / Y_k^1, .. Y_k^m) = c p(X_k) \prod_{\mu=1}^m p(Y_k^\mu / X_k). \tag{29}$$

It is easy to see that, additionally assuming the errors in each of the measurement sets are zero-mean and Gaussian with given covariance matrices, the criterion of the type (9) can be represented as a sum

$$J(X_k) = J^0(X_k) + \sum_{\mu=1}^m J^\mu(X_k), \tag{30}$$

where the summands $J^0(X_k) = N(X_k; \overline{X}_k, \mathbf{P}_x)$ $J^\mu(X_k) = N(Y_k; \mathbf{H}^\mu(X_k), \mathbf{R}_\mu))$ are called the a priori (initial) uncertainty factor and the factors of the corresponding measurements coming, for example, from a

GNSS, an inertial measurement unit, etc. [3, 11, 13]. It is the problem statement aimed at minimizing the criterion (30) that the consideration of the problem often begins with in works involving the FGOA application.

It is also worth mentioning that the a posteriori PDF can be factorized in different ways. In particular, the representation (11) based on the Markov type of the estimated sequence is also a factorization. The fact of using one or another type of factorization, especially when solving nonlinear problems, seems to be very important from the point of view of building efficient algorithms with low computational cost for its solution. For instance, the widely used method of partial analytical integration [30, 31], in fact, is based on factorization in which a co-multiplier (factor) generating a nonlinear block in the algorithm, and a co-multiplier generating a bank of KF are extracted. The specific feature of a problem where a time-invariant subvector can be isolated in the estimated vector, like, for example, in SLAM problems, also seems to be important. Taking this into account is essential when constructing computationally cost-effective SLAM algorithms. However, it should be noted that the issues related to this topic are worth studying separately, beyond the scope of this paper.

VI. CONCLUSIONS

We have investigated the effectiveness of linearization-based algorithms for solving nonlinear estimation problems formulated within the Bayesian approach and based on recursive and nonrecursive iterative schemes for measurements processing.

It has been shown that in contrast to traditional recursive processing schemes using the EKF and a filter with local iterations, the non–recursive schemes provide accuracy close to potential when solving the problems where the a posteriori PDF of the estimated state vector changes from multi-extreme to single-extreme as the measurements accumulate.

A recursive iterative batch linearized smoother has been proposed, in which, at each iteration, all measurements accumulated by the current time point are processed first in a linearized KF, and then the obtained estimates are processed in reverse time. At the same time, at each iteration, the linearization points are adjusted simultaneously for all measurements, and this ensures that the same estimates are obtained as when using a nonrecursive scheme. In other words, compared to the filter with local iterations, this algorithm, as well as the nonrecursive one, provides higher accuracy, close to potential, and proves to be consistent. Moreover, the proposed algorithm does not require the conversion of high-dimensional matrices $nk \times nk$; it is just enough to only convert the matrices with dimension n, which significantly reduces the amount of calculations compared to the nonrecursive algorithm.

The interrelation and differences between the FGOA and algorithms based on linearization and synthesized within the Bayesian approach have been discussed. It is noted that when solving linear estimation problems, the use of FGOA does not provide any advantages compared to traditional KF algorithms.

VII. REFERENCES

[1] Q. Zeng, W. Chen, J. Liu, and H. Wang, "An improved multi-sensor fusion navigation algorithm based on the factor graph," Sensors, vol. 17, p. 641, 2017.

[2] E. Olson, "AXLE: Computationally-efficient trajectory smoothing using factor graph chains," 2021 IEEE International Conference on Robotics and Automation (ICRA), Xi'an, China, 2021, pp. 7443-7448, doi: 10.1109/ICRA48506.2021.9561823.

[3] W. Wen, T. Pfeifer, X. Bai, and L.-T. Hsu, "Factor graph optimization for GNSS/INS integration: A comparison with the extended Kalman filter," Navigation, vol. 68, no. 2, pp. 315–331, 2021.

[4] A.Yu. Filatov, "Scalable algorithms for simultaneous mapping and localization of mobile robot swarms," Synopsis of PhD Thesis in Engineering Science, St. Petersburg, 2021.

[5] X. Wu, B. Xiao, C. Wu, Y. Guo, and L. Li, "Factor graph based navigation and positioning for control system design: A review," Chinese Journal of Aeronautics, vol. 35, no. 5, pp. 25–39, 2022.

[6] Y. Zhai and S. Zhang, "A novel LiDAR–IMU–odometer coupling framework for two-wheeled inverted pendulum (TWIP) robot localization and mapping with nonholonomic constraint factors," Sensors, vol. 22, 4778, 2022, https://doi.org/10.3390/s22134778.

[7] J. Dai, S. Liu, X. Hao, Z. Ren, and X. Yang, "UAV localization algorithm based on factor graph optimization in complex scenes," Sensors, vol. 22, 5862, 2022.

[8] M. Chen, Z. Xiong, J. Xiong, and R. Wang. "A hybrid cooperative navigation method for UAV swarm based on factor graph and Kalman filter," International Journal of Distributed Sensor Networks, vol. 18(1), 2022, doi:10.1177/15501477211064758.

[9] S. Xin, X. Wang, J. Zhang, K. Zhou, and Y. Chen, "A comparative study of factor graph optimization-based and extended Kalman filter-based PPP-B2b/INS integrated navigation," Remote Sens, vol. 15, 5144, 2023, https://doi.org/10.3390/rs15215144.

[10] A.A. Penkovsky, J. Mahmoud, M. Mohrat, and S.A. Kolyubin, "Robust visual-inertial odometry for ground robots in dynamic environments," 30th St. Petersburg International Conference on Integrated Navigation Systems, pp. 49–54, 2023.

[11] D.A. Cherginets and A.A. Vedyakov, "Design of visual-inertial odometry algorithm for a four-legged walking robot with a stereo camera," 31st Saint Petersburg International conference on Integrated Navigation Systems, 2024.

[12] A.R. Abdrashitov, "Relative navigation of a group of unmanned aerial vehicles using Bayesian estimates," Trudy XIX Vserossiiskoi nauchno-prakticheskoi konferentsii "Perspektivnye sistemy i zadachi upravleniya" (Proceedings of XIX Russian National Scientific and Practrical Conference "Advanced Systems and Control Problems"), 2024.

[13] S.H. Quan, S.H. Chen, Y.L. Zhou S. Zhao, H.Z. Hu, and Q. Zhu, "A robust IMU/GNSS/ODO integrated navigation system based on factor graph," 31st Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), 2024.

[14] V.G. Karaulov, A.M. Gruzlikov, and Yu.A. Litvinenko, "Solution of the problem of AUV positioning relative a stationary docking station using a factor graph optimization algorithm," 31st Saint Petersburg International conference on Integrated Navigation Systems, 2024.

[15] O.A. Stepanov, V.P. Zolotarevich, A.V. Motorin, and M.S. Ivanov, "Comparing recursive and nonrecursive estimators in target tracking using bearing-only measurements," 31st Saint Petersburg International conference on Integrated Navigation Systems, 2024.

[16] S. Särkkä, Bayesian Filtering and Smoothing, Cambridge University Press, 2013.

[17] O.A. Stepanov, Osnovy teorii otsenivaniya s prilozheniyami k zadacham obrabotki navigatsionnoi informatsii (Fundamentals of the Estimation Theory with Applications to the Problems of Navigation Information Processing), Part 1, Vvedenie v teoriyu otsenivaniya (Introduction to the Estimation Theory), St. Petersburg: Concern CSRI Elektropribor, 2017.

[18] Y. Bar-Shalom, X. Li, and T. Kirubarajan, Estimation with Applications to Tracking and Navigation, New York: Wiley–Interscience, 2001.

[19] O.A. Stepanov, Osnovy teorii otsenivaniya s prilozheniyami k zadacham obrabotki navigatsionnoi informatsii (Fundamentals of

the Estimation Theory with Applications to the Problems of Navigation Information Processing), Part 2, Vvedenie v teoriyu fil'tratsii (Introduction to the Filtering Theory), St. Petersburg: Concern CSRI Elektropribor, 2017.

[20] O.A. Stepanov, Yu.A. Litvinenko, and A.M. Isaev, "Comparison of recursive and nonrecursive schemes for linearization-based algorithms in solving one class of nonlinear estimation problems,", XIV Vserossiiskoe soveshchanie po problemam upravleniya VSPU-24 (Proc. XIV All-russian Meeting on Control Problems VSPU-24), Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, 2024

[21] O.A. Stepanov, Primenenie teorii nelineinoi fil'tratsii v zadachakh obrabotki navigatsionnoi informatsii (Nonlinear Filtering Theory as Applied to Navigation Data Processing), St. Petersburg: CSRI Elektropribor, 2003.

[22] V.A. Vasil'ev, Yu.A. Litvinenko, O.A. Stepanov, and A.B. Toropov, "Comparison of recursive and nonrecursive filtering algorithms for map-aided navigation, XIV Vserossiiskaya mul'tikonferentsiya po problemam upravleniya (Proc. XIV All-Russian Multiconference on Control Problems), Rostov-on-Don, pp. 125–127, 2021.

[23] O.A. Stepanov, Yu.A. Litvinenko, V.A. Vasil'ev, A.B. Toropov, and M.V. Basin, "Polynomial filtering algorithm applied to navigation data processing under quadratic nonlinearities in system and measurement equations. Part 1. Description and comparison with Kalman type algorithms," Gyroscopy and Navigation, vol. 12, no. 3, pp. 205–223, 2021,

https://doi.org/10.1134/S2075108721030068

[24] B.P. Gibbs, Advanced Kalman Filtering, Least-Squares and Modeling: A Practical Handbook, John Wiley&Sons, Inc., 2011.

[25] J.S. Meditch, Stochastic Optimal Linear Estimation and Control, New York: McGraw Hill, 1969

[26] O.A. Stepanov and A.M. Isaev, A procedure of comparative analysis of recursive nonlinear filtering algorithms in navigation data processing based on predictive simulation, Gyroscopy and Navigation, vol. 14, no. 3, pp. 213–224, 2023, doi: 10.1134/S2075108723030094

[27] A. George and J.W.-H. Liu, Computer Solution of Large Sparse Positive Definite Systems, Englewood Cliffs, NJ: Prentice-Hall Inc., 1981.

[28] A. Aravkin, B. Bell, J. Burke, and G. Pillonetto, "Kalman smoothing and block tridiagonal systems: new connections and numerical stability results," arXiv, 2013, https://doi.org/10.48550/arXiv.1303.5237

[29] F. Dellaert and GTSAM Contributors, "Georgia Tech Smoothing and Mapping Llibrary", https://github.com/borglab/gtsam, 2022.

[30] A. Doucet, N. Freitas, and N. Gordon, Sequential Monte Carlo Methods in Practice, Springer, 2001.

[31] O.A. Stepanov and A.B. Toropov, "Application of sequential Monte Carlo methods using analytical integration procedures to navigation data processing," XII Vserossiiskoe soveshchanie po problemam upravleniya (12th All-Russian Conference on Control), Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, pp. 3324–3337, 2014.

.

Equations of elastic vibrations in the eigenform basis. Three-axis wave solid-state gyroscope*

S.E. Perelyaev,
Ishlinsky Institute for Problems in Mechanics
of the Russian Academy of Sciences (IPMech RAS)
119526, Moscow, Russia,
Prospekt Vernadskogo, 101-1
E-mail: ipm@ipmnet.ru

V. Ph. Zhuravlev,
Ishlinsky Institute for Problems in Mechanics
of the Russian Academy of Sciences (IPMech RAS)
119526, Moscow, Russia,
Prospekt Vernadskogo, 101-1
E-mail: zhurav@ipmnet.ru

A.A. Skripkin
Yuri Gagarin State Technical University
410054, Saratov, Russia,
Politekhnicheskaja Str., 77
E-mail: al1skr@yandex.ru

Abstract - **In the considered well-known publications devoted to the fundamental theory of wave solid gyroscopes in relation to a thin ring, cylinder, and hemisphere, the effect of inertness of elastic Bryan waves has a one-dimensional character: the angular velocity of a body is a scalar characterizing the rotation of an elastic solid body around an axis fixed in space. The generalization of this effect to the spatial case is considered and investigated: an elastic spherically symmetric solid body with a free boundary, on which mass forces act. The flat rotation of the spherical resonator is replaced by a spatial rotation. In such an experiment, the generalized Bryan inertia effect also turns out to be spatial.**
Keywords: Foucault's pendulum, hemispherical wave solid state gyroscope, 3-D resonant gyroscope, 3-D spherical resonant.

I. INTRODUCTION

Recently, a whole class of new gyroscopic devices has appeared, in which the idea of the well-known classical Foucault pendulum is actually realized [1]. This class includes the string gyroscope, ring wave gyroscope [2], hemispherical quartz resonator (HRG) [3...6] or wave solid-state gyroscope (WSG) [7,8], and "quapason" [9], In well-known publications [12,13,14,15], the theory of the HRG with a hemispherical quartz resonator is presented in detail.

Playing the role of a Foucault pendulum, the oscillator with two degrees of freedom is realized as one of the working forms of natural oscillations of an elastic medium possessing axial symmetry. In this case, unlike the classical Foucault pendulum, the rotation of the elastic medium around the axis of symmetry involves the realized form of natural oscillations in rotation relative to inertial space (the exception is the string gyroscope), however, the ratio of the angular velocity of the form relative to the elastic body to the angular velocity of the solid body relative to space is a constant, depending only on the number of the form (mode) of oscillations and almost independent of the number of the form (mode) of oscillations of oscillations and almost independent of the basic structural properties of the material. This coefficient is called the scale factor, or Bryan's coefficient [10], named after the first researcher of oscillations of an elastic body, taking into account its continuous rotation in absolute space with a constant angular velocity. In the eigenspace corresponding to the chosen form of oscillations, all questions of the theory of such an inertial information sensor can be considered in the framework of the same equations analogous to those of the classical Foucault pendulum. For this reason, this whole class of devices is named [15] as generalized Foucault pendulum. This paper addresses the question of how to efficiently observe such a pendulum to make it an accurate integrating gyroscope. It should be noted that both in the case of an elastic ring and elastic hemispherical shell, the physical effect of inertness of mechanical standing waves has a one-dimensional character: the angular velocity $\omega(t)$ is a scalar characterizing the rotation of an elastic solid body around an axis fixed in space.

We propose to consider the generalization of this effect to an arbitrary spatial case, for example, excitation of a standing wave of oscillations with some working pure tone in a stationary elastic solid body (quartz thin-walled, full and elastic spherical resonator). After that, it is necessary to put the three-dimensional solid body into rotation with an arbitrary angular velocity $\omega(t)$ and determine the angular velocity of the spatial standing wave excited on the surface of the spherical resonator (e.g., an ellipsoid of rotation). Such an ellipsoid of rotation at the spatial rotation of the spherical resonator will rotate relative to the body, lagging behind its motion strictly according to a certain law, at which the generalized spatial effect of the Bryan inertness of elastic standing waves on the surface of a full high-gain elastic spherical shell is observed.

II. EFFECT OF INERTNESS OF ELASTIC SHAPES ON A ROTATING HEMISPHERICAL SHELL

The Bryan effect of inertness of elastic waves in a rotating axisymmetric solid body was discovered thirty years earlier than the Sanyak effect, but research into the possibility of creating a gyroscope based on this principle began twenty years later [3-5] than the realization of the first sample of a ring laser gyroscope (RLG). Since then, not only accurate RLG, but also wave solid-state gyroscopes (WSGs/HRG) of medium and high accuracy based on a high-voltage quartz hemispherical resonator and low accuracy using cylindrical metal resonators have been created. High-precision WTGs are developed and manufactured by US and French companies. The first high-precision WTG was created by Northrop Grumman (USA) for aviation applications [2,3]. The combination of high accuracy and unique operational characteristics of HRG-130P for space applications is an undeniable advantage of WTG, but the high

cost of the gyroscope still limits its mass application. Probably that is the reason why this type of gyroscope is not used in gyro-inclinometers for drilling operations. An important achievement of the leading developers of gyroscopes was to provide a method for autocalibration of zero drift and scale factor during gyroscope operation. A sensation of the recent years has become the advent of a high-precision WTG from Safran Electronics & Defence (France).

The accuracy parameters of the serial device Ragys-20, which has found application in the inertial navigation space system s Space Naute, are still inferior to the parameters of the precision device WTG-130P, but the designers have already announced significantly higher accuracy parameters [16]. According to the principle of operation and application in various gyroscopic systems, wave solid-state gyroscopes (WSTGs) are divided into WSTGs operating in the AVS mode (angular velocity sensor - WSTG-AVS), and WSTGs operating in the IG mode - integrating gyroscope (WSTG-IG).

Integrating WTGs have relatively high accuracy (the random drift rate of 0.001 deg/h and an accurate Bryan scaling factor). The Bryan effect and the well-known Lynch-Scott experiment are described in more detail in the monograph [15].

The functioning of the integrating wave solid-state gyroscope (WSTG) is based on the use of the unique inertia properties of mechanical standing waves excited in a moving axisymmetric high-quality shell. Currently, the main manufacturers of WTGs are expanding the field of application of this gyroscope. Modern inertial navigation systems (INS) based on WTGs of different accuracy grades have already been developed for aerospace, land and marine applications.

The feasibility of zero drift at the level of 10^{-4} deg/h has been reported [17]. Such an SINS certainly wins in comparison with the FOF-based SINS by the known criteria of cost, dimensions and accuracy, but the question of the cost remains.

III. EQUATIONS OF ELASTIC VIBRATIONS OF A SPHERICAL SOLID BODY

Let us proceed to the elastic vibrations of a thin spherical shell. Let us take an elastic spherically symmetric solid body $x = (x_1 x_2 x_3)$ with a free boundary on which the mass forces of density f. The principal vector of all forces acting on the elastic volume body $\int_V f\,dm$, without restriction of generality, will be assumed to be equal to zero. Under the action of the principal moment $\int_V r \times f\,dm$, the body under consideration changes its orientation in space (r - radius-vector of an arbitrary point of the body, dm - mass element, V - area occupied by the body). To describe the elastic deformations of a thin spherical shell, we introduce a coordinate system $x_1 x_2 x_3$ related to the elastic body under consideration so that the following conditions are fulfilled:

$$\int_V x\,dm = 0, \quad \int_V r \times x\,dm = 0 \qquad (3.1)$$

where $x = (x_1 x_2 x_3)$ is the elastic displacement of the material point, which in the undeformed state occupied the position r. The conditions (4.1) characterise the coordinate triangle with respect to which the elastic body does not move or rotate on average (for all particles). The following problem is posed: knowing the absolute angular velocity $\omega(t)$ of the trihedron $x_1 x_2 x_3$ in projections on its axes, determine how the waves of

elastic deformations behave. Let us write the Dalembert-Lagrange principle for the body under consideration following [11]:

$$\int_V \left[\ddot{x} + \omega \times (\omega \times (r + x)) + \dot{\omega} \times (r + x) + 2\omega \times \dot{x} + \frac{1}{\rho}\nabla\Pi - f \right] \delta x \times dm = 0 \qquad (3.2)$$

Here ρ is the density depending only on the modulus of the radius vector $|r|$, $\nabla\Pi$ – the gradient of the quadratic functional of the linear theory of elasticity. The coordinates determining the angular position of the body as a whole do not vary; it is assumed that the angular velocity $\omega(t)$ is a known function of time. To choose the generalized coordinates, we consider the special case of $\omega = 0$. In the monograph [16], it is shown that the spectrum of natural vibrations of a free solid body under the conditions (3.1) introduced above is discrete. This means that the increasing sequence of frequencies of natural vibrations is unbounded $\nu_1 \leq \nu_2 \leq \nu_3 \dots$, and the eigenelements $h_1(r), h_2(r), \dots$ corresponding to these sequences of frequencies form an orthonormalised system of functions complete in the configuration space of the problem under consideration:

$$\int_V h_n(r)\, h_1(r)\,dm = \delta_n^l \qquad (3.3)$$

This allows us to introduce independent Lagrangian coordinates describing all degrees of freedom in deforming the body, $\boldsymbol{\omega}(t) \neq 0$ in the general case as follows:

$$x = \sum_{n=1}^{\infty} q_n(t) h_n(r) \qquad (3.4)$$

The problem on the natural vibrations of a spherically symmetric free body admits the $SO(3)$ group, so the spectrum of natural frequencies is degenerate and consists of a sequence of at least threefold frequencies: the configuration space is a direct product of three-dimensional eigenspaces:

$$\nu_1 = \nu_2 = \nu_3 \leq \nu_4 = \nu_5 = \nu_6 \leq \dots$$

configuration space is a direct product of three-dimensional eigenspaces:

$$\{h_1, h_2, h_3\} \times \{h_4, h_5, h_6\} \times \dots$$

We fix the number m of an arbitrary eigenspace and introduce notations for the corresponding generalized coordinates:

$$q_{3m-1} = u, \ q_{3m-2} = v, \ q_{3m} = w \qquad (m = 1,2\dots).$$

Substituting formula (3.4) as well as the expression $\delta x = \sum_{n=1}^{\infty} \delta q_n h_n(r)$ in formula (3.2) and equating the coefficients δq_n at independent variations to zero, we obtain an infinite system of second-order ordinary differential equations with respect to the following form:

$$\ddot{u} + au + bv + cw - v(\dot{\omega}, k_3) + w(\dot{\omega}, k_2) - -2\dot{v}(\omega, k_3) + 2\dot{w}(\omega, k_2) + F_1 + L_1 = 0$$
$$\ddot{v} + bu + bv + ew + u(\dot{\omega}, k_3) - w(\dot{\omega}, k_1) + +2\dot{u}(\omega, k_3) - 2\dot{w}(\omega, k_1) + F_2 + L_2 = 0 \qquad (3.5)$$
$$\ddot{w} + cu + ev + fw - u(\dot{\omega}, k_2) + v(\dot{\omega}, k_1) - -2\dot{u}(\omega, k_2) + 2\dot{v}(\omega, k_1) + F_3 + L_3 = 0$$

In the system of differential equations (3.5) the scalar coefficients have the following form:

$$a = \int_V (h_{3m-2}, \omega)^2 dm - \omega^2,$$
$$b = \int_V (h_{3m-2}, \omega)(h_{3m-1}, \omega) dm,$$
$$c = \int_V (h_{3m-2}, \omega)(h_{3m}, \omega) dm, d = \int_V (h_{3m-1}, \omega)^2 dm - \omega^2,$$
$$e = \int_V (h_{3m-1}, \omega)(h_{3m}, \omega) dm, \; f = \int_V (h_{3m}, \omega)^2 dm - \omega^2,$$
$$F_1 = \int_V (\nabla\Pi, h_{3m-2}) \, dV, \quad F_2 = \int_V (\nabla\Pi, h_{3m-1}) \, dV,$$
$$F_3 = \int_V (\nabla\Pi, h_{3m}) \, dV.$$

L_1, L_2, L_3 are linear functions of the generalized coordinates corresponding to other eigenspaces. The presence of these terms characterises the fact that systems of the type (3.5) for different subspaces are not independent of each other. In deriving equations (3.5), it was assumed for simplicity that the mass forces f are orthogonal to all eigenfunctions:

$$\int_V f h_n(r) dm = 0.$$

This means that in the forces f, there is only a constant component ($\int_V r \times f dm \neq 0$), which ensures the rotation of the body with velocity $\omega(t)$. In expression (3.5) the vector coefficients k_1, k_2, k_3 have the form:

$$k_1 = \int_V h_{3m-1} \times h_{3m} dm, \quad k_2 = \int_V h_{3m} \times h_{3m-2} dm,$$
$$k_3 = \int_V h_{3m-2} \times h_{3m-1} dm.$$

Due to spherical symmetry$(h_{3m-2}, h_{3m-1}, h_{3m})$, the choice of eigenvectors can be made in such a way that the following identities are fulfilled:

$$k_1 = \upsilon(1,0,0), k_2 = \upsilon(0,1,0), k_3 = \upsilon(0,0,1) \qquad (3.6)$$

here:$\upsilon = \pm|k_1| = \pm|k_2| = \pm|k_3| = \pm\left|\int_V h_{3m-1} \times h_{3m} dm\right|$

Using the Cauchy-Bunyakovsky inequality for (3.6), we obtain $0 \leq \upsilon \leq 1$. If we now introduce the notations:

$$z = \begin{pmatrix} u \\ v \\ w \end{pmatrix}, \; \omega = \begin{pmatrix} p \\ q \\ r \end{pmatrix}, \; L = \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix}, \; G = \begin{pmatrix} 0 & -r & q \\ r & 0 & -p \\ -q & p & 0 \end{pmatrix},$$

then equations (3.5) can be rewritten in the following compact vector form [11,15]:

$$\ddot{z} + Az + \upsilon\dot{G}z + 2\upsilon G\dot{z} + L = 0 \qquad (3.7)$$

where A - symmetric matrix of positional forces, consisting of elastic force coefficients F_1, F_2, F_3 and scalar coefficients a, b, c, d, e, f, respectively.

IV. SPATIAL PRECESSION OF A MECHANICAL STANDING WAVE

The vector differential equation (3.7) defines the spatial evolution of the m-th form (mode) of vibrations of a free spherical elastic rigid body caused by the presence of rotation. This evolution is determined by two circumstances. Firstly, the oscillation form itself responds directly to the rotation of the body, which is determined by the presence of the terms G and $\dot{G}$ in (3.7). Secondly, the form of oscillations under consideration is affected by other forms. We note at once that this influence is insignificant, since, for example, when solving equations (3.7) by the averaging method, all terms defined by linear functions of the generalized coordinates L vanish in the first approximation. In addition, the following fact takes place. Note that there exists such a coordinate system $z \rightarrow y$: $z = My$, where M is a time-dependent orthogonal coordinate transformation matrix in which equation (3.7) has a self-adjoint form

when the identity $L=0$ is fulfilled. In this coordinate system, equation (3.7) admits a standing wave type solution. Let us show that. Substituting the equation $z=My$, into the above obtained expression (3.7), we find:

$$\ddot{y} + 2M^T(\dot{M} + \upsilon GM)\dot{y} + M^T(\ddot{M} + 2\upsilon G\dot{M} + \upsilon\dot{G}M + AM)y = 0 \qquad (4.1)$$

We require the following first-order differential equation, which is derived from equation (4.1):

$$\dot{M} = -\upsilon GM \qquad (4.2)$$

and, substituting it into (4.1), we find the following second-order equation:

$$\ddot{y} + M^T(A - \upsilon^2 G^2)My = 0 \qquad (4.3)$$

Thus, if a mechanical standing wave of elastic vibrations with some pure tone is excited in a stationary hollow spherical body and then the body is driven into rotation with an arbitrarily time-varying angular velocity $\omega(t)$, the standing wave will rotate relative to the body itself according to the law of the form (4.2). It should be noted that the first-order differential equation (4.2) is the well-known kinematic Poisson equation. Let us compare equation (4.2) with the known classical kinematic Poisson equation for the solid $\dot{N} = -GN$, in which the orthogonal matrix N determines the orientation of the solid in absolute (inertial) space, whence it obviously follows that the angular velocity of a mechanical standing wave relative to a body is proportional to its angular velocity relative to space:

$$\Omega_0(t) = -\upsilon\omega(t), \qquad (4.4)$$

or for the angular velocity of a mechanical standing wave relative to absolute space, we have the form:

$$\Omega(t) = (1 - \upsilon)\omega(t) \qquad (4.5)$$

The relation (4.4) represents a generalization of the known scalar Lynch-Zhuravlev expression [3,4,11] to the spatial case. Let us emphasize that for a full spherical elastic shell, the Bryan coefficient has a spatial character.

It determines the proportionality of the angular velocity vectors to each other, in contrast to the case of a thin ring, when the angular velocities of plane rotation were considered.

A spatial mechanical standing wave excited on the surface of an elastic high-volume spherical solid body in the form of a three-dimensional ellipsoid of rotation precesses relative to the body with the angular velocity vector proportional to the angular velocity vector of the body itself relative to absolute and/or stationary (inertial) space.

V. CASE OF A COMPLETE SPHERICAL SHELL

The calculation of the Bryan coefficient υ of a closed spherical shell is performed under specific assumptions about the density distribution along the radius $\rho(r)$. Let us consider the case of a thin spherical shell.

In projections on the axes u, v, w directed along the meridian, parallel and normal to the surface of the sphere, the eigenvectors U_1, U_2 have the following form [15]:

$$\boldsymbol{U}_1 = \begin{pmatrix} A_n \left(\frac{dP_n^k}{d\theta} \right) \cos k\,\varphi \\ -A_n k \sin^{-1}\theta\, P_n^k \sin k\,\varphi \\ P_n^k \cos k\,\varphi \end{pmatrix}; \ \boldsymbol{U}_2 = \begin{pmatrix} A_n (dP_n^k / d\theta) \sin k\,\varphi \\ A_n \sin^{-1} k\,\varphi \\ 0 \end{pmatrix}$$

(5.1)

where P_n^k are the adjoint Legendre functions, φ, θ are the spherical coordinates, A_n is a coefficient depending on the number of waves along the meridian and on the oscillation frequency. For the coefficients A_n and oscillation frequency ν in the Lamb approximation we have the following relations (μ is the Poisson's coefficient):

$$A_n = [-2(1+\mu) + ((1-\mu)\nu^2)/2][(1+\mu)n(n+1)]^{-1}$$

Lamb's frequency equation is of the form:

$$(1-\mu)\nu^4 - 2\nu^2(1 + 3\mu + n(n+1)) +$$
$$+4(1+\mu)(n^2 + n - 2) = 0 \qquad (5.2)$$

The vectors (5.1) are two eigenvectors out of three, defining a three-dimensional eigenspace (with n and k fixed). The third one can be obtained from the axes written out by the transformation, but it is no longer necessary to calculate the coefficient k. The transition from the variables u, ν, w to the variables x_1, x_2, x_3 adopted in this paper is expressed by the well-known coordinate transformation formula:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \cos\theta\cos\varphi & \cos\theta\sin\varphi & \sin\theta \\ -\sin\varphi & \cos\varphi & 0 \\ -\sin\theta\cos\varphi & -\sin\theta\cos\varphi & \cos\theta \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix}$$

From where we have an analytical expression of the first vector:

$$\boldsymbol{h}_1 = \frac{1}{\|\boldsymbol{U}_1\|} \cdot$$

$$\cdot \begin{pmatrix} A_n \dfrac{dP_n^k}{d\theta} \cos k\,\varphi \cos\theta \cos\varphi - \dfrac{A_n k}{\sin\theta} P_n^k \cos k\,\varphi \cos\theta \sin\varphi + \\ \quad + P_n^k \cos k\,\varphi \sin\theta \\ -A_n \dfrac{dP_n^k}{d\theta} \cos k\,\varphi \sin\varphi - A_n \dfrac{k}{\sin\theta} P_n^k \sin k\,\varphi \cos\varphi \\ -A_n \dfrac{dP_n^k}{d\theta} \cos k\,\varphi \sin\theta \cos\varphi + \dfrac{A_n k}{\sin\theta} P_n^k \sin k\,\varphi \sin\theta \sin\varphi + \\ \quad + P_n^k \cos k\,\varphi \cos\theta \end{pmatrix}$$

A similar analytical expression can be obtained for the second eigenvector $\boldsymbol{h}_2$. Then we calculate the Bryan coefficient υ of a thin closed (full) spherical shell to use it as an inertial sensor of three quasi-angle rotations of the gyroscope base relative to inertial space:

$$\upsilon = \left| \int_0^{2\pi} \int_0^{\pi} \boldsymbol{h}_1 \times \boldsymbol{h}_2 \sin\theta \ d\theta\, d\varphi \right| =$$

$$= \frac{2\pi k A_n}{\|\boldsymbol{U}_1\|^2} \left| \left\{ \int_0^{\pi} \frac{dP_n^k}{d\theta} P_n^k \sin\theta \ d\theta, 0, \frac{dP_n^k}{d\theta} P_n^k \cos\theta \ d\theta \right\} \right|$$

The integral defining the first component of the written vector is zero. For the second integral, we obtain an analytical expression:

$$\int_0^{\pi} \frac{dP_n^k}{d\theta} P_n^k \cos\theta \ d\theta = -\frac{(n+k)!}{(2n+1)(n-k)!}$$

Taking into account that the norm $\boldsymbol{U}_1$ has the form:

$$\|\boldsymbol{U}_1\|^2 = \frac{2\pi}{(2n+1)} [1 + n(n+1)A_n^2] \frac{(n+k)!}{(n-k)!} \quad (k \neq 0)$$

finally, we find an analytical expression for the Bryan coefficient:

$$\upsilon = k A_n^2 [1 + n(n+1)A_n^2]^{-1}$$

Let us consider as an example the case when $k = n = 2$, $\mu = 0$. According to the above formula (5.2), we have:

$$\nu^2 = 7 + \sqrt{33}, \ A_n = (7 + \sqrt{33})/12 - 1/3.$$

For a thin full spherical shell, the Bryan coefficient [15], equal to approximately one quarter ($\upsilon = 0.25$), is smaller than that obtained for a ring (or cylinder) equal to 2/5. Let us emphasise that for a full spherical shell the Bryan coefficient has a spatial character. It determines the proportionality to each other of the angular velocity vectors of the form (4.4), in contrast to the case of a ring, when the angular velocities of plane rotation were considered. A mechanical standing wave in the form of a three-dimensional ellipsoid precesses relative to the body itself with the angular velocity vector proportional to the angular velocity vector of the body relative to absolute (inertial) space.

VI. THREE-AXIS WAVE SOLID-STATE GYROSCOPE

In publications [11,13,15] it was shown that a full elastic spherical shell can serve as a sensor of inertial information. The use of wave phenomena in spatial elastic media to determine the angular motion of a solid body, we will show on the specific example of a three-degree WSTG.

Indeed, let a mobile object rotate with angular velocity $\omega_0(t)$ relative to inertial space. This is an angular velocity in projections on the axes associated with the moving object and it is determined by the inertial means located on this object. Let the object have an elastic spherical shell rotating with angular velocity $\omega_1(t)$ relative to the object (velocity ω_1 is given in the axes of the object). In addition to the observed (known) velocity $\omega_1(t)$, there is the precession velocity of the standing wave in projections on its same axes - $\omega_2(t)$. We denote the angular velocity of such a wave relative to the shell by $\omega_2^*(t)$. Based on the results obtained in the previous paragraph, we find the relationship between the angular velocities $\omega_0(t)$, $\omega_1(t)$, $\omega_2(t)$.

Recall that the transformation of coordinates from the axes associated with the object to the axes associated with the spherical shell is given by the orthogonal matrix $N(t)$, which defines the position of a solid body in inertial space.

The transformation of coordinates from the axes associated with the spherical shell to the axes associated with the elastic wave is given by the matrix $M(t)$. It follows from the above formula (4.4) that the precession velocity of a spatial

mechanical standing wave relative to a spherical shell in the axes of the shell itself is proportional to the speed of the shell plus the speed of the shell relative to the object itself in the axes of the shell:

$$MN\omega_2^* = -\chi N(\omega_1 + \omega_2)$$

Whence, in the axes related to the moving object, we have the following expression:

$$\omega_2^* = -\chi N^T M^T N(\omega_1 + \omega_2)$$

The angular velocity of the standing wave relative to the moving object is:

$$\omega_2 = \omega_1 + \omega_2^* = \omega_1 - \chi N^T M^T N(\omega_1 + \omega_0)$$

From where the angular velocity of the moving object is found:

$$\omega_0 = \frac{1}{\chi} N^T M N(\omega_1 - \omega_2) - \omega_1 \qquad (6.1)$$

To calculate the angular velocity of the object ω_0 by formula (6.1), in addition to the assumed known velocities $\omega_1(t)$ and $\omega_2(t)$, it is required to know the matrices $N(t)$ and $M(t)$, which are related to the angular velocities by the following known kinematic Poisson equations:

$$\dot{M} = -M[N(\omega_2 - \omega_1)]^\nu, \quad \dot{N} = -N\omega_1^\nu \qquad (6.2)$$

where the symbol ω_1^ν conventionally denotes the matrix representation of the angular velocity of rotation.

The sequence of solving kinematic differential equations of the form (6.1) and (6.2) is as follows:

a) first the following equation is solved:

$$\dot{N} = -N\omega_1^\nu$$

b) the required orthogonal matrix $N(t)$ is found.

Subsequent substitution of this solution into equation

$$\dot{M} = -M[N(\omega_2 - \omega_1)]^\nu,$$

allows us to find the matrix $M(t)$. After that, substitution into expression (6.1) allows us to find the desired velocity ω_0.

If the system outputs not rotation velocities $\omega_1(t)$ and $\omega_2(t)$, but matrices $N(t)$ and $S(t) = M(t)N(t)$, then at first the angular velocity matrix $\omega_1^\nu = -N^T \dot{N}$ is found and after substitution into (6.2) the required angular velocity ω_0 is found.

The problem of determining the angular velocity ω_0 is most simply solved by providing the angular zero velocity for a spherical shell ($\omega_1 \equiv 0$).

Then the orthogonal matrix $N = E$ is a unit matrix. Hence, from expression (6.1) we obtain:

$$\omega_0 = -1/\chi \, (M\omega_2)$$

The condition $\omega_1 \equiv 0$ can be ensured by fixing the sphere at the points at which it touches the vertices of the tetrahedron inscribed in it. Then, if we consider the form of elastic oscillations in which the sphere turns into a triaxial ellipsoid, then the distances between these points change by values of the second order of smallness with respect to the amplitude of elastic

oscillations and in the linear formulation the fixing in them does not constrain the oscillations of the shell. All of the above was considered under the assumption that the shell has spherical symmetry and there are no extraneous perturbations. If the deviations from these assumptions are small, the new effects (e.g., collapse of the standing wave, and/or reduction of its amplitude) can be compensated by introducing an appropriate control. Physically, such control is achieved by applying electrical forces to the shell surface through the introduction of specially arranged electrodes. The calculation of these forces depending on the shape and number of electrodes and on the potential distribution on them in detail is given in the monograph [13]. In this publication, we will limit ourselves to the remark that after the calculation of the necessary forces and projection onto the chosen eigenform (in the sense of calculating the integral $\int_v f \, h_n(r) dm$), the right-hand sides of the system of differential equations of elastic oscillations of the sphere shell [13,16] will contain terms that can be considered to depend on the phase state of the considered eigenform, since knowing this state and using the feedback principle, we can control the values of potentials on the corresponding electrodes at our discretion. Thus, we reduce the above problem to the elucidation of those evolutions of the form of elastic oscillations, which arise under perturbation of the ideal circuit considered above.

Informationally, 3D-WSTG is a three-dimensional inertial sensor whose operation is based on the following basic principles:

- the resonator has perfect axisymmetry to ensure high balancing performance;

- the elastic standing wave is controlled by electrostatic forces generated by electrodes that are in close proximity to the surface of the spherical resonator;

- the resonator is made in the form of a full (closed) sphere, with such a design each gram of resonator material is most effectively involved in the conservation of energy of bending oscillations;

- high-frequency resonator made of amorphous fused quartz provides optimal isotropy (no crystallographic axis) and minimal energy dissipation (low internal friction).

As a result of the performed optimization, the calculated characteristics of the triaxial WSTG are obtained, which, most surprisingly, do not depend on the resonator dimensions. They are related to the quality of the resonator, whose indicator is the goodness of fit (or damping time), as well as to the quality of the wave pattern control, i.e. to the characteristics of the control electronics.

The main disadvantage of the triaxial WSTG is its high cost, mainly due to the extremely high precision of manufacturing the spherical resonator and assembling the complex parts and assemblies that must precisely match the geometry of the resonator's spherical shell to form a uniform gap between these two precision parts of the gyroscope.

VII. PHASE DIVERSITY OF ELASTIC MECHANICAL STANDING WAVES

All further analyses are performed in the framework of the first approximation of the averaging method; therefore, we will consider the following assumptions regarding the three second-order differential equations of the elastic vibrations of the sphere shell obtained in [11,15]. The angular velocity of the body relative to space $\omega(t)$ is considered small in comparison with the frequency of natural elastic vibrations of the sphere shell. The coupling coefficients included in the terms $L(L_1, L_2, L_3)$, which are linear functions of the generalised coordinates corresponding to other eigenspaces, are considered small at the same time.

These terms appear as a result of calculating integrals of the type:

$$\int_v h_k \times h_i \, dm = 0$$

where: h_k, h_i belong to different eigenspaces, and they are really small compared to the same integrals, but in which h_k and h_i are taken from the same eigenspace. In addition, let us add to the system of differential equations of elastic vibrations of the sphere the terms determined by various perturbations and control actions, which we will also consider small. The presence of these terms characterizes the fact that the differential systems of elastic vibrations of a complete spherical shell for different subspaces are not independent of each other. In deriving these equations, it was assumed that the mass forces f are orthogonal to all eigenfunctions:

$$\int_v f \, h_n(r) \, dm = 0$$

This condition means that there is only a constant component in the forces f itself:

$$\int_v f \times h_n(r) \, dm \neq 0,$$

providing rotation of the body itself with angular velocity $\omega(t)$. In the future this restriction will be removed. Considering the above, the infinite system of ordinary differential equations of the second order with respect to generalised coordinates obtained in the monographs [11,15] can be written in the form:

$$
\begin{aligned}
\ddot{u} + u &= \varepsilon U \\
\ddot{v} + v &= \varepsilon V \\
\ddot{w} + w &= \varepsilon W
\end{aligned}
\tag{7.1}
$$

The general solution of the system of differential equations (7.1) in the absence of appropriate control ($\varepsilon = 0$) has the following analytical form:

$$
\begin{aligned}
u &= x_1 cost + x_4 sint \\
v &= x_2 cost + x_5 sint \\
w &= x_3 cost + x_6 sint
\end{aligned}
\tag{7.2}
$$

The velocities of motion along the trajectories (7.2) – changes of initial coordinates u,v,w – are obtained by differentiating expressions (7.2), so that we have:

$$
\begin{aligned}
\dot{u} &= -x_1 sint + x_4 cost \\
\dot{v} &= -x_2 sint + x_5 cost \\
\dot{w} &= -x_3 sint + x_6 cost
\end{aligned}
\tag{7.3}
$$

Next, consider as a substitution of the variables $(u, v, w, \dot{u}, \dot{v}, \dot{w}) \rightarrow (x_1, x_2, x_3, x_4, x_5, x_6)$ in the system of equations (7. 1), leading it to the following general form (after averaging over time t) of equations:

$$\dot{x} = \varepsilon X(x) \tag{7.4}$$

where x is a new phase variable, and the system (7.4) itself defines the evolution in phase space of those oscillation forms observed in the system (7.1) under the condition of no control $\varepsilon = 0$.

When the condition $\varepsilon = 0$ is fulfilled, any trajectory in the system (7.1) corresponds to a fixed point in the system (7.4). The trajectories themselves are defined by parametric equations (7.2) and represent a plane curve, which is generally an ellipse. The equation of the plane in which this ellipse lies is easy to obtain if we note that to exclude $cost$ and $sint$ from the system (7.2) it is necessary to fulfil the conditions of linear dependence:

$$
det \left\| \begin{matrix} u & x_1 & x_4 \\ v & x_2 & x_5 \\ w & x_3 & x_6 \end{matrix} \right\| = 0
\tag{7.5}
$$

Based on (7.5) we obtain the equation we are looking for:

$$
\begin{aligned}
u(x_2 x_6 - x_3 x_5) + v(x_3 x_4 - x_1 x_6) + \\
+ w(x_1 x_5 - x_2 x_4) = 0
\end{aligned}
\tag{7.6}
$$

Standing waves on the surface of a spherical shell are characterized by the simultaneous passage of all its points during oscillations through the equilibrium position, which corresponds to the simultaneous nulling of functions of the form (7.2), and this, in turn, means that the elliptic trajectory in the parameter space (u, v, w) turns into a straight line segment. The plane (7.6) is undefined in this case, which leads to the following conditions on the phase vector:

$$
\begin{aligned}
k_1 = (x_3 x_4 - x_1 x_6) = 0, \quad k_2 = (x_1 x_5 - x_2 x_4) = 0, \\
k_3 = (x_2 x_6 - x_3 x_5) = 0
\end{aligned}
\tag{7.7}
$$

Thus, the set of elastic standing waves of an arbitrarily chosen form of oscillations of a spherical shell, which is in mutually unambiguous correspondence with the set of rectilinear forms of oscillations in the system (7.1) is represented in the phase space of the system (7.4) by a quadratic manifold (7.7), which is the intersection of three spatial conical cylinders.

VIII. EVOLUTIONARY BASIS ON THE VARIETY OF MECHANICAL STANDING WAVES

If in a system of the form (7.1) the initial conditions are chosen so that in the absence of control $\varepsilon = 0$ a rectilinear form of oscillations is realised (these conditions correspond to the position of the initial point in the system (7.4)) on the manifold (7.7), then, when the condition $\varepsilon \neq 0$ is fulfilled, this rectilinear form will undergo one of the following evolutions: destruction of the form, precession of the form, change of the oscillation frequency and change of the oscillation amplitude. To all these types of evolution of the rectilinear shape correspond certain directions in phase space, the finding of which we pay special attention in this section of the paper. The direction of the fastest decay of a standing wave is determined by the normal to the manifold (7.7):

$$e_1 = {dk_1}/{dx} = \{-x_6, 0, x_4, x_3, 0, -x_1\}$$
$$e_2 = {dk_2}/{dx} = \{x_5, -x_4, 0, -x_2, x_1, 0\} \qquad (8.1)$$
$$e_3 = {dk_3}/{dx} = \{0, x_6, -x_5, 0, -x_3, x_2\}$$

If the right part of the system (11) belongs to a linear envelope of three vectors of the form (15), then the evolution of the standing wave consists in its destruction. To determine the directions in the phase space of the system (7.4) along which the standing wave precession occurs, we introduce a spherical coordinate system in the space of initial variables (u, v, w):

$$u = r\cos\varphi \cos\psi$$
$$v = r\cos\varphi \sin\psi \qquad (8.2)$$
$$w = r\sin\varphi$$

Consider the variation in position corresponding to the variation in angle ψ :

$$\delta u = -r\cos\varphi \sin\psi\delta\psi = -v\delta\psi$$
$$\delta v = r\cos\varphi \cos\psi\delta\psi \qquad (8.3)$$
$$\delta w = 0$$

Then, on the other hand, the position variation can be calculated from formulas (7.2):

$$\delta u = \delta x_1 \cos t + \delta x_4 \sin t$$
$$\delta v = \delta x_2 \cos t + \delta x_5 \sin t \qquad (8.4)$$
$$\delta w = \delta x_3 \cos t + \delta x_6 \sin t$$

Equating these relations (8.4) to the analytical expressions (7.7), we obtain:

$$-v\delta\psi \; \delta u = \delta x_1 \cos t + \delta x_4 \sin t = -(x_2 \cos t + x_5 \sin t)\delta\psi$$
$$u\delta\psi = \delta x_2 \cos t + \delta x_5 \sin t = (x_1 \cos t + x_4 \sin t)\delta\psi$$
$$0 = \delta x_3 \cos t + \delta x_6 \sin t$$

Equating $\delta\psi \to 0$ and equating separately the coefficients at $sint$ and $cost$, in the obtained equations, we find:

$$\frac{dx_1}{d\psi} = -x_2; \frac{dx_2}{d\psi} = x_1; \frac{dx_3}{d\psi} = 0; \frac{dx_6}{d\psi} = 0; \frac{dx_5}{d\psi} = x_4; \frac{dx_4}{d\psi} = -x_5$$

In other words, the desired precession direction is given by the following vector:

$$e_4 = \{-x_2, x_1, 0, -x_5, x_4, 0\}$$

For the other two directions of precession (note that the ellipse, as a solid body, has three directions of elementary rotations in total), we introduce a spherical coordinate system using cyclic rearrangement of axes:

To find the direction e_5:

$$u = r\cos\varphi \sin\psi$$
$$v = r\sin\varphi$$
$$w = r\cos\varphi \cos\psi$$

Accordingly, in order to find the direction e_6:

$$u = r \sin\varphi$$
$$v = r \cos\varphi \cos\psi$$
$$w = r \cos\varphi \sin\psi$$

Performing the similar calculations both times, we find the required unit vectors:

$$e_5 = \{x_3, 0, -x_1, x_6, 0, -x_4\}$$
$$e_6 = \{0, -x_3, x_2, 0, -x_6, x_5\}$$

To plot the direction of variation of the variation frequency (8.4), we equate the same variations obtained by time variation in the general analytical expression of the system (7.2):

$$\delta u = -x_1 \sin t\delta t + x_4 \cos t\delta t$$
$$\delta v = -x_2 \sin t\delta t + x_5 \cos t\delta t$$
$$\delta w = -x_3 \sin t\delta t + x_6 \cos t\delta t$$

As a result, we obtain the following sought unit vector:

$$e_7 = \{x_4, x_5, x_6, -x_1, -x_2, -x_3\}$$

To find the direction of variation of the amplitude, we introduce the equation of a six-dimensional sphere of unit radius:

$$S = \frac{1}{2}(x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2 - 1) = 0$$

Then the direction of amplitude change is determined by the gradient S:

$$e_8 = \frac{dS}{dx} = \{x_1, x_2, x_3, x_4, x_5, x_6\}$$

Let us summaries all the results obtained above together. The directions in phase space defining the main evolutions of the rectilinear form of oscillations in configuration space are given by the following vectors:

$$e_1 = \{-x_6, 0, x_4, x_3, 0, -x_1\}$$
$$e_2 = \{x_5, -x_4, 0, -x_2, x_1, 0\} \qquad (8.5)$$
$$e_3 = \{0, x_6, -x_5, 0, -x_3, x_2\}$$

which define the breakdown of the rectilinear form of elastic oscillations:

$$e_4 = \{-x_2, x_1, 0, -x_5, x_4, 0\}$$
$$e_5 = \{x_3, 0, -x_1, x_6, 0, -x_4\} \qquad (8.6)$$
$$e_6 = \{0, -x_3, x_2, 0, -x_6, x_5\}$$

which determine the precession of the rectilinear form of elastic oscillations.

