FTD-TT- 64-1077

TT 64-71672

TRANSLATION

COSMIC RESEARCH

Volume 2 - Number 5 - 1964

FOREIGN TECHNOLOGY DIVISION

AIR FORCE SYSTEMS COMMAND

WRIGHT-PATTERSON AIR FORCE BASE

OHIO

UNEDITED ROUGH DRAFT TRANSLATION

COSMIC RESEARCH

English Pages: 261

PREPARED BY:

TRANSLATION DIVISION
FOREIGN TECHNOLOGY DIVISION
WP-AFB, OHIO.

FTD-TT-64-1077/1+2

Akademiya Nauk SSSR

KOSMICHESKIYE ISSLEDOVANIYA

Tom II

Vypusk 5

Izdatel'stvo Nauka

Moskva 1964

pages 657-813

FTD-TT-64-1077/1+2

TABLE OF CONTENTS

UDK 521.15:517.911

A STUDY OF THE OSCILLATION OF A SATELLITE IN THE PLANE OF AN ELLIPTICAL ORBIT

V.A. Zlatoustov, D.Ye. Okhotsimskiy, V.A. Sarychev, and A.P. Torzhevskiy

Plane oscillations of a satellite on an elliptical orbit under the influence of gravitational moments are considered. Periodic solutions to the equation describing these oscillations are sought. The questions as to the number of periodic solutions and their stability are investigated as functions of the defining parameters – the eccentricity of the orbit and the inertia characteristics of the satellite.

1. For communications, weather, geodetic and certain other types of artificial satellites, an important problem is to seek out methods for passive (without consumption of power and working fluid) orientation to the earth over a long satellite lifetime.

Passive orientation may be achieved, for example, by using the moments acting upon a solid body in a central gravitational field. It is well known that when a solid body moves on a circular orbit, the equilibrium positions corresponding to coincidence of the principal central axes of inertia of the body with the axes of the orbital coordinate system formed by the radius vector, the transversal and the binormal to the orbit include stable equilibrium positions [1, 2]. This circumstance has engendered a number of theoretical investigations and technical proposals concerned with the creation of a gravitational stabilization system for artificial satellites [3-8]. One of the designs of a gravitational stabilization system has recently been successfully put into

practice [9, 10].

On a weakly elliptical orbit, the body's stable equilibrium position almost always enters into stable oscillatory motion with small amplitude and short period, the latter equal to the period of revolution on the orbit.* These oscillations may be regarded as errors of orientation. However, they can easily be calculated and taken into account. In this case, it is expedient to select as the nominal unperturbed motion of the satellite not the equilibrium position in the orbital system of coordinates, as is the case for a circular orbit, but a certain regular periodic motion about this position.

This approach can also be used for elliptical orbits that differ sharply from the circular. The problem reduces to finding stable periodic oscillations of the satellite whose properties are acceptable from the standpoint of using these oscillations as nominal motions for the orientation system.

The object of the present paper consists in making a step toward the solution of this interesting problem in nonlinear mechanics. It will consider oscillations of a satellite in the plane of an elliptical orbit for arbitrary values of the eccentricity. Odd periodic solutions with a period equal to the satellite's period of revolution on its orbit are investigated, and the stability region of these solutions is determined. We see that stable periodic solutions exist for any values of the eccentricity, and that there may be two types of stable solutions in a certain domain of eccentricities. Computer results were used extensively in accomplishing this work.

A number of theoretical results pertaining in particular to proof of the existence of odd periodic solutions have been set forth in [11], which is published in the present issue. We note also that one of the cases of parametric resonance on a weakly elliptical orbit was examined

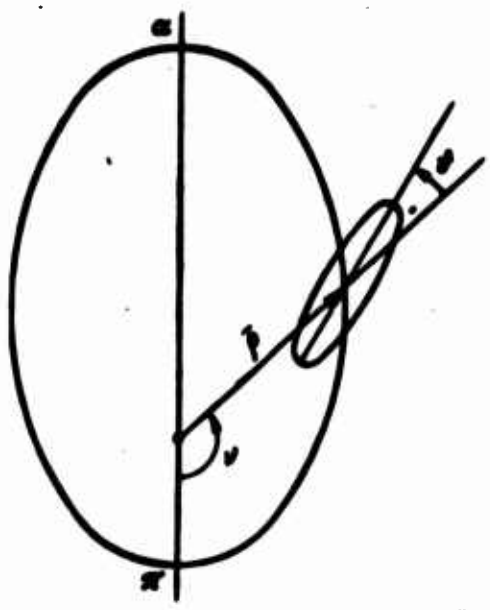


Fig. 1. Satellite on elliptical orbit. α) Apogee; π) perigee; ν) true anomaly; ϑ) angle between axis of satellite and local vertical.

in [12], and that the stability region for period motions of a satellite with almost equal moments of inertia was obtained in [13].

2. The oscillations of a satellite in the plane of an elliptical orbit are described by the following system of differential equations:

$$[d^2(\vartheta+\nu)/dt^2]+\alpha\omega_0^2(1+e\cos\nu)^3\sin\vartheta\cos\vartheta=0,$$
$$d\nu/dt=\omega_0(1+e\cos\nu)^2 \qquad (1)$$
$$(\alpha=3[(B-A)/C],\quad \omega_0=(R/p)\sqrt{g/p}).$$

Here, A, B and C are the principal central moments of inertia of the satellite, R is the radius of the earth, g is the acceleration of gravity at the surface of the earth, p is the parameter of the orbit, e is the eccentricity of the orbit, ν is the true anomaly, t is time and ϑ is the angle between one of the principal axes of the satellite's ellipsoid of inertia and the radius vector of the orbit (Fig. 1). Equations (1) have been derived on the assumption that the motion of the satellite's center of mass does not depend on its oscillations with respect to the center of mass.

If we take the true anomaly ν as the independent variable, then System (1) will be reduced to a single nonlinear differential equation of the second order, as derived in [2]:

$$(1+e\cos\nu)\frac{d^2\vartheta}{d\nu^2}-2e\sin\nu\,\frac{d\vartheta}{d\nu}+\alpha\sin\vartheta\cos\vartheta=2e\sin\nu. \qquad (2)$$

The parameters α and e satisfy the inequalities $|\alpha|\leq 3$, $0\leq e<1$. To a periodic solution of System (1) with a period equal to the period of revolution of the satellite's center of mass on the orbit there corresponds a periodic solution of Eq. (2) with a period 2π (the 2π-per-

iodic solution).

The following is known concerning 2π-periodic solutions of Eq. (2). It has been shown [11], using the results of [14-16], that for all values of the parameters α and e filling the region $E(|\alpha| \leq 3,\ 0 \leq e < 1)$, there exists at least one odd 2π-periodic solution. The region E is divided by a branching curve (bifurcation curve) into two subregions E_1 and E_3. The bifurcation curve proceeds out of the principal resonance point ($\alpha = 1$, $e = 0$), following the tangent to the α-axis. In region E_3, there exist three periodic solutions ϑ_0, ϑ_+, ϑ_-, of which ϑ_0 and ϑ_+ merge on the bifurcation curve and no longer exist as we move into region E_1. A single periodic solution exists in the region E_1.

Workable results on the behavior of the bifurcation curve and the number and type of the periodic solutions can be obtained only in the strip $0 \leq e \ll 1$ by the methods used in [11, 14-16]. The properties of periodic solutions to Eq. (2) have also been investigated in the region $|\alpha| \ll 1$ [13].

On a circular orbit ($e = 0$), Eq. (2) becomes the equation of free oscillations of a mathematical pendulum:

$$\frac{d^2(2\vartheta)}{dv^2} + \alpha \sin(2\vartheta) = 0. \qquad (3)$$

This equation is integrated in elliptical functions. The 2π-periodic solutions to Eq. (3) discussed above take the form

$$3 \geq \alpha \geq 1$$
$$\vartheta_0 = 0,$$
$$\vartheta_+ = \arcsin(k \operatorname{sn} \sqrt{\alpha}\, v), \qquad k^2 = \frac{1}{\alpha}\left(\frac{d\vartheta}{dv}\right)^2_{v=0},$$
$$\vartheta_- = -\arcsin(k \operatorname{sn} \sqrt{\alpha}\, v),$$
$$-3 \leq \alpha < 1$$
$$\vartheta_- = 0.$$

These solutions are generating solutions for eccentricity values other than zero. On a circular orbit, trivial solutions with positive α define the stable equilibrium position.

3. The results set forth above have been used in an investigation of periodic solutions of Eq. (2) over the entire region E. Most attention has been given the odd 2π-periodic solutions.

Finding odd 2π-periodic solutions of Eq. (2) is equivalent to solving the boundary-value problem for this equation with the boundary conditions

$$\vartheta(0) = \vartheta(\pi) = 0.$$

By way of illustration, Fig. 2 presents a diagram of $\vartheta(\pi)$ as a function of $\dot{\vartheta}(0) = d\vartheta/dv|_{v=0}$ for $\vartheta(0) = 0$. The curves have been plotted for $\alpha = 3$ and eccentricity values $e = 0$, 0.446, 0.6 and 0.8. On the figure, we can trace the evolution of the number and magnitude of the roots that correspond to the periodic solutions sought as functions of the orbital eccentricity. Three periodic solutions exist for $\alpha = 3$ and $0 \leq e < 0.446$. The point ($\alpha = 3$, $e = 0.446$) lies on the bifurcation curve. For values of $e > 0.446$ there exists only one periodic solution.

Thus solution of the boundary-value problem reduces to determination of all values of $\dot{\vartheta}(0)$ for which the condition $\vartheta(\pi) = 0$ is satisfied. Results of solving the boundary-value problem are presented in Fig. 3, in which the initial angular velocity $\dot{\vartheta}(0)$ is plotted as a function of the parameter e for a number of values of α. It is evident from the figure that for $-3 \leq \alpha < 1$ there always exists one initial velocity value that realizes periodic solutions, while three may exist for $1 \leq \leq \alpha \leq 3$. The broken line corresponds to the bifurcation curve. Positive values of $\dot{\vartheta}(0)$ correspond to the periodic solutions ϑ_+ and ϑ_0, and negative values to the solution ϑ_-. On the figure, the initial data for the solution ϑ_+ lie above the bifurcation curve. Positive values of $\dot{\vartheta}(0)$ below the bifurcation curve correspond to the solution ϑ_0. Investigation of the region of large eccentricities ($0.9 \leq e < 1$) and negative α has not been completed.

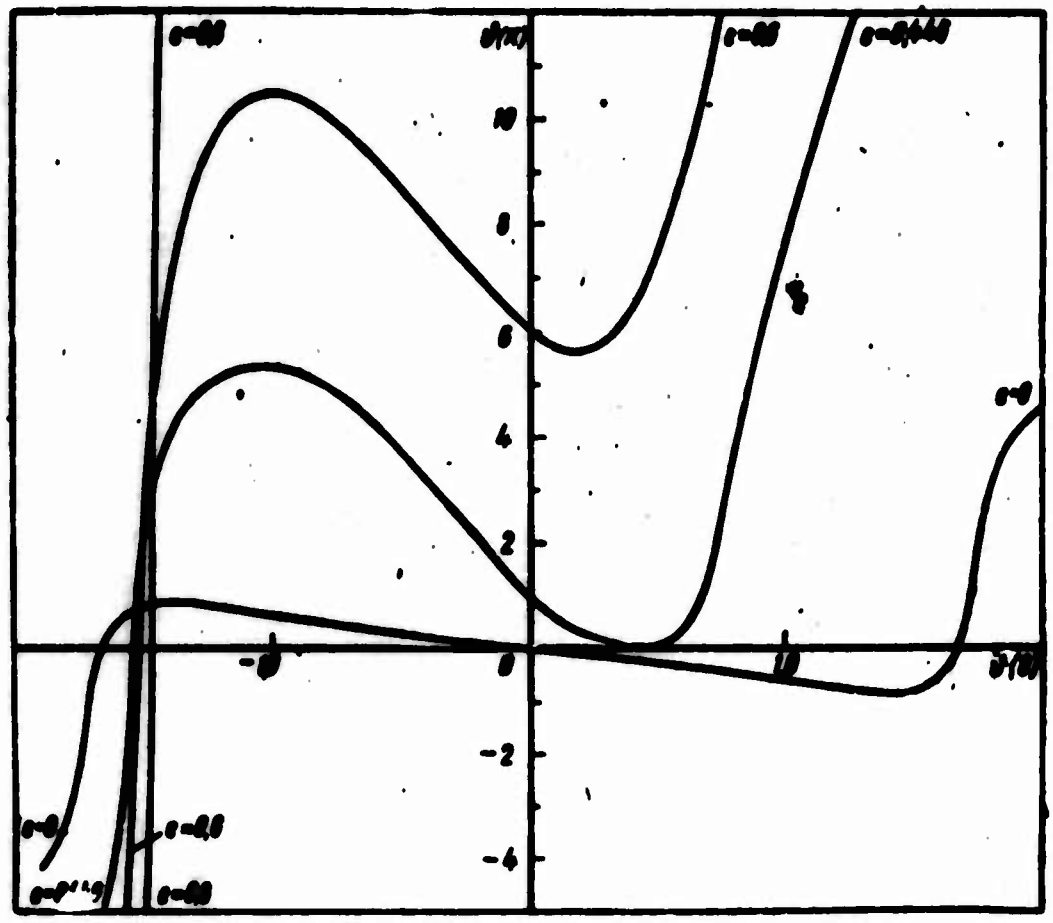


Fig. 2. Dependence of $\dot{\vartheta}(\pi)$ on $\dot{\vartheta}(0)$. $\vartheta(0) = 0$, $\alpha = 3$.

4. With the object of studying the stability of the periodic solutions obtained, let us write an equation in variations for Eq. (2):

$$(1+\varepsilon\cos\nu)\frac{d^2x}{d\nu^2} - 2\varepsilon\sin\nu\frac{dx}{d\nu} + \alpha\varepsilon\cos 2\vartheta^* = 0, \qquad (4)$$

where ϑ^* is the periodic solution being tested for stability and $\underline{x}$ is a small deviation from this solution.

The characteristic equation for the variation equation (4) may be written as follows:

$$\lambda^2 - 2A\lambda + 1 = 0, \text{ where } A = \tfrac{1}{2}[x_1(2\pi) + \dot{x}_2(2\pi)]. \qquad (5)$$

Here x_1 and x_2 are solutions of the equation in variations, solutions that form a fundamental system and satisfy the initial conditions

$$\begin{cases} x_1(0) = 1, \\ \dot{x}_1(0) = 0, \end{cases} \quad \begin{cases} x_2(0) = 0, \\ \dot{x}_2(0) = 1. \end{cases}$$

If $|A| < 1$, the roots of the characteristic equation are complex-

conjugate, and the periodic solution is stable in first approximation. The equation $|A| = 1$ gives the boundary of the stability region of the periodic solution. If $|A| > 1$, the periodic solution is unstable.

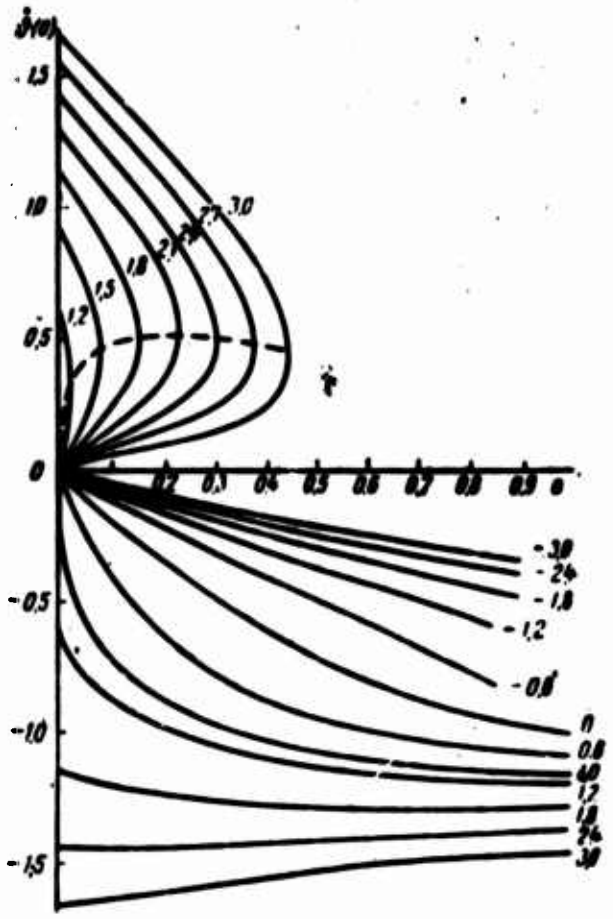


Fig. 3. Initial velocity $\dot{\vartheta}(0)$ corresponding to periodic solutions as a function of the parameters α and $\underline{e}$. The dashed line represents the bifurcation curve.

Values of $x_1(2\pi)$ and $\dot{x}_2(2\pi)$ were determined by numerical integration of Eq. (4) in variations. The results of investigation of the roots of the characteristic equation (5) are presented in Fig. 4, where the boundaries of the stability region of the periodic solutions have been plotted in the α, $\underline{e}$-plane (light lines) with the bifurcation curve (heavy line), which emanates from the point ($a = 1$, $e = 0$). The region E_3 of existence of three periodic solutions is situated to the left of and above the bifurcation curve in Fig. 4. The region E_1, which is situated to the right of and below the bifurcation curve, corresponds to a single periodic solution.

It has been ascertained as a result of analysis that the solution ϑ_+ is always unstable, while the solution ϑ_0 is stable in the region E_3 with the exception of a zone of parametric resonance that begins at point ($\alpha = 9/4$, $e = 0$). The stability region of the periodic solution ϑ_- is more complex. For this solution, the point ($\alpha = 1/4$, $e = 0$) is the initial point of the parametric-resonance region. The transition of the stability region from positive to negative values of α through the resonance point ($\alpha = 0$, $e = 0.682$) is highly interesting. Here, in the stable case, the axis of the minor inertia moment oscillates about the radius vector for $\alpha > 0$, while the axis of the major inertia moment does so for $\alpha < 0$.

Calculation of the boundaries of the stability region for the solution ϑ_- as $e \to 1$ involves a great deal of difficulty, since Eq. (2) has a singularity at $e = 1$, $\nu = (2k + 1)\pi$ ($k = 0, \pm 1, \pm 2, \ldots$). It appears that for $\alpha > 0$, both boundary curves fuse together to approach the vertical tangent to the point ($\alpha = 0$, $e = 1$), while for $\alpha < 0$ they tend asymptotically toward the line $e = 1$.

It follows from the results of [17, 18] that the necessary conditions for stability of periodic solutions of Eq. (2) as obtained on examination in first approximation, are also sufficient for almost all values of the parameters α and e.

5. Solutions of Eq. (2) in the regions $0 \leq e \ll 1$ and $|\alpha| \ll 1$ have also been investigated by asymptotic methods. Thus, for example, the boundaries of the parametric-resonance region of the periodic solution ϑ_0 have been determined in a small neighborhood of the point ($\alpha = 9/4$, $e = 0$). It was found that the resonance effect under consideration is detected only in the third approximation. The parametric-resonance region boundaries are determined by the relationship

$$\alpha = -\frac{9}{4} + \frac{1227}{200}e^2 \pm \frac{74853}{51200}e^3 + \ldots \qquad (6)$$

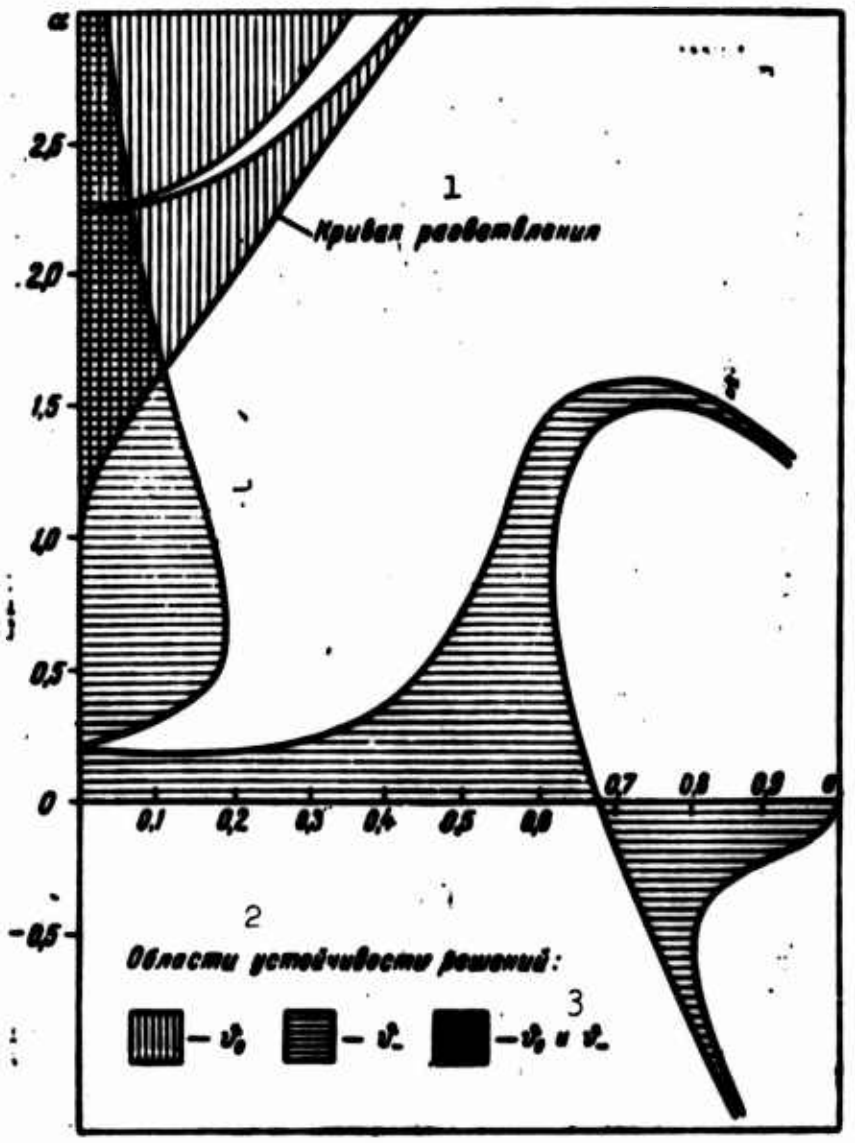


Fig. 4. Regions of existence of one and three periodic solutions and stability regions of the solutions ϑ_0 and ϑ_-. 1) Bifurcation curve; 2) regions of stability of solutions; 3) and.

The paramagnetic resonance region of the solution ϑ_-, which begins at the point ($\alpha = 1/4$, $e = 0$), can now be determined in first approximation. The boundaries of this region take the form [12]:

$$\alpha = \frac{1}{4} \pm \frac{3}{8}e + \ldots \qquad (7)$$

The behavior of the stability-region boundaries for the solution ϑ_- have been investigated in the strip $|\alpha| \ll 1$. The first five terms

of the asymptotic expansion were determined and made it possible to calculate not only the position of the point at which the stability-region boundary of the solution ϑ_- intersects the axis $\alpha = 0$, as was done in [13], but also the slope and curvature of the boundary in the neighborhood of this point.

6. Apart from the periodic motions about the radius vector, great interest attaches to periodic motions relative to a vector fixed in absolute space, for example, about a vector parallel to the major axis of an elliptical orbit. If we denote by ϑ_1 the angle between the satellite's principal axis of inertia and the direction to the perigee from the center of attraction, we may readily obtain from (2), using the relation

$$\vartheta_1 = \vartheta + v, \tag{8}$$

an equation for ϑ_1:

$$(1+e\cos v)\frac{d^2\vartheta_1}{dv^2} - 2e\sin v\frac{d\vartheta_1}{dv} + e\sin(\vartheta_1 - v)\cos(\vartheta_1 - v) = 0. \tag{9}$$

It can be shown by the methods described above that there exists a single 2π-periodic solution for Eq. (9) in the entire region E of the parameters (α, e). The generating solution for it will be $\vartheta_1 \equiv 0$, which obtains for $\alpha = 0$ and corresponds to translational motion in absolute space.

To investigate the stability of these solutions, we may employ Eq. (4) in variations, where it is necessary to set

$$\vartheta^* = \vartheta_1^* - v,$$

Here ϑ_1^* is the periodic (in absolute space) solution to be investigated for stability. The characteristic equation takes the form (5) and its coefficient A may be calculated by the method given above.

In Fig. 5, lines of constant A are represented in the α, e-plane. The stability region corresponds to that part of the plane in which

$|A| < 1$. The boundary of this region, which is identified in Fig. 5 by the heavy curve, consists of segments of the coordinate axes ($A = 1$), part of the upper boundary of the region E and a line $A = -1$ emanating from the point ($\alpha = 0$, $e = 1$) and having a vertical tangent at this point. The part of the region E lying to the right of and above the curve $A = -1$ corresponds to unstable solutions.

The results obtained signify that for elliptical orbits with arbitrary eccentricity there exists a range of values of α for which stable periodic motion of the satellite in absolute space is possible. If $e < 0.465$, then stable oscillations of the satellite are possible for arbitrary values of the parameter α in the interval $0 < \alpha < 3$. With $e > 0.465$, the largest admissible value of α decreases and, in particular, dumbbell-shaped oscillations of the satellite become unstable.

We can also examine periodic solutions in a form more general than (8), for example,

$$\theta_1 = \theta + \frac{m}{2}\nu. \qquad (10)$$

Here the $\underline{m}$ are arbitrary whole numbers. The values of the parameter $\underline{m}$ determine the secular variation of the position of the satellite axis with respect to the radius vector over the orbital revolution period. The value $m = 0$ corresponds to oscillations in the orbital coordinate system, and $m = 2$ to oscillations about the semimajor axis of the orbit (in absolute space). The existence of solutions of Type (10) was indicated in [11]. Their actual construction can be accomplished by solving the boundary-value problem with suitably selected boundary conditions.

7. A revealing characteristic of a self-contained system with one degree of freedom is its phase-plane analysis. For a nonself-contained system, the stroboscopic pattern formed by points of the phase paths at

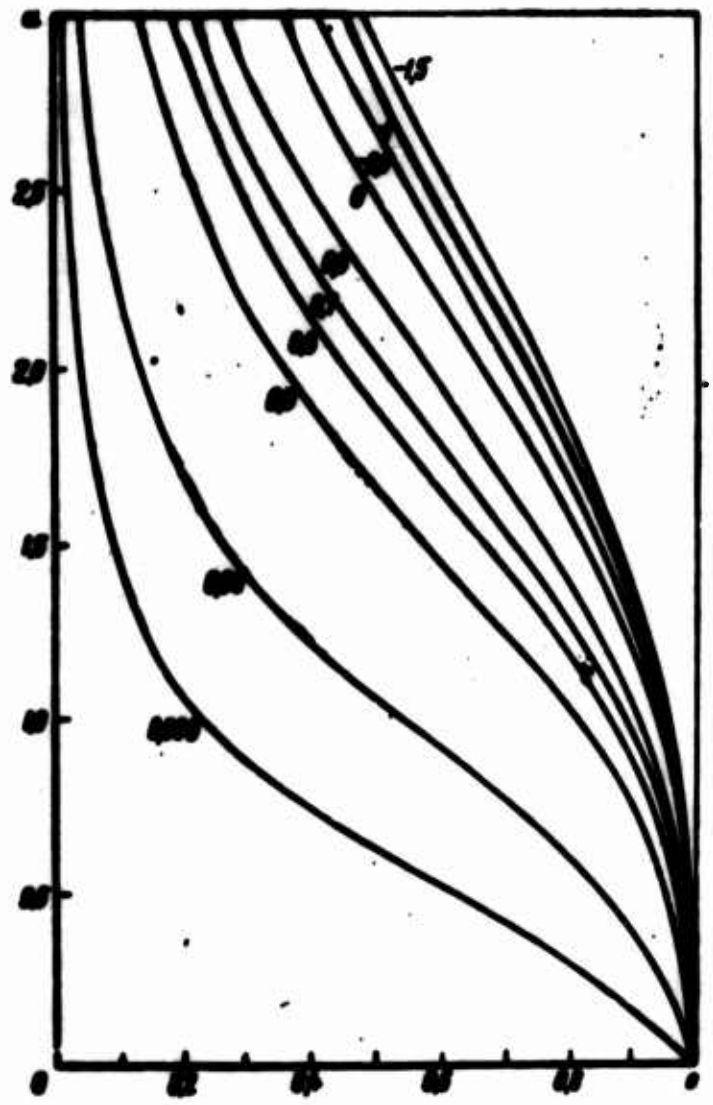


Fig. 5. Lines of A = const and stability region of periodic solution corresponding to oscillations about the direction of the semimajor axis.

discrete points in time differing by a multiple of the system's period performs a similar function. The time shift over a period defines a transformation of the points of the phase plane. A stationary point of such a mapping corresponds to the periodic solution. A periodic solution will be stable if the image of a sufficiently small neighborhood of the stationary point remains small for an arbitrary number of successive mappings.

