Scientific and technical journal

ROCKET-SPACE DEVICE ENGINEERING AND INFORMATION SYSTEMS

Volume 3. Issue 1. 2016

Contents

Vol. 3, Iss. 1, 2016

ROCKET-SPACE DEVICE ENGINEERING AND INFORMATION SYSTEMS
Vol. 3, Iss. 1, 2016, pp. 2–9

Using a Systematic Approach to Solving the Problematic Issues of Functioning of the Automated Complex of Programs for Ballistic and Navigational Support of GNSS Spacecraft Missions

V.V. Betanov[1], **V.K. Larin**[2]

[1]doctor of engineering science, professor, [2]candidate of engineering science
Joint Stock Company "Russian Space Systems", Russia
e-mail: betanov_vv@spacecorp.ru

Abstract. The article describes the general provisions of the systematic approach to the study of the problems arising in the application systems for data processing. A thesis that the investigated systems have a hierarchical structure combining a certain amount of counterparts is taken as a basis. Therefore, the solution for the system under analysis should be in identifying the inconsistencies between the normal operation being on the same or different hierarchy levels and defining the method of solution depending on the problem structure. A step-by-step technology in the form of successive stages of solution is proposed in the frames of a systematic approach. A work of the automated complex of programs for ballistic and navigational support of GNSS spacecraft missions is given as an example of usage of the proposed technology. As a result, the component parts of the complex containing probable errors that disturb normal functioning were detected. Moreover, the methods for their elimination are defined. Conclusions are drawn about the practicability of using a systematic approach in the form of the proposed technological scheme for the analysis of work of the hardware and software objects for data processing in the space industry.

Keywords: system approach, issue, system, subject area

Introduction

The automated complex of programs for ballistic and navigational support (ACP BNS) of spacecraft (S) control is a difficult technical system, which functioning is connected with the usage of the following types of support during the regular work: mathematical, program, information, technical, and other types. Each type of the support is a subsystem consisting of a group of the interconnected elements. Flawless operation of all subsystem elements provides the timely and qualitative solution of complex problems.

ACP BNS operation has shown that in some cases due to various reasons there is a violation of normal work of a complex, which is expressed, as a rule, either in lack of the solution or in obtaining the solution to inadmissible accuracy. The operational analysis of such situations being made gives a chance to establish only the fact of existence of a problem, but not the reasons of its origin.

The article gives a generalized technology of the solution of problematic issues of such kind. The description of an example of using the technology for identification of the possible reasons of abnormal work of ACP BNS is given and methods of their elimination are offered. It is supposed that the procedure of using the generalized technology for solving such situations has an iterative character.

1. The generalized technology for the problems solution

The basic concepts used in this article should be defined: a system approach, problem, system and subject area necessary for logical justification of the offered technology for the solution of problematic issues [1-3].

A system approach in work is the approach to research of an object (a problem, phenomenon, and process) as to the system where elements, internal and external relations are allocated, which influence in the most substantial way the results of its functioning being studied, and the purposes of each of elements are defined from the general mission of the object. In turn, a hierarchically ordered set of questions characterizing a difference between the valid and desirable condition of the object is understood under a problem.

According to the classification by the structure degree, all problems are subdivided into three classes:

- *well-structured,* or quantitatively expressed problems, which lend itself to mathematical formalization and are solved using formal methods;
- *unstructured,* or qualitatively expressed problems, which are described only at the substantial level and are solved by means of informal procedures;
- *semi-structured,* or mixed problems, which contain both qualitative elements and the little-known, uncertain parties, which tend to dominate.

Further, a subject area is a part of the real world considered within this context. The context is an area of research, which is considered an object of some activity.

A system (from ancient Greek σύστημα – a whole made of parts; combination) is a set of interacting or interdependent component parts forming a complex/intricate whole, which quantities surpass the qualities of the forming parts.

At this time a problem solution is discrepancy elimination between a desirable and valid condition of the object. In further reasoning a concept “problems” will be used in relation to an assessment of system functioning of a certain subject area (SA).

It should be noted that a system (as well as a problem) has a hierarchical structure uniting in the whole certain quantity of the interconnected parts. Therefore, the solution for any functioning system can consist (generally) in elimination of the revealed discrepancy of work of one or several parts, which are at one or different levels of a system hierarchy.

In case of finding the solution not for all parts of the system defined as “infected”, it is necessary to make an assessment of a solution degree.

As the generalized technology of a solution for some system, it is possible to consider the following step-by-step sequence of solution stages.

1st step. System decomposition into the largest functionally completed fragments of the first level.

2nd step. Carrying out the analysis of possible discrepancy to normal functioning of the selected fragments (identification of the “infected” parts of the system).

3d step. Problem formulation for the “infected” parts.

4th step. Definition of a structure degree of operation problems of the “infected” parts.

5th step. The choice of a solution method for each “infected” system part (a private problem).

6th step. Finding a private problems solution for the “infected” parts of a system.

7th step. Decomposition of the rest parts with unsolved functioning problems into the elements of the following hierarchy system level (less large) and carrying out actions on steps 1-6.

Decomposition comes to an end in two cases:

- further system decomposition of a functional sign is impossible;

- solutions of private problems for all parts of the last level are found.

8th step. An assessment of a problem solution of system functioning on a cumulative number of the solved private problems at the levels of hierarchical decomposition.

The main methods of the system analysis used when solving problematic issues

The list of the main methods of the system analysis used for the solution of the considered problems [1-3] is given in Table 1.

Table 1. A list of the main methods of a system analysis

Method name	Integrated characteristic
Analytical methods Statistical methods Set-theoretical methods Linguistic methods Semiotic methods Graphic methods	**Formal methods** – methods of the formalized representation of systems
Morphological approach Structurization methods: relevance (objectives) tree, predication graph, etc. Delphi methods Methods of expert estimates Methods of “scenarios” Methods of brainstorming (attack)	**Heuristic methods** – methods directed to activation of using intuition and experience of experts

In most cases formal methods are applied to the solution of the structured problems; heuristic methods are applied to semi-structured and unstructured problems.

Analytical and statistical methods are mostly used out of formal methods; a method of expert evaluations including expert systems, morphological approach and a method of brainstorming are mostly used out of heuristic methods.

The recommendations on the sequence of solution stages depending on the extent of its structurization are provided.

Structured problems

1) Formulation of the purpose.

2) Creation of a mathematical model for the system description in the form of a set of elements connected with each other by certain relations.

3) The analysis of the model regarding search for the “infected” parts, choice of a decision method.

4) An assessment of a solution.

Unstructured problems

1) Formulation of the purpose.

2) The system analysis regarding search for the “infected” parts; a choice of a decision method.

3) Formation of a group of experts and using a brainstorming method.

4) Using a method of expert evaluations, including development of an expert system (taking into account the results of item 3).

5) An assessment of a solution.

Semi-structured problems

1) Formulation of the purpose.

2) Formation of achievement alternatives of the purpose; an assessment of these alternatives by means of the corresponding criteria and a choice of the preferable alternative.

3) The system analysis regarding search for the “infected” parts; a choice of decision methods (formal or heuristic) depending on their structure degree.

4) Searching for a solution of private problems.

5) An assessment of the solution of a common system problem (taking into account the results of item 4).

The following figure gives a technological search scheme for problem solution in the form of a block diagram of a step-by-step technology.

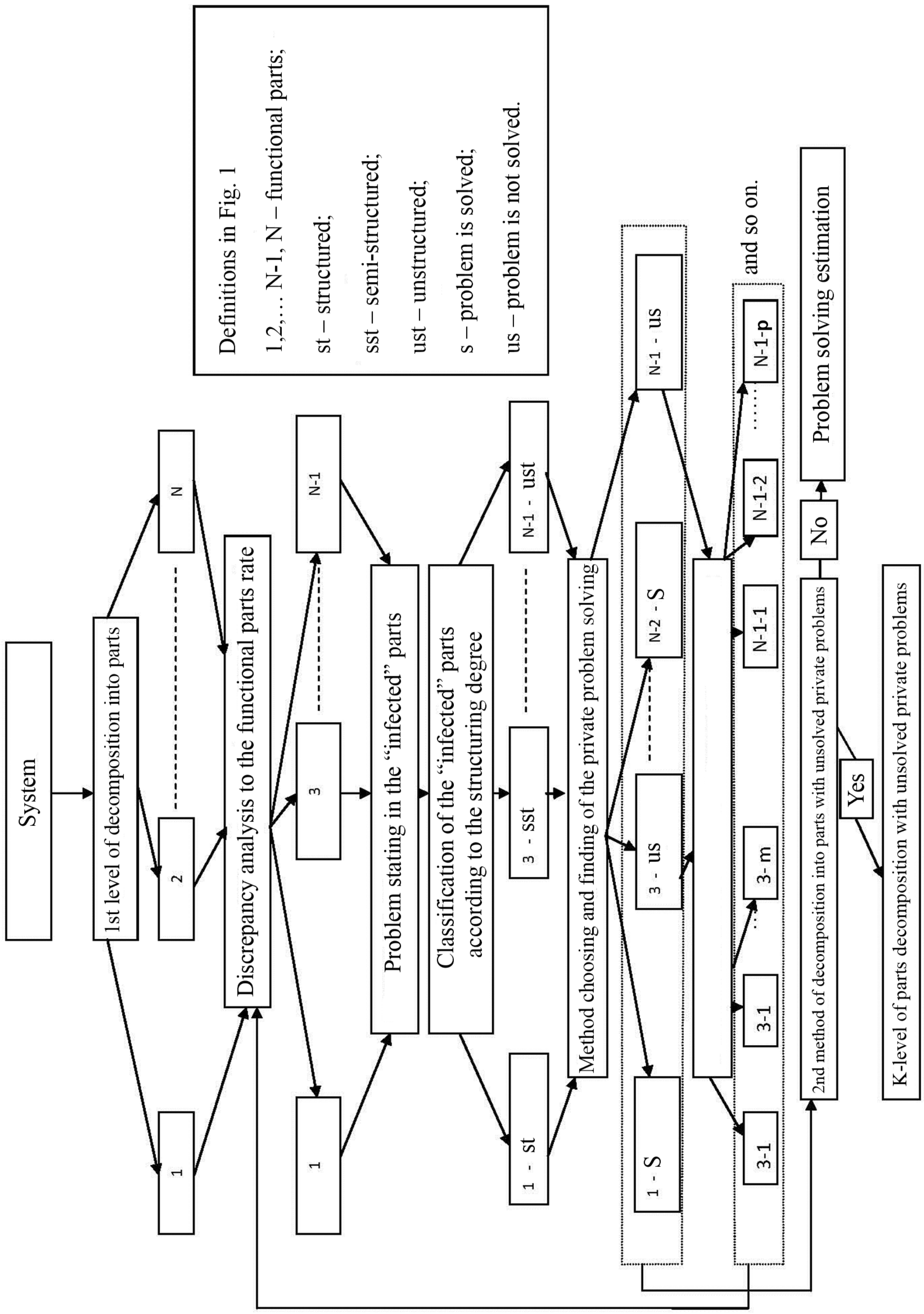


Figure 1. A technological scheme for problem solving

2. Using the generalized technology for searching the ACP BNS operation

As an example, it is possible to consider the problem of unsatisfactory work of ACP BNS, which is expressed in an inadmissible deviation of the current navigation parameters of characteristics of the spacecraft movement calculated on measurements from reference values (the final data provided in the GNSS spacecraft catalogs).

A subject area, which the object of research belongs to, should be specified. In this case it will consist of the following main support types: mathematical (algorithmic), program, technical and information. These support types can serve as functional parts of the first decomposition level.

The analysis of possible discrepancy to normal functioning of the allocated parts has shown that as far as the solution has been found, all support types functioned. However, if the problem was in failure of the technical support (lack of power supply on a server input, mechanical damage of its details, etc.), then it would result in lack of the solution. Nevertheless, as the solution took place, the technical part can be excluded from further consideration. Thus, it is possible to consider algorithmic, program, and information parts to be the “infected” parts.

At the same time, the formulation of a problem remains the same – the unsatisfactory accuracy of the received solution.

All three remaining parts containing possible mistakes (problems) leading to the current situation have in general a structured (an algorithmic part by definition) and semi-structured (program and information parts) character.

Probable existence of a problem in the algorithmic part results in need of its further decomposition into components, namely, into a module of preliminary processing of trajectory measurements (PP) and a determination module of spacecraft movement parameters (a solution of a boundary problem – BP).

The output data of PP is a session measurements table (trajectory measurements of one spacecraft for one tracking station on the set time interval). The number of sessions will be defined by multiplication of at the same time measured types of parameters by the number of measuring points. At the stage of PP the filtration (including rejection) measurements of the current navigation parameters by the set criteria is made. At this, the percent of the defective data has to make a certain part from all measurements accepted in processing and providing convergence of a problem solution. At non-performance of this condition of the measuring information obtained by BP for further calculations will be insufficiently that can lead to inadmissible mistakes in the specified movement of the spacecraft parameters.

Thus, one of the reasons of the existing problem can be in lack of a condition of a necessary minimum of a number the sessions and numbers of measurements in a session, which existence substantially would explain the unsatisfactory solution of a problem.

The module of the solution of a BP is mathematically much more difficult than the PP module. It includes: a model of the spacecraft movement, matrix private derivative of measurements on entry conditions, statistical processing methods of measurements (for example, a method of the ordinary least squares (OLS), methods of integration of the differential equations of the movement, formation and the decision of systems of a large number of the linear equations, interpolation and approximating polynomial, etc. Traditional conditions of the solution of similar problems are well approved and, as a rule, do not cause difficulties. At the same time, as the weakest spot at the solution of a BP it is possible to consider an opportunity of bad conditionality of the matrixes used at the solution of the normal equations for calculation of amendments to parameters of an orbit and to other specified parameters. For OLS it is Gram matrix. An expedient solution of this problem is introduction at this stage of calculations of the criterion of the degree of conditionality of matrixes.

The problem in the information part can be formulated in two variants:

- lack of the whole necessary information or a part of it;

- presence of mistakes in the obtained information for this processing session.

To find the reasons of the problem arising, it is necessary to decompose the information part into the following hierarchy level, namely, information sources: the central data base (CDB) and internal FTP-server of the augmentation and monitoring system (SAM). As the essence of a problem concerns directly information, it is necessary to move to the following level – an information level – distribution of information on sources (Table 2).

Table 2.Sources of the data being used

Source of information	Type of information
CDB	Navigation messages: real-time data, almanac; Reference data: logical power scale, global constants, technical characteristics of signals, parameters of exciting force (for example, charged-coupled device (CCB), spacecraft.
Internal FTP-server	Rinex-files

Probability of existence of mistakes in information, which is contained in sources, should be analyzed. Thus, for CDB:

- navigation messages (real-time data and almanac) are in a supershot, which is transmitted from the spacecraft to ground stations each 2.5 min. At receiving check on reliability is automatically made. These data are used by all the Centers – the participants of GNSS. At the same time mistakes are improbable;

- reference data are registered in the base once, are carefully checked, used at each session of the GNSS spacecraft orbits definition that makes it possible to consider (by the analogy with navigation messages) improbable existence of errors in them;

- parameters of the Earth rotation – the data on the CCD "read" from the external special server that are in the form of the annual massifs containing the daily information (t, x_p, y_p – time and coordinates of poles) used by all participants of GNSS. Mistakes are almost excluded.

Rinex-files of the set type. Their contents include:

1. File of observation data (FOD): time, pseudo-range, phase, and Doppler correction.
2. File of navigation messages (FNM).
3. File of meteorological data (FMD).
4. File of GLONASS navigation messages.
5. File of GEO navigation messages.
6. File of data satellite clocks and receivers (FDW).
7. File of wide area updating information SBAS (FWUI).

From the submitted files, FOD files should be considered, which data are used in a session of information processing and may contain errors. At this FOD includes: time, pseudo-range, phase, and Doppler amendments.

Three variants of existence of private problems are possible:

1. Lack of information on any of parameters.
2. Existence of low-quality data.
3. Presence of an incomplete volume of data.

In the first variant it is possible to allow lack of Doppler measurements; in this case there will be no "a solution on speeds", but the solution of a problem will be received. Lack of information on time or pseudo-range, and also data on a phase causes impossibility to determine the specified orbit parameters on any of spacecraft.

In the second variant it is necessary to separate low-quality information from qualitative. The main criterion of such division is the place corner γ, under which there was a reception of a signal from the spacecraft in a visibility range of the measurement station (MS). It has experimentally been established that at $\gamma \leq 70$ the information obtained by MS is of low quality (a big noise level due to the atmosphere). Observance of this condition when loading information into the base at the first stage of PP will enable one to remove the low-quality information prior to calculations and to provide necessary accuracy.

In the third variant the incomplete volume of information is caused either by passing of the route across the "edges" of visibility ranges of MS or removal of part of information on a condition $\gamma \leq 70$. The most radical solution of this problem is removal of this spacecraft from the processing variant.

The software, as well as algorithmic, is expedient to divide into two main programs of a complex: preliminary processing of measurements and determination of parameters of an orbit (a following level).

Private problems of programs are formulated as follows: PP – formation and recording into the archive an insufficient amount of qualitative sessions of measurements on each spacecraft; BP – determination of parameters of spacecraft orbits with an unsatisfactory accuracy.

It should be accepted that the algorithms of PP and BP transformed into the codes of programs are identical and have no errors. Considering that programs are a set of interdependent modules where both analytical calculations and different information transforms are made, it is necessary to continue the “decomposition” of the software (S) of ACP into software modules. The first level for the software: general modules of the automated complex of programs (ACP), PP, and BP. The decomposition of the software parts at the following bottom level should be continued.

The general modules of ACP:

- settings of operation modes and configuration of a complex (formation of setting files);
- software modules of interaction with the database (DB);
- software modules of interaction with file archives.

Preliminary processing of measuring information (PP):

- formation of measurement sessions of the current navigation parameters (MCNP);
- processing and filtration of sessions MCNP;
- formation of sets of basic lines;
- formation of differential measurements;
- filtration of sessions of differential measurements;
- determination of location (DL) on code measurements of the range;
- statistical assessment of results of DL.

Expeditious specification of orbits parameters of navigation spacecraft (BP):

- solving a boundary problem (BP);
- specification of clock error predictions (CEP).
- formation of the archive of the corresponding files.

Practice of the ACP software operation has shown that one of the main problems in determination of orbits parameters of navigation spacecraft is identification of “jumps” of the phase measurements (processing and filtration of measurement sessions in PP) reflecting violation of reception and loss of the account of an integer of phase cycles in the phase measurement device.

In PP a check of phase measurements by means of the methods based on use of combinations of Melbourne-Vubben and Geometry-Free [2] is realized. However, the specified methods do not solve until the end a problem of “jumps” and in case of hit of such measurements to the BP module lead to solutions with an unsatisfactory accuracy. A way out is development of the expert and diagnostic system (EDS) for a specific case. Now the EDS prototype is developed for an assessment of the location module (PP module) [4-6] that can serve as a technological sample for development of EDS – “jumps”.

In other modules the random errors made when writing programs are possible. Their identification is made at a stage of testing of a complex as the compulsory procedure, which is carried out before using ACP in practice.

Conclusion

The considered approach to the solution of problems of failures in work of ACP BNS permits drawing the following conclusions.

1. For the analysis and the solution of the problems arising during the work of ACP BNS, it is recommended to consider a complex as a system with a hierarchical structure.
2. The generalized technology for the problems solution of work of functionally difficult systems in the form of eight stages, which basic elements are the components received by decomposition of the system on a functional sign, is offered. The degree of their structure is defined and the solution of private problems for the “infected” parts using the methods of the system analysis is proposed.
3. Possible options of problems in work of ACP BNS are considered, the complex components containing probable mistakes in the “infected” modules are specified and methods of their correction are offered.
4. The offered technological diagram for the problems solution within the systems concept can find application in case of evaluation of the work of hardware-software systems of the space branch.

References

1. Tarasenko F.P. *Prikladnoy sistemnyy analiz* [An applied system analysis]. Manual. Tomsk, Izdatel’stvo Tomskogo universiteta, 2004. (in Russian)

2. Genike A.A., Pobedinskiy G.G. *Global’nye sputnikovye sistemy, opredelenie mestopolozheniya i ikh primenenie v geodezii* [Global satellite systems, positioning and their application in geodesy]. Moscow, Kartgeotsentr, 2004. (in Russian)

3. Metodika sistemnogo analiza [A system analysis method]. Paper. Mht. 2014. (in Russian)

4. Betanov V.V., Larin V.K., Pozyaeva Z.A. Baza znaniy dlya programmnogo modulya opredeleniya

mestopolozheniya priemnika [A knowledge base for a program module for receiver location detecting]. *Sb. statey No. 6 "ITMiVT im. S.A. Lebedeva RAN"* [Proceedings No. 6 of JSC "IPMCE"], 2015. (in Russian)

5. Betanov V.V., Larin V.K., Pozyaeva Z.A. K voprosu analiza prichin vozniknoveniya sboev v apparatno-programmnom komplekse utochneniya effemeridno-vremennoy informatsii GNSS [The analysis of error occurrence causes in the GNSS hardware-software complex of ephemeris and time data refinement]. *Nauchno-technicheskiy zhurnal "Raketno-kosmicheskoe priborostroenie i informatsionnye sistemy"* [Scientific and technical journal "Rocket-space device engineering and information systems"]. 2014, Vol. 1, No. 1, pp. 55–60. (in Russian)

6. Betanov V.V., Larin V.K., Pozyaeva Z.A. Prototip ekspertnoy diagnosticheskoy sistemy poiska i korrektsii skachkov v bezraznostnykh fazovykh izmereniyakh [A prototype of the diagnostic expert system for search and correction of surges in non-difference phase measurements]. *Nauchno-technicheskiy zhurnal "Raketno-kosmicheskoe priborostroenie i informatsionnye sistemy"* [Scientific and technical journal "Rocket-space device engineering and information systems"]. 2014, Vol. 1, No. 3, pp. 73–81. (in Russian)

ROCKET-SPACE DEVICE ENGINEERING AND INFORMATION SYSTEMS
Vol. 3, Iss. 1, 2016, pp. 10–24

On Conceptual Fundamentals of Radio Navigation

A.A. Povalyaev

doctor of engineering science
Joint Stock Company "Russian Space Systems", Russia
e-mail: povalyaev_aa@spacecorp.ru

Abstract. Modern educational material and scientific literature dedicated to satellite navigation describes the operating principles of ground positioning systems and satellite radio navigation systems using such terms as pseudorange and pseudodelay. Pseudorange is defined as the product of pseudodelay by light speed. Pseudodelay τpd is defined as τpd , the difference between reception time of the navigation signal in the receiver timescale and the signal transmission time on the satellite timescale. However, the aforementioned sources do not contain any explanation regarding the following issues: how the receiver determines the value of , what the terms "timescale" and "time point on any given timescale" mean, and what the difference between the time according to the timescale and the actual physical time, mentioned in physics textbooks, is. Moreover, many sources define the pseudodelay P, either explicitly or implicitly, as a time interval without explaining whether it is meant to be an interval of the actual time or the time interval within any particular timescale.

Among the above-mentioned systems, nowadays the most complicated and, at the same time, the most advanced ones are the global navigation satellite systems (GNSS). Based on the critical review, the contradictions have been revealed in the paradigm used in modern educational material and scientific literature, which focus on the operating principles of the GNSS. A new paradigm based on defining the concepts of the timescale and satellite clock time is introduced. This new paradigm eliminates the revealed contradictions. A substantial simplification of the system development of the ground positioning systems is suggested based on the newly reintroduced concepts and paradigms.

Keywords: GNSS, pseudorange, pseudodelay, timescale, satellite clock time

Introduction

Nowadays several types of radio navigation systems (RNS), that are related by their conceptual fundamentals, such as global navigation satellite systems (GLONASS, GPS [1-14]), ground-based very low frequency systems (Omega, Alpha, Marshrut [4, 15]), and ground-based long wave systems (Loran-C, Chaika [4]) are operational. By their structure, all these systems are networks of either stationary or mobile radio navigation beacons (RNB) synchronously emitting navigation signals. The timestamps that are carried by these synchronically emitted signals are called the time scale of the system.

Among the above-mentioned systems, the most complicated and, at the same time, the most advanced ones are the global navigation satellite systems (GNSS). Accordingly, the description of the GNSS functioning principles requires using the most complicated conceptual fundamentals. For other RNS, the conceptual fundamentals are simpler and are considered a special case of the GNSS conceptual fundamentals.

1. Review of the concepts used in modern educational and scientific literature to describe the GNSS functioning principles

Sources [1-14] describe the functioning principles of the GNSS using the terms pseudorange and pseudodelay. Pseudorange is defined in all the sources as pseudodelay multiplied by the speed of light. Pseudodelay is defined in [1-14] as $\tau_{pd} = t_r - t^{tr}$, the difference between time t_r of receiving a navigation signal in the timescale of the receiver and time t^{tr} of its emission in the time scale of the navigation satellite. The method by which the receiver learns the value t^{tr}, the meaning of the terms "timescale" and "moment of time in a certain scale", and how the timescale differs from the physical time used in physics textbook is not revealed in [1-14]. Moreover, in the works [1-9], explicitly or not, the pseudorange $\tau_{pr} = t_r - t^{tr}$ is treated as a time interval, without explanation whether or not it is an interval of physical time or a time interval in a certain scale.

Fig. 1 is used, explicitly or not, to describe the GNSS functioning principles in [1-14]. Here, it was taken from the textbook [4]. Similar figures are used for these purposes in [2, 8, 9, 12].

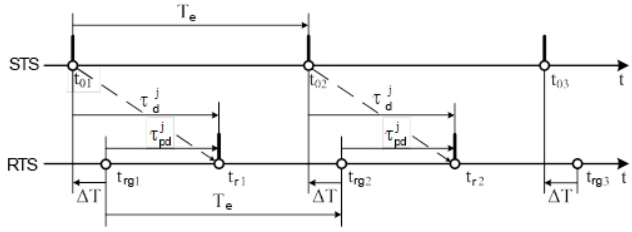


Fig. 1. Description of pseudodelay as a time interval

The GNSS functioning principles and the meaning of pseudodelay are explained by means of the Figures similar to Fig. 1 as follows: navigation satellites emit navigation signals at time points t_{01}, t_{02}, t_{03}, … with an interval of T_e in the system timescale (STS) (i.e., it is implicitly suggested that the navigation signal is a pulse signal). The time reference generator of the navigation receiver generates consequential time points t_{rg1}, t_{rg2}, t_{rg3}, … with the same period T_e defining the time scale of the navigation receiver (RTS). The signals emitted by the satellites at the time points t_{01}, t_{02}, t_{03}, are received at the time points t_{r1}, t_{r2}, t_{r3} …, according to the RTS (i.e., once again a pulse nature of the navigation signal is implied). For convenience, the emission time points t_{01}, t_{02}, t_{03} are connected with the reception moments t_{r1}, t_{r2}, t_{r3}, … with the inclined dashed arrows in Fig. 1.

In general, the RTS is displaced relative to the STS by a value unknown for the navigation receiver – ΔT, as shown in Fig. 1 and defined in [4] as $\Delta T = t_{0i} - t_{rgi}$.

The navigation receiver measures the delays of the satellite signals in its scale, i.e. it assumes that the signals are emitted by the satellites at the reference time points t_{rg1}, t_{rg2}, t_{rg3}, … in the RTS scale, while they are really emitted at the time points t_{01}, t_{02}, t_{03}, … . As a result, the navigation receiver measures not the delay τ_d^j of the signal propagation from the j-th satellite to the navigation receiver, but the pseudodelay τ_{pd}^j

$$\tau_{pd}^j = \tau_d^j + \Delta T , \quad j = \overline{1, J} \tag{1}$$

where J is a total number of satellites tracked by the navigation receiver. Therefore, according to Fig. 1, the pseudodelay τ_{pd}^j in the navigation receiver is formed by measuring the duration of the time interval, which begins at the time points t_{rgi} and ends at the time points t_{ri}.

The pseudodelay (1) multiplied by the light speed c results in pseudorange ρ^j

$$\rho^j = c\tau^j_{pd} = c\left(\tau^j_d + \Delta T\right) = R + c\Delta T =$$
$$= \sqrt{\left(x_{\ddot{i}} - x^j\right) + \left(y_{\ddot{i}} - y^j\right) + \left(z_{\ddot{i}} - z^j\right)} + \Delta R_{\ddot{i}}\ , \qquad (2)$$
$$\mathrm{j} = \overline{1,\mathrm{J}}$$

where x_r, y_r, and z_r are the unknown coordinates of the navigation receiver, x^j, y^j, z^j are the known coordinates of the j-th satellite acquired from its navigation message, and $\Delta R_r = \tilde{n}\Delta T$ is the unknown RTS displacement relative to the STS in meters.

Pseudodelays ρ^j measured by not less than four satellites (J ≥ 4) are used to make a set of equations (2) with four unknown x_r, y_r, z_r, and ΔR_r. The estimates $\hat{x}_r$, $\hat{y}_r$, $\hat{z}_r$, and $\Delta\hat{R}_r$ are found from the solution.

2. Critique of the conceptual model used in the modern educational and scientific literature to explain the GNSS functioning principles

The fundamental model given in Section 1, which for convenience is called the old model, employs the terms, the meaning of which is blurred and, at times, senseless. The use of these terms in the old model leads to contradictions. The following several examples prove this statement.

1. In modern GNSS, signals of navigation satellites are continuous periodic pseudorandom sequences (PRS). What is understood in this case under moments of emission and reception of continuous signals, because such signals are emitted and received at any time point, so can any time point be considered the moment of emission and reception?

2. If the pseudodelay τ^j_d of the signal in Fig. 1 exceeds T_e, the measurement of pseudodelay becomes ambiguous and can be expressed as $\tau^j_{pd} = \tau^j_{pda} + k^j T_e$, $\mathrm{j} = \overline{1,\mathrm{J}}$, where τ^j_{pda} is the ambiguous measurement of the pseudodelay formed by using the one of the time points t_{rgor1}, $t_{or2}t_{rg2}$, $t_{or3}t_{rg3}$…, that is the closest and preceding to the moment of reception, as a reference time point in the RTS; k^j is an indefinite integer. The same ambiguousness appears in a case when the modulus $|\Delta \mathrm{T}|$ of the RTS displacement relative to STS exceeds the period T_e.

Essentially, the ambiguousness of measurements of pseudodelay can be solved with the help of the approximated a priori data on the delay τ^j_d and the displacement of ΔT RTS relative to STS. At this, the total error of the approximated a priori data on the delay τ^j_d and the displacement of ΔT should not exceed $T_e/2$. By means of such a priori data, using the formula (1) a roughly approximated value τ^j_{pdr} can be calculated. Such rough evaluation of the pseododelay produces the following approximated equality $\tau^j_{pdr} \approx \tau^j_{pda} + k^j T_e$. The inaccuracy of this equality does not exceed $T_e/2$. Hence, it is easy to get a formula to calculate an indefinite integer $k^j = \left\langle\left(\tau^j_{pdr} - \tau^j_{pda}\right)/T_e\right\rangle$ $\mathrm{j} = \overline{1,\mathrm{J}}$, where operation $\langle \mathrm{x} \rangle$ means calculating the integer closest to x. In textbook [7], exactly this method for solving the ambiguousness of measurement of pseudodelay in GNSS is described, though it is not used in any real navigation receivers.

3. According to Fig. 1, measurement of pseudodelay is carried out at the time points of receiving the t_{r1}, t_{r2}, t_{r3}, …, signals emitted at the t_{01}, t_{02}, t_{03}, … time points. However, measurement of pseudodelay should be conducted simultaneously for not less than four satellites. Because of the difference in the distances between the satellites, the t^j_{r1}, t^j_{r2}, t^j_{r3}, … time points of receiving signals form the j-th satellite in the navigation field will not coincide with the time points t^k_{r1}, t^k_{r2}, t^k_{r3}, … of receiving signals from the k-th satellite. Therefore, if measurement of pseudodelay for each satellite is carried out at the moment when the signal from this satellite is received, such measurement for different satellites will occur at different time points. What time do the evaluated $\hat{x}_r$, $\hat{y}_r$, $\hat{z}_r$, and $\Delta\hat{R}_r$, found from solving the system of linear equations (2), correspond to?

4. A navigational receiver should measure pseudodelays for all the satellites at uniform time points of t_{meas}. It is possible to use, for example, the reference moments t_{rg1}, t_{rg2}, t_{rg3}, … shown in Fig. 1. The position of these moments on the RTS is defined (is set) by the signal of the generator of the navigation receiver. However, in order to make measurement of the pseudodelays corresponding to different navigational satellites to be carried out in the uniform moments of t_{meas}, it is necessary for the corresponding moments of signal emission from different satellites to differ and precede the moments of measurement t_{meas} for the period of signals propagation from different satellites to the navigation receiver. Further, for convenience, these time points will be called the preceding moments and will be designated as t^j_{p} , where the superfix j is the number of the satellite. Time of signal propagation from different satellites can vary depending on the position of the consumer and the altitude of the satellite orbit. Hence, it is clear that

the assumption about the pulse nature of a navigation signal introduced implicitly cannot be accepted, because satellites cannot emit pulses at the time points preceding the moments of measurement in receivers of all the great number of consumers.

To overcome the contradictions of old conceptual model described above, it is necessary to introduce new concepts considered in the following section.

3. Determination of the semantic content of the concepts “timescale” and “time on a scale”

The contradictions of the old conceptual model of radio navigation revealed above cannot be eliminated without determination of the semantic content of the concepts “timescale” and “time on a scale”. Despite the wide usage of these terms in literature [1–14], the author did not manage to find the definition of their semantic content there. Therefore, it is necessary to define the semantic content of the concepts “timescale” and “time on a scale”.

Further, to eliminate the confusion between the terms “time” and “time on a scale”, instead of the term “time”, we shall use the term “physical time” that means the ideal time that lapses absolutely evenly and that is used in physics textbooks. To designate the physical time, the symbol t is used.

The definition of the semantic content of the concepts “timescale” and “time on a scale” demands the definition of the semantic content of a “phase”, as well as the introduction of distinctions in definitions of the semantic content of the concept of phase. Again, despite the wide usage of the term “phase” in literature, the author did not manage to find the definition of its semantic content. In Textbook [16], a mathematical definition of the concept of phase for a harmonic or in a more general case a quasiharmonic process or signal is given $a(t) = A(t)\cos\varphi(t)$. Here, A(t) is the slow changing signal amplitude, φ(t) is the slow changing phase of the signal (in radians), which is an argument of a harmonic function. The argument φ(t) is determined by the instantaneous angular frequency of the ω(t) signal by the formula

$$\varphi(t) = \int_0^t \omega(x)dx + \varphi_0 \tag{3}$$

where ω(t) = 2πf (t), f(t) is the instantaneous frequency (in Hz). The first item in the right part (3) is defined as the phase increment on the time interval 0 ÷ t, and φ_0 is defined as the initial phase, i.e., the value of the phase φ(t) at the t=0 time moment. The concept of the instantaneous angular frequency ω(t) is the derivative of phase φ(t).

$$\omega(t) = \frac{d\varphi(t)}{dt} \tag{4}$$

For a strictly harmonic frequency signal, ω and f are constants, and the phase changes uniformly or linearly: $\varphi(t) = \omega t + \theta_0$. In case of a quasiharmonic signal, ω(t) is a slowly changing function of physical time t, and the phase changes nonuniformly. Expressions (3, 4) make it possible to geometrically interpret the phase of a quasiharmonic signal as a vector angle φ(t) of variable length A(t) rotating with slowly changing instantaneous angular speed ω(t).

Henceforth we shall use cycle as the more convenient unit of phase. Cycle 1 equals 2π radian.

The expressions (3) and (4) in this case will be transformed to $\varphi(t) = \int_0^t f(x)dx + \varphi_0$, $f(t) = \frac{d\varphi(t)}{dt}$.

In practice, there is often a need to consider varieties of the concept of phase, such as fractional and full phase. The fractional phase $\varphi_{frac}(t)$ is the phase lying within 1st cycle $0 \le \varphi_{frac} < 1$. The full phase $\varphi_{full}(t)$ can accept any values, i.e., contain besides a fractional phase $\varphi_{frac}(t)$ in its structure the integer number of cycles n(t) counted at every moment of physical time t from a starting point defined in advance.

$$\varphi_{full}(t) = \varphi_{frac}(t) + n(t) \tag{5}$$

When measuring a phase there can be situations when the integer n(t), which is a part of the full phase (5) differs from its true value by an uncertain number of cycles. Such full phase is called an ambiguously full phase.

A cyclic interval is an interval of physical time t, during which the full phase goes up by 1 cycle. In case of uneven change of phase, cyclic intervals will have various duration.

A fractional phase of a signal can be deduced from the full phase by adding or subtracting such an integer number of cycles for the result to be ranging from 0 to

1 cycle. It is known that addition of an integer number of 2π (an integer number of cycles) to the argument of the harmonic function does not change the value of this function. In this case, full and submultiple phases are equivalent to each other.

The concept of phase is applicable not only to harmonic or quasiharmonic signals. Fig. 2(b) shows change in time of the pseudorandom sequence (PRS) 11110 00100 11010 at uneven change of its phase, and Fig. 2a shows the schedule of this unevenly changing phase.

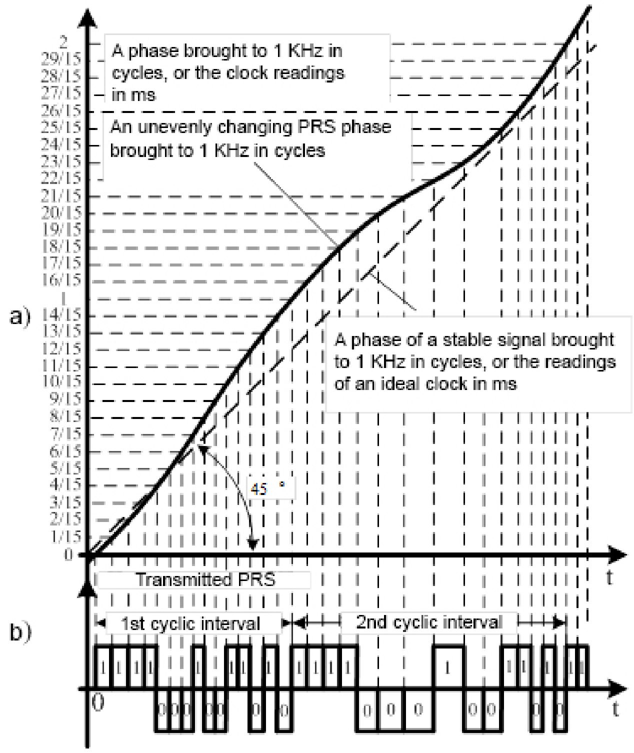


Fig. 2. PRS with uneven phase changing

Fig. 2b shows two identical in structure PRS 11110 00100 11010, located on cyclic intervals of physical time that are different in duration. Each of these cyclic intervals begins and ends at the time of the pulse leading edge corresponding to the first one in the group of four ones in a row in the PRS structure. The phase increment of these PRS on these different cyclic intervals is identical and equal to one cycle (2π rad).

The example shows that in order to define the concept of signal phase, it is necessary to distinguish the concepts of time period and structural period of a signal. Usually, a time period is thought of as a periodically repeating strictly identical interval of physical time. It is necessary to understand that an interval of physical time, on which all structure elements of a signal repeat, is a structural period. This period can have variable duration, but the signal phase increment on it is always equal to 1 cycle. In case of uniformly changing phase, the time intervals, on which the phase increment increases by 1 cycle, become identical, and then the concepts of structural and time periods coincide.

Based on the concept of structural period of a signal, the concept of fractional phase $\varphi_{frac}(t)$ of this signal in cycles is possible to define as the fraction of its structural period (cycle) observed at every moment of physical time t. A full phase of a signal is defined as a sum of the integer number of structural periods (cycles) and the fractional phase of the current structural cycle, which are observed on an interval from the beginning of the count of physical time until the present moment t.

In practice, determination of a quantitative value of physical time t is always carried out by means of a clock, which is understood to be a set of means and actions aimed to determine a quantitative value of physical time as a full phase of some periodically repeating process, which is the foundation of the specified clock. Oscillations of a pendulum, a signal of an electric generator, rotation of the Earth or radiation of atoms when they transit between different energy levels that defines an atomic time, can be used as such process. Hereafter, the process or a signal, which is the underlying operation principle of a clock, will be called a process or a signal of this clock. A quantitative determination of physical time t will be understood as determination of a number for each its moment T(t) that is the value of time at this moment. The specified number T(t) will be called the readings of the corresponding clock for the moment of physical time t under consideration.

Different clocks have different accuracy. The accuracy of a clock is determined by stability of the process of this clock. Therefore, there is a need to distinguish from the known natural processes the most stable one and use the readings of the clock built on its basis as the reference time. According to the present international agreements, the radiation of a cesium atomic beam standard is used as the process of the master clock. By definition, a second as unit of physical time equals 9192631770 periods of radiation corresponding to transition between two super thin levels of the main condition of a cesium-133 atom. However, if one compares the readings of two master clocks using different instances of the device counting the radiation periods of -133 atoms, it becomes evident that in course of time these clocks begin to disagree. This happens because any periodic process used for determination of a quantitative value of physical time has instability and this instability leads to the fact that in course of time even very precise clock disagree. Therefore, readings T(t) of any clock are only approximations to what physical time t is.

Readings of any clock are formed as the sum of their initial setting, the number of full phase increments of the clocks process on an interval of physical time from the moment of the initial setting to the present moment and possible corrections of the clock readings on the same interval of physical time. If readings of a clock are measured in seconds, then on the time interval from the moment of the initial setting a quantitative increment of time is defined as an increment of the full phase of the clock process brought to 1 Hz. This means that the increment of a full phase of the clock process divided by the nominal value of frequency of this process. For example, the increment of a full phase of radiation of the cesium atomic beam standard brought to 1 Hz is defined as the number of increments of the full phase of this radiation divided by 9192631770.

Therefore, readings of clock T(t) is a phase, the value of which is used for the quantitative measurement of physical time. At the time of taking of a clock's readings (i.e., at the time of measurement of the quantitative value of physical time) the phase is treated as time, and the unit of measurement of phase is replaced with the unit of measurement of time.

We shall define the concept of a timescale as moments of physical time t set by the readings of the clock, which are the basis of the scale under consideration [17, 18]. Then, the concept of time on a scale is defined as the readings of the clock, which are the basis of the scale for any moment of physical time t. At the same moment of physical time different clock can have different readings (different time on different scales) and at the different moments of physical time different clock can have identical readings (identical time on different scales). A timescale shift should be understood as the difference of the readings of a clock on one scale and a clock on another one at the same moment of physical time. At the same time, the difference of the clocks readings for the same moment of physical time should not be confused with an interval between the moments of physical time, at which the clock readings are identical. Since any clock is unstable, difference of the clock readings for the same moment of physical time generally is not equal to time interval between the moments of identical readings of this clock.