Variation of the oscillation frequency sets the unit vector:

$$e_7 = \{x_4, x_5, x_6, -x_1, -x_2, -x_3\}$$

The change in the amplitude of the oscillation defines the following unit vector: $e_8 = \{x_1, x_2, x_3, x_4, x_5, x_6\}$.

IX. DESIGN VARIANT OF THE SPATIAL INTEGRATING WAVE SOLID-STATE GYROSCOPE

The spatial integrating wave solid-state gyroscope belongs to a new type of three-dimensional gyroscopic sensors; it is designed to determine the orientation parameters of moving objects of various classes. For this purpose, a high-quality elastic hollow spherical shell is used in the gyroscope design as a sensitive element - resonator. The wave solid gyroscope (WSG) is based on a new physical phenomenon - the inertial properties of elastic standing waves in a solid axisymmetric elastic body, which manifests itself as a lag of the angle of rotation of the standing wave from the rotation of the base. The functioning of the three-dimensional wave solid-state gyroscope (3D-WSG) is based on a completely new physical phenomenon based on the inertia properties of spatial standing waves excited in a spherical axisymmetric elastic body. On this basis, the authors consider a hollow spherical resonator for which the planar rotation of the resonator is replaced by a spatial one. In this case, the generalized Bryan effect, which manifests itself as a lag of three rotation

angles of the spatial standing wave from the base rotation, also appears to be spatial. Therefore, the 3D-WSG design contains a resonator in the form of an axisymmetric thin-walled element capable of vibration, as well as standing wave excitation and control electrodes, information acquisition sensors, and an electronic control unit containing a device for calculating the main gyroscope parameters (angle, oscillation amplitude stabilization, quadrature minimization, and other pendulum parameters). Earlier publications by the authors [23,24] considered two variants of such a spatial WSG design. In one variant of the 3D-WSG design, the resonator is made in the form of a spatial vacuum module consisting of outer and inner hollow spheres of magnetic material with mutually opposite magnetization direction placed concentrically one inside the other /Russian Federation patent No. 2763688, 2021/.

The inner sphere houses the power supply, the electromagnetic centering windings and the control unit. Control electrodes and auxiliary electrode are placed on the outer side of the inner sphere; standing wave position measurement sensors are placed on the inner side of the outer sphere and electromagnetic centering windings are placed on the outer side of the outer sphere. The spatial standing wave is excited in a thin-walled sphere made of a material with low internal friction parameters, enclosing the outside of the inner sphere through non-magnetic bushings. However, this variant of the spatial WSG design has certain disadvantages, which include the low overload capacity of the triaxial gyroscope in the axial and radial directions, due to the great technological complexity and labor intensity of manufacturing the outer and inner hollow spheres made of the selected magnetic material with a mutually opposite direction of magnetization.

In the following variant of the 3D-WSG design, the resonator is made in the form of a spatial vacuum module consisting of outer and inner hollow spheres concentrically placed one inside the other, the outer and inner surfaces of which are respectively covered with monoelectret layers having a charge of the same sign and value. In this case, the control electrodes and auxiliary electrode are located on the monoelectret layer placed on the outer surface of the inner sphere, and the standing wave position measurement sensors are located on the monoelectret layer placed on the inner surface of the outer sphere [24] /Russian Federation patent No. 216847, 2022/.

The inner sphere houses the power supply, electromagnetic centering windings and the 3D-WSG control unit. All control electrodes and one auxiliary electrode are located on the sputtered monoelectret layer placed on the outer surface of the inner sphere of the gyroscope.
Sensors for measuring the position of the spatial standing wave are located on a monoelectret layer placed on the inner surface of the outer sphere, and the spatial standing wave itself is excited in a thin-walled sphere made of material with low internal friction parameters, covering the inner sphere from the outside through non-magnetic bushings. This variant of the spatial WSG

design also has certain disadvantages. Since the placement of the thin-walled sphere around the inner sphere is carried out through a contact connection by means of non-magnetic bushings, the solid wave gyroscope with such a resonator is actually able to function on a moving object as an angular velocity sensor, which limits the potential capabilities of this device arising from the realization of free spatial movement of the generated standing wave.

To eliminate these disadvantages, a principal variant of a new design of a spherical resonator of a spatial integrating wave solid-state gyroscope is proposed, which makes it possible to expand the functional capabilities of the device by realizing a contactless suspension of a thin-walled sphere in which a spatial standing wave is excited [25]. The proposed design of the resonator of the spatial integrating wave solid-state gyroscope also contains concentrically placed one inside the other outer and inner hollow spheres, a thin-walled sphere placed on the outer side of the inner sphere and made covering the inner sphere from its outer side, electromagnetic centering windings, control electrodes and auxiliary electrode, sensors for measuring the standing wave position, as well as a power supply placed in the inner sphere itself. The outer and inner surfaces of the inner and outer spheres respectively are coated with monoelectret layers having a charge of the same sign and magnitude. In this case, the control electrodes and the auxiliary electrode are located on the monoelectret layer placed on the outer surface of the inner sphere, and the standing wave position measurement sensors are located on the monoelectret layer placed on the inner surface of the outer sphere. The outer and inner surface of the thin-walled sphere are coated with monoelectret layers having charge of the same sign and magnitude corresponding to the charge and magnitude of the monoelectret layers of the inner surface of the outer sphere and the outer surface of the inner sphere [25].

X. RESONATOR OF THE SPATIAL INTEGRATING WAVE GYROSCOPE

As usual, to create a fundamentally new functional inertial orientation sensor using the spatial effect of inertness of elastic waves on the sphere, it is necessary to create conditions for stable observation of this phenomenon in a moving coordinate system. Following the patent [25], the device of contactless energy transfer is placed in the inner sphere /RF patent No. 2792218, 2023/ and the device of standing wave excitation and control electrodes are installed on the inner surface of the thin-walled sphere /decision on granting the patent of the Russian Federation from 09.01.2024 on application No. 2792218, 2023/. In addition, we shoulf point out that the contactless transmission of electrical energy is understood as a method of transmission of electrical energy without the use of conductive elements in the electrical circuit, that is, in the absence of a wired electrical contact between the source and the consumer. In this case, for transmission over a short distance, the phenomenon of electromagnetic induction and resonant mode of

operation of the receiving and transmitting devices are currently used, as well as the near electromagnetic field at distances not exceeding one-sixth of the wavelength to exclude its radiation into free space. As a device for contactless transmission of electric power 10 with transmitting elements 7, for example, a wireless charging system can be used (patent of the Russian Federation authors No. 2781948, 2022).

Due to the use in this device contactless energy transfer "superlens" on the basis of special plates from metamaterial with high anisotropy, in which magnetic permeability is negative in one direction, but positive in all other directions and its location in the direction of the receiver from the outside of the transmitting spiral coil parallel to the plane of its winding turns - when interacting with the magnetic fields of the transmitting coil energy transfer is carried out in a narrow cone with the maximum intensity, which allows to significantly reduce the magnetic fields of the transmitting coil. A standing wave excitation device and control electrodes are installed on the inner surface of the thin-walled sphere /Russian Federation patent No. 222947, 2024/.

The principle design of the new spherical resonator is explained below in Fig.1, where it is schematically represented as a spatial vacuum module. In the figure the positions are marked: 1 - outer hollow sphere; 2 - inner hollow sphere; 3 - thin-walled sphere; 4 - electromagnetic centring windings; 5 - power source; 6- monoelectret layer; 7- transmitting elements; 8- auxiliary electrode; 9- standing wave position measurement sensors; 10 - contactless power transmission device; 11 - standing wave excitation devices and control electrodes.
The resonator is made in the form of a spatial vacuum module

Fig.1

and contains concentrically placed one inside the other outer 1 and inner 2 hollow spheres, a thin-walled sphere 3 placed on the outer side of the inner sphere 2 and encompassing the inner sphere 2. The thin-walled sphere 3 is made of a material with low internal friction parameters (i.e., high goodness), such as an amorphous or nanocrystalline metal/alloy. On the outer and inner surface of the thin-walled sphere 3 are applied

monoelectret layers 6 having a charge of the same sign and magnitude corresponding to the charge and magnitude of the monoelectret layers 6 of the inner surface of the outer sphere 1 and the outer surface of the inner sphere 2, representing, for example, a monoelectret film. The outer and inner surfaces of the inner sphere 2 and the outer sphere 1 are coated with monoelectret layers 6 having a charge of one sign and magnitude corresponding to the charge and magnitude of the monoelectret layers 6 of the inner surface of the outer sphere 1 and the outer surface of the inner sphere 2, representing, for example, a monoelectret film. The inner sphere 2 includes a power supply 5 with a DC-AC converter for powering the electromagnetic centring windings 4 installed in the inner sphere 2, and a non-contact power transmission device 10 with transmitting elements 7. The transmitting elements 7 of the non-contact power transfer device 10 and the auxiliary electrode 8 are arranged on a monoelectret layer placed on the outer surface of the inner sphere, and the standing wave position measurement sensors 9 are arranged on the monoelectret layer placed on the inner surface of the outer sphere 1. [25]. On the inner surface of the thin-walled sphere 3, standing wave excitation devices, for example, of piezoelectric type and control electrodes 11 are installed.

The inner 2 and outer 1 sphere are made of non-magnetic material, for example, high-strength composite materials of low density (ρ=1.45...1.6 g/cm3) based on, for example, carbon plastics; or polydicyclopentadiene.

The WSG resonator of the new design works as follows. When the supply voltage is applied (gyroscope is switched on) to the electromagnetic centering windings 4, the spatial vacuum module of the resonator is activated, providing contactless suspension of the inner hollow sphere 2 and the thin-walled sphere 3. To parametrically excite a spatial mechanical standing wave, two diametrically opposed forces are applied to the thin-walled sphere 3, generated by the standing wave excitation devices and control electrodes 11 placed on the inner surface of the thin-walled sphere 3. The standing wave excitation devices with control electrodes 11 placed on the inner surface of the thin-walled sphere 3 function with the help of the contactless energy transfer device 10 and its transmitting elements 7 placed respectively in the inner sphere 1 and on its outer surface. The number of control electrodes is conditioned by the peculiarities of the WSG design and the parameters of the generated standing wave in the resonator, since these electrodes set the generation of the standing wave.

Under the action of the applied forces, the thin-walled sphere 3 will take an ellipsoidal shape. The applied forces at the given moments of time become equal to zero, and the thin-walled closed sphere 3, possessing a certain stiffness, passing through the undeformed state, under the action of inertia forces will deform in the orthogonal direction. Excitation of a spatial standing wave in the thin-walled sphere 3 occurs in the

autogenerator mode at the natural frequency of oscillation [25]. Upon reaching a predetermined oscillation amplitude of the thin-walled sphere 3, the transmitting elements 7 of the contactless power transmission device 11 are switched off, and the auxiliary electrode 8 is switched on, which maintains the constancy of the oscillation amplitude in the thin-walled sphere 3 for any position of the standing wave relative to the base. As the thin-walled sphere 3 oscillates, the overlapping area of the electrodes of the sensors 9 performing the function of measuring the position of the standing wave changes. The output signals of the electrodes of the sensors 9 corresponding to the three components of the input angular velocity vector of the most of the moving object, are fed to the control unit of the WSG. After software mathematical processing of the primary signal in the electronic block of the calculator, the rotation angle/angular velocity of the mobile object on which the WSG is installed is determined. After software mathematical processing of the primary signal in the electronic block of the calculator, the rotation angle/angular velocity of the mobile object on which the WSG is installed is determined.

The proposed fundamentally new design of non-contact suspension of a thin-walled closed spherical resonator of a spatial integrating wave solid-state gyroscope provides a direct possibility of parametric excitation of a spatial mechanical standing wave on the outer surface of the sphere itself, which makes it possible to determine three components of rotational and/or rotational motion of a mobile object.

XI. CONCLUSION

For a full elastic spherical shell, the theoretical Bryan scaling factor is obtained, which is equal to one quarter (k = 0.25). This value of the scaling factor is somewhat smaller than it was previously obtained for a ring (or cylinder) equal to 0.4, or for a hemisphere (k = 0.312).

We emphasize that for a complete spherical shell the Bryan coefficient has a spatial character. It determines the proportionality of angular velocity vectors to each other, in contrast to the case of a ring (cylinder, hemisphere), when the angular velocities of the plane rotation are taken into account. In three-dimensional space, a three-dimensional mechanical standing wave realized in the form of, for example, an ellipsoid of rotation moves relative to an elastic solid body with a vector of angular velocity proportional to the Bryan coefficient (the vector of rotation of the elastic body itself relative to absolute space). For a full elastic spherical shell, the theoretical Bryan scaling factor is obtained, which is equal to one quarter ($k = 0.25$). We emphasise that for a full spherical shell, the Bryan coefficient is spatial in nature. It determines the proportionality of angular velocity vectors to each other, in contrast to the case of a ring (cylinder, hemisphere), when the angular velocities of the plane rotation are taken into account. In three-dimensional space, a three-dimensional mechanical standing wave realized in the form of, for example, an ellipsoid of rotation moves relative to an elastic solid body

with a vector of angular velocity proportional to the Bryan coefficient ($k = 0.25$) to the vector of rotation of the elastic body itself relative to absolute space. In this case, a spatial mechanical standing wave realized on the surface of an elastic two-dimensional sphere in the form of, for example, an ellipsoid of rotation, moves relative to the body with a vector of angular velocity proportional to (the vector of angular velocity of the body itself relative to absolute space. In this case, a spatial mechanical standing wave realized on the surface of an elastic two-dimensional sphere in the form of, for example, an ellipsoid of rotation, moves relative to the body with an angular velocity vector proportional to (the angular velocity vector of the body itself relative to absolute space. Thus, for numerous practical applications we have in one gyroscopic device a very compact three-dimensional inertial angular velocity sensor, which allows us to simultaneously measure the increments of three quasi-coordinates included in the kinematic differential equations, integrating which we solve the classical Darboux problem of spatial orientation of a moving object in three-dimensional space using the angular velocities of the body known from the measurements of the 3-D WTG sensor.

REFERENCE

[1] Foucault L. Demonstration physique du mouvement de la Terre au moyen du pendule // C.r. Acad. sci. Paris, 1851. V. 32. P. 135-138.

[2] Scott W.B. Delco makes low-cost gyro prototype // Aviat. Week. 1982. V. 117, no. 17. pp. 64–72.

[3] Loper E.J., Lynch D.D. Sonic Gyro Fabrication and Testing, Delco Electronics Division, report R77-64, August 1977.

[4] Loper E.J., Lynch D.D. The HRG: A new low-nose inertial rotation sensor. Proc. 16 Jt. Services Data Exchange for Inertial Systems. Los Angeles, CA, 1982.

[5] Loper, E.J., Lynch D.D. Projected system performance based on recent HRG test results // Paper S83–105, IEEE/AIAA 5th Digital Avionics Systems Conference, 31 Oct.-3 Nov., 1983.

[6] Loper E.J., Lynch D.D. Vibratory rotational sensor. Patent EU, no.0141621 A2. 1984.

[7] Zhuravlev, V.F.; Klimov, D.M. About dynamic effects in an elastic rotating ring (in Russian) // Izv. of the USSR Academy of Sciences. MTT. 1983. no. 5. pp. 17-24.

[8] Zhuravlev V.F., Klimov D.M. Wave solid-state gyroscope. (in Russian) Moscow: Nauka, 1985. 125 p.

[9] Leger P. Quapason — a new low-cost vibrating gyroscope // 3nd Intern. Conf. On Integrated Navigation Systems. Saint-Petersburg, 1996. Pt. 1.

[10] Bryan G. H. On the beats in the vibrations of a revolving cylinder or bell // Proc. Camb. Phil. Soc., Vol VII, Nov. 24, pp. 101–111, 1890.

[11] V.F.Zhuravlev, D.M.Klimov. "Applied methods in the theory of vibrations". (in Russian) ISBN:5-02-006627-3, - M.: Nauka. 1988. - 325.P.

[12] Zhuravlev V.F. Theoretical bases of the wave solid-state gyroscope. // Izv. RAS. MTT. 1993. No. 3. pp. 15-26.

[13] Zhuravlev V.F. Principle questions of the theory of new gyroscopic sensors of the family "generalized Foucault pendulum". Problems of mechanics: Collection of articles. To the 90th anniversary of A.Y.Ishlinsky. (in Russian) - Moscow: Fizmazlit, 2003.-832 pp.

[14] Zhuravlev V.F. Wave solid-state gyroscope. Current state, some aspects // Actual problems of aviation and aerospace systems: problems, models, experiment. Kazan, no. 2(33), vol.16, 2011.

[15] Klimov D.M., Zhuravlev V.F., Zhbanov Y.K. Quartz hemispherical resonator (Wave solid-state gyroscope). Moscow: Kim L.A., 2017.

[16] Perelyaev S.E., Bodunov B.P., Bodunov S.B. Wave solid-state gyroscope - a new generation inertial sensor // Proceedings of the XXIV St. Petersburg International Conference on Integrated Navigation Systems. SPb.: JSC "Concern 'TsNII 'Elektropribor", 2017.- P. 287-290.

[17] Matthews.A. «The Operation and Mechanization of the Hemispherical Resonator Gyroscope», 978-1-5386-1647-5/18/©2018 IEEE. pp. 7 – 14.

[18] Zhuravlev. V.F. On the Formation of Feedbacks in the Spatial Oscillator of Van der Pol. // Mechanics of Solids, 2020, vol.55, p.926-931.

[19] Zhuravlev, V.F.; Klimov, D.M. Spatial effect of inertia of elastic waves on the sphere (in Russian) // Izv. RAS. MTT. 2021. No. 3. C. 3-6.

[20] Zhuravlev V., Perelyaev. S. The Generalized Foucault Pendulum is a 3D Integrating Gyroscopes Using the Three-Dimensional Precession of Standing Waves in a Rotating Spherically Symmetric Elastic Solid. // DGON Inertial Sensors and Systems Symposium. IEEE. Publication Year: 2019, Page(s): pp.1 – 12.

[21] Zhuravlev. V.F. On the Formation of Feedbacks in the Spatial Oscillator of Van der Pol, in: Prikladnaya matematika i mekhanika. 2020, Vol. 84, no. 2, pp. 151–157

[22] Perelyaev S.E., Zhuravlev V.F., Bodunov S.B., Bodunov B.P. «New Gyroscopes of the "Generalized Foucault Pendulum" Family: Some Fundamental Issues of Theory and Applied Aspects of Its Implementation in the Engineering Practice of Modern Gyroscopy». 27th Saint Petersburg International Conference on Integrated Navigation Systems, IEEE, Publication Year: 2020.

[23] Perelyaev S.E., Skripkin A.A. Patent No. 2763688 Russian Federation, MPK G01C 19/56 (2021.08). Spatial integrating solid wave gyroscope. Applicant. 05.10.2020; Published on 30.12.2021.

[24] Perelyaev S.E., Skripkin A.A., Cherkasova O.A. Patent No. 216847 RU Russian Federation, MPK G01.C 19/58 (2006/01). Resonator of spatial integrating solid-state wave gyroscope. Appl. 14.11.2022, Published 03.03.2023.

[25] Cherkasova O.A., Perelyaev S.E., Skripkin A.A. Patent No. 222947 Russian Federation, SPC G01C 19/58 (2023.08). Resonator of spatial integrating solid wave gyroscope. Appl. 05.12.2023; Published 24.01.2024.

Quasi-Optimal Angular Acceleration of Spacecraft Obtained on the Basis of the Poinsot Concept

A.V. Molodenkov
Institute of Precision Mechanics
and Control of RAS
Saratov, Russia
molalexei@yandex.ru

Ya.G. Sapunkov
Institute of Precision Mechanics
and Control of RAS
Saratov, Russia
SapunkovYaG@mail.ru

T.V. Molodenkova
Yu.A. Gagarin Saratov State
Technical University
Saratov, Russia
moltw@yandex.ru

Abstract— **The kinematic problem of finding the optimal program angular acceleration of spacecraft under arbitrary boundary conditions for angular position and angular velocity of spacecraft is considered. The quadratic functional of the energy spent on maneuvering a spacecraft with a fixed transition time is used as optimality criterion. Based on the Poinsot concept, which represents an arbitrary angular motion of a rigid body around a fixed point as a generalized conical motion, using the Pontryagin maximum principle, the quasi-optimal analytical solution of the problem is obtained, which is brought to the algorithm. This report presents the theoretically and numerically justified analytical quasi-optimal algorithm for constructing program angular acceleration of spacecraft, which, according to the results, is close to optimal.**

Keywords—spacecraft, rigid body, angular acceleration, arbitrary boundary conditions, quasi-optimal control, explicit solution, algorithm

I. Introduction

When solving many problems of guidance and control of spacecraft, it is necessary to know the optimal angular acceleration of spacecraft in one sense or another under arbitrary given boundary conditions. As a rule, in the literature such acceleration is either calculated numerically or constructed analytically on the basis of polynomials (splines) by representing the orientation quaternion of a spacecraft by polynomials and expressing the angular velocity vector through this quaternion. However, there are no guarantees (proven theorems or considerations from theoretical mechanics) that on the whole set of angular motions of a spacecraft, under any boundary conditions, these analytical solutions will approximate the optimal trajectory of the angular motion of a spacecraft well enough.

The report investigates the classical problem of finding the optimal angular acceleration of spacecraft under arbitrary boundary conditions of the angular position and angular velocity of spacecraft. The quadratic functional of the energy spent on spacecraft maneuvering with a fixed transition time is used as an optimality criterion. The exact solution to the problem of finding the orientation of a rigid body by its known angular velocity (the Darboux problem) in general case is not known. Within the bounds of the classical Poinsot concept, which interprets arbitrary angular motion of a rigid body in terms of precession cones, or otherwise generalized conical motion, a modification of the problem of optimal control of the angular motion of a spacecraft is carried out, and its trajectory is given by explicit expressions in this class of motions containing arbitrary constants and two arbitrary scalar functions (parameters of generalized conical motion).

With respect to these functions and their derivatives, an optimization problem is formulated and solved, in which the second derivatives of these two functions act as controls. At the same time, the generality of the original problem is practically not violated: the known exact solution to the classical problem of optimal angular motion in the case of flat rotation of spacecraft or the new case of regular precession of spacecraft obtained in this report and similar solutions to the problem within the bounds of the considered concept completely coincide; in other cases, in numerical calculations of solution of the initial problem and the proposed analytical solution, relative error between values of the optimization functional, which is the defining characteristic of the problem, is no more than one percent including angular maneuvers of spacecraft at large angles. Therefore, the proposed analytical solution of the problem can be interpreted as quasi-optimal with respect to the classical problem of finding the optimal angular acceleration of a spacecraft. Explicit expressions for the orientation quaternion and the angular velocity vector of a spacecraft are given, an explicit formula for the angular acceleration of a spacecraft is obtained by differentiating the expression of the angular velocity vector, an analytical algorithm for solving the problem is provided. This algorithm can be used on board of spacecraft.

The analytical solution method proposed in the report has previously been successfully applied in dynamic problems of optimal spacecraft reorientation of various configurations with different optimization criteria under arbitrary boundary conditions [1-7].

II. Statement of the classical problem

The angular motion of a spacecraft as a rigid body around the center of mass is described by equations [8]:

$$2\dot{\mathbf{\Lambda}} = \mathbf{\Lambda} \circ \mathbf{\omega}(t) \qquad (2.1)$$

$$\dot{\mathbf{\omega}} = \mathbf{\varepsilon} \qquad (2.2)$$

Here $\mathbf{\Lambda}(t) = \lambda_0(t) + \lambda_1(t)i_1 + \lambda_2(t)i_2 + \lambda_3(t)i_3$ is the quaternion describing the position of a spacecraft in an inertial space; $\mathbf{\omega}(t) = \omega_1(t)\mathbf{i}_1 + \omega_2(t)\mathbf{i}_2 + \omega_3(t)\mathbf{i}_3$ is the angular velocity vector of a spacecraft specified by its projections onto spacecraft-fixed coordinate axes; i_1, i_2, i_3 − the units of the hypercomplex space (imaginary Hamiltonian units), which can be identified with the vectors of a three-dimensional vector space $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$; the symbol " $\circ$ " stands for the quaternion product; $\mathbf{\varepsilon}(t) = [\varepsilon_1(t), \varepsilon_2(t), \varepsilon_3(t)]^T$ − the

vector of angular acceleration. Phase coordinates $\mathbf{\Lambda}, \mathbf{\omega}$ and control $\mathbf{\varepsilon}$ satisfy the requirements of the optimal control problem [8] ($\mathbf{\Lambda}(t), \mathbf{\omega}(t)$ are continuous functions and $\mathbf{\varepsilon}(t)$ is a piecewise continuous function); quaternion $\mathbf{\Lambda}(t)$ is normalized, i.e. $\|\mathbf{\Lambda}\| = \lambda_0^2 + \lambda_1^2 + \lambda_2^2 + \lambda_3^2 = 1$.

Arbitrary boundary conditions for the angular position

$$\mathbf{\Lambda}(0) = \mathbf{\Lambda}_0, \ \mathbf{\Lambda}(T) = \mathbf{\Lambda}_T \qquad (2.3)$$

and angular velocity of a spacecraft

$$\mathbf{\omega}(0) = \mathbf{\omega}_0, \ \mathbf{\omega}(T) = \mathbf{\omega}_T \qquad (2.4)$$

are given. The task is to find the optimal control $\mathbf{\varepsilon}^{\text{opt}}(t)$ that delivers a minimum to the quadratic criterion of energy costs with a fixed time T:

$$J = \int_0^T \mathbf{\varepsilon}^{\text{T}} \mathbf{\varepsilon}\, dt \qquad (2.5)$$

III. Transition to dimensionless variables

Since the main variable of the problem quaternion $\mathbf{\Lambda}$ is a dimensionless quantity, using the formulas

$$t^{\text{dimls}} = t^{\text{dim}}/T, \ \mathbf{\omega}^{\text{dimls}} = \mathbf{\omega}^{\text{dim}}T, \ \mathbf{\varepsilon}^{\text{dimls}} = \mathbf{\varepsilon}^{\text{dim}}T^2, \ J^{\text{dimls}} = J^{\text{dim}}T^3$$

we reformulate the problem in dimensionless variables, while all expressions of the problem statement except for the optimization functional

$$J = \int_0^1 \mathbf{\varepsilon}^{\text{T}} \mathbf{\varepsilon}\, dt \qquad (3.1)$$

will not change. Next, we consider the dimensionless formulation of problem (2.1)–(2.4) and (3.1) $T=1$.

IV. Application of the maximum principle

Applying the Pontryagin maximum principle [8, 9], we introduce the conjugate variables $\mathbf{\Psi}(t)$ (quaternion) and $\mathbf{\varphi}(t)$ (vector) to the phase variables $\mathbf{\Lambda}(t), \mathbf{\omega}(t)$, respectively. The Hamilton-Pontryagin function

$$H = -(\mathbf{\varepsilon}, \mathbf{\varepsilon}) + (\mathbf{\Psi}, \mathbf{\Lambda} \circ \mathbf{\omega})/2 + (\mathbf{\varphi}, \mathbf{\varepsilon}) \qquad (4.1)$$

where "$(\,.\,,\,.\,)$" is the scalar product of vectors.

Equations for conjugate variables

$$\begin{cases} 2\dot{\mathbf{\Psi}} = \mathbf{\Psi} \circ \mathbf{\omega}, \\ \dot{\mathbf{\varphi}} = -\mathbf{vect}(\widetilde{\mathbf{\Lambda}} \circ \mathbf{\Psi})/2 \end{cases} \qquad (4.2)$$

where "$\mathbf{vect}(.)$" is the vector part of a quaternion and the symbol "~" means quaternion conjugation.

Quaternion equations for variables $\mathbf{\Psi}$ and $\mathbf{\Lambda}$ coincide; therefore, their solutions differ by a constant quaternion C:

$$\mathbf{\Psi} = \mathbf{C} \circ \mathbf{\Lambda} \qquad (4.3)$$

Due to this, and taking into account the designation [8]

$$\mathbf{p} = \mathbf{vect}(\widetilde{\mathbf{\Lambda}} \circ \mathbf{\Psi}) = \widetilde{\mathbf{\Lambda}} \circ \mathbf{c}_v \circ \mathbf{\Lambda} \qquad (4.4)$$

where $\mathbf{c}_v$ is the vector part of quaternion $\mathbf{C}$, expressions (4.2) take the form

$$\begin{cases} \mathbf{p} = \widetilde{\mathbf{\Lambda}} \circ \mathbf{c}_v \circ \mathbf{\Lambda}, \\ \dot{\mathbf{\varphi}} = -\mathbf{p}/2 \end{cases} \qquad (4.5)$$

Thus, due to the self-adjointness of the linear differential system of equations (2.1), the dimension of the boundary value problem of the maximum principle is reduced by four [8]. From the maximum condition (4.1), we obtain the continuous structure of the optimal control:

$$\mathbf{\varepsilon}^{\text{opt}} = \mathbf{\varphi}/2 \qquad (4.6)$$

From (2.1), (2.2), (4.5), and (4.6) we have:

$$\dot{\mathbf{\omega}} = \mathbf{\varphi}/2, \quad \ddot{\mathbf{\omega}} = -\mathbf{p}/4, \quad \dddot{\mathbf{\omega}} = -\dot{\mathbf{p}}/4 = \ddot{\mathbf{\omega}} \times \mathbf{\omega}$$

where "$\times$" denotes the vector product.

Therefore, on the entire interval of motion, the optimal angular velocity of a spacecraft must satisfy the third-order differential vector equation

$$\dddot{\mathbf{\omega}} = \ddot{\mathbf{\omega}} \times \mathbf{\omega} \qquad (4.7)$$

Thereby, the solution of the optimal control problem is reduced to the solution of boundary value problem (2.1), (4.7), (2.3), and (2.4).

Given variable $\mathbf{p}$ (4.4), the Hamilton-Pontryagin function (4.1) will be written as follows:

$$H = -(\mathbf{\varepsilon}, \mathbf{\varepsilon}) + (\mathbf{p}, \mathbf{\omega})/2 + (\mathbf{\varphi}, \mathbf{\varepsilon}) \qquad (4.8)$$

V. The exact solution of the problem of finding the optimal angular acceleration of spacecraft in the particular case

We will seek the solution of equations (2.1), (4.7) in the class of conical motions. To do this, we represent the optimal angular velocity of a spacecraft as

$$\mathbf{\omega} = \widetilde{\mathbf{K}} \circ (\mathbf{i}_1 \alpha \sin \Omega t + \mathbf{i}_2 \alpha \cos \Omega t + \mathbf{i}_3 \Omega) \circ \mathbf{K} \qquad (5.1)$$

where $\mathbf{K}$ (quaternion) and α, Ω are undetermined constants. Here

$$\|\mathbf{K}\| = K_0^2 + K_1^2 + K_2^2 + K_3^2 = 1 \qquad (5.2)$$

Note that the quaternion $\mathbf{K}$ makes it possible to rotate the vector in parentheses in (5.1) about a constant axis passing through the fixed point of the spacecraft.

Successively differentiating (5.1) three times with respect to the variable t, we obtain

Successively differentiating (5.1) three times with respect to the variable t, we obtain:

$$\dot{\mathbf{\omega}} = \alpha\Omega\widetilde{\mathbf{K}} \circ (\mathbf{i}_1 \cos \Omega t - \mathbf{i}_2 \sin \Omega t) \circ \mathbf{K} \qquad (5.3)$$

$$\ddot{\mathbf{\omega}} = -\alpha\Omega^2\widetilde{\mathbf{K}} \circ (\mathbf{i}_1 \sin \Omega t + \mathbf{i}_2 \cos \Omega t) \circ \mathbf{K} \qquad (5.4)$$

$$\dddot{\mathbf{\omega}} = \alpha\Omega^3\widetilde{\mathbf{K}} \circ (-\mathbf{i}_1 \cos \Omega t + \mathbf{i}_2 \sin \Omega t) \circ \mathbf{K} \qquad (5.5)$$

Substituting (5.1), (5.4), and (5.5) in (5.7), we verify that the equality holds true. Here, we note that

$$\dddot{\mathbf{\omega}} = \ddot{\mathbf{\omega}} \times \mathbf{\omega} = (\ddot{\mathbf{\omega}} \circ \mathbf{\omega} - \mathbf{\omega} \circ \ddot{\mathbf{\omega}})/2$$

The trajectory of spacecraft motion with angular velocity (5.1) in (2.1) can be found explicitly – it is the regular precession

$$\boldsymbol{\Lambda}(t) = \boldsymbol{\Lambda}_0 \circ \widetilde{\boldsymbol{K}} \circ \exp\{\mathbf{i}_2 \alpha t / 2\} \circ \exp\{\mathbf{i}_3 \Omega t / 2\} \circ \boldsymbol{K} \qquad (5.6)$$

where $\exp\{.\}$ denotes the quaternion exponential function [8].

Expressions (5.1)–(5.6) involve five arbitrary constants $K_0, K_1, K_2, \alpha, \Omega$. Let us satisfy the boundary conditions (2.3), (2.4). Because of the insufficient number of arbitrary constants in solution (5.1), we impose additional requirements on the quantities $|\boldsymbol{\omega}_0|$ and $\boldsymbol{\omega}_T$. The direction of the unit vector of the initial angular velocity $\boldsymbol{\omega}_0^e = \boldsymbol{\omega}_0 / |\boldsymbol{\omega}_0|$ is arbitrary and given. At t=0, we obtain from (5.1):

$$\boldsymbol{\omega}_0 = |\boldsymbol{\omega}_0| \boldsymbol{\omega}_0^e = \widetilde{\boldsymbol{K}} \circ (\mathbf{i}_2 \alpha + \mathbf{i}_3 \Omega) \circ \boldsymbol{K} \qquad (5.7)$$

$$\|\boldsymbol{\omega}_0\| = \|\widetilde{\boldsymbol{K}} \circ (\mathbf{i}_2 \alpha + \mathbf{i}_3 \Omega) \circ \boldsymbol{K}\| = \|\widetilde{\boldsymbol{K}}\| \|\mathbf{i}_2 \alpha + \mathbf{i}_3 \Omega\| \|\boldsymbol{K}\| = \alpha^2 + \Omega^2 \qquad (5.8)$$

At $t=1$ (the terminal time), we obtain from (5.6)

$$\boldsymbol{\Lambda}_T = \boldsymbol{\Lambda}_0 \circ \widetilde{\boldsymbol{K}} \circ \exp\{\mathbf{i}_2 \alpha / 2\} \circ \exp\{\mathbf{i}_3 \Omega / 2\} \circ \boldsymbol{K} \qquad (5.9)$$

here

$$\text{scal}(\widetilde{\boldsymbol{\Lambda}}_0 \circ \boldsymbol{\Lambda}_T) = \text{scal}(\exp\{\mathbf{i}_2 \alpha / 2\} \circ \exp\{\mathbf{i}_3 \Omega / 2\}) \qquad (5.10)$$

where "scal(.)" denotes the scalar part of the quaternion. From (5.2), (5.7) – (5.9) we find $||\boldsymbol{\omega}_0|, \; \alpha, \Omega$, and $\boldsymbol{K}$.

Let us represent (5.7), (5.9) in the form:

$(\mathbf{i}_2 \alpha + \mathbf{i}_3 \Omega) \circ \boldsymbol{K} - \boldsymbol{K} \circ \boldsymbol{\omega}_0 = 0$

$\exp\{\mathbf{i}_2 \alpha / 2\} \circ \exp\{\mathbf{i}_3 \Omega / 2\} \circ \boldsymbol{K} - \boldsymbol{K} \circ \widetilde{\boldsymbol{\Lambda}}_0 \circ \boldsymbol{\Lambda}_T = 0$

or in vector–matrix form using quaternion matrices of m- and n- types [10]:

$$\left(\begin{bmatrix} 0 & 0 & -\alpha & -\Omega \\ 0 & 0 & -\Omega & \alpha \\ \alpha & \Omega & 0 & 0 \\ \Omega & -\alpha & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & -\omega_{01} & -\omega_{02} & -\omega_{03} \\ \omega_{01} & 0 & \omega_{03} & -\omega_{02} \\ \omega_{02} & -\omega_{03} & 0 & \omega_{01} \\ \omega_{03} & \omega_{02} & -\omega_{01} & 0 \end{bmatrix} \right) \begin{pmatrix} K_0 \\ K_1 \\ K_2 \\ K_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad (5.11)$$

$$\left(\begin{bmatrix} m_0 & -m_1 & -m_2 & -m_3 \\ m_1 & m_0 & -m_3 & m_2 \\ m_2 & m_3 & m_0 & m_1 \\ m_3 & -m_2 & m_1 & m_0 \end{bmatrix} - \begin{bmatrix} n_0 & -n_1 & -n_2 & -n_3 \\ n_1 & n_0 & n_3 & -n_2 \\ n_2 & -n_3 & n_0 & n_1 \\ n_3 & n_2 & -n_1 & n_0 \end{bmatrix} \right) \begin{pmatrix} K_0 \\ K_1 \\ K_2 \\ K_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad (5.12)$$

where the elements of the coefficient matrix of the system of linear algebraic equations (SLAE) (5.12) are determined by the components of quaternions:

$\mathbf{m} = \exp\{\mathbf{i}_2 \alpha / 2\} \circ \exp\{\mathbf{i}_3 \Omega / 2\}$,

$m_0 = \cos(\alpha / 2)\cos(\Omega / 2)$, $m_1 = \sin(\alpha / 2)\sin(\Omega / 2)$, (5.13)

$m_2 = \sin(\alpha / 2)\cos(\Omega / 2)$, $m_3 = \cos(\alpha / 2)\sin(\Omega / 2)$

and

$$\mathbf{n} = \widetilde{\boldsymbol{\Lambda}}_0 \circ \boldsymbol{\Lambda}_T \qquad (5.14)$$

Here

$$\|\mathbf{m}\| = 1 \quad (5.15) \qquad\qquad \|\mathbf{n}\| = 1 \quad (5.16)$$

Note that the ranks of the coefficient matrices of systems (5.11) and (5.12) are 2.

Choosing two linearly independent equations in each of systems (5.11) and (5.12), we obtain the SLAE

$$\begin{bmatrix} 0 & \omega_{01} & \omega_{02} - \alpha & \omega_{03} - \Omega \\ -\omega_{01} & 0 & -(\omega_{03} + \Omega) & \omega_{02} + \alpha \\ m_1 - n_1 & 0 & -(m_3 + n_3) & m_2 + n_2 \\ 0 & n_1 - m_1 & n_2 - m_2 & n_3 - m_3 \end{bmatrix} \begin{pmatrix} K_0 \\ K_1 \\ K_2 \\ K_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad (5.17)$$

For the existence of nontrivial solutions of the homogeneous SLAE, the determinant of the coefficient matrix of SLAE must be zero. Thus, taking into account (5.8), (5.9), (5.12), and (5.13), we obtain

$$|\boldsymbol{\omega}_0| = (m_2 \alpha + m_3 \Omega) / (n_1 \omega_{01}^e + n_2 \omega_{02}^e + n_3 \omega_{03}^e) \qquad (5.18)$$

From (5.8), (5.10), and (5.18) with regard to (5.13), we obtain two equations to determine α and Ω :

$$\begin{cases} (\alpha^2 + \Omega^2)(n_1 \omega_{01}^e + n_2 \omega_{02}^e + n_3 \omega_{03}^e)^2 - (m_2 \alpha + m_3 \Omega)^2 = 0, \\ \text{scal}(\widetilde{\boldsymbol{\Lambda}}_0 \circ \boldsymbol{\Lambda}_T) - \cos(\alpha / 2)\cos(\Omega / 2) = 0 \end{cases} \qquad (5.19)$$

The unknown $|\boldsymbol{\omega}_0|$ can be found by (4.18) after solving (5.19).

In (5.17), we replace the last equation with equation (5.2). From the first three equations in (5.17), K_0, K_1, K_2 are expressed in terms of K_3 as

$$K_0 = A_0 K_3 / D, \quad K_1 = A_1 K_3 / D, \quad K_2 = A_2 K_3 / D \qquad (5.20)$$

where

$$A_0 = \omega_{01}[-(\omega_{03} + \Omega)(m_2 + n_2) + (\omega_{02} + \alpha)(m_3 + n_3)] \qquad (5.21)$$

$$A_1 = \omega_{01}\left[(\Omega - \omega_{03})\frac{D}{\omega_{01}^2} - (m_2 + n_2)(\omega_{02} - \alpha) - (m_1 + n_1)\frac{\omega_{02}^2 - \alpha^2}{\omega_{01}}\right] \qquad (5.22)$$

$$A_2 = \omega_{01}[(m_1 - n_1)(\omega_{02} + \alpha) + (m_2 + n_2)\omega_{01}] \qquad (5.23)$$

$$D = \omega_{01}[(m_1 - n_1)(\omega_{03} + \Omega) + (m_3 + n_3)\omega_{01}] \qquad (5.24)$$

From (5.2) and (5.21)–(5.24) we have

$$K_3 = \left[1 + (A_0 / D)^2 + (A_1 / D)^2 + (A_2 / D)^2\right]^{1/2} \qquad (5.25)$$

After determining K_3 from (5.25), K_0, K_1, K_2 are calculated using (5.20). The angular velocity boundary condition at the terminal time T in dimensionless variables ($T=1$) must have the form

$$\boldsymbol{\omega}(T) = \boldsymbol{\omega}_T = \widetilde{\boldsymbol{K}} \circ (\mathbf{i}_1 \alpha \sin\Omega + \mathbf{i}_2 \alpha \cos\Omega + \mathbf{i}_3 \Omega) \circ \boldsymbol{K} \qquad (5.26)$$

Thus, in cases when the angular velocity boundary conditions satisfy constraints (5.18), (5.26) (this implies that the vector $\boldsymbol{\omega}$ belongs, on the whole time interval of motion, to a conical surface that is defined by the arbitrary spacecraft attitude boundary conditions $\boldsymbol{\Lambda}_0, \boldsymbol{\Lambda}_T$ and the arbitrary direction of the angular velocity initial value $\boldsymbol{\omega}_0^e$), the trajectory of the angular motion of spacecraft is in the class of conical motions, and it can be calculated by explicit analytic expressions (5.1) and (5.6).

The optimal angular acceleration from (2/2), (5.1)

$$\boldsymbol{\varepsilon} = \dot{\boldsymbol{\omega}} = \alpha\Omega\widetilde{\boldsymbol{K}} \circ (\mathbf{i}_1 \cos\Omega t - \mathbf{i}_2 \sin\Omega t) \circ \boldsymbol{K} \qquad (5.27)$$

$$|\boldsymbol{\varepsilon}|^2 = \alpha^2 \Omega^2 = const \qquad (5.28)$$

The optimal value of the performance index in dimensionless variables is

$$J = \int_0^1 \boldsymbol{\varepsilon}^2 dt = \alpha^2 \Omega^2 \qquad (5.29)$$

The conjugate variables $\boldsymbol{\varphi}$ and $\mathbf{p}$ can be found using (5.27), (4.5), and (4.4). Thus, the problem is completely solved subject to the given constraints.

Now we describe an algorithm for solving the optimal slew problem for a spacecraft (rigid body) in the class of conical motions in dimensionless variables:

1) according to the given quaternions $\boldsymbol{\Lambda}_0, \boldsymbol{\Lambda}_T$ (2.3), and the unit vector $\boldsymbol{\omega}_0^e$ from (2.4), the time $T=1$, use formulas (5.19), (5.18) (taking into account (5.13), and (5.14)) to calculate α, Ω, and $|\boldsymbol{\omega}_0|$;

2) using $\boldsymbol{\Lambda}_0, \boldsymbol{\Lambda}_T$, α, Ω, and $|\boldsymbol{\omega}_0|$ calculate the components of the quaternion $\mathbf{K}$ by formulas (5.20)–(5.25);

3) using the formula $\boldsymbol{\omega}_0^{\text{calc}} = |\boldsymbol{\omega}_0| \boldsymbol{\omega}_0^e$ and (5.26)

$$\boldsymbol{\omega}_T^{\text{calc}} = \widetilde{\mathbf{K}} \circ (\mathbf{i}_1 \alpha \sin \Omega + \mathbf{i}_2 \alpha \cos \Omega + \mathbf{i}_3 \Omega) \circ \mathbf{K}$$

are calculated the vectors $\boldsymbol{\omega}_0^{\text{calc}}, \boldsymbol{\omega}_T^{\text{calc}}$;

4) the resultant $\boldsymbol{\omega}_0^{\text{calc}}, \boldsymbol{\omega}_T^{\text{calc}}$ are compared with $\boldsymbol{\omega}_0, \boldsymbol{\omega}_T$ defined in (2.4);

5) if the equation at step 4 of the algorithm holds true, then the optimal solution is found in the class of conical motions. The angular velocity of the spacecraft, the trajectory of its angular motion, the vector of the control torque, and the value of the performance index are estimated by (5.1), (5.6), (5.27), and (5.29), respectively;

6) the conjugate variables of the problem $\boldsymbol{\varphi}$ and $\mathbf{p}$ are calculated by (4.4), (4.5), and (5.27).

The optimal solution of the problem obtained in the class of regular conical motions and the previously known [8] exact solution in the class of plane Eulerian rotations (provided $\boldsymbol{\omega}_0, \boldsymbol{\omega}_T \parallel \mathbf{vect}(\widetilde{\boldsymbol{\Lambda}}_0 \circ \boldsymbol{\Lambda}_T)$) for special cases of boundary conditions for the angular velocity of a spacecraft will be used as analytical confirmations in constructing the quasi-optimal solution to the problem of finding the angular acceleration of spacecraft in the class of generalized conical motions.

VI. Modified problem of finding the optimal angular acceleration of spacecraft

The motion of a spacecraft is still described by relations (2.1)–(2.4), while the initial and final values for the angular position and angular velocity of spacecraft are arbitrary and specified. One of the main problems in constructing an analytical solution in the problem of the optimal rotation of a rigid body is the analytical solvability of the classical Darboux problem – the problem of finding the quaternion $\boldsymbol{\Lambda}(t)$ through the angular velocity vector $\boldsymbol{\omega}(t)$.