Invariant sets that map into themselves on transformation after a period may exist on the phase plane. For a stable stationary point in the case in which the system does not have damping, invariant sets may

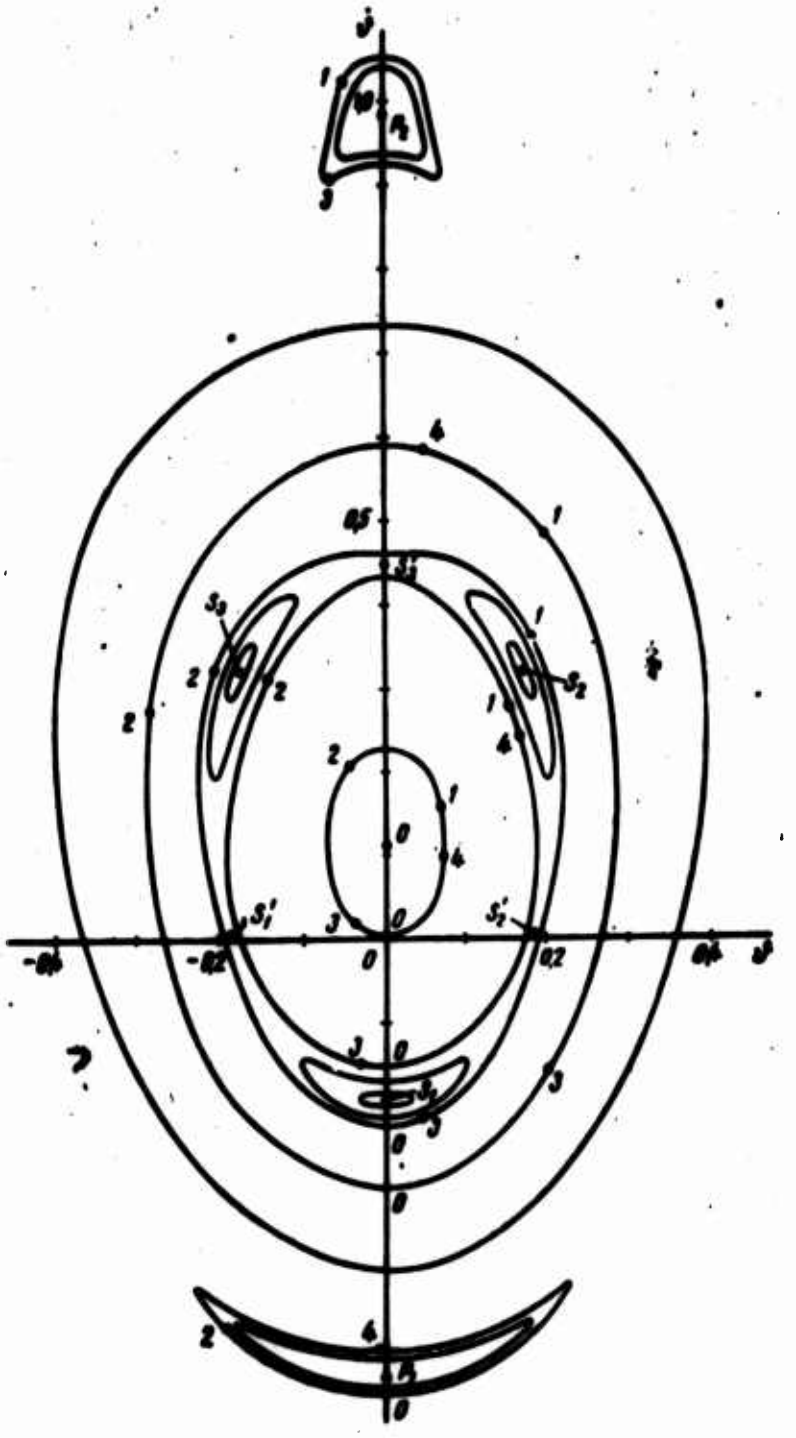


Fig. 6. Stroboscopic pattern of phase paths in the neighborhood of a periodic solution ϑ_0 for $\alpha = 3$ and $e = 0.2$.

form closed curves surrounding the stationary point.

The aggregate of invariant sets has a complex structure. A certain number of points on the phase plane correspond to subharmonic oscillations having a period that is a multiple of the system's period. On mapping after a period, these points map into one another. The number of points is determined by the multiplicity of the subharmonic. In the neighborhood of each of the points corresponding to stable subharmonics

there may be closed curves that, on transformation after a period, become curves surrounding another point of the same subharmonic. After a certain number of mappings, when the points of the subharmonic have returned to their initial positions, the initial curve is mapped into itself.

Plotting the stroboscopic pattern of the phase paths enables us to judge of the behavior of solutions neighborhing the periodic solution and to bound the region of initial data for which the deviation from the periodic solution remains limited.

As an example, Fig. 6 shows the stroboscopic pattern of phase paths for $\alpha = 3$, $e = 0.2$. The angle ϑ and the angular velocity $\dot{\vartheta}$ at perigee are plotted against the axes. The stationary point O corresponds to the stable periodic solution ϑ_0. Points numbered $n = 0, 1, 2, \ldots$, are obtained by transformation after n periods. The initial positions of the phase points are selected on the $\dot{\vartheta}$-axis. The figure represents invariant sets in the neighborhood of a periodic solution. Note the closed curves surrounding the stationary point. Note further the stable subharmonic with period 4π - points P_1 and P_2 mapping into one another after a period 2π. The points S_1, S_2 and S_3 correspond to a stable subharmonic and points S_1', S_2', and S_3' to an unstable harmonic with period 6π. Note also the invariant curves surrounding the stable points P_1, P_2 and S_1, S_2, S_3.

Received

10 June 1964

REFERENCES

1. F. Tisserand. Traite de Mecanique Celeste [Treatise on Celestial Mechanics], II, Paris, 1891.
2. V.V. Beletskiy, collection entitled "Iskusstvennyye sputniki Zem-

11" [Artificial Earth Satellites], No. 3, Izd-vo AN SSSR [Publishing House of the Academy of Sciences USSR], 1959 page 13.

3. D.Ye. Okhotsimskiy, V.A. Sarychev. Collection entitled "Iskusstvennyye sputniki Zemli" No. 16. Izd-vo AN SSSR, 1963, page 5.
4. V.A. Sarychev, collection entitled "Iskusstvennyye sputniki Zemli," No. 16, Izd-vo AN SSSR, 1963 page 10.
5. V.A. Sarychev, Kosmich. Issled. [Cosmic Research], 2, No. 1, 33, 1964.
6. L.J. Kamm. ARS Journal, 32, No. 6, 911, 1962.
7. E.H. Wrench, Report to the XIV Congress of the MAF [sic], Paris, 1963.
8. J.W. West, Report to the XIV Congress of the MAF [sic], Paris, 1963.
9. R.F. Fischell, F.F. Mobley, report TG 514, J. Hopkins University APL, August 1963.
10. R.B. Kershner, Astronautics and Aerospace Engineering, 1, No. 8, 18, 1963.
11. A.P. Torzhevskiy, Komich. Issled., [Cosmic Research], 2, this issue, page 667.
12. V.V. Beletskiy, collection entitled "Iskusstvennyye sputniki Zemli," No. 16, Izd-vo AN SSSR, 1963, page 46.
13. F.L. Chernous'ko, Zh. Vychislit. matem. i matem. Fiziki [Journal of Computer Mathematics and Mathematical Physics], 3 , No. 3, 528, 1963.
14. A. Hammerstein, Acta Math., 54, 117, 1930.
15. R. Iglisch. Monatsh. für Math. und Phys. [Monthly Bulletin for Mathematics and Physics], 39, 173, 1932.
16. E. Schmidt. Math. Annalen [Mathematical Annals], 65, 370, 1908.
17. V.I. Arnol'd, Usp. matem. nauk [Advances in the Mathematical Sci-

ences], 18, No. 6, 91, 1963.

18. Yu. Mozer, collection of translations entitled "Matematika" [Mathematics], No. 4, 3, 1962.

Manuscript Page No.

[Footnote]

2 The resonance regions form an exception.

UDK 521.15:517.911

PERIODIC SOLUTIONS TO THE EQUATION OF PLANE OSCILLATIONS OF A SATELLITE ON AN ELLIPTICAL ORBIT

A.P. Torzhevskiy

The topic is the existence and uniqueness of odd periodic solutions of period $2\pi k$ ($k = 1, 2, \ldots$) to the second-order nonlinear differential equation describing plane oscillations of a satellite on an elliptical orbit. The qualitative behavior of solutions with period 2π is considered as a function of parameter variation; a differential equation is written for a parameter-plane branching (bifurcation) curve, the initial segment of which is constructed in explicit form.

INTRODUCTION

One of the forms of motion of a satellite about its center of mass consists in oscillations of the satellite in the plane of the orbit about the direction of the radius vector, oscillations described by the equation [1]

$$(1+e\cos v)\frac{d^2\delta}{dv^2}-2e\sin v\frac{d\delta}{dv}+\alpha\sin\delta=4e\sin v, \qquad (0.1)$$

where δ is twice the angle between one of the principal axes of the satellite's ellipsoid of inertia that lie in the plane of the orbit and the radius vector of the orbit, v is the true anomaly, $\alpha = 3\frac{B-A}{C}$, A, B and C are the principal central moments of inertia, $|\alpha| \leq 3$ and e is the eccentricity, with $0 \leq e < 1$.

On a circular orbit ($e = 0$) and in the case of a body with dynamic symmetry and the axis of symmetry perpendicular to the plane of the orbit ($\alpha = 0$), Eq. (0.1) is integrated in quadratures. Hence for small e

or α we can find approximate expressions for the solution of this equation, working from the Poincare small-parameter method or the averaging method of N.M. Krylov and N.N. Bogolyubov, as was done in References [2, 3]. Since $e \neq 0$, $\alpha \neq 0$ represents a nonintegrable case, great interest attaches to problems of qualitative investigation of this equation. The present paper considers only odd periodic solutions of period $2\pi k$ ($k = 1, 2, \ldots$), the existence of which can be explained physically by symmetry of the force field with respect to the line of apsides. These solutions satisfy the conditions

$$\delta(0) = \delta(\pi k) = 0. \qquad (0.2)$$

We obtain the boundary-value problem (0.1)-(0.2), which will be investigated below. We note that the solutions being considered represent a particular case of oscillatory motions about a straight line passing through the satellite's center of mass and rotating with arbitrary angular velocity in absolute space (see footnote on page 33).

1. EXISTENCE AND UNIQUENESS OF PERIODIC SOLUTIONS

1. Consider the boundary-value problem

$$\ddot{x} = f(t, x) \qquad (1.1)$$

$$x(0) = x(a) = 0. \qquad (1.2)$$

Let the function f(t, x) be continuous in the strip

$$0 \leqslant t \leqslant a_0 \quad (a_0 > 0), \quad |x| < \infty.$$

Then the problem of (1.1)-(1.2) will be solvable if

$$\int_0^x f(t, u)\,du \geqslant -\frac{c_1}{2}x^2 + c_0 \qquad (1.3)$$

$$c_1 \leqslant (\pi/a)^2; \qquad (1.4)$$

and its solution will be unique when

$$|f(t, x_2) - f(t, x_1)| \leqslant c_0|x_2 - x_1|, \quad c_0 < (\pi/a)^2. \qquad (1.5)$$

Here the c_i are constants.

We shall examine the solution of Problem (1.1)-(1.2) as an extremal of the functional

$$J(x, \dot{x}) = \int_0^{\pi} \left[\frac{1}{2} \dot{x}^2 + \int_0^x f(t, u)\, du \right] dt. \qquad (1.6)$$

It can be affirmed that there exists at least one extremal if we prove that the functional has a lower bound. Let us apply the integral inequality [4]

$$\int_0^{\pi} \dot{x}^2 dt \geqslant (\pi/a)^2 \int_0^{\pi} x^2 dt,$$

which is satisfied if x(t) has a derivative with an integrable square and vanishes at the end points. From (1.6) we obtain

$$J(x, \dot{x}) \geqslant \int_0^{\pi} \frac{x^2}{2} [(\pi/a)^2 - c_1]\, dt + c_2 \geqslant c_2,$$

if

$$c_1 \leqslant (\pi/a)^2.$$

We reduce the boundary-value problem (1.1)-(1.2) to the integral equation

$$x(t) + \int_0^{\pi} K(t, \tau) f[\tau, x(\tau)]\, d\tau = 0,$$

where

$$K(t, \tau) = \begin{cases} \tau[1 - (t/a)], & 0 \leqslant \tau \leqslant t \\ t[1 - (\tau/a)], & t \leqslant \tau \leqslant \pi. \end{cases} \qquad (1.7)$$

Then Condition (1.5) proceeds from the Hammerstein uniqueness theorem [5-6].

Substituting variables as follows:

$$\delta = \frac{x}{1 + e \cos v}, \quad v = t,$$

we transform Problem (0.1)-(0.2) to the form

$$\ddot{x} = f(t, x), \qquad (1.8)$$

$$x(0) = x(\pi k) = 0, \qquad (1.9)$$

with

$$f(t, e) = -[e\cos t/(1+e\cos t)]e - e\sin[e/(1+e\cos t)] + 4e\sin t. \quad (1.10)$$

In analysis of (1.3), we employ the inequality

$$[\sqrt{\rho}e + 2e\sin t/\sqrt{\rho}]^2 \geq 0,$$

where ρ is an arbitrary constant.

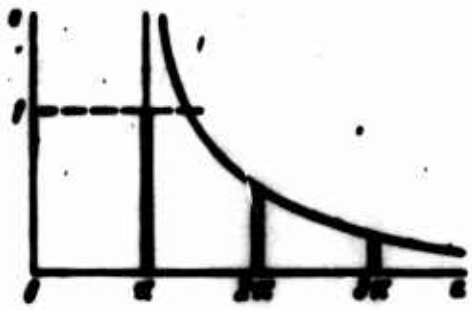


Fig. 1

As a result we get

$$c_1 = [e/(1+e)] + 2\rho.$$

This signifies that with the condition

$$e/(1+e) < 1/k^2 \quad (k = 1, 2, \ldots) \quad (1.11)$$

there exists at least one solution to Problem (1.8)-(1.9). This solution is unique if, according to (1.5),

$$(e + |a|)/(1-e) < 1/k^2 \quad (k = 1, 2, \ldots). \quad (1.12)$$

Inequality (1.11) is represented graphically in the form of a set of vertical settlements with cutoff tops (Fig. 1). Inequality (1.12) is satisfied inside equilateral triangles whose bases lie on the α-axis, while their vertices lie on the e-axis (Fig. 2).*

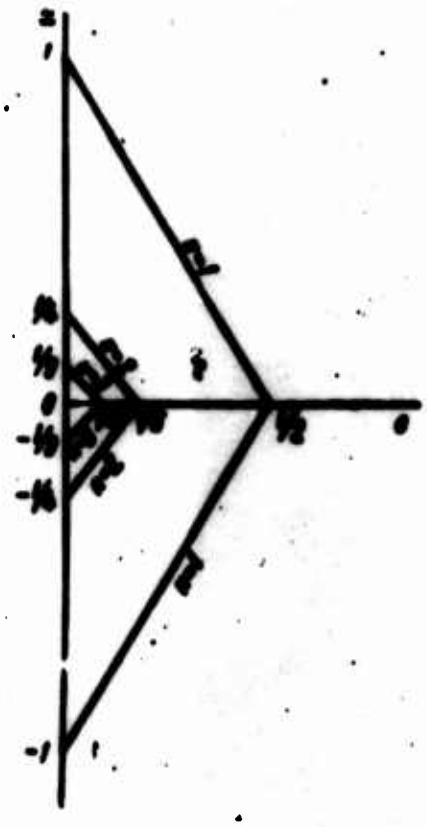


Fig. 2

2. For the sake of simplicity, we shall henceforth study the boundary-value problem that defines the 2π-periodic solution

$$\dot{x} = f(t, x), \quad (1.13)$$

$$x(0) = x(\pi) = 0, \quad (1.14)$$

where f(t, x) is written in the form (1.10).

This problem has at least one solution for all admissible values of the parameters e and α (see (1.11), where $k = 1$).

Let us show that for sufficiently small α there exists a unique solution to Problem (1.13)-(1.14) with $0 \leq e < 1$.

Let us write the equation in deviations for Eq. (1.13):

$$\begin{gathered} \ddot{s} + [a(t) + \varepsilon b(t)]s = 0, \\ s(0) = s(\pi) = 0. \end{gathered} \tag{1.15}$$

Here $s = x_2(t) - x_1(t)$; x_2, x_1 are solutions to Problem (1.13)-(1.14) corresponding to the given values of e and α;

$$a(t) = \frac{e\cos t}{1 + e\cos t}, \quad b(t) = \frac{\cos\left(\frac{t}{1 + e\cos t}\right)}{1 + e\cos t},$$

$\bar{x}$ lies between x_1 and x_2.

Let us reduce the boundary-value problem (1.15) to an integral equation

$$s - \int_0^\pi K(t, \tau)a(\tau)s\,d\tau = \varepsilon \int_0^\pi K(t, \tau)b(\tau)s\,d\tau.$$

Introducing the nomenclature

$$D(t, \tau) = K(t, \tau) + \int_0^\pi \Gamma(t, \xi)K(\xi, \tau)\,d\xi, \tag{1.16}$$

where $\Gamma(t, \tau)$ is the resolvent of the kernel $K(t, \tau)a(\tau)$, we find

$$s = \varepsilon \int_0^\pi D(t, \tau)b(\tau)s\,d\tau.$$

From this,

$$M \leqslant \varepsilon M, \tag{1.17}$$

$M = \max_t |s|$, $\varepsilon = \alpha c$, and c is a constant.

For small α, Inequality (1.17) is possible only with M = 0.

3. Let us study the behavior of solutions to Problem (1.13)-(1.14) in the region of uniqueness of (1.12), i.e., for k = 1.

Property 1. The solution, together with the initial value of the derivative, is negative. For e = 0, x(t) ≡ 0. If e ≠ 0, then x(t) ≢ 0. Let us show that for $\alpha = 0$, $e \neq 0$, $x(t) < 0$ on the interval (0, π). From (1.10) and (1.13) we find

$$\ddot{x} + [e\cos t/(1 + e\cos t)]x = 4e\sin t.$$

Denoting by $\tilde{x}_-$, $\tilde{x}_+$ solutions that satisfy the inequalities

$$\tilde{x}_- \leqslant x(t) \leqslant \tilde{x}_+$$

and the boundary conditions (1.14), we obtain

$$\ddot{\tilde{x}}_+ - \beta^2 \tilde{x}_+ = 4e \sin t, \quad \ddot{\tilde{x}}_- + \beta^2 \tilde{x}_- = 4e \sin t,$$
$$\beta^2 = e/(1-e) < 1/2.$$

Consequently,

$$\tilde{x}_+ < 0, \quad \tilde{x}_- < 0.$$

Let us suppose that the affirmation is false. Then we can find $\underline{e}$ and α such that at a certain point $t_0 \in (0, \pi)$

$$x(t_0) = 0, \quad \dot{x}(t_0) = 0, \quad \ddot{x}(t_0) \leqslant 0,$$

while it follows from (1.13) that

$$\ddot{x}(t_0) = 4e \sin t_0 > 0.$$

Property 2. With fixed $\underline{e}$, the solution, together with the initial value of the derivative, decreases with increasing α as long as $0 > x(t) > -\pi(1 + e \cos t)$.

Let $x_1(t)$ be the solution corresponding to $\underline{e}$ and α and let $x_2(t)$ be the solution corresponding to $\underline{e}$, $\alpha + \varepsilon$ (where ε is small).

Denoting

$$z(t) = x_2(t) - x_1(t),$$

we find from (1.13)

$$\ddot{z} + p(t) z = -\varepsilon \sin\left(\frac{\dot{x}_1(t)}{1 + e \cos t}\right) + R, \qquad (1.18)$$

$$p(t) = \frac{e \cos t + \alpha \cos\left(\frac{x_1(t)}{1 + e \cos t}\right)}{1 + e \cos t},$$

$$R = \varepsilon \cos\left(\frac{x_1(t)}{1 + e \cos t}\right) \frac{z}{1 + e \cos t} + \varepsilon z^2 \ldots$$

Since the solution is continuously dependent on the parameters,

$$R \sim \varepsilon^2.$$

The further reasoning proceeds in the same way as in the proof of Property 1.

2. BIFURCATION CURVE

1. If $e = 0$, then we can write the general solution to Eq. (1.13). It is found that with $e = 0$, $\alpha < 1$, the boundary-value problem (1.13)-(1.14) has a unique solution $x(t) \equiv 0$. In the case of $e = 0$, $\alpha > 1$, there exist the three solutions

$$x(t) = 0, \quad \sin\frac{x}{2} = k\,\mathrm{sn}\sqrt{\alpha}\,t, \quad \sin\frac{x}{2} = -k\,\mathrm{sn}\sqrt{\alpha}\,t,$$

$$k^2 = \frac{\dot{x}^2(0)}{4\alpha} < 1,$$

which we shall denote respectively by Π_0, Π_+, and Π_-. These solutions are the generating solutions in the continuation with respect to the parameters $\underline{e}$ and α. At the point ($e = 0$, $\alpha = 1$), all three solutions fuse (bifurcation point). The bifurcation points lie on analytical bifurcation curves in the parameter plane [7]. The solutions to Problem (1.13)-(1.14) that correspond to points of this curve will be known as multiple or branching solutions. Let us reduce the boundary-value problem (1.13)-(1.14) to the integral equation

$$x - \int_0 K(t, \tau)\left[\frac{e\cos\tau}{1+e\cos\tau}x + \alpha\sin\left(\frac{x}{1+e\cos\tau}\right)\right]d\tau + 4e\sin t = 0, \quad (2.1)$$

where $K(t, \tau)$ is a triangular kernel (see (1.9)), and apply the results of Schmidt on bifurcating solutions [8].

Let $x_2(t)$ be a solution corresponding to $e + \delta$, $\alpha + \varepsilon$, and let $x_1(t)$ be a solution corresponding to $\underline{e}$, α where δ and ε are small.

Introducing the notation

$$z(t) = x_2(t) - x_1(t),$$

we obtain from (2.1)

$$W_{100} + W_{010} + W_{001} + W_{110} + \sum_{k=2}^{\infty} W_{k00} + \sum_{k=1}^{\infty} W_{k01} = 0, \quad (2.2)$$

where $W_{mnp}(z\delta\varepsilon)$ are integroexponential forms of mth degree in $\underline{z}$, nth degree in δ and pth degree in ε, written as

$$W_{100} = z - \int_0^{\pi} K(t, \tau) p(\tau) z d\tau,$$

$$W_{010} = (q(t) x_1(t) + 4 \sin t)\delta - \delta \int_0^{\pi} K(t, \tau) q(\tau) x_1(\tau) d\tau,$$

$$W_{001} = -\varepsilon \int_0^{\pi} K(t, \tau) \sin\left(\frac{x_1(\tau)}{1 + e\cos\tau}\right) d\tau,$$

$$W_{110} = q(t) z\delta - \delta \int_0^{\pi} K(t, \tau) q(\tau) z d\tau,$$

$$W_{k00} = -\frac{\alpha}{k!} \int_0^{\pi} \left. \frac{d^k \sin\left(z + \frac{x_1(\tau)}{1 + e\cos\tau}\right)}{dz^k} \right|_{z=0} K(t, \tau) \left(\frac{z}{1 + e\cos\tau}\right)^k d\tau,$$

$$W_{k01} = -\frac{\varepsilon}{k!} \int_0^{\pi} \left. \frac{d^k \sin\left(z + \frac{x_1(\tau)}{1 + e\cos\tau}\right)}{dz^k} \right|_{z=0} K(t, \tau) \left(\frac{z}{1 + e\cos\tau}\right)^k d\tau,$$

$$p(t) = \frac{e\cos t + \alpha \cos\left(\frac{x_1(t)}{1 + e\cos t}\right)}{1 + e\cos t},$$

$$q(t) = \frac{\cos t}{1 + e\cos t}. \qquad (2.3)$$

If the equation in variations

$$W_{100} = 0 \qquad (2.4)$$

possesses a solution not identically equal to zero, then bifurcation occurs at point (e, α). Denoting the solution of Eq. (2.4) by $\eta(t)$ and by

$$E(t, \tau) = K(t, \tau) p(\tau) + \eta(t)\eta(\tau) \qquad (2.5)$$

the kernel for which Eq. (2.4) has only the trivial solution, we find from (2.2)

$$z - \int_0^{\pi} E(t, \tau) z d\tau = -\eta(t)\lambda - W_{010} - W_{001} - W_{110} -$$

$$- \sum_{k=2}^{\infty} W_{k00} - \sum_{k=1}^{\infty} W_{k01},$$

$$\lambda = \int_0^{\pi} \eta(\tau) z d\tau. \qquad (2.6)$$

If $\Gamma(t, \tau)$ is the resolvent of the kernel $E(t, \tau)$, then

$$z = \beta(t)\lambda + P_{010} + P_{001} + P_{110} + \sum_{k=2}^{\infty} P_{k00} + \sum_{k=1}^{\infty} P_{k01}, \qquad (2.7)$$

where

$$\beta(t) = -\eta(t) - \int_0^t \Gamma(t, \tau)\eta(\tau)d\tau,$$

$$\dot{P}_{mnp} = -W_{mnp} - \int_0^t \Gamma(t, \tau)W_{mnp}\,d\tau.$$

We shall seek a solution Eq. (2.7) in the form of an absolutely and uniformly converging series

$$x = \sum_{m+n+p\geq 1} \lambda^m V^m_{np}\ (\dot{t}\dot{t})\delta^n \varepsilon^p \qquad (2.8)$$

on the condition that

$$\delta \leq k_1,\ \ |\varepsilon| \leq k_2,\ \ \max_t |x| = x \leq h_1,\ \ |\lambda| \leq l_1, \qquad (2.9)$$

where k_1, k_2, h_1, l_1 are sufficiently small.

Here $V^m_{np}(\delta\varepsilon)$ is an integroexponential form of nth degree in δ and pth degree in ε.

Substituting (2.8) into (2.6), we obtain the bifurcation equation [8]

$$\sum_{m=1}^{\infty} L_m\lambda^m + \sum_{m=0}^{\infty}\lambda^m \sum_{n+p\geq 1} \int_0^t \eta(t)V^m_{np}\ (\dot{\delta\varepsilon})dt = 0,$$

$$L_m = \int_0^t V^m_{00}\ (\dot{\delta\varepsilon})\eta(t)dt \quad (m = 1,2,\ldots), \qquad (2.10)$$

where $L_1 = 1$ always.

Schmidt proved the theorem.

If L_m is the first, nonzero coefficient, then the solution $x_1(t)$ branches m times at point (ε, α).

At point $(\varepsilon = 0, \alpha = 1)$, $x_1(t) \equiv 0$. It can be shown by simple calculations that

$$L_2 = 0, \qquad L_3 = 1/\pi^2.$$

Hence the solution $x_1 \equiv 0$ actually does branch three times at point $(\varepsilon = 0, \alpha = 1)$.

Let us find the analytic form of the bifurcation curve in a small neighborhood of the point ($e = 0$, $\alpha = 1$). Let

$$S_1 = \sum_{m=0}^{\infty} L_m \lambda^m,$$

$$S_2 = \sum_{m=0}^{\infty} \lambda^m \sum_{n+p \geq 1} \int_0^{2\pi} \eta(t) V_{mnp}(\theta, t)\, dt.$$

Let us define $l_2 \leq l_1$ in (2.9) so that

$$\sigma_1 = \min_{|\lambda| = l_2} |S_1| \geq c_2 l_2^2, \tag{2.11}$$

where the constant c_2 does not depend on l_2.