4. Description of the GNSS functioning principles based on the new conceptual model

The signals emitted by navigation satellites in the modern GNSS are the high frequency phase-modulated carrier oscillations in the range ~1.2–1.6 GHz. Modulation of the carrier oscillations is carried out by a double-layer signal. The lower layer is a continuous periodically repeating PRS, on which measurement of pseudoranges is carried out. A nominal period of these PRS in the open signals of GLONASS and GPS is equal to 1 ms. The upper layer is formed by binary 20 ms symbols of the navigation message, which inversely modulate periodically repeating PRS of the low layer. Formation of PRS in the onboard equipment of satellites is carried out from a signal of the high-stable atomic frequency standard. A full phase of PRS, emitted by each satellite and interpreted as the readings of clock, sets the onboard time scale (OTS) of this satellite.

According to the definition (5), a full phase $\varphi_{full}(t)$ of PRS for each present moment of physical time t is set by a fractional phase $\varphi_{frac}(t)$ by this PRS and an integer n(t) of the full periods of PRS, which are keeping within an interval from some conditional beginning defined in advance before the present moment t. For example, in GLONASS system such conditional beginning are 00 hours, 00 min 00 sec from January 1, 1996, according to the Moscow standard time defined as UTC (SU)+3 hours. For setting an integer number of cycles n(t), special signals of timestamps and digitization of these timestamps ζ^j are put in navigation messages of satellites. A signal of a timestamp is an a priory defined sequence of pulses in the navigation message. The moment of emergence of the trailing or front edge of a certain pulse in the signal of a timestamp is the timestamp itself. Further, this moment will be called a timestamp moment. For example, in the GLONASS system, a timestamp moment is the moment of the trailing edge of the last pulse of a signal of a timestamp, and in GPS, a stamp moment is the moment of the forward front of the first pulse of a signal of a timestamp. Digitization of a timestamp moment is the readings of clock of the j-th satellite on its board at this moment. Fig. 3 shows characteristic time points in the transmitted (Fig. 3a) and received (Fig. 3b) signal. In Fig. 3a, the moments of the beginning of the periods of PRS in the radiated signal, or in other words the moments of milliseconds according to RTS, are shown with the arrows pointed up. A timestamp moment is italized with a big arrow. The symbol ζ^j shown over a big arrow designates digitization of this timestamp moment. The corresponding time points in the received signal are shown in Fig. 3b in the form of hyphens with crosses. The hyphens focused down on Fig. 3b show some in general case random timestamp moments of RTS some generally. In general, it is not supposed that these time points have any digitization.

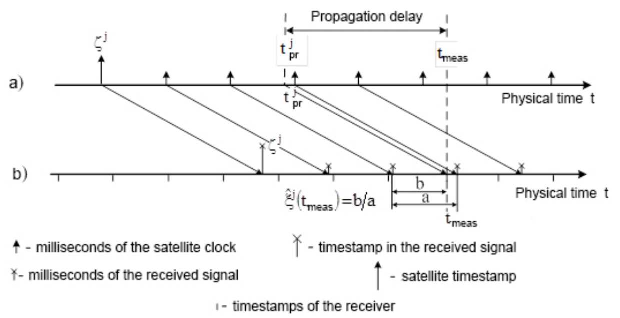


Fig. 3. Timestamps and their digitization in the radiated and received GNSS signals

In the moment of time on RTS marked in Fig. 3b with a symbol t_{meas}, measurement of a fractional phase $\hat{\xi}^j(t_{meas})$ of PRS of the j-th satellite is taken in the navigation receiver. This phase expressed in cycles is equal to the period share b/a in the received signal, which has passed from the beginning of the PRS period until physical time t_{meas}. The value of $\hat{\xi}^j(t_{meas})$ cannot be displayed in Fig. 3, as for this purpose it is necessary to allow a vertical axis, along which the phase (the reading of clock) will be laid off. Such laying off will be made further in Fig. 5 in the form of the clock readings.

As it is seen from Fig. 3, the value of a fractional phase $\hat{\xi}^j(t_{meas})$ measured in the receiver with an accuracy up to an integer of milliseconds and errors of tracking is in agreement with the readings of satellite clock at the time of precedence t_p^j to the measurement moment t_{meas}. The assessment $\hat{T}^j(t_p^j)$ of the complete clock readings of the j-th satellite in seconds at the time of precedence is calculated in the processor of the navigation receiver according to the formula

$$\hat{T}^j(t_{pr}^j) = 10^{-3}\left(\zeta_{m\sec}^j + n^j + \hat{\xi}^j(t_{meas})\right), \quad j = \overline{1, J} \tag{6}$$

where $\zeta_{m\sec}^j$ is the digitization of the last accepted timestamp expressed in milliseconds; n^j is the whole amount of the periods of the accepted PRS lying on a time interval from the last accepted and digitized timestamp until measurement t_{meas}, (in the example, shown in Fig. 3, $n^j = 2$). The actions described above and the funds allocated for this purpose for estimation of the clock readings of the j-th satellite at the time of precedence can be called channel clock of the j-th satellite in the navigation receiver, and the estimates determined by a formula (6) are called readings of this clock. For convenience of the further consideration of the channel clock reading, relating to the time of measurement t_{meas}, will be designated as $T_{chan}^j(t_{meas})$, i.e., $T_{chan}^j(t_{meas}) = \hat{T}^j(t_p^j)$. It is possible to interpret the calculations by a formula (6) as a solution of millisecond ambiguity of estimates $\hat{\xi}^j(t_{meas})$ of the readings of satellite clock.

The channel clock under consideration is schematically shown in Fig. 4 with four small circles. It is obvious that the number of channel clock of the navigation receiver is equal to the number of its channels.

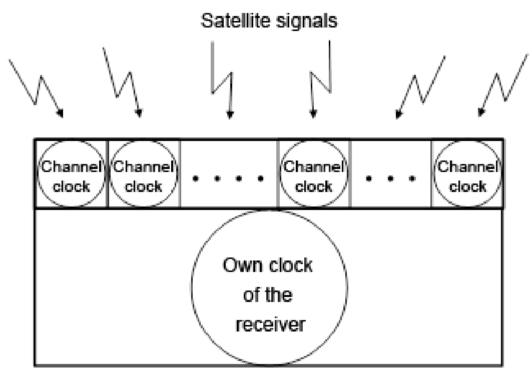


Fig. 4. A model of the navigation receiver as a set of clocks

Apart from the channel clocks in the navigation receiver, its own clock, which is schematically shown in Fig. 4 with a lower big circle, is used. Own clock of the receiver is the clock, which readings define the moments of carrying out measurements, i.e., sets a receiver time scale.

The coordinates of the navigation receiver and the clock reading of the system for a moment t_{meas}. can be determined based only on the readings of the channel clock. The receiver derives the values of the polynomial models coefficients from navigation messages allowing one to calculate the estimates of shifts $\Delta\hat{T}^j_{sys}\left(t^j_p\right)$ of the clock readings of all tracked satellites relative to the clock readings of the system at the time of precedence t^j_{pr}. Further, by means of these estimates, the receiver calculates the estimates of the clock readings of the system for the precedence moments:

$$\hat{T}_{sys}\left(t^j_{pr}\right)=\hat{T}^j\left(t^j_{pr}\right)-\Delta\hat{T}^j_{sys}\left(t^j_{pr}\right)= =T^j_{chan}\left(t_{meas}\right)-\Delta\hat{T}^j_{sys}\left(t^j_{pr}\right), \quad j=\overline{1,J} \tag{7}$$

Using the parameters of mathematical models of satellites movement, transferred in navigation messages and the estimates $\hat{T}_{sys}\left(t^j_p\right)$, the receiver calculates the coordinates $x^j\left(t^j_p\right)$, $y^j\left(t^j_p\right)$, and $z^j\left(t^j_p\right)$ of each j-th satellite for the precedence moment corresponding to this satellite. It should be emphasized that in order to calculate the coordinates of each j-th satellite, the receiver uses not the value t^j_{pr} of physical time for the precedence moment, but the assessment $\hat{T}_{sys}\left(t^j_p\right)$ of the clock readings of the system for this moment, or, otherwise, to calculate the navigation satellites coordinates in GNSS, time according to a system scale is used, but not the value of physical time.

It is possible to write down the following obvious equality for the moments of physical time of t neglecting for simplicity signal delays in the atmosphere:

$$t_{meas}-t^j_p=\frac{R^j}{c}; \qquad j=\overline{1,J} \tag{8}$$

where $t_{meas}-t^j_p$ is a delay of signal distribution; R^j is a distance between the points, which the j-th satellite occupied at the moment of precedence t^j_{pr} and the navigation receiver at the time of measurement t_{meas}; c is velocity of light. Because the clock of the system is very exact, the moments of physical time t_{meas}, t^j_{pr} in (8) can be replaced with the readings $T_{sys}\left(t_{meas}\right)$, $T_{sys}\left(t^j_p\right)$ of the clock of the system at the same moments. Taking this into account, an initial equality (8) can be presented with a high precision in the following form:

$$T_{sys}\left(t_{meas}\right)-T_{sys}\left(t^j_p\right)=\frac{R^j}{c}, \qquad j=\overline{1,J} \tag{9}$$

Replacing in (9) the value $T_{sys}\left(t^j_p\right)$ with the corresponding assessment (7), the following equation for each j-th satellite is received:

$$T_{sys}\left(t_{meas}\right)-\frac{R^j}{c}=\hat{T}^j\left(t^j_p\right)-\Delta T^j_{sys}\left(t^j_p\right), \quad j=\overline{1,J} \tag{10}$$

Expressing in (10) the distance R^j through the coordinates of the navigation receiver $x_r\left(t_{meas}\right)$ $y_r\left(t_{meas}\right)$, $z_r\left(t_{meas}\right)$, at the time of measurement t_{meas} and the coordinate of the j-th satellite $x^j\left(t^j_p\right)$, $y^j\left(t^j_p\right)$, $z^j\left(t^j_p\right)$, at the moment of precedence t^j_{pr}, from (10), the following system of the nonlinear equations concerning unknown $T_{sys}\left(t_{meas}\right)$, $x_r\left(t_{meas}\right)$, $y_r\left(t_{meas}\right)$, $z_r\left(t_{meas}\right)$ is received:

$$T_{sys}\left(t_{meas}\right)-\frac{1}{c}\sqrt{\left(x_r\left(t_{meas}\right)-x^j\left(t^j_{pr}\right)\right)^2+\left(y_r\left(t_{meas}\right)-y^j\left(t^j_{pr}\right)\right)^2+\left(z_r\left(t_{meas}\right)-z^j\left(t^j_{pr}\right)\right)^2}= =\hat{T}^j\left(t^j_{pr}\right)-\Delta\hat{T}^j_{sys}\left(t^j_{pr}\right), \qquad j=\overline{1,J} \tag{11}$$

To find four unknown of $T_{sys}\left(t_{meas}\right)$, $x_r\left(t_{meas}\right)$, $y_r\left(t_{meas}\right)$, and $z_r\left(t_{meas}\right)$, it is necessary to have not less than four equations of a type (11), i.e., to carry out measurements at the same time not less than on four satellites. Solving a system (11) under these conditions, the estimates $\hat{x}_r\left(t_{meas}\right)$, $\hat{y}_r\left(t_{meas}\right)$, and $\hat{z}_r\left(t_{meas}\right)$ of the coordinates of the navigation receiver and the assessment $\hat{T}_{sys}\left(t_{meas}\right)$ of the clock readings of the system at the time of measurement t_{meas}, which can be used further as digitization of time moment, are obtained.

In the system (11), own clock readings of the navigation receiver are not used, i.e., it is not required that timestamps of the receiver should be digitized. However, in practice, usually it is required to carry out navigation definitions not in randomly set measurement moments t_{meas}, but in regular intervals. To count these intervals, it is necessary to use own clock of the navigation receiver shown in Fig. 4 with a big circle. In this case timestamps of the navigation receiver are digitized by readings of its own clock, and instead of the readings of channel clock, a concept of a pseudodelay is employed. At the same time, it is unimportant, how precisely these digitizations coincide with the clock readings of the system $T_{sys}\left(t_{meas}\right)$ at the same moment t_{meas}. The pseudodelay $\tau^j_{pd}\left(t_{meas}\right)$ according to the j-th satellite is determined as a difference of the readings of own clock of the receiver $T_r\left(t_{meas}\right)$ at the time of measurement t_{meas} and the clock readings of the j-th satellite at the time of precedence t^j_{pr} :

$$\tau^j_{pd}\left(t_{meas}\right)=T_r\left(t_{meas}\right)-T^j\left(t^j_p\right), \quad j=\overline{1,J} \tag{12}$$

The initial value $T_r\left(t_{meas}\right)$ can be set randomly, taken from any suitable source, or just calculated according to the following approximate formula:

$$T_r\left(t_{meas}\right)=\zeta^j+0,08\ c \tag{13}$$

where ζ^j is digitization of the next accepted timestamp from any satellite. The error of initial digitization of timestamps of the receiver by a formula (13) does not exceed ± 30 ms.

A pseudodelay assessment $\hat{\tau}^j_{pd}\left(t_{meas}\right)$ formed in the receiver is defined as a difference of the readings of own clock of the receiver and the readings of its channel clock corresponding to the j-th satellite at the time of measurement t_{meas}:

$$\hat{\tau}^j_{pd}\left(t_{meas}\right)=T_r\left(t_{meas}\right)-T^j_{chan}\left(t_{meas}\right)=T_r\left(t_{meas}\right)-\hat{T}^j\left(t^j_p\right), \quad j=\overline{1,J} \tag{14}$$

For any moment of physical time t, the expression (12) can be rewritten as follows:

$$\tau^j_{pd}\left(t\right)=T_r\left(t\right)-T^j\left(t^j_p\right), \quad j=\overline{1,J} \tag{15}$$

where the symbol t^j_{pr} in that case designates the precedence moment to a present situation of physical time t. For any moment of this time, it is possible to introduce the concepts of shifts of the clock readings of the satellite and own clock of the receiver:

$$\Delta T^j\left(t\right)=T^j\left(t\right)-T_{sys}\left(t\right), \quad j=\overline{1,J}$$
$$\Delta T_r\left(t\right)=T_r\left(t\right)-T_{sys}\left(t\right) \tag{16}$$

Using (16), the clock readings of the j-th satellite and the reading of own clock of the receiver $T_r\left(t\right)$, the following can be expressed through the shifts:

$$T^j\left(t\right)=T_{sys}\left(t\right)+\Delta T^j\left(t\right), \quad j=\overline{1,J}$$
$$T_r\left(t\right)=T_{sys}\left(t\right)+\Delta T_r\left(t\right) \tag{17}$$

Substituting (17) into (12), the following expression for pseudodelay is obtained:

$$\begin{aligned}\tau^j_{pd}\left(t_{meas}\right)&=T_r\left(t_{meas}\right)-T^j\left(t^j_{pr}\right)=\\&=T_{sys}\left(t_{meas}\right)-T_{sys}\left(t^j_{pr}\right)+\Delta T_r\left(t_{meas}\right)-\Delta T^j\left(t^j_{pr}\right)=\\&=\Delta T_{sys}\left(t^j_{pr}\div t_{meas}\right)+\Delta T_r\left(t_{meas}\right)-\Delta T^j_{sys}\left(t^j_{pr}\right)\end{aligned} \tag{18}$$

where

$$\Delta T_{sys}\left(t^j_p\div t_{meas}\right)=T_{sys}\left(t_{meas}\right)-T_{sys}\left(t^j_p\right) \tag{19}$$

is increment of the clock readings of the system on the time interval $t^j_{pr}\div t_{meas}$, duration of which is equal to the delay $\tau^j_d\left(t_{meas}\right)=t_{meas}-t^j_p$ of a signal propagation from the point occupied by the j-th satellite in the preceding moment t^j_{pr} until the point occupied by the receiver in the moment of measurement t_{meas}.

Fig. 5 shows the pseudodelays change as the functions of physical time t for two satellites with Nos. j and k.

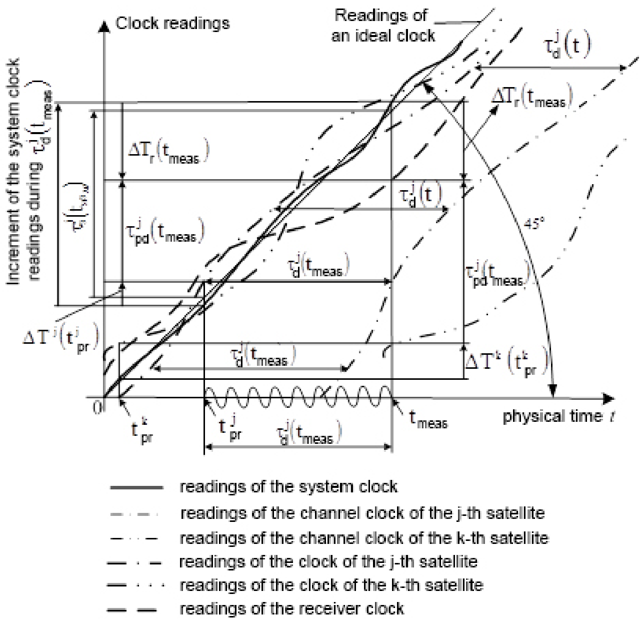


Fig. 5. Change of pseudodelays of the j-th and k-th satellites as the functions of physical time t

A pseudorange assessment $\hat{\rho}^j\left(t_{meas}\right)$ according to the j-th satellite is defined as a pseudodelay assessment $\hat{\tau}^j_{pd}\left(t_{meas}\right)$ (14) multiplied by light speed c:

$$\begin{aligned}\hat{\rho}^j\left(t_{meas}\right)&=c\cdot\hat{\tau}^j_{pd}\left(t_{meas}\right)=\\&=c\left(T_r\left(t_{meas}\right)-T^j_{chan}\left(t_{meas}\right)\right)=, \qquad j=\overline{1,J}\\&=c\left(T_r\left(t_{meas}\right)-\hat{T}^j\left(t_{meas}\right)\right)\end{aligned} \tag{20}$$

From (14) and (20) it is easy to see that at strict synchronism of the clock rate of the receiver and satellites, a pseudodelay assessment becomes a delay assessment, and a pseudorange assessment turns into a range assessment.

Subtracting the readings of own clock of the receiver $T_r\left(t_{meas}\right)$ from the left and right parts (11), the following is received:

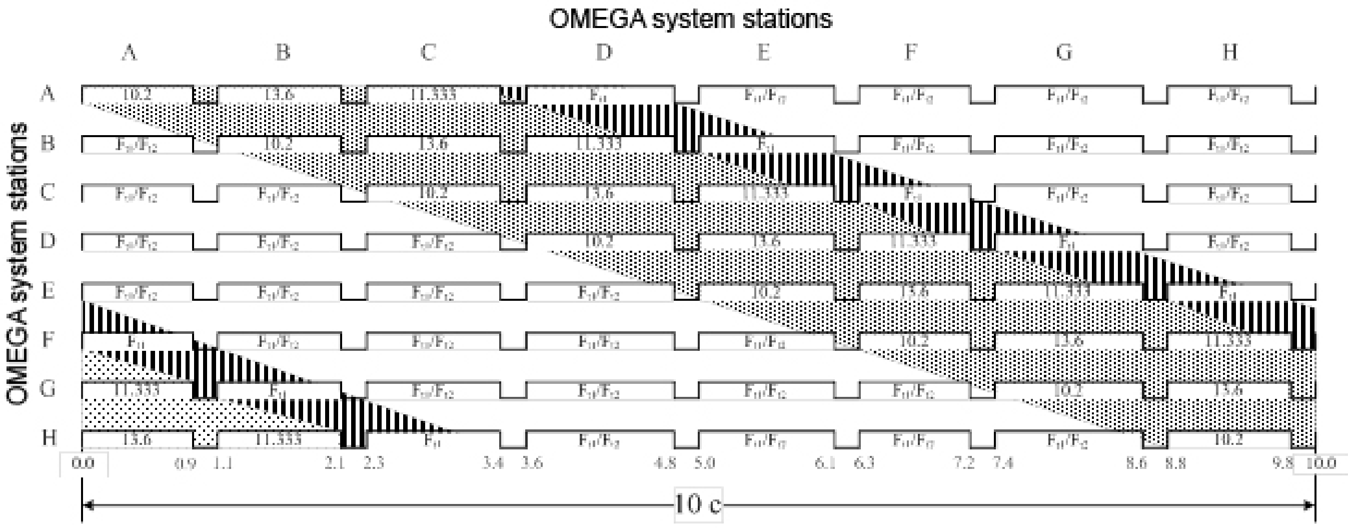


Fig. 6. The structure of the navigation frame RNS OMEGA

$$\sqrt{\begin{pmatrix} x_r(t_{meas})- \\ -x^j(t^j_{pr}) \end{pmatrix}^2 + \begin{pmatrix} y_r(t_{meas})- \\ -y^j(t^j_{pr}) \end{pmatrix}^2 + \begin{pmatrix} z_r(t_{meas})- \\ -z^j(t^j_{pr}) \end{pmatrix}^2} + \Delta R_r(t_{meas}) = \hat{\rho}^j(t_{meas}) + c\cdot\Delta T^j_{sys}(t^j_{pr}),$$
$$j=\overline{1,J} \tag{21}$$

By the definition (16), the contents of the parentheses standing in the left part of the expression (21) is a shift $\Delta T_r(t_{meas})$ of the clock readings of the receiver concerning the clock readings the system at the time of measurement t_{meas}. The following should be introduced to the product of this shift and light velocity c:

$$\Delta R_r(t_{meas}) = c\cdot\left(T_r(t_{meas}) - T_{sys}(t_{meas})\right) = c\cdot\Delta T_r(t_{meas}) \tag{22}$$

where $\Delta R_r(t_{meas})$ is a shift of the readings of the receiver clock concerning the readings of the system clock expressed in meters. The contents of the parentheses in the right part (21) by definition (14) is a pseudodelay assessment, and its multiplication by the light speed according to (20) is a pseudorange assessment. As a result, it is possible to rewrite (22) as the following:

$$\sqrt{\left(x_r(t_{meas}) - x^j(t^j_{pr})\right)^2 + \begin{pmatrix} y_r(t_{meas})- \\ -y^j(t^j_{pr}) \end{pmatrix}^2 + \begin{pmatrix} z_r(t_{meas})- \\ -z^j(t^j_{pr}) \end{pmatrix}^2} + ,$$
$$+c\cdot\begin{pmatrix} T_r(t_{meas})- \\ -T_{sys}(t_{meas}) \end{pmatrix} = c\cdot\begin{pmatrix} T_r(t_{meas})- \\ -\hat{T}^j(t^j_{pr}) \end{pmatrix} + c\cdot\Delta T^j_{sys}(t^j_{pr}),$$
$$j=\overline{1,J} \tag{23}$$

The system of the equations (23) is completely equivalent to the earlier received system (11). The unknown in this system, which are subject to definition, are the receiver coordinates $x_r(t_{meas})$, $y_r(t_{meas})$, and $z_r(t_{meas})$, and a shift $\Delta R_r(t_{meas})$ (22) of the readings of the receiver own clock concerning the readings of the system clock expressed in meters. From (22) it is seen that it is possible to calculate the assessment $\hat{T}_{sys}(t_{meas})$ of the readings of the system clock at the time of measurement t_{meas}, as $T_{sys}(t_{meas}) = T_r(t_{meas}) - \Delta\hat{R}_r(t_{meas})/\tilde{n}$ by means of a shift assessment $\Delta\hat{R}_r(t_{meas})$ and the readings of the receiver clocks $T_r(t_{meas})$.

5. Application of the new conceptual model to ground radio navigational systems

Ground RNS can be considered as a simplified GNSS option. Simplification is that transmitters of navigation signals in these systems are fixed and, therefore, the coordinates of transmitters can be placed in memory of navigation receivers during their production. Timescales of all transmitters of ground RNS, as well as in GNSS, are synchronized with high precision with a system timescale. The structures of the transmitted navigation signals in various ground RNS can differ greatly, but the main principle remains invariable: the phases of the radiated radio navigation signals carry information on the readings of the system clock and set time scales of the received signals in navigation receivers.

It is not possible to consider all types of ground RNS in the article. Therefore, further, as an example, the application of a new conceptual fundamental for the ground super long-wave RNS OMEGA [15] will be considered.

The structure of a navigation frame of the RNS OMEGA is shown in Fig. 6. [4, 15, 19].

Eight stations of the RNS OMEGA are given in the Latin letters A, B, C, D, E, F, G, and H, which with time shift emit radio pulses with an average duration of 1.25 sec., filled with coherent harmonic oscillations at frequencies of 10.2, 13.6, and 34/3 kHz respectively. Navigation radio pulses are highlighted in Fig. 6 with dotted pattern. Radio pulses with unique frequencies used for stations identifications are shown in vertical hatch in Fig. 6. The rest four radio pulses of each line of the navigation frame are used for exchange between the stations [19].

The main navigation frequency in the RNS OMEGA is 10.2 kHz, i. e., a signal phase of this frequency is identified with the readings of the synchronically working clock of the system stations (system clock). However, a signal period of frequency 10.2 kHz equal to 0.09804 ms is very small and, consequently, a fractional phase of this frequency measured in the navigation receiver, carry information on the readings of the system clock of the moment of precedence with an accuracy up to an integer of periods 0.09804 ms. In other words, the readings of the system clock of the precedence moments identified with a fractional phase of frequency 10.2 kHz are measured in the navigation receiver ambiguously. To solve this ambiguity, one applies measurement of fractional phases on difference frequencies: 13.6–10.2 = 3.4 kHz (period of 0.294117 ms) and 34/3–10.2 = 34/30 kHz (period of 0.88235 ms) [4, 15]. The least common multiple of these periods is 60/17 ms, i. e., a time interval, through which fractional phases of harmonic signals of all navigation radio pulses transform into null, equals 60/17 ms [15].

Time moments following the system clock through 60/17 ms are called the RNS OMEGA eras. If measurement of a signal phase of a navigation frame, which has duration of 10 sec (see Fig. 6), is involved to solve ambiguity, then a signal phase at a frequency of 10.2 kHz can be unequivocally measured in the navigation receiver within 30 sec. Solving 30 sec ambiguity is possible by means of normal clock.

It should be noted that in literature [4, 15] when stating the methods of ambiguity solving in the RNS OMEGA, it is said that either ambiguity of range measurement or delays of a signal is solved. Obviously, there is a question about the kind of a range or delay of a signal, if in the noninterrogative systems, which the RNS OMEGA belongs to, a range or delay of a signal cannot be fundamentally measured. Actually, just as in GNSS, ambiguity not of range, but the readings of the clock of the RNS OMEGA stations for the precedence moments is solved. Taking into account that the clock of the stations is synchronized with the system clock, it is possible to speak about disambiguation of the readings of system clock for the precedence moments. The readings of the stations clock is transferred continuously in the phases of harmonic carriers, filling navigation radio pulses. Nevertheless, in literature [4, 15], the concepts of the moments of precedence and the readings of the stations and system clock are not introduced. In addition, it is not specified that phases of harmonic carriers, filling the emitted navigation radio pulses, carry the information on the readings of the stations clock in the precedence moments to the moments of phases measurement of these carriers in navigation receivers. As a result, the concept (the readings of the stations clock for the precedence moments), which ambiguity of measurements is solved, in literature [4, 15] is absent. For this reason, the authors [4, 15] are forced to speak about disambiguation of range measurements.

In addition, in literature [15], it is specified that for disambiguation of pseudoranging measurements, it is necessary to have a priory data not only on the coordinates of the navigation receiver, but also on shift of its timescale. The necessity in having a priory data for solving disambiguation of measurements of pseudoranges in the RNS OMEGA, as well as the need in having such data for solving disambiguation of measurements of pseudoranges in GNSS, about which the manual [7] states, is a mistake. This mistake results from desire of the authors of literature [7, 15] to solve ambiguity of measurements of pseudoranges. Such desire is natural, as in [7, 15] the concepts of the moments of precedence and the readings of the stations (satellites) clock for these moments are not used. However, if to solve ambiguity not of pseudoranges, but that of the readings of the stations (satellites) clock for the precedence moments (that, actually, is done in GNSS), then no a priory data in GNSS are required. In the RNS OMEGA, only rough a priory data on the readings of the system clock for the precedence moments will be required. The errors of these

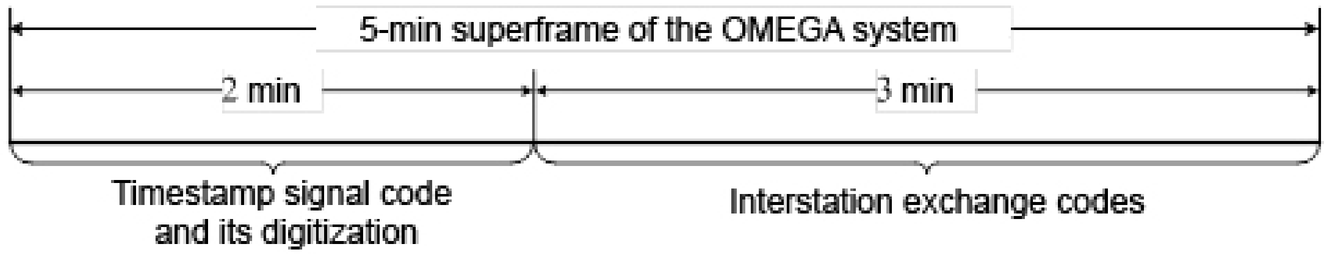


Fig. 7. Five-minute superframe of the RNS OMEGA

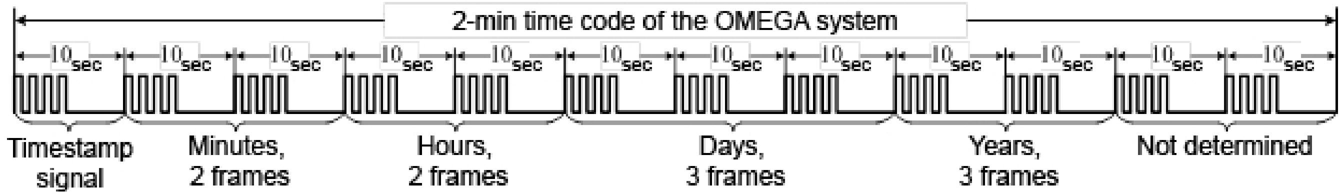


Fig. 8. Two-minute time code of the RNS OMEGA

rough a priory data on the module should not exceed 15 seconds. A priory data with such big errors can be received by means of the regular clock, which is periodically set at the signals of the exact time, being broadcast.

In GNSS, disambiguation of the clock readings of satellites is carried out using timestamps and their digitizations (see Section 4 of this article). In the RNS OMEGA, digitization of timestamps is absent, and the moments of the beginning of 30 sec intervals corresponding to the moments of the RNS OMEGA eras can be used as timestamps. On each such interval, there are three in succession frames. As it was shown earlier, using the measurements of fractional phases on the main and difference frequencies and the readings of external clock, which shift concerning the readings of the system clock does not exceed 15 sec, permits one to solve completely the ambiguity of clock readings of the stations during the precedence moments. After disambiguation solving, unequivocal values of pseudotime delays as differences between the clocks readings of the navigation receiver at the time of measurement and the solved clock readings of the stations within the precedence moments can be formed. At the same time, it is no matter how much a time scale of the navigation receiver is offset concerning a system scale.

If digitizations of timestamps are introduced into the structure of the RNS OMEGA navigation signal, so in this case to solve disambiguation of clock readings of the stations, as well as in GNSS, no a priory data will be required. The paper considers [19] the offers on digitizations of timestamps into the structure of the RNS OMEGA navigation frame. For this purpose, it is offered to introduce a concept of five-minute superframes shown in Fig. 7. In the superframe, the first two minutes are separated for signal transmission of a timestamp designating the beginning of a superframe and digitization of this stamp. The last 3 minutes of a superframe are separated for transfer of codes of interstations exchange.

Each line of the RNS OMEGA navigation frame (see Fig. 6) includes eight radio pulses, four of which were not used for any purposes at the time of publication [19]. These four radio pulses in work [19] are offered to use for transmitting one tenth of a figure by means of a binary code. Unites and nulls of a binary code are offered to transmit via radio pulses frequency change. To do this, two individual frequencies are given for each RNS OMEGA station. On the two-minute time interval, given for transmitting a signal of a timestamp and its digitization, there are 12 frames. Hence, it is possible to transmit 12 decimal digits on this interval. A signal of a timestamp designating the beginning of a superframe is transmitted in the first frame of a superframe. A number of a minute in the hour is transmitted in the frames Nos. 2 and 3. A number of an hour per day is transmitted in the frames Nos. 4 and 5. Three frames Nos. 6, 7, and 8 are separated for transmitting a number of a day of a year. It is proposed to transmit a number of a year of a century in the frames Nos. 9 and 10. Usage of the rest frames Nos. 11 and 12 is not determined. The structure of a two-minute time code of the RNS OMEGA offered in [19] is shown in Fig. 8.

If timestamps of the signals, radiated by the RNS OMEGA stations, are considered as the moments of the beginning of a 30-seconds intervals coinciding with the moments of the RNS OMEGA eras, so each 10th stamp will be digitized. At such digitization, solving the disambiguity of the clock readings of the stations, counting down in seconds from the beginning of the current year (i.e., with an accuracy up to a whole number of years from the century beginning) should be calculated according to the formula similar to (6)

$$\begin{aligned}\hat{T}^{j}\left(t_{pr}^{j}\right) &= N_{day}\cdot 86400+N_{h}\cdot 3600+ \\ &+N_{\min}\cdot 60+N_{30}\cdot 30+\xi_{c}^{j}\left(t_{meas}\right)\end{aligned}, \quad \mathrm{j}=\overline{1,\mathrm{J}} \tag{24}$$

where N_{day} is an amount of days in a year finished up to the beginning of the current superframe; N_h is an amount of hours in the current day finished up to the moment of the current superframe; $N_{\min}$ is an amount of minutes in the current hour finished up to the beginning of the current superframe (the values N_{day}, N_h, $N_{\min}$ are separated from the received superframe); N_{30} is an amount of 30-minutes intervals from the beginning of a current superframe finished up to the moment t_{meas} of carrying out the measurements (the value N_{30} is calculated in the receiver); $\xi_{\tilde{n}}^{j}\left(t_{meas}\right)$ is a phase of the received signal at the frequency 10.2 kHz unequivocally expressed within 30 sec, referring to 1 Hz (the value $\xi_{\tilde{n}}^{j}\left(t_{meas}\right)$ is determined via solving the disambiguity of the measured value of the fractional phase of a signal frequency 10.2 kHz using the measurements of fractional phases at difference frequencies [4, 15]).

Subtracting the value $\hat{T}^{j}\left(t_{p}^{j}\right)$ calculated according to the formula (24) based on the clock readings of the receiver $T_r\left(t_{meas}\right)$ in the moment of the measurement t_{meas}, a univocal value of pseudodelay corresponding to the j-th station of the RNS OMEGA is given. It is possible to offer one more method (an easier one) to solve the disambiguity of measurement of a fractional phase of a signal frequency 10.2 kHz to resume a whole value of the clock readings $T^{j}\left(t_{p}^{j}\right)$ of the stations during the moments of precedence t_{p}^{j}. It should be suggested that the timestamps of the signals radiated by the RNS OMEGA stations are the moments of the beginning of the 10 sec frames shown in Fig. 6. In this case, each 30th stamp will be digitized. Two counters should be included into the receiver's equipment. The first counter will determine a whole number N_f of frames, inserting onto time intervals from the beginning of the superframe until the moment of measurement t_{meas} of a fractional phase of a signal frequency 10.2 kHz. The second counter will determine a whole number $N_{10.2}$ of signal frequency periods 10.2 MHz inserting onto the time interval from the beginning of a frame until the moment of measurement t_{meas}. In this case, solving the disambiguity of clocks readings of a station counted down in seconds from the beginning of the current year (i. e., with an accuracy up to a whole number of years from the beginning of the century) can be carried out with a formula similar to the formulae (6, 24).

$$\begin{aligned}\hat{T}^{j}\left(t_{pr}^{j}\right) &= N_{day}\cdot 86400+ \\ &+N_{h}\cdot 3600+N_{\min}\cdot 60+N_{f}\cdot 10+, \quad \mathrm{j}=\overline{1,\mathrm{J}} \\ &+\frac{N_{10.2}+\xi_{10.2}^{j}\left(t_{meas}\right)}{10.2}\cdot 10^{-3}\end{aligned} \tag{25}$$

It should be noted that in (25) for solving the disambiguity $T^{j}\left(t_{p}^{j}\right)$, the measurements of fractional phases on the measurement frequencies 13.6 and 34/3 kHz are not used. Thus, applying a new fundamental concept of readings $T^{j}\left(t_{p}^{j}\right)$ of stations clock during the moments of precedence t_{p}^{j} enables one to abandon carrying out measurements on the frequencies 13.6 and 34/3 kHz and, hence, ease significantly the RNS OMEGA by introducing digitization of timestamps into the structure of the RNS OMEGA navigation signal and two additional counters into the equipment of the receiver.

The analysis being carried out shows that the concepts introduced in the Sections 3-4 for GNSS are fully acceptable for the ground RNS OMEGA.

Conclusions

Based on the critical review performed in the Sections 1 and 2 of a new conceptual model offered in the Sections 3 and 4, as well as, applying the concepts of a new model for the ground RNS, it is possible to conclude the following:

1. A timescale and clock readings scale of RNP (time according to the RNP scale) in the precedence moments are considered the key concepts of radio navigation. A timescale is the moments of physical time defined by the clock readings that are the base of any scale. Time according to the scale in any moment of physical time is defined as the readings of this clock. Clock is understood as a combination of methods and

actions directed to defining a quantitative value of time according to the scale as reducing to 1 Hz of a complete phase of the periodic process that is a base of a timescale.

2. A concept of pseudodelay (pseudorange) is secondary with regard to the concept of the clock readings (time according to the scale) due to the following:

- A concept of pseudodelay $\tau_{pd}^{j}\left(t_{meas}\right)$ is determined through the concept of the clock readings as a difference between the readings $T_r\left(t_{meas}\right)$ of the clock of a navigational receiver in the moment of measurement t_{meas} and the measurements $T^j\left(t_p^j\right)$ of the clock of the j-th station (satellite) in the precedence moment t_{pr}^{j} to the moment of measurement t_{meas}, i.e. $\tau_{pd}^{j}\left(t_{meas}\right)=T_r\left(t_{meas}\right)-T^j\left(t_p^j\right)$. The clock readings are an independent concept. As it is shown in the Section 4 of the paper, the evaluations of the clock readings of the GNSS satellites in the precedence moments (the readings of a channel clock in the measurement moments) make it possible to carry out all navigation definitions without using a pseudodelay concept.
- In the synchronism mode in the registers of the phases of reference signals, a tracking loop for the phases of the received signals in the navigation receiver, disambiguous evaluations of the stations (satellites) in the moments of precedence are formed. That means that tracking loops of the navigation receiver track not the values of pseudodelays, but the stations (satellites) readings in the precedence moments. The measurements of pseudodelays in the navigation receiver are formed on the secondary base by means of the integration of the receiver's processor codes of its tracking loops.
- In the RNS, solving the disambiguity of pseudodelay measurements is undertaken not directly, as it is stated in [7, 15], but through solving the disambiguity of the clock readings of the stations (satellites) in the precedence moments. To solve this disambiguity, the signals of timestamps and their digitization are employed. At this, to solve the disambiguity no any a priory information is applied. When there are no time stamps digitization in the radiated signals, it is necessary to use a priory information in the form of the readings of the external clock in the precedence moments.
- 3. Applying a new fundamental concept of the readings $T^j\left(t_p^j\right)$ of the stations clock during the precedence moments t_{pr}^{j}, permits one to abandon carrying out measurements on the frequencies 13.6 and 34/3 kHz and, hence, ease significantly the RNS OMEGA by introducing digitization of timestamps into the structure of the RNS OMEGA navigation signal and two additional counters into the equipment of the receiver.

References

1. Solov'ev Yu.A. *Sistemy sputnikovoy navigatsii* [Satellite navigation systems]. Moscow, EKO-TRENDZ, 2000. (in Russian)

2. Lipkin I.A. *Sputnikovye navigatsionnye sistemy* [Satellite navigation systems]. Moscow, Vuzovskaya kniga, 2001. (in Russian)

3. *Informatsionnye tekhnologii v radiotekhnicheskikh sistemakh* [Information technologies in radio engineering systems]. 2nd, revised and enlarged. Moscow, MGTU im. Baumana, 2004. Ed. Fedorov I.B. (in Russian)

4. Bakulev P.A., Sosnovskiy A.A. *Radionavigatsionnye sistemy* [Radio navigation systems]. Moscow, Radiotekhnika, 2005. (in Russian)

5. *GLONASS. Printsipy postroeniya i funktsionirovaniya* [GLONASS. Principles of designing and functioning]. Third edition, revised. Moscow, Radiotekhnika, 2005. Eds. Perov A.I. and Kharisov V.N. (in Russian)

6. Yatsenkov V.S. *Osnovy sputnikovoy navigatsii. Sistemy GPS NAVSTAR i GLONASS* [Satellite navigation fundamentals. GPS NAVSTAR and GLONASS]. Moscow, Goryachaya liniya-Telekom, 2005. (in Russian)

7. Malyshev V.V., Kurshin V.V., Revnivykh S.G. *Vvedenie v sputnikovuyu navigatsiyu* [Introduction to satellite navigation]. Moscow, MAI-PRINT, 2008. (in Russian)

8. Boriskin A.D., Veytsel' A.V., Veytsel' V.A., Zhodzishskiy M.I., Milyutin D.S. *Apparatura vysokotochnogo pozitsionirovaniya po signalam global'nykh navigatsionnykh sputnikovykh sistem: priemniki-potrebiteli navigatsionnoy informatsii* [Equipment for high accuracy positioning on the signals of global satellite navigation systems: receivers-consumers of navigation information]. Moscow, MAI-PRINT, 2010. (in Russian)

9. *Guide to GPS Positioning*. Prepared under the leadership of David Wells. Canadian GPS Associates. Second edition, revised. May 1987.

10. Hoffman-Wellenhof B., Lichtenegger H., and Collins J. *Global Positioning System. Theory and Practice*. Springer-Verlag Wien, New York, 1992.

11. Seeber G. *Satellite Geodesy. Foundations, Methods, and Applications*. Walter de Gruyter, Berlin, New York, 1993.

12. Understanding GPS. Principles and Application. Ed. Elliot D. Kaplan, 1996.

13. Leick A. *GPS Satellite Surveying*. John Wiley and Sons. New York, 1990.

14. Global Positioning System: Theory and Applications. Vols. I and II. Eds. Parkinson B.W. and Spilker J.J. *Progress in astronautics and aeronautics*. Vol. 163. American Institute of Aeronautics and Astronautics, Inc. Washington, DC, 1996.