For the quaternion differential equation (2.1), provided that the angular velocity vector $\boldsymbol{\omega}(t)$ is given by the expression

$$\boldsymbol{\omega}(t) = \mathbf{i}_1 \dot{f}(t) \sin g(t) + \mathbf{i}_2 \dot{f}(t) \cos g(t) + \mathbf{i}_3 \dot{g}(t) \quad (6.1)$$

in which $f(t)$ and $g(t)$ are arbitrary functions of time, the solution satisfying the initial condition (2.3) is known [11]:

$$\boldsymbol{\Lambda}(t) = \boldsymbol{\Lambda}_0 \circ \exp\{-\mathbf{i}_3 g(0)/2\} \circ \exp\{-\mathbf{i}_2 f(0)/2\} \circ$$
$$\circ \exp\{\mathbf{i}_2 f(t)/2\} \circ \exp\{\mathbf{i}_3 g(t)/2\} \quad (6.2)$$

Expression (6.1) and solution (6.2) can be generalized by adding a rotation by a constant angle around some axis. This rotation is specified using the quaternion $\mathbf{K}$, $\|\mathbf{K}\| = 1$. Then the vector $\boldsymbol{\omega}$ and quaternion $\boldsymbol{\Lambda}$ will be determined by the relations

$$\boldsymbol{\omega} = \widetilde{\mathbf{K}} \circ (\mathbf{i}_1 \dot{f}(t) \sin g(t) + \mathbf{i}_2 \dot{f}(t) \cos g(t) + \mathbf{i}_3 \dot{g}(t)) \circ \mathbf{K} \quad (6.3)$$

$$\boldsymbol{\Lambda}(t) = \boldsymbol{\Lambda}_0 \circ \widetilde{\mathbf{K}} \circ \exp\{-\mathbf{i}_3 g(0)/2\} \circ \exp\{-\mathbf{i}_2 f(0)/2\} \circ$$
$$\circ \exp\{\mathbf{i}_2 f(t)/2\} \circ \exp\{\mathbf{i}_3 g(t)/2\} \circ \mathbf{K} \quad (6.4)$$

As can be seen, expressions (6.3), (6.4) include expressions for the angular velocity and trajectory of spacecraft in the exact solution of the problem of finding the optimal angular acceleration of spacecraft when the vector of its angular velocity maintains the constant direction over the entire time interval of the spacecraft movement [8] or performs a regular precession (sec. V). How else one supporting consideration should be noted that using one-to-one variable substitutions [11], the Darboux problem can generally be reduced to solving an equation of type (1.3), where the angular velocity will take the form

$$\boldsymbol{\omega}^*(t) = -(\mathbf{i}_1 \dot{f}(t) \sin g(t) + \mathbf{i}_2 \dot{f}(t) \cos g(t) + \mathbf{i}_3 \dot{g}(t))$$

which differs from (6.1) only in sign. Moreover, an explicit analytical solution to this problem, as in the case of an arbitrary vector $\boldsymbol{\omega}(t)$, is not known. In other words, the proposed structure of angular velocity (6.1) ((6.3)) is closely correlated to the Poinsot concept that any arbitrary angular motion of a rigid body around a fixed point can be considered as some generalized conical motion of a rigid body [12].

We will consider the second derivatives of the functions f and g as the control parameters. Then, if we introduce the notation

$$\dot{f} = f_1, \qquad \dot{g} = g_1 \quad (6.5)$$

we can compose a system of differential equations describing the controlled system:

$$\dot{f} = f_1, \ \ \dot{g} = g_1, \ \ \dot{f}_1 = u_1, \ \ \dot{g}_1 = u_2 \quad (6.6)$$

where f, f_1, g, and g_1 are the phase coordinates, while u_1 and u_2 are the control parameters. We restrict ourselves to the case when quaternion $\mathbf{K}$ is presented in the form of a product:

$$\mathbf{K} = \mathbf{K}_2 \circ \mathbf{K}_1, \ \ \mathbf{K}_1 = \exp\{\mathbf{i}_1 \alpha_1/2\},$$
$$\mathbf{K}_2 = \exp\{\mathbf{i}_2 \alpha_2/2\} \quad (6.7)$$

where α_1 and α_2 are some constants. Note that quaternions $\mathbf{K}_1$ and $\mathbf{K}_2$ determine the rotation of the vector $\boldsymbol{\omega}$ (6.1) around axes $\mathbf{i}_1$ and $\mathbf{i}_2$. Rotation around axis $\mathbf{i}_3$ is included in the formula if we take into account that the function $g(t)$ contains an additive constant.

The conditions ensuring that expressions for $\boldsymbol{\omega}$ and $\boldsymbol{\Lambda}$ (6.3) and (6.4) satisfy the boundary conditions (2.3) and (2.4), taking into account (6.7), can be written as

$$
\begin{aligned}
\widetilde{\mathbf{K}}_1 \circ \widetilde{\mathbf{K}}_2 &\circ (\mathbf{i}_1 f_1(0)\sin g(0) + \mathbf{i}_2 f_1(0)\cos g(0) + \\
&+ \mathbf{i}_3 g_1(0)) \circ \mathbf{K}_2 \circ \mathbf{K}_1 = \boldsymbol{\omega}_0
\end{aligned} \tag{6.8}
$$

$$
\begin{aligned}
\widetilde{\mathbf{K}}_1 \circ \widetilde{\mathbf{K}}_2 &\circ (\mathbf{i}_1 f_1(T)\sin g(T) + \mathbf{i}_2 f_1(T)\cos g(T) + \\
&+ \mathbf{i}_3 g_1(T)) \circ \mathbf{K}_2 \circ \mathbf{K}_1 = \boldsymbol{\omega}_T
\end{aligned} \tag{6.9}
$$

$$
\begin{aligned}
\boldsymbol{\Lambda}_0 \circ \widetilde{\mathbf{K}}_1 &\circ \widetilde{\mathbf{K}}_2 \circ \exp\{-\mathbf{i}_3 g(0)/2\} \circ \exp\{\mathbf{i}_2(f(T) \\
&- f(0))/2\} \circ \exp\{\mathbf{i}_3 g(T)/2\} \circ \mathbf{K}_2 \circ \mathbf{K}_1 = \boldsymbol{\Lambda}_T
\end{aligned} \tag{6.10}
$$

The control angular acceleration of spacecraft corresponding to the solution of the modified optimal control problem is determined from (2.2) by the formula

$$
\boldsymbol{\varepsilon} = \dot{\boldsymbol{\omega}} \tag{6.11}
$$

Then for the controlled system (6.6) we can formulate the following optimal control problem. It is required to find optimal control functions $u_1(t)$, $u_2(t)$, which transfer the controlled system (6.6) from the initial state

$$
f = f(0), \ f_1 = f_1(0), \ g = g(0), \ g_1 = g_1(0) \tag{6.12}
$$

to the final state

$$
f = f(T), \ f_1 = f_1(T), \ g = g(T), \ g_1 = g_1(T) \tag{6.13}
$$

and satisfying relations (6.9)–(6.10), in which α_1, α_2 are the parameters that need to be determined, and minimize the quadratic functional

$$
J = \int_0^T (u_1^2 + u_2^2)dt \tag{6.14}
$$

This optimal control problem will be called the modified problem of finding the angular acceleration of spacecraft.

VII. ANALYTICAL SOLUTION OF THE PROBLEM (6.6)–(6.14)

The Hamilton-Pontryagin function for the optimal control problem [9]

$$
H = -(u_1^2 + u_2^2) + \psi_1 f_1 + \psi_2 g_1 + \psi_3 u_1 + \psi_4 u_2 \tag{7.1}
$$

Where $\psi_1, \psi_2, \psi_3, \psi_4$ are the conjugate variables satisfying the system of equations

$$
\dot{\psi}_1 = 0, \ \dot{\psi}_2 = 0, \ \dot{\psi}_3 = -\psi_1, \ \dot{\psi}_4 = -\psi_2 \tag{7.2}
$$

the general solution of (7.2), which contains arbitrary constants $c_1,...,c_4$, is

$$
\psi_1 = c_1, \ \psi_2 = c_2, \ \psi_3 = -c_1 t + c_3, \psi_4 = -c_2 t + c_4. \tag{7.3}
$$

The maximum condition for Hamilton–Pontryagin function (7.1) implies

$$
u_1 = (-c_1 t + c_3)/2, \quad u_2 = (-c_2 t + c_4)/2 \tag{7.4}
$$

By substituting (7.4) into system (6.6), we find the general solution for the phase coordinates, which contains eight arbitrary constants $c_1,...,c_8$

$$
f = -c_1 t^3/12 + c_3 t^2/4 + c_5 t + c_6, \ g = -c_2 t^3/12 + c_4 t^2/4 + c_7 t + c_8
$$

$$
f_1 = -c_1 t^2/4 + c_3 t/2 + c_5, \quad g_1 = -c_2 t^2/4 + c_4 t/2 + c_7. \tag{7.5}
$$

Due to the fact that in the expression for the function f, constant c_6 enters additively, it can be seen from (6.4) that it has no effect, and therefore c_6 may not be taken into account. Then, to search for nine unknowns $c_1,...,c_5, c_7, c_8, \alpha_1, \alpha_2$ there are nine equations (6.8)–(6.10) (because of the condition $\|\boldsymbol{\Lambda}\| = 1$ in the quaternion equation (6.10), only three scalar equations are independent). Substituting formulas (7.5) into (6.3) and (6.4), we obtain expressions for the angular velocity vector and the spacecraft orientation quaternion. These expressions determine the optimal in the sense of the minimum of functional (6.14) of the angular maneuver of spacecraft constructed in the class of generalized conical motions. Control angular acceleration of spacecraft, based on (6.3), (6.11)

$$
\begin{aligned}
\boldsymbol{\varepsilon} = \dot{\boldsymbol{\omega}} = \widetilde{\mathbf{K}} &\circ [(\mathbf{i}_1(u_1 \sin g + f_1 g_1 \cos g) + \\
&+ \mathbf{i}_2(u_1 \cos g - f_1 g_1 \sin g) + \mathbf{i}_3 u_2] \circ \mathbf{K}.
\end{aligned} \tag{7.6}
$$

The scalar square of the control angular acceleration vector is expressed in terms of the variables of the modified problem as follows:

$$
\boldsymbol{\varepsilon}^2 = u_1^2 + u_2^2 + f_1^2 g_1^2 \tag{7.7}
$$

The modified problem of finding the optimal angular acceleration of spacecraft has thus been completely solved.

If in the modified problem of finding the optimal angular acceleration of spacecraft, the boundary conditions for the angular velocity of a spacecraft are the same as in sec. V, then the solutions of the classical and modified problems will completely coincide. The same can be said when the vectors of boundary conditions with respect to angular velocity $\boldsymbol{\omega}_0, \boldsymbol{\omega}_T$ (2.4) are collinear $\mathbf{vect}(\widetilde{\boldsymbol{\Lambda}}_0 \circ \boldsymbol{\Lambda}_T)$ (flat Eulerian rotation [8]), in these cases, the term $f_1^2 g_1^2$ in (7.7) vanishes and criterion (6.14) corresponds to criterion (2.5) ((3.1)) of the classical problem. In the problem of finding the optimal angular acceleration of spacecraft under arbitrary boundary conditions, assuming that $\int_0^1 f_1^2 g_1^2 dt$ is less than $\int_0^1 \boldsymbol{\varepsilon}^2 dt$, we can omit the last term in (7.7). Then the modified optimal control problem in variables f, g, f_1, g_1, u_1, u_2 with functional (6.14) and expressions (6.3), (6.4), (6.8)–(6.10), (7.5), and (7.6) will correspond to the classical problem the problem of finding the optimal angular acceleration in the class of generalized conical motions. Herewith the modified problem can be considered as a quasi-optimal problem of finding the angular acceleration of spacecraft under arbitrary boundary conditions.

VIII. ANALYTICAL QUASI-OPTIMAL ALGORITHM

The quasi-optimal algorithm for angular acceleration of spacecraft under arbitrary boundary conditions in dimensionless variables has the following form:

1) according to the given values $\boldsymbol{\Lambda}_0, \boldsymbol{\Lambda}_T, \boldsymbol{\omega}_0, \boldsymbol{\omega}_T$ (2.3), (2.4), and T = 1 from (6.7)–(6.10), nine indefinite constants $c_1,...,c_5, c_7, c_8, \alpha_1, \alpha_2$ and are calculated, and the functions f, f_1, g, g_1 are expressed;

2) using (6.7), we calculate the quaternion $\mathbf{K}$;

3) we determine the vector of the angular velocity of spacecraft

$$\boldsymbol{\omega} = \widetilde{\mathbf{K}} \circ (\mathbf{i}_1 \dot{f}(t) \sin g(t) + \mathbf{i}_2 \dot{f}(t) \cos g(t) + \mathbf{i}_3 \dot{g}(t)) \circ \mathbf{K}$$

4) we find the orientation quaternion of spacecraft

$$\boldsymbol{\Lambda}(t) = \boldsymbol{\Lambda}_0 \circ \widetilde{\mathbf{K}} \circ \exp\{-\mathbf{i}_3 g(0)/2\} \circ \exp\{-\mathbf{i}_2 f(0)/2\} \circ$$
$$\circ \exp\{\mathbf{i}_2 f(t)/2\} \circ \exp\{\mathbf{i}_3 g(t)/2\} \circ \mathbf{K}$$

5) angular acceleration vector of spacecraft

$$\boldsymbol{\varepsilon} = \dot{\boldsymbol{\omega}} = \widetilde{\mathbf{K}} \circ [(\mathbf{i}_1 (u_1 \sin g + f_1 g_1 \cos g) +$$
$$+ \mathbf{i}_2 (u_1 \cos g - f_1 g_1 \sin g) + \mathbf{i}_3 u_2] \circ \mathbf{K}.$$

6) from expressions (3.1) and (7.7), the value of the dimensionless optimization criterion for the problem of the optimal reorientation of spacecraft is found.

IX. NUMERICAL EXAMPLES

Comparative results of numerical solutions of the initial (classical) and modified problems of finding the optimal angular acceleration spacecraft are considered.

The boundary value problem of the maximum principle, obtained in sec. IV for the original classical problem, has the form:

$$\begin{cases} 2\dot{\boldsymbol{\Lambda}} = \boldsymbol{\Lambda} \circ \boldsymbol{\omega}, \\ \dot{\boldsymbol{\omega}} = \boldsymbol{\varepsilon}, \\ \dot{\boldsymbol{\varphi}} = -\mathbf{p}/2, \\ \mathbf{p} = \widetilde{\boldsymbol{\Lambda}} \circ \mathbf{c}_v \circ \boldsymbol{\Lambda}, \quad \mathbf{c}_v = \mathbf{const} \end{cases} \tag{9.1}$$

$$\boldsymbol{\Lambda}(0) = \boldsymbol{\Lambda}_0, \quad \boldsymbol{\omega}(0) = \boldsymbol{\omega}_0 \tag{9.2}$$

$$\boldsymbol{\Lambda}(T) = \boldsymbol{\Lambda}_T, \quad \boldsymbol{\omega}(T) = \boldsymbol{\omega}_T \tag{9.3}$$

$$\boldsymbol{\varepsilon}^{\text{опт}} = \boldsymbol{\varphi}/2 \tag{9.4}$$

where to find $\boldsymbol{\varepsilon}^{\text{opt}}, \boldsymbol{\Lambda}^{\text{opt}}, \boldsymbol{\omega}^{\text{opt}}, \mathbf{c}_v$. Conditions at the final moment of time (9.3) must be rewritten in the seven-dimensional phase space $\boldsymbol{\Lambda} \times \boldsymbol{\omega}$:

$$\text{vect}(\boldsymbol{\Lambda}(T) \circ \widetilde{\boldsymbol{\Lambda}}_T) = 0,$$
$$\boldsymbol{\omega}(T) = \boldsymbol{\omega}_T. \tag{9.5}$$

The method for the numerical solution of such boundary value problems of optimal reorientation of a spacecraft (rigid body) is described in [13].

The first line of table 1 shows the values of the orientation quaternion and the angular velocity vector at t=0.5 (in the middle of the optimal motion time interval $[0, T]$, $T = 1$) for boundary conditions for the angular position and angular velocity of spacecraft

$$\boldsymbol{\Lambda}_0 = (0.7951, 0.2981, -0.3975, 0.3478),$$
$$\boldsymbol{\omega}_0 = (0.2739, -0.2388, -0.3) \tag{9.6}$$

$$\boldsymbol{\Lambda}_T = (0.8443, 0.3984, -0.3260, 0.1485),$$
$$\boldsymbol{\omega}_T = (0.0, 0.0, -0.59) \tag{9.7}$$

The components of the angular acceleration vector under the same boundary conditions for the beginning, middle and end of the optimal control process are considered in the first line of the table. 2. In the second rows of the tables, for comparison, the data obtained by solving the modified

problem of searching for the angular acceleration of spacecraft using the analytical algorithm sec. VIII are presented. In this case, the values of the constants $\alpha_1, \alpha_2, c_1, ..., c_5, c_7, c_8$ included in the analytical solution of the modified problem are as follows;

$$\alpha_1 = -0.0421, \ \alpha_2 = -0.2226, \ c_1 = 3.2902$$
$$c_2 = -1.4885, \ c_3 = 2.2113, \ c_4 = -1.45$$
$$c_5 = -0.4156, c_7 = -0.2221, \ c_8 = -0.9216$$

Table 1

λ_0	λ_1	λ_2	λ_3	ω_1	ω_2	ω_3
0.8096	0.3625	-0.3768	0.2668	-0.0502	-0.0114	-0.4937
0.8099	0.3627	-0.3756	0.2673	-0.0488	-0.0098	-0.4938

Table 2

$\varepsilon_1(0)$	$\varepsilon_2(0)$	$\varepsilon_3(0)$	$\varepsilon_1(0.5)$	$\varepsilon_2(0.5)$	$\varepsilon_3(0.5)$	$\varepsilon_1(1)$	$\varepsilon_2(1)$	$\varepsilon_3(1)$
-0.9854	0.7259	-0.4892	-0.2917	0.2087	-0.2878	0.5077	-0.1272	-0.0985
-0.9647	0.7634	-0.4932	-0.3103	0.1687	-0.2847	0.5350	-0.0220	0.1024

The results of solving the two problems are close. The value of the functional (3.1) of the original problem in the example under consideration is 0.4782. The value of the same functional calculated based on the solution of the modified problem is 0.4797. That is, in the example under consideration, the discrepancy between the values of the functional (3.1) for the classical and modified problem is 0.3%. It should be noted that the value of the optimization functional is a defining characteristic of the optimal control problem.

Numerical solutions of the problem of optimal angular motion control were also carried out for cases when the initial state of spacecraft was determined by the ratios (9.6). The final position of spacecraft was set by rotating a spacecraft from the initial position by some angle around the Eulerian axis, the unit vector of which was determined by the coordinates:

$$(0.0450, -0.0752, -0.9962) \tag{9.8}$$

Table 3 shows the components of the quaternions of the final position of a spacecraft for rotations by various values of the Euler angle φ in degrees around vector (9.8).

Table 3

φ°	$\lambda_0(T)$	$\lambda_1(T)$	$\lambda_2(T)$	$\lambda_3(T)$
30	0.8464	0.4065	-0.3185	0.1298
60	0.8401	0.4872	-0.2178	-0.0970
90	0.7766	0.5346	-0.1023	-0.3173
120	0.6601	0.5456	0.0202	-0.5159
150	0.4986	0.5195	0.1414	-0.6794
180	0.3032	0.4579	0.2529	-0.7965

Table 4 shows the values of functional (3.1) in the classical and modified problems of finding the optimal angular acceleration when transferring spacecraft from the initial state (9.6) to the final states $\boldsymbol{\Lambda}_T$ according to table 3 and to the $\boldsymbol{\omega}_T = (0.0, 0.0, -0.59)$, with indicating the difference between the values of these functionals

$\Delta J = J^{\mathrm{modif}} - J^{\mathrm{classic}}$ and the percentage discrepancy $(\Delta J / J^{\mathrm{classic}}) \cdot 100\%$.

Table 4

Functionals and discrepancies in their values	$\varphi = 30°$	$\varphi = 60°$	$\varphi = 90°$	$\varphi = 150°$	$\varphi = 180°$
J^{classic}	0.52385	4.63277	15.31437	56.39081	86.78094
J^{modif}	0.52510	4.63724	15.32882	56.52086	87.51533
ΔJ	0.00125	0.00447	0.01445	0.13005	0.73439
%	0.24	0.10	0.09	0.23	0.85

Table 5 shows similar indicators when the final angular velocity of a spacecraft is determined by the vector

$$\boldsymbol{\omega}_T = (0.0,\ 0.0,\ 0.0)$$

i.e., in this case, spacecraft is transferred to a state of rest.

Table 5

Functionals and discrepancies in their values	$\varphi = 90°$	$\varphi = 120°$	$\varphi = 150°$	$\varphi = 180°$
J^{classic}	24.25074	45.17597	72.66431	106.71186
J^{modif}	24.28745	45.19513	72.77169	107.40843
ΔJ	0.03672	0.01916	0.10738	0.69657
%	0.15	0.04	0.15	0.65

From numerous experiments for various parameters of the spacecraft angular motion control problem, the proximity of solutions to the classical and modified spacecraft angular acceleration search problems follows, which allows us to consider the solution of the modified problem as the quasi-optimal solution to the classical spacecraft angular acceleration search problem.

X. CONCLUSIONS

The analytical quasi-optimal algorithm for solving the kinematic problem of finding the optimal angular acceleration of spacecraft with arbitrary boundary conditions has the theoretical basis and solves the problem of optimal reorientation of spacecraft with good accuracy. At the same time, it does not require a numerical solution of the boundary value problem of the maximum principle or other complex numerical solution and represents ready-made analytical laws of quasi-optimal program control and program trajectory changes that can be installed on board spacecraft. The results obtained can be applied to nanoclass spacecraft that have limitations on computing power.

REFERENCES

[1] A.V. Molodenkov, Y.G. Sapunkov, "Analytical quasi-optimal solution for the problem on turn maneuver of an arbitrary solid with arbitrary boundary conditions," Mechanics of Solids, vol. 54, no. 3, 2019, pp. 474-485.

[2] A.V. Molodenkov, Y.G. Sapunkov, " Analytical Solution of the Optimal Slew Problem for an Axisymmetric Spacecraft in the Class of Conical Motions," Journal of Computer and Systems Sciences International, vol. 55, no. 6, 2016, pp. 969-985.

[3] A.V. Molodenkov, Y.G. Sapunkov, "Analytical Solution of the Minimum Time Slew Maneuver Problem for an Axially Symmetric Spacecraft in the Class of Conical Motions," Journal of Computer and Systems Sciences International, vol. 57, no. 2, 2018, pp. 302-318.

[4] Y.G. Sapunkov, A.V. Molodenkov, "Analytical solution of the problem on an axisymmetric spacecraft attitude maneuver optimal with respect to a combined functional," Automation and Remote Control, vol. 82, no. 7, 2021, pp. 1183–1200.

[5] A.V. Molodenkov, Y.G. Sapunkov, "Analytical quasi-optimal solution of the problem of the time-optimal rotation of a spacecraft," Journal of Computer and Systems Sciences International, vol. 60, no. 4, 2021, pp. 639-653.

[6] A.V. Molodenkov, Y.G. Sapunkov, "Analytical Quasi-Optimal Algorithm for the Programmed Control of the Angular Motion of a Spacecraft," Journal of Computer and Systems Sciences International, vol. 62, no. 3, 2023, pp. 569-580.

[7] Y.G. Sapunkov, A.V. Molodenkov, "Quasioptimal Spacecraft Attitude Control Constructed According to the Poinsot Concept," Aerospace, vol. 10, no. 5, 2023, p. 402.

[8] V.N. Branetz, I.P. Shmyglevskij, Primenenie kvaternionov v zadachah orientacii tverdogo tela (Application of quaternions to rigid body orientation problems). Nauka, Moscow, 1973

[9] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, and E.F. Mishchenko, The Mathematical Theory of Optimal Processes. Nauka, Moscow, 1961; Gordon and Breach, New York, 1986.

[10] Yu.N. Chelnokov, Kvaternionnye i bikvaternionnye modeli i metody mehaniki tverdogo tela i ih prilogeniya (Quaternion and Biquaternion Models and Methods in Mechanics of Solids and Their Applications). Fizmatlit, Moscow, 2006.

[11] A.V. Molodenkov, "On the solution of the Darboux problem," Mechanics of Solids, vol. 42, no. 2, 2007, pp. 167-176.

[12] A.V. Molodenkov, S.E. Perelyaev, "Solution of approximate equation for modified rodrigues vector and attitude algorithm design," Journal of Guidance, Control, and Dynamics, vol. 44, no. 6, 2021, pp. 1224-1227.

[13] Y.G. Sapunkov, A.V. Molodenkov, "Numerical Solution of the Optimal Spacecraft Reorientation Problem," Mekhatronika, Avtomatizatsiya, Upravlenie, no. 6, 2008, pp. 66-70.

Comparing Recursive and Nonrecursive Bayesian Estimation Algorithms in Trajectory Tracking Using Bearing-Only Measurements

O.A. Stepanov
Concern CSRI Elektropribor, JSC, ITMO University,
St. Petersburg, Russia
soalax@mail.ru

V.P. Zolotarevich
Concern CSRI Elektropribor, JSC, ITMO University,
St. Petersburg, Russia
zolotarevich@yandex.ru

M.S. Ivanov
Concern CSRI Elektropribor, JSC, ITMO University,
St. Petersburg, Russia
msivanovitmo@itmo.ru

A.V. Motorin
Concern CSRI Elektropribor, JSC, ITMO University,
St. Petersburg, Russia
motorin.a@mail.ru

The paper considers estimation of the target coordinates and motion parameters using the bearing measurements from the sensor installed on a maneuvering observer. The accuracies of the iterative batch linearized filter, the recursive iterative batch linearized smoother, and the iterative extended Kalman smoother are compared using the Monte Carlo method.

Keywords—nonrecursive algorithm, recursive algorithm, Kalman filter

I. INTRODUCTION

Surveillance systems are widely employed in different applications for recognition of different targets. In marine surveillance systems, noise direction-finding is the main operating mode for the target tracking. This mode features low power consumption, however, it cannot provide range measurement, and uses only bearing measurements for the target tracking. Target tracking using bearing-only measurements is a nonlinear problem due to the nonlinear dependence of the bearings on the estimated coordinates and the problem of their observability. The latter problem is resolved by the observer maneuvering. Due to nonlinearity and high a priori uncertainty of the estimated parameters, the classical extended Kalman filter (EKF) turns out to be ineffective [1]. Thus, robust estimators of target coordinates and motion need to be developed. Particularly, the performance of nonrecursive estimators presents an interest.

The objective of this study is comparing the accuracies of nonrecursive and recursive estimators presented in [2] in bearing-only target tracking. In Part 2, the mathematical problem statement is given. Part 3 describes the iterative batch linearized smoother and its performance. In Part 4, the algorithm is compared with the recursive iterative batch linearized smoother and with iterative extended Kalman smoother.

II. PROBLEM STATEMENT

We track a target moving at nearly constant velocity by the bearing measurements from the moving observer. Assume that the target motion can be presented as a Markov sequence described by the shaping filter

$$x_i^t = \Phi x_{i-1}^t + \Gamma w_i, \ i = \overline{1.N}, \tag{1}$$

where x_i^t is the 4×1 state vector containing the target Cartesian coordinates and absolute velocity projections on the local axes, w_i is the 2×1 vector of discrete normal zero-mean white noise accelerations with the known covariance matrix Q. Matrices of dynamics Φ (4×4) and generating noise Γ (4×2) correspond to Singer model [3–5].

The observer's motion is set to be known:

$$x_i^o = \Phi x_{i-1}^o + \Gamma u_i, \ i = \overline{1.N}, \tag{2}$$

where x_i^o is the 4×1 state vector containing the observer Cartesian coordinates and absolute velocity projections, u_i is the 2×1 known control input vector. Matrices Φ, Γ in models (1), (2) are identical, so we can easily proceed to the relative motion model by setting $x_i = x_i^t - x_i^o$, so that

$$x_i = \Phi x_{i-1} + \Gamma w_i - \Gamma u_i, \ i = \overline{1.N}. \tag{3}$$

The target bearing measurements at each time y_i are presented as

$$y_i = s(x_i) + v_i = arctg\left(\frac{x_i^{(1)}}{x_i^{(2)}}\right) + v_i, \tag{4}$$

where $x_i^{(1)}$, $x_i^{(2)}$ are the target relative coordinates set by the first and second elements of vector x_i , v_i is the normal zero-mean white noise measurement error with the known variance R. At initial time x_i is considered to have normal distribution with the known mathematical expectation $\bar{x}_0$ and covariance P_0 .

It is needed to estimate the vector $X_N = \left(x_0^T, x_1^T, \ldots, x_N^T\right)^T$ by the measurements accumulated by that time $Y_N = \left(y_1, y_2, \ldots, y_N\right)^T$ under the known inputs $\left(u_1^T, u_2^T, \ldots, u_N^T\right)^T$, considering that the state vector at initial time, measurement errors, and generating noise are noncorrelated with each other.

III. ITERATIVE BATCH LINEARIZED FILTER

Linearizing the measurements (4) yields

$$\tilde{y}_i = y_i - s\left(x_{i(l)}\right) + H_i^{x_{i(s)}} x_{i(l)} = H_i^{x_{i(s)}} x_i + v_i \qquad (5)$$

where $x_{i(l)}$ is the linearization point, $H_i^{x_{i(s)}} = \dfrac{ds\left(x_i\right)}{dx_i^T}\bigg|_{x_i = x_{i(s)}}$ is

the (1×4) Jacobian matrix calculated for this point. It can be shown that the mean-square optimal estimate of vector $\hat{X}_N$ and its covariance $\boldsymbol{P}_N$ can be found with the following algorithm [2]:

$$\hat{X}_N^{(j)} = \boldsymbol{P}_N^{(j)}\left(\boldsymbol{F}^T \boldsymbol{D} U_N + \left(\boldsymbol{H}^{(j-1)}\right)^T \boldsymbol{L}\tilde{Y}_N^{(j-1)}\right), \qquad (6)$$

$$\boldsymbol{P}_N^{(j)} = \left(\boldsymbol{F}^T \boldsymbol{D} \boldsymbol{F} + \left(\boldsymbol{H}^{(j-1)}\right)^T \boldsymbol{L}\left(\boldsymbol{H}^{(j-1)}\right)\right)^{-1}, \qquad (7)$$

where j is the iteration number, $\boldsymbol{F}$, $\boldsymbol{D}$, $\boldsymbol{H}$, and $\boldsymbol{L}$ are the known block matrices:

$$\boldsymbol{F} = \begin{vmatrix} E & 0 & \cdots & 0 \\ -\Phi & E & \ddots & \vdots \\ \ddots & \ddots & \ddots & 0 \\ 0 & 0 & -\Phi & E \end{vmatrix}; \boldsymbol{D} = \begin{vmatrix} P_0^{-1} & 0 & \cdots & 0 \\ 0 & \Delta Q^{-1} & 0 & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \Delta Q^{-1} \end{vmatrix};$$

$$\boldsymbol{L} = \begin{vmatrix} R^{-1} & 0 & \cdots & 0 \\ 0 & R^{-1} & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & R^{-1} \end{vmatrix}; \boldsymbol{H}^{(j-1)} = \begin{vmatrix} 0 & H_1^{\hat{x}_1^{j-1}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & H_N^{\hat{x}_N^{j-1}} \end{vmatrix};$$

the composite vector $U_N = \left(\bar{x}_0^T, (\Gamma u_1)^T, (\Gamma u_2)^T, \ldots, (\Gamma u_N)^T\right)^T$, vector $\tilde{Y}_N^{(j-1)} = \left(\tilde{y}_1^{(j-1)}, \tilde{y}_2^{(j-1)}, \ldots, \tilde{y}_N^{(j-1)}\right)^T$, where index $j{-}1$ indicates that the measurements were obtained using the left part of (5) and estimates $\hat{X}_N^{(j-1)}$ from the previous iteration. The block matrix $\boldsymbol{D}$ includes $\Delta Q = \Gamma Q \Gamma^T + E\alpha$, where $0 < \alpha \ll 1$. The fictitious noise $E\alpha$ is added to overcome the degeneracy of matrix $\Gamma Q \Gamma^T$. In the first iteration, a priori mathematical expectations of state vector X_N calculated using $\bar{x}_0$ and model (3) are used as the linearization point. The iterations are needed to update the linearization points. The criterion for stopping the iterations was calculated as

$$\Delta_N = \left(\hat{X}_N^{(j)} - \hat{X}_N^{(j-1)}\right)^T \left(\hat{X}_N^{(j)} - \hat{X}_N^{(j-1)}\right). \qquad (8)$$

Simulation has shown that the algorithm converges in 8–10 iterations.

The performance of nonrecursive iterative linearized batch smoother under the observer circulation by 360° is shown in Fig. 1. The estimated coordinates almost completely match the target trajectory. It can be also easily seen that the position uncertainty range is greatly reduced, i.e., a posteriori uncertainty ellipses are much smaller than a priori ones.

Fig. 1. Simulation of nonrecursive algorithm under the observer complete circulation

IV. RECURSIVE ITERATIVE BATCH LINEARIZED SMOOTHER

The estimates $\hat{X}_N$ of nonrecursive algorithm are smoothed, because they use the whole set of accumulated measurements. At the same time algorithm (6), (7) requires multiple inversions of the 4(N+1)×4(N+1) matrix in the right part of (7), which is computationally intensive. Note that identical estimates can be obtained with the recursive iterative batch linearized smoother tuned to the models (3), (5) [2, 6–9]. The distinctive feature of this smoother is that linearization points are not updated at each step, like in recursive extended Kalman smoother or iterative extended Kalman smoother (I-EKS) [6–9]. The linearization points are selected as in a nonrecursive algorithm: at the first step, a priori mathematical expectations of the state vector X_N are used, then the estimates from the previous steps are used. Therefore, the recursive iterative batch linearized smoother requires as many iterations for updating the linearization points as a nonrecursive algorithm, i.e., it processes the measurements in forward time and the estimates in reverse time for 8–10 times. However, this algorithm implemented in MATLAB needs up to 25 times shorter time than the nonrecursive algorithm considered above to obtain the estimates and their covariances. The agreement between the estimates obtained with the recursive iterative batch linearized smoother and nonrecursive iterative batch linearized filter in simulation conditions is shown in Fig. 2.

Non-updating the linearization point in the measurement process seems only to complicate the algorithm. Thus, I-EKS is a popular algorithm for nonlinear problems [2, 6-9], since it updates the linearization point at each step, thus providing a more accurate estimate at one step of measurement processing in forward time and estimates processing in reverse time. To compare the solutions of the considered problem, we have implemented the I-EKS tuned to the models (3), (5). The unconditional (real) and estimated RMS errors of the observer position for the nonrecursive and recursive algorithms were obtained with the Monte-Carlo method. RMS errors and the estimated Cramer-Rao lower bound for coordinates [10] in this problem are given in Figs. 3 and 4.

Fig. 2. Coordinate estimates for the nonrecursive iterative batch linearized filter and recursive iterative batch linearized smoother.

Fig. 3. RMS error for the horizontal coordinate.

Fig. 4. RMS error for the vertical coordinate.

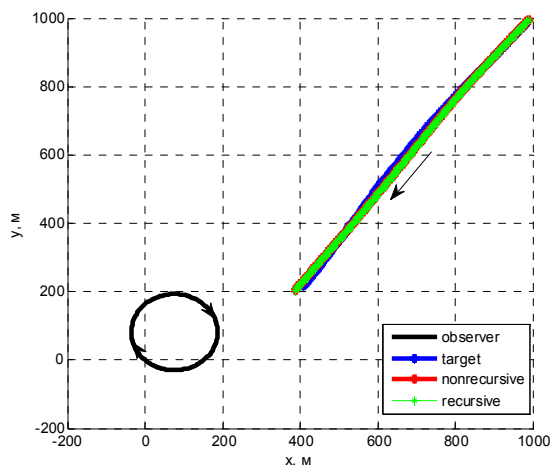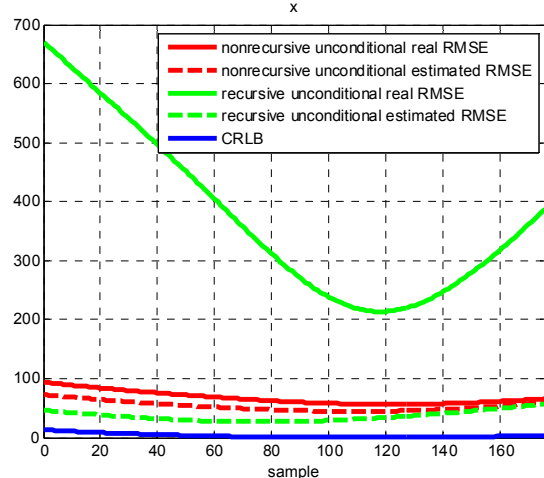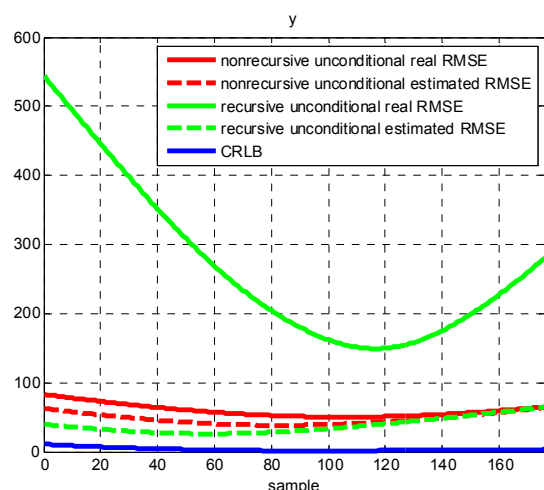The figures clearly show that recursive I-EKS is less accurate than the nonrecursive algorithm. At the same time, I-EKS estimated and real RMS errors do not coincide, i.e., this recursive algorithm, unlike the nonrecursive one, is not consistent [11]. This is due to the multi-extreme nature of a posteriori density. The recursive algorithm generates the Gaussian approximation of a posteriori probability density functions at each measurement, which leads to accumulation of linearization error [12-14]. The loss in accuracy is not compensated by the computational simplicity. It should be mentioned, however, that the Cramer-Rao lower bound cannot be achieved with the nonrecursive algorithm either.

V. CONCLUSIONS

The paper compares nonrecursive and recursive algorithms for estimating the target coordinates and motion parameters based on bearing measurements. It is shown that the iterative nonrecursive iterative batch linearized filter and recursive iterative batch linearized smoother provide the same estimates, the latter being less computationally intensive. At the same time, an I-EKS being recursive smoother with local iterations, where the linearization point is updated at each step, is much less accurate than the nonrecursive one due to accumulation of linearization errors. It is noted that according to the Monte Carlo testing, the nonrecursive iterative batch linearized filter and the recursive iterative batch linearized smoother are consistent, i.e., their unconditional real and estimated covariance matrices agree. As a result of the research, software for modeling the trajectories of an object and an observer, as well as software that implements the each of the considered algorithms was developed.

REFERENCES

[1] K. Jahan, M.M. Naga Lakshmi, P. Suresh, R.L.L. Praveenya, P. Swarjan, Extended Kalman Filter for Bearings-Only Tracking. Journal of Engineering Sciences, 2022, 13(7), 792–795.

[2] O.A. Stepanov, A.M. Isaev, A.V. Motorin, Yu.A. Litvinenko, and V.P. Zolotarevich, Recursive and Nonrecursive Algorithms Applied to Navigation Data Processing: Differences and Interrelation with Factor Graph Optimization Algorithms, 31st St. Petersburg International Conference on Integrated Navigation Systems, 2024.

[3] O.A. Stepanov, Primenenie teorii nelineinoi fil'tratsii v zadachakh obrabotki navigatsionnoi informatsii (Application of Nonlinear Filtering Theory to Navigation Data Processing), St. Petersburg: CSRI Elektropribor, 1998. ISBN 5-900780-45.

[4] R.A. Singer, Estimating Optimal Tracking Filter Performance for Manned Maneuvering Targets, Zarubezhnaya radioelektronika, no. 8, pp. 40–57, 1971.

[5] Y. Bar-Shalom, P.K. Willet, X. Tian, Tracking and Data Fusion: A Handbook of Algorithms. YBS Publishing, 2011.

[6] O.A. Stepanov, Osnovy teorii otsenivaniya s prilozheniyami k zadacham obrabotki navigatsionnoi informatsii (Fundamentals of the Estimation Theory with Applications to the Problems of Navigation Information Processing), Part 1, Vvedenie v teoriyu otsenivaniya (Introduction to the Estimation Theory), St. Petersburg: Concern CSRI Elektropribor, third edition.

[7] H.Van Trees, Detection, Estimation, and Modulation Theory, 1968.

[8] A.P. Sage, J.L. Melsa, Estimation Theory with Applications to Communications and Control, 1971.

[9] S. Sarkka and L. Svensson, Bayesian Filtering and Smoothing. Second Edition. Cambridge University Press, 2023.

[10] O.A. Stepanov, D.A. Koshaev, Applying Cramer-Rao Lower Bound to Nonlinear Estimation Problems, Teoriya i sistemy upravleniya. Izvestiya RAN, 1997, no. 2, pp. 65-72.

[11] O.A. Stepanov and A.M. Isaev, A Procedure of Comparative Analysis of Recursive Nonlinear Filtering Algorithms in Navigation Data Processing Based on Predictive Simulation, Gyroscopy and Navigation, 2023, vol. 14, no. 3.

[12] A.B. Toropov, Filtering Algorithms in Aiding the Marine Navigation System with Nonlinear Measurements, Cand. Sci. dissertation, St. Petersburg, 2013.

[13] V.A. Vasiliev, Yu.A. Litvinenko, O.A. Stepanov, and A.B. Toropov, Comparing Recursive and Nonrecursive Filtering Algorithms in Map-Aided Navigation, in Proceedings of XIV Vserossiiskaya multikonferentsiya po problemam upravleniya (4th Russian Conference on Control Problems), Rostov-on-Don, 2021, pp. 125-127.

[14] O.A. Stepanov, A.M. Isaev, Yu.A. Litvinenko, Comparing Recursive and Nonrecursive Designs of Linearized Algorithms as Applied to One Class of Nonlinear Estimation Problems, Proceedings of the XIV Vserrossiiskogo soveshchaniya po problemam upravleniya (14th Russian Meeting on Control Problems), Moscow, Trapeznikov Institute of Control Problems, 2024, in press.

Design of Visual-Inertial Odometry Algorithm for a Four-Legged Walking Robot with a Stereo Camera

D.A. Cherginets
Faculty of Control Systems and Robotics
ITMO University
St. Petersburg, Russia
cherginetsda@itmo.ru

A. A. Vedyakov
Faculty of Control Systems and Robotics
ITMO University
St. Petersburg, Russia
vedyakov@itmo.ru

Abstract— **This paper considers the problem of estimating the position of a mobile walking four-legged robot. Presented visual-inertial odometry algorithm based on factor-graph optimization that considers the leg kinematics of a walking robot and requires less computational resources. The proposed algorithm is a modification of the VILO algorithm. In the visual part, instead of a monocamera, a stereo camera model is used, which reduces the estimated state vector and eliminates the need for additional identification at the initialization stage. The algorithm performs complex processing of images captured by a stereo camera, measurements of linear accelerations and angular velocities of the robot, force sensors mounted on the foot, as well as angles, angular velocities of the robot's joints. The efficacy of the presented algorithm is demonstrated on data obtained from a real robot. Comparison with the measurements of the motion capture system gives that the odometry algorithm error is less than 2% of the path.**

Keywords—odometry, mobile robotics, position estimation, factor-graph

I. INTRODUCTION

In recent years, the development of mobile robotics applications has been particularly noticeable, which opens up a magnificent opportunity for autonomous tasks in industry. In order to navigate the robot, it is necessary to know its current position and orientation. Simultaneous localization and mapping algorithms are used for this purpose, which include odometry algorithms. In this paper, we propose a visual-inertial-legged odometry algorithm. In essence, it is about synthesizing an algorithm for complex processing of all available information. Nowadays, along with algorithms based on the Bayesian approach, especially as applied to the problems of odometry (algorithms of motion estimation relative to some initial position) for robotic systems, algorithms based on factor-graph optimization [1]. In comparison to approaches based on recurrent generalized Kalman filter, such algorithms have higher accuracy, but require significant computational resources. The most popular algorithms are VILOfDL [2], VILENS [3], WALK-VIO [4], STEP [5] and VILO [6], which are based on the MONO-VINS visual-inertial navigation system [7]. As demonstrated in [6], VILO has the lowest error compared to its counterparts. Consequently, it was selected as the foundation for our algorithm. However, MONO-VINS, upon which the VILO algorithm is based, treats the stereo camera as two mono cameras, which increases the size of the state

vector and, consequently, the computational complexity of the algorithm. It is noteworthy that in order to achieve high accuracy, an initialization process is required, which is not always feasible.

The aim of this work is to design an odometry algorithm for a mobile quadruped walking robot, which is reduced to an integrated processing of images from an on-board stereo camera, measurements from an inertial measurement unit (IMU), encoders in the joints, and force sensors placed on the feet to measure the contact force with the surface.

II. MAIN RESULT

Odometry algorithm based on factor graph optimization keeps track of the N robot state estimations $\hat{X}_k$ and M feature position estimations $\hat{\lambda}_k$.

Let us define a robot state estimate as

$$\hat{X} = \begin{bmatrix} \hat{X}_1 & \hat{X}_2 & \cdots & \hat{X}_N \end{bmatrix},$$
$$\hat{X}_k = \begin{bmatrix} \hat{p}_k & \hat{v}_k & \hat{q}_k & \hat{b}_{a,k} & \hat{b}_{\omega,k} & \theta_k \end{bmatrix}, \quad k \in [1, N],$$

where $\hat{p}_k$ is the robot position, $\hat{v}_k$ is the robot linear velocity, $\hat{q}_k$ is the robot orientation quaternion, $\hat{b}_{a,k}$ and $\hat{b}_{\omega,k}$ are IMU accelerometer and gyroscope biases, θ_k is the robot kinematic parameters.

Let us define a feature state estimate as
$$\hat{\Lambda} = \begin{bmatrix} \hat{p}_{\lambda,1} & \hat{p}_{\lambda,2} & \cdots & \hat{p}_{\lambda,M} \end{bmatrix},$$

where $\hat{p}_{\lambda,k}$ – feature λ_k position.

The measurements of the on-robot sensors form a set of measurements:

$$Z_t = \begin{bmatrix} a_{c,t} & \omega_{c,t} & \phi_{s,t} & \dot{\phi}_{s,t} & f_{s,t} \end{bmatrix},$$
$$\Lambda_t = \begin{bmatrix} v_{\lambda,1,t} & v_{\lambda,2,t} & \cdots \end{bmatrix},$$

where $a_{c,t}$ is the set of IMU linear acceleration measurements, $\omega_{c,t}$ is the set of IMU angular velocity measurements, $\phi_{s,t}$ and $\dot{\phi}_{s,t}$ are sets of joints angle and joints angular velocity, $f_{s,t}$ is the set of force measurements between foot and environment surface, $v_{\lambda,i,t}$ is the set of calculated position of found features.

We denote X as a set of all robot state estimates $\hat{X}$ and all feature state estimates $\hat{\Lambda}$. We also denote Z as the set of all measurements Z_t and Λ_t that coincide with the used states.

Assuming the measurement errors to be independent of each other and distributed according to the Gaussian law with zero mathematical expectation, the robot position estimation problem can be reduced to minimization of the criterion corresponding to the nonlinear least squares method

$$X^* = \min_X \left\{ \sum_k \| r_k(X,Z) \|^2_{\Sigma_k} \right\},$$

where $\| \cdot \|^2_{\Sigma_k}$ is Mahalanobis distance.

Defined by the covariance matrices of measurement errors of the used devices, for the considered problem this criterion can be represented in the following form:

$$
\begin{aligned}
X^* = \min_X \sum_k \Big(& \| r_{0,k}(X,Z) \|^2_{\Sigma_{0,k}} + \| r_{p\lambda,k}(X,Z) \|^2_{\Sigma_{p\lambda,k}} + \\
& + \| r_{p\lambda,0,k}(X,Z) \|^2_{\Sigma_{p\lambda,0,k}} + \| r_{pb,k}(X,Z) \|^2_{\Sigma_{pb,k}} + \\
& + \| r_{plo,k}(X,Z) \|^2_{\Sigma_{plo,k}} \Big),
\end{aligned}
$$

where $r_{0,k}(X,Z)$ is prior information residual function, $r_{p\lambda,k}(X,Z)$ and $r_{p\lambda,0,k}(X,Z)$ are visual measurement residual functions, $r_{pb,k}(X,Z)$ is IMU measurement residual function, $r_{plo,k}(X,Z)$ is leg odometry residual function, $\Sigma_{0,k}$, $\Sigma_{p\lambda,k}$, $\Sigma_{p\lambda,0,k}$, $\Sigma_{pb,k}$, $\Sigma_{plo,k}$ are uncertainty covariance matrices corresponding to their residual functions.

A. Prior Information Residual Function

The prior information factor describes the relationship between the estimated values and a fixed reference frame. The residual function is defined as an error between state estimate $\hat{X}_0$ and the prior value X_{init}:

$$
r_0(X,Z) = \begin{bmatrix} \Phi(T_0^{-1} T_{init}) \\ v_0 - v_{init} \\ b_{a,0} - b_{a,init} \\ b_{\omega,0} - b_{\omega,init} \end{bmatrix},
$$

where T is transformation matrix formed from the quaternion of orientation q, and position p, Φ is re-parameterization operator "lifting" [8].