Further, we introduce $k_3 \leq k_1$, $k_4 \leq k_2$ in such a way that

$$S_2 \leq c_3 k_3 + c_4 k_4 \tag{2.12}$$

where the constants c_3 and c_4 are not functions of l_2 if l_1 is used for their definition.

Denoting

$$k_3 = \alpha_1 \frac{c_2}{2c_3} l_2^2, \qquad k_4 = \alpha_2 \frac{c_2}{2c_4} l_2^2, \tag{2.13}$$

where α_1 and α_2 are regular fractions, we find from (2.11)-(2.13)

$$S_2 \leq \alpha \sigma_1, \quad \alpha = \frac{\alpha_1 + \alpha_2}{2}.$$

Whence, according to the Rushe [Roucher] theorem, in the region

$$\theta \leq k_3, \quad |\varepsilon| \leq k_4, \quad \sigma \leq k_1, \quad |\lambda| \leq l_2 \tag{2.14}$$

the bifurcation equation

$$S_1 + S_2 = 0$$

has three roots, which may coincide. Applying (2.13), we get from (2.14)

$$|\lambda| \leq \frac{1}{\sqrt[3]{c_2}} \sqrt[3]{k_3 c_3 + k_4 c_4} \quad \text{or} \quad |\lambda| \sim \sqrt[6]{(\theta^2 + \varepsilon^2)}.$$

Since

$$V_{10}^0 = \frac{8}{\pi}\sin t, \quad V_{01}^0 = 0, \quad V_{10}^1 = -\frac{3}{4\pi}\sin 2t, \quad V_{01}^1 = -\left(\frac{2}{\pi}\right)^2 \sin t;$$

then the equation of the bifurcation curve can be rewritten with an accuracy to $\sqrt[3]{(\delta^2 + \varepsilon^3)^5}$ in the form

$$\lambda^3 - 2\pi^2\varepsilon\lambda + 4\pi^3\delta = 0. \tag{2.15}$$

Let us write the discriminant of this equation:

$$D = 4\pi^6(\delta^2 - {}^2/_{27}\varepsilon^3).$$

It follows from [9] that only a single real solution of Eq. (2.2) with the same multiplicity corresponds to each root λ of multiplicity $\underline{r}$ of the bifurcation equation.

Thus, there exist, in the region E_3 on Fig. 3, three different solutions to Problem (1.13)-(1.14) ($D < 0$), two of which merge on the bifurcation curve ($D = 0$)

$$\varepsilon^3 = {}^2/_{27}(\alpha - 1)^3, \tag{2.16}$$

and only one solution in the region E_1-E_3 ($D > 0$). From (2.6) with an accuracy to $\sqrt[3]{(\delta^2 + \varepsilon^3)^2}$ we find the solution to Problem (1.13)-(1.14)

$$x(t) = \frac{2}{\pi}\lambda\sin t - \frac{3\varepsilon}{4\pi}\lambda\sin 2t + \frac{2}{27}\left(\frac{\alpha - 1}{\pi}\lambda - 2\varepsilon\right)\sin 3t,$$

where λ is determined from Eq. (2.15). In particular, the solution at the bifurcation curve is written in the form

$$x_0 = 2\sqrt[3]{2\varepsilon}\sin t - {}^3/_4\varepsilon\sqrt[3]{2\varepsilon}\sin 2t + {}^2/_{27}[(\alpha - 1)\sqrt[3]{2\varepsilon} - 2\varepsilon]\sin 3t. \tag{2.17}$$

The relationships

$$\lambda_1 + \lambda_2 + \lambda_3 = 0, \quad \frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_3} = \frac{1}{2\pi}\frac{\varepsilon}{\delta}, \quad \lambda_1\lambda_2\lambda_3 = -4\pi^3\delta$$

indicate negativity of one of the roots and positivity of the two others when $\delta \neq 0$. For $\delta = 0$, the roots of Eq. (2.15) will be

$$\lambda_1 = \pi\sqrt{2\varepsilon} > 0, \quad \lambda_2 = 0, \quad \lambda_3 = -\pi\sqrt{2\varepsilon} < 0,$$

and for $\delta \neq 0$, none of the roots vanishes. Remembering that Π_+ and Π_-

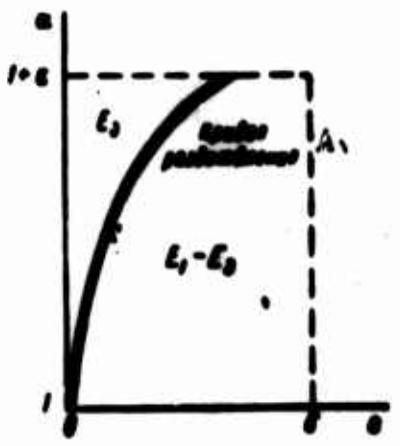


Fig. 3. A) Bifurcation curve.

bifurcate [7], we obtain finally.[sic]

For e = 0, only one bifurcation curve begins from the point (e = 0, α = 1), and it is selected in the form (2.16) in a small neighborhood of this point; the positive double solution (2.17) corresponding to this curve is formed by fusion of the solutions continued from Π_0 and Π_+.

2. Let us establish a number of properties of solutions to Problem (1.13)-(1.14).

Property 1. A solution to Problem (1.13)-(1.14) negative at (0, π) cannot become a solution that assumes a positive value on this interval with variation of e and α. If we posit the contrary, then we find for certain values of e, α a point $t_0 \in (0, \pi)$ at which

$$x(t_0) = 0, \quad \dot{x}(t_0) \leqslant 0,$$

while the right-hand member of Eq. (1.13) is positive at (0, π).

Property 2. If the solution to Problem (1.13)-(1.14) is positive at the point $t_0 \in (0, \pi)$, then it is positive in a neighborhood of it of length $l \geq \pi/\beta$, where $\beta = \sqrt{(e + |a|)/(1 - e)}$.

Conversely, let there be an interval (t_1, t_2) of length $l < \pi/\beta$, such that in it $x(t) > 0$, $x(t_1) = x(t_2) = 0$.

Performing the substitution of variables

$$s = \pi \frac{t - t_1}{t_2 - t_1}$$

we obtain the boundary-value problem

$$\frac{d^2x}{ds^2} = f_1(x, s), \quad x(0) = x(\pi) = 0,$$

where $f_1(x, s) = f(x, t(s))$.

As follows from Article 1, this problem is uniquely solvable and its solution is negative on (0, π) if

$$\left(\frac{t_2-t_1}{\pi}\right)^2 e+\left|\frac{\alpha}{1-e}\right|<1,$$

which results in a contradiction.

3. We note that the solution to Problem (0.1)-(0.2) does not attain the value π, since otherwise we should obtain at some point $v_0 \in (0, \pi)$

$$\ddot{\delta}(v_0)\leqslant 0, \quad \dot{\delta}(v_0)=0, \quad \delta(v_0)=\pi,$$

and from (0.1) we find

$$(1+e\cos v_0)\ddot{\delta}(v_0)=4e\sin v_0>0.$$

Physically, passage through the point π signifies entry into the rotational regime.

Up to the bifurcation curve under consideration, each of the solutions that fuse on it remains single. Otherwise we should have a bifurcation curve intersecting the initial curve at a point ($e \neq 0$, $\alpha \neq 1$), so that more than two solutions would merge at this point. Solutions of the problem (1.13)-(1.14) obtained by continuing Π_0, Π_+, Π_- with varying $\underline{e}$, α will be denoted respectively by x_0, x_+, x_-. Since with $\alpha > 1$, $\underline{e}$ small, $x_0>0$, $x_+>0$, $x_-<0$, we draw the following conclusions from Properties 1 and 2 of this Article:

1) $x_- < 0$ everywhere it exists,

2) $x_+ > 0$ as long as $|\alpha| \leq 4-5e$,

3) merging of x_0 and x_+ occurs over the entire bifurcation curve.

The bifurcation curve cannot terminate at the boundaries ($0 \leq e < 1$, $\alpha = 0$) and ($e = 0$, $0 \leq \alpha \leq 3$) (see Articles 1 and 2).

4. With fixed $\underline{e}$ and increasing α, x_- diminishes together with the initial derivative as long as $0>\dot{x}_->-\pi(1+e\cos t)$, and x_+ increases when $|\alpha|\leqslant 4-5e$.

The existence of a branching solution at point (e, α) is equivalent to the solution of the variation equation (2.4)

$$\ddot{\eta} + p(t)\eta = 0, \qquad (2.18)$$

where p(t) is given by Formula (2.3), with the boundary conditions $\eta(0) = 0$, $\dot{\eta}(0) = 1$, possessing on $[0, \pi]$ more than one zero [5].

This means that the Green's function $T(t, \tau)$ of the operator

$$L = \frac{d^2}{dt^2} + p(t)$$

will be nonnegative on $[0, \pi]$ if there is no bifurcation.

From (1.18) we find

$$s(t) = \int_0^\pi T(t, \tau)[\varepsilon \sin z - 0(\varepsilon^2)]d\tau,$$

from which the proposition to be proven follows directly.

5. Let us simplify the expression for the coefficient L_2, which differs from zero only on the bifurcation curve and can be used to construct it.

Since

$$V_{\varepsilon^1} = \beta(t), \quad V_{\varepsilon^2} = -\frac{\varepsilon}{2}\int_0^\pi D(t, \tau)\beta^2(\tau)\sin\left(\frac{z_1(\tau)}{1 + \varepsilon\cos\tau}\right)d\tau,$$

where $D(t, \tau)$ is written in the form (1.16), then we obtain from (2.10)

$$L_2 = \frac{\varepsilon}{2}\int_0^\pi\int_0^\pi D(t, \tau)\eta(t)\sin\left(\frac{z_1(\tau)}{1 + \varepsilon\cos\tau}\right)\beta^2(\tau)d\tau\,dt.$$

Using the equation adjoint to (2.4), we find from (2.7)

$$\dot{\beta}(t) - \int_0^\pi K(t, \tau)p(\tau)\beta(\tau)d\tau = 0.$$

Hence

$$\beta(t) = \gamma_1\eta(t), \quad \gamma_1 = \frac{1}{\|\eta\|^2}.$$

It follows from (1.16) and (2.5) that

$$D(t, \tau) - \int_0^\pi K(t, \xi)p(\xi)D(\xi, \tau)d\xi = K(t, \tau) + \eta(t)N(\tau),$$

$$\dot{N}(\tau) = \int_0^\pi \eta(\xi)D(\xi, \tau)d\xi.$$

By virtue of the solvability of the equation for D(t, τ), we have

$$\int_0^{\pi} [K(t, \tau) + \eta(t) N(\tau)] \eta(t) dt = 0.$$

From this

$$N(\tau) = -\gamma_1 \gamma_2 \eta(\tau), \quad \gamma_2 = \frac{1}{\|p\eta\|}.$$

Finally, we obtain

$$L_2 = \frac{a}{2} \gamma_1^2 \gamma_2 \int_0^{\pi} \eta^3(t) \sin\left(\frac{x_1(t)}{1 + e\cos t}\right) dt. \qquad (2.19)$$

On the bifurcation curve $L_2 > 0$, since in a small neighborhood of the point $(e = 0, a = 1)$, $x_1(t) > 0$, $\eta = \sin t > 0$, and L_2 does not change sign on the bifurcation curve.

6. Let us find the differential equation of the bifurcation curve.

We rewrite Eq. (2.2), applying the continuous dependence of the solution on the parameters

$$z - \int_0^{\pi} K(t, \tau) p(\tau) z \, d\tau = e \int_0^{\pi} K(t, \tau) \sin\left(\frac{x_1(\tau)}{1 + e\cos\tau}\right) d\tau +$$

$$+ \frac{\delta}{e}\left[\frac{x_1(t)}{1 + e\cos t} - a \int_0^{\pi} K(t, \tau) \sin\left(\frac{x_1(\tau)}{1 + e\cos\tau}\right) d\tau\right] + O(\delta^2 + e^2).$$

For solvability of this equation it is necessary to require that the right-hand member be orthogonal to $\eta(t)$.

From this it follows that

$$\frac{de}{da} = e \frac{\int_0^{\pi} \eta(t) \sin\left(\frac{x_1(t)}{1 + e\cos t}\right) dt}{\int_0^{\pi} \frac{d^2}{dt^2}\left[\frac{x_1(t)}{1 + e\cos t} + a \sin\left(\frac{x_1(t)}{1 + e\cos t}\right)\right] \eta(t) dt}. \qquad (2.20)$$

Using (2.20), we can construct any bifurcation curve as long as some point on this curve has been found.

3. FINDING SOLUTION CONTINUED FROM $x \equiv 0$ FOR $e = 0$

Let us seek a solution to Problem (0.1)-(0.2), introducing the notation $\delta = y$, $\nu = t$, in the form of the formal series

$$y(t)=\sum_{k=1}^{\infty}y_k(t)\varepsilon^k, \tag{3.1}$$

whose coefficients are found as particular solutions to a recurrent system of linear differential equations in the form

$$\begin{gathered}\ddot{y}_1+\alpha y_1=4\sin t,\\ \ddot{y}_k+\alpha y_k=2\sin t\dot{y}_{k-1}-\cos t\ddot{y}_{k-1}+\alpha f_k(y_1,\ldots,y_{k-1})\\ (k=2,3,\ldots).\end{gathered} \tag{3.2}$$

Here $f_k(y_1,\ldots, y_{k-1})$ is a homogeneous form of kth degree in $y_1, \ldots, y_{k-1}$.

The solution to System (3.2) will take the form of functions

$$y_{2k-1}=\sum_{n=1}^{k}a_{2k-1}^{(n)}(\alpha)\sin(2n-1)t,\qquad y_{2k}=\sum_{n=1}^{k}a_{2k}^{(n)}(\alpha)\sin 2nt, \tag{3.3}$$

which is proven by the complete-induction method. Let us write the first few coefficients $a_{2k-1}^{(n)}$ and $a_{2k}^{(n)}$, which are easy to calculate:

$$a_1^{(1)}=\frac{4}{\alpha-1},\qquad a_2^{(1)}=\frac{6}{(\alpha-1)(\alpha-4)},\quad a_3^{(1)}=\frac{\alpha}{\alpha-1}\frac{(a_1^{(1)})^3}{8},$$

$$a_3^{(2)}=\frac{1}{\alpha-9}\left[4a_2^{(1)}-\frac{\alpha}{24}(a_1^{(1)})^3\right];$$

$$a_4^{(1)}=\frac{1}{\alpha-4}\left[\frac{3}{2}(a_3^{(1)}+a_3^{(2)})+\frac{\alpha}{4}(a_1^{(1)})^2a_2^{(1)}\right];$$

$$a_4^{(2)}=\frac{1}{\alpha-16}\left[\frac{15}{2}a_3^{(2)}-\frac{\alpha}{8}(a_1^{(1)})^2a_2^{(1)}\right].$$

It follows from the form of (3.3) that (3.1) is the Fourier series of the function y(t). Hence it converges absolutely and uniformly in the uniqueness region (1.12), where k = 1, and as far as the bifurcation curve with $\alpha > 1$.

An advantage of obtaining the solution in the form of a series in the parameter ε instead of a Fourier series consists in the fact that if the coefficients $y_k(t)$ are determined with sufficient ease from linear differential equations, then the coefficients of the Fourier series will be found from an infinite system of nonlinear algebraic equations with an infinite number of unknowns.

The author thanks V.A. Sarychev and Yu.F. Sotnikova for the great assistance that they rendered in completing this work.

Received

10 June 1964

REFERENCES

1. V.V. Beletskiy, collection entitled "Iskusstvennyye sputniki Zemli" [Artificial Earth Satellites], No. 3, Izd-vo AN SSSR [Academy of Sciences USSR Press], 1959, page 13.
2. V.V. Beletskiy, collection entitled "Iskusstvennyye sputniki Zemli," No. 16, Izd-vo AN SSSR, 1963 page 46.
3. F.L. Chernovs'ko, Zh. vychisl. mat. i matem. fiz. [Journal of Computer Mathematics and Mathematical Physics], 3, No 3, 528, 1963.
4. G.G. Khardi, D.Ye. Littl'vud, G. Polia. Neravenstva [Inequalities], Moscow, Izd-vo inostr. lit. [Foreign Literature Press], 1948 page 219.
5. A. Hammerstein, Acta Math., 54, 117, 1930.
6. F. Trikomi. Integral'nyye uravneniya [Integral Equations] Izd-vo inostr. lit. [Foreign Literature Press] Moscow, 1960, page 257.
7. R. Iglisch. Monatsh. für Math und Phys. [Monthly Bulletins for Mathematics and Physics], 39, 173, 1932.
8. E. Schmidt. Math. Ann., 65, 370, 1908.
9. R. Iglisch, Jahresb. D.M.V. [Annual Report of the German Mathematical Society], 39, 65, 1930.

Manuscript Page No.

[Footnote]

20 The existence conditions (1.11) and uniqueness conditions

(1.12) that we have found for the solutions are easily extended to solutions of Eq. (1.13), which represent a superposition of uniform rotation and oscillations of the form $x(t) = bt + y(t)$, where b is a real number and $y(t)$ is an odd periodic function of period $2\pi k$ ($k = 1, 2, \ldots$).

UDK 517.912.2:521.3

APPROXIMATE CALCULATION OF TRAJECTORY FOR ENTRY INTO ATMOSPHERE

V.A. Yaroshevskiy

II

Use of an approximate equation for the motion of a space vehicle in the atmosphere [1] enables us to examine the influence of lift and initial velocity on the characteristics of atmosphere-entry trajectories.

III. ATMOSPHERE ENTRY TRAJECTORIES OF VEHICLES USING LIFT.

9. The Case of Small Lift-to-drag Ratio

Let us rewrite Eq. (3.17) from [1] in the form

$$y'' = -K + \frac{e^{2x} - 1}{y}, \qquad (9.1)$$

where

$$y = \frac{c_x S}{2m}\sqrt{\frac{R}{\lambda}}\rho, \quad x = \ln\frac{V_{кр}}{V} = \ln\frac{\sqrt{Rg}}{V}, \quad \theta = -\frac{1}{\sqrt{R\lambda}}y'$$

$$n = \sqrt{R\lambda}\sqrt{1 + \left(\frac{c_y}{c_x}\right)^2}\, y e^{-2x},$$

$$L = \sqrt{\frac{R}{\lambda}}\int_{x_{нач}}^{x}\frac{dx}{y}, \quad t = \frac{1}{\sqrt{g\lambda}}\int_{x_{нач}}^{x}\frac{e^x dx}{y},$$

$K = \sqrt{R\lambda}c_y / c_x$ is a quantity proportional to the lift-to-drag ratio (for the earth $\sqrt{R\lambda} \approx 30$). Let the value of K be of the order of a few units and let the angle of entry into the atmosphere be close to zero: $y'(0) = 0, y(0) = 0$. The solution is constructed in the form of a series in powers of $x^{1/2}$:

$$y = \sqrt{\frac{6}{5}}x^{5/2} - \frac{4K}{11}x^3 + \sqrt{\frac{6}{5}}\left(\frac{1}{6} + \frac{K^2}{121}\right)x^{7/2} + \left(\frac{8K}{27\cdot 11} + \frac{16K^3}{9\cdot 1331}\right)x^4 +$$

$$+ \sqrt{\frac{6}{5}}\left(\frac{1}{24} + \frac{13}{6\cdot 19\cdot 121}K^2 + \frac{41K^4}{25\cdot 11^4}\right)x^{9/2} + \dots \qquad (9.2)$$

With $|K| < 3$ (for the earth $0.1 < c_y/c_x < 0.1$), the terms of the series that we have written out represent with sufficient accuracy the solution in the interval of velocities in which the g-force and heat flows reach their maximum values.

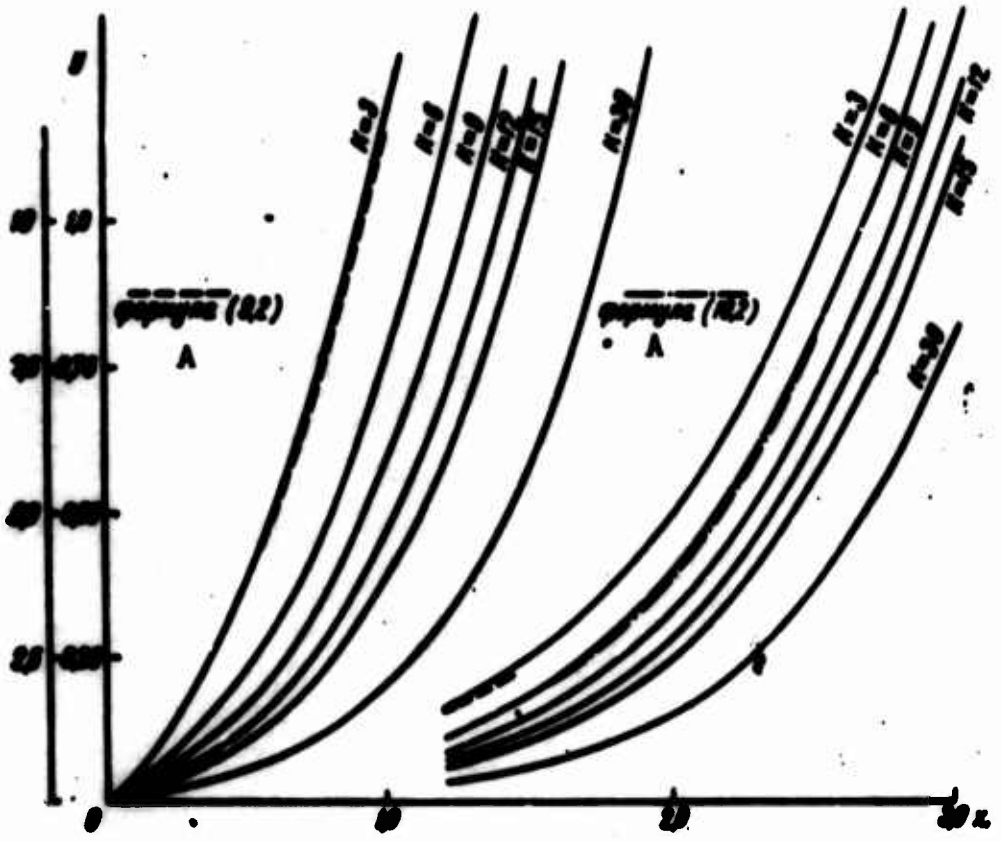


Fig. 1. A) Formula.

If the initial inclination angle of the trajectory is not too small ($y'(0) = c_1 > 1$), the solution is constructed in the form of a series in x:

$$y = c_1x + c_2x^2 + c_3x^3 + c_4x^4 + \dots, \qquad (9.3)$$

where

$$c_2 = \left(\frac{1}{c_1} - \frac{K}{2}\right), \quad c_3 = \frac{1}{2c_1}\left(\frac{K}{2c_1} + 1 - \frac{1}{c_1^3}\right),$$

$$c_4 = \frac{1}{12c_1}\left[\frac{4}{3} - \frac{1}{2c_1}\left(\frac{K}{2c_1} + 1 - \frac{1}{c_1^3}\right)\left(\frac{8}{c_1} - 3K\right)\right].$$

The radius of convergence of this series is not large, and the

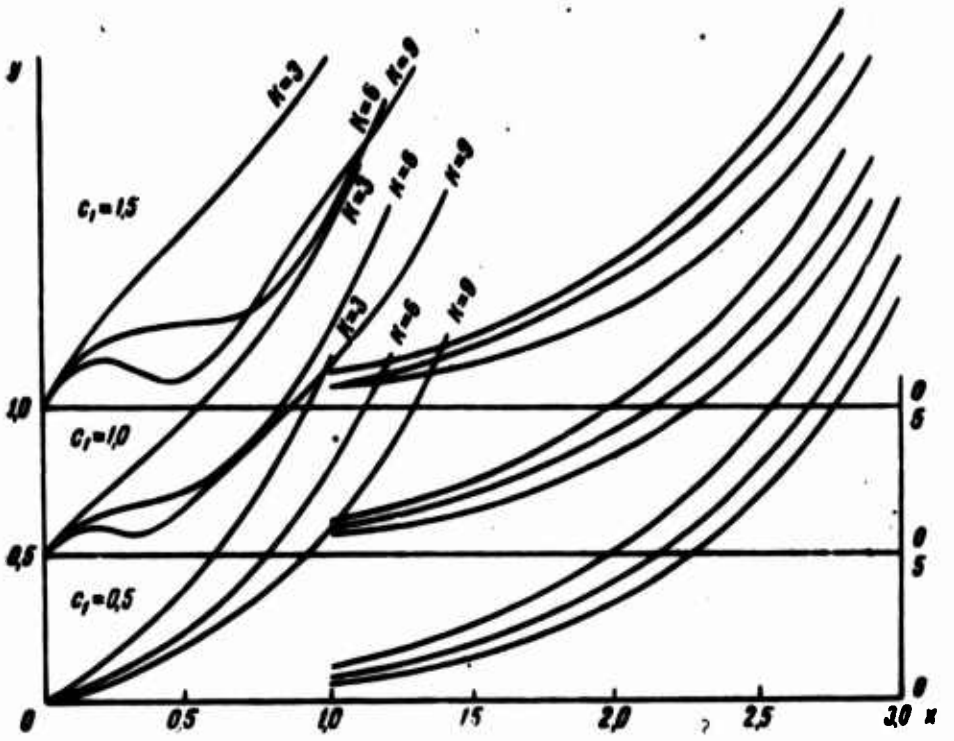


Fig. 2

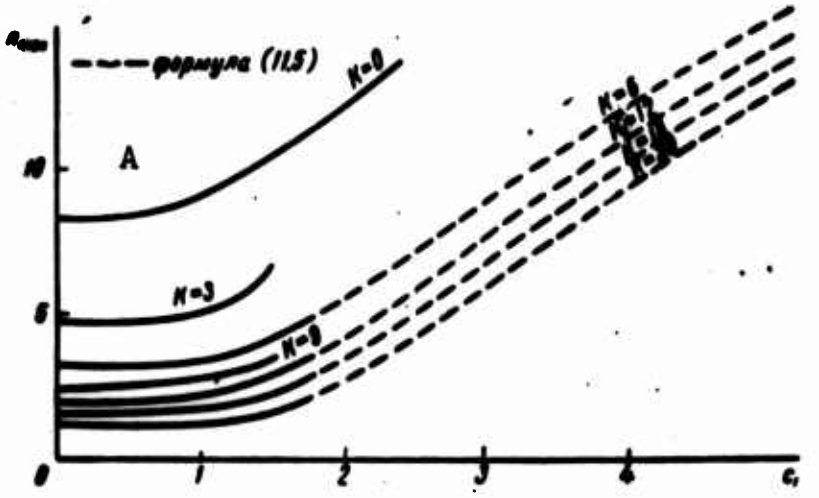


Fig. 3. A) Formula

maximum values of the g-forces and heat flows can be determined by its use only for very small K.

The results of calculations by Eq. (9.1) are presented in Figs. 1-3. As we see, the maximum overload values diminish substantially even for very small positive values of c_y/c_x.

10. Glide Paths

If the vehicle has a high lift-to-drag ratio and the angle of entry into the atmosphere is small, then we have the so-called quasistationary glide trajectory [2-6].

On this trajectory, the term y" in (9.1) is relatively small in magnitude.

Then

$$y \approx y_{pl} = (e^{2x} - 1)/K \qquad (10.1)$$

(the subscript "pl" stands for "planirovaniye" [gliding]).

It is shown with little difficulty that the assumption that the term y" is small is compromised for small $\bar{V}$, where the solution can be represented approximately in the form

$$y \approx [(e^{2x} - 1)/K]f(e^x/K). \qquad (10.2)$$

The function f(z) is a solution to the equation

$$s^4f''(s) + 5s^3f'(s) + 4s^2f(s) = -1 + 1/f(s) \qquad (10.3)$$

for the initial conditions

$$f(0) = 1, \quad f'(0) = 0.$$

As we see, this function is close to 1 only for $e^x/K = s < 0{,}1$ (Fig. 4), i.e., (for the earth), for $\bar{V}(c_y/c_x) > 0.3$.

For small s $f(s) \approx 1 - 4s^2$, and for large s $f(s) \approx 1/s - 1/s^2$.