15. Radionavigatsionnye sistemy sverkhdlinnovolnovogo diapazona [Radio navigation systems of superlong-wave range]. Moscow, Radio i svyaz', 1985. Eds. Olenyuk P.V. and Golovushkin G.V. (in Russian)

16. Gonorovskiy I.S., Demin M.P. *Radiotekhnicheskie tsepi i signaly* [Radio engineering circuits and signals]. Moscow, Radio i svyaz', 1994. (in Russian)

17. Povalyaev A.A. *Sputnikovye radionavigatsionnye sistemy: vremya, pokazaniya chasov, formirovanie izmereniy i opredelenie otnositel'nykh koordinat* [Satellite radio navigation systems: time, readings of clock, formation of measurements and determination of relative coordinates]. Moscow, Radiotekhnika, 2008. (in Russian)

18. Povalyaev A.A., Veytsel' A.V., Mazepa R.B. *Global'nye sputnikovye sistemy sinkhronizatsii i upravleniya dvizheniem v okolozemnom prostranstve* [Global satellite synchronization systems and motion control in near-Earth space]. Moscow, Vuzovskaya kniga, 2012. (in Russian)

19. Fey L. Time Disseminations Capabilities of the Omega System. *Pros. 25th Ann. Symp. Frequency Control*. Electron. Indust. Ass., Washington, DC, pp. 167–170, Apr. 1971.

ROCKET-SPACE DEVICE ENGINEERING AND INFORMATION SYSTEMS
Vol. 3, Iss. 1, 2016, pp. 25–30

Creating 3D Surface Models Using the Resurs-P Spacecraft Images

A.A. Peshkun

Research Center for Earth Operative Monitoring of
Joint Stock Company "Russian Space Systems", Russia
e-mail: peshkun_aa@ntsomz.ru

Abstract. At the present time the Russian Earth remote sensing spacecraft group includes two spacecraft of the Resurs-P type that allow stereo survey of the Earth surface to be performed on one orbit circuit. The Geoton sensor images (panchromatic band with 1-meter resolution) were used to create the models. These models were created automatically by means of the digital photogrammetric station PHOTOMOD. The article contains the information about 3D surface mod-els made from stereo pairs with different convergence angles for the same territory. It is noted that these models are compared with each other from the accuracy and completeness of objects' reflection. The purpose of the comparison is to determine the optimal parameters for the Earth surface stereo survey.

Keywords: Resurs-P, stereo survey, stereo pair, digital surface model

Nowadays two spacecraft of the Resurs-P type (No. 1 was put into orbit on June 25, 2013; No. 2 was put into orbit on December 26, 2014) are being operated in the orbit. They are intended for maps updating, ensuring economic activity of various federal, regional, municipal departments, and other consumers, and for obtaining information in the field of control and environmental protection.

The target-oriented equipment:

- the optical-electronic GEOTON-L1 complex with the system for reception and transmission of information (SRTI) SANGUR-1U (Resurs-P No. 1, Resurs-P No. 2);
- the hyperspectral equipment – HSE (Resurs-P No. 1, Resurs-P No. 2);
- a complex of the wide-range multispectral surveying equipment (WMSE) of high (HR) and average resolution (AR): WMSE-HR and WMSE-AR (Resurs-P No. 1, Resurs-P No. 2).

The scientific equipment:

- a complex of research of galactic beams of ultra-high energies – Nucleon
(Resurs-P No. 2).

Moreover, the onboard AIS receiver designed for radio signals receiving from sea vessels and their automatic identification is installed on board the Resurs-P No. 2 spacecraft.

Table 1. Basic characteristics of the GEOTON-L1 equipment

Characteristics name	Value
Focal distance, mm	4000
Entrance pupil diameter, mm	500
Relative hole	1:8
Field angle, °	5°18′
Photosensitive element size, μm panchromatic spectral	6x6 18x18
Pixel projection onto the Earth surface, m: in panchromatic band in narrow spectral bands	1.0 3.0–4.0
Swath width, km	38
Spectral bands, μm:	
panchromatic	0.62–0.79
blue	0.48–0.53
green	0.54–0.59
red	0.62–0.68
red 2	0.66–0.69
near red	0.70–0.75
near infrared 1	0.72–0.80
near infrared 2	0.81–0.88
Number of simultaneously used spectral bands	1–5
Linear coding bitness of videodata, bit/pixel	10

Table 2. Basic characteristics of the WMSE complex

Characteristics name	Characteristics value	
	WMSE-AR	WMSE-HR
Optical system: focal distance, mm relative hole field angle, °	 40 1:4 54°30′	 200 1:3 11°70′
Swath band, km	441.7	97.2
Pixel projection onto the Earth surface, m: in panchromatic band in narrow spectral bands	 59 118	 12 23.8

Spectral bands, µm:	
panchromatic	0.43–0.7
blue	0.43–0.51
green	0.51–0.58
red	0.60–0.70
IR 1	0.7–0.9
IR 2	0.8–0.9
Photosensitive element size, µm:	
panchromatic	5x5
spectral	10x10
Linear coding bitness of videodata, bit/pixel	12

Table 3. Basic characteristics of the hyperspectral equipment

Characteristics name	Value
Swath, km	30
Pixel projection onto the Earth surface, m	25–30
Spectral bands, µm	0.4–1.1
Channels quantity	not less than 96
Spectral resolution, nm	5–10

Table 4. Orbit parameters

Characteristics name	Value
Type	near-circular sun-synchronous
Height, km	470–480
Inclination, °	97.28
Surveillance periodicity, day	no more than 3

Survey modes

A route survey. Survey in the route mode can be carried out both with a constant value of heel and pitching angles (see Fig. 1) and with the set azimuth (Fig. 2). A spacecraft deviation on a heel and pitching from a nadir is possible up to ±45 °, on yawing is up uto ±60 °. Duration of routes is from 2 to 300 sec.

Route survey

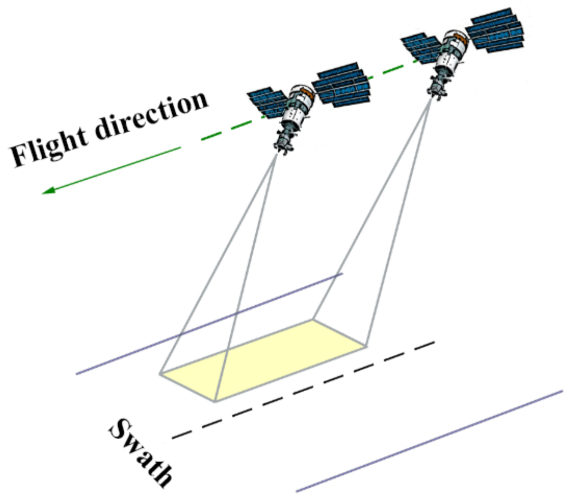


Fig. 1. Survey with a constant heel and pitching

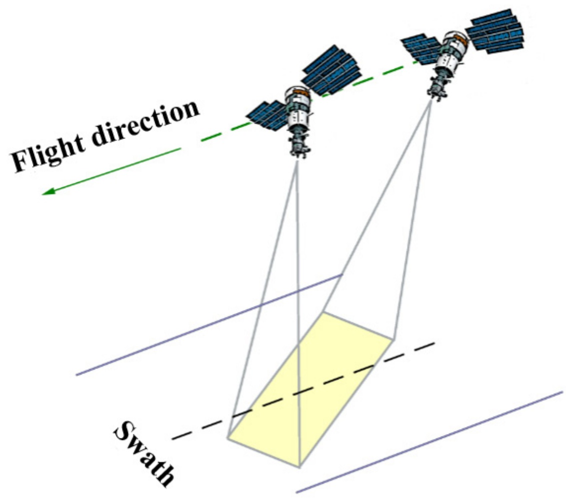


Fig. 2. Survey with the set azimuth

Stereosurvey. Stereosurvey is a receiving a stereopair[1] of images in the photographic way. Stereosurvey is carried out on one circuit with a device deviation on pitching. A length of routes is up to 115 km.

1 A stereopair is a combination of two images of the same object received from two different surveying points.

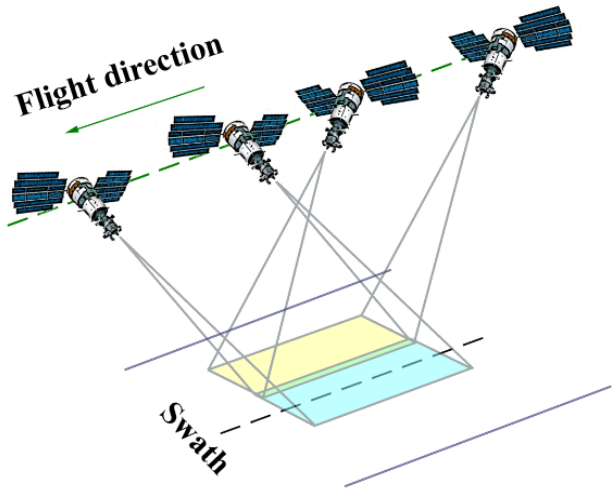


Fig. 3. Aerial survey

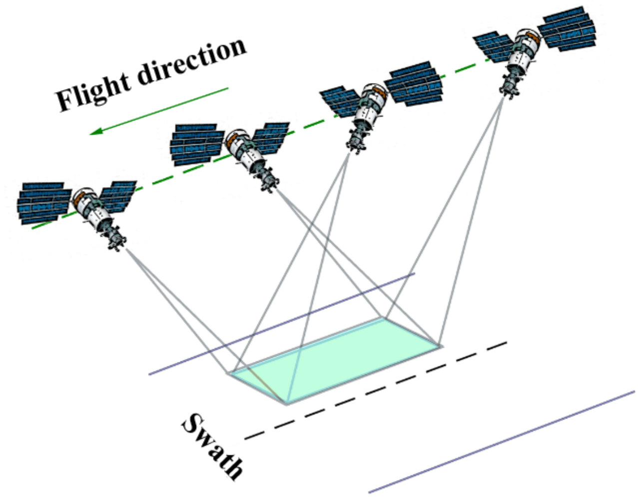


Fig. 4. Stereosurvey

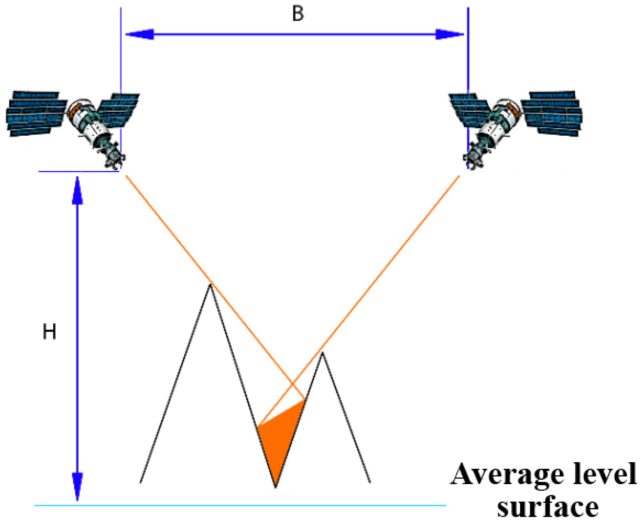


Fig. 5. A bigger *B/H* ratio

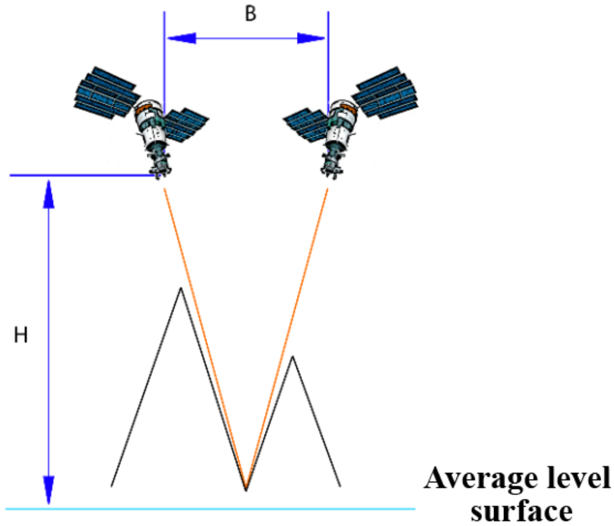


Fig. 6. A smaller *B/H* ratio

Creating a 3D model of the area

The materials of the stereosurvey became the basic data for creation of three-dimensional (3D) models of the area.

The key characteristics of the stereopair is the ratio of photography basis[2] to photography height – *B/H*.

At values of *B/H* ratio close to 1, a convergent angle[3] is about 54°. If stereosurvey is carried out with equal deviations on pitching, then deviation angles in that case are about 27°. Advantages of such stereosurvey parameters are the following: a big convergent angle allows one to increase the accuracy of measurements on a stereopair; a big area of a stereopair. Disadvantages: shadow zones on mountainous areas and areas with high-floor buildings; to create an orthophotomap it is necessary to carry out additional survey with small deviation angles from a nadir or use stereopair images with deviation from a nadir more than 20°.

In the world practice when stereosurvey is on one circuit (with deviations on pitching), the *B/H* ratio is chosen depending on a height difference in the territory under surveying. Therefore, optimum *B/H* ratios are values about 0.5. For mountainous areas, this ratio is reduced, and the ratio is increased for flat ones. Thus, convergent angles vary from 30° to 45°. To create a orthophotomap, stereosurvey can be carried out with different deviation angles of an optical axis from a nadir on pitching (for example, +20° and –10°), so to use a picture with a smaller deviation angle from a nadir for orthotransformation.

2 Photography basis is the distance between two neighbour photography points.

3 A convergent angle of the stereopair is the angle, which is formed by the crossed projecting beams of the stereopair images in the basic planes for the points of the same name.

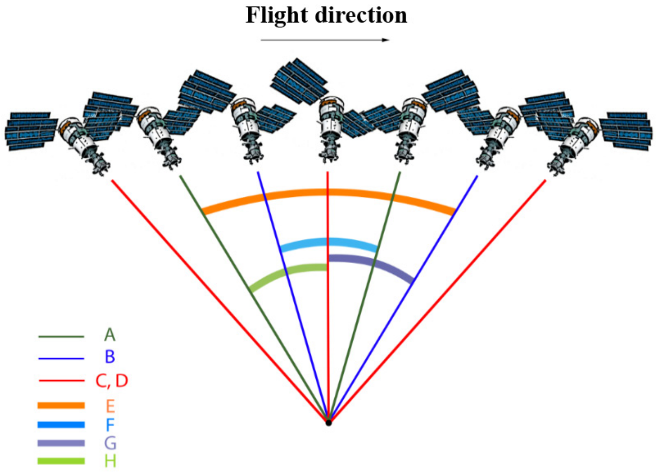


Fig. 7. Formed stereopairs

Table 5. Formed stereopairs parameters

Stereopair	Deviation angle from a nadir, °		Convergent angle, °
	Image No. 1	Image No. 2	
A	32.5	12.0	44.2
B	12.9	36.5	46.3
Triplet			
C	47.2	2.5	46.6
D	2.5	44.1	44.4
Additional stereopairs			
E	32.5	36.5	68.4
F	12.9	12.0	22.7
G	2.5	36.5	37.3
H	32.5	2.5	32.5

For creation of 3D models of the district, stereosurvey with the Resurs-P No. 1 spacecraft with the GEOTON-L1 equipment in the panchromatic range on a test area with various corners of deviations on pitching has been performed.

By results of this survey, four main stereopairs have been created (based on the criterion of survey on one circuit – *A, B, C,* and *D*) and four additional (are picked up for a convergent angle). Eight stereopairs for one territory, but with different convergent angles and ratios of *B/H* have been made in the result. The materials of 1A processing level have been used for photogrammetric processing.

As Fig. 7 and Table 5 show, there are similar stereopairs in the parameters: *A* and *B, C* and *D, G* and *H*. The pictures forming them have deviation corners, similar in size, from a nadir on pitching.

Photogrammetric processing is executed using PHOTOMOD. While processing, a block from seven pictures has been formed. Seven basic and 76 control points are used for performance of block equalizing and specification of external orientation of pictures (see Fig. 8).

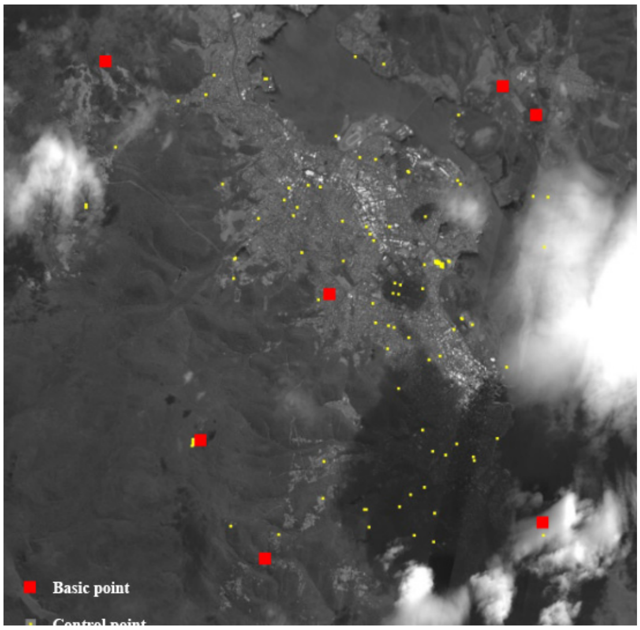


Fig. 8. A disposition scheme of basic and control points

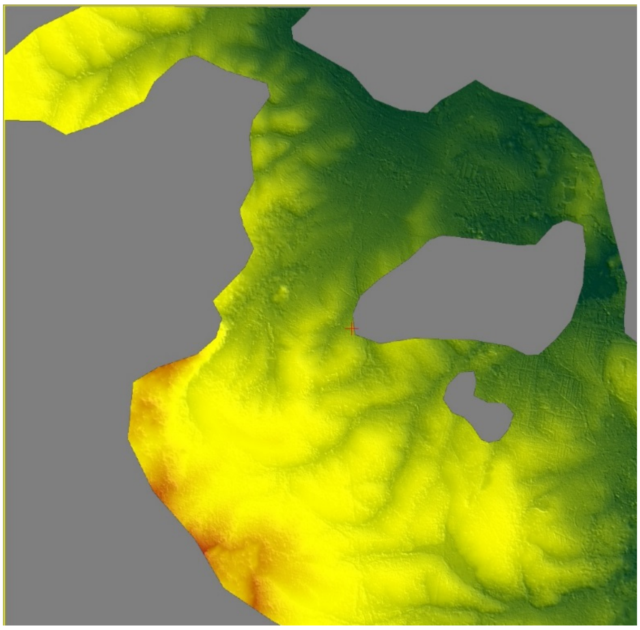


Fig. 9. A fragment of a digital area model No. 5

After adjustment, building of five digital area models is executed (further – DAM). Four DAM are constructed on stereopairs with the smallest errors for height when adjusting (are allocated green in Table 6). One additional DAM (see Fig. 9) is made using a new algorithm of the photogrammetric station PHOTOMOD for creation of dense DAM (employing repeated overlappings) with use of all seven pictures.

Table 6. Block adjustment results from eight stereopairs

Stereopair, No.	Convergent angle, °	Amount of basic points	Amount of control points	Maximum error on basic points in the plan, m	Maximum error on basic points for height, m	Standard deviation on basic points in the plan, m	Standard deviation on basic points for height, m	Maximum error on control points in the plan, m	Maximum error on control points for height, m	Standard deviation on control points in the plan, m	Standard deviation on control points for height, m
A	44.2	4	53	0.71	1.33	0.48	0.83	4.39	6.33	1.88	2.64
B	46.3	5	65	3.10	9.06	2.38	7.20	3.64	10.73	2.24	6.69
C	46.6	4	43	1.15	2.65	0.70	1.50	3.35	5.64	1.92	3.05
D	44.4	4	43	1.13	2.84	0.70	1.68	3.32	6.17	2.00	3.61
F	22.7	3	55	0.84	0.87	0.55	0.64	3.60	10.38	1.58	4.68
E	68.4	5	58	4.02	5.74	3.20	4.46	4.93	7.94	3.23	4.49
G	37.3	5	58	1.34	11.17	1.03	7.69	3.16	13.94	1.79	9.37
H	32.6	4	52	1.16	2.09	0.72	1.20	3.40	5.96	1.78	3.26

Table 7. Estimation of the accuracy of the created digital area models

Stereopair, No.	Amount of control points	Maximum error, m	Standard deviation, m
A	40	5.26	2.00
C	23	6.24	2.76
D	29	8.95	3.54
H	23	5.66	3.65
No. 5	49	12.17	6.13

Conclusion

As Tables 6 and 7 show, the stereopairs with convergent angles about 45° (the *B/H* ratio is about 0.8) and about 33° (the *B/H* ratio is about 0.5) were used for creation of DAM. At the same time, their precision characteristics are comparable. In this situation, the most preferable is a stereopair *H*, as it consists of pictures, one of which is almost not rejected from a nadir (2.5°). It will allow an orthophotomap with a minimum of distortions to be constructed.

In addition, it is worth paying attention to the stereopairs *B* and *G*. These stereopairs have surveys, similar in parameters, respectively, the stereopairs *A* and *H*. However, the orientation accuracy of the stereopairs *B* and *G* is almost twice worse, than that of the stereopairs *A* and *H*. Moreover, the stereopairs *F* and *E* have the maximum errors for height on control points more than seven meters. The stereopairs *B*, *F*, *E,* and *G* are united by presence in them of at least one picture from a stereopair *B*. The stereopairs *A*, *C*, *D,* and *H* do not have pictures from a stereopair *B*. The geometry of pictures of a stereopair *B* can serve as the reason of such result.

The following factors can serve as the main reasons for bad geometry of pictures of a stereopair *B*: spacecraft evolutions while surveying, deviation angles from a nadir, height difference on districts, and overcast. Deviation angles of pictures from a nadir at a stereopair *B* are comparable with deviation angles of pictures from a nadir of a stereopair *A*. However, the results of orientation of a stereopair *A* are significantly better. It means that spacecraft deviation angles from a nadir while surveying on geometry of a picture did not influence greatly, as well as a height difference on districts (the territory for surveying was the same one). The third picture of a triplet (the second picture of a stereopair *D*) was surveyed in conditions similar to conditions of surveying of the second picture of a stereopair *B*, even with a greater deviation angle on pitching. Though, the orientation accuracy of a stereopair *D*, in some parameters, is more than twice better than the orientation accuracy of a stereopair *B*. It confirms the assumption of a low influence on geometry of deviation angles of the spacecraft while surveying and a height difference of the district under surveying.

Stereosurvey was carried out in different days under different meteoconditions. Overcast is present in all seven pictures. In this case, various nature of overcast in days of survey could also become an important factor, which has influenced geometry of pictures of a stereopair *B*. Confirmation or a denial of this assumption requires carrying out a number of additional research and tests, as in this situation there can be additional factors influencing geometry of pictures.

ROCKET-SPACE DEVICE ENGINEERING AND INFORMATION SYSTEMS
Vol. 3, Iss. 1, 2016, pp. 31–35

Analysis of Matrix Photodetectors Application for Scanning Systems

Yu.M. Gektin[1], A.A. Zaytsev[2], A.V. Rozhnev[3], A.M. Solov'ev, M.B. Smelyanskiy

[1]candidate of engineering science, Joint Stock Company "Russian Space Systems", Russia
[2,3]Moscow Institute of Physics and Technology (State University), Russia
e-mail: [1]petrov_sv@spacecorp.ru, [3]sasharozhnyov@gmail.com

Abstract. The paper deals with the design concept of the equipment and its functioning used for remote sensing of water areas. Various constructions of the device have been analyzed according to modern requirements for this class of equipment. The article shows that the usage of matrix photodetectors allows the required signal-to-noise ratio to be obtained. A scheme of the apparatus involving optical-mechanical scanning combined with the matrix photodetector is offered. It is demonstrated that the application of matrix photodetectors with relevant characteristics for a scanning system enables the required signal-to-noise ratio to be achieved for the model of radiation from water surface in defined spectral bands by means of redundant quantity of light-sensitive elements. The interpolation algorithm of signal processing from the photodetector, which permits one to reduce data flow with minimized geometrical distortion, is described. The results of modeling showed that the offered conception of the apparatus would be effective for remote sensing of the World Ocean. A device involving the offered scheme will be sufficient for modern metrological requirements.

Keywords: Earth remote sensing, ocean color, scanning system, matrix photodetector, interpolation algorithm

Joint Stock Company “Russian Space Systems” is faced by a task to create a target equipment – the scanner of color of an open sea surface, which has to be a part of the space complex intended for the solution of problems of hydrometeorological and oceanographic support [1, 2]. To fulfill such tasks including monitoring of the ocean, the scanners, such as SeaWiFS, MODIS, and VIIRS [3] have been already brought to life and checked. The spectral range of the equipment operation of the Earth remote sensing (ERS), which is responsible for the ocean chromaticity, is from 0.4 to 0.9 µm (see Table 1). The orbit parameters (sun-synchronous, altitude of 700-850 km) and geometrical parameters of the equipment (viewing angle is > 100°, resolution at nadir is ~ 1 km) are similar.

Table 1. Comparative characteristics of color scanners in the range of 0.4-0.9 µm.

SeaWiFS		MODIS		VIIRS	
Spectral band, nm	Bandwidth, nm	Spectral band, nm	Bandwidth, nm	Spectral band, nm	Bandwidth, nm
402-422	20	405 - 420	15	402-422	20
433-453	20	438 - 448	10	436-454	18
480-500	20	483 - 493	10	478-498	20
500-520	20	526 - 536	10	-	-
545-565	20	546 - 556	10	545-565	20
-	-	662 - 672	10	-	-
660-680	20	673 - 683	10	662-682	20
745-785	40	743 - 753	10	744-759	15
845-885	40	862 - 877	10	845-884	39

Since the device under development has to solve problems of monitoring of water areas of the World Ocean, it is necessary to consider a so-called “water” radiation model. The equipment under development has to provide the signal-to-noise ratio of > 500 when surveying water areas. The maximum spectral radiance ascending from a water surface is located on the wavelength of 410 nm; the spectral density of radiance on this wavelength is 95 W/m^2×sr×µm (according to the model used during creation of the specialized scanner for research of the World Ocean SeaWiFS [4]). The minimum brightness of the radiation ascending from the water surface is equal to ~7 W/m^2×sr×µm (spectral range with the central wavelength of 860 nm). The calculations executed have shown that none of the known traditional approaches for creating the ERS equipment (namely: using one- and multielement detector in combination with optical-mechanical scanning, multielement (linear) receivers including in the TDI mode, and also the matrix receivers, which are carrying out frame survey of the underlying surface) allows one to meet the specified requirement at the acceptable weight dimension characteristics. The parameters of a “water” model specified above define a task as surveying the objects of small brightness. At the same time, the spectral density of radiance from the top layers of a cloudy surface on the border of the dense atmosphere can reach value of 660 W/m^2×sr×µm (see Fig. 2). Therefore, the problem of creating a scanner operating in all dynamic range of the brightness scenes and providing necessary viewing angle and resolution, and also meeting the imposed metrological requirements regarding the signal-to-noise ratio (> 500 when surveying water areas), is very difficult.

The approach to the solution of the set task is in use of a high-speed low-format matrix CMOS-photodetector in combination with optical-mechanical scanning and accumulation of a digital signal. Use of a matrix photodetector of a bigger dimension is inexpedient for two reasons: first, in the chosen scanning scheme it is necessary to use the photodetector with a high frame rate (time for reading the whole matrix is ~ 1.92 ms), secondly, with increase in dimension of a photodetector geometrical distortions significantly grow when projecting the matrix on the Earth’s surface.

The schematic diagram of the multichannel device based on the principle of the optical-mechanical scanning using the matrix detector operating in the mode of a digital TDA (the time delay and integration mode of a signal), and scanning geometry are given in Fig. 1. The device has to make survey of a water surface with the resolution of 500 m at nadir at an altitude of 820 km in a swath 2800 km (a viewing angle is 110°).

As it is shown in the Figure above, the schematic diagram of the device under development includes a flat scanning mirror, lens, and matrix photodetector. The scanning mirror is one-sided and makes rotation in one direction during scanning with the subsequent reverse.

In the focal plane of the objective there is the matrix photodetector with dimensions 128×128, working at frequency of 520 frames/sec.

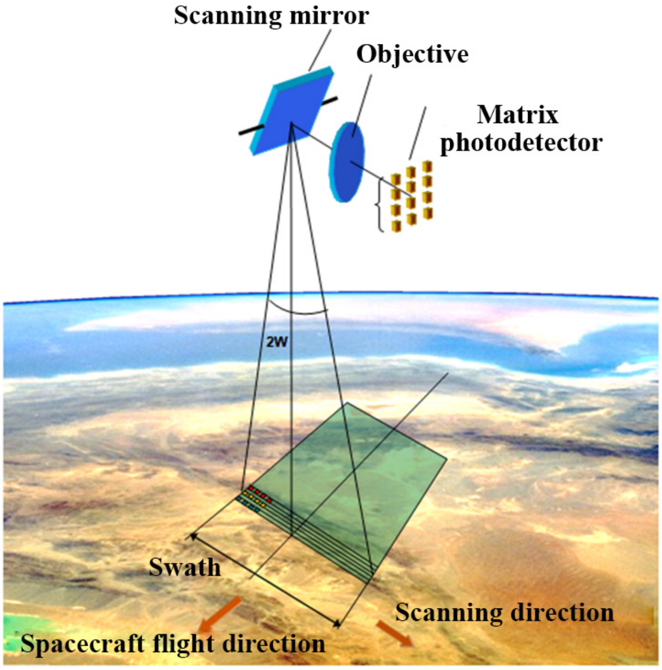


Fig. 1. The schematic diagram of scanning of the underlying surface

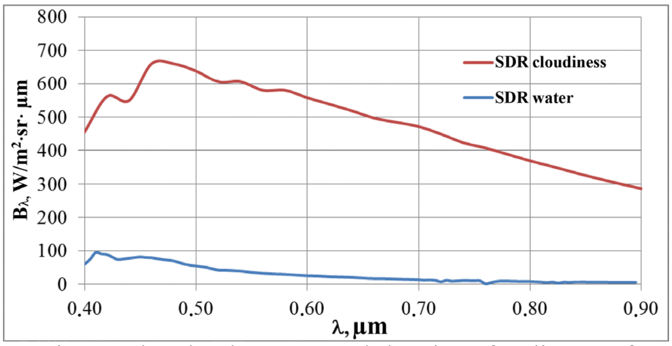


Fig. 2. The absolute spectral density of radiance of the water surface and overcast used by the SeaWiFS developers

The calculation results of the signal-to-noise ratio for the water surface are given in Table 2. It is clear that the matrix of photocells with dimensions 128×128 is capable to provide the necessary signal-to-noise ratio in the set spectral channels on condition of digital joining of readout. If this operation is carried out at ground processing, then information stream from a board is about 820 Mbps. Therefore, there is need to reduce the information stream. To do this it is offered to put the signals received by various matrix photocells, an interpolation method in the signal processing unit (SPU), which is a part of the equipment [5].

The idea of signals integration is in dynamic display of the geometrical ratios taking place when scanning objects in a space of under surveying to the space of random access memory of the onboard signal processing device. Readout at the output of the joining algorithm is equivalent to the readout from a virtual line, geometrically corresponding to the first line of a matrix. Each set of readout from a virtual line is a superposition of the readout received in different time points from all lines of a matrix.

The problem of readout integration is that due to orbital movement of a spacecraft in the focal plane of the scanner objective the movement of the Earth's surface image both in the direction of scanning and perpendicular to it is also taking place. The first, in principle, owning to a right choice of angular speed of optical-mechanical development can be coordinated with the frequency of matrix polling in such a way that during the matrix polling the image of the allocated object moves exactly to one pixel. As for the second, the image shift monotonously increases from the first matrix up to the last, at the same time movement speed depends on the angle of view in a difficult manner [6].

The principle of readout integration is that readout of the 1st, 2nd, …, and 128th lines of a photodetectors matrix are recorded in each column of a memory matrix the dimension 128×128 of the cells at the rate of frames frequency. Thus, in some timepoint the 1st line is recorded in the 1st column, 2nd is recorded in 128th, 3d is recorded in 127th, …, and 128th in 2nd. In the following interrogation cycle 1st is recorded in 2nd, 2nd is recorded in 1st, 3d is recorded in 128th, …, and 128th is recorded in 3d; and etc. At the same time readout from all lines except the 1st are multiplied by 2 weight coefficients (interpolation coefficient) and recorded in consecutive cells of the corresponding column of a matrix of memory (are summarized with contents of these cells). A "ready" counting is consistently read out from each column of a matrix of memory after being recorded with the counting from the last, 128th line of a matrix of photodetectors. Thus, the resultant counting represents the sum with scales 255 realization of single accumulation of an analog signal. In the course of the described algorithm (in real time) the following have to be calculated: a current shift as a function of a line number and the current angular position of an axis of sight, increment of a cell address as a whole part of shift and weight coefficients equal to $(1-d)$ and d, where d is a fractional part of shift.

Table 2. The absolute spectral density of radiance B_l and the signal-to-noise ratio of a water surface model in the range of 0.4–0.9 µm

l, µm	B_l, W/m^2×sr× µm	Signal-to-noise ratio at single readout	Signal-to-noise ratio at integration of signals received for 128 readout cycles
0.410	95	86	945
0.440	78	81	887
0.490	60	75	821
0.550	36	61	673
0.640	21	50	554
0.670	17	46	510
0.745	11	56	618
0.860	7	51	561

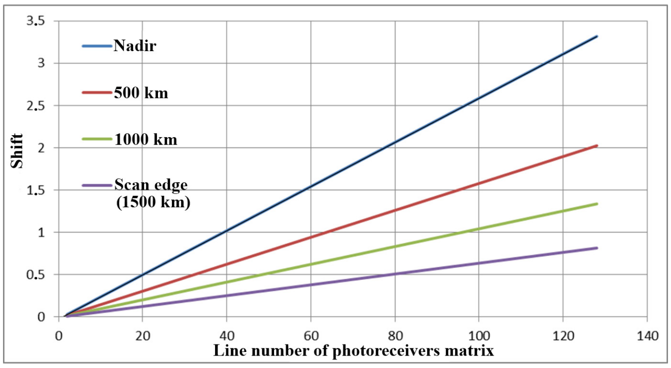


Fig. 3. Sizes of shifts for 4 provisions of an axis of sight regarding a nadir

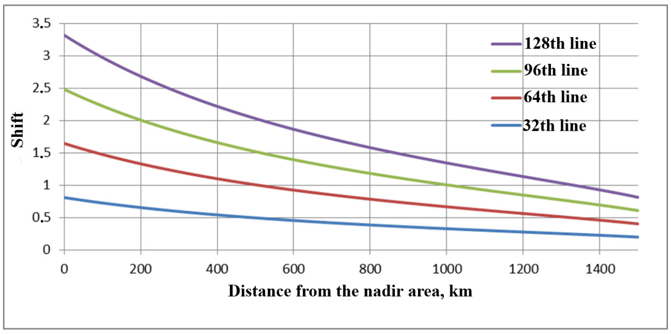


Fig. 4. Dependence of shift on distance relative to a nadir for lines with numbers 32, 64, 96 and 128

The calculation results of shifts of elements coordinates for several positions of an axis of sight concerning a nadir are in the following Figure (Fig. 3): .

The dependence shown in Fig. 4. with a good accuracy (an error no more than 0.6%) is approximated by a two-dimensional function – a linear one depending on a line number of a matrix photodetector (see Fig. 3) and a cubic one depending on the current position of the

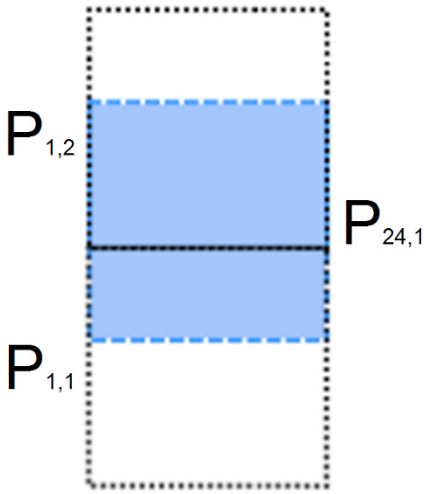


Fig. 5. Elementary counting from the 1st pixel of the 24th line is joined with counting from the 1st and 2nd pixels of the 1st line with the corresponding weight coefficients

axis of sight (see Fig. 4). According to this function, in SPU there should be calculated the address shifts and interpolation coefficients. In Fig. 5 superposition of the 1st and 24th lines of a matrix photodetector is shown.

The reduction of the information stream by 128 times is result of work of the algorithm, as the signal at the SPU output is a convolution on 128 interpolated lines of a matrix. As a result of interpolation, the full optical transfer function (OTF) of the system worsens a little, since OTF of the virtual line received as a result of convolution operation coincides with OTF of a matrix line of photodetectors only at special speed ratios and the direction of the image movement in the focal plane and values of frame frequency of interrogation of a matrix.

In addition, using the linear approach in interpolation algorithm also brings some distortions. However, as the OPF parameters of a "lens – matrix photocell" system are not too high, the marked distortions are admissible.

The described building scheme of the scanning equipment permits one to meet the modern metrological requirements imposed to the ERS equipment intended for determination of water areas chromaticity of the World Ocean. The developed interpolation algorithm makes it possible to reduce considerably the information stream from a spacecraft board without loss of radiometric resolution and with the minimum geometrical distortions.

References

1. Akimov N.P., Zaytsev A.A., Solov'ev A.M. Prakticheskaya realizatsiya sovremennykh metrologicheskikh trebovaniy, pred»yavlyaemykh k perspektivnomu skaneru tsvetnosti KA "Meteor-M" No. 3 dlya issledovaniya akvatoriy [The implementation of modern metrological requirements for the perspective colour scanner of "Meteor-M" No. 3 spacecraft for water areas study]. *Sbornik trudov VI Vserossiyskoy nauchno-tekhnicheskoy konferentsii "Aktual'nye problemy raketno-kosmicheskogo priborostroeniya i informatsionnykh tekhnologiy"* [The proceedings of the VI Russian scientific and technical conference "Current problems of rocket and space device engineering and information technology"], 2013. (in Russian)

2. Romanov A.A. *Osnovy obrabotki i analiza dannykh kosmicheskogo distantsionnogo zondirovaniya okeana* [Data processing and analysis fundamentals of space ocean remote sensing]. Manual. Moscow, MFTI, 2003, 272 p. (in Russian)

3. Remote Sensing of ocean colour in coastal, and other optically-complex, waters. *Reports of the international ocean-colour coordinating group*. IOCCG, Dartmouth, Canada, 2000, No. 3, 140 p. Available at: http://www.ioccg.org/reports/report3.pdf.

4. NASA. Official site of the Sea-viewing Wide Field-of-view Sensor project. Available at: http://oceancolor.gsfc.nasa.gov/SeaWiFS/.

5. Zaytsev A.A. Optimizatsiya kharakteristik fotopriemnoy optiko-elektronnoy sistemy sputnikovoy apparatury, prednaznachennoy dlya resheniya zadach opredeleniya tsvetnosti akvatoriy [Characteristics optimization of the photodetective optoelectronic system of the space equipment dedicated to determination of the water area color]. *Vypusknaya kvalifikatsionnaya rabota bakalavra prikladnykh matematiki i fiziki* [A graduation thesis of the bachelor of applied mathematics and physics]. Moscow, 2012, 45 p. (in Russian)

6. Shovengert R.A. *Distantsionnoe zondirovanie. Modeli i metody obrabotki izobrazheniy* [Remote sensing. Models and methods of image processing]. Moscow, Tekhnosfera, 2010, 560 p. (in Russian)

ROCKET-SPACE DEVICE ENGINEERING AND INFORMATION SYSTEMS
Vol. 3, Iss. 1, 2016, pp. 36–40

Radio Engineering Equipment Control Using a Database "Linter-VS" in OS MSVS

V.M. Vatutin[1], S.A. Dontsov[2], U.V. Efremov

[1]doctor of engineering science, professor, [2]candidate of engineering science
Joint Stock Company "Russian Space Systems", Russia
e-mail: vladmvat@mail.ru, sad_rks@mail.ru, jorgiy@bk.ru

Abstract. According to the growth of modern computers performance, increase of random access memory and memory of disk drives, it is possible to use database management systems (DBMS) for solving the control tasks in complex systems. The article describes the control of a radio engineering complex including, for example, an antenna system, an irradiator control system, a receiver, etc. A control with the use of database tables under "Linter–VS" DBMS installed in OS MSVS is organized. Each device included into the radio engineering complex and also peripheral devices are certain "entities". This entity can be represented in the database in the form of several tables. A table "commands" and table "states" must be present. Auxiliary tables related to a particular entity can be present. It is expected that the controller, located on the side of the controllable device, controls the entity, for example - the antenna system. It is shown that the controller communicates with a database by means of a proxy. The proxy connects a device interface with the controller and makes queries to database tables. The queries to the "commands" table are decoded into controller's commands and transmitted through the device interface. Responses are made through it, which after decoding are transferred into the "states" table. Due to such commands and data transfer, the number of controllable and control entities is practically unlimited. They can be both local and remote. It is stressed that DBMS controls the registration of registered users in the database. Unregistered users are not served by DBMS.

Keywords: automation, radio engineering, complexes, database management system (DBMS), automated control system, GNU/Linux operating system, standard RS-485, linear algorithms

1. Summary data about OS MSVS, Linter–VS and RS–485 interface

OS MSVS, based on GNU/Linux, is a multiuser multitask network operation system (OS). It functions on the hardware platforms Intel, SPARC (Elbrus-90micro), IBM S390 and MIPS (Baget series complexes by Korund-M company), supports the multiprocessor configurations (SMP). It contains mandatory access control means, lists of access control and a role model.

Database management system (DBMS) Linter–VS 7.0 is for creation of information and operating systems for work with various information, and for creation and support of the databases based on the relational model of data. This DBMS is based on the "client-server" architecture. The technology gives real advantages for users and becomes the prevailing way for data processing.

RS-485 (Recommended Standard 485 or EIA/TIA-485-A) is a recommended standard for data transmission on a two-wire half-duplex multipoint consecutive symmetric communication channel. The standard describes only physical levels of signaling (i.e. only the first level of an interrelation model of the open OSI systems). The standard does not describe a program exchange model and exchange protocols. RS-485 was created for expansion of physical capacities of the RS232 interface on binary data transfer.

2. A general interaction scheme

A general scheme of the organization control is given in Fig. 1.

The main knot in the equipment control is the server of the remote database (DB). In tables of the DB, there are high-level control commands of the equipment, such as setting the equipment to specific operating modes in certain time. The commands are the records in the DB which contains columns with a number of an operating mode, time of its performance (time of starting the mode and time of finishing the mode) and variable data (exact values of frequency parameters, data on an azimuth, etc.). Also on this server there are tables of the data received from each equipment (including telemetry, digitized signals from receivers, etc.). An approximate table of high-level commands is provided in Table 1, and reply data are presented in Table 2.

Table 1. An approximate table of high-level commands

Mode number [OC]	Starting mode	Finishing mode	Variable data
1	00:00	01:00	Null
2	08:45	09:45	188
…	…	…	…

Giving the instructions for an operating mode in certain time is regulated by the operating program installed on the server, which, following the algorithm, gives the control server of the equipment definite lines with commands.

However, the server of the remote database is connected with the server of the equipment management not directly. Between them there is a real time communication server and autonomous control of the equipment, which can be a technological personal computer (PC). The mission of this server is in synchronization on operating time of the equipment and giving the commands to it. The autonomous control of the equipment in case of the refusal of the remote control equipment is possible via this server. Thus, two managing entities are generated: a remote control server and independent control server. The organization of an access to the control server of the equipment is arranged so that for it (server) there is no difference in the commands which are written down from the independent or remote control server – selection of the operating knot is made according to the commands of the diagnostic program, which interrogates operability of a communication channel with the remote database.