The covariance matrix $\Sigma_{0,k}$ is defined in accordance with the uncertainties associated with the manner in which the prior information is obtained.

B. Visual Measurement Residual Functions

The stereo camera consists of two rigidly mounted cameras (Fig. 1). Typically, these cameras mounted along horizontal axis with the distance b, which called a baseline. The projections of the point $(x \quad y \quad z)$ in the left and right frames denoted as $(u_l \quad v)$ and $(u_r \quad v)$ which differ only along the u axis. The equation of transformation of features coordinates from left and right frames to space point is presented below:

$$
\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \dfrac{b u_{d,l}}{u_{d,l} - u_{d,r}} \\ \dfrac{b v_d}{u_{d,l} - u_{d,r}} \\ \dfrac{f b}{u_{d,l} - u_{d,r}} \end{bmatrix},
$$

where $u_{d,l}$ and $u_{d,r}$ are undistorted coordinates for the left and right frames respectively, f is focal length.

The projection function is defined as follows:

$$u_l = \frac{fx}{z} + c_x, \quad v_l = \frac{fy}{z} + c_y,$$

where $(c_x \quad c_y)$ is the optical center coordinate.

Fig. 1. Stereo camera model

The visual factor contains two residual functions:

- difference between the measured frame position of the feature $(u \quad v)$ and projection of its estimate:

$$r_{b\lambda}(X,Z) = \begin{bmatrix} u - \hat{u} \\ v - \hat{v} \end{bmatrix}.$$

- difference between the measured space position of the feature and its estimate:

$$r_{b\lambda,0}(X,Z) = p_\lambda - \hat{p}_\lambda.$$

C. IMU Measurement Residual Function

Since the frequency of adding state estimates corresponds to the camera frequency, which is much lower than that of the robot's proprioceptive sensors, IMU measurements must be integrated. To reduce the algorithm computational cost, we used a pre-integration algorithm [8] which allow us to get equations for iterative calculation using only IMU measurements a and ω, as well as their offsets b_a and b_ω:

$$p_{l,i+1} = p_{l,i} + v_{l,i}\delta t_i + \frac{1}{2}R(q_{l,i})(a_i - b_{a,k})\delta t_i^2,$$
$$v_{l,i+1} = v_{l,i} + R(q_{l,i})(a_i - b_{a,k})\delta t_i,$$
$$q_{l,i+1} = q_{l,i}\left(\frac{1}{2}\Omega(\omega_i - b_{\omega,k})\delta t_i\right),$$

where i corresponds to a specific IMU measurement in the interval $t \in [t_k, t_{k+1}]$, δt_i is a period between measurements i and $i+1$, $\Omega(\cdot)$ is a skew-symmetric matrix.

Iterative procedure of the covariance matrix $\Sigma_{pb,k}$ and the updating of p_l, v_l, q_l when the biases estimate b_a and b_ω change are described in [9].

The IMU measurement residual function compares the difference between two consecutive estimates and pre-integrated IMU measurements:

$$r_{pb}(X,Z) = \begin{bmatrix} R(\hat{q}_k)\left(\hat{p}_k - \hat{p}_{k-1} + \frac{1}{2}g\Delta t_k^2 - \hat{v}_{k-1}\Delta t_k\right) - p_l \\ R(\hat{q}_k)(\hat{v}_k - \hat{v}_{k-1} + g\Delta t_k) - v_l \\ 2\varpi(\hat{q}_{k-1}^{-1}\hat{q}_k q_l) \\ b_{a,k} - b_{a,k-1} \\ b_{a,k} - b_{a,k-1} \end{bmatrix},$$

where $\varpi(\cdot)$ extracts the vector part of a quaternion, Δt_k is a period between steps $k-1$ and k.

D. Leg Odometry Residual Function

The measurements of the robot's joint encoders are used to calculate the body velocity estimation $\hat{v}_{b,i}$ [10] assuming that the foot is in contact with the ground and does not slip. We use a pre-integration algorithm, to obtain an iterative calculation:

$$p_{lo,i+1} = p_{lo,i} + R(q_{l,i})\hat{v}_{b,i}\delta t_i.$$

The leg odometry residual function compares the difference between two consecutive estimates and pre-integrated velocity estimation:

$$r_{lo}(X,Z) = \begin{bmatrix} R(\hat{q}_{k-1})(\hat{p}_{1,k} - \hat{p}_{1,k-1}) - p_{lo,1,k} \\ R(\hat{q}_{k-1})(\hat{p}_{2,k} - \hat{p}_{2,k-1}) - p_{lo,2,k} \\ R(\hat{q}_{k-1})(\hat{p}_{3,k} - \hat{p}_{3,k-1}) - p_{lo,3,k} \\ R(\hat{q}_{k-1})(\hat{p}_{4,k} - \hat{p}_{4,k-1}) - p_{lo,4,k} \\ \hat{\theta}_k - \hat{\theta}_{k-1} \end{bmatrix},$$

where Δt_k is a period between steps $k-1$ and k.

III. EXPERIMENTAL RESULTS

The described algorithm was tested on recorded measurements of a Unitree A1 legged robot and a Realsense D435i stereo camera mounted on the front of the robot. We use measurements of linear acceleration and angular velocity of the IMU, angular position and angular velocity of the robot's joints, a force sensor to detect foot contact with the ground and stereo infrared images. The cameras were calibrated to obtain internal and external parameters. The IMU calibration was performed as described in [11]. IMU

noise parameters were calculated using the Allan variance method.

Motive OptiTrack motion capture system was used to obtain ground-truth position. It includes a cascade of eight cameras that track reflective markers and extract geometric relationships between markers to determine body position. It should be noted that a room equipped with a motion capture system does not satisfy the condition described in the manual for achieving the lowest position error. However, the use of a rotating platform with a known size allowed us to ascertain that the positioning error does not exceed 5 mm.

The motion capture system can track the robot in $3\text{m} \times 2\text{m}$ area. To unambiguously determine the robot's location, an array of six markers (in 3 rows of 2 markers each) was used, arranged in such a way that the intersection of the X and Y axes corresponds to the kinematic center of the robot.

Figure 2 shows the robot path captured by the motion capture system and the position estimation using the proposed odometry algorithm on the XY plane. As can be seen from the figure, the shape of the estimation matches the ground-truth position, but it should be noted that there was a drift along the Z-axis, which is not visible in the presented figure.

Figure 3 contains the relative errors of the position estimate together with a line corresponding to 2% of the locomotion distance. The relative error is obtained as the difference of the estimated and actual path, referred to the path length, for the corresponding specific point in time.

Fig. 2. Ground-truth path and it estimate on the XY

Fig. 3. Relative error of the position estimate

IV. CONCLUSION

A modification of the odometry algorithm for a legged robot has been presented, which consists of using a stereo camera model instead of several mono camera models. The algorithm provides complex processing of the stereo camera images, measurements of IMU, joint encoders, and force sensor on the robot's feet. The algorithm was validated in a laboratory environment using pre-recorded measurements. A motion capture system was used to obtain the ground-truth position of the robot. The results of the validation demonstrated that the relative error of position estimation from the travelled path did not exceed 2 percent. Further research will focus on a comparative analysis of the developed method with its analogues under different conditions.

REFERENCES

[1] O.A. Stepanov, A.M. Isaev, A.V. Motorin, Yu.A. Litvinenko, V.P. Zolotarevich, "Recursive and Nonrecursive Algorithms Applied to Navigation Data Processing: Differences and Interrelation with Factor Graph Optimization Algorithms" 31st Saint Petersburg International conference on Integrated Navigation Systems, 2024

[2] Dhedin V., Li H. et al. Visual-Inertial and Leg Odometry Fusion for Dynamic Locomotion // IEEE International Conference on Robotics and Automation (ICRA). 2023

[3] David W., Marco C. et al. Robust Legged Robot State Estimation Using Factor Graph Optimization // Institute of Electrical and Electronics Engineers (IEEE). 2019

[4] Hyunjun L., Byeong-Uk Y. et al. WALK-VIO: Walking-motion-Adaptive Leg Kinematic Constraint Visual-Inertial Odometry for Quadruped Robots // ArXiv. 2021.

[5] Yeeun K., Byeongho Y. et al. STEP: State Estimator for Legged Robots Using a Preintegrated foot Velocity Factor // ArXiv. 2022.

[6] Shuo Y., Zixin Z. et al. Cerberus: Low-Drift Visual-Inertial-Leg Odometry For Agile Locomotion // ArXiv. 2022.

[7] Qin T., Li P. et al. VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator // IEEE Transactions on Robotics. 2018, pp. 1004-1020.

[8] On-Manifold Preintegration for Real-Time Visual–Inertial Odometry / Christian Forster, Luca Carlone, Frank Dellaert, Davide Scaramuzza // IEEE Transactions on Robotics. — 2017. — feb. — Vol. 33, no. 1.

[9] Forster C., Carlone L. et al. IMU Preintegration on Manifold for Efficient Visual-Inertial Maximum-a-Posteriori Estimation. // 2015.

[10] P. Lin, H. Komsuoglu and D. E. Koditschek, "A leg configuration measurement system for full-body pose estimates in a hexapod robot," in IEEE Transactions on Robotics, vol. 21, no. 3, pp. 411-422, June 2005

[11] D. Tedaldi, A. Pretto and E. Menegatti, "A robust and easy to implement method for IMU calibration without external equipments," 2014 IEEE International Conference on Robotics and Automation (ICRA), Hong Kong, China, 2014

Solution of the Problem of AUV Positioning Relative to a Stationary Docking Station Using a Factor Graph Optimization Algorithm

V.G. Karaulov
Concern CSRI Elektropribor JSC, ITMO University
Saint Petersburg, Russia
vladkar0707@gmail.com

Y.A. Litvinenko
Concern CSRI Elektropribor JSC, ITMO University
Saint Petersburg, Russia
ya_litvinenko@mail.ru

O.A. Stepanov
Concern CSRI Elektropribor JSC, ITMO University
Saint Petersburg, Russia
soalax@mail.ru

A. M. Gruzlikov
Concern CSRI Elektropribor JSC
Saint Petersburg, Russia
agruzlikov@yandex.ru

Abstract –The paper is devoted to the problem of determining the coordinates and orientation angles of an autonomous underwater vehicle (AUV) relative to a stationary landing platform (SLP) using a high-frequency short-range hydroacoustic system. This paper proposes and investigates a nonrecursive algorithm for complex measurement processing implemented using optimization methods on graphs, also called factor-graph optimization (FGO) algorithms. The mathematical formulation of the problem within the framework of the Bayesian approach and the algorithm of its solution using FGO are described. The results of the estimation of root mean square (RMS) of estimation of AUV orientation coordinates and angles obtained using predictive simulation modeling are presented.

Keywords - autonomous underwater vehicle, Bayesian approach, nonrecursive algorithm, factor-graph optimization, stationary landing platform.

I. INTRODUCTION

The work discusses problems of determining the coordinates and orientation angles of an autonomous underwater vehicle (AUV) relative to a stationary landing platform (SLP) using a high-frequency short-range hydroacoustic system.

Similar problems were previously considered in [1-4]. In [1-3], a two-stage algorithm for solving the problem is proposed, in which, at the first stage, a machine learning method is used to reduce the domain of a priori uncertainty in the knowledge of initial coordinates and orientation angles, and at the second stage, a system of nonlinear algebraic equations of the pseudorange method is solved, taking into account the results obtained at the first stage. In [4], an algorithm for solving the problem under consideration is also given, but in contrast to [1-3], the initial approximation is proposed to be found by solving the equations without taking into account the measurement noise, and the further solution is realized on the basis of an algorithm synthesized in the framework of the Bayesian approach.

The authors of the above-mentioned works use measurements of pseudoranges between hydrophones on the AUV and radiating beacons on the SLP obtained only in one measurement cycle. In the present work, we propose to solve the problem taking into account the measurements accumulated during the entire motion of the AUV.

II. PROBLEM STATEMENT IN THE BAYESIAN APPROACH

It is assumed that there is a group of n=10 hydrophones operating in a single time scale and located in the head and stern parts of the AUV, also on the SLP. there are m =4 hydroacoustic beacons at the edges of the base (Fig. 1). The beacons emit a tone signal according to a known emission cyclogram. Based on the received signals, it is required to determine the relative coordinates and orientation of the AUV. It is necessary to take into account that the time scales on the AUV and the SLP do not coincide.

The problem under consideration can be formulated within the framework of the Bayesian approach as a recursive filtering problem [5, 6]. For this purpose, let us introduce an eight-dimensional state vector $x_k = \left(x_{0,k}, y_{0,k}, z_{0,k}, K_k, \psi_k, \theta_k, \delta_k, \Delta c_k \right)^T$, where k is the index of a discrete moment of time; $\left(x_{0,k}, y_{0,k}, z_{0,k} \right)$ – coordinates of the AUV center relative to the SLP (the control point of the AUV); $\left(K_k, \psi_k, \theta_k \right)$ – angles of heading, roll, and trim of the AUV relative to the SLP; δ_k – error in pseudorange measurements due to the divergence of time scales; Δc_k – error in knowledge of the speed of sound propagation in water.
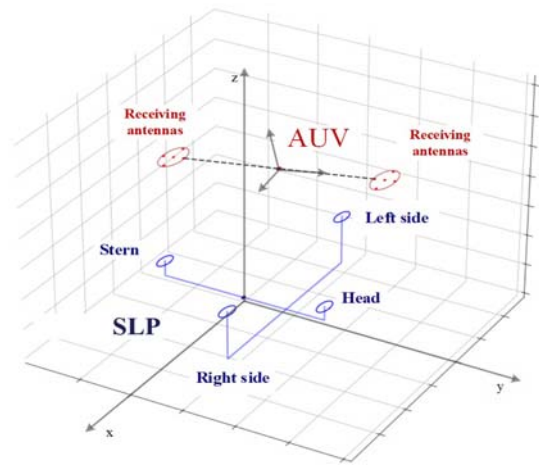

Fig. 1.   Layout of AUV hydroacoustic system and SLP.

In order to formulate the filtering problem, we introduce a linear equation describing the dynamics of variation of the state vector x_k in the form:

$$x_k = \Phi x_{k-1} + \Gamma w_k + Bu_k, \qquad (1)$$

where $\Phi = E_{8 \times 8}$ – dynamics matrix; $\Gamma = E_{8 \times 8}$ – matrix of generating noise; $B = diag\left(\triangle t, \triangle t, \triangle t, 0,0,0,0\right)$ – matrix under input influences; $u_k = \left(\upsilon_{k-1,x}, \upsilon_{k-1,y}, \upsilon_{k-1,z}, 0,0,0,0\right)^T$ – vector of input signals; $\triangle t$ – time between k and k-1 moments of time; $\left(\upsilon_{k-1,x}, \upsilon_{k-1,y}, \upsilon_{k-1,z}\right)$ – AUV velocity for each of the coordinates, which, after recalculation, are output from the navigation system installed on the AUV; w_k – eight-dimensional zero-mean white Gaussian noise, independent of x_0, with known covariance matrix Q.

The goal of the problem to be solved is to obtain an estimate of the state vector x_k from M-dimensional ($M = m \times n$ m, n – number of beacons and number of hydrophones) nonlinear measurements of the pseudoranges between beacons and hydrophones, represented as:

$$y_k = h(x_k) + v_k, \qquad (2)$$

where

$$h(x_k) = \begin{bmatrix} D_{1,k}(x_k) & D_{2,k}(x_k) & D_{3,k}(x_k) & D_{4,k}(x_k) \end{bmatrix}^T;$$

$$D_{i,k}(x_k) = \begin{bmatrix} d_{1,k}^i(x_k) & d_{2,k}^i(x_k) & \dots & d_{j,k}^i(x_k) \end{bmatrix}^T;$$

$$T_k = \begin{bmatrix} T_{1,k} & T_{2,k} & T_{3,k} & T_{4,k} \end{bmatrix}^T; \quad T_{m,k} = \begin{bmatrix} t_{1,k}^m & t_{2,k}^m & \dots & t_{10,k}^m \end{bmatrix}^T;$$

$$d_{j,k}^i(x_k) = \sqrt{\left(x_{0,k} + \tilde{x}_{j,k} - x^i\right)^2 + \left(y_{0,k} + \tilde{y}_{j,k} - y^i\right)^2 + \left(z_{0,k} + \tilde{z}_{j,k} - z^i\right)}$$

$$+ \hat{c}_{0,k} \delta_k + \Delta c_k t_{j,k}^i.$$

In relation (2) $y_k = \hat{c}_{0,k} T_k$; T_k – M-dimensional vector of measured signal detection times at the hydrophone; $i = \overline{1,m}$ – beacon number; $j = \overline{1,n}$ – hydrophone number; $h(x_k)$ – nonlinear M-dimensional function; $d_{j,k}^i$ – true range values; v_k – M-dimensional centered white Gaussian noise independent of w_k and x_0, with known covariance matrix R; $\hat{c}_{0,k}$ – a priori estimate of the sound speed in water, $x^i, y^i, z^i, i = \overline{1,m}$ – known coordinates of beacons in the rectangular coordinate system (CS) associated with the SLP $x_j, y_j, z_j, j = \overline{1,n}$ – known coordinates of hydrophones in the rectangular CS associated with the AUV; $\tilde{x}_j, \tilde{y}_j, \tilde{z}_j, j = \overline{1,n}$ - coordinates of hydrophones after the transition from the CS associated with the AUV to the CS associated with the SLP; $A(K, \psi, \theta)$ – transition matrix [5] from the x, y, z axes to the $\tilde{x}, \tilde{y}, \tilde{z}$.

III. Nonrecursive estimation algorithm

The preliminary analysis has shown that the recurrent algorithm for solving the filtering problem of the generalized Kalman filter type diverges in some cases. To overcome this disadvantage, we move from recursive to nonrecursive formulation of the problem. For this purpose, we introduce a state vector $X_k = \left(x^T_1, \dots, x^T_k\right)^T$, which includes the values of x_i

for all time moments $i = \overline{1,k}$, and a similar vector of measurements $Y_k = \left(y^T_1, \dots, y^T_k\right)^T$.

The problem of nonrecursive estimation can be formulated in the framework of the Bayesian approach as the problem of obtaining an estimate that maximizes the posterior probability density function (MAP) [6]:

$$\hat{X}_k = \arg\max_{X_k} p\left(X_k | Y_k\right) = \arg\max_{X_k} p\left(Y_k | X_k\right) p(X_k) =$$

$$= \arg\max_{\mathbf{X}} p_0\left(x_0\right) \prod_{k=1}^{N} p_k\left(y_k \mid x_k\right) \prod_{k=1}^{N} p_k\left(x_k \mid x_{k-1}\right), \qquad (3)$$

where $p_0\left(x_0\right)$ – is the a priori probability density function (pdf) of the state vector x_0, $p_k\left(y_k | x_k\right)$ is the likelihood function, $p_k\left(x_k | x_{k-1}\right)$ – is the transient pdf. of the state vector x, N is the number of moments in time.

Taking into account the introduced assumption about the Gaussian character of random variables, we can write:

$$p_k\left(y_k \mid x_k\right) \propto \exp\left(-\frac{1}{2}\|y_k - h(x_k)\|_{R_k}^2\right); \qquad (4)$$

$$p_k\left(x_k \mid x_{k-1}\right) \propto \exp\left(-\frac{1}{2}\|x_k - \Phi x_{k-1} - Bu_k\|_{Q_k}^2\right). \qquad (5)$$

The a priori pdf of the state vector is also assumed to be Gaussian, i.e.

$$p_0\left(x_0\right) \propto \exp\left(-\frac{1}{2}\|x_0 - \mu\|_P^2\right), \qquad (6)$$

where P, μ are the a priori covariance matrix and mathematical expectation for the AUV state vector at the initial time instant.

Substituting (4), (5) and (6) into (3) and taking the logarithm, we have:

$$\hat{X}_k = \arg\min_{X_k}\left[\|x_0 - \mu\|_P^2 + \sum_{k=1}^{N}\begin{bmatrix}\|y_k - h(x_k)\|_{R_k}^2 \\ + \|x_k - \Phi x_{k-1} - Bu_k\|_{Q_k}^2\end{bmatrix}\right] \qquad (7)$$

or

$$\hat{X}_k = \arg\min_{X_k} J\left(X_k\right), \qquad (8)$$

where

$$J(X_k) = \begin{bmatrix} \mu^T P \mu + \\ \sum_{k=1}^{N}\begin{bmatrix}\left(y_k - h(x_k)\right)^T R_k^{-1}\left(y_k - h(x_k)\right) \\ + \left(x_k - \Phi x_{k-1} - Bu_k\right)^T\left(\Gamma^T Q_k \Gamma\right)^{-1}\left(x_k - \Phi x_{k-1} - Bu_k\right)\end{bmatrix}\end{bmatrix}. \qquad (9)$$

It is easy to see that criterion (7) coincides with the criterion that is minimized in FGO algorithms [7-9, 15]. According to the terminology traditionally used in FGO, we note that $f_k^{pr} = y_k - h(x_k)$ can be treated as a pseudorange factor, $f_k^{mm} = x_k - \Phi x_{k-1} - Bu_k$ as a motion model factor, and $f_0^{prior} = x_0 - \mu$ as a factor taking into account a priori knowledge about the state vector.

Then we can write criterion (7) in the notations adopted in FGO algorithms:

$$\hat{X}_k = \arg\min_{X_k} \left[\left\| f_0^{prior} \right\|_P^2 + \sum_{k=1}^{N} \left[\left\| f_k^{pr} \right\|_{R_k}^2 + \left\| f_k^{mm} \right\|_{Q_k}^2 \right] \right]. \quad (10)$$

The solution of the problem consists in finding estimates of X_k by solving problem (10).

IV. RESULTS OF APPROBATION ON MODEL DATA

Taking into account the equivalence of criteria (3) and (10) to solve the problem of nonrecursive Bayesian estimation we can use the algorithms used in FGO, such as incremental smoothing and mapping algorithm (ISAM2) [10], dogleg algorithm [11] and Levenberg-Marquardt algorithm [12, 13], and we compare it with the two-stage algorithm [1-3] developed by the authors earlier, the one that processes one-step measurements.

In order to estimate the expected accuracy achieved by the above algorithms and compare it with the accuracy of the previously developed two-stage algorithm, we will perform predictive simulation modeling of the problem of AUV landing on an SLP according to the methodology described in [14].

Let us assume that an AUV is moving vertically downward with a velocity of 14 cm/s (velocity at x, y coordinates is 0) from the initial point with coordinates (0, 0, 40) m; roll and trim angles do not change and are equal to zero; heading angle does not change and is equal to 90 degrees.

We also consider that the RMS of the a priori knowledge of the state vector x_0 at the initial moment of time $x_0 = \left(x_{0,0}, y_{0,0}, z_{0,0}, K_0, \psi_0, \theta_0, \delta_0, \Delta c_0 \right)^T$ for coordinates - 5 m, for angles - 5 degrees; for pseudorange due to inconsistency of time scales - 5 cm; for the error of knowledge of the speed of sound - 1 m/s.

RMS for generating noises we take: for coordinates - 0.25 m, for angles - 1 degree, for pseudorange due to inconsistency of time scales - 1 cm; for inaccuracy of knowledge of sound velocity - 1 cm/s.

Time $\triangle t$ between measurements 0.5 s.

RMS for pseudorange measurements is 15 cm.

In the predictive modeling, a total of 100 realizations with 200 measurements each are obtained.

Based on the simulation results, four algorithms (two-stage algorithm, LM (Levenberg-Marquardt algorithm), DL (dogleg algorithm), and ISAM2) were used to calculate the RMS estimates of the coordinates and orientation angles of the AUVs (Fig. 2).

The open source library GTSAM (Georgia Tech Smoothing and Mapping) [16] was used to solve the nonrecursive estimation problem using the above algorithms.

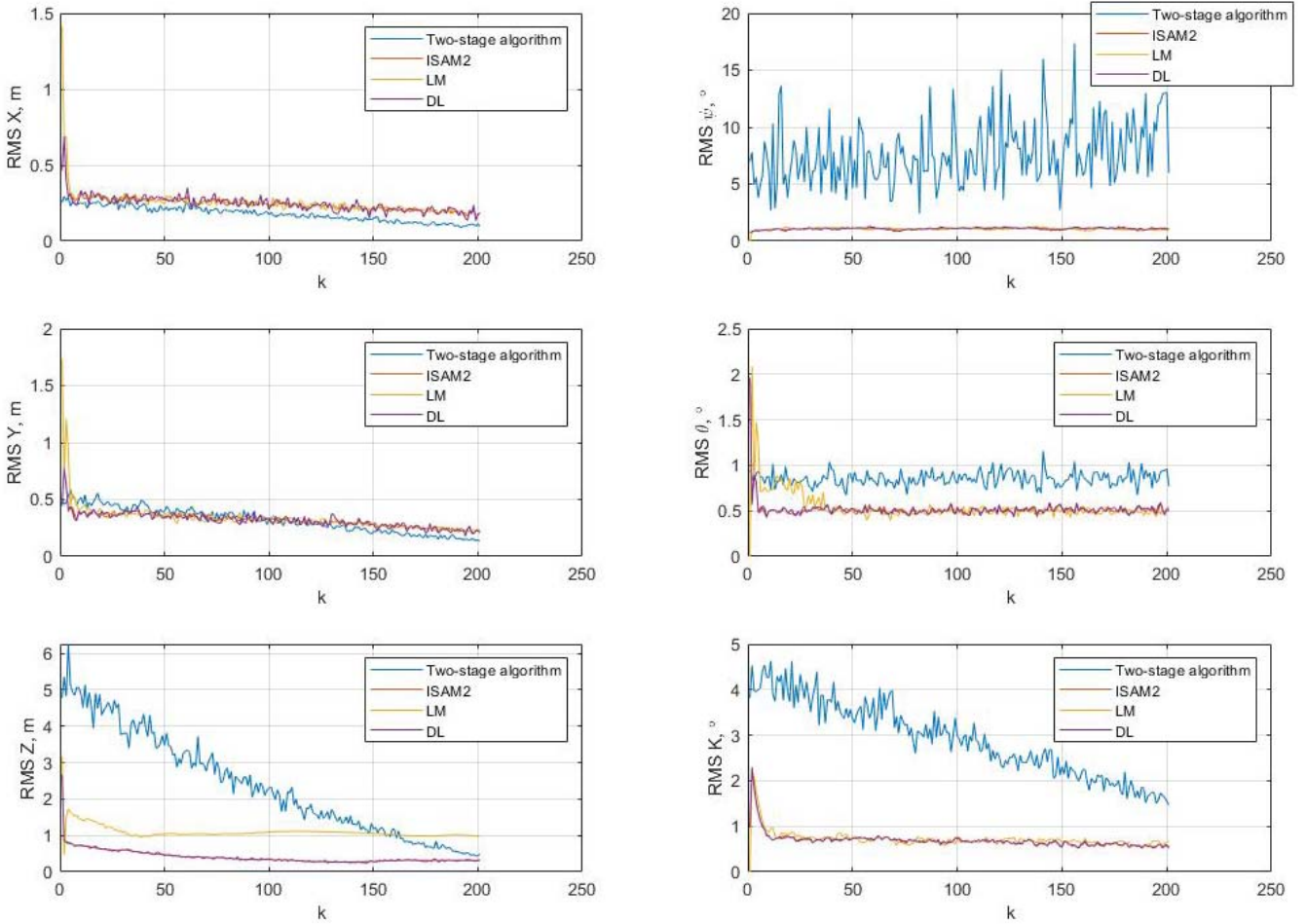Table 1 summarizes the results in terms of accuracy and computation time for the four algorithms.

Fig. 2. Dependence of the RMS of the AUV coordinates and orientation angles on the moments of time k.

TABLE 1. RMS FOR COORDINATES AND ORIENTATION ANGLES FOR THE FOUR ALGORITHMS

	Two-stage algorithm	LM	DL	ISAM2
RMS X, m	0.09	0.19	0.18	0.17
RMS Y, m	0.13	0.21	0.23	0.22
RMS Z, m	0.49	0.97	0.31	0.29
RMS K, deg	1.47	0.67	0.61	0.56
RMS ψ, deg	5.98	1.06	1.07	1.09
RMS θ, deg	0.77	0.55	0.52	0.48
Used time, s	28.23	785.58	903.21	20.25

The validity of the algorithms was not analyzed at this stage. From the results obtained, it follows that the nonrecursive algorithms have higher accuracy compared to the algorithm for processing one-step measurements, but they are computationally expensive. It is also worth noting that the Levenberg-Marquardt algorithm loses in accuracy to the dogleg algorithm. This may be due to the fact that the Levenberg-Marquardt algorithm uses a linear approximation at each iterative step, while the dogleg algorithm uses a piecewise linear approximation, which leads to more efficient convergence. The best result is shown by the ISAM2 algorithm, it uses the dogleg algorithm in terms of graph optimization, but due to the incremental approach it allows us not to recalculate the whole graph when new measurements arrive, i.e. it is not necessary to perform the Jacobian decomposition again for the whole set of accumulated measurements. Thus, the computational load is optimized.

V. CONCLUSION

The paper considers the problem of positioning and determining the AUV orientation angles relative to the SLP. The formulation of the problem to be solved within the framework of the Bayesian approach is given for both recursive and non-recursive schemes of measurement processing. To solve the problem within the framework of the nonrecursive scheme, the algorithms used in FGO problems are used.

The results of RMS estimation of the AUV coordinates and orientation angles with the use of three algorithms (ISAM2, dogleg and Levenberg-Marquardt) are presented. The obtained estimates are compared with the algorithm for processing only one-step measurements.

It is shown that by solving the problem using all the accumulated data, the accuracy is significantly improved, but the computational complexity also increases, which makes it difficult to solve the problem in real time. However, using the ISAM2 algorithm for factor-graph optimization allows us to overcome this problem and work in real time.

In the future it is supposed to analyze the consistency of the algorithms and show the possibility of reducing the amount of computations by applying recursive iterative smoothing algorithms used in Bayesian estimation problems.

REFERENCES

[1] A.M. Gruzlikov, 'Short and Ultra-Short Baseline Navigation of the AUV for Bringing It to the Bottom Docking Device', in 2022, 29th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), 2022, pp. 1–3.

[2] A.M. Gruzlikov, V.G. Karaulov, D.A. Mukhin, N.A. Shalaev "The results of testing the algorithm for positioning and determining the orientation of the underwater vehicle based on data from hydroacoustic beacons" // Izvestia SFedU. Tekhnicheskie nauki, . 2023, no. Pp. 265-275. [In Russian]

[3] V.G. Karaulov, A.M. Gruzlikov, V.P. Zolotarevich 'Recursive Algorithm for AUV Positing and Determining its Orientation', 30th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS). 2023. P. 1-4.

[4] Koshaev, D.A. AUV Relative Position and Attitude Determination Using Acoustic Beacons. Gyroscopy Navig. 13, 262–275 (2022). https://doi.org/10.1134/S2075108722040083.

[5] G. I. Emeliantsev, A. P. Stepanov "Integrated inertial-satellite orientation and navigation systems" St. Petersburg: CSRI Elekropribor, 2016. [In Russian]

[6] O. A. Stepanov "Application of the theory of nonlinear filtering in problems of navigation information processing" St. Petersburg: CSRI Elekropribor, 1998. 370 P. [In Russian]

[7] H.-A. Loeliger "An Introduction to Factor Graphs', IEEE Signal Processing Magazine. 2004. Vol. 21, No. 1. P. 28-41.

[8] F. Dellaert, M. Kaess 'Factor Graphs for Robor Perceprion', Foundations and Trends in Roborics, vol. 6, no. 1-2, pp. 1-139, 2017.

[9] W. Wen, T. Pfeifer, X. Bai, L-T. Hsu "Factor graph optimization for GNSS/INS integration: A comparison with the extended Kalman filter", NAVIGATION. 2021. Vol. 68. P. 315-331.

[10] M. Kaess, H. Johannsson, R. Roberts, V. Ila, J.J. Leonard, and F. Dellaert "iSAM2: Incremental Smoothing and Mapping Using the Bayes Tree", Intl. J. of Robotics Research (IJRR), vol. 31, Feb. 2012, pp. 217-236

[11] M.J.D. Powell "A hybrid method for nonlinear equations". In Robinowitz, P. (ed.). Numerical Methods for Nonlinear Algebraic Equations. London: Gordon and Breach Science. 1970, pp. 87–144.

[12] K. Levenberg, "A method for the solution of certain non-linear problems in least squares", Quarterly of Appl. Math., 2 (1944), 164–168.

[13] D.W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters", J. SIAM, 11 (1963), 431–441.

[14] Stepanov, O.A., Isaev, A.M. A Procedure of Comparative Analysis of Recursive Nonlinear Filtering Algorithms in Navigation Data Processing Based on Predictive Simulation. Gyroscopy Navig. 14, 213–224 (2023). https://doi.org/10.1134/S2075108723030094.

[15] O.A. Stepanov, V.P. Zolotarevich, Y.A. Litvinenko, A.M. Isaev, A.V. Motorin "Recursive and nonrecursive algorithms in navigation information processing tasks. Differences and interrelation with graph optimization algorithms." // 30th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS). 2024.

[16] F. Dellaert and GTSAM Contributors, "Georgia Tech Smoothing and Mapping library", https://github.com/borglab/gtsam, 2022.

Biquaternion Quasi-Optimal Analytical Solution for the Problem of Time-Optimal Programmed Control of Spatial Motion of Spacecraft

Yu.N. Chelnokov

Institute of Precision Mechanics and Control, Russian Academy of Sciences

Saratov, Russia

ChelnokovYuN@gmail.com

A.V. Molodenkov

Institute of Precision Mechanics and Control, Russian Academy of Sciences

Saratov, Russia

molalexei@yandex.ru

M.Yu. Loginov

Institute of Precision Mechanics and Control, Russian Academy of Sciences

Saratov, Russia

mike.loginov@gmail.com

Abstract — **This report describes the analytical solution for a problem of constructing the quasi-optimal time-optimal programmed control for spatial motion of spacecraft of an arbitrary dynamic configuration in the inertial coordinate frame using dual Euler (Rodrigues–Hamilton) parameters and Clifford's biquaternions (dual quaternions). A new model of spatial motion of spacecraft is used to solve the problem. This new model includes differential equations for dual Euler parameters and dual velocities of spacecraft, and dynamic algebraic relations for the vectors of control force and control moment. The boundary conditions for angular and linear positions of spacecraft in an inertial space and for its absolute angular and linear velocities are arbitrary, dual controls are limited in absolute value. It is necessary to derive the quasi-optimal time-optimal coupled control for angular and orbital motion of spacecraft. The solution of the problem is derived in the class of generalized dual coning motions using the method of solving inverse dynamic problems, Kotelnikov-Study transference and the principle of maximum.**

Keywords: spacecraft, programmed quasi-optimal control, spatial motion, dual Euler (Rodrigues–Hamilton) parameters, Clifford's biquaternion, angular (rotational) and translational (orbital) motion.

I. INTRODUCTION

It is known that the general spatial displacement of a free rigid body is equivalent to its screw displacement (Chasles' theorem: any displacement of a free rigid body can be produced by one screw displacement along some axis that is called the axis of screw finite displacement). Thus, motion of a free rigid body (spacecraft, in our case) is a continuous sequence of instantaneous screw displacements. The orientation of the axis of instantaneous screw displacement coincides with the orientation of the axis of instantaneous rotation of a rigid body. Rotation of a body around this axis and translational motion along the axis form (by means of the Clifford's complexity) the so-called dual angle of rotation: the complex composition of rotational and translational motion of a body.

This report describes the derivation of the quasi-optimal analytical solution for a problem of time-optimal programmed control for spatial motion of spacecraft of an arbitrary dynamic configuration in nonlinear dynamic formulation in the inertial coordinate frame using the dual Euler (Rodrigues–Hamilton) parameters and Clifford's biquaternions (dual quaternions) given the arbitrary boundary conditions. The solution is derived in the class of generalized dual coning motions using the method of solving the inverse dynamic problems, Kotelnikov-Study transference and the principle of maximum. A new biquaternion model of spatial motion of spacecraft is used to solve the problem. This model is proposed in works [1–3] that also describe other new biquaternion models of spatial motion of a rigid body and applications of these models for solving the problems of controlling the spatial motion of a rigid body using the feedback principle. Our solution is a biquaternion equivalent of a quaternion solution to a problem of quasi-optimal time-optimal programmed control of rotational motion of spacecraft derived in [4].

II. MATHEMATICAL MODEL OF MOTION OF SPACECRAFT AND THE FORMULATION OF A PROGRAMMED CONTROL PROBLEM

The following notation is used: $\mathbf{r}$ and $\mathbf{v}$ are the radius vector and velocity vector of the center of mass of spacecraft in the inertial coordinate frame ξ; $\boldsymbol{\lambda}$ is the quaternion of orientation of spacecraft in this coordinate frame that consists of Rodrigues-Hamilton (Euler) parameters λ_j ($j = 0, 1, 2, 3$); $\boldsymbol{\omega}$ and $\boldsymbol{\varepsilon}$ are vectors of absolute angular velocity and absolute angular acceleration of spacecraft; $\mathbf{F}_c$ and $\mathbf{M}_c$ are vectors of control force and control moment that are applied to spacecraft; $\mathbf{F} = \mathbf{F}(t, \mathbf{r}, \mathbf{v})$ and $\mathbf{M} = \mathbf{M}(t, \boldsymbol{\lambda}, \boldsymbol{\omega})$ are the resultant vector of other external forces that are applied to spacecraft (forces of gravity, resistance to motion and other forces of interaction between spacecraft and environment) and resultant moment of these forces calculated with respect to the center of mass of spacecraft, considered to be the known functions of time t and variables $\mathbf{r}$, $\mathbf{v}$ and $\boldsymbol{\lambda}$, $\boldsymbol{\omega}$.

The differential equations of motion of spacecraft regarded as a rigid body are as follows:

$$m\left[\dot{\mathbf{v}}_x + \mathbf{K}(\boldsymbol{\omega}_x)\,\mathbf{v}_x\right] = \mathbf{F}_x(t, \mathbf{r}_x, \mathbf{v}_x) + \mathbf{F}_{cx},$$

$$\dot{\mathbf{r}}_x + \mathbf{K}(\boldsymbol{\omega}_x)\mathbf{r}_x = \mathbf{v}_x; \qquad (2.1)$$

$$\dot{\boldsymbol{\omega}}_x = \boldsymbol{\varepsilon}_x = \mathbf{J}^{-1}\left[\mathbf{M}_x(t, \boldsymbol{\lambda}, \boldsymbol{\omega}_x) - \mathbf{K}(\boldsymbol{\omega}_x)\mathbf{J}\boldsymbol{\omega}_x + \mathbf{M}_{cx}\right], \quad (2.2)$$

$$2\dot{\boldsymbol{\lambda}} = \boldsymbol{\lambda} \circ \boldsymbol{\omega}_x; \qquad (2.3)$$

$$\mathbf{r}_x = x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}, \quad \dot{\mathbf{r}}_x = \dot{x}_1\mathbf{i} + \dot{x}_2\mathbf{j} + \dot{x}_3\mathbf{k},$$

$$\mathbf{v}_x = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}, \quad \dot{\mathbf{v}}_x = \dot{v}_1\mathbf{i} + \dot{v}_2\mathbf{j} + \dot{v}_3\mathbf{k};$$

$$\boldsymbol{\lambda} = \lambda_0 + \boldsymbol{\lambda}_v = \lambda_0 + \lambda_1\mathbf{i} + \lambda_2\mathbf{j} + \lambda_3\mathbf{k},$$

$$\dot{\boldsymbol{\lambda}} = \dot{\lambda}_0 + \dot{\boldsymbol{\lambda}}_v = \dot{\lambda}_0 + \dot{\lambda}_1\mathbf{i} + \dot{\lambda}_2\mathbf{j} + \dot{\lambda}_3\mathbf{k};$$

$$\boldsymbol{\omega}_x = \omega_1\mathbf{i} + \omega_2\mathbf{j} + \omega_3\mathbf{k}, \quad \dot{\boldsymbol{\omega}}_x = \dot{\omega}_1\mathbf{i} + \dot{\omega}_2\mathbf{j} + \dot{\omega}_3\mathbf{k},$$

$$\boldsymbol{\varepsilon}_x = \varepsilon_1\mathbf{i} + \varepsilon_2\mathbf{j} + \varepsilon_3\mathbf{k};$$

$$\mathbf{J} = \begin{pmatrix} J_{11} & -J_{12} & -J_{13} \\ -J_{21} & J_{22} & -J_{23} \\ -J_{31} & -J_{32} & J_{33} \end{pmatrix}, \ \mathbf{K}(\boldsymbol{\omega}_x) = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}.$$

In equations (2.1) and (2.2) $\mathbf{r}_x$, $\mathbf{v}_x$, $\boldsymbol{\omega}_x$, $\boldsymbol{\varepsilon}_x$, $\mathbf{F}_x$, $\mathbf{M}_x$, $\mathbf{F}_{cx}$, $\mathbf{M}_{cx}$ are 3x1 column vectors or, later on, quaternions with zero scalar parts constructed from the projections x_k, v_k, ω_k, ε_k, F_k, M_k, F_{ck}, M_{ck} ($k = 1, 2, 3$) of vectors $\mathbf{r}$, $\mathbf{v}$, $\boldsymbol{\omega}$, $\boldsymbol{\varepsilon}$, $\mathbf{F}$, $\mathbf{M}$, $\mathbf{F}_c$, $\mathbf{M}_c$ on the axes of the body-fixed coordinate frame X; m is the mass of a body, $\mathbf{J}$ is the constant matrix of inertia of a rigid body; $\mathbf{K}(\boldsymbol{\omega}_x)$ is the skew-symmetric matrix of angular velocities of a body that corresponds to vector $\boldsymbol{\omega}$; $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ are Hamilton's vector imaginary units; $\mathbf{a}_y$ is the projection of vector $\mathbf{a}$ on the coordinate basis Y ($Y = \xi$, X) defined as quaternion $\mathbf{a}_y = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$, the components of which are projections a_k of vector $\mathbf{a}$ on coordinate basis Y; over dot denotes the derivative with respect to time t, symbol "$\circ$" denotes the quaternion multiplication.

The first matrix equation (2.1) and matrix equation (2.2) are dynamic, and the second matrix equation (2.1) and quaternion equation (2.3) are kinematic equations of spatial motion of spacecraft that is the composition of translational (trajectory) and angular (rotational) motion. They form a system of nonlinear nonstationary differential equations of 13th order in relation to variables x_k, v_k and λ_j, ω_k.

Let us formulate the following problem: to construct the controls

$$\mathbf{F}_{cx} = \mathbf{F}_{cx}(t), \quad \mathbf{M}_{cx} = \mathbf{M}_{cx}(t), \qquad (2.4)$$

that provide the programmed transition of spacecraft from its arbitrarily specified initial state

$$\mathbf{r}_x = \mathbf{r}_x(0), \quad \mathbf{v}_x = \mathbf{v}_x(0), \quad \boldsymbol{\lambda} = \boldsymbol{\lambda}(0), \quad \boldsymbol{\omega}_x = \boldsymbol{\omega}_x(0) \qquad (2.5)$$

to its specified final state

$$\mathbf{r}_x = \mathbf{r}_x(t_1), \quad \mathbf{v}_x = \mathbf{v}_x(t_1), \quad \boldsymbol{\lambda} = \boldsymbol{\lambda}(t_1), \quad \boldsymbol{\omega}_x = \boldsymbol{\omega}_x(t_1) \qquad (2.6)$$

in a time-optimal way in the class of optimal generalized dual coning motions.

The formulated problem is solved using the concept of solving the inversed dynamic problems taking into account the acting gravitational forces $\mathbf{F}_{xgr}$ and discarding the acting disturbing forces and moments. The laws of forming the control force and control moments in accordance with this concept are derived from equations (2.1), (2.2) and are as follows:

$$\mathbf{F}_{cx} = m\left[\dot{\mathbf{v}}_x + \mathbf{K}(\boldsymbol{\omega}_x)\mathbf{v}_x\right] - \mathbf{F}_{xgr}(t, \mathbf{r}_x), \qquad (2.7)$$

$$\mathbf{M}_{cx} = \mathbf{J}\boldsymbol{\varepsilon}_x + \mathbf{K}(\boldsymbol{\omega}_x)\mathbf{J}\boldsymbol{\omega}_x. \qquad (2.8)$$

The required component $\dot{\mathbf{v}}_x = \mathbf{w}_x$ of absolute linear acceleration and required absolute angular acceleration $\boldsymbol{\varepsilon}_x$, which are involved in the first additive members of control laws (2.7) and (2.8), can be constructed from matrix equations (2.9) and quaternion equations (2.10):

$$\dot{\mathbf{v}}_x = \mathbf{w}_x, \quad \dot{\mathbf{r}}_x + \mathbf{K}(\boldsymbol{\omega}_x)\mathbf{r}_x = \mathbf{v}_x; \qquad (2.9)$$

$$\dot{\boldsymbol{\omega}}_x = \boldsymbol{\varepsilon}_x, \quad 2\dot{\boldsymbol{\lambda}} = \boldsymbol{\lambda} \circ \boldsymbol{\omega}_x, \qquad (2.10)$$

that result from equations (2.1)–(2.3).

Parameters $\mathbf{w}_x$ and $\boldsymbol{\varepsilon}_x$, which are involved in these equations, are viewed from now on as the new controls.

Therefore, the problem of constructing the control force $\mathbf{F}_{cx}$ and the control moment $\mathbf{M}_{cx}$ in this formulation is rendered down to constructing the required component $\mathbf{w}_x$ of the absolute linear acceleration and the required absolute angular acceleration $\boldsymbol{\varepsilon}_x$, which act as controls in equations (2.9) and (2.10). The problem of constructing the controls $\mathbf{w}_x$ and $\boldsymbol{\varepsilon}_x$ is common to all moving object that are regarded as rigid bodies, because equations (2.9) and (2.10) hold true for any moving object of such kind. The specific parameters of an object (its mass-inertia characteristics, acting gravitational forces) are taken into account during construction of control force $\mathbf{F}_{cx}$ and control moment $\mathbf{M}_{cx}$ from final relation (2.7) and (2.8).

Let us introduce the kinematic screw $\mathbf{U}$ of the spacecraft, the mapping $\mathbf{U}_x$ of which on the body-fixed basis X is defined by biquaternion

$$\mathbf{U}_x = U_1\mathbf{i} + U_2\mathbf{j} + U_3\mathbf{k} = \boldsymbol{\omega}_x + s\mathbf{v}_x, \quad U_k = \omega_k + sv_k, (2.11)$$

where s is the Clifford complexity with property $s^2 = 0$; $U_k = \omega_k + sv_k$, $k = 1,2,3$ are dual orthogonal projections of kinematic screw $\mathbf{U}$ on the basis X.

Then the vector-matrix differential equations (2.9) and quaternion differential equations (2.10) can be replaced with the two following biquaternion differential equations:

$$\dot{\mathbf{U}}_x = \boldsymbol{\varepsilon}_x + s\mathbf{w}_x = \mathbf{H}_x, \qquad (2.12)$$

$$2\dot{\boldsymbol{\Lambda}} = \boldsymbol{\Lambda} \circ \mathbf{U}_x; \qquad (2.13)$$

$$\boldsymbol{\varepsilon}_x = \varepsilon_1\mathbf{i} + \varepsilon_2\mathbf{j} + \varepsilon_3\mathbf{k}, \quad \mathbf{w}_x = w_1\mathbf{i} + w_2\mathbf{j} + w_3\mathbf{k},$$

$$\mathbf{H}_x = H_1\mathbf{i} + H_2\mathbf{j} + H_3\mathbf{k} = \boldsymbol{\varepsilon}_x + s\mathbf{w}_x, \qquad (2.14)$$

$$H_k = \varepsilon_k + sw_k,$$

$$\boldsymbol{\Lambda} = \Lambda_0 + \boldsymbol{\Lambda}_v = \Lambda_0 + \Lambda_1\mathbf{i} + \Lambda_2\mathbf{j} + \Lambda_3\mathbf{k} = \boldsymbol{\lambda} + s\boldsymbol{\lambda}^0,$$

$$\Lambda_j = \lambda_j + s\lambda_j^0; \ \mathbf{U}_x = \boldsymbol{\omega}_x + s\mathbf{v}_x, \qquad (2.15)$$

in which the phase variables are biquaternion $\mathbf{U}_x$ (the mapping of the kinematic screw $\mathbf{U}$ of spacecraft on the body-fixed basis X) and biquaternion $\boldsymbol{\Lambda}$ of finite displacement of spacecraft in the inertial space, in which the main part (quaternion $\boldsymbol{\lambda}$) defines the orientation of spacecraft in the inertial coordinate frame, and the moment part (quaternion $\boldsymbol{\lambda}^0$) defines the positions of spacecraft in this coordinate frame (the Cartesian coordinates ξ_k ($k = 1, 2,$

3) of the spacecraft's center of mass in this coordinate frame), and the dual control is biquaternion $\mathbf{H}_x$, which is a dual composition of the required absolute angular acceleration $\mathbf{\varepsilon}_x$ and the required component $\mathbf{w}_x$ of the absolute linear acceleration of spacecraft.