Representing the solution in this form enables us to judge deviations of the true trajectory from the quasistationary glide trajectory and decreases in the magnitude of the maximum g-force as compared with the value

$$n_{max}|_{\bar{V}=0} = \sqrt{1 + (c_x/c_y)^2}, \qquad (10.4)$$

calculated for the trajectory (10.1).

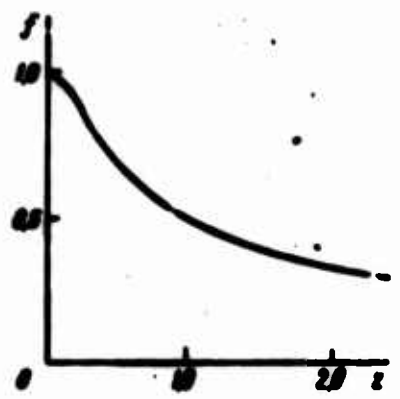


Fig. 4

Formula (10.2) enables us to analyze trajectories with $K \geq 6$ ($c_y/c_x > 0.2$). With $K \geq 30$, we may regard Formula (9.2) as valid. Thus, Formulas (9.2), (10.1) and (10.2) enable us to cover the entire interval of lift-to-drag ratio values (Fig. 1) for small angles of entry into the atmosphere.

The expressions for the distance and time of flight along the

glide trajectory take the form [3]

$$L = \frac{R}{2}\frac{c_y}{c_x}\ln\frac{1-(V/V_{кр})^2}{1-(V_{нач}/V_{кр})^2}, \qquad (10.5)$$

$$t = \frac{1}{2}\sqrt{\frac{R}{g}}\frac{c_y}{c_x}\ln\frac{[1-(V/V_{кр})^2][1+(V_{нач}/V_{кр})^2]}{[1+(V/V_{кр})^2][1-(V_{нач}/V_{кр})^2]}. \qquad (10.6)$$

Small deviations from the quasistationary glide trajectory [7] $\Delta y = y - y_{pl}$ are given by the equation

$$\Delta y'' + \frac{K^2}{e^{2x}-1}\Delta y = 0, \qquad (10.7)$$

whose approximate solution can be determined using the asymptotic method [8]

$$\Delta y \sim \frac{1}{\sqrt{\omega(x)}}\sin\int_0^x \omega(x)\,dx,$$

where

$$\omega(x) = \frac{K}{\sqrt{e^{2x}-1}}.$$

From this it follows firstly that

$$\frac{\Delta y}{y_{нач}} \sim (e^{2x}-1)^{\frac{3}{4}} \to 0 \quad \text{as} \quad x \to \infty,$$

i.e., $\Delta H \to 0$ as $V \to 0$, and, secondly, that the number of complete oscillations n on the trajectory of descent from orbit is about $1/4K$ ($n \approx 7.5c_y/c_x$ for the earth), since

$$n = \frac{\int_0^\infty \omega(x)\,dx}{2\pi} = \frac{K\arccos e^{-x}\Big|_0^\infty}{2\pi} = \frac{K}{4}. \qquad (10.8)$$

11. Segment of First Descent

A distinctive property of the trajectory with large positive values of the lift-to-drag ratio consists in the fact that a tendency arises for the vehicle to ricochet from dense layers of the atmosphere.

With $K \gg 1$, $K \gg 1/c_1$ it is sufficient for calculating the segment of first descent into the atmosphere to restrict ourselves in (9.3) to the

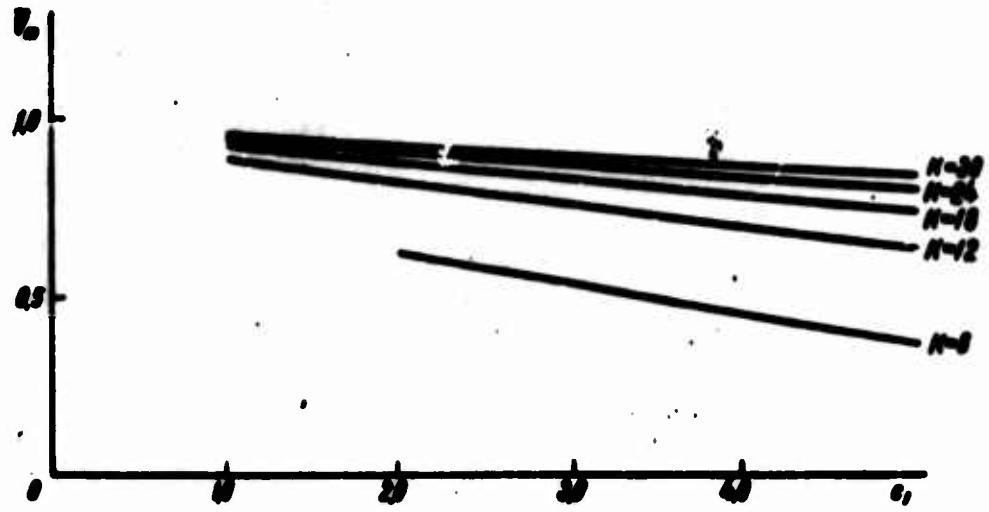


Fig. 5

terms

$$y = c_1 x - Kx^2/2, \qquad (11.1)$$

which is consistent with the formula given in [5]. Here the g-force maximum is attained, for all practical purposes, at the instant at which H_{min} is reached. If K or c_1 is small, then the altitude diminishes monotonically.

With $c_y/c_x > 0{,}3$, $|\theta_{max}| > 2°$ (for the earth), bounces always occur. In this case, the series (9.3) converges rapidly on the segment of first descent into the atmosphere and four terms of this series give a good approximation of the solution. Let us denote the instant corresponding to $H_{min}(\theta = 0)$ by t_m. To determine the velocity $\bar{V}_m$, we use the relation $\theta = 0$:

$$y' = c_1 + 2c_2 x + 3c_3 x^2 + 4c_4 x^3 = 0. \qquad (11.2)$$

From this we obtain in approximation

$$x_m \approx -\frac{c_1}{2c_2} - \frac{3}{8}\frac{c_3 c_1^2}{c_2^3} - \frac{c_1^3}{16c_2^4}\left(9\frac{c_3^2}{c_2} - 4c_4\right). \qquad (11.3)$$

From these values of x_m, we determine $y_m = y(x_m)$ and $\bar{V}_m = e^{-x_m}$ (Figs. 5 and 6). The maximum g-force values on the segment of first descent can also be determined in a similar fashion:

$$x(n_{max}) = x_n = -\frac{\varepsilon_1}{2\varepsilon_2} - \frac{3}{8}\frac{\varepsilon_2\varepsilon_1^2}{\varepsilon_2^3} - \frac{\varepsilon_1^3}{16\varepsilon_2^4}\left(9\cdot\frac{\varepsilon_2^3}{\varepsilon_3} - 4\varepsilon_4\right) + \ldots, \quad (11.4)$$

where

$$\varepsilon_m = \sum_{k=0}^{m-1} c_{m-k}\frac{(-2)^k}{k!}.$$

The maximum total-g-force values are (Fig. 3)

$$n_{max} = 30y(x_n)e^{-2x_n}\sqrt{1 + (c_y/c_x)^2}. \quad (11.5)$$

In calculating the distance and time of flight in the case in which $c_1 > 1$, it is sufficient to select the initial value $x_{nach} = y_{nach}/c_1$ (see [1], Article 6).

Let us again use the series expansions of the functions

$$L' = \sqrt{\frac{R}{\lambda}}\frac{1}{y} = \sqrt{\frac{R}{\lambda}}\left(\frac{f_0}{x} + f_1 + f_2x + f_3x^2\right), \quad (11.6)$$

where

$$f_0 = \frac{1}{c_1} \qquad f_m = -\frac{1}{c_1}\sum_{k=0}^{m-1} f_k c_{m-k+1}.$$

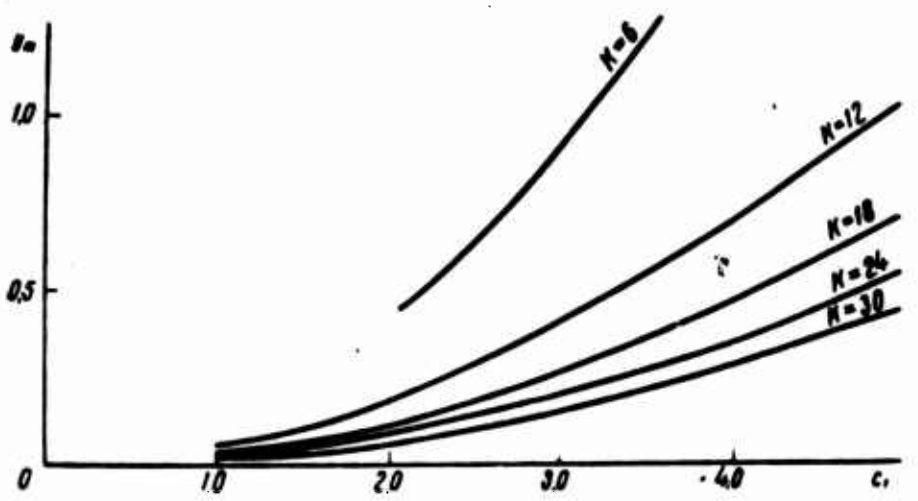


Fig. 6

On integrating, we find that for $x \gg x_{nach}$

$$L = \sqrt{\frac{R}{\lambda}}\left(f_0\ln\frac{c_1x}{y_{нач}} + f_1x + \frac{f_2x^2}{2} + \frac{f_3x^3}{3} + \ldots\right). \quad (11.7)$$

Similarly, we can obtain a formula for the time of flight on the ;ment first descent:

$$t=\frac{1}{\sqrt{g\lambda}}\left(f_0\ln\frac{c_1 x}{y_{max}}+f_1 x+\frac{f_2x^2}{2}+\frac{f_3x^3}{3}+\ldots\right), \qquad (11.8)$$

:re

$$f_m=\sum_{k=0}^{m} f_{m-k}\frac{1}{k!}.$$

One of the most commonly encountered systems for controlling the nosphere-entry trajectory is the following [9-11]: after the vehicle s reached its minimum altitude for the first immersion into the at-sphere, the lift coefficient is changed in such a way as to produce rizontal flight:

$$c_y=\frac{2G}{\rho S}\left(\frac{1}{V^2}-\frac{1}{V_{xp}^2}\right). \qquad (11.9)$$

As the velocity diminishes on the horizontal segment, the value c_y increases until it has become equal to the available $c_{y\ max}$ at point in time t_p. On reaching the speed $V_p=1/\sqrt{1+Ky_m}$ the vehicle ain comes down on a trajectory closely similar to the gliding trajec-ry (10.2); with large c_1 and K,

$$V_p\approx\frac{1}{\sqrt{1+{}^1/_2 c_1^2}},$$

e., it does not depend on K.

The distance and time of flight on the horizontal trajectory are lculated by the formulas

$$L=\sqrt{\frac{R}{\lambda}}\frac{x_p-x_m}{y_m}; \qquad t=\frac{1}{\sqrt{g\lambda}}\frac{e^{x_p}-e^{x_m}}{y_m} \qquad (11.10)$$

ovided that c_x is independent of c_y).

Let us construct the maximum g-force values as a function of the try angle and lift-to-drag ratio for trajectories with a horizontal

flight segment.

On the segment of first descent into the atmosphere, the g-force rises for $x < x_n$, but diminishes with $x > x_n$ and continues to diminish on the horizontal segment with $x_n < x_m < x < x_p$; thereafter, on the last segment, the g-force either diminishes or increases to reach a maximum corresponding to the trajectory of the quasistationary glide.

From this we may conclude that in a certain interval of small entry angles ($0 < |\theta_{nach}| < 3^\circ$ for the earth), the maximum g-force will undergo almost no change, while with large entry angles it is identical to the values of $n_x'x_n$) for the segment of initial descent into the atmosphere (Fig. 3).

The effect of reducing the maximum g-force by the use of positive lift becomes less perceptible at large entry angles.

12. Trajectory With Rebounds

If the lift is not controlled and the parameters c_1 and K are not very small, then, having come down to an altitude $H_{min}(y_m)$, the vehicle will rebound, begin the next stage in its descent, and so forth. A trajectory of this type is known as a rebound trajectory [3, 9, 12-15]. Oscillations about the trajectory of the quasistationary glide take plane in the (H, V) plane.

If

$$\left| \frac{\rho - \rho_{пл}}{\rho_{пл}} \right| \ll 1,$$

we can use the equation in variations, (10.7).

To calculate such trajectories in nonlinear formulation, we use the asymptotic method or the averaging method [16, 17].

Let us examine Eq. (9.1):

$$\varepsilon^2 y'' = -K(x) + \frac{e^{2x} - 1}{y},$$

assuming that the terms $e^{2x} - 1$ and K(x) vary only slightly over the

course of a single oscillation period, for which purpose we formally place the coefficient ε^2 before the term y", with ε the smallness parameter.

We present the solution for y in the form

$$y = y_0(x, \varphi) + \varepsilon y_1(x, \varphi) + O(\varepsilon^2), \qquad (12.1)$$

where

$$\varphi' = \frac{\omega(x)}{\varepsilon} + O(1),$$

i.e., φ is the "fast-shifting" phase and y_0 and y_1 are functions periodic with respect to φ with a period 2π.

Then it is not difficult to show that the function $y_0(x, \varphi)$ satisfies a reference equation with the value of x "frozen" [17]:

$$\omega^2(x)\frac{\partial^2 y_0(x, \varphi)}{\partial \varphi^2} = -K + \frac{e^{2x} - 1}{y_0(x, \varphi)}. \qquad (12.2)$$

For a given amplitude $y_0(x, \varphi)$, the quantity $\omega(x)$ is determined from the condition that the period of the oscillations with respect to φ be equal to 2π.

We shall restrict ourselves to first-approximation determination of the envelopes $y_{max}(x)$ and $y_{min}(x)$, which correspond to the maximum and minimum densities reached in the ricocheting process, i.e., the amplitude values of $y_0(x, \varphi)$ from the periodicity condition of the function $y_1(x, \varphi)$

$$\int_{y_{min}}^{y_{max}} \frac{dy}{dx}\, dy = \text{const}, \qquad (12.3)$$

where dy/dx is determined on solution of the "frozen" equation (12.1):

$$\frac{dy}{dx} = \sqrt{2[Ky_{max} - (e^{2x} - 1)\ln y_{max} - Ky + (e^{2x} - 1)\ln y]}. \qquad (12.4)$$

The value of y_{min} is connected with y_{max} by the relationship

$$K(y_{max} - y_{min}) = (e^{2x} - 1)\ln\frac{y_{max}}{y_{min}}. \qquad (12.5)$$

It follows from (12.3) and (12.4) that

$$\frac{(e^{2x}-1)^{1/2}}{K}h(u)=\text{const}, \tag{12.6}$$

where

$$u=\frac{y_{max}}{y_{ux}}=\frac{Ky_{max}}{e^{2x}-1}>1$$

is the ratio of the amplitude density to the density on the glide trajectory,

$$h(u)=\int_{u_{min}}^{u}\sqrt{u-z+\ln\frac{z}{u}}\,dz, \tag{12.7}$$

and u_{min} and $\underline{u}$ are connected by the relationship

$$u-u_{min}+\ln\frac{u_{min}}{u}=0.$$

For u ~ 1,

$$h(u)\approx\frac{\pi}{2\sqrt{2}}(u-1)^2,\quad u_{min}\approx 2-u,$$

and for u >> 1

$$h(u)\approx\sqrt{2}u^{1/2},\quad u_{min}\approx ue^{-u}.$$

Solving (12.6) for $\underline{u}$, we obtain

$$y_{max}=\frac{e^{2x}-1}{K}h^{-1}\left\{\frac{K}{K_{max}}\left(\frac{e^{2x_{max}}-1}{e^{2x}-1}\right)^{1/2}h(u_{max})\right\}. \tag{12.8}$$

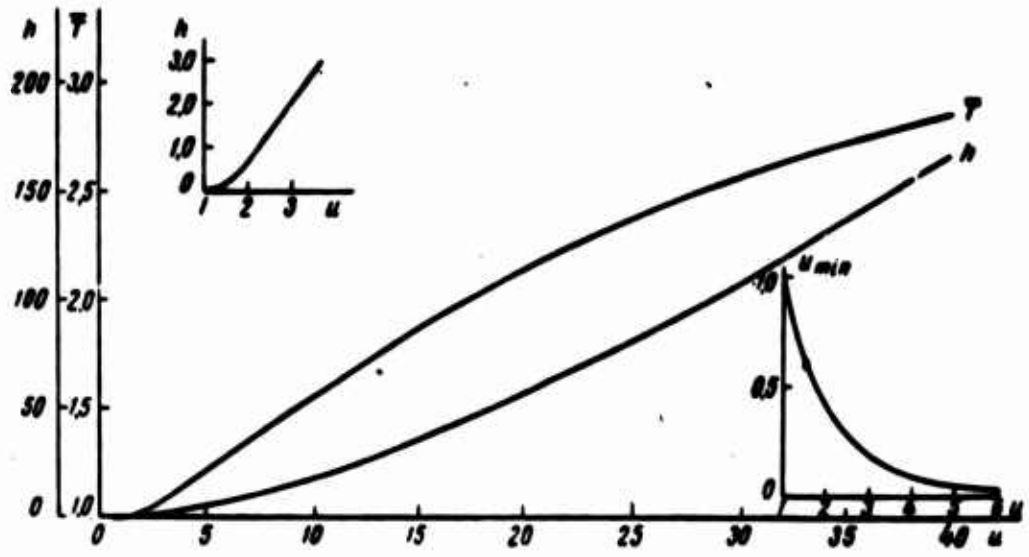


Fig. 7

Taking t_m as the initial point in time, we have

$$s_{max} = s_m,$$
$$u_{max} = \frac{K_{max} y_m}{e^{2s_{max}} - 1}.$$

For a constant lift-to-drag ratio, we can draw the following conclusions:

1. With $s \gg 1$, $s \sim 1/(e^{2s} - 1)$, i.e., $y_{max} \approx$ const, and $y_{min} \sim e^{-2s}$, which takes us to the system examined in [3].

2. As velocity decreases, the rebound trajectory tends to approach the quasistationary glide trajectory $u \to 1$; here, the amplitude $y - y_{pl}$ increases in proportion to $\sqrt{e^{2s} - 1}$, and the amplitude $(y - y_{pl})/y_{pl}$ diminishes in proportion to $(e^{2s} - 1)^{-1/4}$.

The period of the oscillations with respect to $\underline{x}$ is determined by the formula

$$T = 2\int_{y_{min}}^{y_{max}} \frac{dy}{dy/dx} = \frac{2\pi\sqrt{e^{2s} - 1}}{K} T(u), \qquad (12.9)$$

where

$$T(u) \to 1 \quad \text{as } u \to 1, \quad T(u) \approx \sqrt{2u}/\pi \text{ as } u \gg 1.$$

Since $\overline{T}(u) > 1$, we can satisfy ourselves that the number $\underline{n}$ of complete oscillations on a rebound trajectory does not exceed $\frac{1}{4} K$ (for the earth, 7.5 c_y/c_x). The functions h(u) and $\overline{T}(u)$ are shown in Fig. 7. A comparison of the results from calculating the envelope of the rebound trajectory by Formula (12.8) with the results obtained directly from Eq. (9.1) is given in Fig. 8.

In calculating the distance and time of flight on the rebound trajectory, we can get

$$L \approx \sqrt{\frac{\pi}{\lambda}} \left\{ \int_{y_{min}}^{s} \frac{K(s)}{e^{2s} - 1} ds + \frac{1}{e^{2s} - 1} \frac{dy}{dx} \Big|_{y_{min}}^{s} \right\} + O\left(\frac{1}{K}\right) \qquad (12.10)$$

$$t = \frac{1}{\sqrt{g\lambda}} \left\{ \int_{x_{max}}^{x} \frac{K(x)e^x}{e^{2x}-1}dx + \frac{e^x}{e^{2x}-1}\frac{dy}{dx}\Big|_{x_{max}}^{x} \right\} + O\left(\frac{1}{K}\right). \qquad (12.11)$$

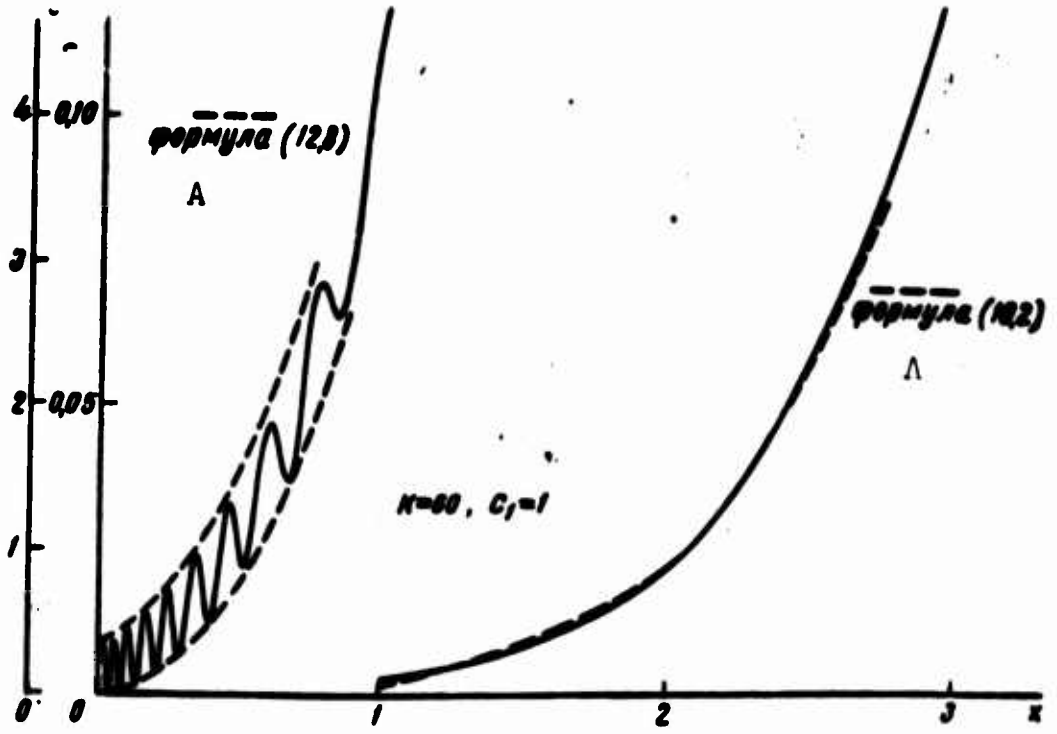


Fig. 8. A) Formula

For the case in which the drag coefficient is variable, we have

$$L \approx R \int_0^{V_{max}} \frac{c_y(V)dV}{c_x(V)V\left(\frac{Rg}{V^2}-1\right)} + \frac{\theta_{max}V_{max}}{g-\frac{V_{max}^2}{R}}. \qquad (12.12)$$

We note that the maximum g-forces on the horizontal-flight trajectories do not exceed the maximum g-forces on rebound trajectories. Further, trajectories with a horizontal segment are found to be more favorable in view of the shorter time and distance of the flight.

The decrease in distance and time on a trajectory with a horizontal segment can be evaluated applying (11.10), (11.11), (12.10) and (12.11) and taking into account that with $x < x_m$, these trajectories do not differ:

$$\Delta L = \sqrt{\frac{R}{\lambda}}\left(\frac{K}{2}\ln\frac{1-V_p^2}{1-V_m^2} - \frac{x_p - x_m}{y_m}\right),$$

$$\Delta t = \frac{1}{\sqrt{g\lambda}}\left[\frac{K}{2}\ln\frac{(1-V_p^2)(1+V_m^2)}{(1+V_p^2)(1-V_m^2)} - \frac{\frac{1}{V_p}-\frac{1}{V_m}}{y_m}\right],$$

IV. ATMOSPHERE-ENTRY TRAJECTORIES WITH VELOCITIES GREATER THAN THE CIRCULAR VELOCITY

13. Calculation of First Phase of Atmosphere-Immersion Segment

If the speed of entry into the atmosphere exceeds the circular velocity, then the initial conditions are more conveniently assigned in terms of the initial velocity $V_E = v_E/\sqrt{Rg}$ and the value of a fictitious perigee parameter $y_p = (c_x S/2m)\sqrt{R}/\lambda\rho_p$, where ρ_p is the density of the atmosphere at the point of a fictitious perigee of the flight trajectory calculated without considering the influence of the atmosphere [18].

The altitude H_p of the fictitious perigee is determined from the relationship

$$\frac{H_E - H_p}{R} \approx \frac{V_E^2 \theta_E^2}{2(V_E^2 - 1)}. \qquad (13.1)$$

The initial conditions for the atmosphere entry trajectory are the values of H_E, V_E and θ_E at the "boundary of the atmosphere."

Applying (13.1), we can obtain initial conditions for (9.1): $x_{nach} = -\ln \bar{V}_E$, $y_{nach} > 0$ is a small quantity,

$$y'_{nach} = \sqrt{2\left(1 - \frac{1}{V_E^2}\right)\ln\frac{y_p}{y_{nach}}}. \qquad (13.2)$$

If $\bar{V}_E \neq 1$, then $y' \to \infty$ as $y \to 0$. Hence it is now possible to construct the solution directly in the form of a power series in $x - x_{nach}$. However, certain relationships can be obtained for determination of the most important parameters.