On the very control server of the equipment there are sets of entities – the programs that are terminal automatic control machines which give commands to the equipment controllers, thereby carrying out a final part of a control algorithm.

On the control server of the equipment there is also a DB, where values of the table of high-level commands are duplicated. These values are analyzed by the program – a finite-state automatic control machine for definition of performance of the mode and time when it should be carried out.

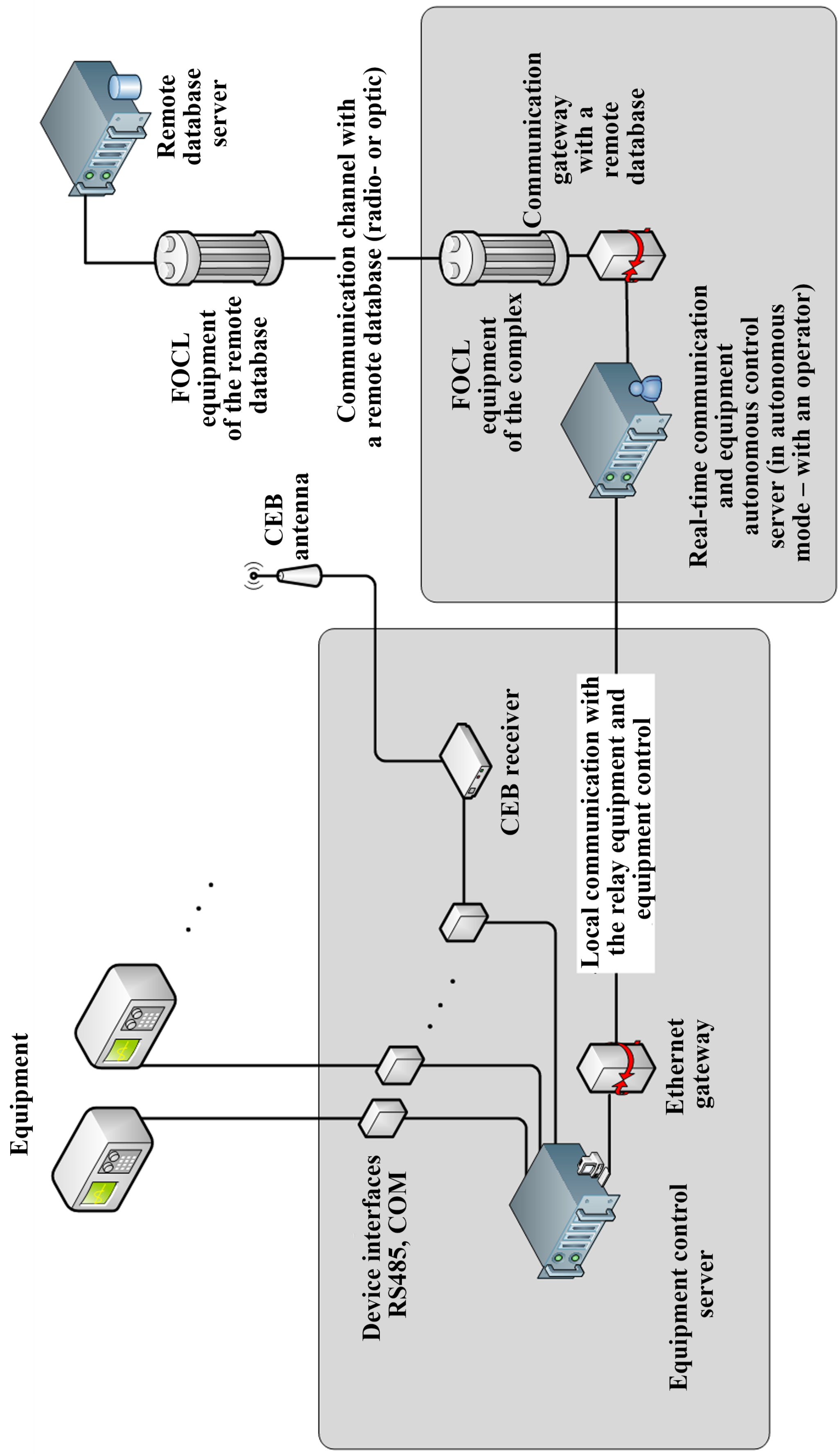


Fig. 1. General scheme for control of RTR

Table 2. An approximate table of response commands

Number [OC]	Good condition	Good condition 2	Result 1	Result 2	Result 3	Performed
1	Yes	Yes	443	98	10	No
…	…	…	…	…	…	

Commands to the controllers are in tables – “dictionaries” of the database, which is on the same PC (the control server of the equipment), which exercises control. The table containing commands to controllers (low-level commands) has also columns with numbers, the sequence of which (from 1 to *n*) presents a linear algorithm of performance of these low-level commands to the controller. The finite-state automatic control machine correlates the columns to numbers of the modes with the corresponding lines from the table of high-level commands. The program, changing according to the set algorithm of commands from the tables of low-level commands, carries out set-up of variable values. The example of low-level commands is given in Table 3.

Table 3. Tables of low-level commands

Number [OC]	Command	Command response	Mode 1	Mode 2	Response mode
1	009EA0	009EA1	null	2	1
2	100000	100001	1	1	null

As a result of fulfilment of every high-level command, each entity will give the response information in the form of the DB table (telemetry, digitized signals from receivers, sign of performance/not performance and so on). These results can be immediately transferred to the DB of the remote server (Table 3).

The operating machine (the control server of the equipment) via RS-485 interfaces in the half-duplex mode will exercise terminal control. The UART controller of the device exercises control of the reception-transmission mode automatically, which is a physical port RS-485.

3. A relations structure of the entities to the database of the control server of the equipment

An access to the data being contained and appearing in the database of the control server is carried out through the authorized connection of an entity according to the unique login and the password (a so-called global role). Authorization is a set of rather strict rights, therefore there are group roles for their variation (for example, one entity can only read data of the certain table, other entity can read, write and delete data). Each such group role has one corresponding global role. Differentiation of the rights is shown in Fig. 2.

4. Advantages and disadvantages of the method of equipment control

Advantages:

1. Using OS MSVS solves a problem of unauthorized data control due to the organization of the mandatory access both on a certain machine and on a complex in general. This is the key moment for radio engineering complexes as at obviously incorrect control of the equipment the probability of its refusal is possible.

2. Using DBMS Linter–VS 7.0 supplements a protection system from an unauthorized access as it has a pam (a method of authorization). This method authorizes the user of DBMS only if this user is registered in the system on the same machine.

3. One of the methods of the organization of data exchange via input-output devices, which is implemented in systems of GNU/Linux architecture (which MSVS OS is), is a so-called teletype (TTY). The teletype is the terminal in the form of the file for standard input. It significantly simplifies transfer and reception of data from any devices at a low level (just at the control level of controllers).

4. The RS-485 interface exchange allows one to exchange data at distance to 1200 meters that is advantage if the equipment is removed from the place of control (high ramps, removal of the equipment for the topological reasons, large-size antennas, etc.). Besides, this protocol supports small exchange speeds (for example, 9600 bauds) that promotes information transfer with minimum losses. The speed of 9600 bauds is sufficient for control of the equipment because low-level commands, as a rule, have the size of no more than 100 bytes (as well as the reciprocal parts containing semantic loading).

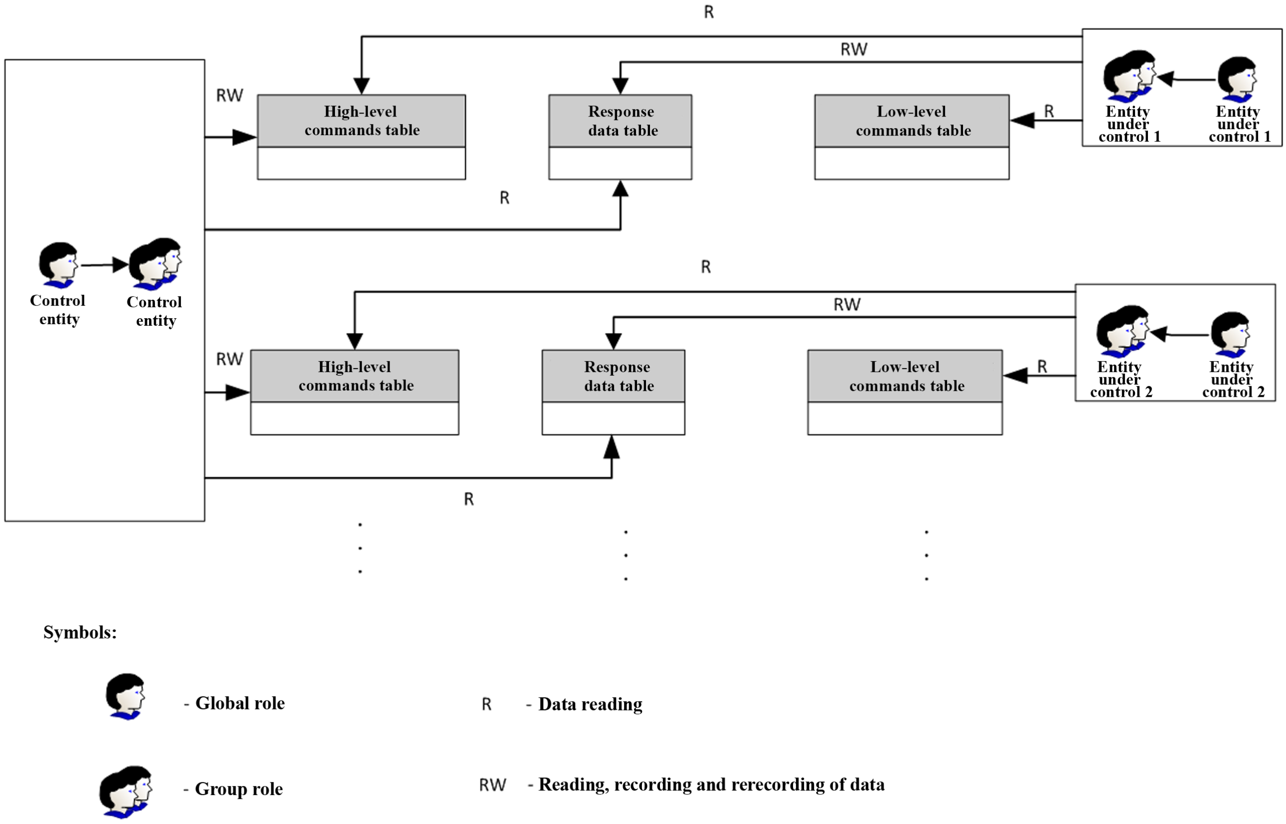


Fig. 2. Rights differentiation of the control server database of the equipment

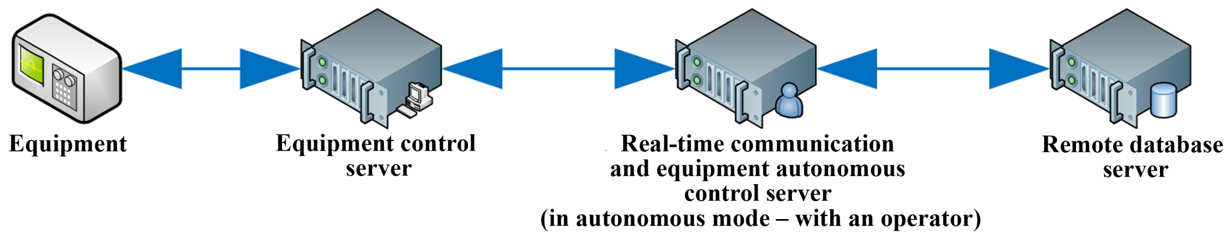


Fig. 3. The path of each command

Disadvantages:

- the path of each command is in four stages: Equipment – Server of equipment control – Server of autonomous equipment control – Server of the remote database (Fig. 3).
- it is quite probable that such quantity of stages reduces system reliability from failures. However, the problem is not very essential, as methods and control algorithms of control load very little modern computer means on which the system of automated control is built.

5. Conclusion

This architecture of automated control systems for the radio engineering equipment (and difficult radio engineering complexes) is universal, as has no strict binding to certain types of the equipment and a variety of the tasks facing a complex. In fact, the system is a system of finite-state automatic control machines and can be easily finished to the level of valve matrixes (instead of the whole PC) that minimizes the system dimensions and will also increase their reliability.

References

1. Worsley J.C., Drake J.A., *PostgreSQL. Dlya professionalov* [PostgreSQL. For professionals]. Saint Petersburg: Piter, 2003, 496 p.

2. Bartunov O., Sigaev F., Napisanie rasshireniy dlya PostgreSQL s ispol'zovaniem GiST. [Writing of extensions for PostgreSQL using GiST]. Available at: http://www.sai.msu.su/~megera/. (in Russian)

ROCKET-SPACE DEVICE ENGINEERING AND INFORMATION SYSTEMS
Vol. 3, Iss. 1, 2016, pp. 41–56

Digital Phasing to Increase the Efficiency of the B-529 Antenna System

S.I. Vatutin[1], O.V. Zaytsev

[1]candidate of engineering science
Joint Stock Company "Russian Space Systems", Russia
e-mail:otd0943_vsi@mail.ru, pochta-ov@mail.ru

Abstract: This article describes the proposals for practical achievements of the ideal signal-to-noise ratio in the receive path of the B-529 antenna system by means of digital phasing of four antennas of this system. A digital equisignal forming method of difference guidance signals from the sum signal from four antennas of the B-529 system is proposed. It is stressed that this method allows to increase twice a signal-to-noise ratio in the antenna guidance path in comparison with a traditional guidance method from sum and difference signals. It is reported that the results can be used when modernizing the B-529 antenna system through transferring its receive path and guidance path to digital processing of radio signals.

Keywords: antenna system, hased array, receiver, digital phasing, intermediate frequency, path-length difference, equisignal method, sum and difference signal.

This article gives a comparative analysis of analog and digital methods of phasing of signals from four antennas of the B-529 system, mounted on a common frame and forming a phased array antenna (see Fig. 1).

Fig. 1. B-529 array antenna

The schematic diagram of a phased B-529 array antenna is presented in Fig. 2.

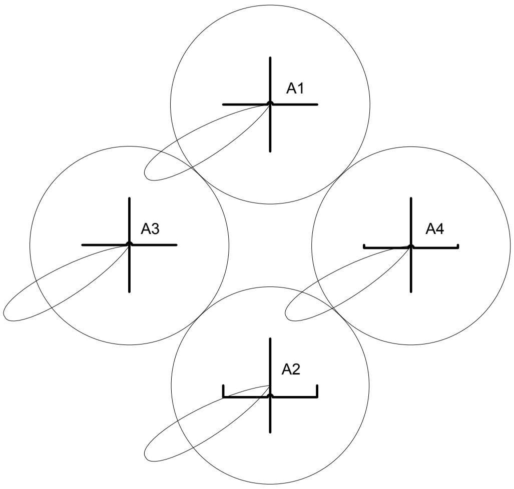


Fig. 2. Schematic diagram of B-529 array antenna

Comparison of the analog and digital methods of processing and phasing of signals from the four antennas of the B-529 system will be conducted under the same conditions, with low-noise amplifiers installed directly on the feed elements of the antennas. Losses in the cables, connected to the hybrid rings from 1 dB in the VHF band up to 2-2.4 dB in the higher UHF band can be neglected.

As seen in Figure 2, the top antenna will be referred to as A1, the bottom one as A2, the left one as A3 and the right one as A4; Ψ is the azimuth, Φ_{SC} is the elevation angle to the direction of the spacecraft, $\Delta\Psi$ is the azimuth deviation, $\Delta\Phi$ is the deviation of the elevation angle from the direction of the SC. During the analysis it is assumed that the array phase center of B-529 is located at the intersection of the lines connecting the phase centers of A1 – A2 and A3 – A4 antennas.

Figure 3 shows the geometric model for calculation of phase incursions in A1 – A4 antennas in relation to the array phase center, occurring when the radiation pattern axis of the antenna declines from the direction of the SC by $\Delta\Psi$ in azimuth and by $\Delta\Phi$ in elevation.

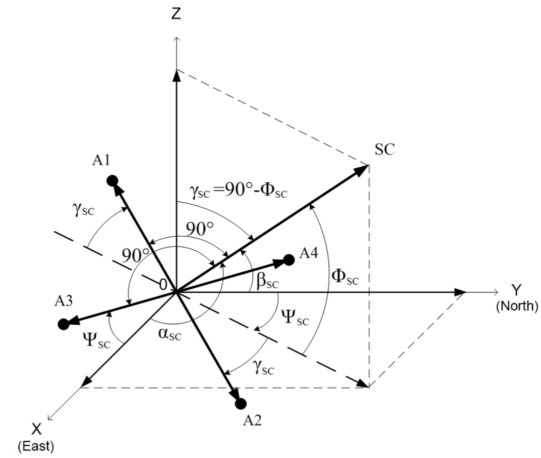


Fig. 3. Geometric model of the B-529 array antenna

The path length difference from the SC to the ΔR_{i0} antenna in relation to the array phase center equals the length of projection of the phase center vector of the antenna on the SC directivity vector and is calculated by the formula for angles between the vectors through directional cosines:

$$\Delta R_{i0} = L_{i0}\cdot[\cos(a_{Ai})\cdot\cos(\alpha_{SC}) + \cos(b_{Ai})\cdot\cos(\beta_{SC}) + \cos(g_{Ai})\cdot\cos(\gamma_{SC})] \quad (1)$$

Here L_{i0} is the distance from the array phase center to the phase center of the antenna Ai. With antenna diameter D = 6 m, $L_{i0} = D/\sqrt{2} = 4.24$ m.

The directional cosine angles in the direction of the SC and the A1 – A4 antennas are determined by the simple formulas:

$$\cos(\alpha_{SC}) = \cos(\Phi_{SC})\cdot\sin(\Psi_{SC});\ \cos(\beta_{SC}) = \cos(\Phi_{SC})\cdot\cos(\Psi_{SC});\ \cos(\gamma_{SC}) = \sin(\Phi_{SC}); \quad (2)$$

$$\cos(a_{A1}) = -\sin(\Phi_{RP})\cdot\sin(\Psi_{RP}); \cos(b_{A1}) = -\sin(\Phi_{RP})\cdot\cos(\Psi_{RP});\ \cos(g_{A1}) = \cos(\Phi_{RP}); \quad (3)$$

$\cos(a_{A2}) = \sin(\Phi_{RP})\cdot\sin(\Psi_{RP});\cos(b_{A2}) =$
$\sin(\Phi_{RP})\cdot\cos(\Psi_{RP});\ \cos(g_{A2}) = -\cos(\Phi_{RP});$ (4)

$\cos(a_{A3}) = \cos(\Psi_{RP});\cos(b_{A1}) =$
$= -\sin(\Psi_{RP});\ \cos(g_{A4}) = 0;$ (5)

$\cos(a_{A4}) = -\cos(\Psi_{RP});\cos(b_{A1}) =$
$= \sin(\Psi_{RP});\ \cos(g_{A4}) = 0;$ (6)

$\Phi_{RP} = \Phi_{SC} + \Delta\Phi; \qquad \Psi_{RP} = \Psi_{SC} + \Delta\Psi;$ (7)

The path length difference ΔR_{ij} and propagation time ΔT_{ij} of the beams between the antennas Ai and Aj, i,j = 1 - 4, i ≠ j is given by:

$\Delta R_{ij} = \Delta R_{i0} - \Delta R_{i0}$ (8)

$\Delta T_{ij} = \Delta R_{ij}/c,$ (9)

where *c* is the speed of light.

The phase shift between the signals of the Ai and Aj antennas on the carrier frequency is given by:

$$\Delta\varphi_{ij} = \omega_c \cdot \Delta T_{ij} = 2 \cdot \pi \cdot f_c \cdot \Delta T_{ij} \quad (10)$$

The acquisition of the sum and difference signals from the vertical and horizontal antennas of the B-529 array in each of the operating bands M1, M2, D1, D2, D4 is performed on the circuit of hybrid ring couplers [3, pp. 94 – 95], that are essentially a coil of coaxial transmission line cable, 3λ/2 long with four ports each placed one quarter wavelength away from each other, as shown in Figure 4. When the port 2 is exited, waves travel in both directions through the ring, at the ports 3 and 1 the signal is in phase, and at the port 4 it is in anti-phase. Thus, the signal power is split equally between the ports 1 and 3, and the port 4 is isolated [3, p. 95].

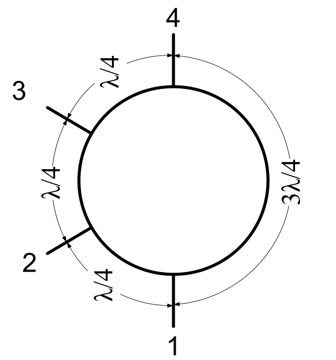


Fig. 4. A hybrid ring on a coaxial cable line

When the port 4 is exited, the waves traveling in both directions through the ring are also in phase at the ports 3 and in anti-phase at the port 2. Therefore, in this case the signal power is split equally between the ports 1 and 3, and the port 2 is isolated. When the ports 2 and 4 are excited simultaneously with an in-phase signal, the waves at the port 3 coming from the ports 2 and 4 are in phase and summarized in amplitude, while the waves at port 1 are in anti-phase and subtracted amplitude.

With the equal power and an amplitude of $\sqrt{2RPs}$ for each of the in-phase signals at the ports 2 and 4, the port 3 receives in-phase signals and the port 1 receives counterphase signals with a power of $P_s/2$ and an amplitude of $\sqrt{RPs}$. Here R is the loading impedance, matched with the iterative impedance of the line. At the port 3 the incoming signals combine and have an amplitude of $\sqrt{\boldsymbol{R}}$ s and a resulting power of $2P_{c,}$ and at the port 1 the signals annihilate and their resulting power is 0. Therefore, the power of the signal at the port 3 equals the sum power of the in-phase signals, fed into the ports 2 and 4. The sum and differential waveform generation in the B-529 antenna array is based on this feature of the hybrid ring.

The noise power P_n of the input ports 2 and 4 is divided equally between the ports 1 and 3, and since the noise signals are independent, they are summed in power at the ports 2 and 4. Therefore, the noise signals with a power of P_n from the ports 2 and 4 give a noise signal of a singular power P_n at the ports 3 and 1.

This results in the doubling of the noise-to-signal ratio when in-phase signals are summed on the hybrid ring: $P_{s3}/P_{n3}=2(P_s/P_n)$.

The schematic diagram of the analogue forming of sum and difference signals of plane-linear polarization individually for the vertical and horizontal polarization from the B-529 antenna array is shown in Figure 5. The diagram features two stages of signal combining with in-phase signals being combined on both stages. Therefore, the signal to noise ratio is doubled on each stage with the resulting signal to noise ratio of $4(P_s/P_n)$.

It is assumed that after forming of the sum and difference signals at the hybrid rings, the further signal processing is performed on modern digital receivers. Since the power of the thermal noise is given by the formula:

$$P_n = k \cdot t^{\circ} \cdot \Delta F_{LNA}\ , \quad (11)$$

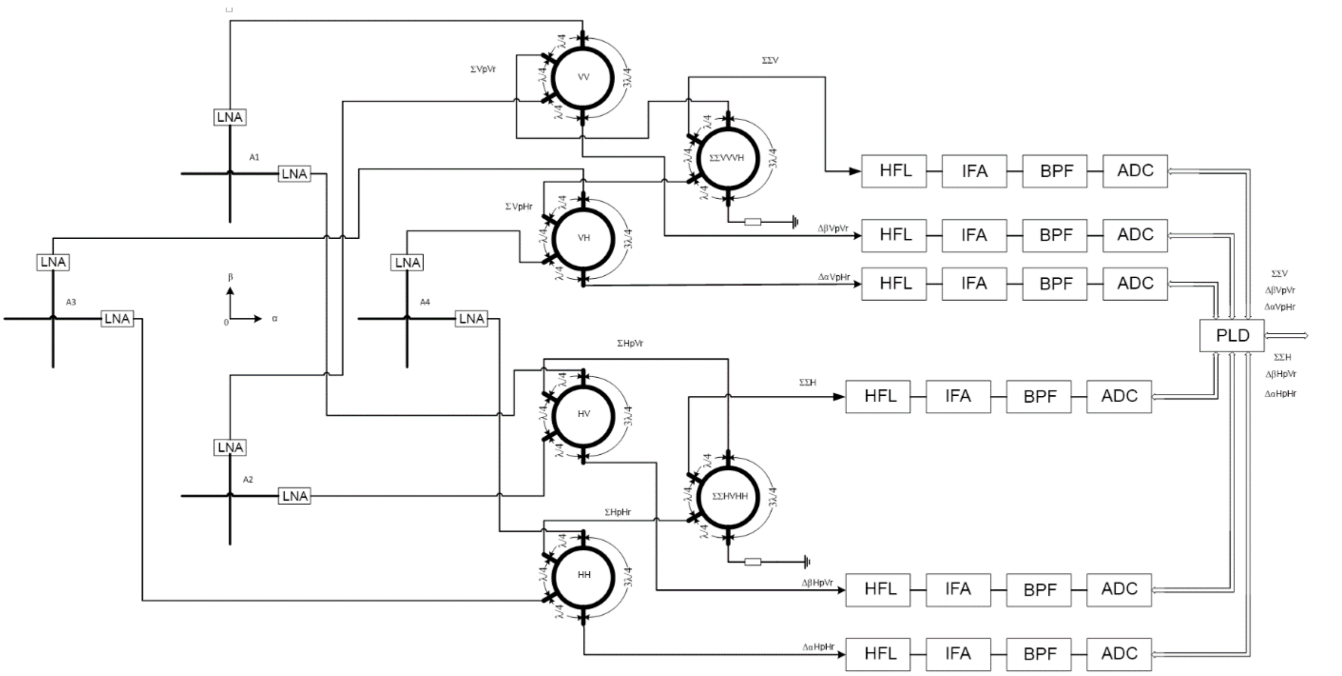


Fig. 5. Schematic diagram of the analogue forming and digital reception of sum and difference signals of plane-linear polarization from the B-529 antenna array

where k is the Bolzmann constant, t° is LNA front end noise temperature, ΔF is LNA band, the bandwidth filter (BWF) is the key factor in determining the signal to noise ratio after the IF amplifier, after which she signal-to-noise ratio equals

$$\left(\frac{P_{s\Sigma}}{P_{n\Sigma}}\right)_{\mathrm{A}} = 4 \cdot \frac{P_S}{P_n} \cdot \frac{\Delta F_{LNA}}{\Delta F_{IFA}}, \tag{12}$$

where ΔF_{IFA} is the bandwidth of the BWF after the IFA. When comparing the analog and the digital methods of summing of the signals from the B-529 antenna array, the signal ratios are the relevant factor, therefore the propagation ratio at the receive path can be assumed to be equal to 1.

The schematic diagram in Figure 6 shows the digital process of signal combining.

Using the digital signal processing method suggests having 8 independent signal channels with vertical and horizontal polarization from each of the four antennas of the B-529. Each of the channels consists of a high-frequency line, an intermediate frequency amplifier (IFA), a band-pass filter (BPF) and an analog to digital converter (ADC).

The digitized signals are fed to the programmable logic device (PLD), which processes them. The signals from the four antennas directed on the object, are as before phased by adjusting the length of the cable from the LNA to the HFL, identical to the analog processing. The hybrid rings are removed from the scheme. Their functions of forming of the sum and difference signals are performed by the PLD.

In this case, when the antenna is directed at the object, in-phase signals for the same polarization (vertical and horizontal) are fed into the ADC. In the PLD the signals with the same polarization are summed by amplitude. With a signal power of P_s and an amplitude of $\sqrt{2R_s}$ the amplitude of the resulting signal will be equal $4\sqrt{2R_s}$ and the power will be $16P_s$. The noise signals are summed by power, therefore, with a noise power in one of the channels of P_s at the LNA input, the noise power at the output of the band-pass filter of the IFA it will equal $P_s(\Delta F_{IFA}/\Delta F_{LFA})$, and after summing of the four signals it becomes $4P_s(\Delta F_{IFA}/\Delta F_{LFA})$. Therefore, the signal-to-noise ratio for the linear polarization after the summing of the signals from the four antennas equals

$$\left(\frac{P_{s\Sigma}}{P_{n\Sigma}}\right)_{D} = 4 \cdot \frac{P_S}{P_n} \cdot \frac{\Delta F_{LNA}}{\Delta F_{IFA}} \tag{13}$$

From (2) and (3) it is apparent that with the antenna array perfectly directed on the source of the signal will have equal signal-to-noise ratio for both digital and analog processing:

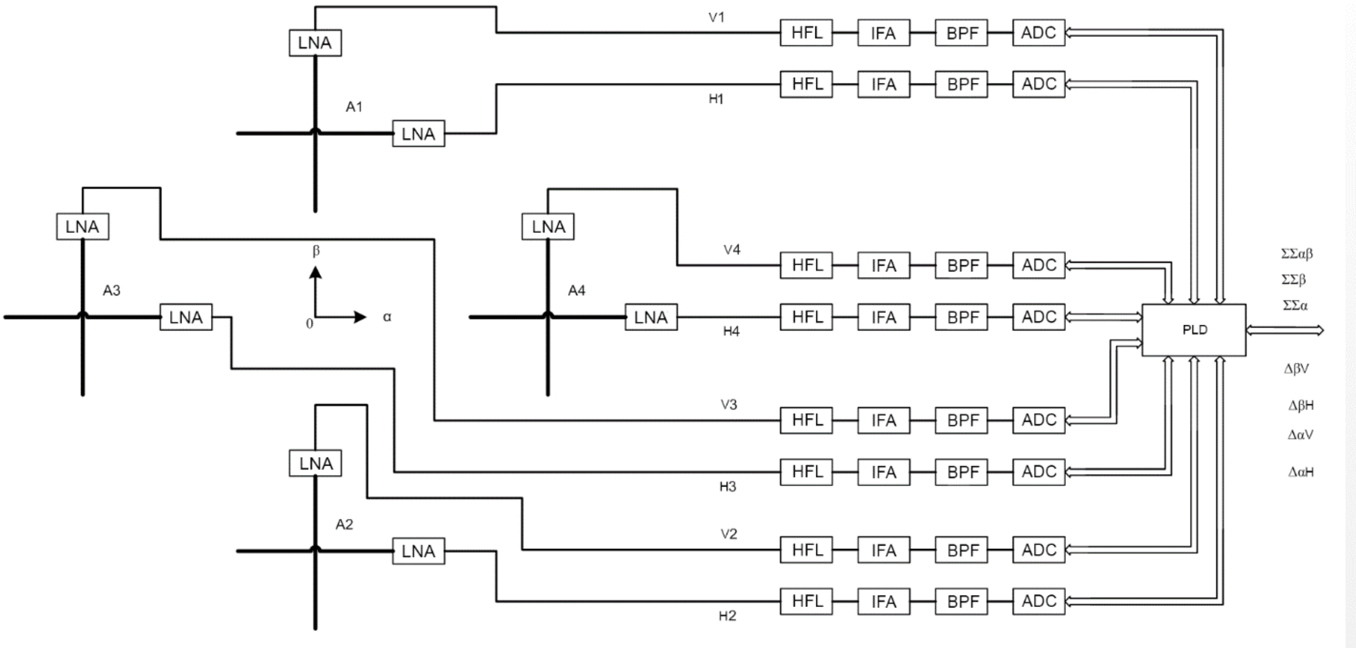


Fig. 6. Schematic diagram of digital signal processing RTS-9

$$\frac{\left(\frac{P_{s\Sigma}}{P_{n\Sigma}}\right)_D}{\left(\frac{P_{s\Sigma}}{P_{n\Sigma}}\right)_A} = 1 \tag{14}$$

Since the total antenna area of the four 6-meter dish antennas of the B-529 system is equal to the total area of the 12-meter TNA-57 antenna, the sum data signal range of the properly phased B-527 antenna array should be commeasurable with the range of the TNA-57 antenna.

Let us compare the characteristics of the analog and digital autotracking of the B-529 antenna array. It is imperative to consider here an example of an inaccurate antenna array positioning on the source of the radio signal.

If the signal power of a single antenna of the antenna array perfectly directed on a spacecraft is P_s, then the signal power from a signal antenna at the difference output port of the hybrid ring is $P_s/2$ with an amplitude of

$$U_m = \sqrt{RP_s} \tag{15}$$

The amplitude of a signal from a single antenna, deviated by Θ angle, equals

$$U_{m\Theta} = U_m \cdot F(\Theta), \tag{16}$$

where F(Θ) is the normalized pattern of a single parabolic antenna of the B-529 antenna array, which for the estimated calculations can be approximated as

$$F(\Theta) = \exp[-a \cdot (\Theta/\Theta_{0.5})^2] \tag{17}$$

Here $\Theta_{0.5}$ is half the width of the radiation pattern at 0.5 of the signal power and 0.707 of the signal amplitude and a = 0,346574.

With simultaneous deviation from the direction to the object by ΔΨ angle in elevation and by ΔΦ angle in azimuth the resulting deviation angle Θ is calculated from the relation for the angle between the vector of the radiation pattern and the vetor of the direction to the SC:

$$\cos\Theta = \cos(\alpha_{RP}) \cdot \cos(\alpha_{SC}) + \cos(\beta_{RP}) \cdot \cos(\beta_{SC}) + \\ + \cos(\gamma_{RP}) \cdot \cos(\gamma_{SC}), \tag{18}$$

where, in accordance with (2) and (7):

$$\cos(\alpha_{RP}) = \cos(\Phi_{SC} + \Delta\Phi) \cdot \sin(\Psi_{SC} + \Delta\Psi); \tag{19}$$

$$\cos(\beta_{RP}) = \cos(\Phi_{SC} + \Delta\Phi) \cdot \cos(\Psi_{SC} + \Delta\Psi); \tag{20}$$

$$\cos(\gamma_{RP}) = \sin(\Phi_{SC} + \Delta\Phi); \tag{21}$$

Hence,

$$\Theta = \arccos(\cos\Theta) \tag{22}$$

With the deviation of the axis of the phased antenna array from the direction to the SC, a phase shift $\Delta\varphi_{ij}$ occurs between the signals received by the antennas Ai and Aj. The phase shift is computed using the formulas (1) to (10).

The vector diagram for the difference signal of the antennas is presented on Fig. 7.

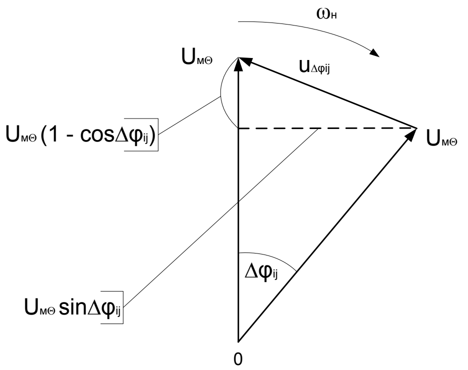


Fig. 7. The vector diagram for the difference signal of the antennas

When the antenna array is deviated from the object by ΔΨ in tilt, and by ΔΦ in azimuth, with the Ai and Aj antenna signals shifted in phase by Δφij, the modulus of the amplitude of the difference signal UΔφij in accordance with the vector diagram in Fig. 9 is calculated by the obvious formula:

$$U_{\Delta\varphi ij} = \sqrt{\begin{matrix}(U_{\text{м}\Theta} - U_{\text{м}\Theta} \cdot cos\Delta\varphi ij)^2 + \\ +(U_{\text{м}\Theta} \cdot sin\Delta\varphi ij)^2\end{matrix}}$$

Whence

$$U_{\Delta\varphi ij} = U_{\text{м}\Theta}\sqrt{2(1 - cos\Delta\varphi ij)} \tag{23}$$

The power of the difference signal of the Ai and Aj antennas after the analog subtraction, taking (15) and (16) into account, equals:

$$\begin{aligned} P_{\text{c}\Delta ij\text{A}} &= \frac{U^2_{\Delta\varphi ij}}{2R} = \frac{U^2_{\text{м}\Theta} \cdot 2(1 - cos\Delta\varphi ij)}{2R} = \\ &= \frac{U^2_{\text{м}} \cdot \left(F(\Theta)\right)^2(1 - cos\Delta\varphi ij)}{R} = \\ &= \frac{R \cdot P_{\text{c}} \cdot \left(F(\Theta)\right)^2(1 - cos\Delta\varphi ij)}{R} = \\ &= P_{\text{c}} \cdot \left(F(\Theta)\right)^2(1 - cos\Delta\varphi ij) \end{aligned} \tag{24}$$

As mentioned before, the noise with a power of P_n from each of the two single antennas at the sum and difference output ports of the hybrid ring, add up to a single power of P_n. Therefore, the signal-to-noise power ratio for the difference signals from the Ai and Aj antennas deviated from the direction to the SC is calculated by the formula:

$$\frac{P_{\text{s}\Delta ij}}{P_{\text{n}\Delta ij}} = \frac{P_{\text{s}}}{P_{\text{n}}} \cdot \left(F(\Theta)\right)^2(1 - cos\Delta\varphi ij) \tag{25}$$

After the BPF of the IFA, the signal-to-noise power ratio for the difference signals from the Ai and Aj antennas deviated from the direction to the SC becomes equal to:

$$\left(\frac{P_{\text{s}\Delta ij}}{P_{\text{n}\Delta ij}}\right)_{\text{BPF IFA}} = \frac{P_{\text{s}}}{P_{\text{n}}} \cdot \frac{\Delta F_{\text{LNA}}}{\Delta F_{\text{IFA}}} \cdot \left(F(\Theta)\right)^2 \cdot \cdot (1 - cos\Delta\varphi ij) \tag{26}$$

Whence, the normalized values for the signal-to-noise ratio of the difference signal in relation to the signal-to-noise ratio of the signal from a single antenna after the BPF of the IFA with the analog subtraction:

$$\frac{\left(\frac{P_{\text{s}\Delta ij}}{P_{\text{n}\Delta ij}}\right)_{\text{BPF IFA}}}{\left(\frac{P_{\text{s}}}{P_{\text{n}}}\cdot\frac{\Delta F_{\text{LNA}}}{\Delta F_{\text{IFA}}}\right)} = = (\text{F}(\Theta))^2(1 - cos\Delta\varphi ij) \tag{27}$$

With the digital forming of the difference signal the signal-to-noise ratio, that equals P_s/P_n, remains the same up to the BPF of the IFA, and after the filter it equals $(P_s/P_n)(\Delta F_{LNA}/\Delta F_{IFA})$. With the digital subtraction the noise are summed by power, but, unlike with the hybrid bridge, the power of the signal isn't divided in two, therefore, instead of (15) we have:

$$U_c = \sqrt{2RPs} \tag{28}$$

Whence

$$U_s^2 = 2RP_s \tag{29}$$

The amplitude of the signal of a single antenna with the antenna array declined by Θ degrees equals

$$U_{s\Theta} = U_s \cdot F(\Theta), \tag{30}$$

By analogy with (16), (23) and (24) with regard of (28) – (30) with the digital signal processing method, before the BPF, a difference signal with twice the power is received:

$$P_{s\Delta ijD} = 2P_s \cdot \left(F(\Theta)\right)^2(1 - cos\Delta\varphi ij) \tag{31}$$

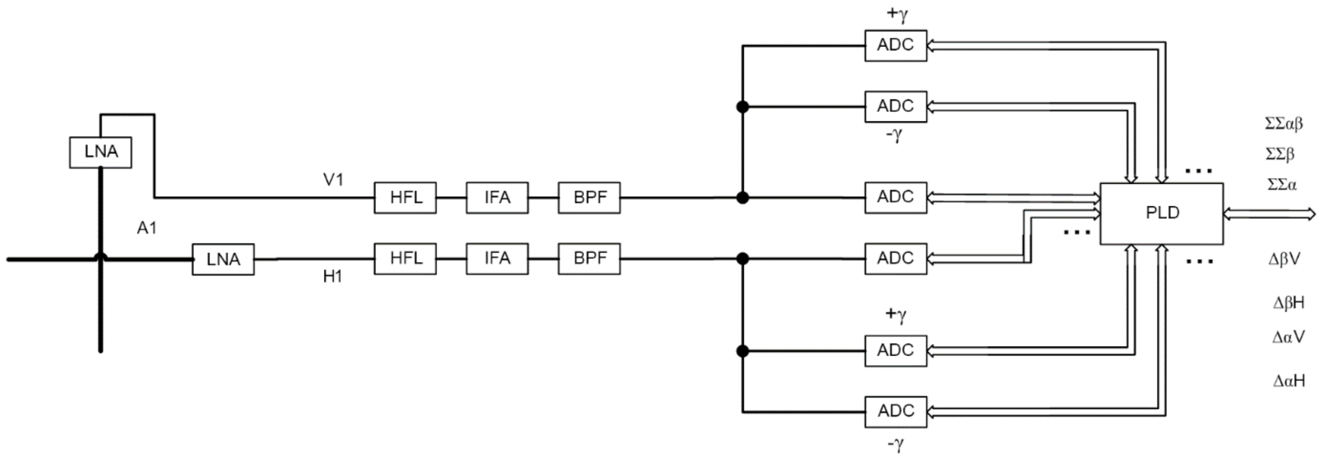


Fig. 8. Schematic diagram of digital signal processing of one of the four RTS-9 antennas for the difference auto-tracking of the sum signals.

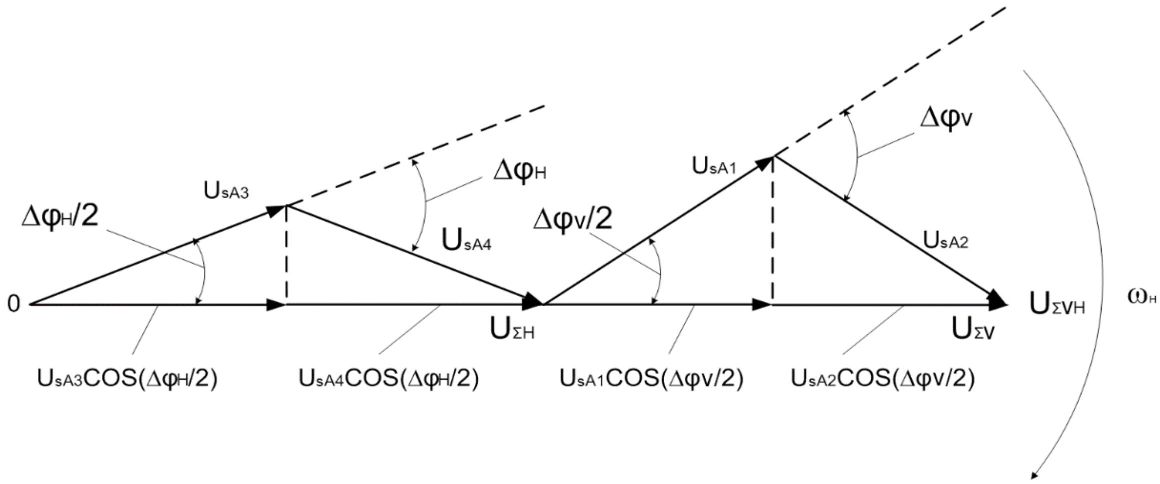


Fig. 9. Vector diagram for digital forming a sum signal of the vertical and horizontal rows of a deviated antenna array

However, the noise will also double in power. As a result, the same normalized values of signal-to-noise ratio of the difference signal are received, as compared to the signal-to-noise ratio of a single antenna signal after the BPF of the IFA, processed with the analog method:

$$\frac{\left(\frac{P_{s\Delta ij}}{P_{n\Delta ij}}\right)_{D\,BPF\,IFA}}{\left(\frac{P_s}{P_n}\cdot\frac{\Delta F_{LNA}}{\Delta F_{IFA}}\right)} = (\mathrm{F}(\Theta))^2(1-cos\Delta\varphi ij) \quad (32)$$

Therefore, the direct conversion of the method of difference signal forming from analog to digital processing will not improve the signal-to-noise ratio in the hypothetical case of a perfect analog processing. However, aside from being implemented with the perfect characteristics of the methods used for analog phasing, the digital processing enables new possibilities for the improvement of the signal-to noise ratio while forming a difference signal.