The Cartesian coordinates ξ_k of the center of mass of a body (i.e. the projections of radius vector $\mathbf{r}$ of the center of mass of spacecraft on the axes of the inertial coordinate frame ξ) and the projections x_k of this vector on the axes of the body-fixed coordinate frame X can be found from the components of quaternions $\mathbf{\lambda}$ and $\mathbf{\lambda}^0$ using the following formulas (where the overline denotes conjugation):

$$\mathbf{r}_\xi = \xi_1\mathbf{i} + \xi_2\mathbf{j} + \xi_3\mathbf{k} = 2\mathbf{\lambda}^0 \circ \overline{\mathbf{\lambda}}, \quad \mathbf{r}_x = x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k} = \overline{\mathbf{\lambda}} \circ \mathbf{\lambda}^0.$$

We will define the problem of deriving the required component $\mathbf{w}_x$ of the absolute linear acceleration and the required absolute angular acceleration $\mathbf{\varepsilon}_x$, which are involved as controls in the equations (2.9) and (2.10), with the following biquaternion formulation: it is necessary to construct the modulus-limited biquaternion programmed control $\mathbf{H}_x = \mathbf{\varepsilon}_x + s\mathbf{w}_x$, that in a time-optimal way repositions the spacecraft, the motion of which is defined by equations (2.12) and (2.13), from its arbitrarily specified initial state

$$\mathbf{\Lambda}(0) = \mathbf{\lambda}(0) + s\mathbf{\lambda}^0(0), \quad \mathbf{U}_x(0) = \mathbf{\omega}_x(0) + s\mathbf{v}_x(0) \quad (2.16)$$

to its specified final state

$$\mathbf{\Lambda}(t_1) = \mathbf{\lambda}(t_1) + s\mathbf{\lambda}^0(t_1), \quad \mathbf{U}_x(t_1) = \mathbf{\omega}_x(t_1) + s\mathbf{v}_x(t_1) .(2.17)$$

The boundary conditions $\mathbf{\lambda}^0(0)$ and $\mathbf{\lambda}^0(t_1)$ are derived from the specified initial and final values of quaternions $\mathbf{\lambda}$ and $\mathbf{r}_\xi$ using the formula

$$\mathbf{\lambda}^0 = (1/2)\mathbf{r}_\xi \circ \mathbf{\lambda} =$$
$$= (1/2)\left(\xi_1\mathbf{i} + \xi_2\mathbf{j} + \xi_3\mathbf{k}\right) \circ (\lambda_0 + \lambda_1\mathbf{i} + \lambda_2\mathbf{j} + \lambda_3\mathbf{k}).$$

After solving the formulated problem, in the constructed control $\mathbf{H}_x = \mathbf{\varepsilon}_x + s\mathbf{w}_x$ it is necessary to separate the main part $\mathbf{\varepsilon}_x$ and the moment part $\mathbf{w}_x = \dot{\mathbf{v}}_x$. After this separation, the laws of constructing the control force $\mathbf{F}_{cx}$ and control moment $\mathbf{M}_{cx}$ are derived in accordance with the concept of solving the inverse dynamic problems using formulas (2.7) and (2.8).

III. SOLVING THE PROBLEM OF QUASI-OPTIMAL PROGRAMMED CONTROL

With the help of Kotelnikov-Study transference, which makes it possible to generalize quaternion formulas for angular motion control onto biquaternion formulas for general spatial motion control of a rigid body, the quasi-optimal time-optimal algorithm is derived for programmed control of spatial motion of spacecraft. This algorithm is constructed by generalizing the approximate analytical solution of the quaternion problem of rotating spacecraft in a time-optimal way under arbitrary boundary conditions for angular position and angular velocity [4], which, in its turn,

is based on the exact solution to the modified problem of turning a rigid body in an optimal way, which was earlier introduced by Ya.G. Sapunkov [5]. Let us describe this approach briefly.

In the case when a free rigid body performs spatial generalized screw coning motion, for which the kinematic screw $\mathbf{U}_x$ of a body has the form

$$\mathbf{U}_x = U_1\mathbf{i} + U_2\mathbf{j} + U_3\mathbf{k} = \mathbf{\omega}_x + s\mathbf{v}_x =$$
$$= \left(\dot{F}(t)\sin G(t)\right)\mathbf{i} + \left(\dot{F}(t)\cos G(t)\right)\mathbf{j} + \dot{G}(t)\mathbf{k}, \quad (3.1)$$

where $F(t), G(t), \dot{F}(t), \dot{G}(t)$ (the dual parameters of generalized screw coning motion) are arbitrary differentiable functions of time, the biquaternion kinematic equation (2.13) has an analytical solution (derived by the authors of this report)

$$\mathbf{\Lambda}(t) = \mathbf{\Lambda}_0 \circ \mathbf{exp}\{-(G(0)/2)\mathbf{k}\} \circ \mathbf{exp}\{-(F(0)/2)\mathbf{j}\} \circ$$
$$\circ \mathbf{exp}\{(F(t)/2)\mathbf{j}\} \circ \mathbf{exp}\{(G(t)/2)\mathbf{k}\} \quad (3.2)$$

that is the dual equivalent of a quaternion solution [4, 5]; "$\mathbf{exp}\{.\}$" denotes the biquaternion exponent.

The screw motion of a free rigid body (spacecraft) described by relations (3.1), (3.2) can be additionally generalized by introducing an arbitrary dual turn in the inertial coordinate frame through a constant dual angle around a given axis. This turn is defined by an arbitrary constant biquaternion $\mathbf{K} = \mathbf{\kappa} + s\mathbf{\kappa}^0$, $\|\mathbf{K}\| = 1$ (the right parts of formulas (3.1), (3.2) are extended by left and right multiplication by biquaternions $\overline{\mathbf{K}}$ and $\mathbf{K}$ respectively, where $\overline{\mathbf{K}} \circ \mathbf{K} = \mathbf{K} \circ \overline{\mathbf{K}} = 1$).

Formulas (3.1), (3.2) include all the known exact quaternion analytical solutions to the classical problem of optimal turn of a dynamically symmetric rigid body, when the angular velocity vector throughout whole time period is constant in direction or circumscribes a circular cone [6], and the dual equivalents of these solutions. The proposed structure of the kinematic screw (3.1), in the quaternion case − of the angular velocity vector of spacecraft, is consistent with the Poinsot concept (which states that any arbitrary angular motion of a rigid body around a fixed point can be viewed as a certain generalized coning motion of a rigid body) generalized onto the arbitrary spatial motion of a free rigid body in the class of screw coning motions.

The boundary conditions of the problem (2.16), (2.17) are satisfied by using relations (3.1), (3.2) for the initial and final moments of motion.

In relation to dual functions $F(t), G(t)$ and their first and second derivatives (that are viewed as the new dual controls) the optimization problem is formulated by the authors of this report, which includes the dual restrictions for the second derivatives of these functions. This problem is solved analytically using the Pontryagin's maximum principle; the control force $\mathbf{F}_{cx}$ and the control moment $\mathbf{M}_{cx}$ are derived using the derived biquaternions $\mathbf{\Lambda}(t)$ and $\mathbf{U}_x(t)$, and by solving the inverse dynamics problem of motion of a free rigid body. The derived analytical solution is the quasi-optimal time-optimal solution for the problem of

optimal spatial motion (maneuvering) of spacecraft. Dual algebra modelling software was developed to mathematically model the controlled motion of spacecraft using the derived control laws. Mathematical modelling demonstrated the efficiency of the developed theory of quasi-optimal programmed control of the spatial motion of spacecraft.

REFERENCES

[1] Yu.N. Chelnokov, "Controlling the Spatial Motion of a Rigid Body Using Biquaternions and Dual Matrices," Mechanics of Solids, vol. 56, no. 1, 2021, pp. 13–33.

[2] Yu.N. Chelnokov, "Synthesis of Control of Spatial Motion of a Rigid Body Using Dual Quaternions," Journal of Applied Mathematics and Mechanics, vol. 83, no. 5-6, 2019, pp. 704–733.

[3] Yu.N. Chelnokov, "Synthesis of Control of Spatial Motion of a Rigid Body Using Dual Quaternions," Mechanics of Solids, vol. 55, no. 7, 2020, pp. 59–80.

[4] A.V. Molodenkov, Y.G. Sapunkov, "Analytical Quasi-Optimal Solution of the Problem of the Time-Optimal Rotation of a Spacecraft," Journal of Computer and Systems Sciences International, vol. 60, no. 4, 2021, pp. 639-653.

[5] A.V. Molodenkov, Y.G. Sapunkov, "Solutions of the Optimal Turn Problem for a Spherically Symmetric Body under Arbitrary Boundary Conditions in the Class of Generalized Conical Motion," Mechanics of Solids, vol. 49, no. 5, 2014, pp. 495-505.

[6] A.V. Molodenkov, Y.G. Sapunkov, "Analytical Solution of the Minimum Time Slew Maneuver Problem for an Axially Symmetric Spacecraft in the Class of Conical Motions," Journal of Computer and Systems Sciences International, vol. 57, no. 2, 2018, pp. 302-318.

Biquaternion Solution to the Problem of Energy-Optimal Control of the Spacecraft Spatial Motion

I.A. Pankratov

National Research Saratov State University,
Institute for Problems of Precision Mechanics
and Control RAS
Saratov, Russia
PankratovIA.mechanic@gmail.com

Yu.N. Chelnokov

Institute for Problems of Precision Mechanics
and Control RAS
Saratov, Russia
ChelnokovYuN@gmail.com

Abstract — In this report the problem of energy-optimal program control of the spacecraft spatial motion in an inertial coordinate system is solved in a nonlinear formulation, using new biquaternion differential equations of spatial motion of a rigid body. The spacecraft is considered as a free rigid body of arbitrary dynamic configuration performing spatial motion equivalent to the composition of angular (rotational) and translational (orbital) motions. The boundary conditions for the angular and linear positions of the spacecraft, as well as for its angular and linear velocities are arbitrary. The control time is given and fixed; the dual vector control function (composition of angular and linear accelerations) is not limited by the dual module. A differential boundary value problem of the twenty-eighth order was obtained within the framework of the concept of solving inverse problems of dynamics using the Pontryagin maximum principle. Examples of numerical solutions of the boundary value problem are given for the cases when spacecraft mass distribution corresponds to a spherically symmetric rigid body or to the International Space Station (ISS). At the same time, the difference between the initial and final orientations of the spacecraft is large in angular measure and small in linear displacement (the problem of spatial maneuvering of the spacecraft). The analysis of the obtained numerical solutions is given; their characteristic properties and patterns are revealed.

Keywords — spacecraft, spatial motion, inverse problem of dynamics, maximum principle, energy-optimal control, quaternion, biquaternion.

I. STATEMENT OF THE PROBLEM

Let us consider a free rigid body, for example, a spacecraft capable to perform an arbitrary spatial motion relative to an inertial coordinate system, equivalent to a spatial helical motion (the composition of the translational motion of the body together with its arbitrarily selected point and the rotation of the body around this point). The body is under the action of an arbitrary main vector and the main moment of external forces, including the vector of the controlling force and the vector of the controlling moment. Biquaternion equations of spatial motion of a rigid body used to solve the problem with the help of the concept of solving inverse problems of dynamics and the maximum principle have the form [1-3]:

$$
\begin{aligned}
&\dot{\mathbf{U}}_x = \boldsymbol{\varepsilon}_x + s\mathbf{w}_x = \mathbf{H}_x, \qquad 2\dot{\boldsymbol{\Lambda}} = \boldsymbol{\Lambda} \circ \mathbf{U}_x; \\
&\boldsymbol{\varepsilon}_x = \varepsilon_1\mathbf{i} + \varepsilon_2\mathbf{j} + \varepsilon_3\mathbf{k}, \quad \mathbf{w}_x = w_1\mathbf{i} + w_2\mathbf{j} + w_3\mathbf{k}, \\
&\mathbf{H}_x = H_1\mathbf{i} + H_2\mathbf{j} + H_3\mathbf{k}, \quad H_k = \varepsilon_k + sw_k; \\
&\boldsymbol{\Lambda} = \Lambda_0 + \boldsymbol{\Lambda}_v = \Lambda_0 + \Lambda_1\mathbf{i} + \Lambda_2\mathbf{j} + \Lambda_3\mathbf{k} = \boldsymbol{\lambda} + s\boldsymbol{\lambda}^0, \\
&\Lambda_j = \lambda_j + s\lambda_j^0.
\end{aligned} \qquad (1)
$$

In these formulas, the phase variables are the biquaternion $\mathbf{U}_x = \boldsymbol{\omega}_x + s\mathbf{v}_x$ (mapping of the kinematic spacecraft screw $\mathbf{U}$ to the associated basis X) and the biquaternion $\boldsymbol{\Lambda} = \boldsymbol{\lambda} + s\boldsymbol{\lambda}^0$ of the finite displacement of the spacecraft in inertial space, whose main part (quaternion $\boldsymbol{\lambda}$) characterizes the orientation of the spacecraft in the inertial coordinate system, and the moment part (quaternion $\boldsymbol{\lambda}^0$) – the location of the spacecraft in this coordinate system (Cartesian coordinates ξ_k ($k = 1, 2, 3$) of the spacecraft center of mass in this coordinate system), and dual control is the biquaternion $\mathbf{H}_x = \boldsymbol{\varepsilon}_x + s\mathbf{w}_x$, it is a dual composition of the required absolute angular acceleration $\boldsymbol{\varepsilon}_x$ and the required component $\mathbf{w}_x$ of the spacecraft absolute linear acceleration.

It is necessary to construct a biquaternion program control $\mathbf{H}_x = \boldsymbol{\varepsilon}_x + s\mathbf{w}_x$, that provides a program translation of the spacecraft, whose motion is described by equations (1) from its arbitrary given initial state

$$
\boldsymbol{\Lambda}(0) = \boldsymbol{\lambda}(0) + s\boldsymbol{\lambda}^0(0), \qquad \mathbf{U}_x(0) = \boldsymbol{\omega}_x(0) + s\mathbf{v}_x(0) \qquad (2)
$$

into its specified final state

$$
\boldsymbol{\Lambda}(T) = \boldsymbol{\Lambda}_\mathrm{k} = \boldsymbol{\lambda}_\mathrm{k} + s\boldsymbol{\lambda}_\mathrm{k}^0, \qquad \mathbf{U}_x(T) = \mathbf{U}_x^\mathrm{k} = \boldsymbol{\omega}_x^\mathrm{k} + s\mathbf{v}_x^\mathrm{k}. \qquad (3)
$$

It is necessary to minimize the functional

$$
J = \int_0^T \left[\alpha_1\left(\varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2\right) + \alpha_2\left(w_1^2 + w_2^2 + w_3^2\right) \right] dt, \qquad (4)
$$

characterizing the energy consumption for the spacecraft transfer to the final state (3). The end time of the process T in functional (4) is fixed and given. There are no restrictions on the components of the required angular and linear acceleration vectors (controls).

It is known that the norms of quaternion $\boldsymbol{\lambda}$ and biquaternion $\boldsymbol{\Lambda}$ are equal to one (they are the first partial integrals of the problem). Therefore, the right boundary condition imposed by the first of the relations (3) on the biquaternion $\boldsymbol{\Lambda}$ of the spacecraft orientation and location is replaced by the equality of the vector parts of the main and moment parts of the biquaternion product $\widetilde{\boldsymbol{\Lambda}}(T) \circ \boldsymbol{\Lambda}_\mathrm{k}$ to zero:

$$
\mathrm{vect}\left[\widetilde{\boldsymbol{\lambda}}(T) \circ \boldsymbol{\lambda}_\mathrm{k}\right] = \mathbf{0}, \quad \mathrm{vect}\left[\widetilde{\boldsymbol{\lambda}}(T) \circ \boldsymbol{\lambda}_\mathrm{k}^0 + \widetilde{\boldsymbol{\lambda}}^0(T) \circ \boldsymbol{\lambda}_\mathrm{k}\right] = \mathbf{0}. \qquad (5)
$$

Vector equations (5) are equivalent to six scalar equations (in (5) and further, the wave above quaternion means the conjugate quaternion, and vect $[...]$ is the vector part of the quaternion). This replacement of the right quaternion boundary condition increases the efficiency of the numerical solution of the problem.

The Cartesian coordinates ξ_k of the spacecraft center of mass (i.e., the projections of the radius vector $\mathbf{r}$ of the spacecraft center of mass on the axis of the inertial coordinate system ξ) and the projections x_k of this vector on the axis of the coordinate system X associated with the spacecraft are found through the components of the quaternions $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}^0$ according to the formulas [4, 5]

$$\mathbf{r}_\xi = \xi_1\mathbf{i} + \xi_2\mathbf{j} + \xi_3\mathbf{k} = 2\boldsymbol{\lambda}^0 \circ \widetilde{\boldsymbol{\lambda}}, \quad \mathbf{r}_x = x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k} = \widetilde{\boldsymbol{\lambda}} \circ \boldsymbol{\lambda}^0.$$

After solving the formulated problem we have to separate the main $\boldsymbol{\varepsilon}_x$ and moment $\mathbf{w}_x = \dot{\mathbf{v}}_x$ parts in the constructed control $\mathbf{H}_x = \boldsymbol{\varepsilon}_x + s\mathbf{w}_x$.

After this separation, the laws of the formation of the control force $\mathbf{F}_{cx}$ and the control moment $\mathbf{M}_{cx}$ are obtained in accordance with the concept of solving inverse problems of dynamics by the known analytical formulas.

Note that the problem of constructing controls $\mathbf{w}_x$ and $\boldsymbol{\varepsilon}_x$ has general nature for all moving objects considered as a rigid body, since equations (1) are valid for any such moving object. The specifics of the object (its mass-inertia and other characteristics, acting external disturbing forces and their moments) are taken into account when we have to construct the control force $\mathbf{F}_{cx}$ and the control moment $\mathbf{M}_{cx}$ based on finite ratios (in their right parts, in general, we include the main vector $\mathbf{F} = \mathbf{F}(t, \mathbf{r}, \mathbf{v})$ of other external forces acting on the rigid body (forces of gravity, resistance to motion and other external perturbing forces) and the main moment $\mathbf{M} = \mathbf{M}(t, \boldsymbol{\lambda}, \boldsymbol{\omega})$ of these forces, calculated relative to the center of mass of the body, assumed by known functions of time t and variables $\mathbf{r}$, $\mathbf{v}$ and $\boldsymbol{\lambda}$, $\boldsymbol{\omega}$).

II. TRANSITION TO THE DIMENSIONLESS VARIABLES

Let's replace the dimensional variables of the problem to dimensionless ones by the following formulas (quaternion $\boldsymbol{\lambda}$ is dimensionless) [6]:

$$I^{mag} = \left((I_1^2 + I_2^2 + I_3^2)/3\right)^{1/2}, \quad I_k^{dl} = I_k / I^{mag};$$
$$\boldsymbol{\omega}_x^{dl} = T\boldsymbol{\omega}_x, \qquad\qquad t^{dl} = t/T,$$
$$\mathbf{M}^{dl} = \left(T^2/I^{mag}\right)\mathbf{M}; \qquad \boldsymbol{\lambda}^{0dl} = \boldsymbol{\lambda}^0/L,$$
$$\boldsymbol{\varepsilon}_x^{dl} = T^2\boldsymbol{\varepsilon}_x, \qquad\qquad w_x^{dt} = \left(L/T^2\right)w_x.$$

Then the dimensionless quaternion phase equations corresponding to the biquaternion phase equations (1) will take the form (the upper indices "dl" are omitted here and further, and the dot denotes the derivative in dimensionless time):

$$\dot{\boldsymbol{\omega}}_x = \boldsymbol{\varepsilon}_x, \qquad\qquad \dot{\mathbf{v}}_x = \mathbf{w}_x,$$
$$2\dot{\boldsymbol{\lambda}} = \boldsymbol{\lambda} \circ \boldsymbol{\omega}_x, \qquad 2\dot{\boldsymbol{\lambda}}^0 = \boldsymbol{\lambda} \circ \mathbf{v}_x + \boldsymbol{\lambda}^0 \circ \boldsymbol{\omega}_x. \qquad (6)$$

Note that the dimensional and dimensionless equations coincide in their form and the end time of the control process in functional (4) equals to 1.

Let's assume that the main vector of other external forces (in the expression for the controlling force) is due to the gravitational attraction of the spacecraft to the Earth, and the main moment of other external forces equals to zero. Then

the expressions for the controlling force and moment in dimensionless form will be written as follows

$$\mathbf{F}_{cx} = \mathbf{w}_x + \mathbf{K}(\boldsymbol{\omega}_x)\mathbf{v}_x + N_G \cdot \mathbf{r}_x / r^3,$$
$$\mathbf{M}_{cx} = \mathbf{J}\boldsymbol{\varepsilon}_x + \mathbf{K}(\boldsymbol{\omega}_x)\mathbf{J}\boldsymbol{\omega}_x. \qquad (7)$$

Here $N_G = G \cdot M_0 \cdot T^2 / L^3$ is a dimensionless parameter of the problem, G is the gravitational constant, and M_0 is the mass of the attracting body (Earth).

Note that the dimensionless expression for the control force in (7), taking into account that the quaternion $\boldsymbol{\lambda}$ is normalized, can be rewritten as follows

$$\mathbf{F}_{cx} = \mathbf{w}_x + \mathbf{K}(\boldsymbol{\omega}_x)\mathbf{v}_x + N_G \cdot \widetilde{\boldsymbol{\lambda}} \circ \boldsymbol{\lambda}^0 / \left\|\boldsymbol{\lambda}^0\right\|^{3/2}. \qquad (8)$$

III. THE LAW OF OPTIMAL CONTROL

We have solved this problem using the maximum principle of L.S. Pontryagin. First let us introduce additional dimensionless biquaternion variables $\boldsymbol{\Psi} = \boldsymbol{\varphi} + s\boldsymbol{\chi}$ and $\mathbf{M} = \boldsymbol{\mu} + s\boldsymbol{\mu}^0$. These variables are conjugate to the phase variables $\mathbf{U}_x = \boldsymbol{\omega}_x + s\mathbf{v}_x$ and $\boldsymbol{\Lambda} = \boldsymbol{\lambda} + s\boldsymbol{\lambda}^0$. Here $\boldsymbol{\varphi}$, $\boldsymbol{\chi}$ are vectors (quaternions with zero scalar parts), and $\boldsymbol{\mu}$, $\boldsymbol{\mu}^0$ are quaternions with non-zero scalar parts.

In this case, the Hamilton-Pontryagin function has the form

$$H = -\left[\alpha_1\left(\varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2\right) + \alpha_2\left(w_1^2 + w_2^2 + w_3^2\right)\right] + (\boldsymbol{\varphi}, \boldsymbol{\varepsilon}) + (\boldsymbol{\chi}, \mathbf{w}) +$$
$$+ (\boldsymbol{\mu}, \boldsymbol{\lambda} \circ \boldsymbol{\omega}_x)/2 + (\boldsymbol{\mu}^0, \boldsymbol{\lambda} \circ \mathbf{v}_x + \boldsymbol{\lambda}^0 \circ \boldsymbol{\omega}_x)/2,$$

where «(. , .)» is the scalar product of the corresponding three-dimensional and four-dimensional variables.

The system of equations for conjugate variables has the form:

$$2\dot{\boldsymbol{\varphi}} = \text{vect}\left[(\boldsymbol{\lambda}_v - \boldsymbol{\lambda}_0) \circ \boldsymbol{\mu} + (\boldsymbol{\lambda}_v^0 - \boldsymbol{\lambda}_0^0) \circ \boldsymbol{\mu}^0\right]$$
$$2\dot{\boldsymbol{\chi}} = \text{vect}\left[\boldsymbol{\lambda} \circ \boldsymbol{\mu}^0\right] \qquad (9)$$
$$2\dot{\boldsymbol{\mu}} = \boldsymbol{\mu}^0 \circ \mathbf{v}_x + \boldsymbol{\mu} \circ \boldsymbol{\omega}_x, \qquad 2\dot{\boldsymbol{\mu}}^0 = \boldsymbol{\mu}^0 \circ \boldsymbol{\omega}_x.$$

The law of energy-optimal control (i.e., the law of control satisfying the necessary optimality condition) has the form:

$$\boldsymbol{\varepsilon}_x^{opt} = \boldsymbol{\varphi}/(2\alpha_1), \qquad\qquad \mathbf{w}_x^{opt} = \boldsymbol{\chi}/(2\alpha_2). \qquad (10)$$

The control (10) maximizes Hamilton-Pontryagin function in variables $\boldsymbol{\varepsilon}_x$, $\mathbf{w}_x$.

Transversality conditions that do not contain indefinite Lagrange multipliers have the form:

$$t = 1, \qquad\qquad (\boldsymbol{\lambda}_k, \boldsymbol{\mu}^0) = 0,$$
$$-\lambda_{k0}^0(\text{vect }\boldsymbol{\lambda}_k, \text{vect }\boldsymbol{\mu}^0) + \lambda_{\kappa0}(\text{vect }\boldsymbol{\lambda}_k^0, \text{vect }\boldsymbol{\mu}^0) + \qquad (11)$$
$$+ \lambda_{k0}(\text{vect }\boldsymbol{\lambda}_k, \text{vect }\boldsymbol{\mu}) = 0.$$

Thus, the problem of energy-optimal control of the spacecraft spatial motion is reduced to a boundary value problem with a movable right end of the trajectory described by a system of twenty-eight order nonlinear differential equations (6), (9), (10) and twenty-six boundary conditions

(2), (3), (5). Also we should add two transversality conditions (11).

IV. EXAMPLES OF NUMERICAL SOLUTION OF THE PROBLEM

The algorithm proposed by us for the numerical solution of the problem is based on the Levenberg-Marquardt algorithm, which is a combination of the modified Newton method and the gradient descent method [7]. Also the Runge-Kutta method was used for integration of phase and conjugate differential equations.

Let's compare the results of the obtained solution of the energy-optimal control problem for rigid bodies (spacecraft) with different mass distribution, assuming the boundary conditions given in [6]:

$$\lambda(0) = (0.7951, 0.2981, -0.3975, 0.3478),$$
$$\omega_x(0) = (0.2739, -0.2388, -0.3);$$
$$\lambda_k = (0.8443, 0.3984, -0.3260, 0.1485), \quad \omega_x^k = (0, 0, -0.59).$$

Body 1. Spherically-symmetric rigid body: $I_1 = I_2 = I_3 = 1.0$ (dimensionless variables).

Body 2. International Space Station (ISS) (early version [8]) as an arbitrary rigid body: I_1=4853000 kg·m², I_2=23601000 kg·m², I_3=26278000 kg·m², (dimensional moments of inertia) or $I_1 = 0.2358$, $I_2 = 1.1466$, $I_3 = 1.2766$ (dimensionless values).

In order to compare our calculations with the results obtained in [6], it is necessary to put the initial and final values of the linear velocity vector of the spacecraft equal to the zero vectors.

In this case, the initial and final values of the components of the radius vector of the spacecraft center of mass (in the inertial coordinate system) must also coincide. Let the dimensionless components of the radius vector correspond to [9]: $\xi_1^0 = \xi_1^k$=23399727.8 m, $\xi_2^0 = \xi_2^k$=23962416.6 m, $\xi_3^0 = \xi_3^k = -18801552.4$ m.

It is also necessary to take $w_x^{opt} = 0$, then the linear velocity of the spacecraft will be zero during the entire time of its movement, i.e. the spacecraft will perform only rotational motion. Let the coefficients in the functional (4) are equal to $\alpha_1 = 1$ and $\alpha_2 = 0$.

In this formulation of the problem, the type of mass distribution of the spacecraft is taken into account only at the end, when we calculate the control moment, and it does not participate in the formation of optimal control. Therefore, the results of calculations in dimensionless variables correspond to the calculations given in [6] for the case of a spherically symmetric body.

In the absence of translational displacement for the case of spherical mass distribution, the values of the components of the orientation quaternion λ and the angular velocity vector ω at the intermediate point and angular acceleration ε at the beginning, middle and end of the motion are close to those indicated in Tables 1, 2 from paper [6; p. 102]. The value of the integral of the square of the control moment (the minimized functional in [6]) is equal to 0.4779, this value is close to the number 0.4782 from Table 8 in paper [6].

Let the scale factors for distance and time correspond to [10]: $L = 37000000.0$ m, $T = 11272.855470$ sec. Then, taking into account the known values for the mass of the Earth $M_0 = 5.9722 \cdot 10^{24}$ kg, and the gravitational constant $G = 6.67408 \cdot 10^{-11}$ N·m²·kg⁻², the dimensionless parameter of the problem will be equal to $N_G = 0.99997 \approx 1.0$.

Similar calculations were carried out for the case when the mass distribution of the spacecraft corresponds to the ISS. The value of the integral of the square of the control moment is equal to 0.35778, this value is less than a percentage higher than the number 0.35522 from Table 8 in paper [6].

Note that when there is no translational displacement, the values of the components of the dimensionless radius vector do not affect the laws of change of phase variables and optimal control, as well as the end time of the controlled process. In this case, only the moment part of the Λ biquaternion is multiplied by a constant quaternion (differential equations for main and moment parts of the Λ biquaternion coincide in their form up to the notation), and the initial values of the conjugate variables change. Also in this case, the radius vector of the spacecraft in the inertial coordinate system will have constant modulo and direction. The control force will be a constant vector in the inertial coordinate system ξ. In this coordinate system, the controlling force is a vector equal in modulo, but opposite in direction to the vector of gravitational attraction. In this case, the projections of the gravitational force on the coordinate axes associated with a spacecraft will change during the spacecraft angular motion.

Now let us consider the general case when the initial and final values of the components of the spacecraft radius vector and its velocity vector do not coincide. Let the coefficients in the functional (4) are equal to $\alpha_1 = 1$ and $\alpha_2 = 1$ respectively, and the dimensionless components of the radius vector and the velocity vector correspond to the dimensional parameters of the orbit from the book [11]: for the initial position of the spacecraft $\xi_1^0 = -12194795.0$ m, ξ_2^0=21779195.0 m, ξ_3^0=8278547.0 m, $\dot{\xi}_1^0 = -1080.750$ m/sec, $\dot{\xi}_2^0 = -1849.256$ m/sec, $\dot{\xi}_3^0 = 3274.225$ m/sec; for the final position of the spacecraft $\xi_1^k = -12110249.0$ m, ξ_2^k=21643438.0 m, ξ_3^k=8744787.0 m, $\dot{\xi}_1^k = -1063.392$ m/sec, $\dot{\xi}_2^k = -1905.728$ m/sec, $\dot{\xi}_3^k = 3247.462$ m/sec.

In the case of a spherical mass distribution, the laws of change of the components of the rotation quaternion and the angular velocity vector of the spacecraft in the presence of translational displacement practically have not changed due to the smallness of the linear displacement of the spacecraft.

Note that the components of the spacecraft radius vector, the vector of its linear velocity and the moment part of the biquaternion are harmonic functions of time. In this case, each of the components of the linear velocity vector changes sign twice approximately at $t = 0.21$ and $t = 0.78$ (the moments of sign change of each of the three components of the specified vector are close to each other, but do not

coincide). The points of time at which one of the components of the linear velocity vector of the spacecraft is zero are the points of extremum of the corresponding component of the spacecraft radius vector.

Each of the components of the optimal angular acceleration vector of the spacecraft is less than one in modulo (their numerical values coincide with the values of the corresponding component of the optimal control moment), but modulo of the specified vector is greater than one during the first 20% of the spacecraft motion time. At the same time, the laws of change of the first and third components of the specified vector are close to linear laws. At $t = 0.715$ these components practically coincide with each other; this point is the point of the minimum modulo of the spacecraft optimal angular acceleration vector. The modulo of the optimal linear acceleration vector decreases linearly to almost zero during the first half of the active motion of the spacecraft and also increases linearly further, exceeding six dimensionless units, both at the beginning and at the end of the motion. At the same time, the components of the optimal linear acceleration vector are close to linear functions. The first and second components of the specified vector decrease over the entire period of active spacecraft motion, and the third component increases. The first component of the control force vector is a decreasing function, and the third component is an increasing one (both are close to linear functions). Both of these components change the sign once during the spacecraft motion. At the same time, the second component of the control force vector changes according to the harmonic law, remaining positive throughout the spacecraft motion. The moment of changing the sign of the second component of the optimal acceleration vector is the inflection point of the second component of the control force vector.

In this paper we also present the calculation results for the case when the mass distribution of the spacecraft corresponds to the ISS. At the same time, the laws of change of the required quantities remained the same as in the previous case, with the exception of the control moment. The initial values of the first and third components of the control moment vector increased in modulo, and the initial value of the second component decreased in modulo. The initial value of the control moment modulo decreased by 12%, its final value decreased by 2.3 times.

We should note non-uniqueness of the numerical solution of the boundary value problem of optimal control of the spacecraft spatial motion associated with the nonlinearity of the differential equations of the problem. Under the same boundary conditions, different solutions for the laws of motion, control and behavior of conjugate variables were obtained in the formulation of the boundary value optimal control problem. The one solution that corresponds to the lower value of the functional (4) was chosen.

V. Conclusion

We have constructed a biquaternion theory of energy-optimal control of the spacecraft spatial motion. We have given examples of numerical solutions to this control problem for the case of spacecraft spherical symmetry and the case when the spacecraft mass distribution corresponds to the ISS. The characteristic properties and regularities of these solutions have been revealed. The combination of two methods for solving boundary value problems (modified Newton method and gradient descent method) has made it possible to increase the accuracy of the numerical solution of the corresponding boundary value problem of optimal control.

References

[1] Yu.N. Chelnokov, "Upravlenie prostranstvennym dvizheniem tverdogo tela s ispol'zovaniem dual'nykh kvaternionov (Control of spatial motion of a rigid body using dual quaternions)," XII Vserossiiskii s"ezd po fundamental'nym problemam teoreticheskoi i prikladnoi mekhaniki (Ufa, 20–24 avgusta 2019 g.): Sbornik trudov v 4 tomakh. V. 1: Obshchaya i prikladnaya mekhanika. RITs BashGU, Ufa, 2019. pp. 288-290. DOI: 10.22226/2410-3535-2019-congress-v1

[2] Yu.N. Chelnokov, "Synthesis of control of spatial motion of a rigid body using dual quaternions," Mechanics of Solids, 2020, vol. 55, no. 7, pp. 59–80. DOI: 10.3103/S0025654420070080

[3] Yu.N. Chelnokov, "Controlling the spatial motion of a rigid body using biquaternions and dual matrices," Mechanics of Solids, 2021, vol. 56, no. 1, pp. 13-33.

[4] Yu.N. Chelnokov, "On integration of kinematic equations of a rigid body's screw motion," Applied mathematics and mechanics, 1980, vol. 44, no. 1, pp. 19-23.

[5] Yu.N. Chelnokov, "Kvaternionnye i bikvaternionnye modeli i metody mekhaniki tverdogo tela i ikh prilozheniya: Geometriya i kinematika dvizheniya (Quaternion and biquaternion models and methods of solid mechanics and their applications: Geometry and kinematics of motion)," Fizmatlit, Moscow, 2006.

[6] Ya.G. Sapunkov, A.V. Molodenkov, "The new algorithm of quasi-optimal reorientation of a spacecraft," Izvestiya of Saratov University. Mathematics. Mechanics. Informatics, 2023, vol. 23, iss.4, pp. 95–112.

[7] N.N. Moiseev, "Chislennye metody v teorii optimal'nyh sistem (Numerical methods in the theory of optimal systems)," Nauka, Moscow, 1971.

[8] Yu.R. Banit, M.Yu. Belyaev, T.A. Dobrinskaya, N.I. Efimov, V.V. Sazonov, V.M. Stazhkov, "Opredelenie tenzora inertsii mezhdunarodnoi kosmicheskoi stantsii po telemetricheskoi informatsii," Preprint № 57, IPM im. M.V. Keldysha RAN. Moscow, 2002, 19 p.

[9] T.V. Bordovitsyna, "Sovremennye chislennye metody v zadachah nebesnoi mekhaniki (Modern numerical methods in problems of celestial mechanics)," Nauka, Moscow, 1984.

[10] I.A. Pankratov, Ya.G. Sapunkov, Yu.N. Chelnokov, "Quaternion models and algorithms for solving the general problem of optimal reorientation of spacecraft orbit," Izvestiya of Saratov University. Mathematics. Mechanics. Informatics, 2020, vol. 20, iss. 1, pp. 93-104.

[11] T.V. Bordovitsyna, V.A. Avdiushev, "Teoriya dvizheniya iskusstvennykh sputnikov Zemli. Analiticheskie i chislennye metody (Theory of motion of artificial Earth satellites. Analytical and numerical methods)," Izdatel'skii Dom Tomskogo gosudarstvennogo universiteta, Tomsk, 2016.

Study of the Stabilizing Properties of a Jet Stream in the Angular Motion of Spacecraft

Anton Doroshin
Theoretical mechanics dep.
Samara National Research University
Samara, Russia
https://orcid.org/0000-0001-6310-0963

Mikhail Krikunov
Laboratory of dynamics and motion control of aircraft
Samara National Research University
Samara, Russia
https://orcid.org/0000-0002-0379-7372

Abstract—In order to study the qualitative behavior of spacecraft (SC) as a body with variable mass, one can use the phase trajectory curvature analysis method, which is outlined by the apex of the longitudinal axis of the vehicle, being a thrust axis, and study the evolution of the curvature value. The present paper presents a study of a jet stream properties during the angular motion of a rocket taking into account the K. Magnus effect and modification the phase trajectory curvature analysis method for the case when the moments of forces applied to the vehicle depend on the components of the angular velocity vector, in particular, when the moments are created by Coriolis forces arising from the jet.

Keywords—spacecraft, variable structure, phase trajectory curvature analysis method, qualitative analysis.

I. INTRODUCTION

Studying the dynamics of the angular motion of spacecraft with variable composition with changing mass and moments of inertia over time is still one of the most important problems of space flight mechanics. The fundamental aspects of the problem were previously considered in [1]–[6], as well as the applied developments of it [7]–[9], and also as part of the study of additional aspects of dynamics [10].

In this work, based on fundamental results and previously obtained well-known models, the dynamics of the angular motion of spacecraft (SC) with variable composition in the boost phase of the trajectory is analyzed, taking into account the dynamic effect created by the movement of jet particles from the nozzles and combustion channels of jet engines, which was previously considered by K. Magnus [5]. As shown in [5], this dynamic effect (Magnus effect) generates Coriolis forces and moments, which, as a rule, have a damping effect on spatial oscillations, leading to a decrease in the opening of the precession cone, which is a positive stabilizing factor in the dynamics of rockets and spacecraft with variable composition. However, the noted Magnus effect can lead to destabilization of the angular motion of spacecraft, when nutational oscillations increase and the precession cone unfolds. Demonstration and study of such "negative" manifestations of the Magnus effect is the main goal of this article, demonstrating the novelty of the research in terms of identifying new properties of motion following from the known models. Moreover, it is

necessary to note the importance and relevance of the problem of analysis and synthesis of precessional motion studied in the article, still attracting the attention of famous researchers in rigid body dynamics [11]–[15].

II. MATHEMATICAL MODELING OF THE MOTION DYNAMICS OF VARIABLE STRUCTURE SPACECRAFT

Let us consider the motion of spacecraft with variable composition during the boost phase of its trajectory. The equations of motion in the coordinate system (CS) associated with the SC are written as follows [9]:

$$\begin{cases} A(t)\dot{p} + (C(t) - B(t))qr = M_x, \\ B(t)\dot{q} + (A(t) - C(t))rp = M_y, \\ C(t)\dot{r} + (B(t) - A(t))pq = M_z, \end{cases} \quad (1)$$

where $A(t)$, $B(t)$, $C(t)$ are the main moments of inertia in the coordinate system associated with the SC; p, q, r — components of the angular velocity vector; M_x, M_y, M_z are components of the external and reactive forces moments applied to the spacecraft.

In his monograph [5], K. Magnus as an example considered a very important phenomenon for rocket technology that reveals the damping effect of a jet stream, namely, equations (1) were considered for a free gyroscope with variable composition (a system with a jet engine) taking into account the moments of reactive forces along with the absence of other external influences. For an expanded description of the Magnus effect and its possible negative destabilizing influence, we shall write the differential of the Coriolis force from a small particle of a jet stream moving along the combustion channel with a relative speed $\vec{v}_r = [0, \ 0, \ V]^{\text{T}}$ to the associated $Cxyz$ axes (Fig. 1):

$$dF_{coriolis} = -2dm \cdot \vec{\omega}_e \times \vec{v}_r = -dm \begin{bmatrix} \vec{i} & \vec{j} & \vec{k} \\ p & q & r \\ 0 & 0 & V \end{bmatrix} =$$

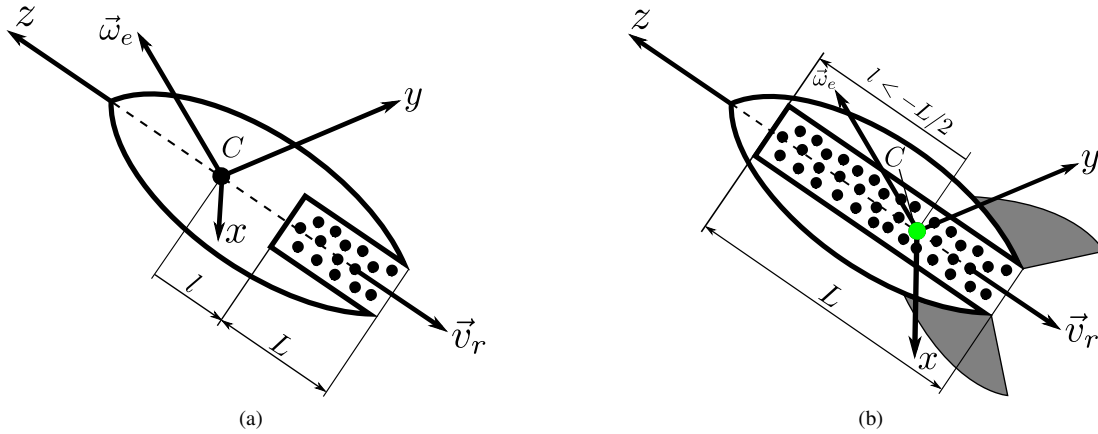$$= -\left[\vec{i}(qV) + \vec{j}(-pV) + \vec{k} \cdot 0\right] \cdot 2dm.$$

Fig. 1: To determine the damping effect of the jet stream.

Let us write down the differential moment of the Coriolis forces:

$$d\vec{M}_{coriolis} = \vec{r} \times d\vec{F}_{coriolis} =$$
$$= [0, 0, -(l+\xi)]^{\mathrm{T}} \times [-qV, pV, 0] \cdot 2dm =$$
$$= - \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 0 & 0 & l+\xi \\ -qV & pV & 0 \end{vmatrix} \cdot 2dm =$$
$$= -2 \underbrace{\gamma d\xi}_{dm} \left[-\vec{i}pV(l+\xi), -\vec{j}qV(l+\xi), 0 \right]^{\mathrm{T}}, \quad (2)$$

where γ is linear mass density of a jet stream, the value l is a distance between the upper edge of the combustion channel and the center of mass (Fig. 1a).

Having integrated (2) along the entire combustion channel, we obtain the integral moment from the Coriolis forces:

$$\vec{M}_{coriolis} = -\int_0^L 2\gamma V \begin{bmatrix} -p(l+\xi) \\ -q(l+\xi) \\ 0 \end{bmatrix} d\xi =$$
$$= -2\gamma V \begin{bmatrix} -p \\ -q \\ 0 \end{bmatrix} \frac{(l+\xi)^2}{2} \Big|_{\xi=0}^{\xi=L} =$$
$$= -2\gamma V \begin{bmatrix} -p \\ -q \\ 0 \end{bmatrix} ((l+L)^2 - l^2) = -D \begin{bmatrix} -p \\ -q \\ 0 \end{bmatrix} V,$$

where

$$D = \gamma \left((l+L)^2 - l^2 \right). \quad (3)$$

Taking into account the fact that the emission of particles occurs in the opposite direction to the z axis, the velocity component V will be negative, which can be written $V = -v_r$, $v_r > 0$.

Thus, the moment of Coriolis forces from the particles of a jet stream will have the form:

$$\vec{M}_{coriolis} = Dv_r \begin{bmatrix} -p & -q & 0 \end{bmatrix}^{\mathrm{T}}. \quad (4)$$

Let's substitute (4) into (1) and write the equations of motion for a body/spacecraft with variable composition:

$$\begin{cases} A(t)\dot{p} + (C(t) - B(t))qr = -Dv_r p, \\ B(t)\dot{q} + (A(t) - C(t))rp = -Dv_r q, \\ C(t)\dot{r} + (B(t) - A(t))pq = 0. \end{cases} \quad (5)$$

Similar to [5], we shall consider the base case of the Magnus effect. Let us consider that $A \equiv B$, then, performing elementary transformations, (5) can be written as:

$$\begin{cases} \dot{p} + \dfrac{Dv_r}{A(t)}p - r_0 \left(1 - \dfrac{C(t)}{A(t)} \right) q = 0, \\ \dot{q} + \dfrac{Dv_r}{A(t)}q - r_0 \left(1 - \dfrac{C(t)}{A(t)} \right) p = 0, \\ \qquad\qquad\qquad\qquad\qquad r = r_0. \end{cases} \quad (6)$$

We will introduce the following notation:

$$\mu(t) = Dv_r/A(t), \qquad \nu(t) = r_0 \left(1 - C(t)/A(t) \right),$$

then the first two equations (6) will take the form:

$$\begin{cases} \dot{p} + \mu(t)p - \nu(t)q = 0, \\ \dot{q} + \mu(t)q + \nu(t)p = 0. \end{cases} \quad (7)$$

Following K. Magnus [5], on the basis of introducing the complex variable $\omega^* = p + iq$, we can reduce (7) to an equation of a new complex variable with an exact solution:

$$\omega^* = \omega_0^* \exp\left[-\int_0^t (\mu + i\nu)dt + \varphi_0 \right].$$

The last solution shows the transverse components of the angular velocity vector p and q are monotonically decreasing within the studied case, and the angular velocity vector itself is approaching the Cz axis of the associated SC, which means folding into the precession cone with damping of nutational oscillations under the assumption that the value $\mu(t)$ always remains positive. In other words, Magnus gyroscope with variable composition (SC with variable composition) is

damped by the engine jet stream while tending to suppress nutational oscillations as well as to collapse the precession cone towards the z axis.

In order to compare the conclusions, we will use the results of work [9], presenting the following dynamic equations:

$$
\begin{cases}
\dot{F} = -\dfrac{r}{AB}\left[B\left(C-B\right)\cos^2 F + \right. \\
\qquad \left. + A\left(C-A\right)\sin^2 F\right] + \dfrac{M_x}{A}\cos F - \dfrac{M_y}{B}\sin F, \\
\dot{G} = -G\dfrac{r}{AB}\sin F \cos F \left(B-A\right)\left(C-A-B\right) + \\
\qquad + \dfrac{M_x}{A}\sin F + \dfrac{M_y}{B}\cos F, \\
\dot{r} = -G^2\dfrac{B-A}{C}\sin F \cos F + \dfrac{M_z}{C}.
\end{cases} \quad (8)
$$

With the polar substitution of variables being used [7]–[9]:

$$
p = G(t)\sin F(t), \qquad q = G(t)\cos F(t). \quad (9)
$$

Let us consider the case of dynamic symmetry of the body $A \equiv B$, then:

$$
\dot{G} = \frac{f_G}{A(t)}, \quad \dot{F} = \frac{(C(t)-A(t))r + f_F}{A(t)}, \quad \dot{r} = 0, \quad (10)
$$

where

$$
f_G = M_x \sin F + M_y \cos F, \, f_F = M_x \cos F - M_y \sin F. \quad (11)
$$

The Coriolis forces of the particles of the jet stream we will use as the only moments of forces (4), thus:

$$
M_x = -Dv_r p, \qquad M_y = -Dv_r q. \quad (12)
$$

Substituting (12) and (9) into (11), we will have:

$$
f_G = -Dv_r G, \qquad f_F = 0, \quad (13)
$$

and after substituting (13) into (10) we will have:

$$
\dot{G} = -Dv_r G/A(t), \quad \dot{F} = (1 - C(t)/A(t))\, r_0, \quad r = r_0. \quad (14)
$$

Let's integrate the first equation (14):

$$
G = G_0 \exp\left[-\int_0^t Dv_r/A(t)\,dt\right]. \quad (15)
$$

From solution (15) it obviously follows the same dynamic fact that the amplitude of the angular velocities p and q decreases with time when the value of D is positive. Under the assumptions of Magnus, this fact completely confirms the stabilizing properties of the jet stream [5].