Such an approach to solution of the problem is used in [13, 14, 19, 20]. Thus, we shall assume that the velocity $\bar{V}_m$ corresponding to the minimum perigee altitude of immersion into the atmosphere is not too greatly different from $\bar{V}_E$. We obtain from (9.1)

$$\frac{y'^2}{2}\Big|_{x_{nach}} = -Ky\Big|_{y_{nach}} + \int_{y_{nach}}^{y} \frac{e^{Kx_{nach}} - 1}{y}\,dy + \int_{y_{nach}}^{y} \frac{e^{Kx} - e^{Kx_{nach}}}{y}\,dy. \qquad (13.3)$$

Taking the initial conditions, we can rewrite this last relationship in the form

$$\frac{y'^2}{2} = -Ky + (e^{2x_{nach}} - 1)\ln\frac{y}{y_0} + \int_0^y \frac{e^{2x} - e^{2x_{nach}}}{y}\,dy \qquad (13.4)$$

or, with $x = x_m$, when $y'(x_m) = 0$,

$$(1 - e^{2x_{nach}})\ln\frac{y_m}{y_0} = -Ky_m + \int_0^{y_m} \frac{e^{2x} - e^{2x_{nach}}}{y}\,dy. \qquad (13.5)$$

Since the last term in (13.4) and (13.5) is of the order of $x - x_{nach}$, it can be calculated in approximation, with an accuracy to $0[(x - x_{nach})^2]$ on the assumption that

$$e^{2x} - e^{2x_{nach}} \approx 2e^{2x_{nach}}(x - x_{nach})$$

$$y' \approx \sqrt{2(1 - e^{2x_{nach}})\ln\frac{y_m}{y} + 2K(y_m - y)}. \qquad (13.6)$$

Integrating by parts, we obtain

$$\int_0^{y_m} \frac{e^{2x} - e^{2x_{nach}}}{y}\,dy \approx 2e^{2x_{nach}} \int_0^{y_m} \frac{\ln\frac{y_m}{y}\,dy}{\sqrt{2(1 - e^{2x_{nach}})\ln\frac{y_m}{y} + 2K(y_m - y)}} =$$

$$= \sqrt{\frac{2}{1 - e^{2x_{nach}}}}\, e^{2x_{nach}} y_m \int_0^1 \frac{\ln\frac{1}{s}\,ds}{\sqrt{\ln\frac{1}{s} + \eta(1 - s)}},$$

where

$$\eta = \frac{Ky_m}{1 - 1/V_z^2}$$

or

$$\ln\frac{y_z}{y_m} = \eta - \frac{\sqrt{2}\,y_m}{\left(1 - \frac{1}{V_z^2}\right)^{1/2} V_z^2} f_1(\eta), \qquad (13.7)$$

where

$$f_1(\eta) = \int_0^1 \frac{\ln\frac{1}{s}\,ds}{\sqrt{\ln\frac{1}{s} + \eta(1 - s)}}.$$

The value of $\bar{V}_m = e^{-x_m}$ is determined by means of the approximate formula

$$x_m \approx x_{max} + \int_0^{y_m} \frac{dy}{\sqrt{2K(y_m - y) + 2(1 - e^{2x_{max}}) \ln \frac{y_m}{y}}}$$

or

$$\ln \bar{V}_m = \ln \bar{V}_B - f_3(\eta) \frac{y_m}{\sqrt{2\left(1 - \frac{1}{\bar{V}_B^2}\right)}},$$

where

$$f_3(\eta) = \int_0^1 \frac{ds}{\sqrt{\ln \frac{1}{s} + \eta(1 - s)}}. \qquad (13.8)$$

For $\eta \geq 0$, the functions $f_1(\eta)$ and $f_2(\eta)$ can be approximated with high accuracy by semiempirical formulas similar to [20]

$$f_1(\eta) = \frac{\sqrt{\pi}}{2\sqrt{1 + \frac{\pi}{[8(1 - \ln 2)]^2}\eta}} \approx \frac{\sqrt{\pi}}{2\sqrt{1 + \frac{\pi}{6}\eta}},$$

$$f_2(\eta) \approx \frac{\sqrt{\pi}}{\sqrt[3]{1 + \frac{\pi}{4}\eta}}.$$

The maximum g-force is reached at a value $x_n < x_m$. To determine this value, let us write the solution for y in the neighborhood of $x = x_m$:

$$y \approx y_m + y''(x_m)\frac{(x - x_m)^2}{2} + \ldots =$$

$$= y_m - \left\{K + \frac{1 - e^{2x_m}}{y_m}\right\}\frac{(x - x_m)^2}{2} + \ldots.$$

We find from the maximum condition of the function ye^{-2x} that

$$x_n = x_m - \sqrt{\frac{1}{4} + \frac{2y_m}{K + (1 - e^{2x_m})/y_m}} + \frac{1}{2}, \qquad (13.9)$$

$$n_{max} = \sqrt{R\lambda}\sqrt{1 + \left(\frac{c_y}{c_x}\right)^2}\, y(x_n)e^{-2x_n}.$$

To justify the use of the approximate formulas (13.7), (13.8) and (13.9), let us consider two extreme cases: K = C and $Ky_m \gg 1$. In the former case, utilization of these formulas is equivalent to taking into account the following terms in the expansions

$$y_m = y_p(1 + \alpha_1 y_p + \ldots), \qquad \bar{V}_m = \bar{V}_E(1 + \beta_1 y_p + \ldots),$$
$$n_{max} = n(x_m)[1 + \gamma_2 y_p^2 + \ldots].$$

As we see, accounting for the difference $x_m - x_n$ (the coefficient γ_2) results in a correction of the second order of smallness in the determination of n_{max}, although this correction may still be found significant, particularly for $\bar{V}_m$ close to unity; hence it should be taken into account in making practical calculations. In the latter case, with $Ky_m \gg 1$, taking the second terms in the right members of Relationships (13.7) and (13.8) into account together with the difference between x_m and x_n in determining the maximum g-force is equivalent to allowance for errors with relative magnitudes of the orders of $1/K\gamma\overline{Ky_m}$, $\sqrt{y_m/K}$ and y_m/K, respectively. Consequently, as K increases, all of these corrections diminish and the correction in the determination of $\bar{V}_m$ is the most significant among them.

14. Segment of Climbout From Atmosphere

We shall use the term "atmosphere climbout segment" to denote the climbing segment that follows after the point $x = x_m$. Approximate formulas for this segment have been derived for two cases: 1) the vehicle's speed of climbout from the atmosphere, $\bar{V}_{Ex}$, is markedly in excess of the circular velocity; 2) this speed is equal to the circular velocity.

There is practically no difference between the technique for calculating the climbout segment and that presented above, except that the sign before the second term in Eq. (13.7) changes:

$$\ln\frac{y_{p_1}}{y_m} = \eta_1 + \frac{\sqrt{2}\,y_m}{(1 - 1/\bar{V}_{Ex}^2)^{1/4}\,\bar{V}_{Ex}^3} f_1(\eta_1), \qquad (14.1)$$

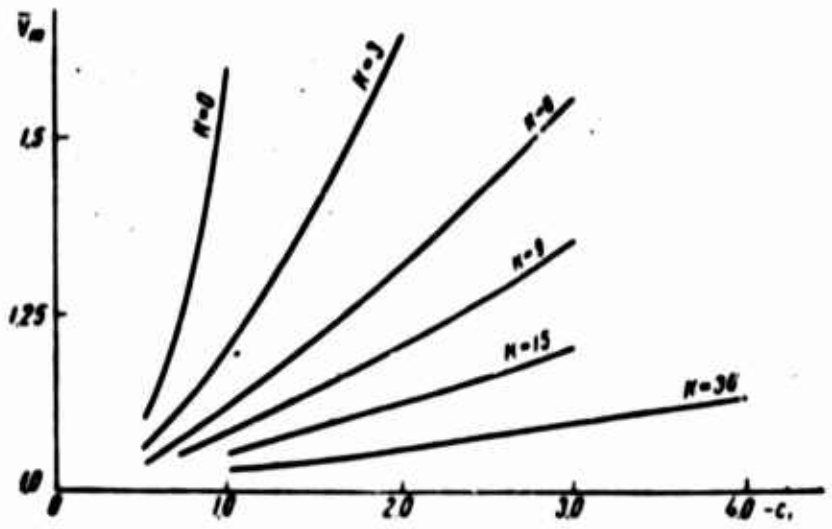


Fig. 9

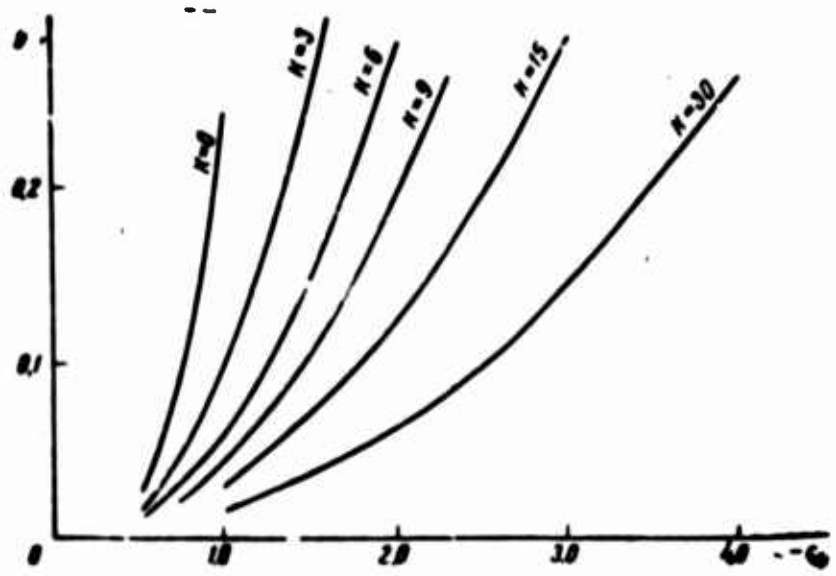


Fig. 10

where y_{p_1} is the new value of the fictitious-perigee parameter after climbout from the atmosphere:

$$\eta_1 = \frac{K y_m}{1 - 1/V_{ex}^2},$$
$$\ln V_{ex} = \ln V_m - \frac{y_m}{\sqrt{2\left(1 - \frac{1}{V_{ex}^2}\right)}} f_2(\eta_1). \qquad (14.2)$$

Formulas (14.1) and (14.2) are valid for the condition that the value of V_{Ex} is not too close to 1.

In the latter case, the technique employed is that of constructing the solution in the form of a series in powers of $\underline{x}$ for $x < 0$, beginning at the point $x = 0$, $y = 0$, at which, by definition, the trajectory

of climbout from the atmosphere terminates.

Values of the derivative at the moment of climbout, $c_1 = y'(0) = -\sqrt{R\lambda}\,\theta_{Ex} < 0$ and values of K are assigned. The resulting series is identical to Series (9.3), except that we are considering values of $x < 0$ and $c_1 < 0$. The minimal immersion altitude and other parameters are determined in the same way as in Subsection 10. The results of the calculation are presented in Figs. 9 and 10. With $K \geq 0$ and small $K < 0$, they give adequate accuracy if $n_{max} < 20$. By "gluing" the values given for (x_m, y_m) to the values of (x_m, y_m) determined using Relationships (13.7) and (13.8), we can find trajectories corresponding to the vehicle's climbout from the atmosphere at transcircular speed (injection of the vehicle into a stationary orbit around the planet). With high negative K, the trajectories corresponding to the altitude-gaining and climbout segment are close to the trajectories of quasistationary gliding (Subsection 11):

$$y \approx \frac{e^{2x} - 1}{K}.$$

15. Determination of Atmosphere Entry Corridor for Vehicles with Lift

The width of an entry corridor is defined as the difference between the altitudes of the fictitious perigee of a trajectory corresponding to the "limit of capture" (escape from the atmosphere at transcircular speed) with the condition that negative lift is used and that of the trajectory on which the maximum g-force coincides with the maximum permissible value, with the condition that positive lift is used [18]:

$$\Delta H_p = H_{p_{max}} - H_{p_{min}} = \frac{1}{\lambda} \ln \frac{\rho_{p_{max}}}{\rho_{p_{min}}} = \frac{1}{\lambda} \ln \frac{y_{p_{max}}}{y_{p_{min}}}. \qquad (15.1)$$

The value of $y_{p_{max}}$ is determined by the value assigned to the maximum permissible g-force n_{max} with Relationships (13.7), (13.8) and (13.9). In the general case, with $c_y/c_x = 0$ or with very small values of the

lift-to-drag ratio K < 3, the maximum g-force of the order of 10 that is attainable in the first immersion into the atmosphere may be exceeded on the subsequent segment of the motion (with subcircular velocity), so that the value of $y_{p_{max}}$ must be determined taking the absolute g-force maximum into account. With K > 3, however, the absolute g-force maximum in motion along the lower boundary of the atmosphere is reached on the very first immersion into the atmosphere or, on the other hand, subsequent g-force maxima can be made smaller than the first by controlling the lift, with the result that Formulas (13.7), (13.8) and (13.9) are found to be adequate to define the lower boundary of the corridor.

With even greater K > 7.5 and moderately large $y_m(n_{max})$, we may consider that

$$y_p \approx y_m e^{\beta K y_m/(1-1/\bar{V}_E^2)},$$

$$n_{max} \approx \sqrt{R\lambda}\sqrt{1+(c_y/c_x)^2}\,\bar{V}_m^2 y_m,$$

$$\bar{V}_m \approx \bar{V}_E e^{(-\alpha\beta y_m/\sqrt{(1-1/\bar{V}_E^2)+(\alpha K y_m)^2})},$$

where the exponent in the last formula is small.

We can derive the following formula on this basis:

$$y_{p_{max}} \approx \frac{n_{max}}{\sqrt{R\lambda}\sqrt{1+(c_y/c_x)^2}\bar{V}_E^2} e^{(\mu+\nu\sqrt{(1+\mu^2)\nu^2+(\mu^2+1)/\alpha^2 K^2})}, \qquad (15.2)$$

where

$$\mu = \frac{n_{max}}{(\bar{V}_E^2-1)\sqrt{1+(c_x/c_y)^2}}, \qquad (15.3)$$

$$\nu = \frac{n_{max}}{\bar{V}_E\sqrt{\bar{V}_E^2-1}\sqrt{R\lambda}\sqrt{1+(c_y/c_x)^2}}. \qquad (15.4)$$

The upper boundary of the corridor or the value of $y_{p_{min}}$ is determined from the condition that the vehicle be captured by the atmosphere. In this case, when the available negative lift-to-drag ratio is not too small (|K| > 7.5), $y_{p_{min}}$ is determined comparatively simply: it can be shown that with this value $\bar{V}m$ is close to the value $\bar{V}_E$, and that the

A ΔH_p, км

n_{max}	$\bar{V}_E = 1,2$		$\bar{V}_E = 1,6$		$\bar{V}_E = 2,0$		c_y/c_x
	[18]	ф-ла (15.6) B	[18]	ф-ла (15.6) B	[18]	ф-ла (15.6) B	
5	37,2	41	12,3	12,7	2,8	4,5	
10	84,3	77,6	26,6	25,8	12,3	14	0,25
20	190	156	53,5	49,3	27,8	27,4	
5	84	85	29,7	30,5	16,5	17,4	
10	163	163,1	54,5	54,6	31	31,9	1,0
20	337,5	324,4	102,3	100,2	56,8	57,2	
5	100	104,5	34,9	37,8	19,4	22	
10	185,6	192,6	62	65,5	35,6	36,7	4,0
20	369	379	114,8	117,3	65,1	67,9	

A) ΔH_p, km; B) formula.

negative lift with $x = x_m$ should compensate the difference between the centrifugal force and the force of gravitation. From this it follows that at the limit of capture $\eta = -1$ or

$$y_{p_{min}} \approx \frac{1 - 1/\bar{V}_E^2}{|K|e}. \qquad (15.5)$$

Applying (15.2) and (15.5), we can derive an approximate formula for the width of the entry corridor:

$$\Delta H_p \approx \frac{1}{\lambda}\left\{\mu + \ln\mu + 1 + \sqrt{2\pi}\frac{1+\mu}{\sqrt{1+{}^1/_4\pi\mu}}\nu\right\}, \qquad (15.6)$$

where μ and ν are determined by Formulas (15.3) and (15.4). A comparison of the results from calculation by Formula (15.6) with the "exact" results of [18] is given in the table. For very large $c_y/c_x > 2$ $\nu \to 0$,

$$\mu \to \frac{n_{max}}{\bar{V}_E^2 - 1},$$

and then

$$\Delta H_{p_{пред}} \approx \frac{1}{\lambda}\left\{\frac{n_{max}}{\bar{V}_E^2 - 1} + \ln\frac{n_{max}}{\bar{V}_E^2 - 1} + 1\right\} \qquad (15.7)$$

When μ drops to values smaller than 1, the entry corridor vanishes very quickly, since with $\mu < 1$, satisfaction of the condition that the difference between the centrifugal and gravitational force in movement along the capture boundary be compensated results in the appearance of a g-force that exceeds the maximum permissible value.

V. THE INVERSE PROBLEM: DETERMINATION OF THE COEFFICIENTS c_y AND c_x FROM A GIVEN TRAJECTORY

16. Determination of c_y

The problem of determining c_y from a given trajectory in the (H, V)-plane with the condition that c_x be independent of c_y is solved as follows.

From the equations

$$\frac{dV}{dt} = -\frac{X(V,H)}{m} - g\sin\theta, \tag{16.1}$$

$$\frac{dH}{dt} = V\sin\theta, \tag{16.2}$$

where

$$X(V,H) = \frac{c_x S\rho V^2}{2},$$

we find

$$\sin\theta = -\frac{X(V,H)}{mV(g/V + dV/dH)},$$

and

$$\cos\theta\left(H, V, \frac{dV}{dH}\right).$$

Combining Eq. (16.2) with the equation

$$V\frac{d\theta}{dt} = \frac{c_y S\rho V^2}{2m} - \cos\theta\left(g - \frac{V^2}{R+H}\right), \tag{16.3}$$

we obtain

$$c_y = \frac{mV^2\frac{d\cos\theta}{dH} - m\left(g - \frac{V^2}{R+H}\right)\cos\theta}{\frac{S\rho V^2}{2}} =$$

$$= c_y\left(V, H, \frac{dV}{dH}, \frac{d^2V}{dH^2}\right) = c_y(V).$$

If we limit ourselves to a simplified equation of motion on the main segment [1], we get

$$c_y = \left\{ \frac{1/\bar{V}^2(x) - 1}{y(x)} - \frac{d^2y(x)}{dx^2} \right\} \frac{c_x(1)}{\gamma \bar{R} \cdot \lambda},$$

where

$$x = -\int_1^{\bar{V}} \frac{d\bar{V} c_x(1)}{\bar{V} c_x(\bar{V})}, \quad y = \frac{c_x(1) S}{2m} \sqrt{\frac{\bar{R}}{\lambda}} \rho,$$

and the assigned relationship H(V) has been transformed into a relationship y(x).

If the trajectory is given in the form $\rho = \rho(V, t)$ or $y = y(x, t)$, we first solve the equation [1]

$$\frac{dt}{dx} = \frac{1}{\sqrt{g\lambda}} \frac{1}{y(x, t) V(x)} \tag{16.4}$$

from which we find t(x) and y[x, t(x)], after which the problem is the same as the preceding one.

17. Determination of c_x

Let us determine the relationship $c_x(V)$ from a given trajectory $\rho(V)$ for the condition that c_y be independent of c_x. Using the assumptions listed in [1], we write the simplified equations of motion, introducing the independent variable s(ds = Vdt):

$$\frac{dV}{ds} = -\frac{S}{2m} c_x(V) V \rho(V), \tag{17.1}$$

$$\frac{d^2H}{ds^2} = \frac{S}{2m} c_y(V) \rho(V) - \frac{g}{V^2} + \frac{1}{R}. \tag{17.2}$$

At the same time,

$$H = -\frac{1}{\lambda} \ln \frac{\rho(V)}{\rho_0},$$

from which, applying (17.1), we get

$$\frac{d^2H}{ds^2} = -\frac{S^2}{4m^2\lambda} \frac{d}{dV} \left[\frac{d\rho(V)}{dV} c_x(V) V \right] c_x(V) V \rho(V). \tag{17.3}$$

Equating the right-hand members of (17.2) and (17.3), we find that

$$c_x(V) = \frac{2m\sqrt{2\lambda}\sqrt{\int_{V_0}^{V}\left[-\frac{S}{2m}c_y(V)\rho(V)+\frac{g}{V^2}-\frac{1}{R}\right]\frac{[d\rho(V)/dV]dV}{\rho(V)}}}{SV[d\rho(V)/dV]}$$

The problem of determining $c_x(V)$ for $c_y = 0$ and an assigned relationship $n_x(V)$ is solved in a similar manner. For example, for the case $n_x = \text{const}$, we obtain

$$c_x(V) = \frac{2mn_xg}{SV^2\rho[H(V)]};$$

where H(V) is determined from the equation

and

$$\frac{d^2H}{dz^2} = -\frac{V_{sp}^2}{4n_x^2g}\left(\frac{1}{z}-1\right),$$

$$z = \left(\frac{V}{V_{sp}}\right)^2.$$

Received

1 June 1964

REFERENCES

1. V.A. Yaroshevskiy, Kosmich. issled. [Cosmic Research], 1, No. 4, 507, 1964.
2. K.A. Ehricke, Astronautica Acta, 1, No. 3, 137, 1955.
3. A.J. Eggers, H.J. Allen, S. E. Neice, NACA Report, No. 1382, 1959.
4. C. Gazley. Vistas in Astronautics, 1, 8, 1958.
5. D.R. Chapman, NACA TR, R-11, 1959.
6. W.H. T. Loh. ARS Journal, 30, No. 2, 201, 1960.
7. I.S. Campbell, ARS Journal, 29, No. 7, 525, 1959.
8. A. Erdeyi, Asimptoticheskiye razlozheniya [Asymptotic Expansions], Fizmatgiz [State Publishing House for Physicomathematical Literature], 1962.
9 A. Ferry, L. Feldman, W. Daskin, Jet Propulsion, 27, No. 11, 1184,

1957.

10. D.I. Kepler, ARS Paper, No. 786, 1959.
11. B.A. Galman, Planetary and Space Sciences, 4, 399, 1961.
12. T.R. Nonweiler, Aeronautics, 38, No. 5, 41, 1958.
13. E. Bendor, C. Kattler, A. Krenkel, Proceedings of the Fifth Symposium on Ballistic Missile and Space Technology, IV, 89, 1960.
14. K. Wang, Lu Ting, ARS Journal, 30, No. 6, 565, 1960.
15. L. Lees, F.W. Hartwig, C.B. Cohen, ARS Paper, No. 785, 1959.
16. V.M. Volosov, Zh. vychisl. mat. i matem. fiz. [Journal of Computer Mathematics and Mathematical Physics], 3, No. 1, 3, 1963.
17. G.Ye. Kuzmak. Prikl. matem. i mekh. [Applied Mathematics and Mechanics], 21, No. 2, 262, 1957.
18. D.R. Chapman. NASA TR, R-55, 1959.
19. W.H. I. Loh. ISA Paper, No. 61-116, 1961.
20. Ph. Levine, IAS Paper, No. 62-162, 1962.

Manuscript Page No.	[Transliterated Symbols]
35	кр = kr = kriticheskiy = critical
35	нач = nach = nachal'nyy = initial
38	пл = pl = planirovaniye = gliding
55	пред = pred = predel'nyy = limiting

UDK 629.197.7

CALCULATING SPACE-VEHICLE TRANSFER TRAJECTORIES BETWEEN COPLANAR CIRCULAR ORBITS IN A NEWTONIAN GRAVITATIONAL FIELD

V.A. Il'in

The sphere-of-influence method is used in an examination of the problem of determining space-vehicle trajectories of transfer between coplanar circular orbits in a Newtonian gravitational field. Simple relationships are derived between the eccentricities and focal parameters of single- and double-impulse transfer trajectories that satisfy the condition of constancy of the transfer characteristic velocity or of the angle between the initial and final radius vectors. Examples are given in which the results obtained are used to solve a number of problems of interorbital transfer and rendezvous of circumterrestrial spacecraft and problems of interplanetary transfers.

In problems of interorbital transfer, the time of motion and angular displacements of the space vehicle are examined as the basic conditions defining the trajectories of the space craft. In contrast to the classical problems of celestial mechanics, however, the energy considerations of transfer, which are usually specified in the form of the characteristic velocity, emerge as one of the determining factors in problems of astrodynamics.

To calculate the characteristics of a plane orbital transfer, it is necessary to determine two elements of the orbit, for example, the focal parameter p and the eccentricity e. In most studies that have been devoted to interorbital transfers, the technique for calculating such transfers is based on the Euler-Lambert equation [1-3], whose use leads to considerable complication of the energy relationships.

The present paper proposes a technique for calculating interorbital transfers by spacecraft without using the Euler-Lambert equation – a method based on direct consideration of the restrictions imposed on the characteristic velocity and the angular range of the flight.

STATEMENT OF PROBLEM

Let us consider the transfer of a vehicle through space from an ISZ [Artificial Earth Satellite] orbit to an ISP [Artificial Planet Satellite] orbit on the following assumptions.

1. The planets' orbits are circular and coplanar, and the transfer trajectory lies in the plane of the planets' orbits.

2. The motion of the vehicle is examined successively in the sphere of influence of the earth, on the heliocentric segment and in the sphere of influence of the planet.

3. In examining the heliocentric segment of the transfer, the initial and final points of the transfer arc are considered to coincide with the centers of the corresponding planets.

4. For acceleration and deceleration of the vehicle at the earth and the planet, respectively, thrust impulses are imparted to the apparatus at certain points on the satellite orbits.

5. The characteristic velocity

$$\Delta v_{\text{зад}} = \int_0^t |\mathbf{a}|\, dt,$$

where $\vec{a}$ is the acceleration of the vehicle by gravity, is adopted as the transfer energy characteristic.

It is noted that if we exclude the influence of the gravitational fields of the earth and the planet, then the scheme under consideration will correspond to transfer between circular orbits in a Newtonian gravitational field.

TRANSFERS WITH CONSTANT CHARACTERISTIC VELOCITY

Let us consider transfer from the orbit of an artificial planet satellite to a planetocentric hyperbola or back with impulse acceleration of the space vehicle, assuming that the vector $\vec{v}_{sf}$ of the vehicle's planetocentric velocity is assigned at the planet's sphere of influence (Fig. 1). As we know, the minimum velocity increment is achieved by imparting the thrust impulse at the pericenter of the orbit in the direction of the vehicle's velocity impulse.

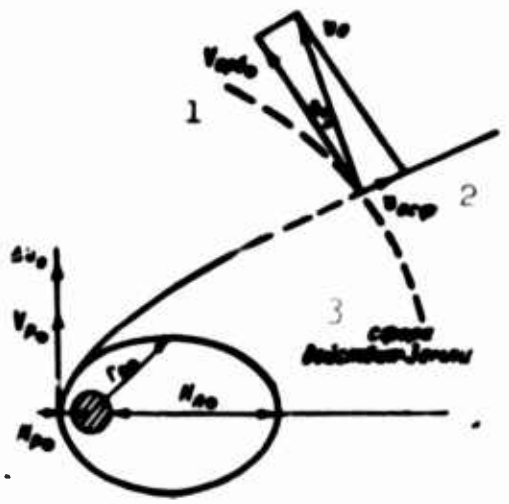


Fig. 1. 1) $\vec{V}_{orb_\bullet}$; 2) $\vec{v}_{Osf_\bullet}$; 3) earth's sphere of influence.

As a result, the expression for the velocity increment at the orbit pericenter of an artificial earth satellite may be presented in the form

$$\Delta v_{0_\bullet} = \sqrt{v_{p_\bullet}^2 + V_{cp_\bullet}^2 + v_{c\Phi_\bullet}^2 - 2\frac{K_\bullet}{r_{c\Phi_\bullet}}} - v_{p_\bullet}, \qquad (1)$$

where

$$V_{cp}^2 = V_\bullet^2 \frac{1}{1+(H_{cp_\bullet}/R_\bullet)}, \quad v_{p_\bullet}^2 = V_{cp_\bullet}^2 \frac{1+(H_{A_\bullet}/R_\bullet)}{1+(H_{P_\bullet}/R_\bullet)} \qquad (2)$$

In Formulas (1) and (2), $K_\bullet = fM_\bullet$, $M_\bullet$ is the mass of the earth, f is the gravitational constant, $R_\bullet$ is the mean radius of the earth, $V_\bullet$ is the 1st cosmic [circular] velocity, $V_\bullet = (K_\bullet/R_\bullet)^{1/2}$, $H_{P_\bullet}$, $H_{A_\bullet}$ are altitudes of perigee and apogee of the ISZ orbit,

$$H_{cp_\bullet} = (H_{P_\bullet} + H_{A_\bullet})/2,$$

$V_{sr_\bullet}$ is the velocity of the ISZ on a circular orbit with altitude H_{sr}, $v_{p_\bullet}$ is the perigee velocity of the ISZ, $r_{sf_\bullet}$ is the radius of the earth's

sphere of influence and $v_{sf_\bullet}$ is the geocentric velocity of the vehicle at the boundary of the earth's sphere of influence. For a circular artificial satellite orbit $v_{p_\bullet} \equiv V_{sr_\bullet}$.

Let us now express the quantity v_{sf} in terms of the parameters p, e of the space vehicle's heliocentric trajectory segment. From the vector equality (Fig. 1) $\vec{v}_{sf} = \vec{v} - \vec{V}_{orb}$, we obtain

$$v_{c\bullet}^2 = V_{op\bullet}^2 + v^2 - 2V_{op\bullet}v\cos\theta, \qquad (3)$$

where V_{orb} is the velocity of the planet's motion along a heliocentric circular orbit of radius R_{orb} and θ is the angle of inclination of the vector $\vec{v}$ to the local transversal (direction along the vector $\vec{V}_{orb}$).

Expressing the energy integrals and the moment of inertia in terms of p and e and substituting the resulting expressions in (3), we get

$$v_{c\bullet}^2 = V_{op\bullet}^2\{3 - 2\sqrt{p/R_{op\bullet}} + [(e^2-1)/p]R_{op\bullet}\}.$$

Finally, we shall have the following expression for determing the values of v_{sf} at the initial and final points:

$$v_{c\bullet}'^2 = (f/n)\{3 - 2\sqrt{p'}/n + n[(e^2-1)/p']\}. \qquad (4)$$

In Formula (4), $p' = p/R_{orb_\bullet}$ is the dimensionless focal parameter, $v'_{sf} = v_{sf}/v_{orb_\bullet}$ is the dimensionless velocity at the sphere of influence, and $n = R_{orb}/R_{orb_\bullet}$ is the ratio of the radii of the initial and final orbits.

We note that to determine v'_{sf} on the initial orbit, i.e., $v'_{sf_\bullet}$, it is sufficient to set $n = 1$ in (4).

Let us now consider a two-impulse flight from an ISZ orbit to an ISP orbit, for example, an earth-to-Mars transfer.