It is suggested that the forming of the difference guidance signal should be implemented for the sum signal from all of the four antennas of the B-529. For this purpose, it is suggested that the beam of the antenna array is oscillated horizontally and vertically and the difference signals of the combined beam of the array between the corresponding deviated positions are calculated. In other words, the beam central line is used. The schematic diagram of the suggested technical realization is presented in Figure 8.

This figure demonstrates the line of devices for a single antenna only. The schemes of the signal processing for the other 3 antennas are identical. Here, in addition to the ADC for the forming of the signal with vertical and horizontal polarization without a phase shift, two additional ADC are introduced to form a signal leading and lagging in phase by a certain angle μ, the selection of which will be discussed later. When the antenna is deviated from the direction to the object, alongside with the difference signals, upon which the principle of autotracking is based, being generated, there occurs a reduction in the power of the sum signal. The vector diagram for digital forming a sum signal of the vertical and horizontal rows of a deviated B-529 antenna array are presented in Figure 9.

The vector diagram in the Fig. 9 considers that in accordance with (1)–(10) the phase shifts of the signals of the A1 and A2 antennas related to the phase array center are equal in modulus and opposite in sign. The same is correct to the phase shifts in relation to the phase center of the antennas A3 and A4. This means that the pairwise sums of the signals from the A1, A2 antennas and A3, A4 antennas have the same phase, equal to the phase of the sum signal of all four antennas with any deviation of the radiation pattern axis from the direction to the spacecraft.

According to Figure 9, the amplitude of the digital sum signal is given by:

$$U_{\Sigma VH} = 2U_{s\Theta}\left[\cos\left(\frac{\Delta\varphi V}{2}\right) + \cos\left(\frac{\Delta\varphi H}{2}\right)\right] \qquad (33)$$

The resulting power of the vertical and horizontal array rows:

$$P_{s\Sigma VH} = \frac{U_{\Sigma VH}^2}{2R} = \frac{\left[2U_{s\Theta}\left\{\cos\left(\frac{\Delta\varphi V}{2}\right) + \cos\left(\frac{\Delta\varphi H}{2}\right)\right\}\right]^2}{2R} =$$

$$= \frac{4 \cdot \cancel{2 \cdot R} \cdot P_s [F(\Theta)]^2 \left[\cos\left(\frac{\Delta\varphi v}{2}\right) + \cos\left(\frac{\Delta\varphi h}{2}\right)\right]^2}{\cancel{2R}}, \qquad (34)$$

$$P_{s\Sigma VH} = 4P_s[F(\Theta)]^2\left[\cos\left(\frac{\Delta\varphi V}{2}\right) + \cos\left(\frac{\Delta\varphi H}{2}\right)\right]^2$$

Since with the digital signal processing the noise in the four lines are independent, they are summed by power, after the BPF IFA:

$$P_{nVH} = 4P_n \cdot \frac{\Delta F_{IFA}}{\Delta F_{LNA}} \qquad (35)$$

Hence the signal-to-noise ratio after the BPF IFA:

$$\frac{P_{s\Sigma VH}}{P_{nVH}} = \frac{P_s}{P_n} \cdot \frac{\Delta F_{LNA}}{\Delta F_{IFA}} \cdot \qquad (36)$$

$$\cdot \left(F(\Theta)\right)^2 \left(\cos\left(\frac{\Delta\varphi V}{2}\right) + \cos\left(\frac{\Delta\varphi H}{2}\right)\right)^2$$

its normalized value related to the signal-to-noise ratio for the signal from a single antenna after the BPF IFA equals:

$$\frac{\left(\frac{P_{s\Sigma VH}}{P_{nVH}}\right)}{\left(\frac{P_s}{P_n}\cdot\frac{\Delta F_{LNA}}{\Delta F_{IFA}}\right)} = \qquad (37)$$

$$= (F(\Theta))^2\left(\cos\left(\frac{\Delta\varphi V}{2}\right) + \cos\left(\frac{\Delta\varphi H}{2}\right)\right)^2$$

Now let us evaluate the potential possibilities of the digital processing while forming a difference auto-tracking signal by the equisignal zone method based on the sum signal from all four antennas.

Let us examine the forming of the difference signal with the beam electronically declined in vertical direction. First, let us discuss the situation when the axis of the radiation pattern is deviated from the direction to the SC by ΔΨ in azimuth and by ΔΦ in elevation, presented on Figure 10.

In this situation the A1 and A3 antennas are extended towards the source of the signal. In other words, the B-529 is vertically declined towards the reduction of the elevation by ΔΦ angle, and horizontally by ΔΨ angle towards the increased azimuth angle.

Now we shall deflect the sum beam vertically towards the increasing elevation, introducing latency for the signal from the A1 antenna by the phase angle μ and lead for the signal from the A2 antenna by the same phase angle μ. This leads to some amount of compensation of the original declination of the antenna array towards the reduced elevation and the increase in the amplitude of the sum signal, as shown in Figure 10.

Now we shall decline the sum beam vertically towards the reducing elevation introducing the lead for the signal from the A1 antenna by the same phase angle μ and the lag for the A2 antenna by still the same phase angle μ. This leads to a greater deviation of the antenna array towards the reducing elevation in addition to the ΔΦ angle, and to even greater reduction of the amplitude of the sum signal.

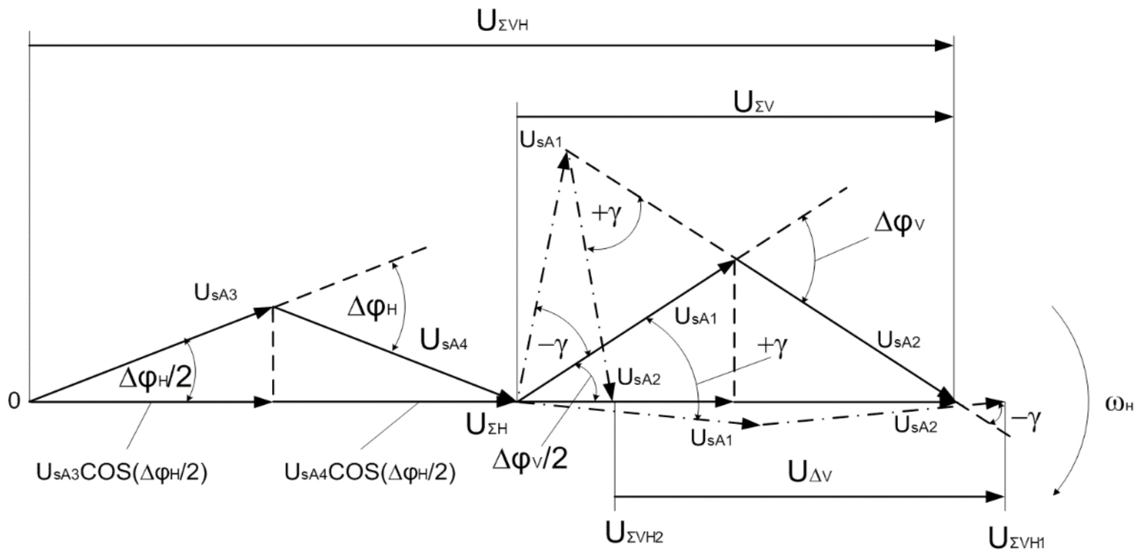


Fig. 10. Forming of the difference signal by the sum signal of the vertical row of the antenna array with a vertical deviation towards the reducing elevation.

Now we shall subtract the second sum signal from the first one to acquire the difference signal for the compensation of deviation of the antenna signal from the vertical direction to the object.

According to Fig. 10:

$$U_{\Sigma VH1} = U_{s\Theta}\left[2\cos\left(\frac{\Delta\varphi H}{2}\right) + \cos\left(-\frac{\Delta\varphi V}{2} + \mu\right) + \right.$$
$$\left. + \cos\left(+\frac{\Delta\varphi V}{2} - \mu\right)\right] \tag{38}$$

$$U_{\Sigma VH2} = U_{s\Theta}\left[2\cos\left(\frac{\Delta\varphi H}{2}\right) + \cos\left(-\frac{\Delta\varphi V}{2} - \mu\right) + \right.$$
$$\left. + \cos\left(+\frac{\Delta\varphi V}{2} + \mu\right)\right] \tag{39}$$

$$U_{\Delta V} = U_{\Sigma VH1} - U_{\Sigma VH2} =$$
$$= 2U_{s\Theta}\left[\cos\left(\frac{\Delta\varphi V}{2} - \mu\right) - \cos\left(\frac{\Delta\varphi V}{2} + \mu\right)\right] \tag{40}$$

The difference signal of the vertical guidance depends only on the sum signal of the two antennas of the vertical row. Therefore, in order to prevent the accumulation of additional noise from the orthogonal row of the array, when forming the vertical auto-tracking difference signal, it is advisable to use the sum signal only from the vertical antennas, and from the horizontal ones for the horizontal auto-tracking.

Figure 11 shows the procedure of forming of the difference vertical guidance signal with the antenna array declined towards the increasing elevation. The difference signal in this case is given by Formula (40), but it changes its sign to the opposite, since the angle of deviation ΔΦ changes its sign to the opposite, and, consequently, the phase shift Δφv changes its sign.

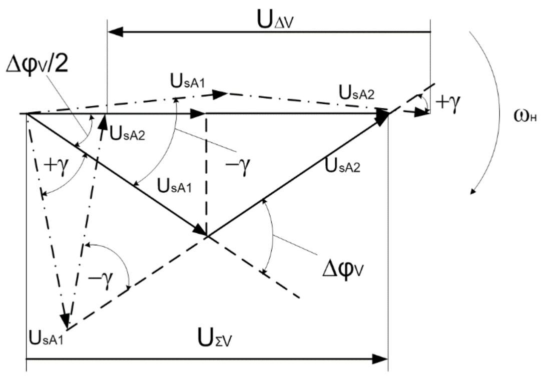


Fig. 11. Forming of the difference signal by the sum signal from the vertical row of the array with vertical deviation towards the increasing elevation.

Figure 12 shows, that without any vertical deviation of the antenna array, the difference signal equals 0 with any oscillating angle μ. Now let us evaluate the signal-to-noise ratio for the difference signal by the sum signal of the vertical row of the antenna array by the equisignal zone method.

From (40) the formula for the power of the difference signal by the sum signal of the vertical row of the array is deduced:

$$P_{s\Delta V} = 4P_s[F(\Theta)]^2\left[\cos\left(\frac{\Delta\varphi 12}{2} - \mu\right) - \cos\left(\frac{\Delta\varphi 12}{2} + \mu\right)\right]^2 \tag{41}$$

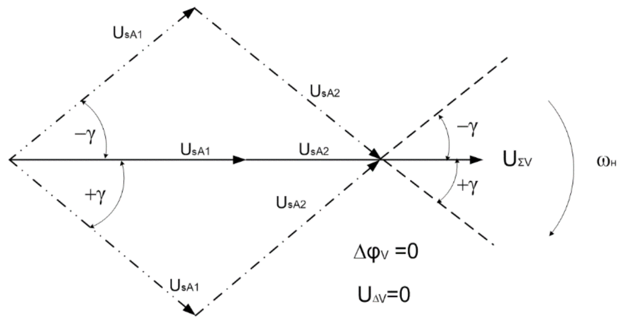


Figure 12. Forming of a zero difference signal by the sum signal of the vertical row of the array without vertical deviation – equisignal zone.

Since the noise from the LNA for the each of the “vertical” antennas twice independently affects the forming of the difference signal (when combining and subtracting), the total power of the noise after the BPF IFA equals:

$$P_{s\Delta V} = 4P_s \cdot \frac{\Delta F_{IFA}}{\Delta F_{LNA}} \tag{42}$$

From (41) and (42) the formula for the signal-to-noise ratio when forming the difference signal by the sum signal of the vertical row of the antennas by the equisignal zone method is derived:

$$\frac{P_{s\Delta V}}{P_{n\Delta V}} = \frac{P_s}{P_n} \cdot \frac{\Delta F_{LNA}}{\Delta F_{IFA}} \cdot$$
$$\cdot [F(\Theta)]^2 \left[\cos\left(\frac{\Delta\varphi 12}{2} - \mu\right) - \cos\left(\frac{\Delta\varphi 12}{2} + \mu\right)\right]^2 \tag{43}$$

Whence, the normalized value of the signal-to-noise ratio of the difference signal by the sum signal of the vertical row of the antenna array by the equisignal zone method, in relation to the signal-to-noise ratio of the input signal of a single antenna after the BPF IFA equals:

$$\frac{\left(\frac{P_{s\Delta V}}{P_{n\Delta V}}\right)}{\left(\frac{P_s}{P_n} \cdot \frac{\Delta F_{LNA}}{\Delta F_{IFA}}\right)} =$$
$$= (F(\Theta))^2 \left[\cos\left(\frac{\Delta\varphi 12}{2} - \mu\right) - \cos\left(\frac{\Delta\varphi 12}{2} + \mu\right)\right]^2 \tag{44}$$

The signal-to-noise ratio for the difference auto-tracking signals processed with digital and analog equisignal method is deduced by dividing (44) by (27):

$$\frac{\left(\frac{P_{\Delta V}}{P_{n\Delta V}}\right)_D}{\left(\frac{P_{\Delta V}}{P_{n\Delta V}}\right)_A} =$$
$$= \frac{\left[\cos\left(\frac{\Delta\varphi 12}{2} - \mu\right) - \cos\left(\frac{\Delta\varphi 12}{2} + \mu\right)\right]^2}{1 - cos\Delta\varphi 12} = 2 \cdot (sin\mu)^2 \tag{45}$$

Obviously, this equation attains a maximum at = 2 with μ=π/2 and sinμ=1.

Hence, with the deviation angle μ=π/2, the signal-to-noise ratio is two times higher when using the equisignal digital method as compared to the original analog method of acquisition of the difference auto-tracking signal on the hybrid ring couplers. This is correct providing two vertical and two horizontal antennas are used to form the difference signal. The downside is that three times as many ADC is required (24 rather than 8).

Notably, the design of the B-529 antenna array allows it to rotate the guidance axis by 45° and use the guidance method forming the sum and difference signals following the (A1 + A3) – (A4 + A2) scheme when the axis is declined by + 45° and the (A1+A4) – (A3+A2)) scheme when the axis is declined by – 45° (see Fig. 2).

Before we discuss this “diagonal” tracking method, it is important to note that according to (1) – (10) since the B-529 antenna array is symmetrical, the phase shifts are equal:

$$\Delta\varphi_{13} = \Delta\varphi_{42} \text{ and } \Delta\varphi_{14} = \Delta\varphi_{32} \tag{46}$$

Using the same method, as before, it is not hard to show that the normalized values of signal-to-noise ratio for the sum-difference signal (A1+A3)-(A4+A2) with analog processing in relation to the signal-to-noise-ratio of the signal from a single antenna after the BPF IFA when subtracted by the analog method:

$$\frac{\left(\frac{P_{s\Delta\varphi 13-42}}{P_{n\Delta\varphi 13-42}}\right)_{A\ BPF\ IFA}}{\left(\frac{P_s}{P_{Ш}} \cdot \frac{\Delta F_{LNA}}{\Delta F_{IFA}}\right)} =$$
$$= 2(F(\Theta))^2 \left[\cos\left(\frac{\Delta\varphi 13}{2}\right)\right]^2 (1 - cos\Delta\varphi 14) \tag{47}$$

With digital processing after BPF IFA the normalized signal-to-noise ratio is the same:

$$\frac{\left(\frac{P_{s\Delta\varphi13-42}}{P_{n\Delta\varphi13-42}}\right)_{D\ BPF\ IFA}}{\left(\frac{P_s}{P_n}\cdot\frac{\Delta F_{LNA}}{\Delta F_{IFA}}\right)} = \tag{48}$$

$$= \left(\frac{P_{s\Delta\varphi13-42}}{P_{n\Delta\varphi13-42}}\right)_{A\ BPF\ IFA} \Big/ \left(\frac{P_s}{P_n}\cdot\frac{\Delta F_{LNA}}{\Delta F_{IFA}}\right)$$

Therefore, the digital signal processing method makes the perfect position control of a phased antenna array possible.

It is of interest to compare the signal-to-noise ratio (47) when guiding by the sum-difference signal from all four antennas to the signal-to-noise ratio (44) when guiding by the equisignal method for two antennas with an oscillating angle $\mu = \pi/2$:

$$\frac{4\cancel{(F(\Theta))^2}\left[\sin\left(\frac{\Delta\varphi\nu}{2}\right)\right]^2}{\cancel{2\cdot(F(\Theta))^2}\left[\cos\left(\frac{\Delta\varphi13}{2}\right)\right]^2\cdot\cancel{2}\cdot\left[\sin\left(\frac{\Delta\varphi14}{2}\right)\right]^2} = \tag{49}$$

$$= \frac{\left[\sin\left(\frac{\Delta\varphi12}{2}\right)\right]^2}{\left[\cos\left(\frac{\Delta\varphi13}{2}\right)\right]^2\left[\sin\left(\frac{\Delta\varphi14}{2}\right)\right]^2}$$

The numerator attains a maximum 1 at $\Delta\varphi_{12} = \pm\pi$. The denominator attains a maximum at $\Delta\varphi_{13} = 0$ or 2π и $\Delta\varphi_{14} = \pm\pi$, i.e. when there is no deviation along the A1-A3 antennas. Since the signal-to-noise ratio for the equisignal method for two antennas does not depend on the lateral deviation of the radiation pattern, and the signal-to-noise ratio for the sum-difference method from all four antennas of the array decreases with a lateral deviation of the pattern, then with identical angular deviation, the signal-to-noise ratio of the equisignal method is never lower than of the sum-difference signal from all four antennas.

Furthermore, a modification of the equisignal method for all four antennas is possible, that will double the signal-to-noise ratio than the sum-difference tracking method. The essence of the modification is to scan the radiation pattern of the array by a certain ν angle in both directions with a ± 45° inclination and extract the difference signal.

By the same method, it can be shown that the normalized signal-to-noise ratio of the difference signal from the four antennas by the equisignal method after the BPF IFA is given by:

$$\frac{\left(\frac{P_{s\Sigma1234\Delta\varphi14}}{P_{n1234\Delta\varphi14}}\right)}{\left(\frac{P_s}{P_n}\cdot\frac{\Delta F_{LNA}}{\Delta F_{IFA}}\right)} =$$

$$= 2\cdot\left(F(\Theta)\right)^2\left[\cos\left(\frac{\Delta\varphi13}{2}\right)\right]^2\cdot \tag{50}$$

$$\cdot\left[\cos\left(\frac{\Delta\varphi14}{2}+\nu\right)-\cos\left(\frac{\Delta\varphi14}{2}-\nu\right)\right]^2$$

By dividing (67) by (47) let us compare the normalized signal-to-noise ratios for the difference signal (A1+A3) – (A4+A2) and the difference signal from the four antennas by the equisignal method:

$$\frac{\left(\frac{P_{s\Sigma1234\Delta\varphi14}}{P_{n1234\Delta\varphi14}}\right)}{\left(\frac{P_{s\Delta\varphi13-42}}{P_{n\Delta\varphi13-42}}\right)} =$$

$$= 2\cdot\left(F(\Theta)\right)^2\left[\cos\left(\frac{\Delta\varphi13}{2}\right)\right]^2\left[\cos\left(\frac{\Delta\varphi14}{2}+\nu\right)-\right. \tag{51}$$

$$\left.-\cos\left(\frac{\Delta\varphi14}{2}-\nu\right)\right]^2\Big/\ 2\left(F(\Theta)\right)^2\left[\cos\left(\frac{\Delta\varphi13}{2}\right)\right]^2\cdot$$

$$\cdot\ (1-\cos\Delta\varphi14)\ = 2\cdot(\sin\nu)^2$$

This expression, obviously, attains a maximum 2 at $\nu = \pi/2$, with $\sin\nu = 1$. Hence, by selecting the declination angle $\nu = \pi/2$, the signal-to-noise ratio of the equisignal digital method can be doubled in power, as compared to the original analog hybrid ring coupler auto-tracking method.

The results received are illustrated by the plots in Figures 13 – 21, showing the signal-to-noise ratios of the sum and difference signals for the analog and digital processing at a carrier frequency of 150 MHz at different sections of the radiation pattern.

Figures 13 – 16 feature the plots of signal-to-noise ratio against deviations in elevation and azimuth that illustrate the influence of the elevation angle of the target on the deviation in azimuth. With target elevation angles close to zenith, the azimuth antenna guidance is lost with any tracking method. This does not occur when error-based tracking in mutually perpendicular directions at an angle of 45° is employed, as shown in figures 17 – 20.

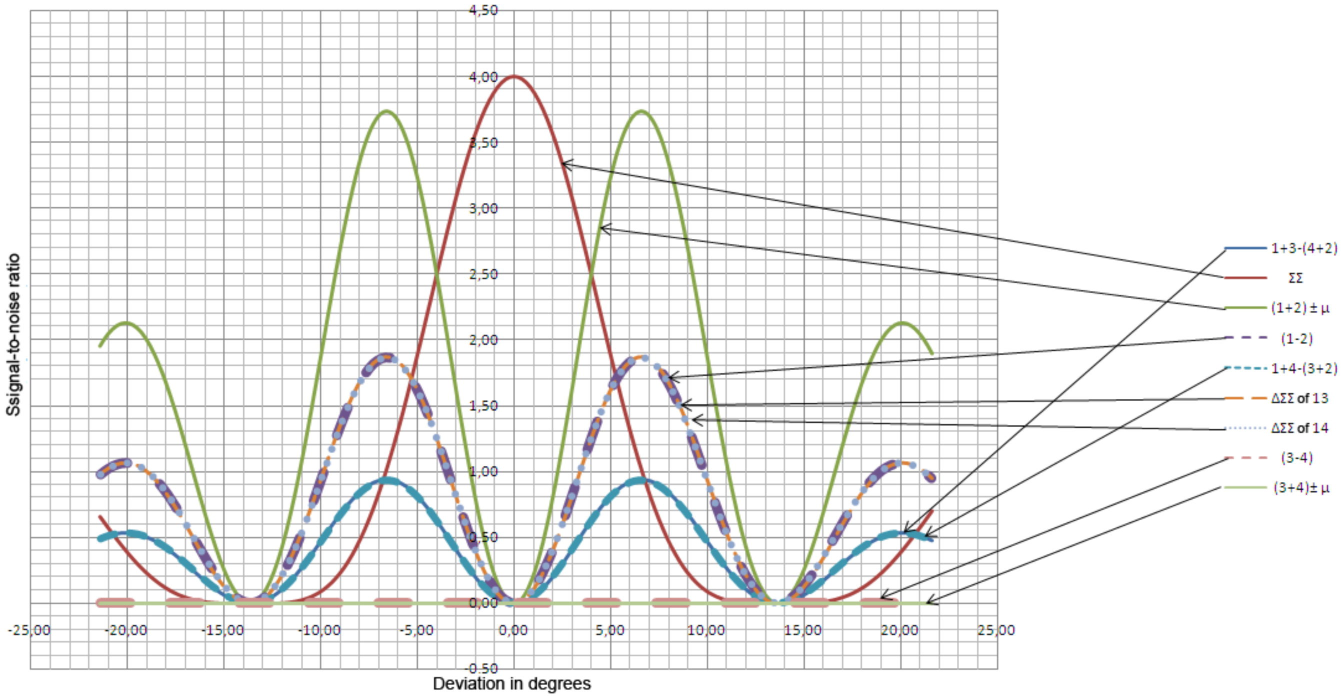


Fig. 13. Elevation error. Does not depend on azimuth or elevation.

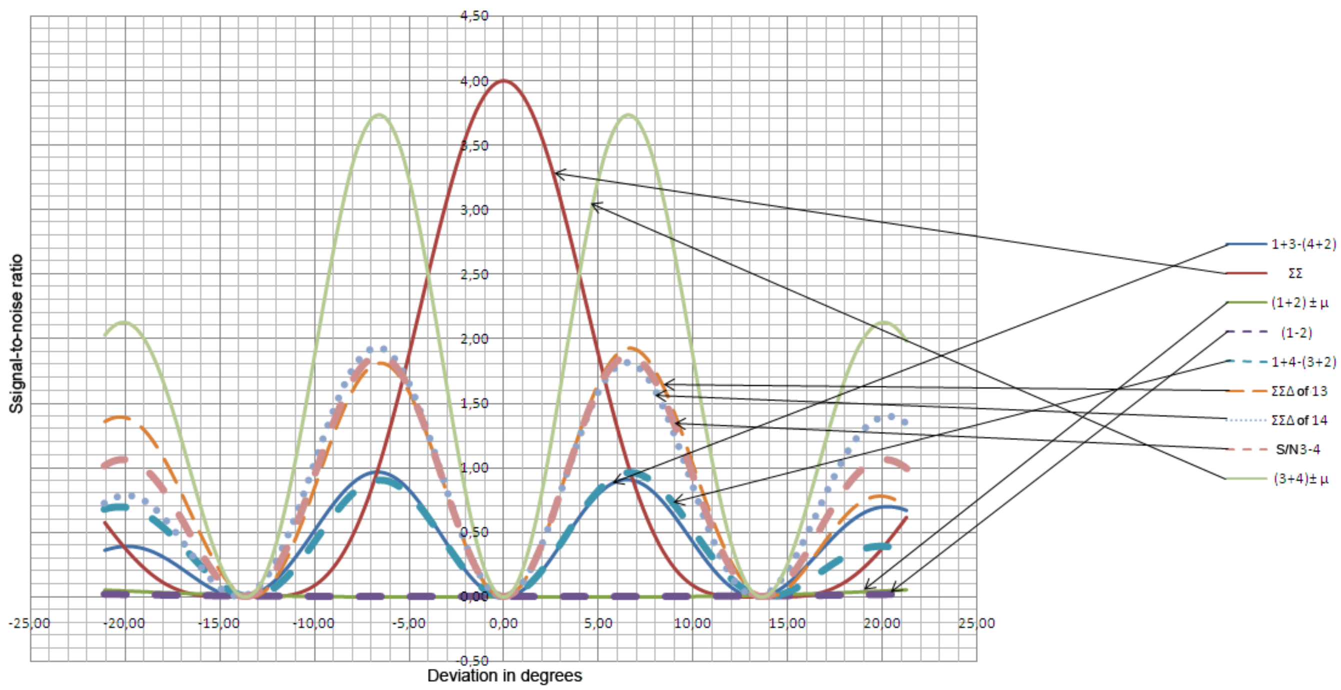


Fig. 14. Azimuth error. 10° elevation.

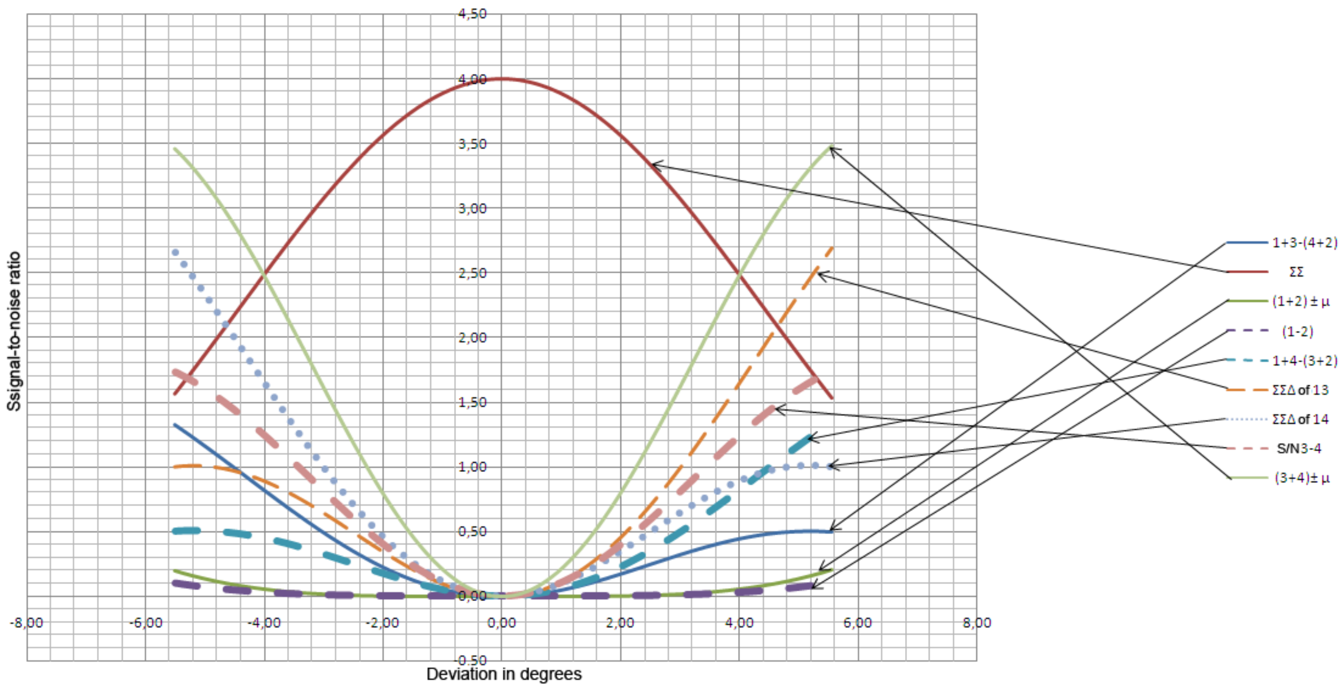


Fig. 15. Azimuth error. 75° elevation.

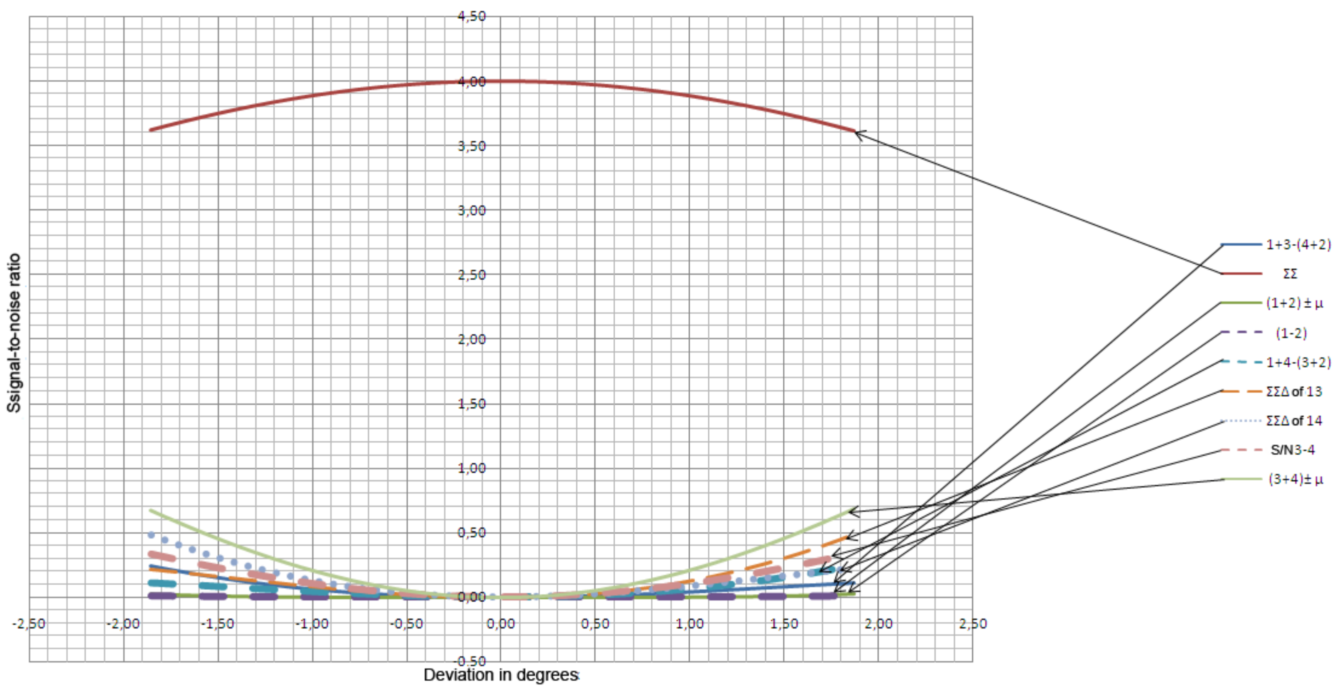


Fig. 16. Azimuth error. 85° elevation.

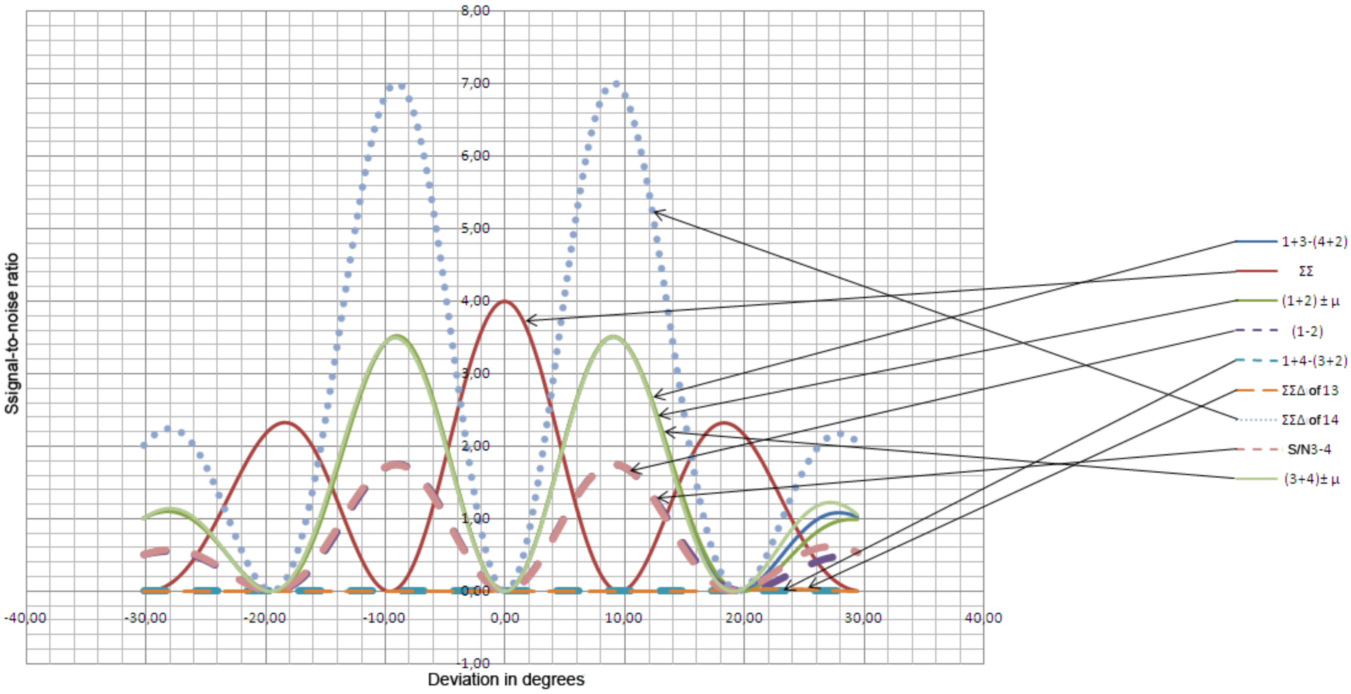


Fig. 17. Direction error for the A1-A4 antennas. 10° elevation. Does not depend on azimuth.

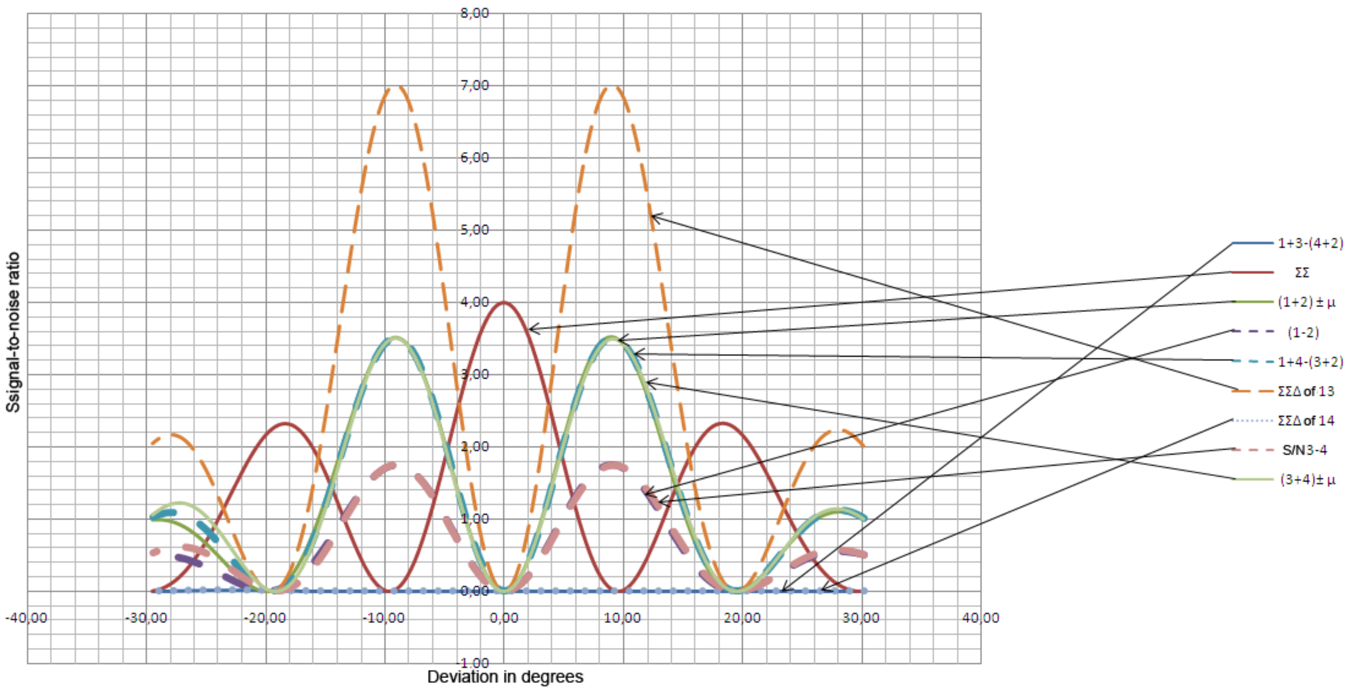


Fig. 18. Direction error for the A1-A3 antennas. 10° elevation.

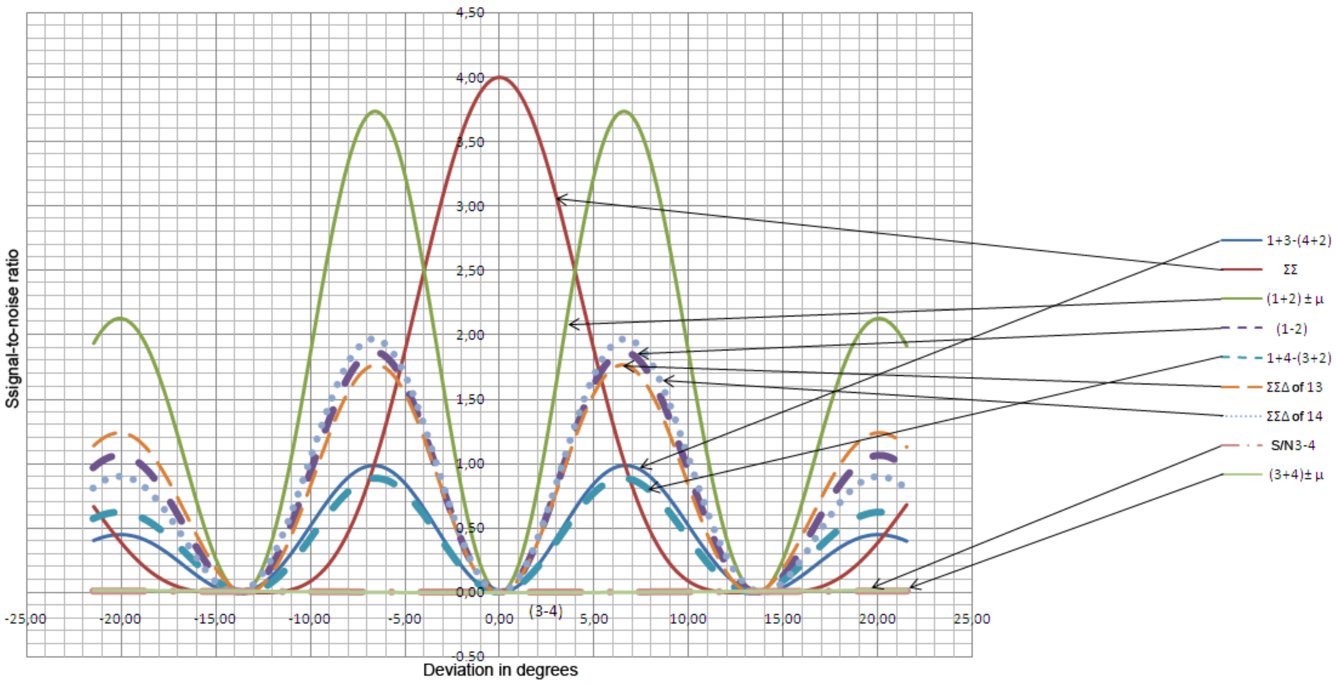


Fig. 19. Direction error for the A1-A4 antennas. 89° elevation.