In a more detailed version, to take into account the nuances associated with the type of time dependence of moments of inertia, it is necessary to use the evaluation method for the evolution of the phase trajectory curvature [9], which allows to write the function $P(t)$:

$$
P(t) = \frac{\dot{F} - r_0}{G}\,\frac{d}{dt}\left(\frac{\dot{F} - r_0}{G}\right). \quad (16)
$$

After differentiating and substituting the expressions for $\dot{F}$, $\dot{G}$ and $\ddot{F}$, we have the following:

$$
P(t) = \frac{Cr_0^2}{AG^2} \cdot \left(\frac{d}{dt}\left(\frac{C}{A}\right) + \frac{Dv_r C}{A^2}\right).
$$

It is easy to see that the sign of the function $P(t)$ is determined by the expression in parentheses, therefore the condition for the positivity of the function $P(t)$ is:

$$
\dot{C}A + C(Dv_r - \dot{A}) > 0 \quad \text{or} \quad \dot{C}/C > (\dot{A} - Dv_r)/A. \quad (17)
$$

The positivity of the function $P(t)$ indicates that the apex curvature of the longitudinal axis of the spacecraft is constantly increasing, i.e. the apex moves along a twisting spiral with a folding cone of precession. Conditions (17) are taking into account the influence of Coriolis forces of jet particles. When this factor is excluded from consideration, while $Dv_r = 0$, conditions (17) will be completely consistent with the conditions mentioned earlier [7]–[9].

In other words, if the following condition is met:

$$
D = \gamma\left((l + L)^2 - l^2\right) > 0 \quad (18)
$$

the result (17) coincides with the well-known case considered by K. Magnus [5] for variable-mass gyroscopes (in applications to rockets), when the jet stream properties are creating stabilizing and damping moments of Coriolis inertia forces inside the combustion channel through additional pressure on the channel walls resulting from the relative movement of particles. It is worth noting that a similar analysis of moments of forces damping in the right sides of (5) was also studied in the aspects of the dynamics of coaxial systems of bodies with variable composition [7].

As we have already noted the indicated positive effect on the jet particles dynamics with suppression of the precession cone [5], [7], will take place in cases of inertia and mass configuration of a body with variable composition that satisfies condition (18). A similar case is visualized in Fig. 1a.

Moreover, condition (18), as well as conditions $A(t) > C(t)$, $\dot{A} < 0$, $\dot{C} < 0$ will be satisfied for elongated shapes of bodies with preferably longitudinal structure and a combustion channel located behind the center of mass. It is the case that was studied in [5], and it is precisely this situation that is basically typical for liquid-propellant rockets, with fuel particles burning in the lower part of the rocket (in the combustion chamber), and the rocket nozzle is quite short and located at the bottom. Also we shall assume here that movement of particles in fuel lines into the combustion chamber can be considered as slow compared to their movement in the nozzle during combustion, and therefore no noticeable Coriolis effects are created in the lines.

However, condition (18) may not be satisfied. In this way, new approaches to the design of rocket and space systems and the emergence of new ideas for types and layouts of rocket engines, for example [16], suggest that the center of mass can geometrically shift inside the body in the downward direction. This can create an negative lever arms for distributed moments of Coriolis forces immediately or in the process of

changing inertia and mass configuration when the value of D becomes negative (Fig. 1b). This situation is also quite common for the solid propellant rockets, where the entire combustion channel is almost a single nozzle, in with particles moving from the front to the nozzle exit with high pressure and velocity, creating distributed Coriolis moments of force, both "below" the center of mass and "above" center of mass, when negative integral effects can occur, creating a swing of the nutation angle instead of damping precessional oscillations.

In connection with the above, let us consider the effect of the moment of Coriolis forces in a more general formulation, taking into account the fact that the value of D can become negative, while the moments of inertia A and B are not equal. Let us consider an example of calculation for the case of burning a solid propellant charge, when $l < -L/2$, for example, for $l = -0.52L$, when $A \neq B$, for example, $B = 0.9A$. In this case, the spacecraf's center of mass and the center of mass of the charge will be at different points (Fig. 2).
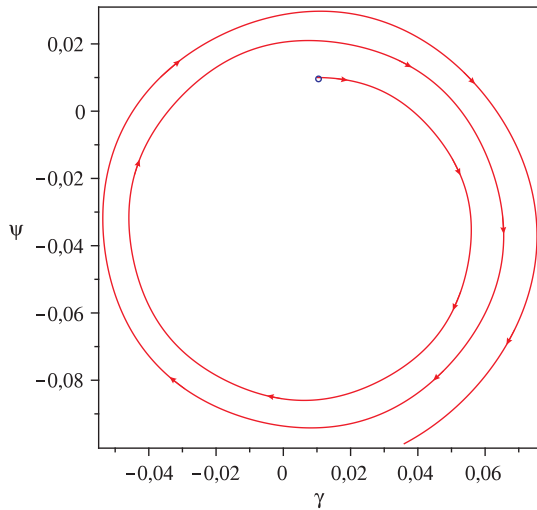

Fig. 2: Thrust vector hodograph for a solid propellant cylindrical charge.

The remaining parameters of the system are: $L = 2$, $H = 2$, $v_r = 100$, $t_k = 12$, $R_1 = 0.4$, $R_2(t) = 0.025t$, $\rho = 780$.

Fig. 2 shows a hodograph of the longitudinal axis apex and the thrust vector for the complete system, taking into account the Coriolis effects of the relative motion of the jet particles. As follows from the integration of the complete system, there is a "negative" Magnus effect, when the precession cone increases its opening and we shall see an unwinding spiral of the longitudinal axis hodograph.

III. ANALYTICAL SYNTHESIS

Let's write the function $\langle P \rangle$. Similar to [9], we will use the system:

$$\begin{cases} \dot{F} = (1/2)r\left(\alpha - \beta \cos 2F\right) + \eta_1, \\ \dot{G} = (1/2)Gr\beta \sin 2F + \eta_2, \\ \dot{r} = (1/2)G^2\xi_3 \sin 2F + \eta_3, \end{cases} \quad (19)$$

where

$$\alpha = \xi_1 + \xi_2, \quad \beta = \xi_1 - \xi_2;$$
$$\xi_1 = B/A - C/A, \quad \xi_2 = A/B - C/B,$$
$$\xi_3 = A/C - B/C; \quad (20)$$
$$\eta_1 = (M_x/A)\cos F - (M_y/B)\sin F,$$
$$\eta_2 = (M_x/A)\sin F + (M_y/B)\cos F, \quad \eta_3 = M_z/C.$$

Substituting (12) and (9) into (20), we will have:

$$\eta_1 = -(1/2)Dv_rG\sin 2F, \quad \eta_3 = M_z/C,$$
$$\eta_2 = -\frac{1}{2}Dv_rG\left[\left(\frac{1}{A} + \frac{1}{B}\right) + \left(\frac{1}{A} - \frac{1}{B}\right)\cos 2F\right]. \quad (21)$$

After substituting (21) into (19) and averaging over F we will have:

$$\begin{cases} \langle \dot{F} \rangle = (1/2)r\alpha, \\ \langle \dot{G} \rangle = -(1/2)Dv_rG\left(1/A + 1/B\right), \\ \langle \dot{r} \rangle = M_z/C. \end{cases} \quad (22)$$

In this case, it is possible to construct an averaged curvature function, which needs to be written explicitly:

$$\langle P \rangle = \frac{\langle \dot{F} \rangle - \langle r \rangle}{\langle G \rangle} \cdot \frac{d}{dt}\left[\frac{\langle \dot{F} \rangle - \langle r \rangle}{\langle G \rangle}\right]. \quad (23)$$

To explicitly write (23), we will substitute (22) into (23) and obtain the following refined form, which also takes into account the Magnus effect:

$$\langle P \rangle = \frac{r^2(\alpha - 2)^2}{4G^2}\left[\frac{M_z}{Cr} + \frac{\dot{\alpha}}{\alpha - 2} + \frac{1}{2}Dv_r\left(\frac{1}{A} + \frac{1}{B}\right)\right]. \quad (24)$$

To analyze and synthesize the dynamics of the system, we will use the refined (24), allowing us to draw the following conclusions:

1) Formula (24), in contrast to a similar expression obtained in [9], contains an additional term $(1/2)Dv_r(1/A + 1/B)$, which determines the action of the Magnus effect.

2) The case studied by Magnus [5] with the damping effect of a jet stream turns out to be a special case of dynamics, following from expression (24).

3) The Magnus case [5] is described only when the change in the inertia and mass configuration of the system is keeping the value of D positive throughout the entire time of the system motion. However, as it was shown, it may happen that while $D < 0$, Magnus effect will no longer have a damping effect on the dynamics of angular motion; on the contrary, the opening of the nutation cone will increase.

4) Refined dynamics of a system with moments of forces depending on the components of the angular velocity vector were studied in a fundamental sense earlier [7] under the assumption of dynamic symmetry of solid bodies. In present work, these features are taken into account from the perspective of applied research in connection with the analysis/synthesis of rocket and

space systems dynamics with a generalization to the case of dynamic asymmetry.

5) Formula (24) is also applicable for the case when there is no dynamic symmetry.

IV. CONCLUSIONS

Thus, for a detailed study of the influence of variable inertia and mass configuration of a body/spacecraft with variable composition with solid rocket engines, taking into account the Magnus effect in any of its implementations (positive or negative), you can use (24). Using the equation, it is possible to analyze and/or synthesize the dynamics of a system, considering the methods and types of solid fuel charges internal placement, and the laws of its combustion (volumetric, surface, end, channel, etc.), which will ultimately allow us to obtain the required time-dependent moments of inertia in connected axes, contributing to the implementation of the positive dynamics of issuing reactive impulses with subsequent natural focusing of the direction of operation of the jet engine as the precession cone collapses inward.

REFERENCES

[1] R. Grammel, *Der Kreisel: seine theorie und seine Anwendungen*. Vieweg+ Teubner Verlag, 1920.

[2] I. V. Meshchersky, *Works on Mechanics of Variable-Mass Bodies*, 2nd ed. Moscow: Gostechizdat, 1952.

[3] F. R. Gantmakher and L. M. Levin, *The Flight of Uncontrolled Rockets*. Oxford: Pergamon Press, 1964.

[4] A. A. Kosmodem'ianskii, *A Course in Theoretical Mechanics, Part 2*. Jerusalem: Published for the National Science Foundation by the Israel Program for Scientific Translations, 1963.

[5] K. Magnus, *Kreisel: theorie und anwendungen*. Springer-Verlag, 1971.

[6] L. Cveticanin, *Dynamics of bodies with time-variable mass*. Springer, 2016.

[7] A. V. Doroshin, "Evolution of the precessional motion of unbalanced gyrostats of variable structure," *Journal of Applied Mathematics and Mechanics*, vol. 72, no. 3, pp. 259–269, 2008.

[8] ——, "Analysis of attitude motion evolutions of variable mass gyrostats and coaxial rigid bodies system," *International Journal of Non-Linear Mechanics*, vol. 45, no. 2, pp. 193–205, 2010.

[9] A. V. Doroshin and M. M. Krikunov, "The generalized method of phase trajectory curvature synthesis in spacecraft attitude dynamics tasks," *International Journal of Non-Linear Mechanics*, vol. 147, p. 104246, 2022.

[10] J. Wolny and R. Strzałka, "Momentum in the dynamics of variable-mass systems: Classical and relativistic case," *Acta Physica Polonica A*, vol. 135, no. 3, pp. 475–479, 2019.

[11] F. L. Chernousko, L. D. Akulenko, and D. D. Leshchenko, *Evolution of motions of a rigid body about its center of mass*. Springer, 2017.

[12] V. S. Aslanov, *Rigid Body Dynamics for Space Applications*. Butterworth-Heinemann, 2017.

[13] A. P. Markeev, "Stability in the regular precession of an asymmetrical gyroscope in the critical case of fourth-order resonance," *Doklady Physics*, vol. 63, no. 7, pp. 297–301, 2018.

[14] D. Leshchenko, S. Ershkov, and T. Kozachenko, "Evolution of a heavy rigid body rotation under the action of unsteady restoring and perturbation torques," *Nonlinear Dynamics*, vol. 103, no. 2, pp. 1517–1528, 2021.

[15] ——, "Rotations of a rigid body close to the lagrange case under the action of nonstationary perturbation torque," *Journal of Applied and Computational Mechanics*, vol. 8, no. 3, pp. 1023–1031, 2022.

[16] A. V. Elkin, E. S. Zemerev, V. I. Malinin, L. L. Khimenko, and F. R. Khaziakhmetova, "Rocket engine on granular solid fuel," *PNRPU Aerospace Engineering Bulletin*, vol. 64, pp. 16–24, 2021.

Controlling the Sensing Element of a Pendulum Compensation Accelerometer
Using a Static Modal Combined Controller H_2/H_∞

V.M. Nikiforov

Academician Pilyugin Scientific-Production Center of Automatics and Instrument-Making
Moscow, Russia

A.V. Solov'ev

Academician Pilyugin Scientific-Production Center of Automatics and Instrument-Making
Moscow, Russia

M.L. Simakov

Academician Pilyugin Scientific-Production Center of Automatics and Instrument-Making
Moscow, Russia

A.A. Gusev

Academician Pilyugin Scientific-Production Center of Automatics and Instrument-Making
Moscow, Russia

K.A. Andreev

Academician Pilyugin Scientific-Production Center of Automatics and Instrument-Making
Moscow, Russia

A.V. Shevchenko

Academician Pilyugin Scientific-Production Center of Automatics and Instrument-Making
Moscow, Russia

Abstract — **The paper considers the quality assurance of control of the sensitive element of a compensating pendulum accelerometer based on minimization of norms representing the measure of the energy spent on control and the worst transfer coefficient from the disturbing influence to the controlled parameter. The control laws of a sensitive element which is part of the accelerometer feedback are synthesized. Conclusions are made on the provision of control quality parameters with different controllers.**

Keywords — pendulum compensation accelerometer, minimization of norms, modal combined controller, static controller, dynamic controller, linear matrix inequalities.

I. Introduction

The increasing operational requirements for the functions of technical facilities is leading to more requirements being imposed on control system equipment. Sensor control quality of system control units of various technical facilities significantly affects their performance. In turn, improving sensor control quality results in the application of control laws based on modern alternative approaches, such as minimization of the norms H_2 and H_∞. The operating quality of technical control systems is characterized by various quality functionals that impose restrictions on control actions, state vector variables and control time. In some instances, with rated control parameter spaces, the norms can be used as quality functionals.

Norms of transfer matrices allow evaluating the influence of the input signal magnitude on the output signal magnitude. In the case that the input signal is a disturbing influence, the quality can be estimated based on the norm value, that is, the lower the norm, the more the disturbing influence is suppressed and, accordingly, the higher the control quality. Controller synthesis is performed through minimization of the transfer function norm.

The objective of this study is to improve the sensor control quality of a pendulum accelerometer through minimizing the norms H_2 and H_∞**.**

The object of the study is a pendulum compensation accelerometer represented by an extended mathematical model [1, 2].

The subject of the study is control law synthesis through minimization of the norms H_2 and H_∞ by a pendulum compensation accelerometer sensor forming part of the accelerometer feedback when exposed to external disturbing influences.

II. Task Formulation

The task of this work can be formulated as follows: for a pendulum compensation accelerometer represented by an extended mathematical model, it is necessary to synthesize the controller in the feedback of the accelerometer based on minimization of norms H_2 and H_∞, which provides quality parameters stabilization time $13 \cdot 10^{-3} \leq t_y \leq 15 \cdot 10^{-3}$ and overshoot $\sigma \leq 6.5$ %, as well as the degree of stability, the degree of speed, etc.

III. Synthesis of controllers

Methods for optimising linear control systems according to Hardy space norms H_2 and H_∞ are used to build controllers in various technical systems. State-space models for a closed-loop system are used to solve problems of synthesising static or dynamic controllers with a quality functional in the form of norms H_2 and H_∞. The controller synthesis is performed based on minimizing the norm H_2, based on minimizing the norm H_∞, or based on minimizing the norm H_2/H_∞. The norm value H_2 represents a measure of the energy spent on control; the norm value H_∞ is the worst transfer coefficient from the disturbing influence to the controlled parameter.

The norm value H_2 is defined as the spent reaction energy to the signal in the form of delta function through the frequency characteristic according to the formula:

$$\|W\|_{H_2}^2 = \frac{1}{2\pi} trace \int_{-\infty}^{\infty} W(\sigma + j\omega) W(\sigma + j\omega)^* d\omega.$$

The norm value H_∞ is defined as the maximum value of the frequency ω of the maximum singular value of frequency characteristics according to the formula:

$$\|W\|_\infty = \sup_{\omega \in \mathbb{R}} \sigma_{max}(W(j\omega)).$$

One way to synthesize a controller is by using linear matrix inequalities [3, 4]. The composition of the linear matrix inequality system depends on the minimized norm.

Another way to synthesize a controller is by using an algorithm based on solving the Riccati equations (II-Riccati method).

The equations of Riccati I and Riccati II in matrix form when minimized H_2 have the form:

$$A^T P + PA + C_2^T C_2 - \left(PB + C_2^T D_2\right)\left(D_2^T D_2\right)^{-1}\left(D_2^T C_2 + B^T P\right) = 0,$$

$$QA^T + AQ + EE^T - \left(QC_1^T + ED_1^T\right)\left(D_1 D_1^T\right)^{-1}\left(D_1 E^T + C_1 Q\right) = 0.$$

The equations of Riccati I and Riccati II in matrix form when minimized H_∞ have the form:

$$A^T P + PA + C_2^T C_2 + PEE^T P / \gamma^2 - \left(PB + C_2^T D_2\right)\left(D_2^T D_2\right)^{-1}\left(D_2^T C_2 + B^T P\right) = 0,$$

$$QA^T + AQ + EE^T + QC_2^T C_2 Q / \gamma^2 - \left(QC_1^T + ED_1^T\right)\left(D_1 D_1^T\right)^{-1}\left(D_1 E^T + C_1 Q\right) = 0.$$

When synthesizing a modal control law based on norm H_2/H_∞ minimization using linear matrix inequalities, it is necessary that the roots of the characteristic equation be located in the D-domain. The D-domain is bounded by curves defined based on linear matrix inequalities.

As a result of the synthesis for static controllers, the following were obtained:

- controller gains,
- norm values H_2 and H_∞.

As a result of the synthesis of the dynamic controller, the following were obtained:

- controller support matrices,
- norm values H_2 and H_∞.

IV. MODELING RESULTS

Mathematical modelling of a pendulum compensation accelerometer was carried out in a dynamic modelling environment with synthesised static and dynamic controllers. Graphs of changes in the angle of rotation of the sensing element of the pendulum compensation accelerometer, changes in the angular velocity of the sensing element and changes in the singular value of the closed system of the pendulum compensation accelerometer were obtained.

Graphs of changes in the angle of rotation of the sensing element, changes in the angular velocity of the sensing element and changes in the singular value of the closed system of the pendulum compensation accelerometer with the dynamic modal combined controller are shown in Figs. 1-3.

Graphs of changes in the angle of rotation of the sensing element, changes in the angular velocity of the sensing element and changes in the singular value of the closed system of the pendulum compensation accelerometer with the static modal combined controller are shown in Figs. 4-6.
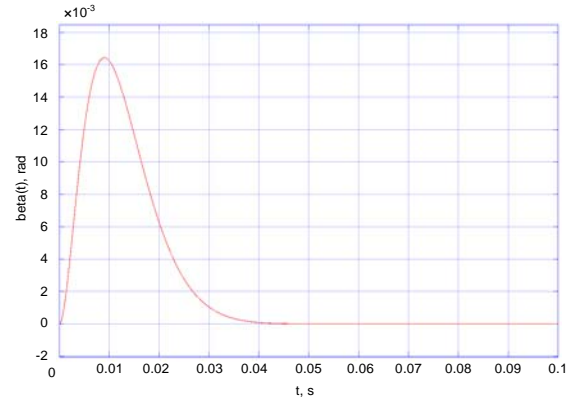

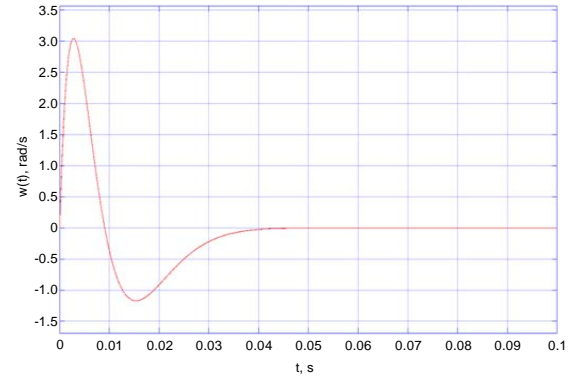Fig. 1. Changes in the angle of rotation of the sensing element

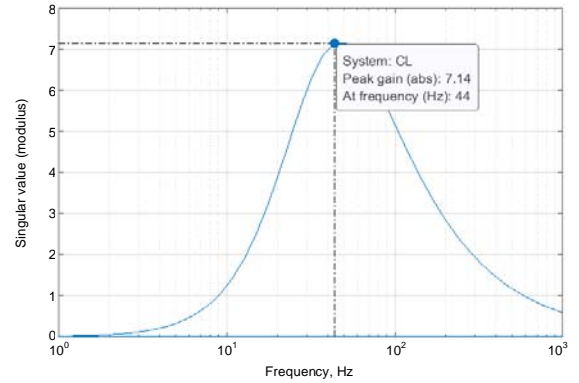Fig. 2. Changes in the angular velocity of the sensing element

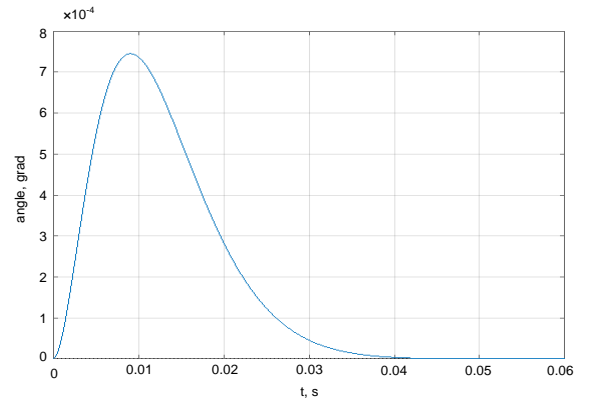Fig. 3. Changes in the singular value of the closed system

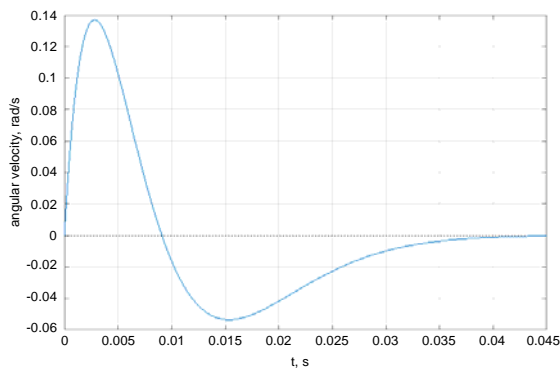Fig. 4. Changes in the angle of rotation of the sensing element

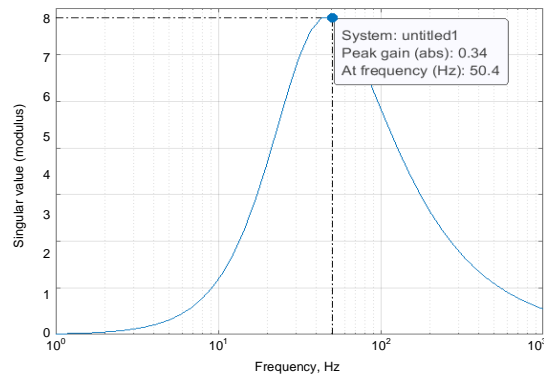Fig. 5. Changes in the angular velocity of the sensing element

Fig. 6. Changes in the singular value of the closed system

For the dynamic modal combined controller, the norm H_2 = 113.80. For the static modal combined controller, the norm H_2 = 43.89.

In the case of controller synthesis through minimizing the norms H_2/H_∞, the poles of a closed system of a pendulum compensation accelerometer lying in the required D-region were obtained.

A static modal combined controller based on H_2/H_∞ provides high performance and requirements for control quality parameters with minimal energy costs compared to other controllers.

V. CONCLUSION

The results allow drawing the following conclusions:

- a static modal combined controller based on H_2/H_∞ provides high performance and requirements for control quality parameters of a pendulum compensation accelerometer sensor with minimal energy costs compared to other controllers;

- controllers synthesised through minimizing individual norms H_2 and H_∞ do not provide high performance and requirements for control quality parameters of a pendulum compensation accelerometer sensor;

- the most resource-demanding ones during synthesis are the dynamic controller when minimizing the norm H_2 and the dynamic controller when minimizing the norm H_∞, obtained by solving the Riccati II equation;

- poles of the closed-loop pendulum accelerometer system with controllers synthesised through minimizing the norm H_2/H_∞ are located in the required D-region;

- reducing the norm H_2 will lead to an increase in the norm H_∞ and vice versa, which is a consequence of the contradiction between the energy spent on suppressing external disturbances and the worst transfer coefficient from the disturbing influence to the controlled parameter.

REFERENCES

[1] Nikiforov, V.M., Smirnov, E.S., YUrlov, F.A., CHajkovskij, M.M., Zolotuhin, S.S., Expanded mathematical model of a pendulum accelerometer with a flint suspension, Trudy FGUP "NPTsAP". Systemy i pribory upravleniia, ISSN 1991-5950, 2016, no. 1, pp. 27-35.

[2] Nikiforov, V.M., Gusev, A.A., Zolotukhin, S.S, Zhukova, T.A., Nizhegorodov, A.A., Identification of pendulous accelerometer mathematical model taking into account parametric uncertainty.//24th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), St. Petersburg, CSRI Elektropribor, 2017, pp. 329–332.

[3] Nikiforov, V.M., Gusev, A.A., Zolotukhin, S.S, Zhukova, T.A., Nizhegorodov, A.A., Synthesis of a pendulous accelerometer feedback controller using linear matrix inequalities and construction of invariant ellipsoids.//24th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), St. Petersburg, CSRI Elektropribor, 2017, pp. 326-328.

[4] Nikiforov, V.M., CHajkovskij, M.M., Gusev, A.A. Andreev, K.A., Anohin, A.S., Stihareva, N.P. Improving the quality of the transient process of the compensation pendulum accelerometer with LMI control.//25th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), St. Petersburg, CSRI Elektropribor, 2020", pp. 310-312.

Derivation of Dynamics Equations for the Hemispherical Resonator of a Solid-State Wave Gyroscope Using the D'Alembert-Suslov Method

V.M. Kotlov

State Research Institute of Aviation Systems
Moscow, Russia
e-mail: vadimkot366@yandex.ru

S.E. Perelyaev

Ishlinsky Institute for Problems in Mechanics, RAS
Moscow, Russia
e-mail: ipm@ipmnet.ru

Abstract—**The paper considers the transition from the theory of the ring wave solid-state gyroscope developed in [1] to the theory of the hemspherical wave solid-state gyroscope, based on the monograph [2], including the condition of inextensibility of the middle surface of the quartz hemisphere. To obtain inertial information, we replace the new physical principle introduced in [1] of the inertial properties of elastic waves with a more general principle of inertial motion of elastic deformations of a hemispherical resonator. The latter follow from the solution of synthesized differential equations of dynamics.**

Keywords—holonomic system, nonholonomic system, solid-state wave gyroscope (SWG), angular rate sensor (ARS), integrating the SWG.

I. INTRODUCTION

In the presented study, the problem of synthesis of dynamic equations of a uniaxial hemispherical quartz resonator of a solid-state wave gyroscope (SWG) is stated and solved.

In order to derive dynamic equations, we used the well-known method proposed by V.F. Zhuravlev [1], modified for an SWG with a ring resonator and presented by the authors at the 30th St. Petersburg International Conference on Integrated Navigation Systems in 2023 [3].

In this study, a system of modified differential dynamical equations of motion of a hemispherical resonator is derived using a new kind of waves (solitons) as bending vibrations of the hemispherical SWG resonator itself, which can serve as an inertial information sensor like a conventional inert solid.

The well-known classical method of V.F. Zhuravlev as applied to the theory of a hemispherical resonator of the SWG is considered in the same formulation as in the classical work on the theory of a thin ring resonator [1], i.e., under the condition of inextensibility of the middle surface of a hemispherical quartz resonator based on the fundamental linear theory of shells by S.P. Timoshenko [4].

The equations for the dynamics of motion of a hemispherical SWG are obtained on the basis of classical holonomic mechanics with the use of a simple Hooke's law: the resonator material elastically bends in proportion to the applied force and is described by a system of six second-order differential equations each.

From the solution of the equations of motion obtained for the kinetic energy of a hemispherical resonator, it follows that elastic deformations move along the resonator at a constant speed relative to the fixed axes like a rigid body without changing its configuration. This phenomenon in continuous media is called "soliton", or solitary wave. Moving elastic deformations under the influence of overloads behave like point solid bodies.

The time moments of the intersection of "solitons" with points of a rotating resonator, spaced from its center at a known distance, characterize the angle of rotation of the resonator base relative to the fixed axis. To obtain inertial information, the physical principle introduced in the classical work (inertial properties of elastic waves) can be replaced with a more general one – the principle of inertial properties of a new type of elastic deformation waves of the resonator (solitons).

II. FORMULATION OF THE PROBLEM

The purpose of this work is to synthesize a system of dynamic equations of motion of a hemispherical resonator using a new type of mechanical waves (solitons). In the classic works by D.M. Klimov and V.F. Zhuravlev [1,2], the equations for the dynamics of motion of a hemispherical SWG were obtained in the form of the well-known equations of the mechanician V.V. Novozhilov, who used the theory of thin-walled elastic shells. A plate (shell) is an elastic body whose size, called the thickness of the shell, in one of the directions is significantly smaller than the other two. The surface of the shell passing through the middle of its thickness will be called the median surface. The presence of a degenerate coordinate in a thin plate, that is, a coordinate along which the dimensions of the elastic body are small, makes it possible to synthesize approximate mathematical models with a number of independent variables that is one less than is usually accepted in the original problem of the classical theory of elasticity.

In the classical linear theory of shells by the mechanician S.P. Timoshenko: "We expand the vector function *R*, which determines the position of a point on an equidistant surface, into a Taylor series in the neighborhood of the middle surface" [4]. When moving from a thin ring to a hemisphere, you can use the Hamilton–Ostrogradsky variational equation, which allows you to apply the results obtained in [3] for a thin elastic ring.

III. DERIVATION OF THE EQUATIONS OF MOTION OF A HEMISPHERICAL RESONATOR USING THE LAGRANGE METHOD

To synthesize the equations of motion of a hemispherical resonator using the Lagrange method, it is necessary to determine the kinetic energy density *T* and potential energy

density U of a hemispherical quartz resonator and write out the corresponding coupling equation (substitution neg). The equations of motion of the hemisphere follow from the condition that the variation of the integral of the resulting density of the Lagrange function L is equal to zero:

$$\int_{\theta 1}^{\theta 2} \int_{\varphi 1}^{\varphi 2} \int_{t1}^{t2} \delta[L] dt d\varphi\, d\theta =$$
$$= \int_{\theta 1}^{\theta 2} \int_{\varphi 1}^{\varphi 2} \int_{t1}^{t2} \delta[T - U] dt d\varphi\, d\theta = 0 \qquad (3.1)$$

where: d[ ] - operation of taking a differential, δ[ ] - operation of taking a variation; L= T-U, density T – kinetic, U – potential energy density of the hemispherical resonator.

In accordance with Hamilton's principle of stationary action [5], the equations of motion of the hemisphere follow from the condition that the variation of the integral of the resulting density of the Lagrange function (3.1) is equal to zero.

Let us write down the expressions for the density of kinetic and potential energies of the hemispherical resonator under study:

$$2T = (\dot{r} - v\,\Omega\,cos[\theta])^2 \; +$$
$$+(\dot{v} + (r + R)\,\Omega\,cos[\theta] - u\,\Omega\,sin[\theta])^2 +$$
$$+(\dot{u} + v\,\Omega\,sin[\theta])^2 \qquad (3.2)$$
$$2U = ((rs - v\,cos[\theta]) +$$
$$+(vs + ((r + R)\,cos[\theta] - u\,sin[\theta]))^2 + (us +$$
$$v\,sin[\theta])^2) \qquad (3.3)$$

where: $\dot{v} = \frac{\partial}{\partial \tau} v$, $\dot{r} = \frac{\partial}{\partial \tau} r$, $\dot{u} = \frac{\partial}{\partial \tau} u$, $vs = \frac{\partial}{\partial \varphi} v$, $rs = \frac{\partial}{\partial \varphi} r$, $us = \frac{\partial}{\partial \varphi} u$; R, φ, θ – radius and angular coordinates of the hemisphere; r, v, u – small deformations of a hemispherical point along the radius and perpendicular to it (see Appendix 1, Fig. 1); $\tau = \kappa t$– normalized time; $\kappa^2 = \frac{EI}{\rho SR^4}$, ρ, E– density and Young's modulus of the hemisphere material;

Here I – moment of inertia of the cross-section of the shell of a hemispherical resonator relative to the neutral axis; S – cross-sectional area of the shell of a hemispherical resonator.

Using simple Hooke's law, it can be shown that the material of a hemispherical resonator bends elastically in proportion to the force applied to it. In this case, the coupling equation is the classical kinematic condition for the inextensibility of the middle surface of the hemispherical resonator under consideration:

$$(rs - v\,cos[\theta])^2 + (vs + ((r + R)\,cos[\theta] -$$
$$-u\,sin[\theta]))^2 + (us + v\,sin[\theta])^2 = (R\,cos[\theta])^2 \quad (3.4)$$

where

$$rs = \frac{\partial r[\tau, \varphi, \theta]}{\partial \varphi}, vs = \frac{\partial v[\tau, \varphi, \theta]}{\partial \varphi}, us = \frac{\partial u[\tau, \varphi, \theta]}{\partial \varphi}$$

By introducing new variables $\{I_1, J_1, K_1\}$, we represent the inextensibility condition in the following linear form

$$(rs - v\,cos[\theta]) = I_1$$
$$vs + ((r + R)\,cos[\theta] - u\,sin[\theta]) = J_1 \quad (3.5)$$
$$us + v\,sin[\theta] = K_1$$

In the new variables introduced above, the inextensibility condition (3.4) will look like

$$I_1{}^2 + J_1{}^2 + K_1{}^2 = (R\,cos[\theta])^2 \qquad (3.6)$$

This condition will be fulfilled identically for any values of the new angles $\Psi = \Psi[\varphi]$, $Q = Q[\varphi, \theta]$ associated with the variables $\{I_1, J_1, K_1\}$ as follows:

$$\begin{cases} I_1 \to \rho_1\,sin[\Psi]sin[Q] \\ J_1 \to \rho_1\,sin[\Psi]cos[Q] \\ K_1 \to \rho_1\,cos[\Psi] \end{cases} \qquad (3.7)$$

Putting together relations (3.5, 3.7), we obtain that the inextensibility condition (3.4) is given in linear form by the substitution neg:

$$\text{neg} =$$
$$= \begin{cases} rs \to v\,Cos[\theta] + \rho_1\,Sin[\Psi]Sin[Q] \\ vs \to -((r + R)Cos[\theta] - uSin[\theta]) + \rho_1 Sin[\Psi]Cos[Q] \\ us \to -v\,Sin[\theta] + \rho_1\,Cos[\Psi] \end{cases}$$
$$(3.8)$$

where: $\rho_1 = R\,cos[\theta]$.

In accordance with the classical linear theory of shells by S.P Timoshenko, we expand expression (3.3) for potential energy U in the Taylor series in the vicinity of the middle surface. When the hemispherical resonator itself deviates from the middle surface, the potential energy density U will change along the angle φ:

$$\begin{cases} r_{s\varphi} - v_\varphi\,cos[\theta] = I_{1\varphi} \\ v_{s\varphi} + (r_\varphi\,cos[\theta] - u_\varphi\,sin[\theta]) = J_{1\varphi} \\ u_{s\varphi} + v_\varphi\,sin[\theta] = K_{1\varphi} \end{cases} \qquad (3.9)$$

where respectively: $r_{s\varphi} = \frac{\partial rs}{\partial \varphi}$, $v_{s\varphi} = \frac{\partial vs}{\partial \varphi}$, $u_{s\varphi} = \frac{\partial us}{\partial \varphi}$;

$$I_{1\varphi} = \frac{\partial}{\partial \varphi} I_1, J_{1\varphi} = \frac{\partial}{\partial \varphi} J_1, \; K_{1\varphi} = \frac{\partial}{\partial \varphi} K_1$$

When deviating from the middle surface of a hemispherical resonator, the potential energy density U will change along the angle θ:

$$\begin{cases} r_{s\theta} - v_\theta\,cos[\theta] + v_\theta\,sin[\theta] = I_{1\theta} \\ v_{s\theta} + (r_\theta\,cos[\theta] - u_\theta sin[\theta]) - \\ \quad -(r_\theta\,sin[\theta] + u_\theta cos[\theta]) = J_{1\theta} \\ u_{s\theta} + v_\theta sin[\theta] + v_\theta cos[\theta] = K_{1\theta} \end{cases} \qquad (3.10)$$

where: $r_{s\theta} = \frac{\partial rs}{\partial \theta}$, $v_{s\theta} = \frac{\partial vs}{\partial \theta}$, $u_{s\theta} = \frac{\partial us}{\partial \theta}$,

$$I_{1\theta} = \frac{\partial}{\partial \theta} I_1, \; J_{1\theta} = \frac{\partial}{\partial \theta} J_1, K_{1\theta} = \frac{\partial}{\partial \theta} K_1.$$

In the vicinity of the middle surface of the shell of a hemispherical resonator, the potential energy U will have the following analytical form:

$$2U = \delta^2 (Rcos[\theta])^2 + \kappa^2 (I_{1\varphi}{}^2 + J_{1\varphi}{}^2 + K_{1\varphi}{}^2) +$$
$$+ \kappa^2 (I_{1\theta}{}^2 + J_{1\theta}{}^2 + K_{1\theta}{}^2) \qquad (3.11)$$

where: $\delta^2 = \frac{E}{\rho R^2}$, $\kappa^2 = \frac{EI}{\rho SR^4}$.

IV. IV. EQUATIONS OF DYNAMICS OF A HEMISPHERICAL RESONATOR

Using the D'Alembert method, modified by the mechanician G.K. Suslov [6] for systems with non-integrable couplings, we obtain the required system of equations for the dynamics of a hemispherical resonator. The dynamics equations of such a resonator are a system of six second-order differential equations each (see Appendix 2). The first three differential equations of the system - second-order equations, provided that the angles θ, φ are constant, have the following form:

$$\begin{cases} \dfrac{\ddot{r}}{\cos[\theta]} = \left(2\dot{v}\Omega + v\dot{\Omega} + \Omega^2\left((r+R)\cos[\theta] - u\sin[\theta]\right)\right) \\ \dfrac{\ddot{u}}{\sin[\theta]} = -\left(2\dot{v}\Omega + v\dot{\Omega} + \Omega^2\left((r+R)\cos[\theta] - u\sin[\theta]\right)\right) \\ \ddot{v} = \left(v\Omega^2 - 2\Omega(\dot{r}\cos[\theta] - \dot{u}\sin[\theta]) - \dot{\Omega}\begin{pmatrix}(r+R)\cos[\theta] \\ -u\sin[\theta]\end{pmatrix}\right) \end{cases} \quad (4.1)$$

where: $\ddot{r} = \dfrac{\partial}{\partial\tau}\dot{r},\ \ddot{u} = \dfrac{\partial}{\partial\tau}\dot{u}, \ddot{v} = \dfrac{\partial}{\partial\tau}\dot{v},\ \dot{\Omega} = \dfrac{\partial}{\partial\tau}\Omega$

The second two differential equations of the second order, provided that the variables τ and θ are constant, will be look like this:

$$\begin{cases} \dfrac{d}{d\varphi}Q = Q_s,\ \ \dfrac{d}{d\varphi}\Psi = \Psi_s \\ \dfrac{d}{d\varphi}(Q_s \sin[\Psi]^2) = 0, \dfrac{d}{d\varphi}\left(\Psi_s^{\,2} + Q_s^{\,2}\sin[\Psi]^2\right) = 0 \end{cases} \quad (4.2)$$

The third differential equation of the second order for constant values of the variables τ, φ has the form:

$$\begin{cases} \dfrac{d}{d\theta}Q = Q_\theta, \\ \dfrac{d[Q_\theta \cos[\theta]^2]}{d\theta} = 0 \end{cases} \quad (4.3)$$

V. Solution of the System of Equations for the Dynamics of a Hemispherical Resonator

Solving a system of equations of the form (4.1), (4.2), (4.3) gives the wave motion of a hemispherical resonator:

$$\{r[\tau, \Delta\varphi, \Delta\theta] + R, v[\tau, \Delta\varphi, \Delta\theta], u[\tau, \Delta\varphi, \Delta\theta]\} =$$
$$= \{\Delta X_{no} + V_{xn}\tau, \Delta Y_{no} + V_{yn}\tau, \Delta Z_{no} + V_{zn}\tau\}\cdot$$
$$\cdot\begin{pmatrix} \cos[\varphi + \psi[t]] & -\sin[\varphi + \psi[t]] & 0 \\ \sin[\varphi + \psi[t]] & \cos[\varphi + \psi[t]] & 0 \\ 0 & 0 & 1 \end{pmatrix}\cdot \boldsymbol{C} +$$
$$+ f[\Delta\varphi, \Delta\theta] \quad (5.1)$$

where: $f[\Delta\varphi, \Delta\theta] = \{r_o + R_o, v_o, u_o\}\cdot Exp[C\ \Delta\varphi\ ] +$

$$+ \rho_1 \int_0^{\Delta\varphi} (\{\sin[\Psi]\sin[Q], \sin[\Psi]\cos[Q], \cos[\Psi]\}\cdot$$

$$\cdot Exp[C\ \Delta\varphi_2\ ])\cdot d\Delta\varphi_2 \quad (5.2)$$

$$\rho_1 = R\,Cos[\theta],\ \boldsymbol{C} = \begin{pmatrix} 0 & -\cos[\theta] & 0 \\ \cos[\theta] & 0 & -\sin[\theta] \\ 0 & \sin[\theta] & 0 \end{pmatrix}$$

here: $\Psi = \Psi[\Delta\varphi - \Delta\varphi_2], Q = Q[\Delta\varphi - \Delta\varphi_2, \Delta\theta]$;

$r_o, v_o, u_o -$ initial parameter values;

$\{\Delta X_{no}, \Delta Y_{no}, \Delta Z_{no}, V_{xn}, V_{yn}, V_{zn}\} -$ constants.

VI. Conclusions

From the differential equations for the dynamics of motion of a hemispherical resonator synthesized in this study, it follows that is possible

- to develop both an inertial rotation angle sensor (integrating gyroscope) and an ARS without excitation of mechanical standing waves of elastic vibrations of the edge of a hemispherical resonator;

- to use the obtained analytical expressions to process the measured parameters to determine the angular velocity and/or the integral of the angular velocity of the gyroscope base;

- to form solitary mechanical standing waves (solitons) as bending vibrations of the edge of a hemispherical resonator, which can serve as a sensor of inertial information, like a conventional inert solid body.

The new results from the theory of a uniaxial hemispherical SWG resonator include:

a) the proof of the integrability of the kinematic condition of inextensibility of the hemispherical resonator's middle surface and simplification of the approach to deriving differential equations for the dynamics of motion of a uniaxial inertial sensor;

b) confirmation of the holonomy condition of inextensibility of the hemispherical resonator's middle surface, which makes it possible to obtain equations for the dynamics of its motion using the symbolic mathematics program package for engineering calculations "Wolfram Research Mathematica" 5.2.0 (see Appendix 1);

c) obtaining equations for the dynamics of motion of a uniaxial hemispherical SWG in the form of six ordinary second-order differential equations, independent of each other in time and angles, which makes it possible to practically use ordinary differential equations of the Lynch-Zhuravlev type;

d) the inertial information that determines the gyroscopic effect depends on the inertial motion of the deformations of the hemispherical resonator, which provides a justification for the inertial properties of standing waves.

From the analysis of the differential equations for the dynamics of motion of a hemispherical resonator obtained in this study, it follows that is possible

- to develop a new inertial rotation angle sensor (integrating gyroscope) and angular velocity sensor (AVS) without excitation of mechanical standing waves of elastic vibrations of the resonator edge;

- to form solitary mechanical waves (solitons) as bending vibrations of the hemispherical resonator edge, which can serve as a sensor of inertial information, like an ordinary inert solid body.

References

[1] Zhuravlev V.F., Klimov D.M. Wave solid-state gyroscope. - Moscow: Nauka, 1985.

[2] Klimov D.M., Zhuravlev V.F., Zhbanov Y.K. Quartz hemispherical resonator (wave solid-state gyroscope). - Moscow: Kim L.A., 2017.

[3] Kotlov V.M., Perelyaev S.E. Derivation of the Equations of Motion Dynamics of Ring Wave Solid-State and Pendulum Gyroscopes by the Method of Holonomic Mechanics // 30th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS) 2023,pp.456-459.

[4] Rickards R.B. Finite element method in the theory of shells and plates - Riga: Zinatne, 1988.

[5] Elsholtz L.E. Differential equations and calculus of variations. - Moscow: Nauka, 1965.

[6] Suslov G.K. Theoretical Mechanics. M.,L: GITTL, 1946.

Appendix 1

Derivation of the system of equations (4.1) when using a computer system of analytical transformations [*Wolfram Research Mathematica* 5.2.0].