The characteristic velocity of this transfer is

$$\Delta V_{01} = \Delta v_0 + \Delta v_1, \qquad (5)$$

where Δv_0 is determined by Expression (1) and the impulse velocity incre-

ment at the pericenter of the Mars-satellite orbit, Δv_1, is given by the analogous relationship

$$\Delta v_1 = \sqrt{v_{p\sigma}^2 + V_{cp\sigma}^2 + v_{c\Phi\sigma}^2 - 2(K_\sigma/r_{c\Phi\sigma})} - v_{p\sigma}. \qquad (6)$$

Substituting (1) and (6) in (5) and introducing the nomenclature

$$\begin{aligned} A &= (\Delta V_{01} + v_{p\oplus} + v_{p\sigma})/V_{cp\oplus}, \\ B^2 &= \left[v_{p\oplus}^2 + V_{cp\oplus}^2 - 2(K_\oplus/r_{c\Phi\oplus})\right]/V_{cp\oplus}^2, \\ C^2 &= \left[v_{p\sigma}^2 + V_{cp\sigma}^2 + 2(K_\sigma/r_{c\Phi\sigma})\right]/V_{cp\oplus}^2, \end{aligned} \qquad (7)$$

we recast Equality (5) in the form

$$A = \sqrt{B^2 + v_{c\Phi\oplus}^2} + \sqrt{C^2 + v_{c\Phi\sigma}^2},$$

from which, using (4), we may obtain

$$e^2 = 1 + a_1 p' + a_{1/2} p'^{1/2} + a_2 p'^2, \qquad (8)$$

where

$$a_1 = \frac{1}{4}\left(A\alpha + \frac{\beta}{A}\right)^2 - B^2 - 3, \quad a_{1/2} = 2 - \alpha\gamma - \frac{\beta\gamma}{A^2}, \quad a_2 = \frac{\gamma^2}{A^2}. \qquad (9)$$

$$\alpha = 1 + \frac{B^2 - C^2}{A^2}, \quad \beta = 3\left(1 - \frac{1}{n}\right), \quad \gamma = 1 - \frac{1}{n^{1/2}}. \qquad (10)$$

Relationship (8) is an equation of isoenergetic curves – lines of constant characteristic velocity of two-impulse transfers between the spheres of influence of two arbitrary planets the ratio of whose orbital radii is $\underline{n}$, in the plane of the parameters (p', e) for the heliocentric segment of the transfer.

If it is assumed that the orbits of the artificial satellites are at infinitely great distances from the centers of the planets, for example, if for the earth and Mars

$$H_{p\oplus} = H_{A\oplus} = H_{p\sigma} = H_{A\sigma} = \infty,$$

then $\Delta V_{01} = v_{c\Phi\oplus} + v_{c\Phi\sigma}$ and in Formulas (9)-(10) we must set

$$A = \Delta V_{01}/V_{cp\oplus}, \quad B = C = 0, \quad \alpha = 1. \qquad (11)$$

Relationships analogous to (8) can also be obtained for single-impulse transfers between orbits. Thus, in the case of a thrust impulse

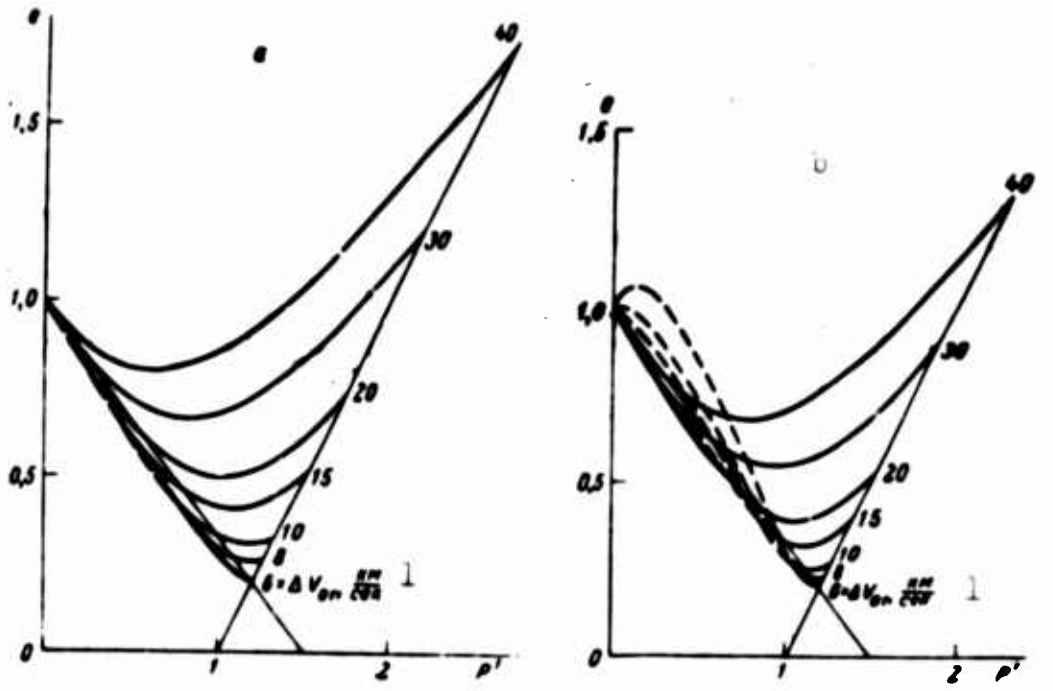


Fig. 2. 1) km/sec.

at the ISZ orbit ($\Delta V_{01} = \Delta v_0$, $\Delta V_1 \equiv 0$) the equation of the lines $\Delta V_{01} =$ = const takes the form

$$e^2 = 1 + a_1' p' + 2p'^{3/4}, \qquad (12)$$

where

$$a_1' = A'^2 - B^2 - 3, \qquad A' = (\Delta v_0 + v_{p\oplus}) / V_{op\oplus}, \qquad (13)$$

and in the case of a thrust impulse at an artificial Mars satellite orbit ($\Delta V_{01} = \Delta v_1$, $\Delta v_0 \equiv 0$)

$$e^2 = 1 + a_1'' p' + a_{3/4}'' p'^{3/4}, \qquad (14)$$

where

$$a_1'' = A''^2 - C^2 - \frac{3}{n}, \quad a_{3/4}'' = \frac{2}{n^{3/4}}, \quad A'' = \frac{\Delta v_1 + v_{p\sigma}}{V_{op\sigma}}. \qquad (15)$$

We note that a transfer between circular orbits in the field of attraction of a given planet is analogous to transfer between ISP orbits infinitely remote from the centers of attraction.

In investigating the resulting equations, we shall limit ourselves to the case of transfer to a superior planet ($n > 1$)* and examine a two-impulse earth-to-Mars transfer. The region of admissible parameters for earth-to-Mars transfers on the plane (p', e) is bounded [4] by the

line

$$e = p' - 1, \tag{16}$$

which corresponds to conic sections tangent at the pericenter to the orbit of the earth, and by the line

$$e = 1 - (p'/n), \tag{17}$$

which corresponds to ellipses tangent to the orbit of Mars at the apocenter.

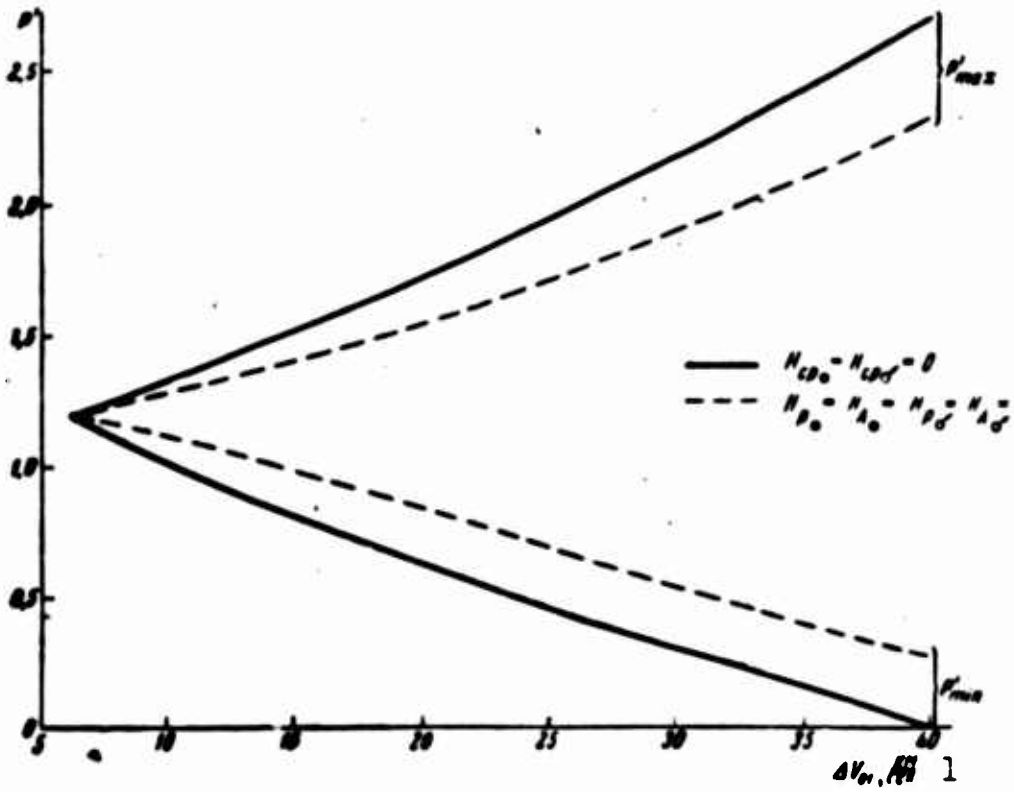


Fig. 3. 1) km/sec. ср = sr = average.

Figure 2 shows the results of calculating e = e(p') curves with ΔV_{01} = const from Eq. (8) for the two extreme cases of orbit height

$$H_{cp\,\odot} = H_{cp\,\sigma} = 0 \text{ (Fig. 2a)},$$
$$H_{P\,\odot} = H_{A\,\odot} = H_{P\,\sigma} = H_{A\,\sigma} = \infty \text{ (Fig. 2b)}$$

for the condition

$$r_{cp\,\odot} = r_{cp\,\sigma} = \infty.$$

It follows from the above data that the nature of the curves ΔV_{01} = = const is practically independent of the heights of the ISZ and ISP orbits in the region of p', e values under consideration. An analysis in-

dicates that the segments of the curves drawn with the dashed line in Fig. 2 have no physical significance.

The curves (8) intersect the boundary lines (16) and (17) at points for which the values of $p'^{1/2}$ are determined from the following equations:

for the boundary $e = p' - 1$

$$(a_2 - 1)p' + a_{1/2}p'^{1/2} + a_1 + 2 = 0, \qquad (18)$$

for the boundary $e = 1 - (p'/n)$

$$[a_2 - (1/n^2)]p' + a_{1/2}p'^{1/2} + a_1 + (2/n) = 0. \qquad (19)$$

Analysis of the roots of Eqs. (18) and (19) indicates that of the two real roots for each equation it is necessary to take the positive root closest to the point of intersection of Lines (18) and (19), which corresponds to a Hohmann transfer [4]. Let us denote the squares of these roots by p'_{max} (ΔV) and p'_{min} (ΔV), respectively. Diagrams of values of p'_{max} (ΔV) and p'_{min} (ΔV) are given in Fig. 3 for an earth-to-Mars transfer.

Similar results are also obtained for single-impulse transfers by use of Relationships (12) and (14).

TRANSFERS WITH CONSTANT ANGULAR RANGE

Transfer between two circular orbits may be accomplished via one of the four conic-section arcs shown in Fig. 4. Following Reference [4], let us designate transfers via the arcs 01, 0*01, 011* and 0*011* as transfer routes A, B, C and D respectively. The last two routes C and D occur only for elliptical transfers.

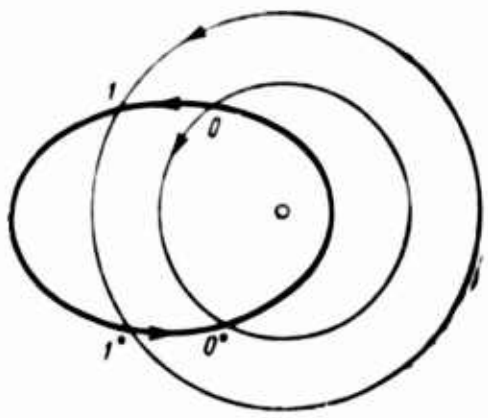


Fig. 4

The change in the true anomaly in a transfer from point 0 to point 1 by route A is (Fig. 4)

$$\Delta\eta_{01}^{(A)} = \eta_1 - \eta_0,$$

where $0 \leq \eta_0,\ \eta_1 \leq 180^\circ$ are the true anomalies at points 0 and 1:

$$\begin{aligned} \eta_0 &= \arccos[(1/e)(p' - 1)], \\ \eta_1 &= \arccos[(1/e)((p'/n) - 1)]. \end{aligned} \tag{20}$$

For routes B, C and D, $\Delta\eta_{01}$ is defined by the relationships

$$\begin{aligned} \Delta\eta_{01}^{(B)} &= \Delta\eta_{01}^{(A)} + 2\eta_0 = \eta_1 + \eta_0, \\ \Delta\eta_{01}^{(C)} &= \Delta\eta_{01}^{(A)} + 2(\pi - \eta_1) = 2\pi - (\eta_0 + \eta_1), \\ \Delta\eta_{01}^{(D)} &= 2\pi - \Delta\eta_{01}^{(A)}. \end{aligned}$$

Considering route A, we have

$$\cos\Delta\eta_{01} = \cos\eta_1\cos\eta_0 + \sin\eta_1\sin\eta_0,$$

from which, applying (20), we obtain

$$e^2 = b_0 + b_1 p' + b_2 p'^2, \tag{21}$$

where

$$\begin{aligned} b_0 &= 2[(1 - \cos\Delta\eta_{01}) / \sin^2\Delta\eta_{01}], \\ b_1 &= -2[1 + (1/n)][(1 - \cos\Delta\eta_{01}) / \sin^2\Delta\eta_{01}], \\ b_2 &= [1 + (1/n^2) - (2/n)\cos\Delta\eta_{01}] / \sin^2\Delta\eta_{01}. \end{aligned} \tag{22}$$

We arrive at the same results for routes B, C and D.

The relationship obtained is an equation of isogons – lines of equal transfer angles $\Delta\eta_{01}$ = const between circular orbits.

As in our investigation of the lines ΔV_{01} = const, we shall restrict ourselves to the case $n > 1$ and examine an earth-to-Mars transfer. The results of calculation of the curves $e = e(p')$ for $\Delta\eta_{01}$ = const are presented for this case in Fig. 5 (each curve $\Delta\eta_{01}$ = const also corresponds to values $360^\circ - \Delta\eta_{01}$).

It is evident from the graph that in the general case, a curve $\Delta\eta_{01}$ = const is tangent to the boundaries of the region of admissible transfer trajectory parameters.

The curve $\Delta\eta_{01}$ = const touches the line (16) at $\cos\Delta\eta_{01} < 1/n$; the value of p' at the point of tangency $p'_{max}(\Delta\eta)$ is determined by the relationship

$$p'_{max}(\Delta\eta) = (1 - \cos\Delta\eta_{01}) / [(1/n) - \cos\Delta\eta_{01}]. \qquad (23)$$

As $\cos\Delta\eta_{01} \to 1/n$ $p'_{max}(\Delta\eta) \to \infty$ and for $\cos\Delta\eta_{01} > 1/n$ the curve $\Delta\eta_{01}$ = const has no points in common with the Line (16). As for the Line (17), the curve $\Delta\eta_{01}$ = const is tangent to it for all values of $\cos\Delta\eta_{01}$, at the points

$$p'_{min}(\Delta\eta) = (1 - \cos\Delta\eta_{01}) / [1 - (\cos\Delta\eta_{01}/n)]. \qquad (24)$$

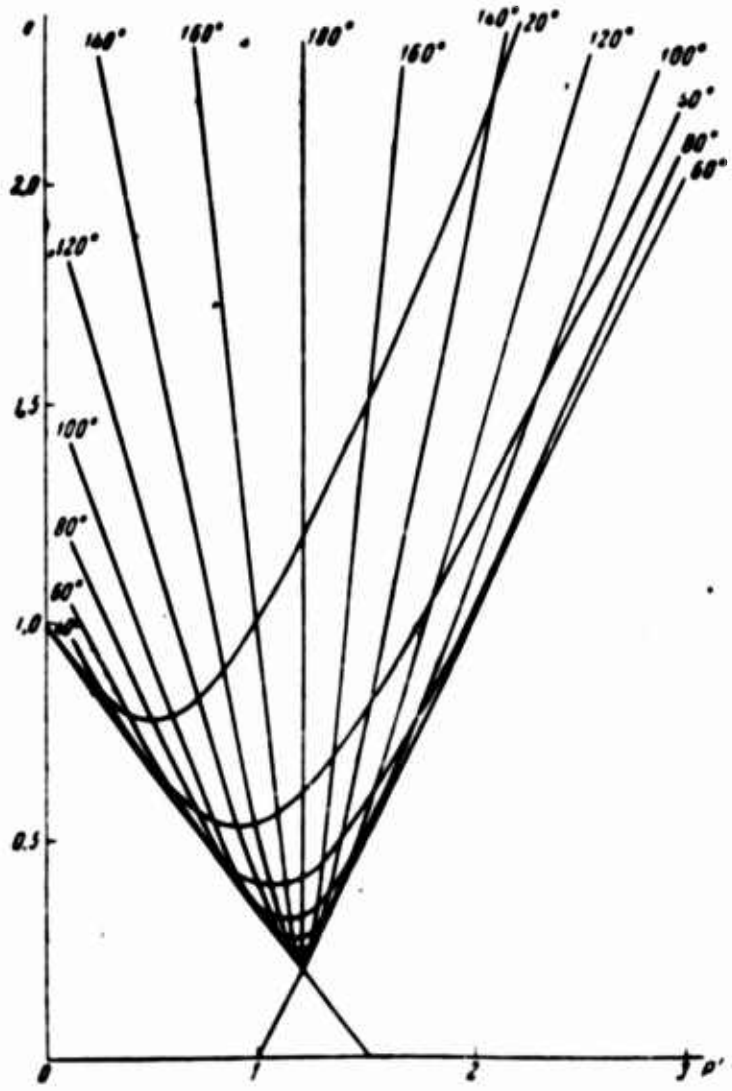


Fig. 5

With $\Delta\eta_{01} = 180^\circ$ (Hohmann semiellipse), the curve $\Delta\eta_{01}$ = const degenerates into a vertical straight line

$$p' = p'_{zon} = 2n/(n+1).$$

Then

$$p'_{min}(\Delta\eta) = p'_{max}(\Delta\eta) = p'_{zon}.$$

A diagram of the curves $p'_{max}(\Delta\eta)$, $p'_{min}(\Delta\eta)$ for earth-to-Mars transfer is presented in Fig. 6 (the curve corresponds to values $0 \leq \Delta\eta_{01} \leq 180^\circ$ and $360^\circ - \Delta\eta_{01}$).

We note that the singularities that appear in Eqs. (21) and (22) at $\Delta\eta = 0, 180, 360^\circ$ stem from consideration of radial transfers whose geometrical locus is the semiaxis $p' = 0$, $e \geq 1$. This case is not of practical interest and will not be examined here.

Let us establish correspondence between the various transfer arcs and points on the curves $\Delta\eta_{01} = \text{const}$.

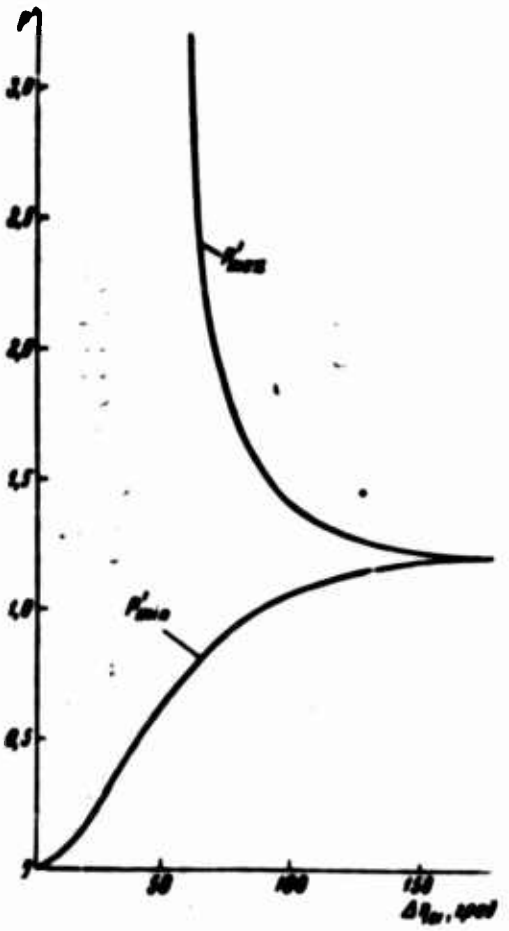


Fig. 6. 1) Degrees.

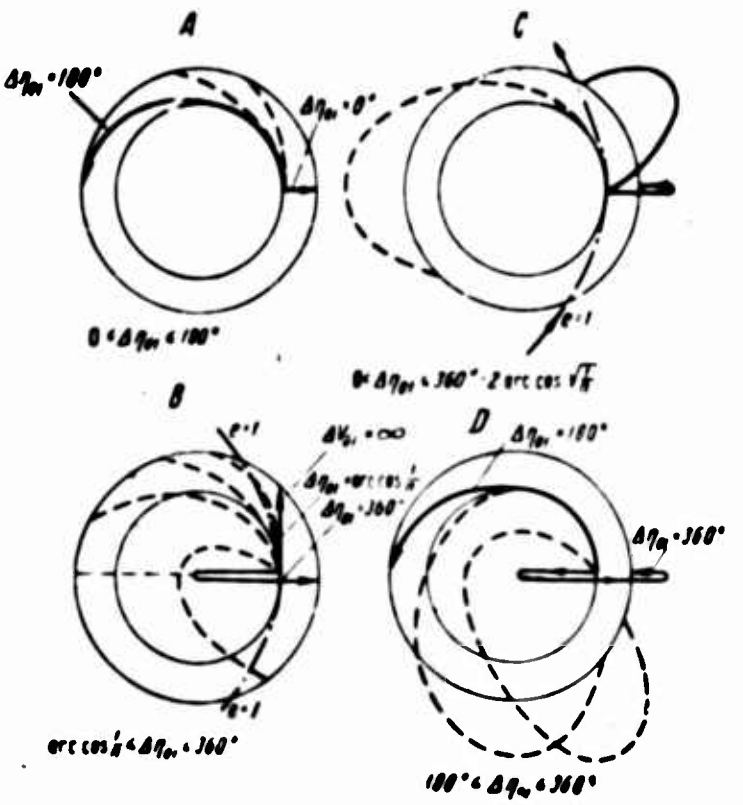


Fig. 7

The change in the true anomaly for the various transfer arcs is enclosed within the following limits (Fig. 7):

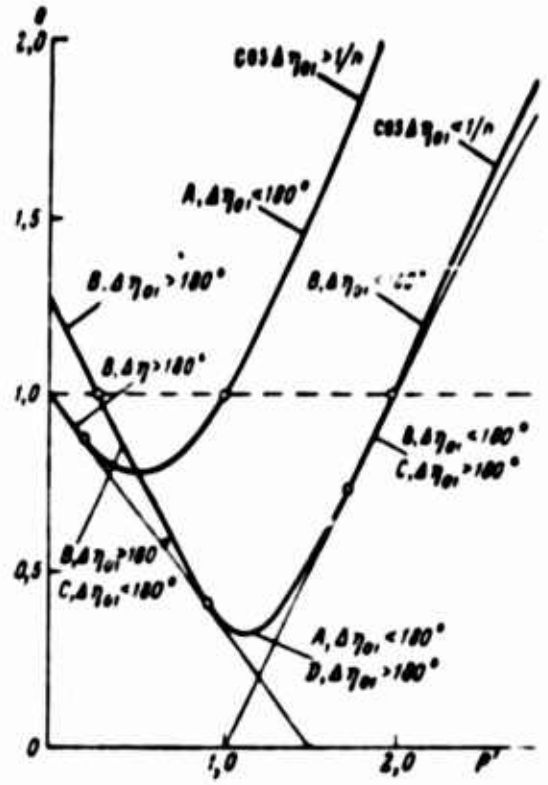


Fig. 8

Range of Variation of $\Delta\eta$ for Various Routes

2 маршрут / 1 участок кривой	$0 \le p' \le p'_{min}(\Delta\eta)$		$p'_{min}(\Delta\eta) \le p' \le p'_{max}(\Delta\eta)$		$p'_{max}(\Delta\eta) \le p' < \infty$	
	$e<1$	$e>1$	$e<1$	$e>1$	$e<1$	$e>1$
A	—	—	$0 \div 180°$		—	—
B	$180 \div 360°$		—	—	$\arccos 1/n \div 180°$	
C	$0 \div 180°$	—	—	—	$180 \div 360° - 2\arccos\sqrt{1/n}$	—
D	—	—	$180 \div 360°$	—	—	—

Note: The Table has been compiled for values of $\cos\Delta\eta < 1/n$. For $\cos\Delta\eta \ge 1/n$, $p'_{max}(\Delta\eta)$ is replaced by ∞, and the segment of the curve $p'_{max}(\Delta\eta) \le p' < \infty$ is lacking.

1) Segment of curve; 2) route.

A: $0 \leqslant \eta_{01} \leqslant 180°$.

B: $\arccos(1/n) \leqslant \Delta\eta_{01} \geqslant 360°$.

C: $0 \leqslant \Delta\eta_{01} \leqslant 360° - 2\arccos\sqrt{1/n}$.

D: $180° \leqslant \Delta\eta_{01} \leqslant 360°$.

Since the coefficients b_i depend on $\sin^2\Delta\eta_{01}$ and $\cos\Delta\eta_{01}$ the curves $e = e(p'; \Delta\eta_{01} = \text{const})$ for values $0 \le \Delta\eta_{01} \le \pi$ and $\pi \le \Delta\eta_{01}' = 2\pi - \Delta\eta_{01} \le 2\pi$ do not differ from one another. It is obvious that one

to one correspondence will then be retained for transfers A and D between $\Delta\eta_{01}$ and cos $\Delta\eta_{01}$ and, consequently, between points of the curve $e = e(p'; \Delta\eta_{01} = \text{const})$ and $\Delta\eta_{01}$. For one to one correspondence to be maintained between points of the curve $e = e(p'; \Delta\eta_{01} = \text{const})$ and $\Delta\eta_{01}$ for transfers B and C, the corresponding curves must have two branches each: one for values $0 \leq \Delta\eta_{01} \leq \pi$ and another for values $\pi \leq \Delta\eta_{01} \leq 2\pi$. The results of a study of the correspondence between the values of $\Delta\eta_{01}$ and the various segments of the curves $e = e(p', \Delta\eta_{01} = \text{const})$ are given in Fig. 8 and in the Table.

We note that the results obtained under this heading are in fact independent of examination of transfers between circular orbits and are valid when only the angle between two radius vectors of a material point moving in a Newtonian field is given.

APPLICATION OF ISOENERGETIC CURVES AND ISOGONS IN THE DYNAMICS OF SPACE FLIGHT

1. Let us examine a single-impulse transfer with assigned energy $\Delta V_{01} = \text{const}$ and a semimajor axis $\underline{a}$.

An annular transfer between the orbits of the earth and the planet with a period that is a multiple of the earth's orbital period, with the result that the space vehicle re-encounters the earth, may serve as an example of such transfer. Using (12), we obtain for the constant of the energy integral

$$h' = \frac{h}{V^2_{\text{орб}_\oplus}} = \frac{e^2-1}{p'} = a_1' + 2\sqrt{p'}.$$

Since, for an ellipse,

$$a' = a / R_{\text{орб}_\oplus} = -1/h',$$

we obtain finally

$$p' = \frac{1}{4}\left(a_1' + \frac{1}{a'}\right)^2,$$

with the necessary condition

$$a_1' + \frac{1}{a'} < 0.$$

2. Let us consider a two-impulse earth-to-planet tra sfer along arc B with an assigned position of the pericenter.

Such a problem arises in consideration of "quick" earth-planet-earth transfers [5].