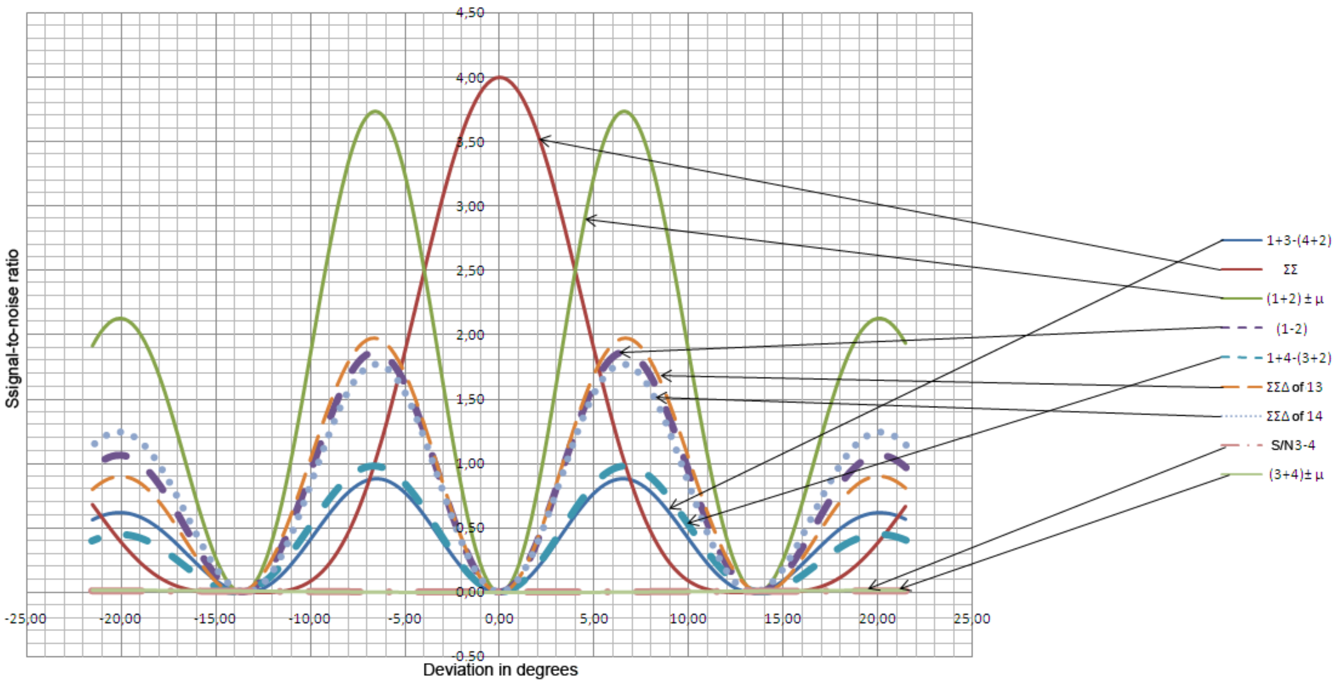


Fig. 20. Direction error for the A1-A3. 89° elevation.

As seen in Figures 19 – 20, the guidance by the difference signals in mutually perpendicular directions at an angle of 45° is not lost even in zenith, besides, the signal-to-noise ratio of the difference signal from all four antennas by the method of the equisignal zone is twice as high as the signal-to-noise ratio of the difference signal of the pairs of antennas angled at 45°, i.e. (A1+A3) – (A4+A2) and (A1+A4) – (A3+A2) signals. The downside is the complicated processing and angle control for the introduced phase shifts during the analog-to-digital conversion, since the phase shift by 90° when scanning the sum beam at 45° is determined by the phase shifts in azimuth and elevation, and those phase shifts consequently depend on the elevation of the target.

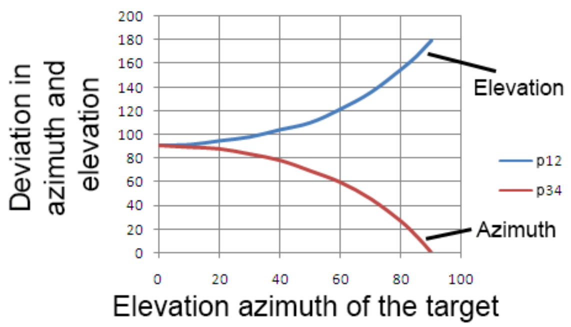


Fig. 21. Errors on the analog-to-digital converter in azimuth and elevation when scanning the radiation pattern by ± 90° at 45° angle.

Conclusion

1. Digital processing provides the same signal-to-noise ratio as an ideal analog processing, i.e. digital radio signal processing allows for perfect guidance of a phased antenna array for the information transmission sum signal from four antennas of the B-529 system.

2. For the difference signal, the equisignal auto-tracking method with digital processing when applied properly, can give a gain in the signal-to-noise power by 2 times compared the analog method in ideal conditions, thus increasing the range of the spacecraft-ground radio link in $\sqrt{2}$ times.

References.

1. Grishmanovskiy V.A., StepanovV.S. Raschetdal'nosti, vremenividimosti KA i energeticheskikh kharakteristik radioliniy dlya peredachi telemetricheskoy informatsii [The calculation of distance, visibility time of a spacecraft and radio links energy characteristics for telemetry data transmission].*Uchebno-metodicheskoeposobie k prakticheskim zanyatiyam po kursu "Komandno-izmeritel'nyy kompleks upravleniya"* [A study guide for tutorials for the course "Command and measurement control complex"]. VIKI im. A.F. Mozhayskogo, 1985. (in Russian)

2. *Energeticheskie kharakteristiki kosmicheskikh radioliniy* [Energy characteristics of space radio links]. Edited by O.A. Zenkevich. Moscow: Sov. Radio, 1972. (in Russian)

3. Voskresenskiy D.I., Gostyukhin V.L., Maksimov V.M., Ponomarev L.I. *Ustroystva SVCh i antenny* [SHF devices and antennas]. Edited by D.I. Voskresenskiy. 2nd edition, revised and enlarged. Moscow: Radiotekhnika, 2006, 376 p.: il. (in Russian)

ROCKET-SPACE DEVICE ENGINEERING AND INFORMATION SYSTEMS
Vol. 3, Iss. 1, 2016, pp. 57–65

System and Technical Aspects of Development of the Ground-based Automated Control Complex for Spacecraft of Scientific and Socio-economic Purposes and Measurements till 2025

M.Yu. Kislyakov[1], N.S. Logachev[2], A.M. Petushkov[3]

[1]candidate of engineering science, [2]doctor of military science, [3]candidate of engineering science
Joint Stock Company "Russian Space Systems", Russia
e-mail: [1]kisl60@mail.ru, [2]vn_i_op@mail.ru, [3]alexandr.petushkov@gmail.com

Abstract. The peculiarities of spacecraft control perspective technologies and main trends in their practical usage, as well as arising problems, are examined. The following spacecraft control technologies are considered: ballistic and navigational provision, information-telemetric support, control via satellite relays, and integration of spacecraft data communication channels. Moreover, a preliminary plan for perspective technologies realization and the image of a perspective ground-based automated control complex for spacecraft of scientific and socioeconomic purposes and measurements is substantiated.

Keywords: ground-based automated control complex for spacecraft of scientific and socioeconomic purposes, information support technology for spacecraft control.

Introduction

According to the project of Russian federal space program for the 2016-2025 period, it is planned to substantially expand the composition of the orbital groups of spacecraft of scientific and socioeconomic purposes and measurements, to carry out a number of principally new space projects, including interplanetary, to broaden the international space cooperation. The solution of these tasks will require significant development of the existing ground-based automated control complex for spacecraft of scientific and socioeconomic purposes and measurements (GBACC SC SSPM).

To meet this challenge it is possible to use either extensive or intensive approach to GBACC SSPM development. Solely extensive development, such as increasing the number of ground facilities, cannot meet the requirements for a modern spacecraft control complexes, it is imperative to search for ways of intensive development based on the introduction of advanced spacecraft and measurements control technologies.

Presently, a number of trends to improve satellite control processes that significantly affect architecture of the future GBACC SSPM is being developed and implemented. The most important ones are:

the use of consumer navigation equipment [1];

development of information-telemetric support [1];

utilization of relay technologies:

- using relay satellites (RS) in geostationary orbits [1,3];
- using RS in low orbits [2,4,5];

the use of network control technologies [3,4,5];

integration of data transmission channels [4].

A detailed examination of these trends is the subject of the research into the design of spacecraft control systems, this article will examine the implementation of these particular trends and their influence on the architecture of future GBACC SC SSPM.

1. The use of consumer navigation equipment

Ballistic and navigational spacecraft flight support includes measurement of current navigational parameters (MCNP) of the spacecraft's movement. Currently, the main BNS technology is the direct MCNP from ground stations. The advantage of the ground-based measurements are their maturity, simplicity, and reliability. The drawback is a heavy workload of ground facilities. Thus, the Resource ERS satellites requires 6-7 daily MCNP sessions with geographically dispersed control and measurement stations on 2-3 adjacent circuits, that is 50-60% of the total number of communication sessions with the spacecraft on the daily interval. The use of consumer navigational equipment (CNE) significantly reduces the number of communication sessions and reduces the requirements for the number of control facilities and their spatial layout.

Using radio-navigation field of fully-deployed GLONASS system allows globally and continuously determine the position of the spacecraft in orbit on altitudes from 200 to 2000 km, with critical errors of up to of tens of meters and using cm/s units for the components of the velocity vector, which limits the application of the CNE in the spacecraft in highly elliptical orbits, geostationary orbit, during a flight to the Moon, as well as in upper stage rockets, when delivering a spacecraft to the geostationary orbit.

The navigation support of spacecraft in orbits of over 2000 km can be performed by discrete radio navigational field. Space-time discretization of a radio navigation field is caused by determinancy of antenna patterns of navigational satellites, as well as the impossibility of radio emission into the upper hemisphere, and radio shadow due to the screening effect of the Earth.

Under these conditions, for SC on elliptical orbits with a perigee under 200 km and apogee over 2000 km, it is advisable to use a technology that combines the determination of spacecraft movement parameters based on individual of the radio navigational field during the perigee phase with the prediction of movement parameters during the apogee phase of the flight, followed by a refinement of movement parameters based on the discrete radio navigational field, that can be generated by even one navigation satellite.

For the particularly precise work, requiring knowledge of the center of mass of the consumer spacecraft with margin of error of several meters (for example, when docking, carrying out scientific experiments, etc.), the differential method of navigation can be used.

The widespread use of GLONASS / GPS navigation systems will provide significant economic benefits, since even a significant increase in the number of satellites in a constellation and increased requirements for accuracy and timeliness of orbit determination, does not require more ground-based measuring devices, reducing the number of measurement sessions of current navigation parameters and the volume of measurement information in spacecraft computer.

Thus, the analysis of technical system features of navigational and ballistic support of advanced satellites allows us to conclude that in the near-term outlook ground-based automated control complexes will employ a combined NBS technology that will utilize both GLONASS / GPS signals and ground-based MCNP. The use of ground-based facilities, primarily, will be necessary for the spacecraft, flying on highly ecliptic, geostationary, lunar and interplanetary orbits, as well as for upper-stage rockets.

To further expand the scope of consumer navigation equipment on board of a spacecraft it is necessary to further develop the theory of navigation definitions in discontinuous field.

2. Improving information-telemetric support

Currently, large amounts of telemetry data are one of the main factors that increase the load of ground-based radio equipment, communications equipment, and ultimately MSRS. The introduction of advanced methods of information-telemetric support of spacecraft will increase the capacity of ground-based automated control complexes for spacecraft of scientific and socioeconomic purposes and measurements.

ITS technology is currently being implemented by using the primary mode of telemetry, involving the transfer of complete telemetry data streams from spacecraft through a standalone radio channel or CIS channel, as well as the use of generic control information transmitted through the CIS channel.

The basis for further development is the introduction of new antenna systems, compression of telemetric data, processing of telemetric data on board of a SC, the use of packet telemetry.

The basis of modern technologies to reduce the redundancy of transmitted messages is comprised of algorithms of syntactic and semantic compression of telemetric data.

Syntax compression involves increasing the amount of information carried by each transmitted character of telemetric data. It is based on unconventional presentation of data telemetry as their images that reduce the structural redundancy of telemetric data. Semantic compression reduces the temporal redundancy of transmitted data. It is based on aperture methods of reducing of telemetric data redundancy, associated with the establishment of thresholds, which need to be reached for data to be considered a significant result of telemetry, and, therefore, transmitted.

Combining different compression methods in conjunction with computing and algorithms for reverse recovery of telemetry, which are represented in finite fields, it is one of the approaches to the construction and improvement of advanced on-board telemetry systems.

Compression of telemetric data allows to reduce the amount of information transmitted by 5-10 times and to increase by 5-6 dB the equivalent power of the radio link. This means that the required reception reliability will be provided with a decrease in the requirements to the effective surface of the antenna systems by 8-10 times, which will create conditions for the use of smaller diameter antenna systems. The ability to use small-sized antenna systems creates conditions for reliable reception of telemetric data when deploying small radio systems on mobile vehicles, thereby increasing the performance, mobility and adaptability of ground telemetry complexes. In addition, the compression of telemetric data allows relaying of reduced telemetric messages using MKSR, conventional satellite relays, as well as transmission of the results of telemetry through standard telephone channels.

Creating variously based mobile means for receiving and processing telemetric data will significantly improve the efficiency of deployment of information support facilities for carrier vehicle launches, which is especially important for the information support of launches from Vostochny Cosmodrome.

Thus, it is now possible to significantly improve the information-telemetric support, as well as the characteristics of ground-based facilities for receiving and processing of telemetric data. It is necessary to harmonize approaches to the implementation of promising methods and deadlines for their implementation.

3. Utilization of relay technologies for SC control

Ground-based control and measurement stations (CMS), that provide navigation-ballistic, information-telemetric and command support solutions must be separated by latitude and longitude to ensure the maximum possible continuous zone of radio visibility over the territory of Russian Federation for all possible inclinations of spacecraft orbits and rocket flight paths.

Expanding the range of spacecraft orbit inclinations, transfer of a significant proportion of starts to Vostochny Cosmodrome potentially requires the creation of a new ground-based tracking stations, i.e. extension of GBACC SC SSPM infrastructure. A possible solution to this problem, the creation of variously based mobile measuring stations, has been mentioned above.

Another solution is creation of a spacecraft control and information gathering network using data relay satellites. This communication technology will make the processes of spacecraft control and collection of measurements from carrier vehicles independent on the zones of radio visibility of ground-based facilities; the technological cycles of spacecraft control, collection of measurements from carrier vehicles will become adaptive.

The inclusion of RS into the spacecraft control loop has a number of technical problems that determine the modes of their application. The following problems are determinative:

establishment and maintaining of the communication channel(s) between the SC and the RS;

establishment and maintaining of the communication channel(s) between the SC and the ground-based station;

establishment and maintaining of the communication channel(s) between the RS (if intersatellite link is in use);

data flow control in communication channels.

3.1 Application of relay spacecraft control technologies using relay satellites in geostationary orbits

It is sufficient to use only 3 RS spaced by approximately 120 degrees in a geostationary orbit for global coverage. A typical RS has antennas for receiving both low-data-rate signals from SC in multistational access mode (MAM), and high-data-rate signals in individual access mode (IAM).

In MAM mode uses antennas with circular pattern, providing virtually instantaneous link establishment between RS and SC. Low data rate of this link makes it suitable only for short message exchange, in particular, for messages such as “Contact the GCC”.

The IA mode uses a spot beam antenna, which makes it necessary to calculate target designations for antennas, to take into account their rotational capacity, the process of mutual targeting of antennas, link establishment, antenna rotation during the communication session, termination of a data exchange session. Additional link establishment and termination activities reduce the effective data exchange time. The number of IA channels is limited, so it is necessary to individually allocate the time of communication sessions in this mode.

A major shortcoming of employment of RS in a geostationary orbit is the necessity of high-energy radio systems and a significant latency in data transmission channels.

These problems while establishing a physical data link make it imperative to further develop data transfer procedures of intersatellite links: employment of typed messages; combination of command and software, telemetric, and target data streams; the use of batch processing of messages; the use of data compression techniques and other ways to improve the exchange rate.

So far only a few countries have found the solutions to the challenges of designing such systems (see Table 1).

The USA has the most experience; their TDRSS system has been in operation since 1983. Presently, the third generation of RS is developed and deployed. The system has a global coverage, ability to work in different frequency bands (except the optical), communication devices of SC and carrier rockets are perfected. CP are able to communicate with each other, thus ensuring the communication of client SC with ground stations in near-real-time mode.

China has developed and deployed a national “Tianlian-1” system, it was tested in the orbit. The improved “Tianlian 2” system is under development.

Russian Federation has deployed the “Luch-5” RS system, but to date no spacecraft (except for ISS ROS), rocket carriers or boosters are equipped with the required client equipment.

In 2016 European space community plans to deploy its EDRS relay system [3], originally for communication with the satellites over Europe, the Mediterranean and North Africa. By 2019 it is planned to launch two additional satellites to provide global coverage. The onboard relay equipment, including laser-based, has passed testing in space.

Table 1 Status of development of relaying systems based on RS in GSO

Name (country)	Start of use	Number of SR (coverage)	Client SC altitude range	SR-SR Link	SC-RS range	No. of channels in IA mode	No. of channels in MA mode		Development stage
							RS to SC	SC to RS	
TDRSS (USA) (Third generation)	1983	3 (global)	Up to 10000 km	+	S, Ku, Ka	2 - S 2 - Ku or 2 - Ka	2 to 300 Kbit/s	5 to 1.5 Mbit/sec	In use
“Tianlian” (China)	2008	3 (global)	n/a*	-	n/a	n/a	n/a	n/a	In use
“Tianlian-2” (China)	n/a	3 expected (global)	n/a	n/a	n/a	n/a	n/a	n/a	In development
«Luch-5» (Russia)	2016	3 (global)	Up to 2000 km	-	S, Ku	S- up to 5 Mbit/s Ku - up to 150 Mbit/s	4 channels up to 256 kbit/s		deployed
“Luch-5M” (Russia)	2021	1	Up to 2000 km	-	S, Ku, Ka Laser communication	same 600 Mbit/s	same		In development
EDRS (ESA)	2016-2017 2019-2020	2(Europe) 4(global)		-	-S, Ku, Ka Laser communication	n/a 1.8 Gbit/s	n/a	n/a	Demonstration tests in space

*n/a - no data available

Fig. 1 Orbital formation of the O3b system: 16 satellites in a circular equatorial orbit with H = 8000 km.

3.2 Application of relay spacecraft control technologies using relay satellites in low Earth orbits

Currently, work is underway to create LEO multi-satellite communication systems and broadband information exchange. By design, these systems must work with low-energy handset type terminals, or a small-sized mobile and stationary transceivers, providing them with global continuous communication with any area of the Earth. To ensure global coverage, some systems establish the intersatellite relay links between the spacecraft of the system within the orbital plane and spacecraft in different planes.

Examples of such systems are the “Iridium”, “INMARSAT”, “Globalstar”, “Orbcom”, “Messenger”, etc.

Such systems potentially make it possible to maintain control of low-altitude spacecraft without the use of powerful transceiver equipment and ground-based CIS, which is especially important for small spacecraft. Lower, compared to the CP in the GSO, orbital altitude can significantly reduce the data transmission latency and simplify the solution of technical problems associated with such latency. The information exchange protocols, used in these systems, are available to a wide range of developers and consumers.

In 2005 JSC “Russian Space Systems” developed and launched TNS-type spacecraft for testing small spacecraft control using low-orbit communication systems. The results were satisfactory, new nanosatellites are being prepared for launch [2].

Google is presently the pacesetter in the field of low orbit relay systems. It currently deploys the O3b orbital system [5], consisting of 16 SC in the equatorial orbit at an altitude of 8000 km (Fig. 1). The O3b constellation will provide broadband connectivity within the 45/-45 degrees south and north latitude. One of the anticipated areas of application is providing small satellites with communication with ground stations. A higher the orbit, compared to the aforementioned, allows to extend the variety of spacecraft orbits, that can take advantage of this virtual SDTS to communicate with the MCC and consumers, similarly to the TNS-type SC.

Creation of low-orbit satellite networks for SC control and target data transfer is an important development direction for small spacecraft control systems, which has a great potential. It is necessary to accelerate the perfection of low-altitude communication systems for command of small-sized spacecraft. So far few real steps were made in this direction.

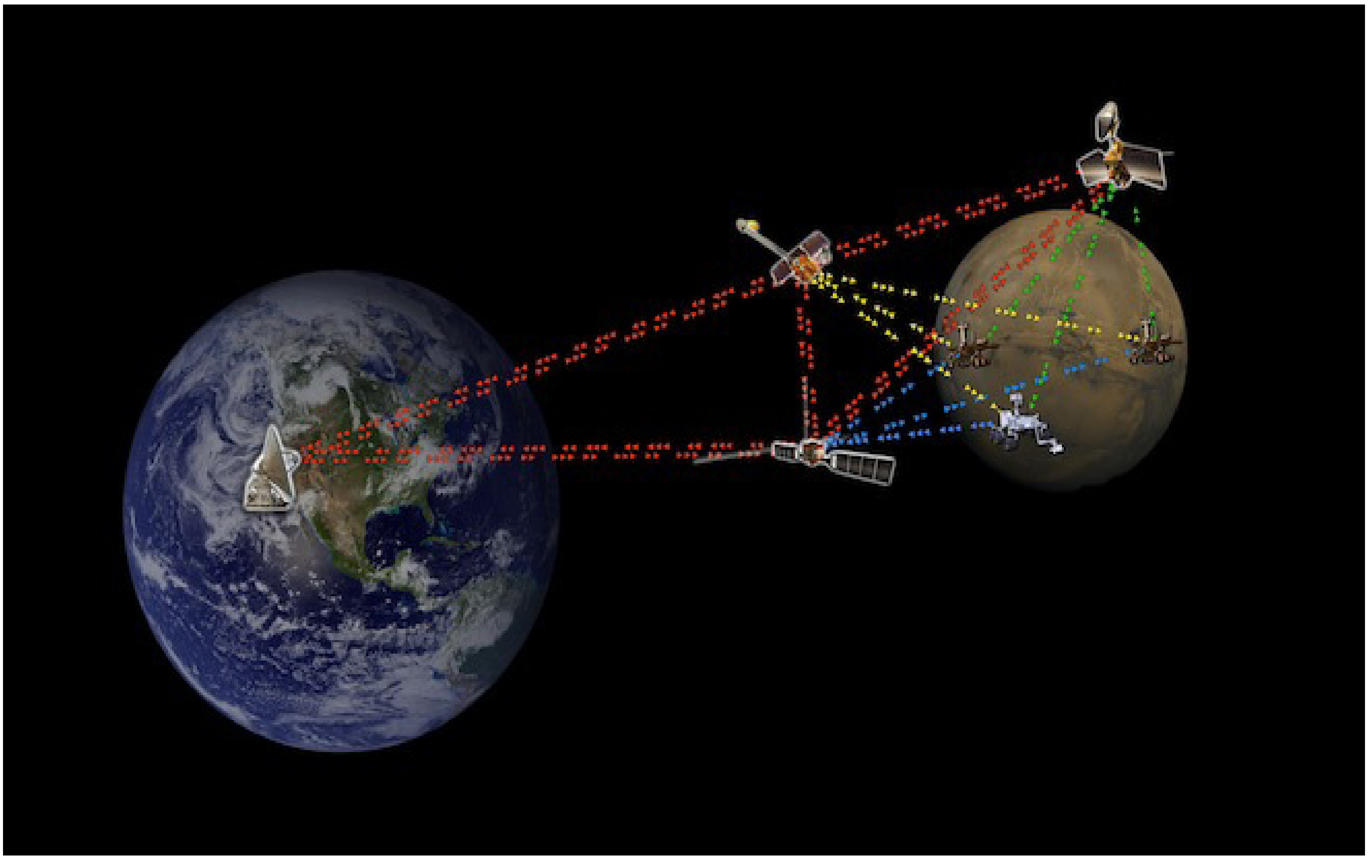


Fig. 2 Network design variant for information exchange with the lunar modules

4. Network spacecraft control technologies

The introduction of the relay spacecraft control and target information (TI) transmission technology leads to the following conclusion: the spacecraft control and TI transmission system becomes an analogue of a distributed computing network in which it is advisable to use standardized communication protocols, rather than create dedicated hardware and software. Consequently, the next important direction of spacecraft control technology development is creation and implementation of common standards (formats, protocols) for information exchange between on-board and ground-based equipment, regardless of the type of information transmitted, with the transition, in the long term, to packet-switching [4]. The introduction of the network protocols would eliminate the need of specialized ground-based facilities for processing and transmitting of various types of information, significantly reducing the number of specialized facilities, while increasing the employment of multipurpose facilities. The entire set of tools of the integrated GBACC SC SSPM and on-board SC control systems will be unified into a single computer network conforming to the multi-tiered architecture of open systems.

The results of THC-type spacecraft testing showed that the technical implementation of network protocols is viable and that it makes it possible to create fundamentally new technologies for small spacecraft control and development of control complexes.

Another trend in small spacecraft development is the creation of multi-satellite (cluster) systems [4], in which, to put it into computer networks terms, one spacecraft acts as a server and communicates with the MCC, while the others become clients of the local network. In its essence, it is a counterpart of the widespread local area networks. The application of this idea to the data exchange and spacecraft control systems is promising, at least in terms of increasing the efficiency of information exchange between the interacting SC, reducing the number of ground-based control stations and the number of required communication sessions. This is ensured by the ability, if the network is sufficiently developed, for any ground station to establish contact with any satellite, regardless of its location in space, without waiting for its

Table 2 Suggested trends of implementation of the primary and interrelated advanced spacecraft control and information exchange technologies.

Name of the advanced technology of spacecraft control	Suggested trends of implementation of the advanced spacecraft control technologies
Using the on-board consumer navigation equipment GNS GLONASS / GPS for MCNP for the transmission of information as a part of TMD. Integrating circuits of trajectory and telemetry control.	On-board control complexes for future space and rocket vehicles, boosters, complexes of measurement, information collecting and processing facilities for launch vehicles, ground-based spacecraft control complexes.
Utilization of relay technologies for SC control and communication	Relay systems, on-board and ground-based control complexes for future rocket and space equipment, future UCMS, consumer relay equipment.
Utilization of the on-board systems for control, diagnostics and automated recovery of SC equipment.	On-board control complexes for future rocket and space equipment.
Development of a system for forming and sending the “Call Earth” signals using relay systems.	On-board and ground-based control complexes for future SC.
Implementation of coordinate-time control in place of program-time control.	On-board control complexes for future spacecraft, boosters, launch vehicles, MCC.
Using the target radio link to deliver command-program information on board of the spacecraft and TMD transmission.	On-board control complexes for future spacecraft, Earth remote sensing data receiving stations, satellite communications stations, MCRS information receiving and transmitting stations.
Synchronization, phasing and correction of the on-board time scale based on the signals from GNS Glonass / GPS using on-board GCE.	On-board control complexes for future rocket and space equipment.
Intellectualization of information processing and decision-making on control levels of automatic SC control systems.	On-board control complexes for future SC, hardware-software complexes of MCC.
Streamlined spacecraft control and measurement technologies.	Ground-based control complexes, boosters, complexes of measurement, information collecting and processing facilities.
Batch telemetry, TMD compression.	Future onboard relay systems, onboard CMS, future GCMS, CMS, CSC.
Automated trajectory measurements by telemetry signals using the GCMS.	Future NPRS. On-board control complexes for rocket and space equipment.
Using existing low-altitude satellite space communications systems (“Inmarsat”, “Globalstar”, “Orbcom”, “O3b” and others.) and VHF stations for data exchange with rocket and space equipment.	On-board and ground-based control complexes for small SC.
The use of mobile command and measuring facilities to the increase the coverage and control and management efficiency.	Future mobile measuring and command stations.
Application of navigational pseudolites, located on the Earth’s surface to improve the accuracy and efficiency of the spacecraft control in GSO and HEO using GNS GLONASS / GPS.	GLONASS, On-board control complexes for future SC.
Creating a local radio navigation field for navigational satellites in high orbits	GLONASS, On-board control complexes for future SC.
Using the very long baseline interferometry for NBO of interplanetary spacecraft.	Ground-based control complex for ISC.
The use of network technologies of information exchange.	Service communication and data transmission system, the GCC and the SCU of future satellites.

entry into the zone of the immediate radio visibility, including using relay channels, i.e. using the asynchronous mode of SC control and data transfer (Fig. 2).

Analysis of the problems of space communication has led to the conclusion about the need to develop fundamentally new protocols [5]. Interplanetary networking protocols are called Bundle protocols. The biggest difference between the Bundle protocols and TCP/IP is that the transmitted information packets are not lost if they fail to reach the destination; they are accumulated and stored in special nodes until it is possible to resume the transmission.

DeepSpaceNetwork, developed by NASA (USA), which is used to communicate with the spacecraft beyond the Earth, already supports Bundle protocols. The International Space Station also has a number of nodes to support these protocols and is in fact already a part of the interplanetary Internet.

Two Martian satellites, MarsReconnaissanceOrbiter and MarsOdyssey, support a prototype version of the software needed to build such networks. The two Mars rover vehicles, Opportunity and Curiosity, also use these protocols.

Practical implementation of networking information exchange methods of spacecraft control becomes a feasible task, especially for the spacecraft that will be controlled through relay satellites. We propose to start working on this task after 2020.

5 Integration of data transmission channels

Challenges in the aforementioned specific areas encourage setting high level tasks as well: if the data is transferred through network protocols in packet data exchange mode, what does prevent the unification of individual radio channels (CPI, TMD, MCNP, TI) into a single channel?

Therefore, one of the promising development directions of GBACC SC SSPM, considering the need to ensure the sufficient promptness of spacecraft control, consists in a significant increase in the data rate of all communication channels, including CIS channels. However, increasing the volume of transmitted information makes it necessary to expand the radio bandwidth, which, as a consequence, leads to a deterioration in noise resistance [1]. A possible solution is implementing the batch processing and routing of data streams mentioned above. Table 2 contains generalized suggestions for the implementation of the primary and interrelated advanced spacecraft control and information exchange technologies.

Conclusion

The analysis of SC control technologies has shown that in the development of GBACC SC SSPM in 2015-2016 the greatest efforts should be directed to the modernization of existing facilities, introduction of new facilities, implementation of the existing SC control, data transmission and processing technologies.

As a result of implementation of these technologies an efficient GBACC SC SSPM should be created, which should utilize modern standardized control and measurement facilities and meet world-class standards.

References

1. Kukushkin S.S., Rudakov V.B., Makarov M.I. Analiz problem sozdaniyaperspektivnykhtekhnologiyizmereniyiupravleniyakosmicheskimisredstvami [A problem analysis of development of the perspective measurements and control technologiesof spacecraft].*Raketno-kosmicheskaya tekhnika. Informatsionnye sistemy I tekhnologii. Nauchye trudy. V 2-kh tomakh* [Rocket and space engineering. Information systems and technologies. Proceedings in two volumes]. Vol. 1. Moscow: NII KS im. A.S. Maksimova, 2012, pp. 86-108. (in Russian)

2. Limanskaya T.V., Sergeev A.S. Odnopunktnoe upravlenie gruppirovkoy malorazmernykh KA [One-station control of the small-sized spacecraft constellation] *Uspekhi sovremennoy radioelektroniki* [Modern radioelectronics achievements]. Moscow: Radiotekhnika, 2013, №1, pp. 78-82. (in Russian)

3. *European data relay system (EDRS*). Available at: http://www.esa.int/spaceinimages/2014/06/european_data_relay_system_EDRS. (in Russian)

4.*Nizkoorbital'naya kosmicheskaya sistema personal'noys putnikovoy svyazi i peredachi dannykh* [A low-orbit space system for personal satellite communication and data transmission] Pod red. General'nogo konstruktora mnogofunktsional'noy kosmicheskoy sistemy personal'noy sputnikovoy svyazi I peredachi dannykh, prezidenta OAO "Sputnikovaya sistema "Gonets"A.I. Gal'kevicha [Edited by A.I. Gal'kevich, designer general of the multifunctional space system of the personal satellite communication and data transmission, president of "GONETS Satellite Company public corporation"]. Tambov: OOO "Izdatel'stvo Yulis", 2011, 169 p. (in Russian)

5. *Kosmicheskiy mezhplanetnyy internet* [Space interplanetary Internet]. Available at: http://www.rubroad.ru/ magazine/hardware/3389-google-pridumala-kosmicheskij-mezhplanetnyj-internet. (in Russian)

ROCKET-SPACE DEVICE ENGINEERING AND INFORMATION SYSTEMS
Vol. 3, Iss. 1, 2016, pp. 66–72

Correction of Temperature Error of Pressure Piezoelectric Sensors for Space Technology Products

V.P. Malanin[1], V.V. Kikot, P.N. Efimov

[1]*candidate of engineering science*
Joint-Stock Company "Research institute of physical measurements", Russia
e-mail: ait@pnzgu.ru, inbox@post.su, nik2@niifi.ru

Abstract. The article presents the questions of correction of temperature error of piezoelectric pressure sensors in conditions of transient temperature and thermal shock. It is noted that when the piezoelectric pressure sensor is operating in thermal shock conditions, it is important to measure the main parameter by means of the piezoelectric sensor: the pressure and piezoelectric element temperature at a single point of space and at the same time to eliminate the influence of the temperature gradient. That makes difficult to use additional sensors to measure the temperature of the actuating medium. It is proposed to use the impedance parameters of working and vibrocompensation piezoelectric elements as information source about the temperature of piezoelectric elements. The scheme of secondary transmitter of output signals of working and vibrocompensation piezoelectric elements is presented; its description and algorithm for obtaining the correction signals are described. The results showed that the change of the conversion efficiency of the pressure piezoelectric sensor of the piezoelectric elements form the PZT-83G piezoelectric material in the range of - 180 to + 200 °C is approximately 35%. Moreover, time dependence of output signals from the working and vibrocompensation piezoelectric elements of the pressure piezoelectric sensor when exposed by thermal shock of liquid nitrogen is presented. It is shown that using the proposed correction method of measurement errors from the transient temperature of the actuating medium can reduce measurement error of dynamic pressure from the thermal shock.

Keywords: pressure sensors, piezoelectric element, temperature error, membrane, equivalent circuit, impedance, conversion efficiency

Sensors are the basis of automatic control systems for space applications. Sensors perceive not only the data on the measured value, but also the impact of wide range of influencing factors that emerge during the operation in the so-called harsh environments. The sensor equipment, used for this purpose, is exposed to an extremely concentrated and complex effect of destabilizing factors, such as change of pressure, high levels of vibration and shock, acoustic noise, transient temperatures of the actuating medium, thermal shocks. The influence of the quick-changing and acoustic pressures of transient cryogenic temperatures, which affect the sensing elements of the primary transducers of piezoelectric sensors, on the metrological characteristics of the sensors is one of the most important factors that determine whether the use of these sensors is possible in each specific case.

Development and improvement of methods of temperature error correction in piezoelectric sensors are the tasks of many aviation and space technology development organizations, since a distinctive feature of the sensor equipment usage in these fields is the impact of transient temperature on the sensors. The initial conversion of the dynamic pressure under powerful and quick-changing temperature effects in ranges from -253 to +300° c causes temperature transients in sensors and, as a consequence, increased measurement errors of dynamic input non-electrical values during the transients. The transients in piezoelectric sensors under a thermal shock can last from seconds to tens of minutes, depending on the specific characteristics of the individual sensor units. Ensuring stability and accuracy of measurement of dynamic pressures during that time is a serious and urgent problem. To date, two well-known techniques of temperature sensor error correction have formed. The first technique is based on reducing the power of the effecting destabilizing factor, while the second one suggests reducing the sensitivity of the sensor's metrological characteristics to temperature.

References [1-3] describe the search for the balance between the universality of use, reliability of the developed sensors and their metrological characteristics. The pursuit of universality, for example, imposes restrictions on the metrological characteristics, resulting in complication of sensors operation because of the necessity of additional structural elements (leading piping to reduce the temperature of the actuating medium, adapters, structural units with intermediate temperature-controlled environments, through which the pressure is sensed), as well as the need to perform additional configuration of a sensor for the desired measuring range of dynamic pressures, etc. The distinction of this approach lies not in designing sensors for systems, but, on the contrary, in designing systems for sensors. If, however, the main objective is to improve the reliability and metrological characteristics of sensors, then, for example, for different dynamic pressure measuring points in a device with controlled parameters of the actuating medium it is possible to use sensors of various designs and conversion methods to achieve the best metrological characteristics and high reliability. Various challenges of designing the sensors and solutions to them have generated a pool of constructive scientific and technical designs, that are currently in use.

It is also important to take into account the mutual influence of constructive elements, with their physical and electrical parameters changing as the temperature of the actuating medium changes. The thermoelastic displacement in the body of a sensor and it's connected elements, the differences in the dimension changes due to different linear extension thermal coefficients of the materials of the sensor's elements, the tension and deformations of the membrane material, pyroelectric effects in the piezoelectric element of the sensor are just a few consequences of changes in the temperature of the actuating medium. The commonly known constructive designs for reducing the impact of temperature on the dynamic pressure measurement accuracy are:

1) increasing the mass of a sensor to compensate for the impact of the thermal shock, the downside of this solution is the increased mass of the sensor and decreased operation speed;

2) coating the sensor membrane with a layer of silicone rubber, ceramics or installation of asbestos cloth soaked in low viscosity oil in front of the membrane , which affects the sensitivity of the sensor;

3) the use of various modifications of membranes.

For example, when measuring high pressures, it is possible to employ a membrane in the form of a thin plate, that is manufactured together with the sensor casing, but such a design usually requires either a threadless connection at the sensor installation point, or an additional external installation casing in order to prevent the indirect influence of tightening on the metrological characteristics of the sensor. The sensors exposed to thermal shocks feature a membrane in the form of a thin plate, that connects to the sensor casing by welding, but welding does not provide the sufficient reliability for the

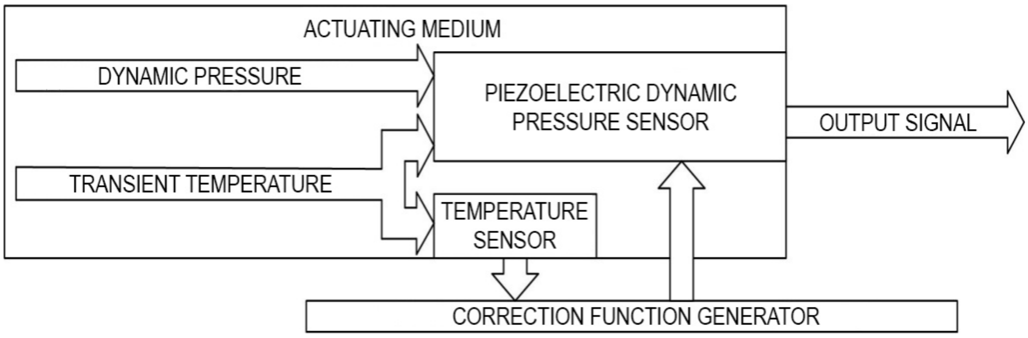


Fig. 1. Structure of a piezoelectric dynamic pressure sensor with a temperature sensor

prolonged use of the sensor in some aggressive environments. To increase the service life of the sensors when measuring high pressures at the temperatures of up to 800 °, membranes in the form of pistons are employed, but another force-transferring element of the construction reduces the sensitivity of the sensor.

One of the design options for reducing the impact of temperature on dynamic pressure sensors is the temperature control that can be implemented, for example, by introducing the means of forced cooling into the sensor design using channels in the casing of the sensor for coolant passage. Temperature control provides the best thermal stability of the piezoelectric sensors, but the small operating temperature range, relatively large dimensions and high power consumption limit the use of such sensors.

Temperature compensation is also a widely used method of circuit temperature error correction. In contrast to temperature control, when thermal compensation is employed, the generated compensatory effect is applied to the output signal of the sensor under the impact of temperature as a destabilizing factor, which results in changes to the output signal, caused by temperature changes, converging to zero. The typical structure of piezoelectric sensor for dynamic pressures with temperature compensation employed is presented on Figures 1 and 2.

When measuring dynamic pressure using a temperature measurement channel for the temperature of the actuating medium, the temperature sensor can be installed separately from the dynamic pressure sensor. When the temperature sensor is installed at a different location, the temperature differences at the temperature measurement point and the pressure measurement point as well as the different rate of temperature change of the sensing elements of the pressure sensor and the temperature sensor under thermal shock impact the accuracy of the pressure measurements. Performance of temperature sensors is determined by their response rate, which, depending on the model, varies from 0.05 to 20 seconds, whereas the duration of the transients in dynamic pressure sensors, including temperature changes of their piezoelements, as mentioned above, can reach tens of minutes. This results in the decrease in the precision of the dynamic pressure measurement, more complicated definition and implementation of the corrective function, as well as the increase in the amount of work required during the initial tune-up of the system before operation.

Application of circuit techniques for reducing the effects of temperature impacts is devoid of the weaknesses of constructive and technological methods. Development of nano- and micro-electromechanical technologies, miniaturization of sensing elements, as well as the electrical circuits of secondary transducers, make it possible to combine circuit and structural design methods for the development of sensors, allowing the creation of fundamentally new designs of sensors and control systems. This is facilitated by the emergence of new piezoelectric materials, such as langatate and langasite, which are able to withstand severe thermal shock. For example, it has become possible to locate both dynamic pressure and temperature sensors and, when necessary, the circuits of secondary transducers a in a single casing. This also improves the weight dimension characteristics of the measuring systems in which these sensors are included. An advantage of this arrangement of the temperature sensor is measuring temperature and dynamic pressure at a single point without spatial separation of the channels to measure them. The accuracy of pressure measurement in this case is indirectly affected by the differences in the physical characteristics of materials of the elements versus the temperature.

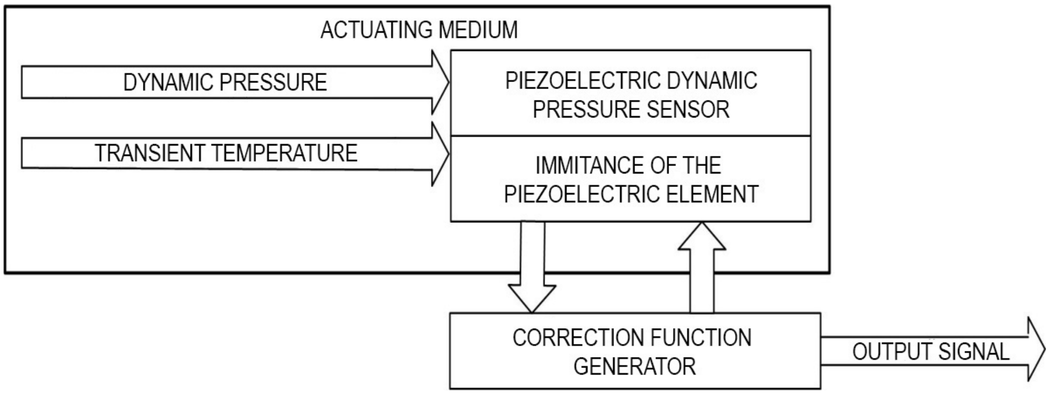


Fig. 2. Structure of a piezoelectric dynamic pressure sensor without a temperature sensor, with the immitance of the piezoelectric element used for temperature measurements

Another promising method for correcting the temperature-induced error in dynamic pressure measurement is using the parameters of piezoelectric elements for measuring temperature - the impedance analysis of piezoelectric elements. For instance, the dependence of the electrical capacity of the piezoelectric element of dynamic pressure sensor on temperature is used for generating the temperature error correction signal. Specifically, the dependence of voltage drop amplitude of the high-frequency current signal on the complex resistance of the piezoelectric element, which is, in turn, expressly dependent on temperature, can be used for temperature measurement. During the sensor's operation the instantaneous measured values indicate the temperature and the correction value for its sensitivity changed by temperature. The practical implementation of said design is the temperature error correction device developed by the authors. Its block diagram is shown in Fig. 3. This method of correction, based on impedance analysis of the piezoelectric element, has a number of advantages:

a) reduced weight dimension characteristics of the sensor due to use of a single primary measurement transducer for dynamic pressure and temperature measurement;

b) increased accuracy of measurement due to no spatial separation of the measuring points of dynamic pressure and temperature within the monitored object;

c) simplified circuitry of the secondary transducers, forming the correction signal, due to identical physical and electrical parameters of the sensor elements for measuring dynamic pressure and temperature;

g) improved metrological characteristics and performance due to reduced response time of the sensor;

d) possibility to upgrade existing designs that include piezoelectric dynamic pressure sensors.