Substituting kinetic energy into the computer program for calculating Euler's equation "Euler Equations"

$$2T = (r1 - v\,\Omega\,Cos[\theta])^2$$
$$+ \left(v1 + \Omega\big((r + R)Cos[\theta] - uSin[\theta]\big)\right)^2$$
$$+ (u1 + v\Omega Sin[\theta])^2$$

when replacing variables:

$r\to r[t, \varphi, \theta], v \to v[t, \varphi, \theta],$

$u\to u[t, \varphi, \theta],$

$r1\to r^{(1,0,0)}[t, \varphi, \theta], v1 \to v^{(1,0,0)}[t, \varphi, \theta],$

$u1\to u^{(1,0,0)}[t, \varphi, \theta]$

using the program for calculating the Euler equation:

«Calculus`Variational Methods`Euler Equations»

$$\frac{1}{2}\big((r^{(1,0,0)}[t, \varphi, \theta] - v[t, \varphi]\Omega[t]Cos[\theta])^2 +$$

$$+(v^{(1,0,0)}[t, \varphi, \theta] + \Omega[t]((r[t, \varphi, \theta] + R)Cos[\theta] -$$

$$-u[t, \varphi, \theta]Sin[\theta]))^2 + (u^{(1,0,0)}[t, \varphi, \theta] +$$

$$+v[t, \varphi, \theta]\Omega[t]Sin[\theta])^2),$$

$$\{r[t, \varphi, \theta], v[t, \varphi, \theta], u[t, \varphi, \theta]\}, \{t, \varphi, \theta\}]$$

we obtain three differential equations, which, when changing variables:

$$r[t, \varphi, \theta] \to r, v[t, \varphi, \theta] \to v, u[t, \varphi, \theta] \to u$$
$$r^{(1,0,0)}[t, \varphi, \theta] \to r1, v^{(1,0,0)}[t, \varphi, \theta] \to v1,$$
$$u^{(1,0,0)}[t, \varphi, \theta] \to u1$$
$$r^{(2,0,0)}[t, \varphi, \theta] \to r2, v^{(2,0,0)}[t, \varphi, \theta] \to v2,$$
$$u^{(2,0,0)}[t, \varphi, \theta] \to u2$$

give the required three expressions of the Euler equations:

$$\{-r2 + 2v1\Omega Cos[\theta] + v\Omega1Cos[\theta] + \Omega^2 Cos[\theta](rCos[\theta] + RCos[\theta] - uSin[\theta]) = 0;$$

$$-u2 - 2v1\Omega Sin[\theta] - v\Omega1Sin[\theta] - \Omega^2 Sin[\theta](rCos[\theta] + RCos[\theta] - uSin[\theta]) = 0\};$$

$$-v2 + v\Omega^2 - 2r1\Omega Cos[\theta] + 2u1\Omega Sin[\theta]$$
$$- \Omega1(rCos[\theta] + RCos[\theta] - uSin[\theta])$$
$$= 0$$

The resulting three equations are considered under the condition

$\varphi = Const, \theta = Const,$

in which they have the following analytical form:

$$\frac{d}{dt1}r = r1, \frac{d}{dt1}v = v1, \frac{d}{dt1}u = u1$$

$$\frac{1}{Cos[\theta]}\frac{d}{dt1}r1 = (2v1\,\Omega + v\,\Omega1 + (r + R)\Omega^2),$$

$$\frac{1}{Sin[\theta]}\frac{d}{dt1}u1 = -(2v1\,\Omega + v\,\Omega1 + (r + R)\Omega^2)$$

$$\frac{d}{dt1}v1 = (v\Omega^2 - 2\Omega(r1Cos[\theta] - u1Sin[\theta]) - \Omega1((r$$
$$+ R)Cos[\theta] - uSin[\theta]))$$

where (see Zhuravlev [1] page 16 f. (1.23) и f. (1.24)):

$$t1 = \kappa1\,t, \qquad \Omega[t1] = \frac{\Omega[t]}{\kappa1}, \Omega1[t1] = \frac{d}{dt1}\Omega[t1]$$

Appendix 2

Derivation of the analytical expression for the kinetic energy of movement of deformations of a hemispherical resonator
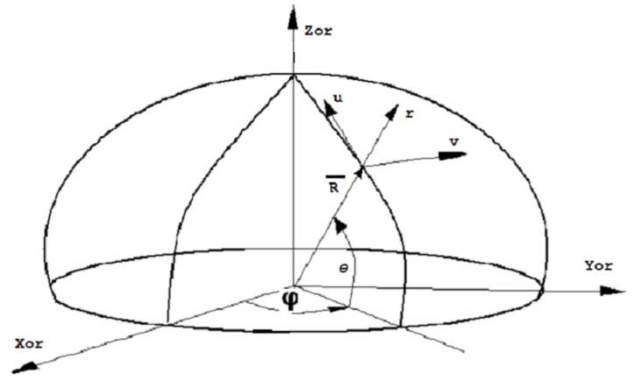

Fig. 1.

$\overline{Ror}$ — radius vector of points of a hemispherical resonator

in projections on the axis of the resonator base:

$$\overline{Ror} = \{\underset{Xor}{RCos[\theta]Cos[\varphi]}, \underset{Yor}{RCos[\theta]Sin[\varphi]}, \underset{Zor}{RSin[\theta]}\} \quad (A2.1)$$

In accordance with Fig. 1, the deformation $\overline{dRor}$ of the resonator along the axes $\{\bar{r}, \bar{v}, \bar{u}\}$ will be:

$$\overline{dRor} = \{r, v, u\}.\begin{pmatrix} Cos[\theta]Cos[\varphi] & Cos[\theta]Sin[\varphi] & Sin[\theta] \\ -Sin[\varphi] & Cos[\varphi] & 0 \\ -Cos[\varphi]Sin[\theta] & -Sin[\theta]Sin[\varphi] & Cos[\theta] \end{pmatrix}$$
(A2.2)

Assuming that the radius vector of the deformed resonator along the axes of the resonator base will be:

$$\overline{Rr} = \overline{Ror} + \overline{dRor} \qquad (A2.3)$$

And, taking into account that (A2.1) can be represented as

$$\overline{Ror} = \{R, 0,0\}.$$

$$\cdot\begin{pmatrix} Cos[\theta]Cos[\varphi] & Cos[\theta]Sin[\varphi] & Sin[\theta] \\ -Sin[\varphi] & Cos[\varphi] & 0 \\ -Cos[\varphi]Sin[\theta] & -Sin[\theta]Sin[\varphi] & Cos[\theta] \end{pmatrix} \quad (A2.4)$$

on the right side of (A2.3) we obtain the expression:

$$\overline{Rr} = \{R + r, v, u\}.$$

$$\cdot\begin{pmatrix} Cos[\theta]Cos[\varphi] & Cos[\theta]Sin[\varphi] & Sin[\theta] \\ -Sin[\varphi] & Cos[\varphi] & 0 \\ -Cos[\varphi]Sin[\theta] & -Sin[\theta]Sin[\varphi] & Cos[\theta] \end{pmatrix} \quad (A2.5)$$

The rotation of the axes of the resonator base $\{Xor, Yor, Zor\}$ relative to the inertial axes $\{Xn, Yn, Zn\}$ around the Zn axis is given by the matrix

\	Xn	Yn	Zn	
Xor	$Cos[\psi[t]]$	$Sin[\psi[t]]$	0	
Yor	$-Sin[\psi[t]]$	$Cos[\psi[t]]$	0	(A2.6)
Zor	0	0	1	

Let the radius vector of the deformed resonator in projections onto the inertial axes be:

$$\overline{Rn} = \{\overset{Xn}{\Delta Xn}, \quad \overset{Yn}{\Delta Yn}, \quad \overset{Zn}{\Delta Zn}\} \qquad (A2.7)$$

The connection between $\overline{Rn}$ and $\overline{Rr}$ according to formulas (A2.3, A2.7) is:

$$\overline{Rn} = \overline{Rr}.$$

$$\cdot \begin{pmatrix} Cos[\psi[t]] & Sin[\psi[t]] & Sin[\theta] \\ -Sin[\psi[t]] & Cos[\psi[t]] & 0 \\ 0 & 0 & 1 \end{pmatrix} = \{R + r, v, u\}.$$

$$\cdot \begin{pmatrix} Cos[\psi[t]] & Sin[\psi[t]] & 0 \\ -Sin[\psi[t]] & Cos[\psi[t]] & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad (A2.8)$$

Considering that the matrix

$$\begin{pmatrix} Cos[\theta]Cos[\varphi] & Cos[\theta]Sin[\varphi] & Sin[\theta] \\ -Sin[\varphi] & Cos[\varphi] & 0 \\ -Cos[\varphi]Sin[\theta] & -Sin[\theta]Sin[\varphi] & Cos[\theta] \end{pmatrix} =$$

$$= \begin{pmatrix} Cos[\theta] & 0 & Sin[\theta] \\ 0 & 1 & 0 \\ -Sin[\theta] & 0 & Cos[\theta] \end{pmatrix} \cdot \begin{pmatrix} Cos[\varphi] & Sin[\varphi] & 0 \\ -Sin[\varphi] & Cos[\varphi] & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

we obtain that $\overline{Rn}$ according to (A2.7,A2.8) will be:

$$\overline{Rn} = \{\Delta Xn, \Delta Yn, \Delta Zn\} =$$

$$\{R + r, v, u\} \cdot \begin{pmatrix} Cos[\theta] & 0 & Sin[\theta] \\ 0 & 1 & 0 \\ -Sin[\theta] & 0 & Cos[\theta] \end{pmatrix}.$$

$$\cdot \begin{pmatrix} Cos[\varphi + \psi[t]] & Sin[\varphi + \psi[t]] & 0 \\ -Sin[\varphi + \psi[t]] & Cos[\varphi + \psi[t]] & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad (A2.9)$$

The speed of movement of the points of the deformed resonator relative to the inertial axes is obtained by differentiating $\overline{Rn}$ with respect to time:

$$\frac{d}{dt}\overline{Rn} = \left\{\frac{d}{dt}\Delta Xn, \frac{d}{dt}\Delta Yn, \frac{d}{dt}\Delta Zn\right\} = \{Vxn, Vyn, Vzn\} \qquad (A2.10)$$

Taking into account the analytical expression (A2.10), we obtain:

$$\frac{d}{dt}\overline{Rn} = \{r1, v1, u1\}.$$

$$\cdot \begin{pmatrix} Cos[\theta] & 0 & Sin[\theta] \\ 0 & 1 & 0 \\ -Sin[\theta] & 0 & Cos[\theta] \end{pmatrix} \cdot \begin{pmatrix} Cos[\varphi + \psi[t]] & Sin[\varphi + \psi[t]] & 0 \\ -Sin[\varphi + \psi[t]] & Cos[\varphi + \psi[t]] & 0 \\ 0 & 0 & 1 \end{pmatrix} +$$

$$+ \{R + r, v, u\} \cdot \begin{pmatrix} Cos[\theta] & 0 & Sin[\theta] \\ 0 & 1 & 0 \\ -Sin[\theta] & 0 & Cos[\theta] \end{pmatrix} \cdot \begin{pmatrix} Cos[\varphi + \psi[t]] & Sin[\varphi + \psi[t]] & 0 \\ -Sin[\varphi + \psi[t]] & Cos[\varphi + \psi[t]] & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\cdot \begin{pmatrix} 0 & \Omega & 0 \\ -\Omega & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

From where the desired expression of kinetic energy will be equal to the:

$$2T = Vxn^2 + Vyn^2 + Vzn^2 = (r1 - v\,\Omega\,Cos[\theta])^2 +$$

$$+ (v1 + \Omega\,((r + R)\,Cos[\theta] - u\,Sin[\theta]))^2 +$$

$$+ (u1 + v\,\Omega\,Sin[\theta])^2$$

Relative navigation of aircraft using an optical surveillance system

Alexander Knyazhsky

Department of Aerospace
Measuring and Computing Systems
St. Petersburg State University of
Aerospace Instrumentation
St. Petersburg, Russian Federation
knjagskij@mail.ru

Alexander Nebylov

Department of Aerospace
Measuring and Computing Systems
St. Petersburg State University of
Aerospace Instrumentation
St. Petersburg, Russian Federation
nebylov@aanet.ru

Vladimir Nebylov

Department of Aerospace
Measuring and Computing Systems
St. Petersburg State University of
Aerospace Instrumentation
St. Petersburg, Russian Federation
vladnebylov@gmail.com

Abstract – **A principle of relative navigation of aircraft is proposed, which makes it possible to determine the relative position of aircraft by processing the image, taking into account the previously known geometry of the location of light sources on the body of the observed aircraft. The performance of the proposed principle was confirmed by computer simulation.**

Keywords— observation, optical system, docking, aircraft, relative motion

I. INTRODUCTION

The principle of operation of an optical observation system for the implementation of relative navigation of aircraft is proposed, which consists in installing LEDs of various colors on the body of the observed aircraft, specifying the geometry of the location of these LEDs and their colors in the control system of the observing aircraft, obtaining a video image, extracting coordinates from it and colors of the LEDs (in the coordinate system of a two-dimensional image), comparing the geometry of the observed location of the LEDs with the given one and calculating, based on their deviations and the coordinate scale, the position of the observing aircraft relative to the observed aircraft.

II. OPERATING PRINCIPLE OF OPTICAL SURVEILLANCE SYSTEM

The proposed principle can be implemented as follows:

1. set the colors and other characteristics of the glow of the LEDs, by which the location of the observing aircraft is determined, relative to the observed aircraft, and the geometry of their location, for example, in such a way that all points have different colors and glow intensity, flicker frequency, or other glow characteristics determined infrared camera. To make LED color identification easier, the optical camera uses rainbow colors;

2. obtain a video image of the observed aircraft, for example, by combining images obtained using video cameras operating in the optical and infrared ranges;

3. determine the colors and coordinates of the LEDs in the aircraft image, for example, using neural network image processing algorithms;

4. calculate the distances in the image between the LEDs and the angles of inclination of the segments between them, relative to the horizontal axis;

5. using the calculated angles of inclination of the segments between the LEDs in the image, find the distance

to the observed aircraft and the angles of rotation of the observing aircraft relative to the observed aircraft.

Taking into account the fact that the video camera must generate a one-to-one relationship between the object and the image, the distance between the aircraft is calculated as the product of the focal length and the ratio of the reference distance to the observed one. The reference distance is calculated taking into account the tilt angles in the image, calculated using known trigonometric formulas, based on the relative and observed positions of the LEDs.

III. DEVICE DESCRIPTION

Figures 1 and 2 show a device for implementing a method for docking an aircraft using an optical surveillance system and the following designations are introduced: 1 - platform; 2 - docking elements of the carrier aircraft; 3 - dampers; 4 - aircraft carrier control system; 5 - control panel; 6 - automatic control system for docking the docking aircraft with the carrier aircraft; 7 – receiver-transmitter of aircraft carrier signals; 8 - docking elements of the docked aircraft; 9 – control systems of the docked aircraft; 10 – signal transceiver on the docking aircraft; 11 - LEDs emitting specified colors; 12 - LED monitoring and control systems; 13 - optical video camera; 14 - infrared video camera; 15 - module for complex video image processing.
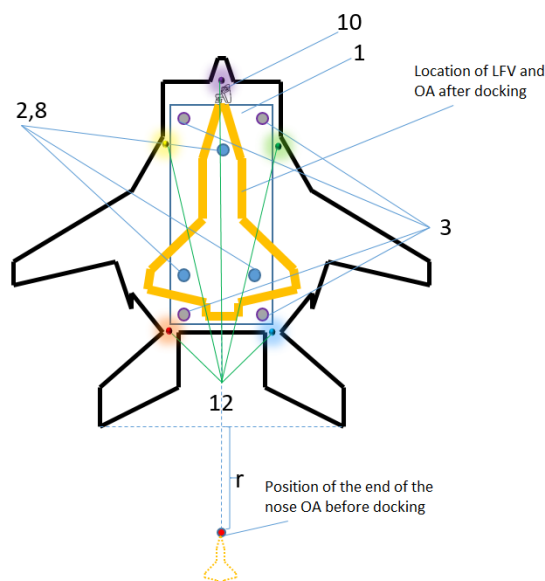

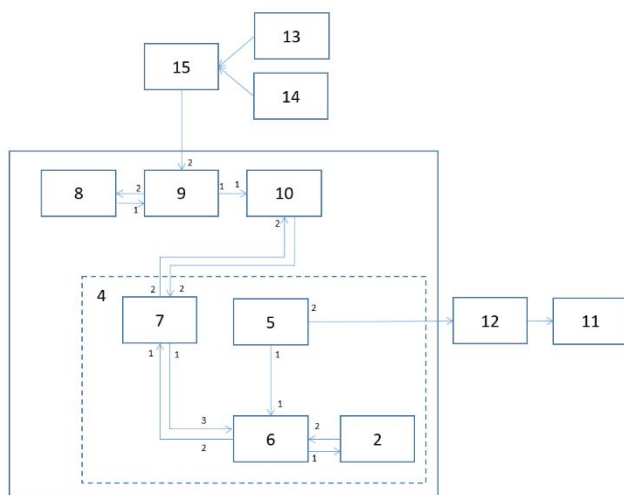Fig. 1. Device for implementing a method for docking OA using an optical surveillance system

Fig. 2. Scheme of device interaction for implementing a method for docking aircraft using an optical surveillance system

As docking elements on the carrier aircraft and the docked aircraft, you can use, for example, high-strength heat-resistant metal grips that automatically engage each other when approaching each other with a given accuracy from a certain direction. A damper must be located between the body of the vehicle and the grip of the docking element to absorb the shock from the docking of the aircraft. The fixation of the docking aircraft on the body of the carrier aircraft must be carried out in such a way that friction and heating between the bodies and docking elements do not destroy their structures. The automatic docking control system can be a computing device that, after receiving a docking start signal from the control panel, begins to receive a signal about the position of the docking aircraft, the location of the carrier aircraft, and when the spatial and angular coordinates of the docking aircraft relative to the carrier aircraft fall within the established ranges, automatically captures it using the docking elements of the carrier aircraft. The LEDs should glow in different colors so that they can be used to clearly determine the position of the docking aircraft relative to the carrier aircraft.

IV. SIMULATION

To assess the possibility of using the proposed operating principle of an optical observation system for relative navigation of an aircraft, computer simulation was carried out in Python. During the simulation, the problem of landing an aerospace aircraft on an aircraft with a dynamic principle of supporting, for example, an ekranoplan, was solved. Four LEDs were located at the rear of the aircraft, two more on the upper front of the fuselage, and two on the sides. At the same time, the LEDs located on the right side were duplicated so that they could be recognized by an infrared camera without determining the color. The simulation was carried out taking into account the principles of aircraft motion set forth in [1-3].

The dynamic characteristics of the aircraft were specified by limiting their angular velocities and their first derivatives, as well as acceleration during aerodynamic braking of the aircraft and the delay time of the control response. There was no need to establish upper limits on aircraft ground speeds

and accelerations due to the fact that only their reduction was allowed, with the exception of random fluctuations caused by control errors.

The following assumptions were made: the total error of control and external influence for the track angle, trajectory inclination angle and ground speed is distributed according to the Gaussian distribution law with zero mathematical expectation. The oscillations of the trajectory of a docked aircraft are greater in power and frequency of oscillations of the trajectory of the carrier aircraft, due to the lower mass of the docked aircraft. Since the wind affects both aircraft almost equally, the control error of the docking aircraft relative to the carrier aircraft caused by the wind can be neglected.

The coordinates of the location of the LEDs and their colors (codes) were specified relative to the center of mass of the carrier aircraft. It was assumed that the surveillance system reliably determines the location of the LEDs with a probability of 90%.

Model parameters: ground speed of the orbital aircraft (OA) and low-flying vehicle (LFV) during docking 90 – 120 m/s; LFV movement height is 5-8 m with a height fluctuation of 1 m; the speed of approach of the OA and the LFV before docking is 2-4 m/s; the total error of control and external influence of the OA along the path angle has a standard deviation of 1-4 degrees; the total error of control and external influence of the OA on ground speed has a standard deviation of 0.5-2 m/s; the total error of control and external influence of the LFV along the path angle has a standard deviation of 1-4 degrees; the total error of control and external influence of the LFV in terms of ground speed has a standard deviation of 0.5-2 m/s; the specified angle of inclination of the OA trajectory at the main landing site is 20 - 40 degrees; the specified angle of inclination of the OA trajectory during landing before entering a horizontal trajectory is 2-5 degrees.

The simulation was carried out by sequentially alternating the specified model parameters with a step of 0.1 degrees in the directional angle and 0.1 m/s in the ground speed. Thus, the model passed discretely through the entire state space.

Figures 3-6 present the graphs describing the implementation of the landing trajectory of the OA and the movement of the LFV with the following parameters: ground speed of the OA and LFV during docking 100 m/s; the speed of approach of the OA and the LFV before docking is 2 m/s; the total error of control and external influence of the OA along the path angle has a standard deviation of 3 degrees; the total error of control and external influence of the OA on ground speed has a standard deviation of 1 m/s; the total error of control and external influence of the LFV along the path angle has a standard deviation of 3 degrees; the total error of control and external influence of the LFV in terms of ground speed has a standard deviation of 1 m/s; the specified angle of inclination of the OA trajectory at the main landing site is 30 degrees; the specified angle of inclination of the OA trajectory during landing before entering the horizontal trajectory is 3 degrees.

Figures 3a and 3b show examples of changing the heights of the LFV and OA trajectories to carry out their docking. The LFV moves straight with a given track angle, ground speed and flight altitude. Before reaching the altitude of the

start of the docking maneuver (about 200 m), the OA is lowered at an angle of 30 degrees, after which, in order to increase the accuracy of docking and increase the time for aiming, it reduces the angle to 3 degrees. After reaching the docking altitude, the OA catches up with the LFV at a speed greater than the LFV speed by 2-4 m/s, trying to maintain the docking altitude. After capturing the OA by the manipulator and docking, the height of the OA is stabilized by a given value relative to the height of the LFV. In the first implementation option (Fig. 1a), the OA LFV caught up and docked within 10 s. During this time, the LFV and OA traveled about 1000 m (in the interval r = [3200, 4200] m). In the second option (Fig. 1b), the OA LFV caught up and their docking took approximately 5 seconds, during which time a distance of about 500 m was covered (in the interval r = [4200, 4800] m). Thus, the docking time and the approach trajectory are probabilistic quantities. The docking time depends on the accuracy of control of the OS, the LFV and the docking device, primarily the manipulator. It is easiest to grasp with a large manipulator.
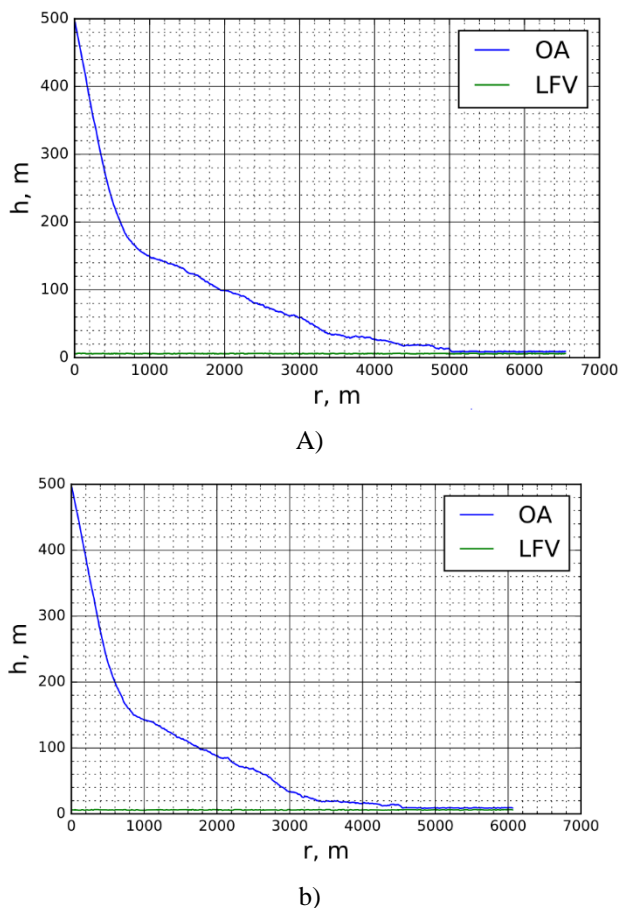

A)

b)

Fig. 3. Dependencies of heights of OA (OA – orbital aircraft) and LFV (low-flying vehicle) from passed distances
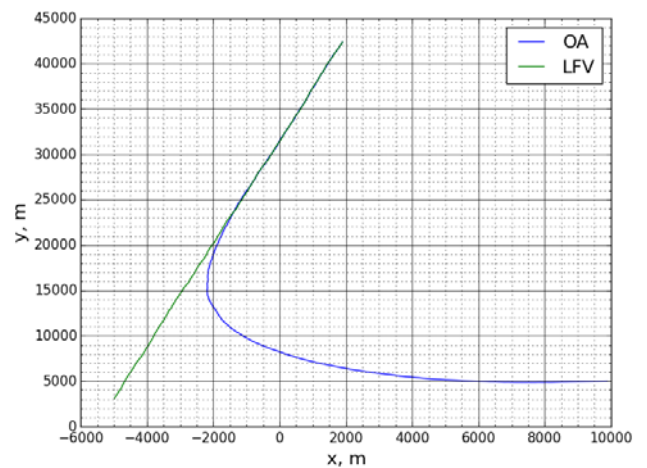

Fig. 4. Two-dimensional trajectories of OA and LFV at rapprochement and docking

Figure 4 shows two-dimensional trajectories of the OA and LFV in the horizontal plane. From the scale of the graph it is clear that before the OA reaches the altitude of the start of the docking maneuver, the OA makes a turn in such a way that the track angles of the OA and the LFV, and their coordinates in the horizontal plane, are as close as possible. Also, the OA assumes a speed close to the speed of the LFV, so that before docking it takes a position behind the LFV and begins to catch up with it at a relative speed of 2-4 m/s. After docking the OA with the LFV, their track angles and velocities are equal, since the OA is on the deck of the LFV.
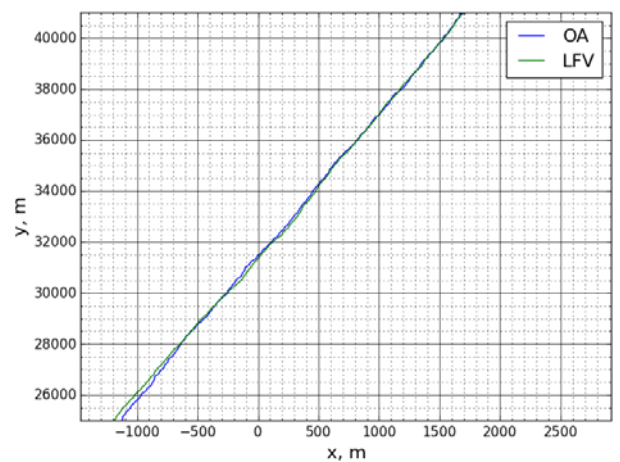

Fig. 5. Two-dimensional trajectories of the OA and LFV during approach and docking (enlarged fragment)

Figure 5 shows an enlarged fragment of the trajectories of the OA and LFV in the horizontal plane.
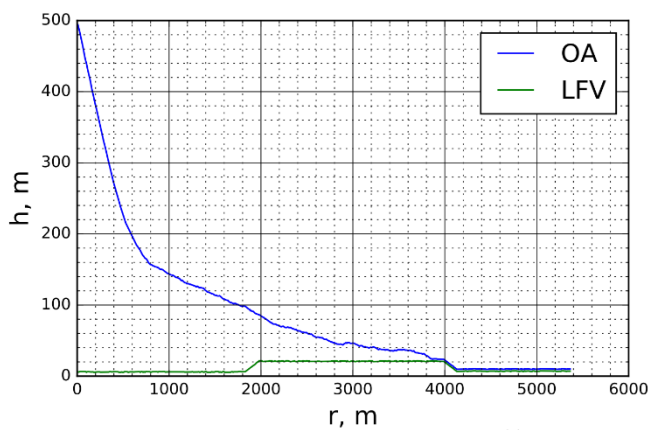
Fig. 6. An example of changes in the altitudes of the LFV and OA trajectories when the LFV approaches before docking

Figure 6 shows an example of changing the heights of the LFV and OA trajectories when the LFV approaches a height of 20 m before docking, in order to dampen the impact on the underlying surface under the LFV. This maneuver is advisable to carry out when there is a risk of an accident when touching the underlying surface. For example, if the LFV is moving over a swampy or non-water surface.

Simulation showed the possibility of landing the OA using an LFV using the proposed principle, with the specified ranges of motion parameters of the OA and LFV, and their errors. For successful docking, oscillations of the OA trajectory relative to the LFV trajectory with a probability of 95% within 10 seconds must be within a range equal to the distance between the manipulator grips. For example, with a distance between the manipulator grips of 5 m, the maximum permissible standard deviation of the OA control error at the trajectory angles is 3 degrees.

CONCLUSIONS

Preliminary logical processing of the results made it possible to reliably determine the relative position of the LEDs in 99% of cases. By means of computer simulation, the possibility of using the proposed operating principle of an optical surveillance system for docking aircraft was confirmed. By using an optical surveillance system instead of a heavy and expensive radio system, the weight and cost of on-board equipment can be reduced.

REFERENCES

[1] A. Nebylov, "Measurement of flight parameters near the sea surface", GAAP. St. Petersburg, 1994. 307 p.

[2] A. Nebylov, "Aerospace Sensors", Encyclopedia. Momentum Press, New Jersey, USA, 2009, 560 p.

[3] A. Nebylov, J. Watson, "Aerospace Navigation Systems", J. Wiley & Sons. UK, 2016 , 371p

Features of the Synthesis of a Robust Radio-Inertial Integrated Speed Meter

A. Nebylov

International Institute for Advanced Aerospace
Technologies, State University of Aerospace
Instrumentation Saint-Petersburg, Russia
nebylov@aanet.ru

V. Nebylov

International Institute for Advanced Aerospace
Technologies, State University of Aerospace
Instrumentation Saint-Petersburg, Russia
vladnebylov@gmail.com

Abstract—**The ability to accurately measure ground speed is necessary for any controlled aircraft, since it allows dead reckoning and coordinate control even in the absence of information from the satellite navigation system for some reason. A typical solution to this problem is the joint processing of information from a Doppler radio sensor (DRS) and accelerometer with a longitudinal axis of sensitivity.**

Key words —aircraft, accelerometer, radio sensor, invariant integrated speed meter, robustness.

For the DRS error, reduced to its input and measured in m/s, the spectral density S is known as $S_{vI}(\omega) = S_{vI}$, measured in m^2/s. For the error of the accelerometer as a meter of the derivative of the speed reduced to its input and having the dimension of speed, the maximum value of the variance of the first derivative is known, i.e. acceleration D_1 is measured in m^2/s^4.

The measured flight speed $V(t)$ has a variance limited above by a given value D_v with a certain mathematical expectation taken as the conventional zero of the speed scale and not affecting the measurement accuracy. Indeed, the known mathematical expectation is easy to compensate (make zero) a priori.

Let us consider the optimization problem of the described ground speed meter (Fig. 1) according to the criterion of the highest guaranteed accuracy.
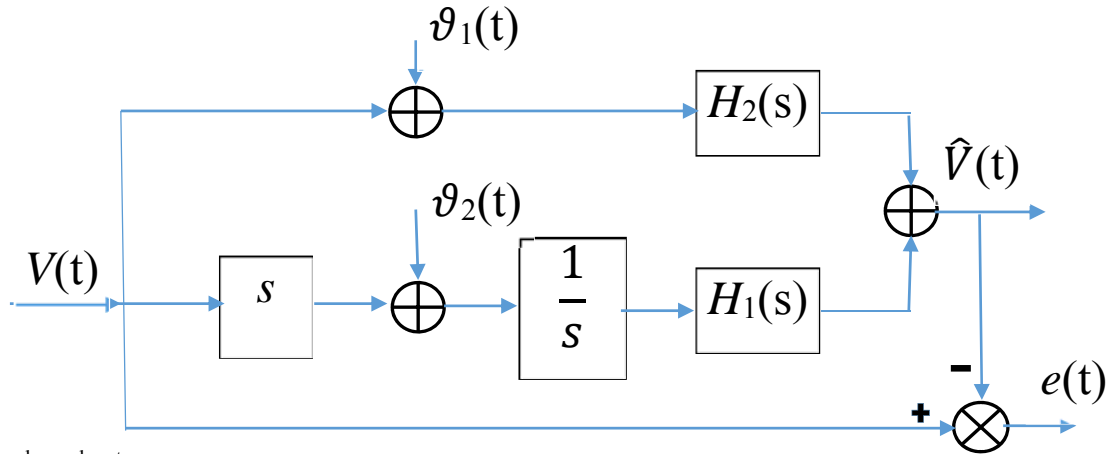

Fig. 1. Ground speed meter

The optimality criterion in this case is the minimum upper bound of the variance of the centered measurement error $e(t) = V(t) - \hat{V}(t)$

where $V(t)$ is the measurement result.

Since the highest of the orders of limited derivative actions is equal to unity, following [1,3,4], we choose the order of the denominators of the transfer functions $H_1(s)$ and $H_2(s)$ to be unity, i.e. $n=1$. This determines the type of transfer functions of the DRS channel $H_1(s)$ and the accelerometer channel $H_2(s)$:

$$H_1(s) = b_{10}/(1+a_1s),\ H_2(s) = b_{21}s/(1+a_1s).$$

Then the transfer function connecting the Laplace images of the measured speed and the dynamic measurement error will be

$$H_3(s) = 1 - H_1(s) - H_2(s) = \left[1 - b_{10} + (a_1 - b_{21})s\right]/ /(1+a_1s).$$

The optimization objective function is taken in the form

$$\overline{D}_e = D_{e1} + \overline{D}_{e2} + \overline{D}_{eg} \to \min,$$

$$D_{e1} = S_{v1}\frac{1}{2\pi}\int_{-\infty}^{\infty}\left|H_1(j\omega)\right|^2 d\omega = \frac{b_{10}^2 S_{v1}}{2a_1}$$

Where error variance D_{e1} along the location sensor channel, $\overline{D}_{e2}$ and $\overline{D}_{eg}$ - upper bounds of error variances along the accelerometer channel e_2 and dynamic error e_g, for which, following the theory of robust systems [1,2], we write the formulas

$$\overline{D}_{e2} = b_{21}^2 D_{v2}, \overline{D}_{eg} =$$
$$= D_g \left[\max\left\{1 - b_{10}, 1 - b_{21}/a_1\right\} \right]^2.$$

Having subjected the resulting function of three variables to analytical research to a minimum $\overline{D}_e(a_1, b_{10}, b_{21})$, you can find expressions for the optimal values of the meter parameters

$$a_1^0 = \left(\frac{S_{v1}}{4 D_{\dot{v}2}}\right)^{1/3},$$

$$b_{10}^0 = \left(\frac{3 D_{\dot{v}2}^{1/3} S_{v1}^{2/3}}{4^{2/3} D_g} + 1\right)^{-1},$$

$$b_{21}^0 = a_1^0 b_{10}.$$

For example, with S_{v1} = 1m²/s, D_{v2} = 10^{-3}m²/s⁴, D_g=0,5m²/s² we get a_1^0 = 6.30s, b_{10}^0 = 0.808, b_{21}^0 = 5.09s, $\overline{D}_e(a_1^0, b_{10}^0, b_{21}^0)$ = 0.096 m/s. For comparison, let's find the minimum possible value $\overline{D}_e$ in the case of a priori imposition of the invariance condition $H_1(s) + H_2(s) = 1$: $\overline{D}_e(a_1^0, 1, a_1^0)$ = 1.19 $D_{v2}^{1/3} S_{v1}^{2/3}$ = 0.119 m²/s². It can be seen that accepting the invariance condition leads to a loss in potential accuracy, since this (and any other additional condition) causes a deviation from the main requirement during optimization.

Note that the optimal parameter values lead the transfer function $H_3(s)$ to the form $H_3(s) = 1 - b_{10}^0$ ($0 < b_{10} \leq 1$ always), the corresponding inertia-free link. Physically, this is explained by the lack of a priori information about the frequency composition of the random process $V(t)$, and therefore it is advisable to make all its spectral components equal in the formation of a dynamic measurement error. If $D_g \to \infty$, that is, there is no information at all about the measured coordinate, then $b_{10}^0 \to 1$, $H_3(s) \to 0$ and the invariance condition is satisfied. The same result, but with limited high measurement accuracy, is obtained with $S_{v1} \to 0$, when there is no error of the location sensor (in this case only the location sensor is used), and when $D_{v2} \to 0$, when the accelerometer error is quasi-constant. In the latter case, the optimal parameters provide zero steady-state error only for an infinitely long observation time.
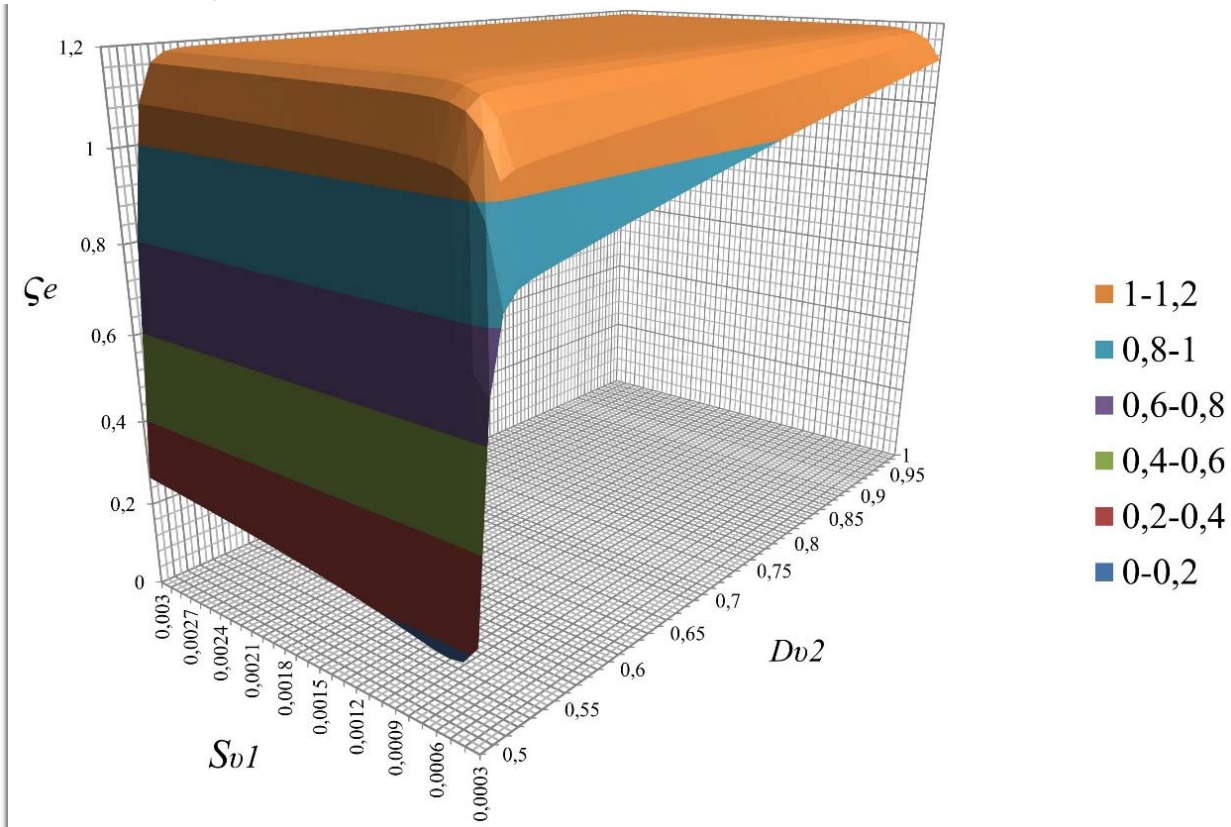

Fig. 2. Plot of the resulting root mean square error in an invariant measuring system

A three-dimensional plot of the resulting root mean square error in an invariant measuring system $\varsigma_e = \sqrt{\overline{D}_e}$ from the values of S_v and D_1 is shown in Fig. 2. It can be seen that the dependence of the accuracy of the integrated meter on the accelerometer error is stronger than on the DRS error.

The accuracy of indicators found in the article is guaranteed, and the meter itself is robust. This is the practical value of the article.

CONCLUSION

A typical example of the synthesis of a two-component robust flight parameter meter with incomplete a priori information is considered. It is shown what to do if for one of the sensors (usually inertial) there is no spectral description, but it is possible to limit the variance of the derivative error. The concept of an upper estimate for the resulting root mean square error is introduced. It is convenient to make the

minimum of this value as an optimality criterion. It is shown that when the dynamics of change in the measured flight parameter is weak, it is inappropriate to require the fulfillment of the invariance condition.

If the order of the system under consideration is high, the study can be carried out numerically.

REFERENCES

[1] Nebylov A.V. Ensuring control accuracy. M. Nauka – Fizmatlit, 1994, 304 p.

[2] Watson D., Nebylov A. V. (editors). Aerospace navigation systems. J. Wiley and Sons. UK, 2016, 371 p.

[3] Besekersky V.A. and others. Radioautomatics. M., Higher School, 1985, 271 p.

[4] Tupysev V.A., Stepanov O.A., Loparev A.V., Litvinenko Yu.A. Guaranteed Estimation in the Problems of Navigation Information 3rd IEEE Multi-conf. on Systems and Control. Saint Petersburg. 2009. pp. 1672–1677.

Numerical-Analytical Determining of Equilibrium Positions of Cubesat Nanosatellite under Aerodynamic and Gravitational Moments

E.V. Barinova
Inter-University Department
of Space Research
Samara National Research University
Samara, Russia
L5545@yandex.ru

I.A. Timbai
Inter-University Department
of Space Research
Samara National Research University
Samara, Russia
timbai@mail.ru

E. Mironov
Inter-University Department
of Space Research
Samara National Research University
Samara, Russia
mironovevgeniy01@gmail.com

Abstract—**For conducting many of the mission objectives, it is necessary to provide the required attitude of nanosatellites. An important task is determining the equilibrium positions of the nanosatellite relative to its center of mass. One of the important differences in the CubeSat nanosatellites dynamics is that the aerodynamic drag force depends on two angles of orientation - angles of attack and proper rotation, which is due to the rectangular parallelepiped shape. This paper proposes an algorithm for numerical-analytic determination of equilibrium positions of a dynamically asymmetric CubeSat nanosatellite when the center of pressure is displaced from the center of mass along three axes.**

Keywords—**CubeSat nanosatellite, aerodynamic moment, gravitational moment, angle of attack, angles of precession, angle of proper rotation**

I. INTRODUCTION

For successful implementation of most flight missions, a certain orientation of the nanosatellite (NS) is required. Due to the small size of a nanosatellite and low energy reserves, it is advisable to provide the necessary orientation using passive or combined stabilization systems [1-3]. Most nanosatellites are launched into low circular orbits, where gravitational and aerodynamic moments are predominant, therefore it is advisable to use both moments to ensure passive stabilization of motion relative to the center of mass. When creating stabilization systems, an important issue is to determine the equilibrium positions of nanosatellites relative to the center of mass and to study the nature of uncontrolled motion in their vicinity.

The equilibrium positions relative to the center of mass under the action of gravitational and aerodynamic moments for satellites with a shape close to spherical have been studied in detail in the literature [4-6]. For such satellites, the aerodynamic drag force does not depend on the attitude of the satellite relative to the incoming airflow. A symbol-numerical method for determining all equilibrium positions of the spherical satellite in the orbital coordinate system when the satellite's center of pressure is displaced relative to the center of mass by three coordinates and the three main

moments of inertia are not equal to each other, has been proposed in [6].

An important feature of the CubeSat nanosatellites is that they have the shape of a rectangular parallelpiped and, as a result, the aerodynamic drag force depends on the satellite's attitude relative to the oncoming airflow (angles of attack and proper rotation).

In works [7-8] the authors obtained analytical expressions for determining the equilibrium positions for two particular cases: displacement of the center of pressure of a dynamically symmetric NS from its center of mass along three coordinates and an NS with three different principal moments of inertia when the center of pressure is displaced from the center of mass along one of the principal axes of inertia. It has been shown that 8, 12 or 16 equilibrium positions are possible for a dynamically symmetric NS, and 8, 12, 16, 20 or 24 for a dynamically asymmetric NS.

In development of previous research, this paper considers the problem of determining the equilibrium positions for the more general case where the center of pressure of a dynamically asymmetric CubeSat is displaced relative to the center of mass along three coordinates. In addition, since a body-fixed coordinate system oriented along the structural axes of the CubeSat is used for it layout, including payload placement, it is important to take into account the presence of products of inertia.

In this paper we have developed an algorithm for numerical-analytical determination of the equilibrium positions of the CubeSat nanosatellite under the action of aerodynamic and gravitational moments when it moves in low circular orbits, considering the products of inertia. We have studied the NS which shape is a rectangular parallelepiped, for example, CubeSat 1-3U, 6U.

II. MATHEMATICAL MODEL

To describe motion of the nanosatellite around the center of mass, we used two reference frames: the trajectory reference frame $OXYZ$ (which coincides with the orbital one in

the case of circular orbit) and the body-fixed $Oxyz$ reference frame, the Ox axis of which is directed along the longitudinal axis of the NS, the Oy and Oz axes are perpendicular to its lateral sides. To define the relation between the trajectory and the body-fixed reference frames, we used three Euler angles: α is the spatial angle of attack, ψ is the angle of precession, φ is the angle of proper rotation. The coefficients of the rotation matrix from the trajectory frame to the body-fixed frame are the following:

$$b_{11} = \cos\alpha, \quad b_{12} = \sin\alpha\sin\psi, \quad b_{13} = -\sin\alpha\cos\psi,$$
$$b_{21} = \sin\alpha\sin\varphi, \quad b_{22} = \cos\varphi\cos\psi - \cos\alpha\sin\varphi\sin\psi,$$
$$b_{23} = \cos\varphi\sin\psi + \cos\alpha\sin\varphi\cos\psi, \quad b_{31} = \sin\alpha\cos\varphi,$$
$$b_{32} = -\sin\varphi\cos\psi - \cos\alpha\cos\varphi\sin\psi,$$
$$b_{33} = -\sin\varphi\sin\psi + \cos\alpha\cos\varphi\cos\psi.$$

The gravitational moment projections on the body-fixed axes take the form:

$$M_{gx} = 3\frac{\mu}{R^3}(J_{xy}b_{13}b_{33} + J_{yz}(b_{33}^2 - b_{23}^2) -$$
$$-J_{xz}b_{13}b_{23} + (J_z - J_y)b_{23}b_{33}),$$
$$M_{gy} = 3\frac{\mu}{R^3}(-J_{xy}b_{23}b_{33} + J_{xz}(b_{13}^2 - b_{33}^2) +$$
$$+J_{yz}b_{13}b_{23} + (J_x - J_z)b_{13}b_{33}),$$
$$M_{gz} = 3\frac{\mu}{R^3}(J_{xy}(b_{23}^2 - b_{13}^2) + J_{xz}b_{23}b_{33} -$$
$$-J_{yz}b_{13}b_{33} + (J_y - J_x)b_{13}b_{23}),$$

(1)

where J_x, J_y, J_z are the NS axial moments of inertia; J_{xy}, J_{xz}, J_{yz} are the NS products of inertia; μ is the gravitational constant of the Earth; R is the distance from the center of mass to the center of attraction.

The aerodynamic moment for the nanosatellite, which has the shape of rectangular parallelepiped, can be represented in the body-fixed axes in the following form:

$$M_{ax} = -c_0 qS \cdot \tilde{S}(\alpha,\varphi) \cdot (\Delta y b_{31} - \Delta z b_{21}),$$
$$M_{ay} = -c_0 qS \cdot \tilde{S}(\alpha,\varphi) \cdot (\Delta z b_{11} - \Delta x b_{31}),$$
$$M_{az} = -c_0 qS \cdot \tilde{S}(\alpha,\varphi) \cdot (\Delta x b_{21} - \Delta y b_{11}),$$

(2)

where c_0 is the drag coefficient, which can take the values from 2 to 3, depending on the physical properties of the gas and the surface of the nanosatellite, for design studies, it is assumed to be 2.2; $q = \rho V^2 / 2$ is the velocity head; ρ is the current atmospheric density; V is the NS flight velocity; $\tilde{S}(\alpha,\varphi) = |\cos\alpha| + \sin\alpha\left(\frac{l_x}{l_z}|\sin\varphi| + \frac{l_x}{l_y}|\cos\varphi|\right)$ is the area of the NS projection on the plane perpendicular to the velocity vector of the oncoming flow, divided by the characteristic area S [9]; $(\Delta x, \Delta y, \Delta z)$ are the coordinates of the center of pressure (geometric center) with respect to the center of mass; l_x, l_y, l_z are the lengths of the corresponding edges of the nanosatellite.

The equations of the NS motion under the aerodynamic and gravitational moments in the circular orbit can be written in the following way:

$$J_x\dot{\omega}_x - J_{xy}\dot{\omega}_y - J_{xz}\dot{\omega}_z - J_{yz}(\omega_y^2 - \omega_z^2) -$$
$$-J_{xz}\omega_x\omega_y + J_{xy}\omega_x\omega_z + (J_z - J_y)\omega_y\omega_z = M_{gx} + M_{ax},$$
$$J_y\dot{\omega}_y - J_{xy}\dot{\omega}_x - J_{yz}\dot{\omega}_z - J_{xz}(\omega_z^2 - \omega_x^2) +$$
$$-J_{xy}\omega_y\omega_z + J_{yz}\omega_x\omega_y + (J_x - J_z)\omega_z\omega_x = M_{gy} + M_{ay},$$
$$J_z\dot{\omega}_z - J_{xz}\dot{\omega}_x - J_{yz}\dot{\omega}_y - J_{xy}(\omega_x^2 - \omega_y^2) +$$
$$-J_{yz}\omega_x\omega_z + J_{xz}\omega_y\omega_z + (J_y - J_x)\omega_x\omega_y = M_{gz} + M_{az},$$

(3)

$$\omega_x = \dot{\psi}\cos\alpha + \dot{\varphi} + \omega_0 b_{12},$$
$$\omega_y = \dot{\psi}\sin\varphi\sin\alpha + \dot{\alpha}\cos\varphi + \omega_0 b_{22},$$
$$\omega_z = \dot{\psi}\cos\varphi\sin\alpha - \dot{\alpha}\sin\varphi + \omega_0 b_{32}.$$

(4)

Here, ω_x, ω_y, ω_z are the projections of the absolute angular velocity vector on the body-fixed axes; $\omega_0 = \sqrt{\mu/R^3}$ is the orbital angular velocity in the circular orbit.