Using the relationships

$$r_p' = p'/(1+e), \quad r_p' = r_p/R_{орб}$$

and (8), we obtain the following equation for determination of p':

$$\left(a_2 - \frac{1}{r_p'^2}\right)p' + a_{1/2}p'^{1/2} + a_1 + \frac{2}{r_p'} = 0.$$

3. Let us consider the problem of determining the minimal velocity impulse required to rendezvous an ISZ moving in a circular orbit of radius R_0 with an ISZ also moving on a circular orbit, but of radius R_1, with a specified angular distance.

Fig. 9 shows the typical shape of the $\Delta\eta_{01}$ = const and ΔV_{01} = const curves (for a single-impulse transfer with n ≈ 1). It is not difficult to perceive that with $\Delta\eta_{01}$ = const the minimum of ΔV_{01} for $\Delta\eta_{01} < 180^\circ$ is reached on arc A, and for $\Delta\eta_{01} > 180^\circ$ on arc D, whose parameters coincide with those of arc A, at the moment of tangency between the curves $\Delta\eta_{01}$ = const and ΔV_{01} = const.

The solution of the problem reduces to investigation of the following equation of fourth degree in $p'^{1/2}$:

$$b_2 p'^2 - 2p'^{3/2} + (b_1 - a_1')p' + b_0 - 1 = 0, \qquad (25)$$

which is obtained from (12) and (21). It follows from the shape of the curves $\Delta\eta 01$ = const and ΔV_{01} = const that Eq. (25) has no more than two real roots in the region of the parameters (p', e) under consideration. The solution sought corresponds to the point at which these roots become a single multiple root.

4. A most interesting problem in which application of Relationships 8) and (12) is found to be highly effective is that of earth-to-planet-o-earth transfers. This problem has been considered in a number of apers, among which we cite [1, 2 and 5].

The general scheme of a four-impulse transfer of a spacecraft to a lanet of the solar system with return to the earth is shown in Fig. 0.

The total angle of rotation of the vehicle's radius vector about he sun is

$$\Delta\eta_{01} + \Delta\eta_{23} + \omega_1\Delta t_\Sigma,$$

here ω_1 is the mean motion of the destination planet and Δt_Σ is the parking time" on the ISP orbit. During this time, the earth revolves hrough an angle

$$\omega_0 T_\Sigma = \omega_0(t_{01} + t_{23} + \Delta t_\Sigma),$$

here ω_0 is the mean motion of the earth and T_Σ is the time of the space xpedition.

The condition for the vehicle's return to the earth at point 3 is ritten in the form

$$\Delta\eta_{01} + \Delta\eta_{23} + \omega_1\Delta t_\Sigma = \omega_0(t_{01} + t_{23} + \Delta t_\Sigma) \pm 2\pi k \quad (\text{with } k = 0.1,\ldots),$$

rom which the parking time is

$$\Delta t_\Sigma = \frac{\Delta\eta_\Sigma - \omega_0 t_\Sigma}{\omega_0 - \omega_1} + kT_{sin} \quad (k = 0, \pm 1, \ldots), \qquad (26)$$

ere

$$\Delta\eta_\Sigma = \Delta\eta_{01} + \Delta\eta_{23}, \quad t_\Sigma = t_{01} + t_{23},$$

$T_{sin} = 2\pi/|\omega_0 - \omega_1|$ is the synodic period of the planet.

It follows from (26) that

$$T_\Sigma = t_\Sigma + \Delta t_\Sigma = \frac{\Delta\eta_\Sigma - \omega_1 t_\Sigma}{\omega_0 - \omega_1} + kT_{sin} \quad (k = 0, \pm 1, \ldots). \qquad (27)$$

For subsequent treatment, it will be convenient to represent Relationships (26) and (27) in the form

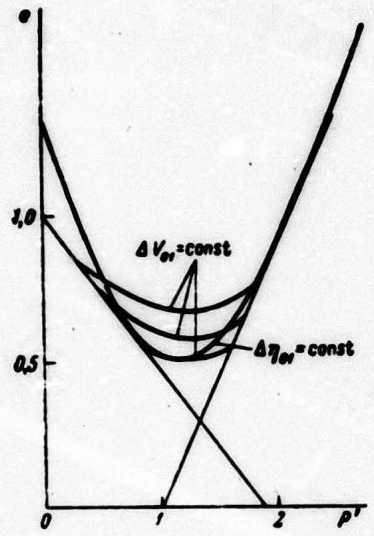


Fig. 9

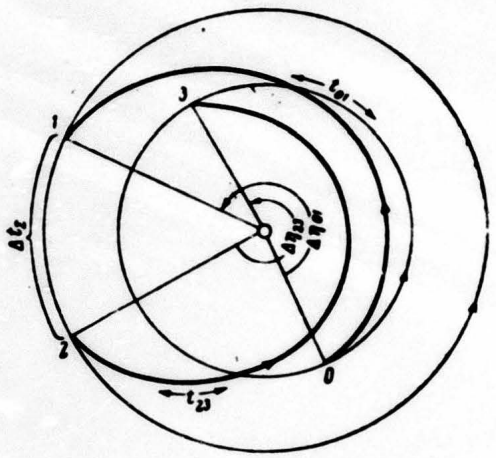


Fig. 10

$$\Delta t_\Sigma = \Delta t_{01}(p_{01}', e_{01}) + \Delta t_{23}(p_{23}', e_{23}) + kT_{\text{com}} \quad (28)$$

$$T_\Sigma = T_{01}(p_{01}', e_{01}) + T_{23}(p_{23}', e_{23}) + kT_{\text{com}}, \quad (29)$$

where Δt_{01}, Δt_{23} and T_{01}, T_{23} are respectively the same as the first terms in the right members of Equalities (26) and (27) with the subscript Σ changed to 01 or 23 and $p' = p/R_{orb_\oplus}$.

Let us formulate the problem of optimum earth-to-planet transfers as follows:

With a given resultant characteristic velocity of the vehicle

$$\Delta V_\Sigma = \Delta V_{01}(p_{01}', e_{01}) + \Delta V_{23}(p_{23}', e_{23}) = \text{const} = \Delta V_\Sigma^* \quad (30)$$

and provided that the space vehicle remains on the ISP orbit for a specified time

$$\Delta t_\Sigma = \Delta t_{01}(p_{01}', e_{01}) + \Delta t_{23}(p_{23}', e_{23}) + kT_{\text{com}} = \text{const} = \Delta t_\Sigma^*, \quad (31)$$

find a combination of flight-out and flight-back routes and parameters for the corresponding conic sections (p_{01},, e_{01}) and (p_{23},, e_{23}) such that the duration of the space transfer will be minimal:

$$T_\Sigma = T_{01}(p_{01}', e_{01}) + T_{23}(p_{23}', e_{23}) + kT_{\text{com}} \to \min. \quad (32)$$

The solution of this problem on an EVM [Electronic Computer] re-

duces to solution of systems of transcendental equations of the form

$$\begin{aligned} \Delta V_{01}(p_{01}', e_{01}) &= \Delta V_{01}^{\circ} \\ \Delta t_{01}(p_{01}', e_{01}) &= \Delta t_{01}^{\circ} \\ \Delta V_{23}(p_{23}', e_{23}) &= \Delta V_{\Sigma}^{\circ} - \Delta V_{01}^{\circ} \\ \Delta t_{23}(p_{23}', e_{23}) &= \Delta t_{\Sigma}^{\circ} - \Delta t_{01}^{\circ} \end{aligned} \tag{33}$$

with a subsequent search for the minimum of the function

$$\begin{aligned} T_{\Sigma}(\Delta V_{01}^{\circ}, \Delta t_{01}^{\circ}) = T_{01}(\Delta V_{01}^{\circ}, \Delta t_{01}^{\circ}) + \\ + T_{23}(\Delta V_{\Sigma}^{\circ} - \Delta V_{01}^{\circ}, \Delta t_{\Sigma}^{\circ} - \Delta t_{01}^{\circ}) + kT_{\text{сим}} \end{aligned}$$

for the condition $\Delta V_{\Sigma *} = \text{const}$, $\Delta t_{\Sigma *} = \text{const}$ with ΔV_{01*}, Δt_{01*} varied.

We changed the problem into a form such that it is completely unnecessary to solve systems of transcendental equations of the type of (33).

We note that by virtue of the property of isoperimetry, the initial problem is equivalent to the problem described by the equation system

$$\begin{aligned} \Delta V_{\Sigma} &= \Delta V_{01} + \Delta V_{23} \to \min, \\ \Delta t_{\Sigma} &= \Delta t_{01} + \Delta t_{23} + kT_{\text{сим}} = \Delta t_{\Sigma}^{\circ}, \\ T_{\Sigma} &= T_{01} + T_{23} + kT_{\text{сим}} = T_{\Sigma}^{\circ}. \end{aligned} \tag{34}$$

Examining the last two relationships of (34) on the basis of (26) and (27) as a system of linear equations in $\Delta\eta_{\Sigma}$ and t_{Σ}, we can always (the system determinant $\Delta = 1/(\omega_0 - \omega_1) \neq 0$) set $\Delta\eta_{\Sigma *}$ and $t_{\Sigma *}$ in correspondence with the assigned $\Delta t_{\Sigma *}$ and $T_{\Sigma *}$ by the formulas

$$\Delta\eta_{\Sigma}^{\circ} = \omega_0 T_{\Sigma}^{\circ} - \omega_1 \Delta t_{\Sigma}^{\circ} - 2\pi k, \quad t_{\Sigma}^{\circ} = T_{\Sigma}^{\circ} - \Delta t_{\Sigma}^{\circ}.$$

Thus, the system formed by the last two equalities in (34) is equivalent to the system of equalities

$$\begin{aligned} \Delta\eta_{\Sigma} &= \Delta\eta_{01} + \Delta\eta_{23} = \text{const} = \Delta\eta_{\Sigma}^{\circ}, \\ t_{\Sigma} &= t_{01} + t_{23} = \text{const} = t_{\Sigma}^{\circ}, \end{aligned}$$

and the isoperimetric problem (34) is equivalent to the problem

$$\begin{aligned} \Delta V_\Sigma &= \Delta V_{01} + \Delta V_{23} \Rightarrow \min, \\ \Delta\eta_\Sigma &= \Delta\eta_{01} + \Delta\eta_{23} = \Delta\eta_\Sigma^\circ, \\ t_\Sigma &= t_{01} + t_{23} = t_\Sigma^\circ. \end{aligned} \tag{35}$$

But, on the basis of the isoperimetry property, the latter is equivalent to the problem defined by the following system of relationships:

$$\begin{aligned} \Delta V_\Sigma &= \Delta V_{01} + \Delta V_{23} = \Delta V_\Sigma^\circ, \\ \Delta\eta_\Sigma &= \Delta\eta_{01} + \Delta\eta_{23} = \Delta\eta_\Sigma^\circ, \\ t_\Sigma &= t_{01} + t_{23} \Rightarrow \text{extr}, \end{aligned} \tag{36}$$

where the sense of the extreme of t_Σ can be established on the basis of analysis of the properties of actual earth-planet-earth transfer families.

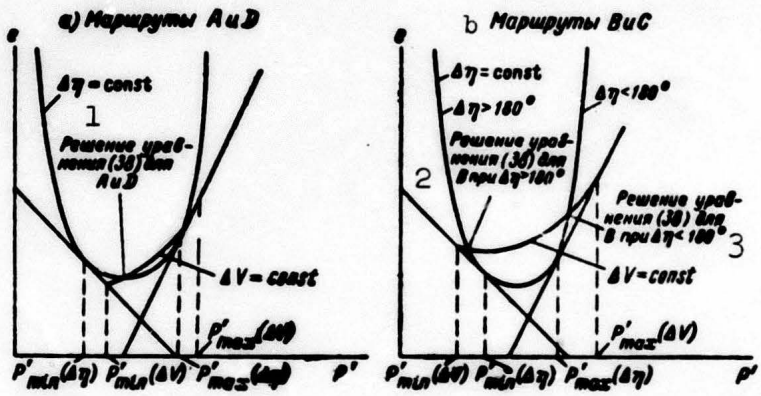


Fig. 11. a) Routes A and D; b) routes B and C. 1) Solution of Eq. (38) for A and D; 2) solution of Eq. (38) for B with $\Delta\eta > 180°$; 3) solution of Eq. (38) for B with $\Delta\eta < 180°$.

Determination of transfers that satisfy the first two conditions in (36) reduces to determining the roots of the equation system

$$\begin{aligned} \Delta V_{01}(p_{01}', e_{01}) &= \Delta V_{01}^\circ, \\ \Delta\eta_{01}(p_{01}', e_{01}) &= \Delta\eta_{01}^\circ, \\ \Delta V_{23}(p_{23}', e_{23}) &= \Delta V_\Sigma^\circ - \Delta V_{01}^\circ, \\ \Delta\eta_{23}(p_{23}', e_{23}) &= \Delta\eta_\Sigma^\circ - \Delta\eta_{01}^\circ. \end{aligned} \tag{37}$$

However, on the basis of Relationships 3) and (21), the roots of the equation system (37) are found among the roots of an equation of fourth degree in $p'^{1/2}$:

$$(a_2 - b_2)p_{01}'^2 + a_{3/2}p'^{3/2} + (a_1 - b_1)p' + 1 - b_0 = 0. \tag{38}$$

Using one of the regular methods, for example, the Ferrara method, for solution of Eq. (38), we can use (38) to obtain immediately $\underline{p}'$ and, consequently, $\underline{e}$ that satisfy the first two conditions of (36).

The roots of Eq. (38) that have physical significance are identified with the aid of the following criteria, which were derived on the basis of an examination of the properties of lines $\Delta V_{01} = \text{const}$, $\Delta\eta_{01} = \text{const}$ on the plane of the parameters (p_{01}, e_{01}) (Fig. 11).*

Routes A and D (Fig. 11a)

$$\max\{p'_{\min}(\Delta V),\ p'_{\min}(\Delta\eta)\} \leqslant p' \leqslant \min\{p'_{\max}(\Delta V),\ p'_{\max}(\Delta\eta)\} \quad (39)$$

Route B (Fig. 11b)

$$\text{a)}\quad \arccos\frac{1}{n} \leqslant \Delta\eta_{01} \leqslant 180^\circ \quad (40)$$

$$p'_{\max}(\Delta\eta) \leqslant p' \leqslant p_{\max}(\Delta V)$$

$$\text{б)}\quad 180^\circ \leqslant \Delta\eta_{01} \leqslant 360^\circ \quad (41)$$

$$p'_{\min}(\Delta V) \leqslant p' \leqslant p'_{\min}(\Delta\eta)$$

Route C

For $0 \leq \Delta\eta_{01} \leq 180^\circ$ the constraint on $\underline{p}'$ is identical to (41), and for $180^\circ \leq \Delta\eta_{01} \leq 360^\circ - 2\arccos(1/n)^{1/2}$ it coincides with (40).

It follows from the form of the curves in Fig. 11 that for routes B and C with $\Delta V_{01} = \text{const}$, $\Delta\eta_{01} = \text{const}$ no more than one transfer is defined, and no more than two for A and D.

Solving Eqs. (37) for assigned $\Delta V_{\Sigma*}$, $\Delta\eta_{\Sigma*}$, we represent t_Σ in the form of a function of ΔV_{01*}, $\Delta\eta_{01*}$:

$$t_\Sigma(\Delta_{01}^\circ, \Delta\eta_{01}^\circ) = t_{01}(\Delta V_{01}^\circ, \Delta\eta_{01}^\circ) + t_{23}(\Delta V_\Sigma^\circ - \Delta V_{01}^\circ, \Delta\eta_\Sigma^\circ - \Delta\eta_{01}^\circ).$$

From here on, determination of the optimum transfer reduces to using known methods to find points on the $(\Delta V_{01}, \Delta\eta_{01})$ plane at which t_Σ

reaches the appropriate extreme.

5. Of particular interest is the problem of earth-planet-earth transfers using deceleration in the earth's and planet's atmosphere's, since the amount of energy required for earth-planet-earth space epeditions can be reduced substantially in this case.

One of the differences between this problem and that considered above is the necessity of accounting for g-force and heat-flow limitations as the vehicle sinks into the atmosphere. This limitation can be realized, for example, by specifying velocities for the vehicle at the sphere's of influence of the earth and the planet.

Let us examine as an example a three-impulse earth-planet-earth transfer with deceleration in the atmosphere of the earth. It is obvious that the statement of this problem can be reduced to the problem described by Relationships (30)-(32) with the additional equality

$$v'^2_{c\Phi_3} = 3 - 2\sqrt{p'} + \frac{e^2 - 1}{p'} = v'^2_{c\Phi_3}, \qquad (42)$$

which corresponds to assignment of the velocity at the earth's sphere of influence as it is approached at point 3 (Fig. 10).

Obviously, the present problem can be reduced to Problem (36) with application of Equality (42).

Considering a single-impulse planet-earth transfer and applying Relationships (14), (15) and (42), we obtain the following expression for the focal parameter of the orbit:

$$\sqrt{p'_{23}} = \frac{A^{*2} - C^2 - v'^{*2}_{c\Phi_3} + 3[1 - (1/n)]}{2[1 - (1/n^{3/2})]}. \qquad (43)$$

Calculating p_{23}, from (43) and e_{23} from (14), we can find all characteristics of the planet-earth transfer, including $\Delta\eta_{23}$ and t_{23}. Calculation of the characteristics of a two-impulse earth-planet transfer

follows the technique set forth above for values of ΔV_{01}, $\Delta\eta_{01}$ defined by the equalities

$$\Delta V_{01} = \Delta V_2^* - \Delta V_2, \quad \Delta\eta_{01} = \Delta\eta_2^* - \Delta\eta_{23},$$

where $\Delta V_2(p_{23}, e_{23})$ is the impulse at the planet for the planet-earth transfer.

Thus, for an assigned value of v_{sf3}, we express all characteristics of the transfer and, in particular, t_Σ as a function of ΔV_2. Thereafter, determination of the optimum transfer reduces to using numerical methods on an EVM [Electronic Computer] to find the extreme of t_Σ with respect to ΔV_2.

6. As a last example, let us examine the problem of passing arcs of a conic section between two points on the orbits of the earth and the destination planet, as determined by the date of departure from the ISZ orbit, t_0, and the date of arrival t_1 in the neighborhood of the destination planet.

Assigning the quantities t_0 and t_1 determines the radius vectors $\vec{r}_0$ and $\vec{r}_1$ and, consequently, all characteristics of the transfer, including the angle $\Delta\eta_{01}$.

Turning now to the motion of the space vehicle in the transfer plane, we can treat it as a transfer along an arc of specified type between circular orbits with radii $\vec{r}_0$ and $\vec{r}_1$ and write

$$\Delta\eta_{01} = \Delta\eta_{01}(n_{01}, p_{01}', e_{01}). \tag{44}$$

where

$$t_{01} = t_{01}(n_{01}, p_{01}' e_{01}) = t_1 - t_0,$$
$$n_{01} = \frac{r_1}{r_0} = \text{const}, \quad p'_{01} = \frac{p_{01}}{r_0}. \tag{45}$$

Applying Eq. (21) and excluding e_{01} from (45), we reduce the problem of determining the conic-section parameters to solution of a single transcendental equation in p_{01}:

$$t_{01} = t_1 - t_0 = t_{01}(n_{01}, p_{01}'). \tag{46}$$

The above procedure substantially simplifies solution of a number of interorbital space-transfer problems. The example considered also indicates that Eq. (21) can be used effectively in problems of determining the orbit of a body from two positions, which was investigated in [3].

Received

2 January 1964

REFERENCES

1. R.H. Battin, J. Aerospace Sci., 26, No. 9, 545, 1959.
2. J.V. Breakwell, R.W. Gillespie and S. Ross, J. Aerospace Sci., 31, No. 2, 201, 1961.
3. P.Ye. El'yasberg, Collection entitled "Iskusstvennyye sputniki Zemli" [Artificial Earth Satellites], No. 13. Izd-vo AN SSSR [Academy of Sciences USSR Press], 1962, page 3.
4. M. Vertregt, J. Brit. Interpl. Soc., 16, No. 6, 326, 1958.
5. K.A. Ehricke, Trans. ASME, Ser. B., J. Eng. for Ind., 83, No. 1, 1, 1961.

Manuscript Page No.

[Footnotes]

65 *We note that analysis of the relationships obtained for the case $n < 1$ is not necessary in practice, since, instead of transfer to an inferior orbit we can always consider the return transfer to a superior orbit.

78 *The criteria presented here are valid for $n > 1$, i.e., for transfers to superior planets of the solar system. In considering transfers to inferior planets, it is expedient to consider the "return" transfers from the inferior planet to the earth in order to retain the technique as set forth without any changes.

Manuscript Page No.	[List of Transliterated Symbols]
61	ИСЗ = ISZ = iskusstvennyy sputnik Zemli = artificial earth satellite
61	ИСП = ISP = iskusstvennyy sputnik planety = artificial planet satellite
61	хар = khar = kharakteristicheskiy = characteristic
62	орб = orb = orbita = orbit
62	сф = sf = sfera = sphere
62	ср = sr = sredniy = average
69	хом = khom = khomanskiy = Hohmann's, of Hohmann
74	син = sin = sinodicheskiy = synodic
80	ЭВМ = EVM = elektronnaya vychislitel'naya mashina = electronic computer

UDK 629.197.7

CONCERNING A PRIORI ERROR ESTIMATES IN DETERMINING PARAMETERS BY THE METHOD OF LEAST SQUARES

M.L. Lidov

The problem of minimizing the error in the determination of a given trajectory parameter (by selecting a certain battery of measurements from a given aggregate) is considered with the assumption of the most unfavorable relationship between the measurement error correlation coefficients.

The maximum probability method, or the method of least squares is used to determine flight trajectory parameters from a given set of measurements [1].

The method of least squares is used in cases in which the errors of the parameters to be measured are actually correlated, but the correlation coefficients are unknown. In the latter case, the actual error of determination of any given trajectory parameter will also depend on the actual correlation between the errors of measurement.

To estimate the error of parameter determination by the method of least squares, it is helpful, along with the case of independent measurement errors, to consider another (extreme for this problem) case, in which the errors of measurement are bounded only in absolute magnitude, and any arbitrary correlation among them is possible. On these assumptions, the present paper will consider the problem of minimizing errors in the determination of a given trajectory parameter (by selecting a certain battery of measurements from a given aggregate), assuming that the least favorable relationship prevails among the correlation coeffi-

cients of the measurement errors.

In the method of least squares, the vector $a = (a_1, \ldots, a_m)$ is determined by minimization of the function

$$I(a) = \sum_{i=1}^{n} \left(\frac{D_{i\,\text{изм}} - D_i}{\sigma_i^*} \right)^2, \qquad (1)$$

where $D_i = D_i(a)$ is a function of the vector $\underline{a}$, $D_{i\,\text{изм}}$ is the measured value of the quantity D_i and σ_{i*} is the a priori maximum error of the ith measurement.

Let us use a^* to denote the true value of $\underline{a}$. Then

$$D_{i\,\text{изм}} = D_i(a^*) + \sigma_i, \qquad D_i(a) = D_i(a^*) + \operatorname{grad}_a D_i \Delta a + \ldots. \qquad (2)$$

where σ_i is the absolute value of the measurement error and Δa is the error in the determination of the vector $\underline{a}$.

From (1) and (2), in linear approximation,

$$I = \sum_{i=1}^{n} (\rho_i - b_i \Delta a)^2, \qquad (3)$$

$\left(\rho_i = \frac{\sigma_i}{\sigma_i^*}, \quad b_i = \operatorname{grad} \frac{D_i}{\sigma_i^*}\right.$ is an m-dimensional vector).

Let us denote the ith measurement by b_i and the error of the ith measurement by ρ_i. Errors $\delta \underline{l}$ of determination of the parameter $\underline{l}$ are related to errors in the determination of the vector $\underline{a}$; in linear approximation

$$\delta l = c \Delta a, \quad c = \operatorname{grad}_a l. \qquad (4)$$

Let us denote by $\mathfrak{B}$ the given set of measurements b_i, which contain, together with the vector b_i the vector $-(b_i : \beta_k)$, the set of all possible combinations of $\underline{r}$ measurements $b_i \in \mathfrak{B}$

$$\beta = \beta_m \cup \beta_{m+1} \cup \ldots.$$

The problem considered is that of determining

$$\min_{\beta} \max_{\rho_i} \delta l \qquad (5)$$

on the assumption that*

$$|\rho_i| \leqslant 1 \qquad (6)$$

From the minimum condition of I with respect to $\underline{a}$ we obtain from (3)

$$\Delta a = (B^*B)^{-1}B^*\rho, \qquad (7)$$

where $B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ is a square n x m matrix and $\rho = \begin{pmatrix} \rho_1 \\ \vdots \\ \rho_n \end{pmatrix}$ is a vector. The following statement is evident from (4), (6) and (7):

The maximum of $\delta\underline{l}$ with respect to ρ_1 [i = 1, ..., n) is reached at the vertices of an n-dimensional cube with $|\rho_1| = 1$.

a) Let us consider the case in which the number of measurements $\underline{n}$ is equal to the dimension of the space M. Let b_1, b_2, ..., b_m be elements of $\mathfrak{B}$ that satisfy the conditions:

1) linear independence;

2) the vector $\underline{c}$ can be decomposed into vectors b_1 with positive coordinates b_1 with positive coordinates λ_1:

$$c = \lambda_1 b_1 + \ldots + \lambda_m b_m, \quad \lambda_i > 0 \quad (i = 1, 2, \ldots, m) \qquad (8)$$

In this case, we obtain from (7) and (8)

$$\max_{\rho_i} \delta l = \sum_{i=1}^{m} \lambda_i, \qquad \rho_i = 1 \quad (i = 1, 2, \ldots, m). \qquad (9)$$

In case a), the problem of finding

$$\min_{\beta_m} \max_{\rho_i} \delta l$$

is a solved problem of linear programming [2]. The simplex method can be used to minimize $\max_{\rho i} \delta\underline{l}$ in β_m. According to [2] (for a specific form of the functional (9)), the following theorem applies.

Theorem 1. For the measurements b_1, ..., b_m to realize

$$\min_{\beta_m} \max_{\rho_i} \delta l \qquad (10)$$

it is necessary and sufficient that for any $b_i \in \mathfrak{B}$ the condition

$$\left| \sum_{i=1}^{m} x_{ij} \right| \leqslant 1, \tag{11}$$

be satisfied; here, the x_{ij} are the coordinates of the analysis of vector b_j into vectors of the basis $b_1, \ldots, b_m$

$$b_j = x_{1j}b_1 + \ldots + x_{mj}b_m. \tag{12}$$

b) Below we present the proof of the following theorem for the general case of $n > m$ measurements.

Theorem 2. For $n \geq m$,

$$\min_{\beta_n} \max_{\rho_i} \delta l \geqslant \min_{\beta_m} \max_{\rho_i} \delta l. \tag{13}$$

Proof. Let $b_1, \ldots, b_m$ be measurements for which

$$\min_{\beta_m} \max_{\rho_i} \delta l, \qquad |\rho_i| \leqslant 1 \qquad (i = 1, 2, \ldots, m).$$

is realized.

Let us consider an arbitrary set of measurements

$$b_1', \ldots, b_n' \quad (b_i' \in \mathfrak{B}) \quad (i = 1, 2, \ldots, n). \tag{14}$$

Let $\rho_1 \ldots, \rho_n$ ($|\rho_i| \leq 1$) be the errors for which max δl is realized in the measurements (14). Then we have for arbitrary Δ_1 with $|\Delta_1| \leq 1$

$$\delta l(\rho_1, \ldots, \rho_n, b_1', \ldots, b_n') \geqslant \delta l(\Delta_1, \ldots, \Delta_n, b_1', \ldots, b_n'). \tag{15}$$

Let

$$b_i' = x_{1i}b_1 + x_{2i}b_2 + \ldots + x_{mi}b_m \ (i = 1, 2, \ldots, n). \tag{16}$$

Let us consider

$$I = \sum_{i=1}^{n} (\Delta_i - b_i'\Delta a)^2. \tag{17}$$

We take

$$\Delta_i = \sum_{j=1}^{m} x_{ji} \qquad (i = 1, 2, \ldots, n). \tag{18}$$

According to Theorem 1, $|\Delta_1| \leq 1$.