Application of the impedance analysis method in conjunction with other circuitry and structural methods, aimed at improving the accuracy of dynamic pressure measurement and correction of temperature errors, will allow the newly designed sensors meet the ever-increasing requirements for reliability of control systems.

The main circuit signal processing method for correction of errors caused by the influence of destabilizing factors is implemented in the secondary transducer. It involves sending the data on the affecting destabilizing factor directly to the primary transducer and computation of the output correction signal using the sensors microcontroller. When piezoelectric sensors operate under a thermal shock, in order to eliminate the influence of a temperature gradient, it is important to measure the main parameter, pressure, and the temperature, affecting the piezoelectric element at a single point in space and at the same time, making it difficult to use additional sensors with spacial and temporal separation of the measuring channels for measuring the temperature of the actuating medium.

Authors consider a piezoelectric element model represented in equivalent circuit form, which consists of a power supply, which has internal resistance, and the complex resistance of the piezoelectric element. Complex resistance of a piezoelectric element is composed of the piezoelectric element leakage resistance and the capacity reactance of the piezoelectric element. For the piezoelectric element temperature measurement it is proposed to use the dependence of voltage drop amplitude of the high-frequency harmonic current of the current

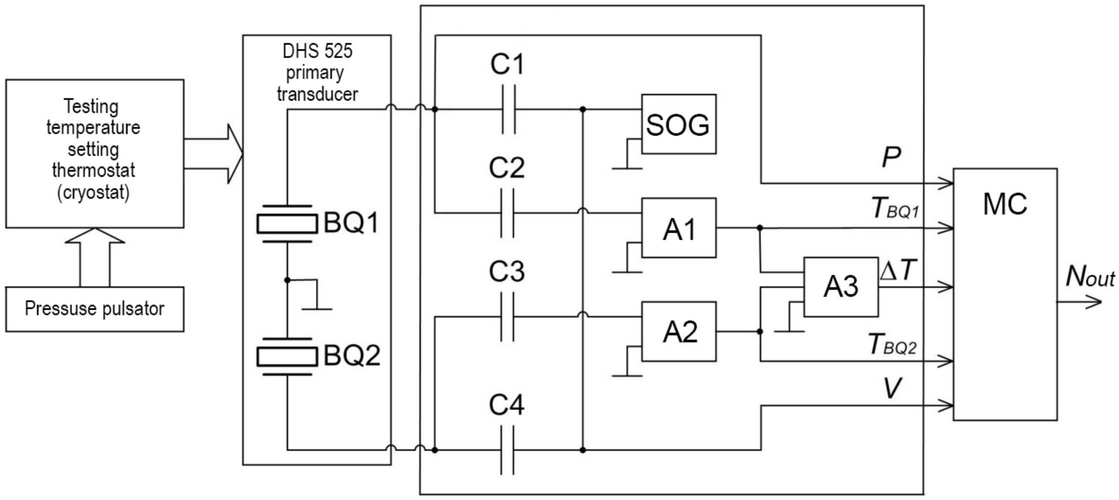


Fig. 3. Measurement diagram.

signal of the current source on the complex resistance of the piezoelectric element. An experiment to study the temperature characteristics of the actual and the compensating sections of the piezoceramic module PM 7-02V produced of material 83g-PZT (piezoelectric element) sensing element (SE) pressure sensor DHS 525 (sensor) was carried out using the measuring circuit shown in Figure 2, where *BQ1* and *BQ2* are the actual and the compensating sections of the piezoelectric element, respectively, *C1* - *C4* are the capacitors used to separate the pressure and temperature measurement channels.

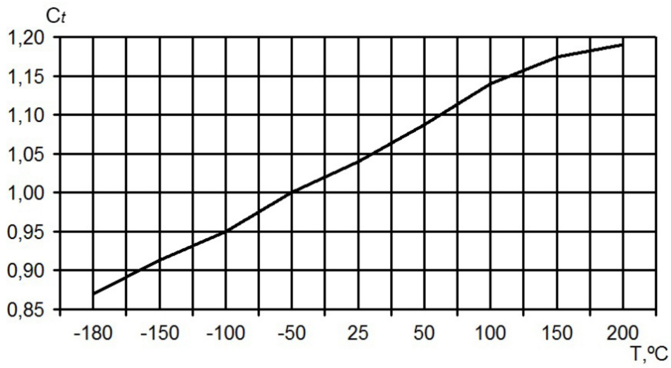


Fig. 4. The dependence of temperature effect coefficient *C*t on temperature T.

Pressure pulsator is used to set the pressure pulsations for the determination of the coefficient of the effect of temperature *K*t on the sensor's conversion efficiency *K*. The sinusoidal oscillation generator *SOG* feeds a sinusoidal currency signal at a frequency of 1 MHz and a crest voltage of 12V to the actual *BQ1* and the compensating *BQ2* sections of the piezoelectric element. Under the influence of temperature changes in sections *BQ1* and *BQ2* of the piezoelectric element of the sensor SE, which are set by a thermostat (cryostat), the amplitude of the voltage drop on the impedance of the *BQ1* and *BQ2* sections the piezoelectric element change their values. The signal of the dynamic pressure ΔP, picked up from the actual electrode section *BQ1*, in which the temperature impact *T* and vibration *V* is not compensated, is fed to the first input of the microcontroller MC. Vibration signal *V* picked up from the electrodes on the vibrocompensated section *BQ2*, in which the temperature impact *T* is not compensated, is fed to the fifth input of the microcontroller MC. Temperature signals T_{BQ1} and T_{BQ2} picked up from sections *BQ1* and *BQ2* of the piezoelectric element are amplified by the first A1 and the second A2 high-frequency amplifiers for the voltage drop across the impedance of piezoelectric elements *BQ1* and *BQ2* and fed to the second and fourth inputs of the microcontroller MC, respectively. Also, the signals T_{BQ1} and T_{BQ2} are utilized to produce with a using a high-frequency amplifier the voltage drop difference for the production of ultrasonic signal ΔT of the temperature difference between the actual *BQ1* and the vibrocompensating *BQ2* sections of the piezoelectric element when exposed to thermal shock. Signal ΔT is fed to the third input of the microcontroller MC The microcontroller MC performs the correction of the dynamic signal pressure ΔP and vibration *V*, impacted by the changes in temperature *T* of the actuating medium, using the signals of the dynamic pressure ΔP, temperature *T*, vibration *V*, temperature difference in the

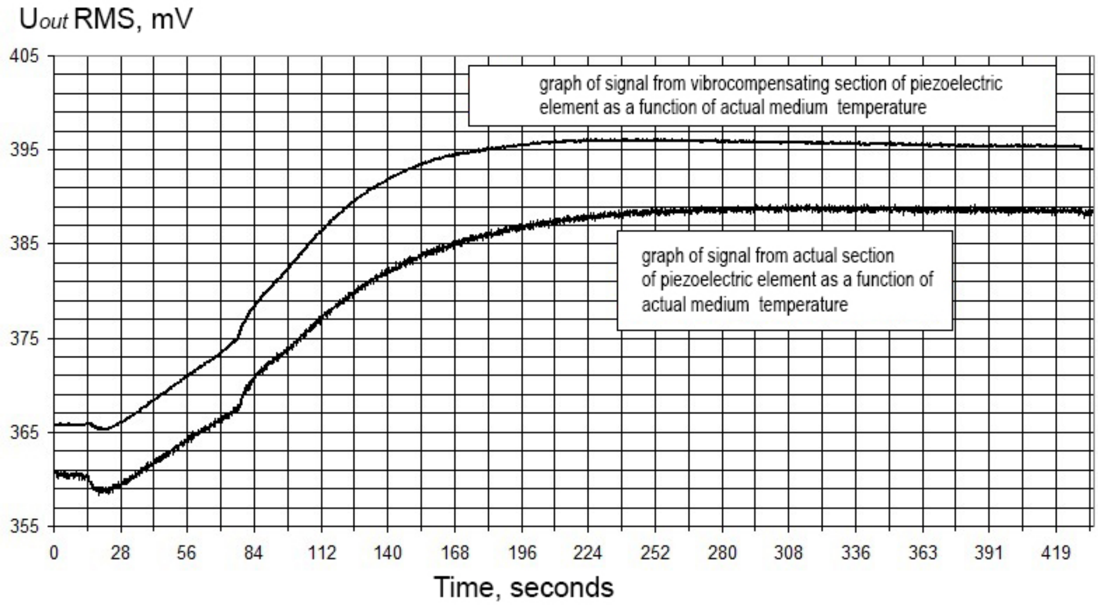


Fig. 5. The dependence of the output signals of the actual *BQ1* and compensating *BQ2* sections of the piezoelectric element on temperature;

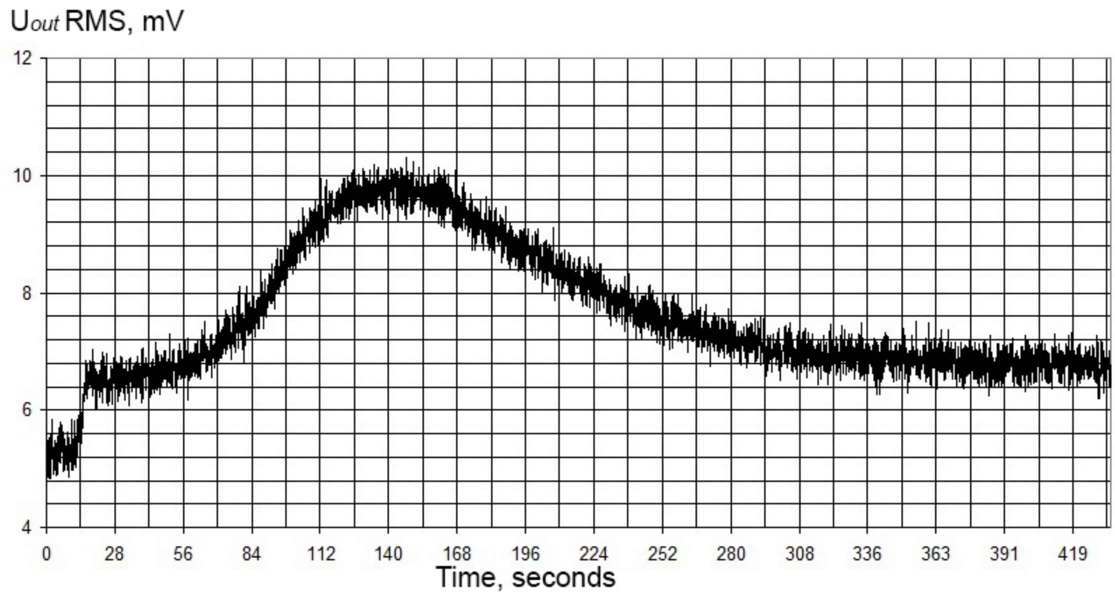


Fig. 6. The dependence of the difference in output signals from the piezoelectric element sections on temperature.

sections of the piezoelectric element ΔT. The microcontroller MC outputs a digital signal *N*out with data on the dynamic pressure Δ*P* and temperature *T* of the actuating medium as well as vibration *V*, which affects the sensor's SE. Figure 3 shows the attained dependence of coefficient C_T of the effect of the temperature of the actuating medium in the temperature range from -180 to +200 ° C for the conversion coefficient C_c of the DHS 525 sensor›s SE. Figure 4 shows that the deviations of the conversion factor of the sensor›s SE, resulting from the effect of temperature gradient, are approximately 35% throughout the range of operating temperatures.

The time dependence of the output signals of sections of the piezoelectric element SE on the temperature gradient of the actuating medium from +23 to -196°C, was also deduced and shown in Figure 5.

Also the time dependence of the difference between the output signals of the piezoelectric element sections was determined, which is shown in Figure 6.

As can be seen from Figures 4 and 5, the change rates of the output signals of the piezoelectric element sections affected by temperature gradient are different for a period of time. This happens because of the uneven heating of the sensor's SE during the temperature change and leads to a measurement error in the output signal of the sensor's SE. This error is compensated by using the measurement circuit shown in Fig. 3.

The practical implementation of the correction of temperature error in piezoelectric sensors based on the method of impedance analysis of the piezoelectric element will improve measurement accuracy of the mass-produced, as well as the newly designed piezoelectric dynamic pressure sensors, and will expand the operating temperature range.

References

1. *Proektirovanie datchikov dlya izmereniya mekhanicheskikh velichin* [Sensors designing for mechanical quantities measurement]. Moscow, Mashinostroenie, 1979, 482 p. Ed by Osadchiy E.P. (in Russian)

2. *Datchiki. Preobrazovateli. Sistemy. Katalog OAO "NIIFI"* [Sensors. Converters. Systems. The catalogue of JSC "Research institute of physical measurements"]. Penza, Penzenskaya Pravda, 2011, pp. 70-98. (in Russian)

3. Malanin V.P., Shamrakov A.L., Kikot V.V. Ustroystvo korrektsii temperaturnoy pogreshnosti p'ezoelektricheskikh datchikov davleniya [The correction device temperature error piezoelectric pressure sensors]. *Izmerenie. Monitoring. Upravle-nie. Kontrol'* [Measurement. Monitoring. Management. Control]. Penza, PGU, 2014, №3 (9), pp. 71-76. (in Russian)

ROCKET-SPACE DEVICE ENGINEERING AND INFORMATION SYSTEMS
Vol. 3, Iss. 1, 2016, pp. 73–84

The Design of the LNA based on the Domestic ECB Using CAD AWR

N.N. Petukh, Yu.S. Belousov, A.P. Garbuzenko, K.O. Daryushkin

Joint Stock Company "Russian Space Systems", Russia
e-mail: petukhnikolay@gmail.com

Abstract. The purpose of this work was to design a low noise amplifier (LNA) based on the domestic electronic component base (ECB). The LNA is in the form of microassembly. This microassembly was designed on the base of the transistor of 3П398 type (manufactured in Veliky Novgorod). The results of this research showed that the microassembly based on the domestic ECB meets the requirements of the receiver. The paper presents the design solutions used in the creation of the X-band LNA. Moreover, the results of calculations characteristics of amplification, noise temperature, matching from input and output and assessment of stability are presented. The design work was made by means of CAD (computer-aided design) "MWO AWR". As a result of this study, this LNA was compared to foreign analogues and it turned out that this product is at the same level as foreign analogues used in space technology. In addition, a method of increasing the lifetime of circuit operation of transistors is demonstrated.

Keywords: LNA, VSWR, transistor, HEMT, ECB, radiation, lifetime

Introduction. State of a problem

As operational requirements to the low noise amplifiers (LNA), manufactured according to the technology of hybrid integrated circuit (HIC) grow, the attention to the accuracy of calculation of electric characteristics and both the design of the amplifier in general and each of its elements in the frequency range from a direct current to the boundary frequency of the used transistors (f_b) considerably increases. Moreover, at rather small working bandwidths due to increase in accuracy of calculations there is a possibility of fuller use of properties of transistors and the LNA scheme to meet contradictory requirements for k_n (noise factor), k_a (amplification factor), VSWR (voltage standing wave ratio), and work stability.

Nowadays the following main requirements to the LNA parameters for the majority of communication systems of centimeter range are present:

k_a = 20–30 dB,

k_n = 0.8–1.2 dB,

VSWR = 1.2–1.4.

In general, monolithic microcircuits (MMC) and modules of many foreign firms meet these requirements. The examples of their characteristics are given in Table 1 and in Fig. 1.

Table 1. Low noise VHF-microcircuits of foreign production

Name	Characteristics			
	Gain, dB	k_n, dB	VSWR	R1, dBm
HMC753LP4E Hittite	14	2	2.7	15
HMC903 Hittite	19	1.6	2.3	16
CHA3666-QAG UMS	21	1.8	2	16
CGY2120XUH/C1 Ommic	13.2	0.5	3	12
AMF-5F-04000800-07-10P Miteq	50	0.7	2	10

The unification purposes are performed in monolithic devices owing to simplification of the structure of the matching circuits in the wide range of frequencies. The parameters of the given products of UMS and Miteq companies are unique in their own way; however, their working frequency range is slightly lower than necessary for the device under development.

Apart from political (imposing of sanctions) and the financial (high price) reasons influencing a possibility to use foreign MMC, there are also purely technical reasons according to which their application is not considered to be an optimal solution.

Certain technological aspects of installation of the microcircuits without output UMS and Hittite, in particular, soldering control and counteraction to diffusion processes in low capacitance gaps of operation in the conditions of space production are considered to be for a long time to be technical reasons. Big dispersion power (about 3 W) for very small sizes of the case in case of using MMC Miteq and micron thickness of the transmission line inside MMC are a problem, because it bears potential unreliability in the long term. The requirement for providing the acceptable loadings out of the required working system bandwidth, but inside the band pass of a chip, is also critical.

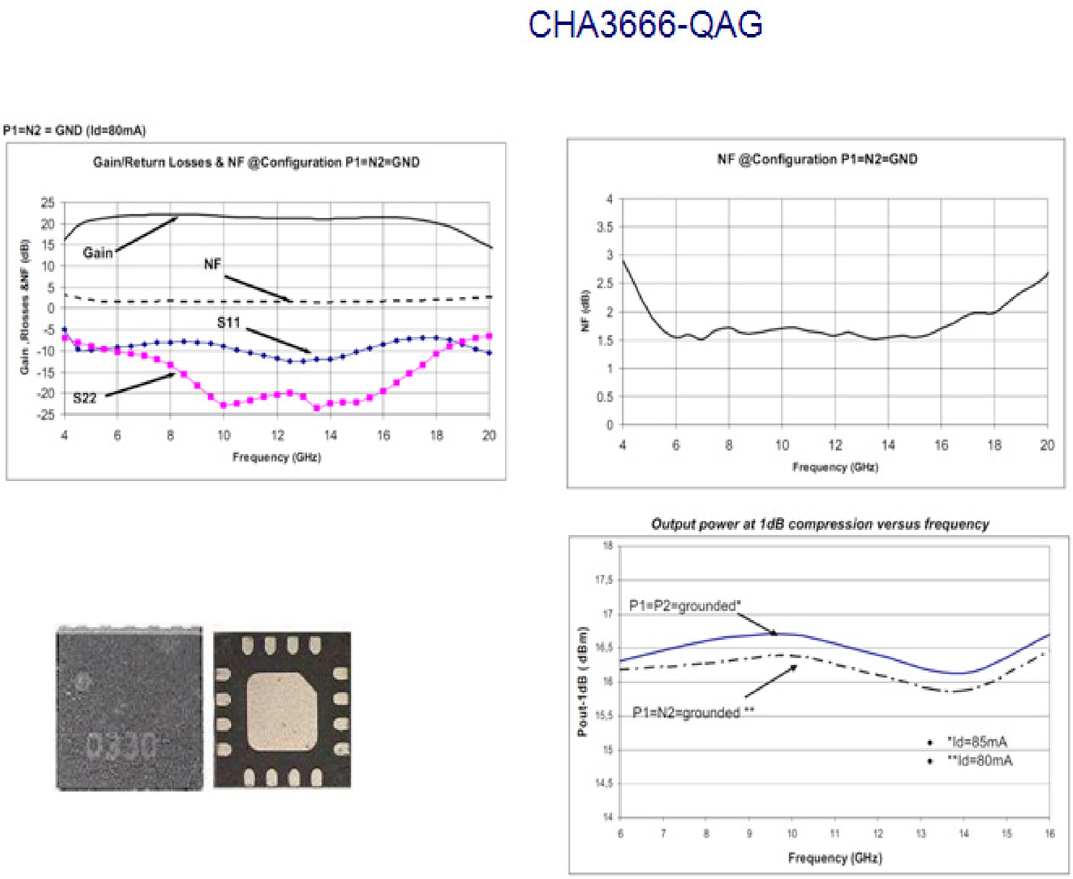
With collapse of the USSR, the production of the Russian electronic component base (ECB) has been almost turned down.

During the last 20 years, the ECB (from 80 to 90%) used onboard was delivered from abroad. Often the components of the Industrial type, after carrying out the corresponding tests, were used for cost reduction. In especially responsible cases, the components of the Space type, acquired at very high price, were applied. Due to the last geopolitical risks, such approach becomes less justified.

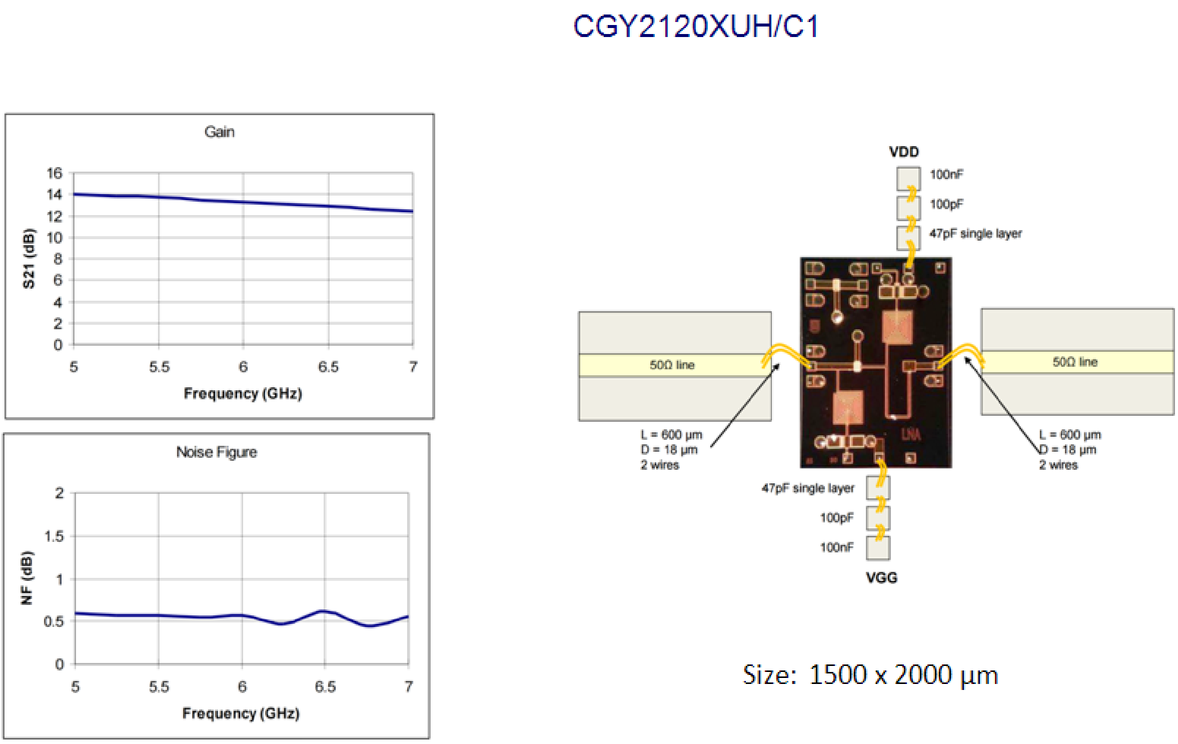
A calculation method of the LNA using domestic ECB

In this work one of the key elements of a satellite communication system – the unit of the X-band LNA built entirely on the domestic modern ECB and meeting the parameters of the world requirements – is considered.

As a rule, the range of the used frequencies in space communication lines does not exceed 10% of the transistor working bandwidth. Thereof, potential properties of the transistor can be implemented in preferable way. This can be observed in the form of the tendency of the nowadays in foreign developments when LNA designing for the solution of priority tasks.



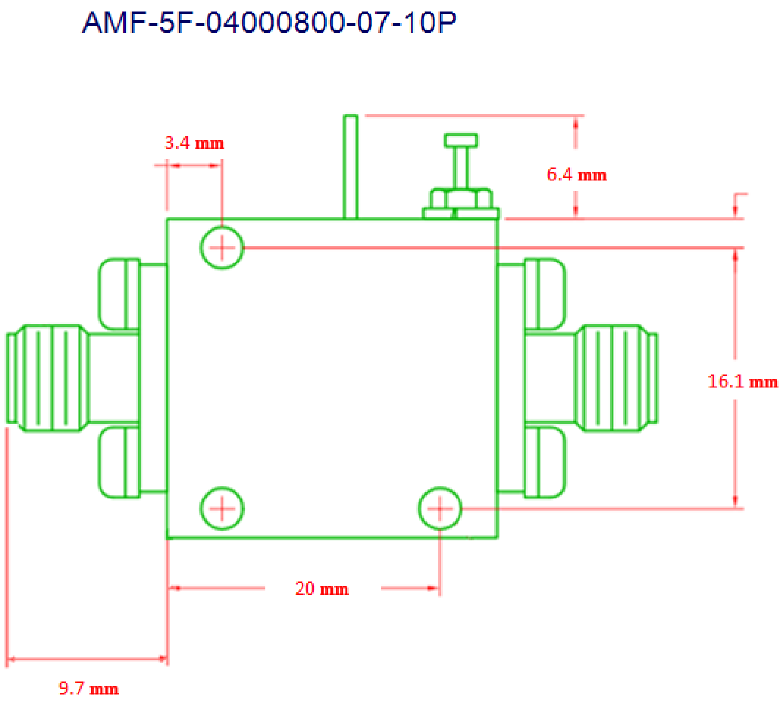
a



b

Fig 1. Parameters detailing of low noise MMC of foreign production: *(a)* copulate amplifier of UMS company, *(b)* amplifier made in crystal of Ommic company.

Description:	Amplifier
Specifications at	23 °C:
Frequency:	4 to 8 GHz
Gain:	50 dB min.
Gain Flatness:	1.5 dB+/- max.
Noise Figure:	0.7 dB max.
Noise Temperature:	50.7 K max.
VSWR In:	2:1 max.
VSWR Out:	2:1 max.
P1dB Out:	10 dBm min.
Voltage:	15 V nom.
Current:	200 mA nom.
Outline Drawing:	131581-5
Operating Temp:	-40 to 75 °C



c

Fig 1. *(c)* amplifier in assembly of Miteq company.

The choice of the transistor for implementation of the long-term project does not give a wide field for activity.

This work describes the possibility of the HIC LNA development using HEMT transistors of domestic production (Planet Argall, Veliky Novgorod) and application of the dialogue mode with CAD MWO AWR. Types and the main parameters of transistors are given in Table 2 [1].

The following provisions are offered to achieve the acceptable result:

1. In the ranges of frequencies that are 2-3 times smaller than the boundary frequency of the used transistors in HIC on VHF there are difficulties with autoshift at frequencies smaller than f_b due to parasitic resonances of a parallel type in blocking capacitors. This leads to instability of the amplifier and unacceptable unevenness of K_n and K_a characteristics, if the resonance is near to a working band. For example, the K10-71 capacitor with capacity C = 5.1 pF, dimensions 1.5 x 1.5 x 0.2 mm has the reactance presented in Fig. 2.

Using such capacitor in an autoshift chain in the two-stage LNA of the X-band leads to emergence of instability and sharp increase in K_n (see Figs. 3a, b).

Therefore, an autoshift is applied, as a rule, only in LNA working at those frequencies where it is possible to choose the capacitor for grounding the source of the transistor with the first consecutive resonance getting to a working frequency band, and the parallel resonance has to be outside f_b.

2. The transistor case considerably transforms its small-signal parameters and leads to decrease in achievable characteristics. If measurement of

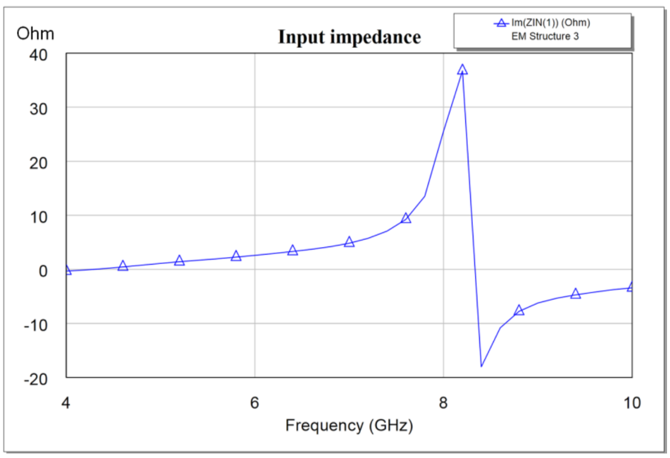


Fig. 2. An example of the spurious resonance of the parallel type in the blocking capacitor

Table 2. Low noise VHF transistors of dual application

Name of the item	Δf_p, GHz	f_{meas}, GHz	Electrical parameters values (T = 25 ± 10 °C)					Case
			$K_{Ш\ min}$, dB (max)	Kap_{opt}, dB (min)	S, mA/B (min)	P_{output}, mW (min)	P (dispersion), мВт	
3П 398А-2	4–18	8	0.4 (type)	K_{yPmax} 12.9	60	–	50	023 
3П 398Б-2	4–18	12	0.45 (type)	K_{yPmax} 12.9	30	–	50	
3П 373А-2	1–8	4	0.4	11.5	30	–	100	
3П 373Б-2			0.5	11				
3П 373В-2			0.6	10				
3П 374А-2	4–18	12	0.85	9	15	–	35	
3П 374Б-2			1	10				
3П 374В-2			1.2	8.5				
3П 397А-2	0.1–6	0.1–6	0.3	16	30	30	200	022 
3П 398В-2	12–25	18	0.95 (type)	K_{yPmax} 11.3	24		50	
3П 385А-2	12–25	18	0.8	9.5	15	–	35	
3П 385Б-2			1	10				
3П 385В-2			1.2	8.5				
3П 618А-2	0.5–4	0.5–1	0.3	18	60	500	1000	
3П 618Б-2		2	0.5	15		250	500	
3П 386А-2	18–30	25	1.05	8	10	5	30	010 
3П 386Б-2			1.25	7.5				
3П 386В-2			1.5	7				
3П 398Г-2	25–35	30	0.8 (type)	K_{yPmax} 9.3	20	–	30	
3П 618В-2	1–10	4	0.7	12	60	150	300	
		8	1.5	6		100		
3П 389А-2	25–40	37	2.5	6	5	–	20	
3П 389А-5			2	6.5				

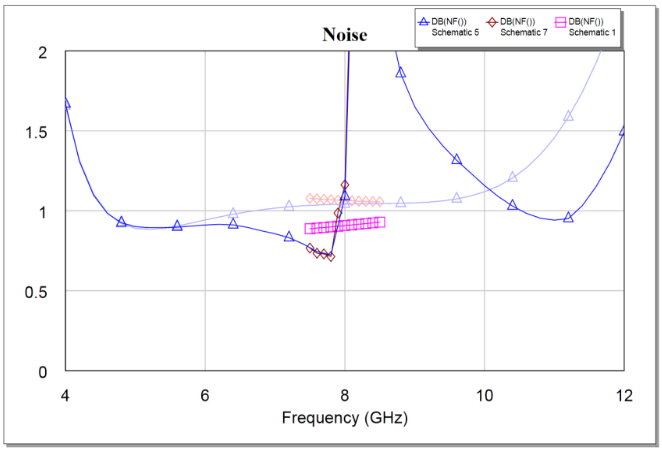


Fig. 3a. Noise growth of the amplifier due to reactance of the blocking capacitor

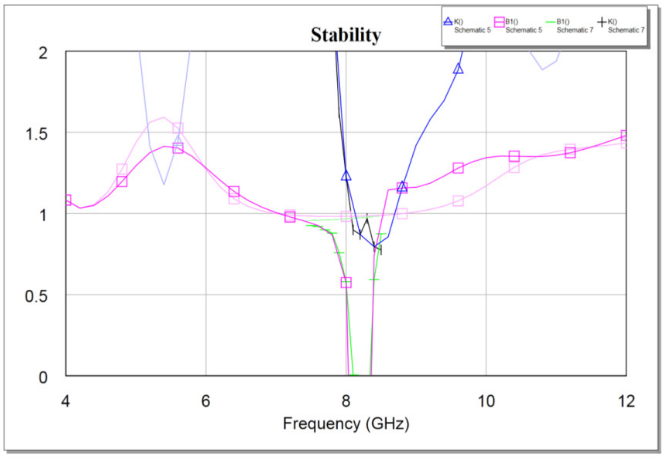


Fig. 3b. Emerging an amplifier instability due to reactance of the blocking capacitor

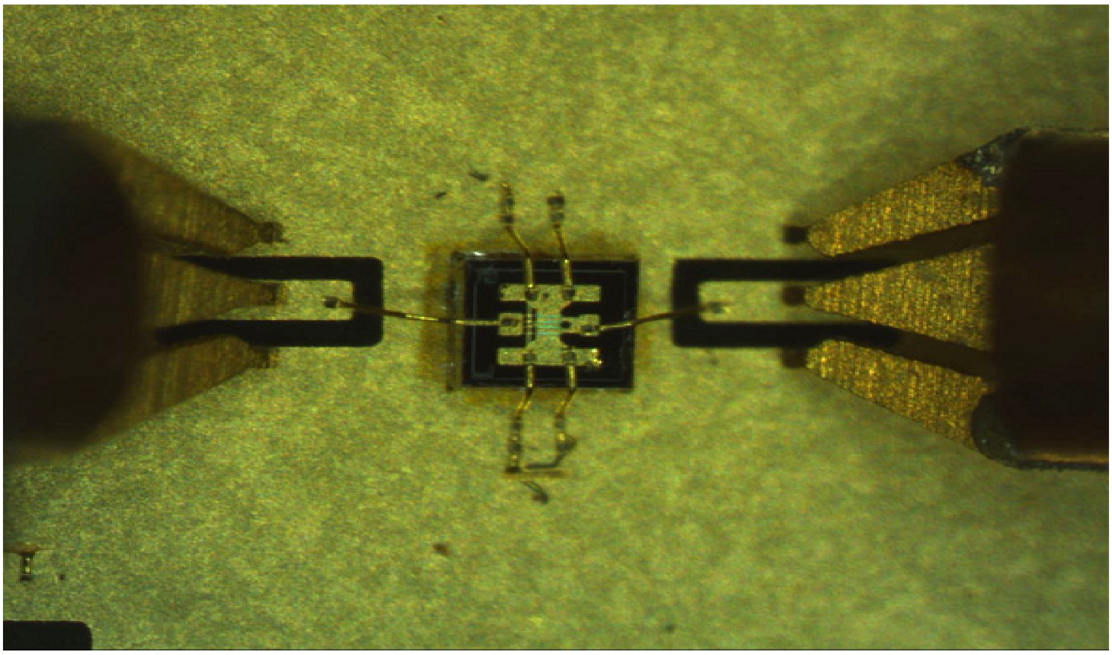


Fig. 4. 3П398Б-5 crystal parameters change by the probe

S-parameters of a crystal is made at a very high modern level (by means of probe stations, see Figs. 4 and [2]), so the measurement of the case parameters is carried out by the simplified technique that can lead to calculation errors at high frequencies.

Fig. 5 shows the characteristics of the two-stage LNA based on cased and uncased 3П398Б-2 and 3П398Б-5 transistors

In the uncased variant gain is 3 dB more; K_n is 0.2 dB less; VSWR is considerably lower. Difference, that is even more essential, is in the three-stage LNA. The decision to use uncased transistors is based on these reasons.

3. Their noise matrixes for analytical calculation of K_n are not given in reference data on parameters of domestic transistors. As a rule, the size of $k_{n\ min}$ measured in the coordination mode at one frequency is given. Therefore, to calculate the characteristics of the LNA it is expedient to use a small-signal HEMT model presented in CAD as FETN. Initial values of the FETN parameters are very close to parameters of real HEMT of C-, X-, and Ku-bands. Use of the installed programs of optimization allows one to provide coincidence of these parameters with the set accuracy. At the same time, it is necessary to exclude precisely known, for example, Lg, Ld, Ls, etc. from the varied FETN parameters. In Fig. 6, the measured

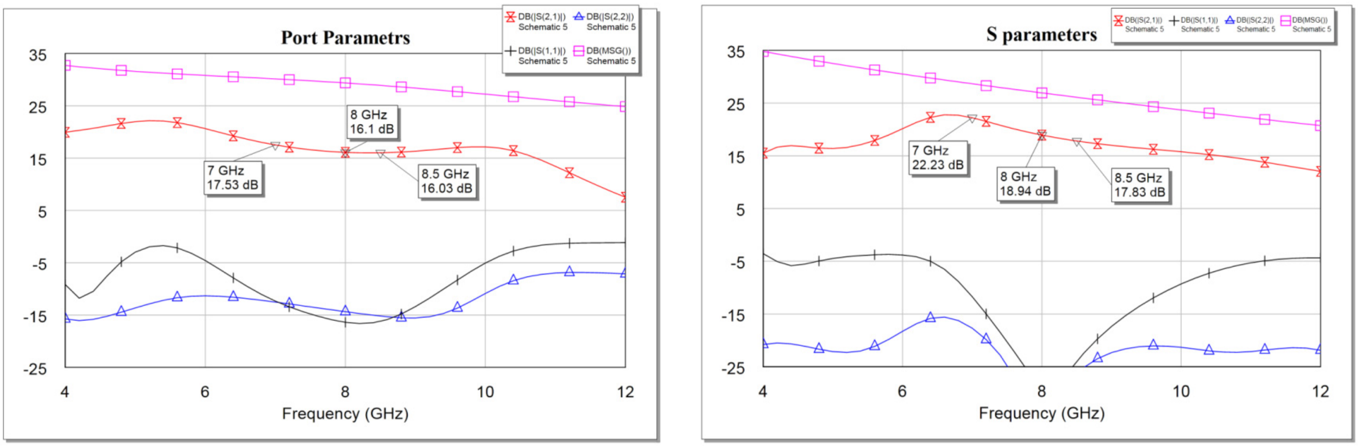


Fig. 5. Comparison of the optimal parameters of two-stage amplifiers based on cased transistors *(a)* and on crystals *(b)*

S-parameters of the transistor 3P398A-5 and the optimized FETN S-parameters are presented. Extraction of the T_d and T_g noise parameters was made in a 50-Ohm path manually in the FETN mode coordination.

4. Input and output chains of each transistor were calculated on transistor impedances in one-stage version of the amplifier at the initial stage of design.

As the working frequency band of the three-stage LNA did not exceed 20%, so quarter-wave stubs were used for power isolation that provided possibility of installation behind the stubs of the resistors, providing a necessary stock of stability at extra band frequencies due to introduction of dissipative losses at suppression of their noise inside the band.

In the course of calculations it has become clear that a very essential role for complex optimization of the amplifier is played by the inductance size in a source of transistors (that is hard to achieve when using cased transistors parameters). In the first and second stage consecutive feedback on current in the form of a piece of a high-resistance transmission line in a source is applied. In the first stage the size of feedback was chosen proceeding from realization of necessary stability and coordination in noise parameters. In the second stage the feedback size, electric distance between the 1 and 2 stage and chains of coordination were chosen based on the realization of K_u stability and maximization. In the third stage parameters selection of coordination chains, decoupling and attenuation provided stability and correction of unevenness of the characteristic, as well as the size of K_u.

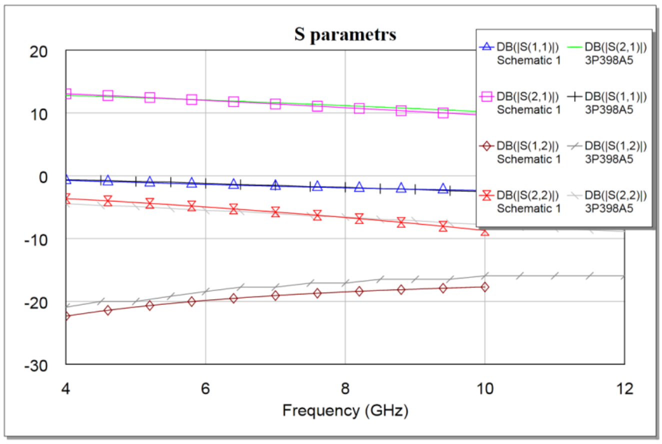


Fig. 6. The measured and optimized parameters ratio of the 3П398Б-5 transistor

On the basis of the above mentioned, a three-stage LNA with the central frequency of 8 GHz and a frequency band more than 15%, K_u = 28 dB, K_n = 1.1 dB, VSWR less than 1.3 has been developed. The scheme and layout of the LNA are given in Figs. 7 and 8.

In Fig. 9, calculation frequency characteristics of k_a, VSWR, k_n, and stability factors on amplitude *K* and on the phase B_f are shown.

The experience of the previous calculations of similar amplifiers on foreign transistors and their realization show the high extent of coincidence of calculations with the results obtained.