The equilibrium position is a position which the nanosatellite will be holding all the time, in the case when at the initial moment of time, it was in that position and the velocities of all its points were zero [10]. According to the definition, the rate of change of the angles is zero ($\dot{\alpha} = 0$, $\dot{\psi} = 0$, $\dot{\varphi} = 0$) and (4) can be written as:

$$\omega_x = \omega_0 \cdot b_{12},$$
$$\omega_y = \omega_0 \cdot b_{22},$$
$$\omega_z = \omega_0 \cdot b_{32}.$$

(5)

Substituting (5) into system (3) and adding six conditions for the orthogonality of the direction cosines, we obtain a system for determining the equilibrium positions in the elements of the rotation matrix

$$\omega_0^2(-J_{yz}(b_{22}^2 - b_{32}^2) - J_{xz}b_{12}b_{22} + J_{xy}b_{12}b_{32} +$$
$$+(J_z - J_y)b_{22}b_{32}) = M_{gx} + M_{ax},$$
$$\omega_0^2(-J_{xz}(b_{32}^2 - b_{12}^2) - J_{xy}b_{22}b_{32} + J_{yz}b_{12}b_{32} +$$
$$+(J_x - J_z)b_{32}b_{12}) = M_{gy} + M_{ay},$$
$$\omega_0^2(-J_{xy}(b_{12}^2 - b_{22}^2) - J_{yz}b_{12}b_{32} + J_{xz}b_{22}b_{32} +$$
$$+(J_y - J_x)b_{12}b_{22}) = M_{gz} + M_{az},$$
$$b_{11}^2 + b_{21}^2 + b_{31}^2 = 1,$$
$$b_{12}^2 + b_{22}^2 + b_{33}^2 = 1,$$
$$b_{13}^2 + b_{23}^2 + b_{33}^2 = 1,$$
$$b_{22}b_{23} + b_{32}b_{33} + b_{12}b_{13} = 0,$$
$$b_{22}b_{21} + b_{32}b_{31} + b_{12}b_{11} = 0,$$
$$b_{21}b_{23} + b_{33}b_{31} + b_{13}b_{11} = 0.$$

(6)

It is impossible to find an analytical solution of the system (6) in general form. To solve it, in this paper we have developed a numerical-analytical algorithm for determining the equilibrium positions of the CubeSat nanosatellite under the action of aerodynamic and gravitational moments. Below is the description of the steps of this algorithm.

Firstly, the initial data are entered into the program. That is the NS mass, the lengths of the NS sides, the coordinates of the center of mass relative to the geometric center, the components of the inertia tensor, the altitude of the circular

orbit and the atmospheric model which is selected from those proposed. According this model, the atmospheric density is calculated.

If the characteristics of the nanosatellite correspond to special cases for which analytical solutions are found, the equilibrium positions are determined using the formulas obtained in [7-8]. If no special cases are found, the system (6) is numerically solved with respect to the directional cosines b_{ij}.

The obtained solution has the following form:

$$\left(b_{11}^k \quad b_{12}^k \quad b_{13}^k \quad b_{21}^k \quad b_{22}^k \quad b_{23}^k \quad b_{31}^k \quad b_{32}^k \quad b_{33}^k \right).$$

Here $k = \overline{(1, K)}$ is number of directional cosines combination; K is the number of the last obtained combinations of directional cosines.

Then we translate the obtained combinations of rotation matrix elements into Euler angles by the following formulas:

$$\alpha_n^k = \arccos(b_{11}^k),$$

$$\varphi^k = arctg\left(\frac{b_{21}^k}{b_{31}^k} \right) + \frac{\pi}{2}\left[1 - sign\left(b_{31}^k \right) \right],$$

$$\psi^k = arctg\left(-\frac{b_{12}^k}{b_{13}^k} \right) + \frac{\pi}{2}\left[1 - sign\left(-b_{13}^k \right) \right]. \tag{7}$$

Since the initial system of equations is essentially non-linear and includes absolute values, the combinations of angles obtained by formulas (7) are substituted into the initial system (6) to check compliance with the equilibrium condition. As all calculations are performed numerically, the equilibrium condition is assumed to be the case when the difference between the left and right sides of the equations of system (6) is less than the machine epsilon.

This algorithm was implemented in Wolfram Mathematica 12, machine epsilon for this math package is considered to be $\varepsilon = 10^{-16}$.

III. CALCULATION OF EQUILIBRIUM POSITIONS FOR CUBESAT 3U NANOSATELLITE

As an example, the equilibrium positions for the CubeSat 3U nanosatellite with the following characteristics were determined using the developed algorithm:

$$m = 3.6\ kg;\ l_x = 0.34\ m;\ l_y = 0.1\ m; l_z = 0.1\ m;\ S = 0.01\ m^2;$$

$$J_x = 0.0081\ kg \cdot m^2;\ J_y = 0.0379\ kg \cdot m^2;\ J_z = 0.0371\ kg \cdot m^2;\ c_0 = 2.2;$$

$$J_{xy} = -0.00008\ kg \cdot m^2;\ J_{xz} = 0.00034\ kg \cdot m^2;\ J_{yz} = 0.00006\ kg \cdot m^2;$$

$$\Delta x = -0.01165\ m;\quad \Delta y = -0.00046\ m;\quad \Delta z = 0.00397\ m.$$

Fig. 1 – 3 show the dependence of equilibrium positions on the angles of attack, precession and proper rotation as a function of flight altitude. The atmospheric density was calculated according to GOST 4401-81 [11]. According to these graphs, it is possible to determine the altitudes at which the numbers of equilibrium positions for this nanosatellite is changed.
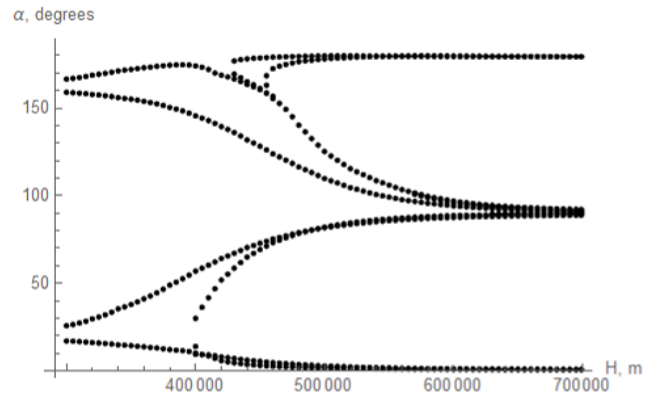

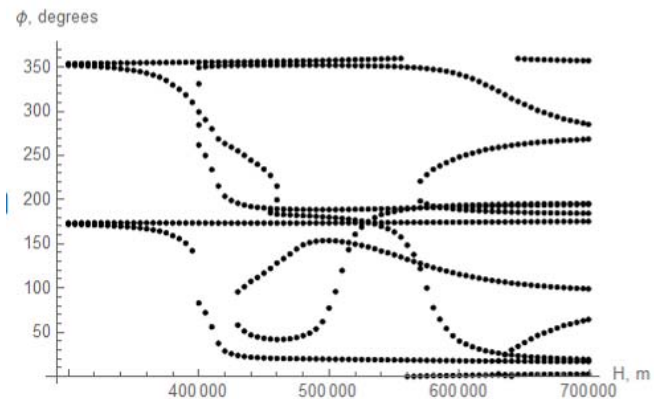Fig. 1. Equilibrium positions by angle of attack α for different altitudes

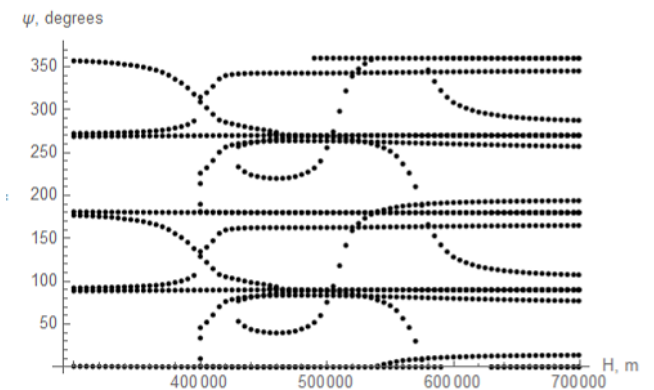Fig. 2. Equilibrium positions by angle of proper rotation φ for different altitudes

Fig. 3. Equilibrium positions by angle of precession ψ for different altitudes

Table 1 shows all the equilibrium positions of the CubeSat 3U NS at the altitude of 500 km obtained by the developed algorithm.

TABLE 1. RELATIVE EQUILIBRIUM POSITIONS OF THE NANOSATELLITE AT THE ALTITUDE OF H=500 KM

№	$\psi,°$	$\varphi,°$	$\alpha,°$	№	$\psi,°$	$\varphi,°$	$\alpha,°$
1	255.6	77.4	179.8	9	180.1	352.2	82.1
2	75.6	77.4	179.8	10	0.1	352.2	82.1
3	268.3	180.2	178.1	11	360	173.8	81.7
4	88.3	180.2	178.1	12	180	173.8	81.7
5	269.5	153.8	125.2	13	263.5	188.8	2.8
6	89.5	153.8	125.2	14	83.5	188.8	2.8
7	270	357.8	109.8	15	342.6	19.7	1.6
8	90	357.8	109.8	16	162.6	19.7	1.6

Therefore, in this paper we have developed an algorithm for numerical-analytical determination of the equilibrium positions of the CubeSat nanosatellite under the action of aerodynamic and gravitational moments when moving in low circular orbits. This algorithm can be useful in the development of stabilization and angular motion control systems, as it allows us to more fully see the picture of the uncontrolled motion of the NS under the influence of moments of external forces.

REFERENCES

[1] Belokonov, I. V., Timbai, I. A., and Nikolaev, P. N., Analysis and Synthesis of Motion of Aerodynamically Stabilized Nanosatellites of the CubeSat Design, Gyroscopy and Navigation, 2018, Vol. 9, No. 4, pp. 287–300.

[2] Rawashdeh, S.A., Attitude analysis of small satellites using model-based simulation, International Journal of Aerospace Engineering, 2019, vol. 2019.

[3] Chesi, S., Gong, Q., Romano, M., Aerodynamic three-axis attitude stabilization of a spacecraft by center-of-mass shifting, J. Guid. Control. Dyn., vol. 40, no. 7, pp. 1613–1626, 2017.

[4] Sarychev, V.A., Gutnik, S.A., Dynamics of an axisymmetric satellite under gravitational and aerodynamic torques, Cosmic Research, 2012, V. 50, no. 5. pp. 367–375.

[5] Sarychev, V.A., Mirer, S.A., Degtyarev, A.A., Duarte, E., Investigation of equilibria of a satellite subjected to gravitational and aerodynamic torques, Celestial. Mech. Dyn. Astron., 2007, vol. 97, no. 4, pp. 267–287.

[6] Sarychev, V. A., Gutnik, S.A., Satellite Dynamics under the Influence of Gravitational and Aerodynamic Torques. A Study of Stability of Equilibrium Positions, Cosmic Research, 2015, vol. 54, no. 5, pp. 388–398.

[7] Barinova, E.V., Timbai, I.A., Determining of Equilibrium Positions of CubeSat Nanosatellite under the Influence of Aerodynamic and Gravitational Moments, 27th Saint Petersburg International Conference on Integrated Navigation Systems, ICINS 2020 – Proceedings, St. Petersburg, Concern CSRI Elektropribor, 2020.

[8] Barinova, E.V., Timbai, I.A., Study of Relative Equilibrium Positions of a Dynamically Symmetric Cubesat Nanosatellite under Aerodynamic and Gravitational Moments, 26th Saint Petersburg International Conference on Integrated Navigation Systems, ICINS 2019 – Proceedings, St. Petersburg, Concern CSRI Elektropribor, 2019.

[9] Barinova E.V., Belokonov I.V., Timbay I.A., Choice of Design Parameters of a CubeSat 6U for Providing Three-axis Passive Stabilization, Proceedings of the International Astronautical Congress, IAC, 2022, Vol. 2022-September.

[10] Gantmakher, F.R., Lektsii po analiticheskoi mekhanike [Lectures on analytical mechanics], Nauka, Moscow, 1966.

[11] GOST 4401-81. Standard atmosphere parameters, Moscow, 1981.

Studying the Effectiveness of Using Nanosatellite Pre-spin Technology to Improve Maneuvering Accuracy

L.I. Sinitsyn
Samara National Research University
Samara, Russia
sinicyn.li@ssau.ru

I.V. Belokonov
Samara National Research University
Samara, Russia
belokonov.iv@ssau.ru

Abstract— **This study examines the problem of ensuring the uniaxial stabilization of a nanosatellite equipped with a jet engine during the correction impulse delivery. A study was conducted on the effect of preliminary rotation of a nanosatellite on the accuracy of the formation of inspection traffic in a probabilistic formulation. The SamSat-M nanosatellite with an electrothermal propulsion system, developed at Samara University, was chosen as the subject of this study.**

Keywords— nanosatellite, maneuvering, orbit correction error, inspection motion, uniaxial stabilization, Monte Carlo method

I.   INTRODUCTION

The current level of development in space technology is characterized by a sharp increase in interest in nanosatellites equipped with propulsion systems. By the beginning of 2024, the Nanosats database contains statistics on more than 2.5 thousand launched nanosatellites, of which about 220 have an on-board propulsion system. At the same time, there is an annual increase in geometric progression in the number of nanosatellite projects with propulsion systems: in 2019, less than 20 projects were recorded, in 2021 about 40 projects, and in 2023 more than 80 projects.

Along with the number of missions of maneuvering nanosatellites, their complexity also increases, and so, the requirements for maneuver accuracy are increasing. The maneuvering error, or trajectory correction error, is associated with the nanosatellite movement relative to the mass center in the correction pulse delivery process, which leads to "smearing" of the thrust vector. In this case, only part of the impulse is spent accelerating the nanosatellite in the direction required for the maneuver.

The main reason for the error in the implementation of the corrective impulse is associated with a change in the angular orientation of the thrust vector during the operation of the propulsion system. A change in the angular orientation of the thrust vector occurs because the thrust force does not pass through of the nanosatellite mass center and during the realization of the corrective impulse a disturbing moment arises [1].

Reducing the disturbing torque of the thrust force by refining the design of the nanosatellite and ensuring the direction of the thrust force vector through the mass center is not always an achievable goal under the conditions of unavailability of technologies and restrictions on the cost and time of manufacturing a nanosatellite. Thus, to increase the

accuracy of trajectory correction, it is advisable to use an active angular motion control system that maintains the required orientation of the thrust vector throughout the entire correction impulse.

II.   PROBLEM STATEMENT

This study examines a maneuver that ensures the movement of a nanosatellite equipped with a propulsion system in the inspection missions when an active spacecraft moves relatively to another passive spacecraft. The passive spacecraft is assumed to move in a circular orbit of radius R, which is called the reference orbit. The nanosatellite moves in such a way that its motion relatively to the passive spacecraft forms a closed trajectory called an inspection ellipse [2-3]. In this study, the influence of disturbing acceleration is not considered.

Assuming the gravitational field is formed by the Earth with a spherically symmetrical mass distribution, the small distance between the nanosatellite and the passive spacecraft compared to the value of R, and considering assumption the linearized system of equations of motion of the active spacecraft has the form [4]:

$$\ddot{x} + 2\omega_o \dot{y} = 0, \ \ddot{y} - 2\omega_o \dot{x} - 3\omega_o^2 y = 0, \ \ddot{z} + \omega_o^2 z = 0, \qquad (1)$$

where x, y, and z are the coordinates of the nanosatellite in the orbital coordinate system, the origin of which coincides with the center of mass of the passive spacecraft, its Ox axis is directed along the orbital speed of the passive spacecraft, the Oy axis is directed along the geocentric radius vector of the passive spacecraft, and the Oz axis complements the system to the right; ω_o – angular velocity of orbital motion.

The solution to system (1) has the form [5]:

$$\begin{pmatrix} \boldsymbol{r} \\ \boldsymbol{v} \end{pmatrix} = A \times \begin{pmatrix} \boldsymbol{r}_0 \\ \boldsymbol{v}_0 \end{pmatrix}, \qquad (2)$$

where $\begin{pmatrix} \boldsymbol{r} \\ \boldsymbol{v} \end{pmatrix}$, $\begin{pmatrix} \boldsymbol{r}_0 \\ \boldsymbol{v}_0 \end{pmatrix}$ are the current and initial column vectors of coordinates $\boldsymbol{r}(x, y, z)$ and speed $\boldsymbol{v}(\dot{x}, \dot{y}, \dot{z})$; A – 6×6 matrix with elements:

$$A_{11} = 1, \ A_{12} = 6\left(\sin\left(\omega_o t\right) - \omega_o t\right), \ A_{14} = \frac{4}{\omega_o}\sin\left(\omega_o t\right) - 3t,$$

$$A_{15} = -\frac{2}{\omega_o}\left(1 - \cos\left(\omega_o t\right)\right), \ A_{22} = 4 - 3\cos\left(\omega_o t\right),$$

$$A_{24} = \frac{2}{\omega_o}\left(1 - \cos\left(\omega_o t\right)\right), A_{25} = \frac{1}{\omega_o}\sin\left(\omega_o t\right),$$

$$A_{33} = \cos\left(\omega_o t\right), A_{36} = \frac{1}{\omega_o}\sin\left(\omega_o t\right),$$

$$A_{42} = -6\omega_o\left(1 - \cos\left(\omega_o t\right)\right), A_{44} = 4\cos\left(\omega_o t\right) - 3,$$

$$A_{45} = -2\sin\left(\omega_o t\right), A_{52} = 3\omega_o\sin\left(\omega_o t\right), A_{54} = 2\sin\left(\omega_o t\right),$$

$$A_{55} = \cos\left(\omega_o t\right), A_{63} = -\omega_o\sin\left(\omega_o t\right), A_{66} = \cos\left(\omega_o t\right),$$

the remaining elements of the matrix are equal to 0.

It is necessary to test the influence of the error in the implementation of the impulse of a nanosatellite (active spacecraft), which forms the inspection motion, on the deviation from the unperturbed trajectory for one orbital period. In this case, the matrix of system (2) is reduced to the following form:

$$A = \begin{pmatrix} 1 & -12\pi & 0 & -\dfrac{6\pi}{\omega_o} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

The initial state vector $\begin{pmatrix} r_0 \\ v_0 \end{pmatrix}$ is a random vector obtained by numerical simulation of the movement of the mass center during maneuvering, considering the main sources of errors in the implementation of the corrective impulse, which include the nozzle throat radius, the nozzle exit section radius, and the angular and linear deviation of the nozzle from the longitudinal axis. It is assumed that all elements of the state vector are uncorrelated random variables.

The mean of distribution and variance of the nanosatellite state vector after one orbital period are determined using linear transformations of vector random variables [6]:

$$\mathrm{M}\begin{bmatrix} r \\ v \end{bmatrix} = A \times \mathrm{M}\begin{bmatrix} r_0 \\ v_0 \end{bmatrix},$$

$$K\begin{bmatrix} r \\ v \end{bmatrix} = A \times K\begin{bmatrix} r_0 \\ v_0 \end{bmatrix} \times A^T, \tag{3}$$

where $\mathrm{M}\begin{bmatrix} r \\ v \end{bmatrix}$, $\mathrm{M}\begin{bmatrix} r_0 \\ v_0 \end{bmatrix}$, $K\begin{bmatrix} r \\ v \end{bmatrix}$, $K\begin{bmatrix} r_0 \\ v_0 \end{bmatrix}$ are the means of distributions and correlation matrices of the state vector after one orbital period and at the initial moment of time, and A^T is the transposed matrix.

III. METHOD FOR ASSESSING THE EFFECTIVENESS OF NANOSATELLITE PRE-SPIN TECHNOLOGY

To assess the influence of preliminary spinning of the nanosatellite, the probabilistic characteristics of the movement of the nanosatellite mass center due to a thrust impulse were determined, taking into account the main

sources of errors in the implementation of the corrective impulse for two cases: with and without preliminary rotation.

Pre-spin (preliminary rotation) of a nanosatellite can be achieved in various ways, for example, using a reaction wheel [7] or magnetic coils [8]. In this study, the spinning mechanism of the nanosatellite is not considered.

The motion model used to determine the state vector includes the equations of motion of the mass center and relative to the mass center in a fixed coordinate system (fixed coordinate system coincides with the body coordinate system at a time when the nanosatellite propulsion system has not turned on yet):

$$\begin{pmatrix} \dot{\boldsymbol{r}}_A \\ \dot{\boldsymbol{v}}_A \\ \dot{\boldsymbol{\omega}} \\ \dot{\boldsymbol{\lambda}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{v}_A \\ \boldsymbol{u} \\ I^{-1}\left(\boldsymbol{\omega} \times I\boldsymbol{\omega} + \boldsymbol{M}\right) \\ \dfrac{1}{2}\Omega\boldsymbol{\lambda} \end{pmatrix}, \tag{4}$$

where $\boldsymbol{r}_A\left(x_A, y_A, z_A\right)$, $\boldsymbol{v}_A\left(\dot{x}_A, \dot{y}_A, \dot{z}_A\right)$ are the column vectors of coordinates and velocities of the nanosatellite mass center in a fixed coordinate system; I is the diagonal matrix of axial moments of inertia of the nanosatellite; $\boldsymbol{\omega}\left(\omega_x, \omega_y, \omega_z\right)$ is the angular velocity of the nanosatellite relative to the fixed coordinate system; $\boldsymbol{u}$ is the acceleration from the operation of the propulsion system, expressed in a fixed coordinate system; $\boldsymbol{M}$ is the disturbing torque of the thrust force in the coordinate system associated with the nanosatellite; $\boldsymbol{\lambda}\left(\lambda_0, \lambda_1, \lambda_2, \lambda_3\right)$ is the quaternion of orientation of the associated coordinate system relative to the fixed coordinate system, Ω is the 4×4 matrix:

$$\Omega = \begin{pmatrix} 0 & -\omega_x & -\omega_y & -\omega_z \\ \omega_x & 0 & \omega_z & -\omega_y \\ \omega_y & -\omega_z & 0 & \omega_x \\ \omega_z & \omega_y & -\omega_x & 0 \end{pmatrix}.$$

Preliminary spin is considered by specifying the initial angular velocity of motion of the associated coordinate system relative to the longitudinal axis of the nanosatellite, along which the reactive force is generated.

The acceleration resulting from the operation of the propulsion system in a fixed coordinate system is determined using the following expression:

$$\boldsymbol{u} = -C^T \boldsymbol{F} / m,$$

where C^T is the transposed matrix of direction cosines for the transition from a body coordinate system to a fixed coordinate system; m is the mass of the nanosatellite; $\boldsymbol{F}$ is the thrust force in the body coordinate system.

The matrix of direction cosines C for the transition from a fixed to a body coordinate system is defined as follows:

$$C = \begin{pmatrix} \lambda_0^2 + \lambda_1^2 - \lambda_2^2 - \lambda_3^2 & 2\left(\lambda_1\lambda_2 + \lambda_0\lambda_3\right) & 2\left(-\lambda_0\lambda_2 + \lambda_1\lambda_3\right) \\ 2\left(-\lambda_0\lambda_3 + \lambda_1\lambda_2\right) & \lambda_0^2 - \lambda_1^2 + \lambda_2^2 - \lambda_3^2 & 2\left(\lambda_2\lambda_3 + \lambda_0\lambda_1\right) \\ 2\left(\lambda_1\lambda_3 + \lambda_0\lambda_2\right) & 2\left(-\lambda_0\lambda_1 + \lambda_2\lambda_3\right) & \lambda_0^2 - \lambda_1^2 - \lambda_2^2 + \lambda_3^2 \end{pmatrix}.$$

Projections of the thrust force $\boldsymbol{F}$ are determined by considering the phases of bootstrap operation, the residual effects of the thrust, the geometry of the nozzle, and its angular deviation from the longitudinal axis according to the formulas given in [9].

The disturbing torque of the thrust force in the motion model (4) is determined by the vector product $\boldsymbol{M} = \boldsymbol{d} \times \boldsymbol{F}$, where $\boldsymbol{d}$ is the displacement of the center of the nozzle exit section relative to the nanosatellite mass center.

To determine the probabilistic characteristics of the mass center movement using the Monte Carlo method, multiple integration of equation (4) is performed over the time interval of thrust formation. When modeling motion, the random parameters are the nozzle throat radius, the nozzle exit section radius, and the angular deviation of the nozzle from the longitudinal axis. It is assumed that these parameters are uniformly distributed random variables. The boundaries of the distributions were set by considering the technological capabilities of manufacturing a nanosatellite with a propulsion system [9].

Using transformation (3), the efficiency of the preliminary rotation of the nanosatellite is estimated.

IV. EVALUATION OF THE EFFECTIVENESS OF USING PRE-SPIN OF A NANOSATELLITE TO INCREASE MANEUVERING ACCURACY

The SamSat-M nanosatellite of the CubeSat-3U format, which is being developed at Samara University, was chosen as the object of study. SamSat-M is equipped with an electrothermal propulsion system capable of generating a thrust of approximately 0.1 N for 3 s. To estimate the probabilistic characteristics of the mass center movement acquired by the nanosatellite because of the thrust impulse delivery, modeling was performed with a sample volume of 4000, with a probability of 0.95 ensuring an estimation error of no more than 2%. The motion simulation parameters are shown in Table I.

TABLE I. PARAMETERS FOR SIMULATING THE MOTION OF THE SAMSAT-M NANOSATELLITE

Parameter	Value
Number of simulations	4000
Integration step, s	0,1
Displacement vector of the center of the nozzle exit section relative to the center of mass of the nanosatellite in the body coordinate system, mm	(-185; 4; 0)
Angles of nozzle deviation from the longitudinal axis in two planes, degrees	-0,5 … +0,5
Nozzle throat radius, mm	0,15 … 0,25
Nozzle exit section radius, mm	2,03 … 2,05
Nanosatellite mass, kg	4

When forming the inspection motion, we assumed that the thrust impulse was produced at a point with the coordinates $\boldsymbol{r}_0 = \left(\mathrm{M}[x_0], -\mathrm{M}[\dot{x}_0]/(2\omega_o), \mathrm{M}[z_0] \right)^T$ of the orbital coordinate system, while the speed of the nanosatellite was specified as a vector $\boldsymbol{v}_0 = \left(\mathrm{M}[\dot{x}_0], \mathrm{M}[\dot{y}_0], \mathrm{M}[\dot{z}_0] \right)^T$. In the case of an ideally produced impulse, this point corresponds to the initial

conditions of unperturbed motion of the following form:
$$\boldsymbol{r}_0 = \left(0, -\dot{x}_0/(2\omega_o), 0 \right)^T, \quad \boldsymbol{v}_0 = \left(\dot{x}_0, 0, 0 \right)^T.$$

Figure 1 shows a scheme for producing an impulse from the SamSat-M nanosatellite to form movement along an inspection trajectory relative to a passive spacecraft.
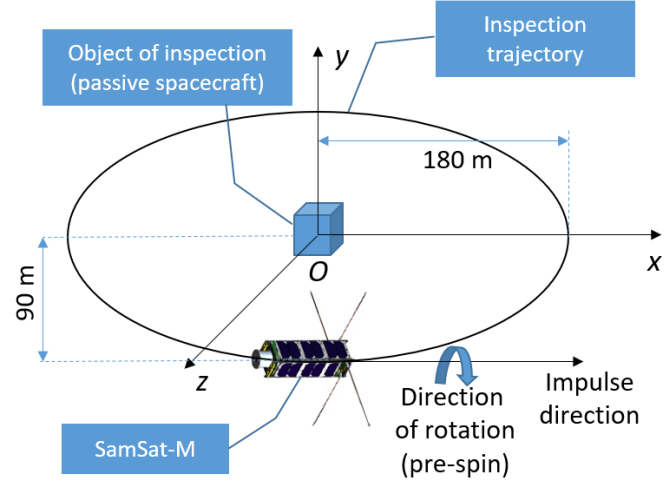

Fig. 1. Scheme for producing an impulse for the formation of an inspection movement

A histogram of speed increment values in the required direction is shown in Fig. 2.
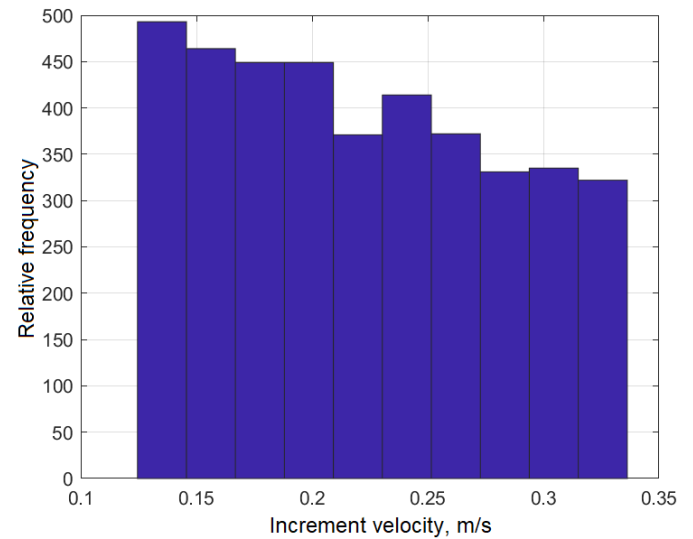

Fig. 2. Histogram of nanosatellite velocity increment values in the required direction in the absence of preliminary spin

Because of the simulation, the following estimates of the probabilistic characteristics were obtained: in the absence of preliminary rotation, the mean of distribution of the velocity increment in the required direction was 0.196 m/s and the standard deviation was σ = 0,055 m/s. If, before producing a thrust impulse, the nanosatellite was rotated 1 revolution per second relative to the thrust delivery axis, the mean of distribution and the standard deviation were 0.197 m/s and σ = 0.053 m/s, respectively. At the same time, the standard deviation of the unwanted speed increment in a direction perpendicular to the required one is reduced by an order of magnitude.

V. Conclusion

The accuracy of thrust impulse implementation influences the efficiency of performing a particular maneuver. The rotation of the nanosatellite relative to the axis along which the thrust impulse is produced allows us to average the arising disturbing moments of the thrust force, thereby reducing the error in the implementation of the thrust impulse.

This study investigated the effect of preliminary rotation of the nanosatellite before producing a thrust impulse on the efficiency of the inspection motion formation maneuver. The effectiveness of preliminary rotation is assessed by determining the probabilistic characteristics of the movement of the center of mass of a nanosatellite participating in relative motion as an active spacecraft over a time interval equal to one orbital period.

The influence of pre-spin on the efficiency of inspection traffic formation was studied using the example of a nanosatellite with an electrothermal propulsion system SamSat-M of CubeSat-3U format, which was developed at Samara University. It is shown that the lack of special means to reduce the effect of smearing the thrust impulse does not allow high-precision maneuvering to form an inspection movement. Pre-spin up to 1 revolution per second makes it possible to reduce the standard deviations of the state vector parameters by almost an order of magnitude.

References

[1] **Khromov A.V.** Interaction of the corrective propulsion system with the spacecraft orientation system, Electromechanical Issues, 2012, vol. 127, pp. 27-32.

[2] **Shcherbakov M.S., Medvedev S.A.** Study of the possibility of application of osculating ellipses of relative motion in the inspection problem, Mathematical modeling, computer and natural experiment in natural sciences, 2022, no. 4, mathmod.esrae.ru/40-157

[3] **Shcherbakov M.S., Uskova E.A.** Choosing the motion initial conditions and sustainability research inspection motion in small elliptic orbits, Proceedings of young scientists and specialists of Samara University, 2020, no. 1 (16), pp. 44-50.

[4] **Appazov R.F., Sytin O.G.** Metody proyektirovaniya trayektoriy nositeley i sputnikov Zemli (Methods for designing trajectories of carriers and Earth satellites). M.: Nauka. Gl. red. fiz.-mat. lit., 1987.

[5] **Narimanov G.S., Tikhonravov M.K.** Osnovy teorii polota kosmicheskikh apparatov (Fundamentals of the theory of spacecraft flight). M.: «Mashinostroyeniye», 1972.

[6] **Lebedev A.A., Bobronnikov V.T., Krasil'shchikov M.N., Malyshev V.V.** Statisticheskaya dinamika i optimizatsiya upravleniya letatel'nykh apparatov (Statistical dynamics and optimization of aircraft control). M.: Mashinostroyeniye, 1985.

[7] **Belokonov I.V., Sinitsin L.I.** Analysis of Reaction Wheel Control Strategies to Provide Gyroscopic Stabilization of the Thrust Direction During Nanosatellite Maneuvering, Cosmonautics and rocket Engineering, 2022, no 3 (126), pp. 17-29.

[8] **Ovchinnikov M.Yu., Roldugin D.S.** Recent advances in the active magnetic control of satellites, Spacecrafts and technologies, 2019, no. 2 (28), pp. 73-86.

[9] **Sinitsin L.I., Belokonov I.V.** Influence of deviations in manufacturing of electrothermal propulsion system on nanosatellite maneuvering accuracy, Vestnik of Samara University. Aerospace and Mechanical Engineering, 2021, vol. 20, no. 1, pp. 29-45.

On Improving the Accuracy of the Acoustical Navigation System with a Short Baseline for AUV Group Positioning

F.S. Dubrovin
IMTP FEB RAS
Vladivostok, Russia
f_dubrovin@mail.ru

A.Y. Rodionov
IMTP FEB RAS
Vladivostok, Russia
deodar1618@yandex.ru

A.F. Scherbatyuk
IACP FEB RAS
Vladivostok, Russia
alex-scherba@yandex.ru

Abstract— The report considers the approach of passive synthesis for acoustic aperture in a short base line navigation system, in which, instead of forming a beam, a direct assessment of the AUV coordinates and an indirect assessment of the bearing on the beacon are performed. It is assumed that there is no synchronization of the beacon clock and the AUV clock. The synthesis process uses the AUV motion model. The report describes the proposed and implemented estimation algorithm based on a particle filter, as well as some results for computer modeling of its operation.

Keywords- acoustical navigation system, group navigation, autonomous underwater vehicle

I. INTRODUCTION

The acoustical navigation system with a short base line (APS SBL) is used to solve the problem of determining the coordinates of a mobile object (in this case – autonomous underwater vehicle (AUV)) relative to the location of the beacon. The linear dimensions of the short base line located along the AUV body and including at least two APS SBL receivers are limited by the size of the AUV body and do not exceed 3-5 meters. The moments of signal reception by receivers (t_1, t_2) are recorded on board the AUV. The time interval $\Delta t_{12} = t_1 - t_2$ is proportional to the distance difference between the receivers and the beacon. The distance difference is connected with the bearing θ to the beacon and the value of the base line a between the APS SBL receivers on the AUV by expression:

$$\Delta t_{12} = (a*cos\ \theta)/c \ \text{ or } \ \theta = arccos\ (c*\Delta t_{12}/a) \quad (1)$$

where c is the velocity of the sound signal propagation in the water, θ is the angle between the line directed along the AUV body and the line directed at the beacon [1].

Assuming that a is measured with high accuracy, the error in the bearing estimation is determined according to the expression:

$$\delta\theta = -(\delta c*\Delta t_{12} + c*\delta\Delta t_{12})/(a*sin\theta), \quad (2)$$

and it follows that the bearing error decreases with an increase in the base distance a between the receivers. It also depends on the bearing value θ and increases indefinitely at angles of multiples for π.

In accordance with (2), an increase in the base distance a between the receivers leads to a decrease in the bearing error, but it is limited by the size of the AUV. To artificially increase the base distance, it can be used the technique of passive acoustic aperture synthesis, which consists of the following. Assume that the motion of the underwater vehicle on a certain part of the trajectory can be considered uniform and rectilinear. The virtual linear aperture is formed due to the motion of the AUV at a constant speed during several cycles of operation for the APS SBL. In this case, virtual hydrophones are formed using phase correction. This technique is called the Extended Towed Array Method (ETAM).

Although ETAM has been proven to work quite well [2], it has two limitations. Firstly, since the corrections are based directly on hydrophone measurements, they do not work at a very low signal-to-noise ratio (SNR). The second limitation is that even with a favorable SNR, the maximum achievable aperture is limited by the fact that the error of the phase correction coefficient increases with increasing number of repetitions.

Assuming that there is no synchronization of the beacon clock and the AUV clock, an algorithm for passive synthesis of the acoustic aperture is proposed for direct estimation of the AUV coordinates and indirect estimation of the bearing to the beacon. As with ETAM, it is a recursive process that uses the AUV motion model. In Part 2 of the report, the proposed estimation algorithm using a particle filter is described, and in part 3 some results of computer modeling of its operation are presented.

II. DESCRIPTION OF THE ALGORITHM

Consider the coordinates estimating of one AUV from a group. Taking into account the fact that the vertical coordinate (depth) is measured with the required accuracy, for simplicity of presentation, assume that the AUV moves in a horizontal plane rectilinearly and uniformly. At the initial moment of time $t=0$, AUV is on the surface and its position is known with the required accuracy. Without generality limitation, it can be assumed that the beacon is located at a fixed point with coordinates (x_a, y_a, z_a). It is assumed that the AUV and beacon clocks are not synchronized, but the period of navigation signals radiation by the beacon is known.

The task is to obtain estimates of the AUV location, provided that measurements of propagation times and propagation velocity of acoustical signals, as well as measurements of the course of AUV motion contain errors that are assumed to be random. Consider the use of a Particle Filter to solve this problem.

Let's form a cloud of N particles evenly distributed in 3-dimensional space (x, y, v), where x and y are horizontal coordinates, and v is the velocity of AUV motion in the horizontal plane. The weights of all particles at the initial

moment of time are assumed to be the same and equal to $\omega_i =1/N$. Suppose that the AUV moves rectilinearly and uniformly along a known course φ.

The AUV moves in the interval between the cycles of the APS SBL, τ is the known period of its operation. Each particle represents a model of AUV and moves like AUV. Since the location is calculated with errors, the location error can increase indefinitely over time.

At the stage of extrapolation of the particle filter operation, the predicted state vector (x^e_k, y^e_k) is calculated. To do this, the estimates obtained in the previous step (x^p_{k-1}, y^p_{k-1}) are used, as well as the estimated motion of the AUV:

$$x^e_k = x^p_{k-1} + \tau*v*cos(\varphi), \qquad y^e_k = y^p_{k-1} + \tau*v*sin(\varphi), \quad (3)$$

where $v = const$ is the unknown estimated velocity, and φ is the measured course of the AUV.

The following measurements are available in the k-th cycle of the APS SBL operation:

- the difference in the time of the signal arrival in the k-th cycle to the first and second receivers of the APS SBL on the AUV:

$$\Delta t_{12} = (d_{1k} - d_{2k})/c, \quad (4)$$

where d_{1k} and d_{2k} are the distances from the beacon to the first and second receiver of the APS SBL on the AUV in the k-th cycle of the APS SBL, respectively;

- the difference $\Delta t_{k,k-1}$ of the time for arrival of the signal in ($k-1$)-th and k-th cycles to the first receiver APS SBL on the AUV:

$$t_k = t^s_k + d_{1k}/c, \; t_{k-1} = t^s_{k-1} + d_{1k-1}/c, \; t^s_k - t^s_{k-1} = \tau, \quad (5)$$

where t^s_k and t_k are the times of emission and reception of the signal in the k-th cycle, t^s_{k-1} and t_{k-1} are the times of emission and reception in the ($k-1$)-th cycle, from where by subtract we have:

$$\Delta t_{k,k-1} = (d_{1k} - d_{1k-1})/c + \tau. \quad (6)$$

At the correction stage, to form an estimate of the AUV location, the measured time differences of the arrival for signals (Δt_{12}, $\Delta t_{k,k-1}$) and the calculated time differences of the arrival for signals (ΔT_{12}, $\Delta T_{k,k-1}$) for all particles are used, taking into account their current coordinates.

III. SOME RESULTS OF AUV POSITIONING MODELING BASED ON APS SBL DATA

The assessment of the accuracy and operability of the navigation algorithm for determining the location of the AUV was carried out within the framework of numerical modeling using a particle filter. The model experiments were performed in the IDLE environment (Python). During the simulation, the period of operation for APS SBL was 10 seconds. The speed of the acoustical signal propagation was assumed to be equal to 1500 m/s. A cloud of 1000 particles was used in the simulation.

The simulation was performed using the Monte-Carlo method, while each experiment was performed in a cycle 100 times with different random errors and based on the results obtained, the average value and variance of errors in estimating the location and velocity of the AUV were calculated. The initial area of uncertainty for the AUV location was a square with a side of 20 m, and the initial area of uncertainty for the speed of motion was an interval (0, 2.0) m/s. It was assumed that there was no current velocity in the working area.

It was assumed that the measurements for the course and arrival times of the signals contained random errors. In a situation where, during modeling, random errors in measuring the AUV course were in the range of (-0.5, 0.5) degrees, and random errors in measuring the arrival times of signals were in the range of (-10^{-4}, 10^{-4}) seconds, the average value of the AUV location error was 6.88 m, the standard deviation was 2.26 m, the average value of the velocity determination error was 0.014 m/s and its standard deviation was 0.011 m/s. In the case when random errors in measuring the AUV course were in the range (-0.5, 0.5) degrees, and random errors in measuring the arrival times of signals lay in the range (-10^{-3}, 10^{-3}) seconds, the average value of the AUV location error was 34.36 m, the standard deviation was 5.01 m, the average value of the velocity determination error was 0.094 m/s and its standard deviation was 0.031 m/s.

Figure 1 shows the model trajectory of the AUV motion during 50 cycles of the APS SBL operation, which was obtained in one of the experiments. The AUV started from a point with coordinates (-3000, -2000, 500) m. The beacon was located at a point with coordinates (0, 0, 0) m. Parameters of the AUV motion in the horizontal plane: the speed is 1 m/s, the course is 45 degrees. The distance traveled was about 500 meters and the travel time was about 8.3 minutes.

Errors in estimating the trajectory of the AUV during 50 cycles of the APS SBL operation using the described algorithm are shown in Figure 2. In this experiment, random errors in measuring the arrival times of signals lay in the range of (-10^{-4}, 10^{-4}) seconds.

The error in calculating the location of the AUV without correction for the specified time of motion, with an estimate of the speed from the measurements of the thruster rotation with an error of about 20% of the real speed value and a random error in measuring the course of 0.5 degrees, exceeds 100 meters. At the same time, in this case, due to the lack of synchronization of the AUV clock and the beacon clock, it is not possible to directly determine the location of the AUV using the APS SBL.

REFERENCES

[1] Borodin V., Smirnov G., Tolstyakova N. et al. Hydroacoustic navigation aids. L: Shipbuilding, 1983, 263 p.

[2] YuWang; ZaiXiaoGong; RenheZhang. Phase correction in extended towed array method. J. Acoust. Soc. Am. 141, 3916 (2017).
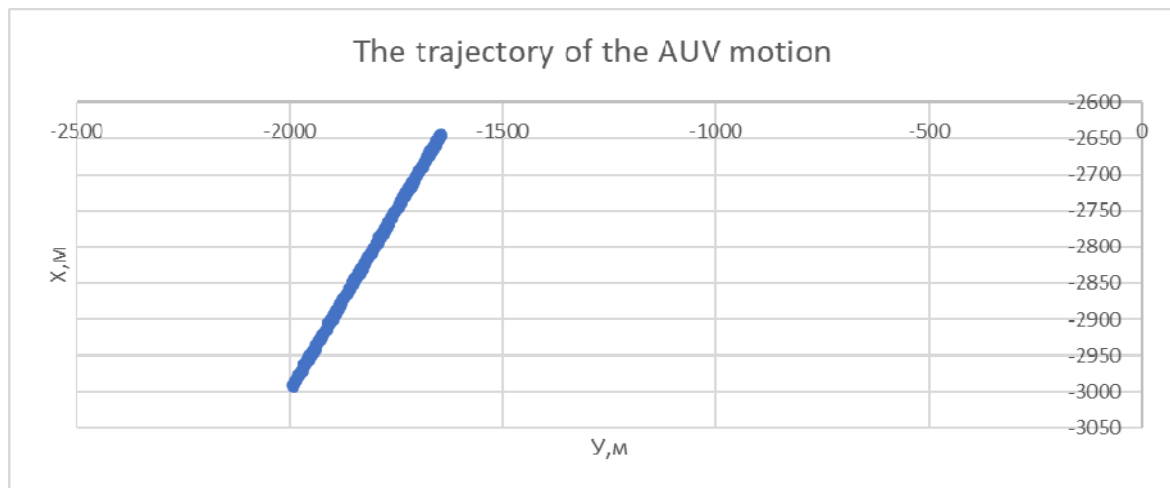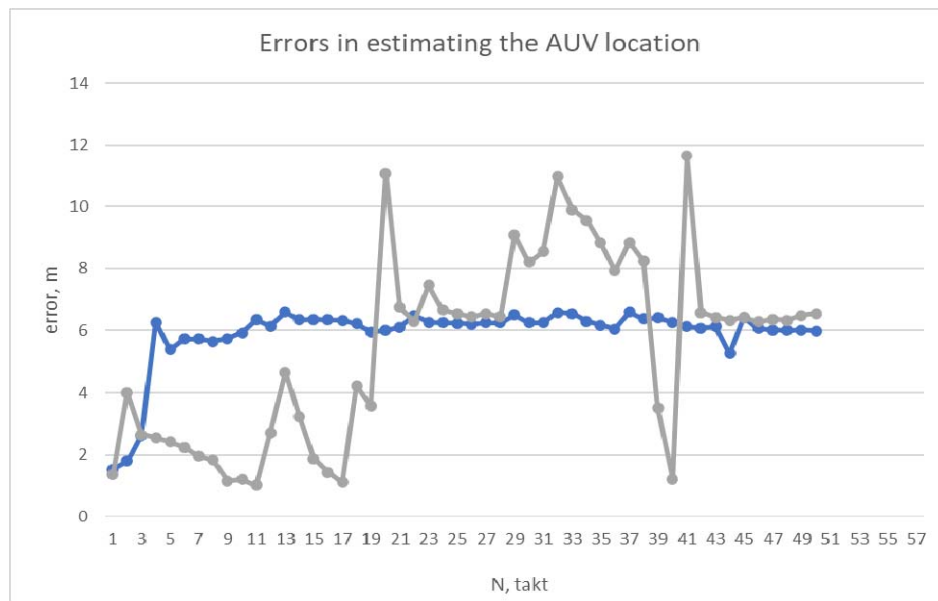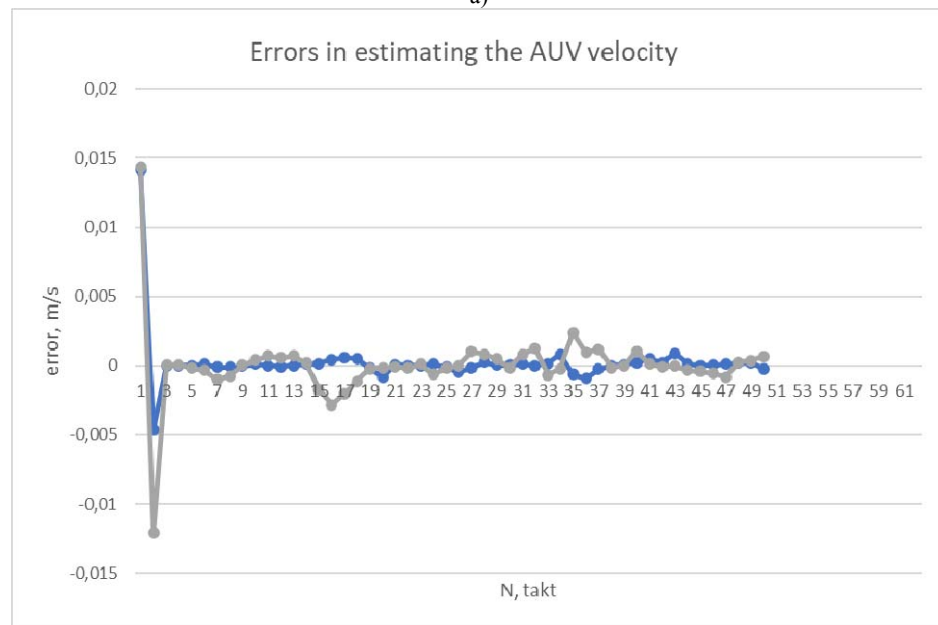
Fig. 1. The trajectory of the AUV (the beacon is located at a point with horizontal coordinates 0.0, 0.0).

a)

b)

Fig. 2. Errors in estimating the AUV location (a) and velocity (b) (blue curve – for errors in measuring the course in the range of ± 0.1 degrees, gray curve – for errors in measuring the course in the range of ± 0.5 degrees).

31ST SAINT PETERSBURG INTERNATIONAL CONFERENCE
ON INTEGRATED NAVIGATION SYSTEMS

27–29 May 2024

Верстка *А. А. Зуева*

_____