Based on the received data, the LNA device (Fig. 10) is functionally completed including two identical microassemblies of the amplifier, microstrip filter and two stripline isolators serving as input loading for amplification microassemblies. A waveguide isolator

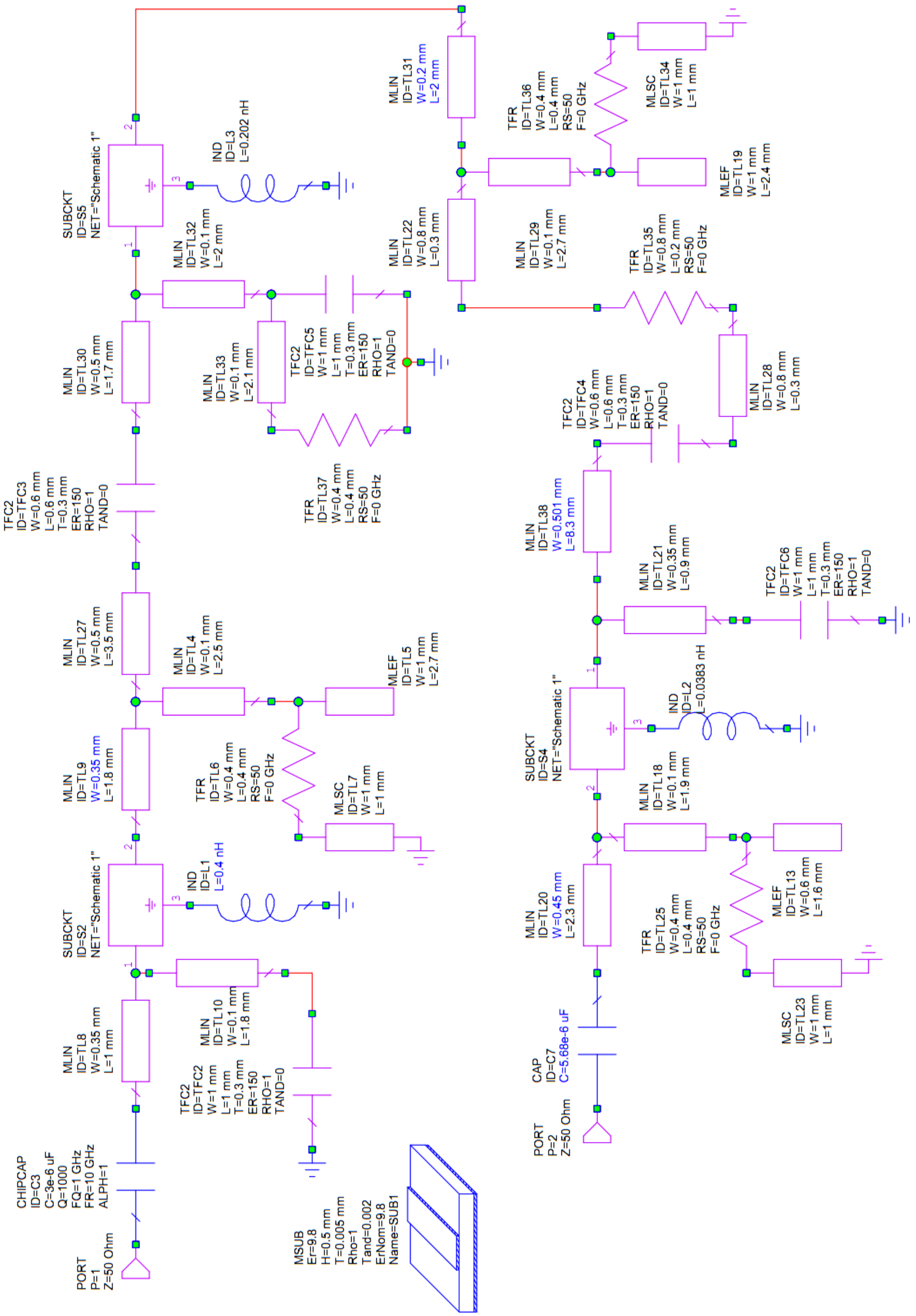


Fig. 7. Presentation of calculating elements of the LNA layout in AWR

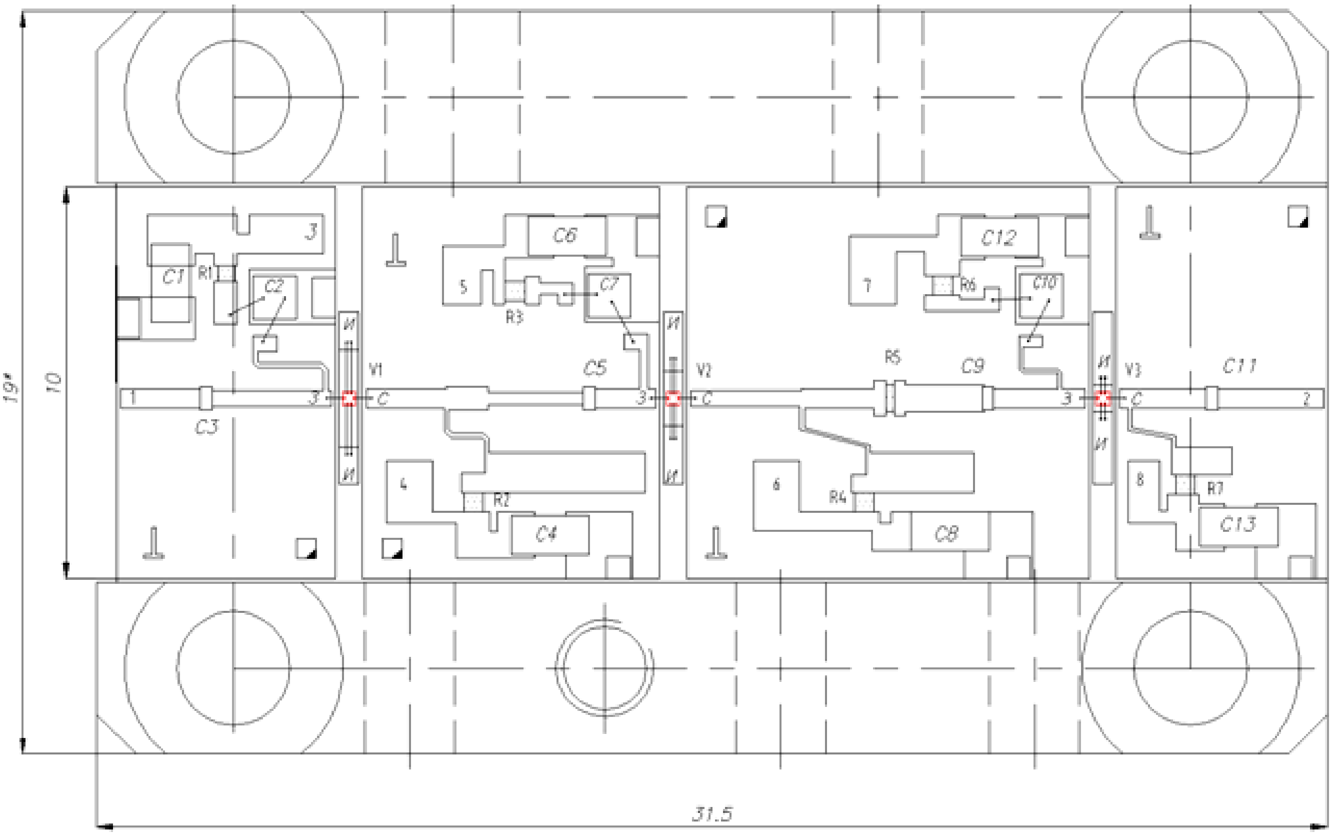


Fig. 8. The LNA layout implementation in the microassembly type

bringing a minimum of losses in the general highway provides output loading. Full autonomy to the device is provided by the source of secondary power supply with dimensions 60×40×8 mm established on the opposite side from a radiopath.

The LNA characteristics received during modulation are given in Fig. 11

The following measures are proposed to increase the operation reliability of the device:

1. There is no need of introduction of exotic technologies in the superhigh frequency band. Only strict implementation control of already available is necessary.

2. Time of no-failure operation of the transistor 3P398B-5 stated in the specifications is 50 000 hours under hard operation conditions. When solving the creation problem of satellites with the 15-year service life, the specified resource has to be 140 000 hours. In this regard, it is offered to apply triple redundancy of input blocks. That will require either purchase or own development of the three-position electromechanical switch operated by voltage pulses.

3. Moreover, increase in both transistors and a LNA unit service life is possible when the established threshold level of current consumption is achieved using the timely start of relaxation process of a working half-set.

A method to increase the operation recourse of the VHF transistors for special applications

Ensuring the required radiation resistance of the onboard electronic equipment to the space ionizing radiation is one of the most important solvable problems of spacecraft creation with long terms of service life (10–15 years).

Duration of spacecraft lifetime directly depends on resistance of the electronic component base (ECB) being used to special factors of the space (S).

In technical requirements for the transistor being used, the time of no-failure operation is 50 000 hours at the ambient temperature up to +85 °C. According to the specification on the system, the maximum working temperature is +50 °C. This circumstance objectively promotes increasing mean-time-between-failures; however, it requires additional tests and coordination with the producer of the element base.

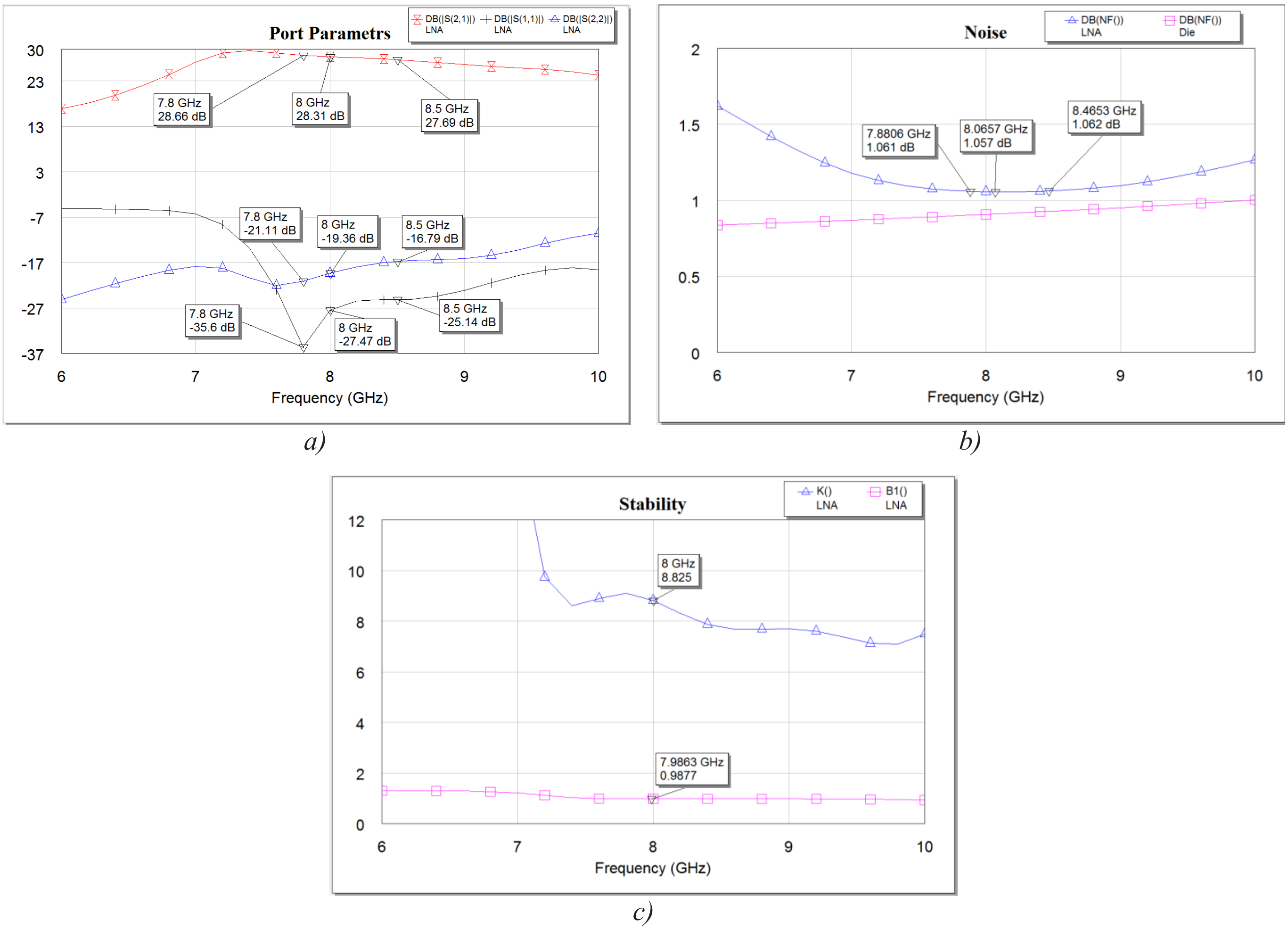


Fig. 9. Design characteristics of the LNA microassembly: S-parameters *(a)*, noise characteristics *(b)* and stability *(c)*

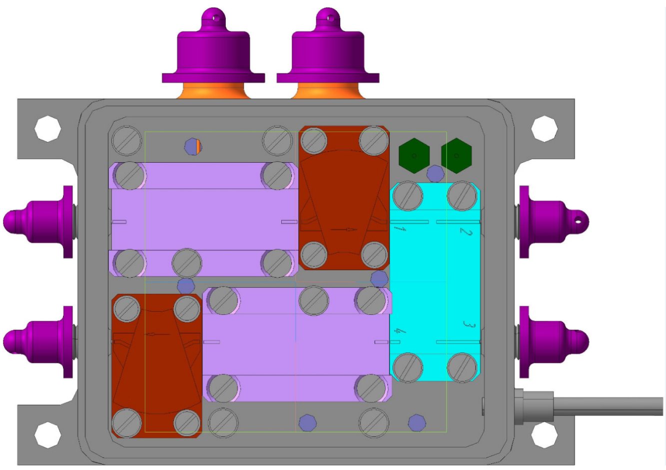


Fig. 10. An external appearance of the LNA

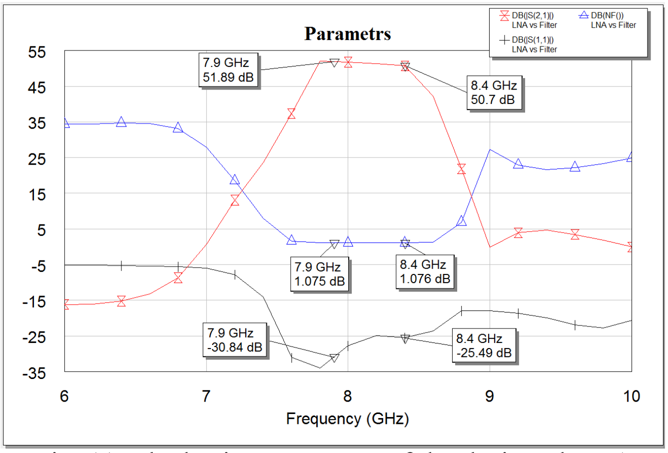


Fig. 11. The basic parameters of the designed LNA

The interests in ensuring reliability of onboard devices require reservation at the level of devices. For unconditional implementation of the customer requirements, it is offered to use triplicating for the operating time for the LNA unit.

In turn, this decision assumes using a three-position electromechanical switch operated by voltage pulses.

The research data of special space factors impact on the operation physics of semiconductor devices indicate a gradual increase in current consumption by electronic

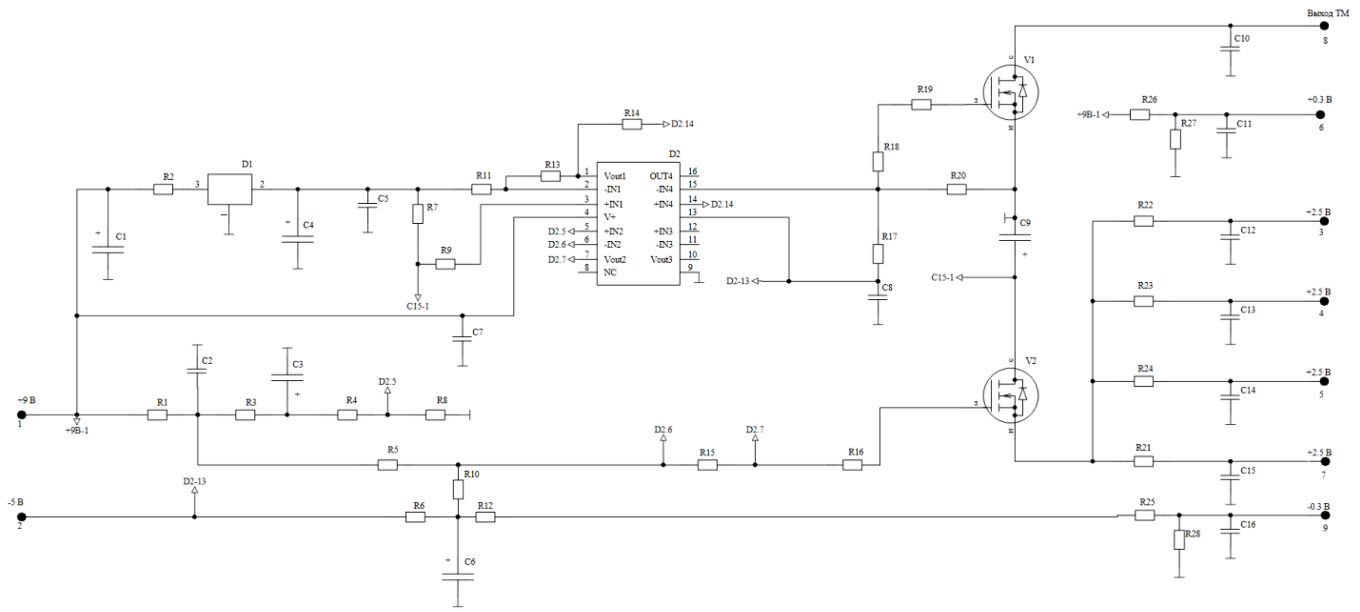


Fig. 12. A fragment of power supply scheme to the LNA with the protection function on the accumulated dose. Scheme active elements: D1 — a linear voltage regulator 1325ЕР1У; D2 — a microcircuit of operation control of the electronic equipment 1114СК1У; V1, V2 — MOSFET- transistors 2П525А-9

devices in course of time. As the statistics shows, failures in the space conditions are caused largely by excess of the maximum values of the absorbed dose of the ionizing radiation specified in element technical requirements. The heavy charged particles (HCP) influence on operation reliability of arsenide-gallium transistors is negligible. The radioelement parameters drift involves the ionizing space radiation influence on the semiconductor structure and, in fact, involves a critical dose excess indicator in relation to the ECB element. At the same time, the element degradation extent under the influence of the ionizing radiation is defined by a ratio between accumulation processes and relaxation processes (annealing). Annealing partially allows products to be brought to a working condition, that is to neutralize a taken positive charge by the radiation-induced electrons from a conductivity region. The element restoration speed is defined by temperature optimum values duration of annealing.

This article offers the design solution that enables failures to be prevented in the equipment via tracking the moment of consumption current excess by active elements under the influence of special factors and giving the relevant telemetric information to the decisive device. The LNA scheme entirely developed on the domestic ECB meets the requirements of the priority (according to time) feed of negative bias on locks of microwave transistors and also identifies failures due to potential jumpers break. The similar circuit decision can be employed in the majority of the reserved radio frequency blocks of a retransmitter. A fragment of the scheme is presented in Fig. 12.

The R1 resistor is connected in series into the feed bar of the LNA. Depending on the current consumption, voltage drop to R1 (about 60 mV) is amplified by the first two stages of the operational DC amplifier D1. As current consumption by amplifiers grows up to 25% (at the corresponding excess of a threshold at the input of the second operational amplifier (OpAmp) stage up to 5 mV), the positive output voltage of the operational amplifier (+ 5 V) is transformed into negative (-5 V) one. Thus, the V1 transistor, being used for telemetry of the LNA block operability, will pass from an open state into the closed mode. It will serve as a fault signal, according to which the system will carry out the following switching: an input signal and voltage supply into a reserve half-set of the LNA, and the working half-set into the relaxation mode. Removal of potential from the switched half-set allows its working capacity to be restored eventually either completely or substantially, depending on time of the subsequent relaxation.

Conclusions

1. By means of the CAD system based on serially released low-noise transistors of a domestic production, the layout of the three-stage amplifier (with parameters at the level of the international standards), which has stability, internal coordination, amplification factor, and noise temperature, is calculated.

2. The production technology of amplifiers of a similar class in the superhigh frequency band (which was optimized during decades) enables one to realize in practice the offered scheme of the LNA without special improvements. Costs of production can be significantly reduced at the order of the lot of the average sizes products.

3. The proposed design concept making it possible to prolong a service life of devices in the conditions of influence of space factors can be used when developing the equipment for spacecraft with the long term of active existence.

References

1. Dmitriev V. *Otechestvennye SVCh-komplektuyushchie na arsenide galliya. Komponenty i tekhnologii* [Domestic UHF-spare parts based on gallium arsenide. Components and technologies]. 2011, No. 6. (in Russian)

2. Kozlovskiy E.Yu., Seleznev B.I. Modelirovanie SVCh maloshumyashchego pHEMT tranzistora s primeneniem SAPR Microwave Office [Modelling of the UHF low-noise pHEMT transistor with the use of CAD Microwave Office]. *Vestnik Novgorodskogo gosudarstvennogo universiteta* [Vestnik NovSU]. 2012, No. 68. (in Russian)

3. Arykov V.S., Barov A.A., Kondratenko A.V. Monolitnaya integral'naya skhema maloshumyashchego usilitelya diapazona 8–12 GGts na osnove GaAs PHEMT–tekhnologii [An integrated monolithic circuit of the low-noise amplifier of 8–12 GHz based on GaAs PHEMT technology]. *I Vserosiyskaya nauchno-tekhnicheskaya konferentsiya "Elektronika i mikroelektronika SVCh"* [I All-Russian scientific and technical conference "Electronics and Microelectronics"]. Saint Petersburg, 2012. (in Russian)

4. Mokerov V.G., Babak L.I., Fedorov Yu.V., Cherkashin M.V., Sheerman F.I., Bugaev A.S., Kuznetsov A.L., Gnatyuk D.L. Razrabotka komplekta monolitnykh maloshumyashchikh usiliteley Kh-diapazona na osnove 0.15 mkm GaAs pHEMT-tekhnologii [The design of the set of the monolithic low-noise amplifiers of the K-band based on 0.15 μm GaAs pHEMT technology]. *Doklady TUSURa* [Proceedings of TUSUR]. No. 2 (22), p. 1, December 2010. (in Russian)

5. Mikhaylovich S.V., Fedorov Yu.V., Bugaev A.S., Galiev R.R., Yachmenev A.E., Shcherbakova M.Yu. Postroenie masshtabiruemoy shumovoy modeli MHEMT na GaAs s Lg ot 50 do 250 nm [On the making of scalable noise model of GaAs MHEMT with Lg of 50 to 250 nm]. *Doklady TUSURa* [Proceedings of TUSUR]. No. 2 (24), p. 2, December 2011. (in Russian)

6. Boyko K.V., Noykin Yu.M., Zargano G.F. Maloshumyashchiy usilitel' na PTSh [A low-noise amplifier on MESFET]. *Metodicheskie ukazaniya k vypolneniyu spetsial'nogo laboratornogo praktikuma "Nelineynye tverdotel'nye ustroystva SVCh" (spetsial'nost' 071500, radiofizika i elektronika) Chast' XV* [Recommended practices for implementation of specialized laboratory course "nonlinear solid devices" (specialty 071500, radiophysics and electronics). Part XV]. Rostov-on-Don, 2001. (in Russian)

7. Iovdal'skiy V.A., Vinogradov V.G., Zemlyakov V.E., Lapin V.G., Kilichenkov R.B., Grinberg D.S., Gerasimenko S.V. Modelirovanie konstruktsiy GIS MShU SVCh-diapazona [The modelling of structures of GIS LNA of the microwave band]. *Materialy Mezhdunarodnoy nauchno-tekhnicheskoy konferentsii INTERMATIC–2012* [The materials of the international scientific and technical conference INTERMATIC–2012]. Moscow: MIREA, 2012, p. 4. (in Russian)

8. Petukh N.N., Belousov Yu.S., Garbuzenko A.P. Metodika proektirovaniya maloshumyashchego usilitelya povyshennoy nadezhnosti s ispol'zovaniem otechestvennoy EKB [The design technique of the low-noise amplifier of advanced reliability using the domestic ECB]. *Materialy VII Vserossiyskoy nauchno-tekhnicheskoy konferentsii "Aktual'nye problemy raketno-kosmicheskogo priborostroeniya i informatsionnykh tekhnologiy"* [The materials of VII All-Russian scientific and technical conference "Current problems of rocket and space device engineering and information systems"]. Moscow, RUDN, 2015. (in Russian)

ROCKET-SPACE DEVICE ENGINEERING AND INFORMATION SYSTEMS
Vol. 3, Iss. 1, 2016, pp. 85–90

On Creating the World's First Strategic Long-Range Missile R-7 and Its Control System

V. K. Startsev
candidate of engineering science

The development of the R-7 (8K71) rocket was made according to the Government Resolution dated May 20, 1954.

The chief manufacture OKB-1 (special research bureau) at NII-88 (chief designer S.P. Korolev) and co-manufactures: OKB-456 (engine, chief designer V.P. Glushko), NII-885 (control system, chief designers M.S. Ryazansky and N.A. Pilyugin), NII-10 (giroinstruments, chief designer V.I. Kuznetsov), GSKB "Spetcmash" (Barmin's Spetcmash bereau) (ground-based equipment, chief designer V.P. Barmin) were assigned according to the Resolution.

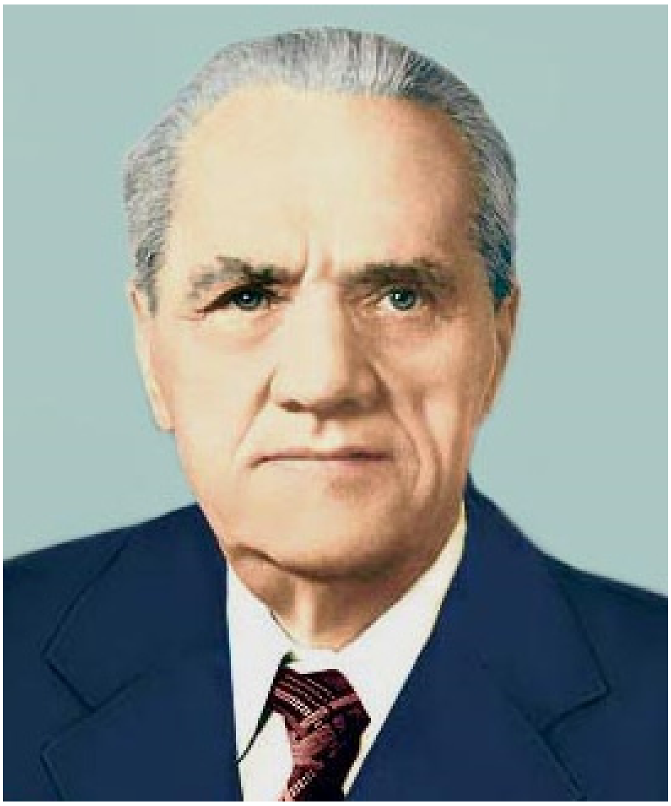
The preliminary design on the R-7 rocket complex was completed in July 1954.

The commission of experts headed by M.V. Keldysh, the president of the USSR Academy of Sciences, was formed for consideration of the preliminary design. The commission included prominent scientists and representatives of the customer. The commission of experts concluded that the materials of the preliminary design proved correctness of the choice of the schematic diagram for the rocket, its propulsion systems, and flight control system with the ground-based equipment and could be the basis for further efforts.

On November 20, 1954, the preliminary design of the R-7 rocket was approved by the Council of Ministers of the USSR.

Ryazansky Mikhail Sergeyevich

Deputy Director General for Science at NII-885 during 1955-1985, Chief Designer of the R-7 radio control system



Pilyugin Nikolay Alekseyevich

Chief Designer of the R-7 autonomous control system during 1955-1963

The R-7 was a two-stage vehicle built according to the cluster configuration. Its first stage consisted of four side units symmetrized around the central unit, which, in its turn, was the second stage (Fig. 1).

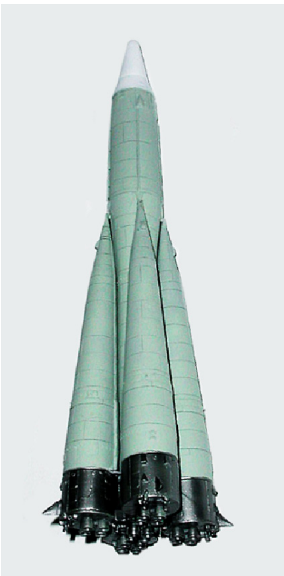


Fig. 1. The R-7 general appearance

The central stage included an instrument section, oxidizer tanks (cooled oxygen), fuel (kerosene) tanks, and propulsion.

The development of the design documentation for the missile system and its components began in 1953.

The first flight model of the rocket was sent to the newly-built research test site Tyuratam (Baikonur) at the end of 1956.

The creation of the test site was made according to the Resolution of Council of Ministers of the USSR of 12.02.1955, which was called “On the new research test site for the Ministry of Defence”. The Resolution provided:

− creating in the period of 1955–1958 a research test site of the Ministry of Defence for flight development of the R-7 components and equipment, “Burya” and “Buran”, with headquarters of the test site located in the Kyzylorda and Karaganda regions of the Kazakh Soviet Socialist Republic (KSSR) in the region between Kazalinsk and Zhosaly, an area of falling of the parts of products in the Kamchatka region of the Russian Soviet Federative Socialist Republic (RSFSR) at the cape Ozoyrny, an area of falling of the first stages of the R-7 in the territory of the Aktobe region of the KSSR near the Tangiz Lake;

− preparing in a three-months period the organization actions on construction of the specified research test site .

The choice of the place for the research test site was a very difficult task since the preliminary design of the R-7 involved obligatory existence of a radio control system. The system demanded a symmetric arrangement of the main and mirror points of a radio system on both sides from the launching pad at the distance of 250 km, which was defined by the need to ensure the accuracy of measurement of a side deviation of the rocket during a boost phase.

The location options for the research test site in Transcaucasia and in the Far East were rejected.

The launch-site data acquisition and measurements system consisted of the ground stations equipped with radio engineering and optical systems for trajectory parameters measurement was created as a part of the research test site on the extensive territory on the R-7 flight route and on the fields for head parts falling.

Building of launch and technical facilities, automobile and railroads, houses and social and cultural life institutions was developed to solve these tasks.

People who in severe climatic conditions had to perform tasks of the government in short terms were the main force.

On January 15, 1955, the first group of military builders under the command of the senior lieutenant I.N. Denezhkin arrived at the Tyuratam station. The first building of the inhabited area for test engineers, the Zarya settlement (the present city of Baikonur), was put already on May 15, 1955.

In December 1956, plane flights of all launch-site data acquisition and measurements system stations were carried out, and the special commission made the conclusion about their normal functioning and readiness for service.

The difficulties, which military builders and test engineers met during creation of the research test site and its measurement complex, were connected not only with severe climatic conditions, but also with extremely short deadlines of construction and lack of experience of building similar objects.

In March 1957, the first R-7 rocket for carrying out development flight tests (DFT) arrived to a technical position of the test site.

On April 10, 1957, the meeting of the State commission on carrying out the DFT took place. The state commission was approved by the Council of Ministers of the Soviet Union. V.M. Ryabikov from GU Glavspetsmontazh of the Ministry of Medium Machine-Building Industry of the USSR was appointed its chairman, and S.P. Korolev became its engineering manager.

On May 15, 1957, the State commission signed the act of acceptance of a launch pad and readiness of the test site for the first launch of the R-7.

The first missiles launches (on May 15, on June 11, and on July 12, 1957) were emergency, generally because of malfunctions of the propulsion system.

The fourth launch on August 21, 1957, was successful, and the rocket reached the targeted region for the first time.

A combined (radio- and autonomous) control system for the strategic ballistic missile R-7 providing the set deviations of a rocket head part was created in NII-885. At the same time, the integrator of the apparent velocity of the autonomous system was adjusted on the border of the area of possible switching off the engine corresponding to flight of the head part from the purpose on range.

Two new complexes were formed in NII-885 for solving this task: the Complex No. 1 for development of an autonomous control system headed by N.A. Pilyugin and the Complex No. 2 for creation of a radio control system headed by M.S. Ryazansky.

The Complex No. 1 included four departments:

- for stabilization machine development (department head G.P. Glazkov);
- for design and theoretical efforts (department head M.S. Khitrik);
- for development of complex schemes, power supply system and onboard cable system (department head Ya.S. Zhukov);
- for creation of element means (department head M.S. Doynikov).

The Complex No. 2 included two departments:

- for development of low frequency equipment and a complex in general (department head Ye.Ya. Boguslovsky);
- for development of high frequency equipment and antenna systems (department head M.I. Borisenko).

Chief designers N.A. Pilyugin and M.S. Ryazansky were the outstanding scientists, both of them were included into the legendary Council of Chief designers created by S.P. Korolev in Germany.

The strategic R-7 rocket had considerable design differences in comparison with small range missiles. Due to this fact the autonomous control system was faced to solve a number of tasks of the accounting of the parameters, which influence the control accuracy in earlier developed control systems were not so considerable.

As a result, the autonomous control system besides traditional functions on stabilization of the rocket body in space, stabilization of the center of mass with respect to the planned trajectory and control function on flying range provided continuous velocity vector control of the rocket and regulation of process of fuel tank draining. It was implemented in the created systems of normal and side stabilization of the rocket center of mass, regulation system of phantom velocity (PVR), tank draining system (TDS), as well as in regulation system consumption and ratio of fuel's components. The autonomous system operated these processes during the flight of the first stage, cluster staging, and flight of the second stage. The radio system made correction of the rocket movement in the side direction and control in range by outputting the side commands and a command for of the engine switching off. Control engine chambers and air vanes were the executive bodies of the control system.

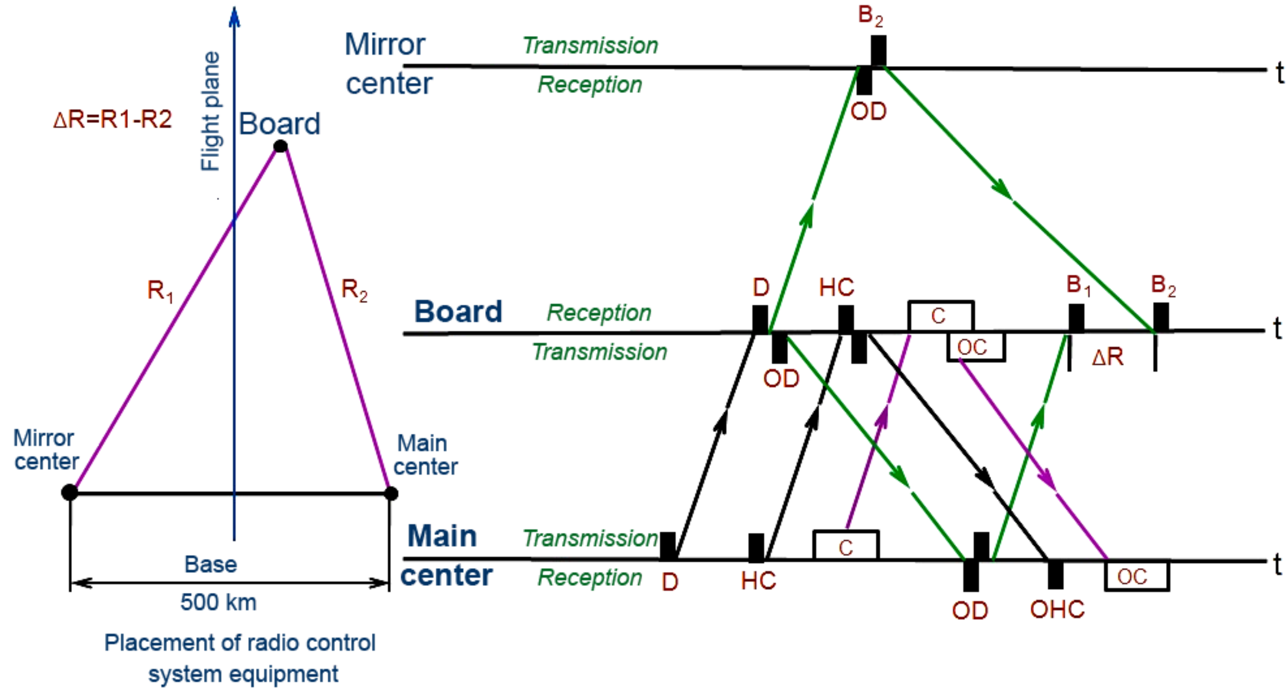


Fig. 2. Scheme of radio signal transmission in the radio control system of the R-7

A special modeling equipment with standard devices was installed at the Complex No. 1 stand for development tests of control processes and the choice of optimum interaction of the autonomous and radio control systems. An analog type computer was installed in the equipment to simulate dynamic characteristics of the rocket as a subject of control.

Subsequently, the simulating installation was used for carrying out functional control of standard devices before their delivery to the head enterprise.

The system of radio control was a pulsed and ranging system incorporating a ground and onboard equipment. The ground equipment was placed on the main and mirror points. The onboard receiving and transmitting equipment was placed in the instrument section of the rocket and was interfaced to the autonomous control system.

Radio control in range was made based on measurement of six movement parameters.

Measurement of movement parameters and command transmission of control in radio system was effected using a common pulsed multichannel radio communication line operating in the X-band of radio waves with coded signals.

Fig. 2 shows a scheme of radio signal transmission in the radio control system of the R-7.

Composition of the parameters being measured:

1. Radial velocity $\dot{R}$. Radial velocity measurement was carried out by discrimination of Doppler frequency increment of one of the spectrum harmonics of complex signals in the form of a narrow pulse packet sent from the main center and then retransmitted back.

2. Radial range R. Radial range measurement was effected by the method of active radiolocation.

3. Side deviation from the flight plane ΔR. The measurement was executed as difference of distances between the rocket and main and mirror centers.

4. Velocity change of side deviation $\Delta\dot{R}$. The measurement was a result of differentiation of the side deviation value being measured.

5. Elevation and speed of its change β and β. The measurement was carried out with a special radio direction finder built based on the method of amplitude-pulse modulation.

Side radio correction of the flight was carried out onboard the rocket by development of the signals corresponding to the measured side deviation and side velocity with respect to the flight plane with the amendment considering influence of the Earth rotation. To do this the results of measurements of a corner of the place β and the velocity of his change β made by a radio direction finder were used. These signals arrived in the amplifier-converter of the automatic machine of side stabilization of the autonomous system and were transformed into the commands sent to the mechanisms of turn of the rocket in the angle of yaw.

Originally, the system of radio control possessed an essential disadvantage: it did not allow one to carry out change of the flight direction of the rocket in the acceptable sector of angles.

Later, the equipment of the main point station involved a station of side control with a computer determining the size of side velocity, which needs to be reported to the rocket by the time of the engine switching off to compensate the accumulated disturbances at the end of an active lag of the flight to satisfy this requirement.

A calculation algorithm for problem solving of side velocity value included all six movement parameters being measured.

The station of side control formed the commands arriving aboard the rocket for giving it the corresponding angle of yaw.

The commands of side control had a discrete character and were designated as "Big left", "Small left", "Zero", "Big right", and "Small right".

The solution of a control algorithm in range was executed by the computing device located on the main station, which based on the current information on the movement coordinates of the rocket, gave a command for the engine switching off at the moment when a combined value of this information reached the set value. Switching off the engine for reduction of influence of dynamic disturbances occurred in two stages – on preliminary and main commands.

A ground equipment of the main station of the radio system was located in the settlement of Tartugay in the KSSR and placed at 13 stations having various functional purposes (receiving and transmitting stations, station of coder-decoders, station of timers, station of radial speed discrimination, station of computing devices, station of locational guidance of antennas, transmitting-receiving antennas, and radio direction finder).

The mirror station of the radio system was located in Sary Shagan settlement in the KSSR and served for retransmission of the radio signals radiated by the onboard equipment. The equipment of the mirror station was placed on 6 stations.

The onboard equipment of a radio control system incorporated the transferring and reception devices, onboard antennas (one was directed to the main station, another was directed to the mirror station), coder-decoder, device for antenna control and device of interface to the autonomous control system.

A preliminary flight development of the autonomous control system was made on the P5RД rockets, and radio systems flight development occurred on the P5P rockets at their launch from the missile research test site Kapustin Yar.

The ideological foundation for creation of the radio control system was laid by the outstanding scientists of NII-20 who made the significant contribution to development of domestic telemechanics and radiolocation equipment, subsequently transferred to NII-885: B.M. Konoplev, M.S. Ryazansky, E.M. Manukyan, E.Ya. Boguslavsky. Considerable efforts were executed on the research of questions of radio wave propagation by the employees of NII-885 E.F. Dubovitskaya, K.I. Gringauz, Yu.S. Pavlov, M.I. Borisenko, and a number of the enterprises of the USSR Academy of Sciences.

Chief designers of radio control system at NII-885 were:

- E.Ya. Boguslavsky, M.I. Borisenko, F.I. Tokarev, G.A. Vilkov, A.M. Trakhtman, S.P. Peshnev, K.K. Zykov, R.A. Chigirev, E.A. Rozenman, G.Ya. Guskov, B.G. Sergeyev, V.A. Grishmanovsky (ground equipment);
- N.E. Ivanov, I.Ya. Sytin, E.P. Molotov, V.G. Buryak, V.F. Grushetsky, V.P. Kuzovkin, Yu.F. Makarov, T.D. Fatkina (onboard equipment).

With completion of the R-7 rocket development, including a rocket control system, following persons received high government awards: M.S. Ryazansky, N.A. Pilyugin, E.M. Manukyan, P.A. Tunik (laureates of the Lenin Prize); E.Ya. Boguslavsky, M.I. Borisenko, G.P. Glazkov and workers of the pilot-producing plant V.I. Ryabov, S.I. Mikhaylov (the Hero of Socialist Labour). In total, 304 employees of NII-885 were awarded with orders and medals.

Considering a great scientific, practical and social value of the efforts performed on creating a new technology, chief designers M.S. Ryazansky and N.A. Pilyugin became corresponding members of the USSR Academy of Sciences. Twelve employees of NII-885 were awarded a doctor of engineering science degree without defending a thesis.

During the R-7 and its control system creation a number of scientific and technical problems that laid a fundamental scientific and technical basis for further improvement of rocket and space developments was solved.

The R-7 and its control system became the basic for creating a number of modifications.

Thus, the two-stage R-7 rocket – carrier vehicle "Sputnik" (8К71ПС and 8А91) enabled three first Earth artificial satellites to be launched allowing one to start research of the near Earth space.

Three-stage launchers 8К72 with the E stage ("Vostok") and 11А511 with the I stage ("Soyuz") enabled one to start far space and Moon research, launch "Vostok" and "Voskhod" manned spacecraft and later "Soyuz" spacecraft. The four-stage rocket 8К78 with I and L stages ("Molniya") permitted far space and Moon research to be expanded, as well as to carry out the flights of automatic interplanetary probes to the Mars and Venus.

The R-7 rocket and its modifications enabled one to expand diversified space research and create conditions for the applied usage of rocket and space technology for benefit of science, defence and national economy. Later these missiles were used for launching spacecraft "Zenit", "Meteor", "Electron", "Progress", etc.

Nowadays research and experiments for scientific and national economy purposes are continued.

V.K. Startsev, Candidate of Engineering Science, Laureate of the

Award of the Russian Federation Government in the Field of Science and Technology, Director of the History and Technology Museum

References:

1. *Vekhi istorii. 65 let OAO "Rossiyskie kosmicheskie sistemy"* [History landmarks. The 65th anniversary of Joint Stock Company "Russian Space Systems"]. Moscow, Izdatel'stvo "Media Pablisher", 2011, 128 p.

2. *Kosmonavtika. Entsiklopediya* [Cosmonautics. Encyclopedia]. Moscow, "Sovetskoe izdatel'stvo", 1985, 528 p.

3. B.E. Chertok. *Rakety i lyudi* [Rockets and people]. Moscow, "Mashinostroenie", 1996, 446 p.

4. *Baykonuru – 50. Istoriya kosmodroma v vospominaniyakh veteranov* [The 50th anniversary of Baikonur. History of the cosmodrome in memories of the veterans]. Moscow, OAO "Tipografiya "Novosti", 2005, 833 p.

5. V.I. Favorskiy, I.V. Meshcheryakov. *Kosmonavtika i raketno-kosmicheskaya promyshlennost'* [Cosmonautics and rocket and space industry]. Vol. 1, 2. Moscow, "Mashinostroenie", 2003, 344, 429 p.