Substituting Δ_1 and b_1, from (16) and (18) into (17), we obtain

$$I = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} x_{ji}(1 - b_i \Delta a) \right)^2.$$

The minimum of $\tilde{I}$ with respect to Δa equal to zero (since $I \geq 0$) is determined by the system of equations

$$1 - b_i \Delta a = 0 \ (i = 1, 2, \ldots, m).$$

By (9) and the definition of b_1,

$$\delta l(\Delta_1, \ldots, \Delta_n, b_1', \ldots, b_n') = \min_{\beta_m} \max_{\rho_i} \delta l \leqslant \delta l(\rho_1, \ldots, \rho_n, b_1', \ldots, b_n') = \min_{\beta_n} \max_{\rho_i} \delta l.$$

Received

2 March 1964

REFERENCES

1. E.L. Akim and T.M. Eneyev, Kosmich. issled. [Cosmic Research], No. 1, 5, 1963.

2. D.V. Dantsig, Collection entitled "Metody resheniya obshchey zadachi lineynogo programmirovaniya" [Methods of Solving the General Problem of Linear Programming], Gostekhizdat [State Publishing House for Technical and Theoretical Literature], Moscow, 1963.

Manuscript Page No. [Footnote]

85 *Condition (6) permits arbitrary correlation among the errors of the measurements.

Manuscript Page No. [Transliterated Symbol]

84 из = iz = izmerennyy = measured

UK 521.312

APPLICATION OF V.V. RUMYANTSEV'S THEOREM OF STABILITY WITH RESPECT TO SOME OF THE VARIABLES IN PROBLEMS OF CELESTIAL MECHANICS

V.G. Demin

Using V.V. Rumyantsev's theorem on stability with respect to some of the variables, the stability of the semimajor axes of the orbits of celestial bodies is demonstrated in those problems of celestial mechanics in which small perturbing functions are conservative in nature.

To describe the motion of a celestial body, which is assumed to be a material point, in the presence of conservative perturbing forces, we shall utilize Poincare's first system of canonical elements

$$L = \sqrt{f(m_0 + m)a}, \quad \rho_1 = \sqrt{f(m_0 + m)a}(1 - \sqrt{1 - e^2}),$$
$$\rho_2 = \sqrt{f(m_0 + m)a(1 - e^2)}(1 - \cos i), \qquad (1)$$
$$\lambda = nt + \varepsilon, \quad \omega_1 = -\pi, \quad \omega_2 = -\Omega,$$

where f is the gravitational constant, m_0 and m are the masses of the mutually attracting bodies, a is the semimajor axis of an osculating elliptical orbit, e is its eccentricity, i is its inclination, n is the mean motion, Ω is the longitude of the node, π is the longitude of the pericenter and ε is the mean longitude of the epoch.

The differential equations of perturbed motion written in these variables will take the form

$$\frac{dL}{dt} = \frac{\partial S}{\partial l}, \quad \frac{d\rho_1}{dt} = \frac{\partial S}{\partial \omega_1}, \quad \frac{d\rho_2}{dt} = \frac{\partial S}{\partial \omega_2},$$
$$\frac{d\lambda}{dt} = -\frac{\partial S}{\partial L}, \quad \frac{d\omega_1}{dt} = -\frac{\partial S}{\partial \rho_1}, \quad \frac{d\omega_2}{dt} = -\frac{\partial S}{\partial \rho_2}, \qquad (2)$$

with the characteristic function S determined by the equality

$$S = \frac{f^2(m_0 + m)^2}{2L^2} + \mu R(L, \rho_1, \rho_2, \lambda, \omega_1, \omega_2, \mu). \qquad (3)$$

The second summand in Relationship (3) represents the perturbat i function, in which μ denotes the small parameter. Equations (2) admit of the generalized energy integral

$$S = \text{const.} \qquad (4)$$

With $\mu = 0$, Eqs. (2) define an unperturbed elliptical orbit.

Integration of differential equations of the form (2), which is accomplished in celestial mechanics by one of the methods of successive approximation, is usually complicated by the small denominators, which give rise to secular terms. The question arises as to whether the secular terms are due to the physical nature of the problem or whether they are in some cases a consequence of inadequacies of the methods employed. Of particular interest from the standpoint of cosmogony is the nature of the time variation of the semimajor axis a. If a remains bounded throughout the entire time of motion, it can be affirmed that the motion is Lagrange stable. It is known, however, that mixed terms appear among terms of second order with respect to the perturbing masses in perturbations of the semimajor axes of planetary orbits, while the third-order terms contain purely secular terms. Hence we cannot draw any conclusions regarding the boundedness of the planetary motions from analytical motion theory. Considerable progress has been made by V.I Arnold, who succeeded in proving the metric analogue to the celebrated Laplace theorem – an analogue consisting in the statement that planetary orbits will be stable for most (in the sense of the measure) initial conditions. Let us show that for celestial-mechanics problems in which a generalized energy integral exists the question of Lagrange stability of the orbits can be solved simply by use of V.V. Rumyantsev's theorem of stability with respect to some of the variables (1).

To investigate the stability of an unperturbed Keplerian motion

$$L = L^{(0)},\ \rho_1 = \rho_1^{(0)},\ \rho_2 = \rho_2^{(0)},\ \lambda = \lambda^{(0)},\ \omega_1 = \omega_1^{(0)},\ \omega_2 = \omega_2^{(0)} \qquad (5)$$

let us supplement System (2) with the equation

$$d\mu / dt = 0 \qquad (6)$$

and assume, furthermore, that R is a bounded function analytical with respect to L and μ in a certain neighborhood of unperturbed values of the Poincare elements.

Let us introduce the following nomenclature for the perturbations of the variables:

$$x_1 = L - L^{(0)},\ x_2 = \mu,\ x_3 = \rho_1 - \rho_1^{(0)},\ x_4 = \rho_2 - \rho_2^{(0)},$$
$$x_5 = \lambda - \lambda^{(0)},\ x_6 = \omega_1 - \omega_1^{(0)},\ x_7 = \omega_2 - \omega_2^{(0)}. \qquad (7)$$

Then we can indicate the first two integrals for the differential equations of perturbed motion, which we shall not write out in the new variables:

$$V_1 = \frac{f^2(m_0+m)^2}{2}\left[\frac{1}{(L_0+x_1)^2} - \frac{1}{L_0^2}\right] + x_3 R,$$
$$V_2 = x_5. \qquad (8)$$

Let us show that the motion will be Lyapunov stable with respect to the quantity L, using V.V. Rumyantsev's theorem for the proof. For this purpose, we construct the Lyapunov function in the form of a bundle of integrals:

$$V = V_1^2 + AV_2^2, \qquad (9)$$

where A is an arbitrary constant. Expanding the right-hand member of Formula (9) in series in powers of the perturbations x_1 and x_2, we obtain

$$V = \frac{f^4(m_0+m)^4}{L_0^6} x_1^2 - \frac{2f^2(m_0+m)}{L_0^3} R_0 x_1 x_3 + (R_0^2 + A) x_3^2 + \ldots \qquad (10)$$

The condition that the function V be sign-determinate with respect to the quantities x_1 and x_2 when the latter are rather small is written in the form

$$\frac{\mu^4(m_0+m)^2}{L_0^6}A>0 \qquad (11)$$

from which it is evident that when A > 0, the function V satisfies Condition (11). Since by virtue of the differential equations of perturbed motion, the derivative of V is identically equal to zero, the motion will be stable with respect to L. It follows from Formulas (1) that the motion will also be stable with respect to the semimajor axis of the orbit.

As can easily be shown, the result obtained has bearing on the problem of motion of a satellite of a slightly flattened spherical planet, on the three-dimensional circular three-body problem (in the case of motion in the vicinity of one of the attracting masses), on the problem of satellite motion in the gravitational field of a slowly rotating solid body whose central ellipsoid of inertia is close to spherical, and so forth.

If the perturbing function S satisfies the condition

$$\frac{\partial S}{\partial l}=\frac{\partial S}{\partial \omega_1}+\frac{\partial S}{\partial \omega_2}, \qquad (12)$$

which, in the problem of satellite motion, signifies the existence of axial symmetry in the perturbing force field, then Eq. (2) admits of yet another first integral in addition to (4):

$$L-\rho_1-\rho_2=\text{const.} \qquad (13)$$

The integral (13) corresponds to the area integral. In this case it can be proven by reasoning analogous to that given above that the motion of the celestial body in the presence of constantly operating perturbations that satisfy Condition (12) is stable with respect to the Delone canonical elements L and H. One consequence of this result is that for ρ_2 = 0 (i = 0), the trajectory is included within a circular annulus whose boundaries are only slightly deformed by small perturba-

tions of the form (12).

Received

20 March 1964

REFERENCE

1. V.V. Rumyantsev, Vestnik MGU, seriya mat., mekh., fiz., khim. [Herald of the Moscow State University, Series for Mathematics, Mechanics, Physics and Chemistry], No. 4, 1957.

UDK 521.312:517.912.2

ON THE STABILITY OF SATELLITE ORBITS WITH CONSTANTLY OPERATING PERTURBATIONS

V.G. Demin

The Kolmogorov-Arnold method is used to prove stability of planet-satellite orbits and the conditionally periodic nature of the motion on the assumption that the satellite's motion is perturbed by constantly operative small forces of a conservative nature.

1. New methods for qualitative analysis of Hamiltonian systems [1] are opening new possibilities for the investigation of a number of problems of classical and celestial mechanics. Here, the nature and complexity of the investigation of the dynamic problem will depend essentially on the properties of the characteristic function H_0 that defines the unperturbed motion. The analysis becomes particularly complicated in those cases in which degeneration occurs, when

$$\det\left[\frac{\partial^2 H_0}{\partial p_i \partial p_j}\right] = 0. \qquad (1)$$

In the choice of the two-body problem as the unperturbed problem, which is traditional for celestial mechanics, Relationship (1) is identically satisfied, with the intrinsic degeneracy combining with limit degeneracy.

If, however, we do not wish to bow to rooted tradition, we can, in many cases, break the Hamiltonian of the perturbed problem down into two parts in such a way as to eliminate both intrinsic and limit degeneracy. The classical problem of two fixed centers [2, 3], the generali-

zed problem of two fixed centers [4, 5] and one of its limit variants [6] may be adopted as the undisturbed problem for this purpose in various problems of celestial mechanics. Below we set forth the limit version of the problem of two stationary centers, as considered in [6] for the requirements of qualitative satellite motion analysis.

2. Let us study the motion of a satellite, assumed to be a material point, in a coordinate system with fixed axis directions, taking the planet's axis of rotation as the third axis, while the basic plane is parallel to the equatorial plane of the planet. We shall place the coordinate origin at one of the umbilical points of inertia that the planet would have if we disregarded its equitorial flattening.

In the system of coordinates selected, the gravitational potential of the planet is expressed by the formula

$$V = \frac{fm}{r} + \frac{fm\delta}{r^3} + R(r, \varphi, \lambda). \qquad (2)$$

Here f is the gravitational constant, m is the mass of the planet, δ is the vertical coordinate of the umbilical point of inertia r, φ and λ are the spherical coordinates of the satellite in the selected reference system and R is the perturbation function. We note that the nature of the perturbing forces is nonessential for the investigations to follow, so that the function R may incorporate both the disturbing effect of the planet's figure and other disturbing factors. Let us assume that the perturbations are quite small. We also point out that the first two terms of Formula (2) take full account of perturbations due to the second spherical harmonic in the expansion of the planet's potential.

Selecting, as the generalized coordinates, the quantities

$$q_1 = r, \quad q_2 = \varphi, \quad q_3 = \sin\lambda \qquad (3)$$

we shall have the following expression [7] for the total integral of the Hamilton-Jacobi equation in the unperturbed problem under consideration,

according to the Stockel principle:

$$W = -\alpha_1 t + \sqrt{2}\alpha_3 \arcsin q^3 + \sqrt{2}\int \sqrt{\alpha_1 q_1^2 + fmq_1 + \alpha_2}\frac{dq_1}{q_1} + \\ + \sqrt{2}\int \sqrt{(\alpha_2 + fm\delta \sin q_2)\cos^2 q_2 - \alpha_3}\frac{dq_2}{\cos q_2}, \tag{4}$$

where α_1, α_2 and α_3 are arbitrary constants of integration.

In qualitative analysis of solutions, it is more convenient to use the Hamiltonian unperturbed-motion system with a special selection of the canonical variables:

$$\frac{d\eta_i}{dt} = 0, \quad \frac{d\xi_i}{dt} = -\frac{\partial H_0}{\partial \eta_i}, \tag{5}$$

where $H_0(\eta_1, \ldots, \eta_n)$ is the Hamiltonian of the perturbed problem. Action-angle variables form such a system of canonical variables.

The unperturbed problem under consideration here represents a particular case of a dynamic system according to Stekkel'. Consequently, the satellite's motion will be conditionally-periodic, and the action-angle variables will be expressed in terms of elementary periods ω_{ij}, which, by virtue of (4), are determined by the relationships [7]:

$$\omega_{11} = \frac{1}{\sqrt{2}}\int_{a_1}^{b_1}\frac{q_1 dq_1}{\sqrt{P_1(q_1)}}, \quad \omega_{12} = -\frac{1}{\sqrt{2}}\int_{a_1}^{b_1}\frac{dq_1}{q_1\sqrt{P_1(q_1)}}, \quad \omega_{13} = 0,$$

$$\omega_{21} = 0, \quad \omega_{22} = \frac{1}{\sqrt{2}}\int_{a_2}^{b_2}\frac{\cos q_2 dq_2}{\sqrt{P_2(\cos q_2)}}, \quad \omega_{23} = -\frac{1}{\sqrt{2}}\int_{a_2}^{b_2}\frac{dq_2}{\cos q_2\sqrt{P_2(\cos q_2)}}, \tag{6}$$

$$\omega_{31} = 0, \quad \omega_{32} = 0, \quad \omega_{33} = \frac{1}{\sqrt{2\alpha_3}}\int_{-1}^{1}\frac{dq_3}{\sqrt{1-q_3^2}}.$$

In Formulas (6), a_1 and b_1 are roots of the equation

$$P_1(q_1) = \alpha_1 q_1^2 + fmq_1 - \alpha_2 = 0, \tag{7}$$

between which the varible q_1 is included in the motion process; a_2, b_2 are analogous roots of the equation

$$P_2(\cos q_2) = \cos^2 q_2(fm\delta \sin q_2 + \alpha_2) - \alpha_3 = 0. \tag{8}$$

The canonical variables ξ_1 are connected with the elementary periods by the formulas

$$\xi_1 = \frac{\pi}{\omega_{11}}(\beta_1 - t),$$
$$\xi_2 = \frac{\pi}{\omega_{22}}\left[\beta_2 - \frac{\omega_{12}}{\omega_{11}}(\beta_1 - t)\right], \qquad (9)$$
$$\xi_3 = \frac{\pi}{\omega_{33}}\left\{\beta_3 - \frac{\omega_{23}}{\omega_{22}}\left[\beta_2 - \frac{\omega_{12}}{\omega_{11}}(\beta_1 - t)\right]\right\},$$

where β_1 denote canonical constants adjoint to α_1. We shall not write out the explicit expressions for the canonical variables η_1, since they are not needed for the reasoning to follow. We note only that the generating function of the contact transformation is written as follows:

$$S = \alpha_1 t, \qquad (10)$$

where α_1 should be regarded as a function of the new variables η_1, η_2 and η_3.

The canonical equations of the perturbed motion in these variables take the form

$$\frac{d\xi_i}{dt} = \frac{\partial K}{\partial \eta_i}, \quad \frac{d\eta_i}{dt} = -\frac{\partial K}{\partial \xi_i}, \qquad (11)$$

where the new characteristic function K is determined by the equality

$$K = R + \frac{\partial S}{\partial t} = R + \alpha_1(\eta_1, \eta_2, \eta_3). \qquad (12)$$

In the new variables, the equations of unperturbed motion will be

$$\frac{d\eta_i}{dt} = 0, \quad \frac{d\xi_i}{dt} = -\frac{\partial \alpha_1}{\partial \eta_i} = \omega_i(\eta_1, \eta_2, \eta_3). \qquad (13)$$

In Eqs. (13), we have used the nomenclature

$$\omega_1 = -\frac{\pi}{\omega_{11}}, \quad \omega_2 = \frac{\pi\omega_{12}}{\omega_{11}\omega_{22}}, \quad \omega_3 = \frac{\pi\omega_{23}}{\omega_{11}\omega_{22}\omega_{33}}. \qquad (14)$$

3. Our further discussion will be based on the theorem of A.N. Kolmogorov [1], which can be formulated as follows.

Let the Hamiltonian function (12) be analytic in a certain region $\eta \in G, |\mathrm{Im}\xi| < \rho$ and let the nondegeneracy condition

$$\det\left|\frac{\partial^2 a_1}{\partial\eta_i\partial\eta_j}\right| \neq 0,$$

be satisfied, while $|R| < M$. Then with sufficiently small M, points of the region G, with the exception of the set of measure small with M, will move conditionally periodically along tori approximating tori η = = const.

To apply this theorem, as can be seen easily enough, it is sufficient to show that

$$\det\left|\frac{\partial\omega_i}{\partial\eta_j}\right| \neq 0. \tag{15}$$

Calculation of the Jacobian (15) can be simplified considerably if, instead of the canonical variables η_1, we use the elements a, p and i that we introduced in Reference [6]:

$$a_1 = -\frac{fm}{2a}, \quad a_2 = \frac{fmp}{2}, \quad a_3 = \frac{fm}{2}(p + 2\delta\sin i)\cos^2 i. \tag{16}$$

Then

$$\det\left|\frac{\partial\omega_i}{\partial\eta_j}\right| = \frac{D(\omega_1, \omega_2, \omega_3)}{D(a, p, i)}\frac{D(a, p, i)}{D(\eta_1, \eta_2, \eta_3)}. \tag{17}$$

Since the transformation performed is nonsingular, we shall have instead of (15)

$$\frac{D(\omega_1, \omega_2, \omega_3)}{D(a, p, i)} \neq 0. \tag{18}$$

Using Formulas (6) and (16), we can obtain

$$\begin{gathered}\omega_{11} = \frac{\pi}{n}, \quad \omega_{12} = \frac{\pi}{\sqrt{fmp}} \quad \omega_{22} = \sqrt{\frac{2}{fm\delta(\lambda_1 - \lambda_3)}}\,\mathbf{K}(k), \\ \omega_{23} = \frac{1}{\sqrt{2fm\delta(\lambda_1 - \lambda_3)}}\left\{\frac{1}{1-\lambda_1}\Pi(n', k) + \frac{1}{1+\lambda_1}\Pi(n'', k)\right\},\end{gathered} \tag{19}$$

where $n^2 = fm/a^3$, $\vec{K}(k)$ is a complete elliptical integral of the 1st kind, $\Pi(n', k)$, $\Pi(n'', k)$ are complete elliptical integrals of the 3rd kind,

$$k^2 = \frac{\lambda_1 - \lambda_2}{\lambda_1 - \lambda_3}, \quad n' = \frac{\lambda_1 - \lambda_2}{1 - \lambda_1}, \quad n'' = \frac{\lambda_2 - \lambda_1}{1 + \lambda_1} \tag{20}$$

and the quantities λ_1, λ_2 and λ_3 are roots of Eq. (8), which is transformed to

$$2\delta\lambda^3 + p\lambda^2 - 2\delta\lambda - (p\sin i - 2\delta\cos^2 i)\sin i = 0. \qquad (21)$$

It is evident from (14) and (19) that ω_1 does not depend on p and i, so that

$$\frac{D(\omega_1, \omega_2, \omega_3)}{D(a, p, i)} = \frac{\partial\omega_1}{\partial a}\frac{D(\omega_2, \omega_3)}{D(p, i)}. \qquad (22)$$

Expanding ω_2 and ω_3 in Taylor series in powers of the quantity δ/p, which is of the order of $3\cdot10^{-2}$ for near artificial earth satellites and diminishes with increasing parameter p, we shall have

$$\omega_2 = n\left[1 + \frac{\delta^2}{p^2}\left(1 - \frac{7}{4}\sin^2 i\right)\right] + \text{higher-order terms} \qquad (23)$$

$$\omega_3 = -n\left(1 - \frac{\delta}{p}\sin i\right) + \text{higher-order terms}$$

Then, as is easily seen, the Jacobian

$$\frac{D(\omega_1, \omega_2, \omega_3)}{D(a, p, i)} = \frac{2n^2\delta^2\cos i}{p^4}\frac{\partial\omega_1}{\partial a} + \text{higher-order terms} \qquad (24)$$

will be nonzero with $i \neq 90^\circ$ and for sufficiently small values of the quantities δ. And this means that the motion of the planet's satellite is conditionally periodic motion and proceeds along tori approximating the unperturbed tori

$$\eta_i = \text{const},$$

where the quantities η_i coincide for $\delta = 0$ with the Poincare canonical elements.

In the event that the perturbing function R is not a function of the longitude λ, we can obtain a stronger result: the motion of the satellite will be conditionally periodic and Lagrange stable for arbitrary initial conditions. For the proof, it is first necessary to ignore the cyclical coordinate λ and then repeat reasoning analogous to that given above.

The author acknowledges his indebtedness to the staffs of the

Celestial Mechanics and Gravimetry Department of the Moscow State University and the Theoretical Mechanics Faculty of the Patrice Lumumba University of Friendship between Peoples for the discussion of this paper.

Received
20 March 1964

REFERENCES

1. V.I. Arnol'd, In collection "Problemy dvizheniya iskusstvennykh nebesnykh tel" [Problems of the Motion of Artificial Celestial Bodies], Izd-vo AN SSSR [Academy of Sciences USSR Press], 1963, page 7.
2. V.G. Demin, Astronom. zh. [Astronomical Journal], 37, No. 6, 1068, 1960.
3. V.G. Demin, Astronom. zh. [Astronomical Journal], 38, No. 1, 157, 1961.
4. Ye.P. Aksenov, Ye.A. Grebenikov and V.G. Demin, Collection entitled "Iskusstvennyye sputniki Zemli" [Artificial Earth Satellites], No. 8. Izd-vo AN SSSR [Academy of Sciences USSR Press], 1961, page 64.
5. Ye.P. Aksenov, Ye.A. Grebenikov and V.G. Demin, Astronom. zh. [Astronomical Journal], 40, No. 2, 363, 1963.
6. V.G. Demin, Sobshcheniya GAISH [Communications of the Shternberg State Institute of Astronomy], No. 125, 3, 1962.
7. C.L. Charlier, Die Mechanik des Himmels [The Mechanics of the Heavens], Vol. I, II, Leipzig, 1902, 1907.

UDK 537.591 (047)

CERTAIN PROBLEMS AND FUTURE PROSPECTS FOR INVESTIGATIONS OF COSMIC RAYS

N.L. Grigorov, I.D. Rapoport, I.A. Savenko and G.A. Skuridin

This is a consideration of the basic problems involved in the operation of an ionization calorimeter – an instrument used to measure the energy of cosmic-ray particles and the relationship between the basic parameters of the ionization calorimeter and the conditions of its utilization.

In addition, there is a consideration of the possibilities of employing an ionization calorimeter to study a number of the characteristics of reactions between atomic nuclei and cosmic-ray particles exhibiting energies of 10^{11}-10^{13} ev [electron-volts], for the study of the composition of high-energy primary cosmic-ray particles (10^{11}-10^{14} ev), for the study of the electron component of primary cosmic rays, and for the study of high-energy γ-radiation.

INTRODUCTION

During the past 30 years, cosmic rays have been subjected to intensive research. The interest in these rays is based primarily on two factors.

First of all, primary cosmic radiation represents a flow of atomic nuclei (with H and He nuclei predominating), exhibiting tremendous energies up to 10^{19} ev, and accordingly this represents the only means of studying nuclear interactions at high and superhigh energies.

In the second place, since cosmic rays are the output of certain accelerating mechanisms "operating" in the galaxies, they carry infor-

mation with respect to the characteristics of these mechanisms. In their motion from the source, the particles of cosmic rays are subject to the action of magnetic fields and they undergo nuclear reactions with atoms from the interstellar medium, generating secondary radiation (π-mesons whose decay yields an electron-positron component and high-energy γ-quanta). Therefore the study of the primary cosmic rays – the study of their energy spectrum, their chemical composition, of the features of propagation through the Galaxy, and of the electron-photon component – yields information on the processes of cosmic-ray acceleration, on the electromagnetic properties of the interstellar medium, on the distribution of cosmic rays within the Galaxy, i.e., it provides important information about one of the components of the Galaxy, playing an active role in the evolution of the latter.

The particularly fruitful period of about 8 years (1945-1953) of intensive investigations into nuclear reactions involving cosmic-ray particles reached its conclusion with a number of fundamental discoveries. It was established that new particles – primarily π-mesons – are generated in reactions between atomic nuclei and primary cosmic particles exhibiting energies of 10^9-10^{10} ev. A whole Pleiades of new ("strange") particles was uncovered. The basic energy characteristics of the nuclear reactions at moderate energies were clarified. As a result of these discoveries there came into being a new branch of physics – the physics of elementary particles.

The last ten years of investigations in the area of cosmic rays have been concentrated primarily on the solution of the following problems: the study of the characteristics of cosmic-particle reactions with atomic nuclei at energies of 10^{11}-10^{13} ev; the study of processes involved in interaction at superhigh energies; the study of the energy spectrum of primary cosmic rays; the study of certain astrophysical

problems involving cosmic rays, i.e., the motion of cosmic-ray particles of various energies within the Galaxy and the solar system; the study of the electron component of primary cosmic rays, and the search for γ-rays.

In our opinion the extremely modest results produced by investigations in connection with cosmic radiation during the past ten years, at least so far as this concerns the study of the mechanism of interactions between atomic nuclei and high-energy particles, are associated with the fact that the investigators, on observing the interaction of high-energy particles and determining the parameters which could be measured, in none of the available methods ever knew with sufficient accuracy the energy of the primary particle responsible for the observed interaction. As a result, in the overwhelming majority of cases the experimental data were never unequivocally interpreted.

In order radically to change the situation in the investigation of high- and superhigh-energy cosmic rays and significantly to raise the effectiveness of their utilization in the study of nuclear reactions in the energy region of 10^{11}-10^{15} ev, it appears to us that two conditions must be satisfied:

1. Experiments to study primary cosmic rays of high and particularly of superhigh energy (10^{14}-10^{15} ev) must be conducted beyond the limits of the atmosphere.

2. A method must be employed to make possible the determination of the energy of the primary cosmic-ray particles with the required accuracy over the entire range under study (10^{11}-10^{15} ev).

The first condition may be satisfied by conducting experiments aboard artificial satellites of the earth.

The second condition may be satisfied by using new principles of measuring particle energies, i.e., principles which could be applied

over the entire range of energies to 10^{15} and even 10^{16} ev.

The existing classical method of measuring the energy of a charged particle on the basis of the distortion of its trajectory in a powerful magnetic field enables us, in conjunction with a Wilson chamber, to determine energy in the region below 10^{10} ev. The use of spark chambers, providing considerably greater accuracy in the measurement of particle-trajectory curvature, apparently makes it possible to expand the region of measureable energies to 10^{11} ev, while the utilization of tremendous magnets weighing hundreds and thousands of tons or the use of solenoids with superconductive windings will permit raising the limit of measureable energies to 10^{12} ev. Thus, these methods are completely useless for the solution of problems involving the measurement of particle energies of 10^{15} ev.

Apparently, the only available means at the present time for the solution of this problem involves the utilization of the new method proposed by N.L. Grigorov, in 1954; this is the method of the ionization calorimeter. This method was first employed in 1957 in the Cosmic-Ray Laboratory of the MGU [Moscow State University] by N.L. Grigorov, I.D. Rapoport, and V.S. Murzin [1] and has now become the basic (at least in the USSR) method used in cosmic-ray research at mountain peaks. As will become evident from the following, this method is quite flexible, i.e., it permits with equal accuracy the measurement of energies of charged and neutral particles, and also makes it possible to combine this measuring instrument with the most varied of particle-observation facilities: the Wilson chamber, the spark chamber, and even with nuclear photoemulsions. A particular advantage of this method in connection with cosmic rays is the possibility of constructing extremely "fast" installations.

There are as yet no special descriptions of such instruments in

the literature. We believe it expedient to fill this gap and to discuss the basic principles involved in the construction of ionization calorimeters, as well as to describe their structural features in connection with their operating conditions. Section I of this article is devoted to that undertaking.

I. THE IONIZATION CALORIMETER

1. Operating Principle

Extensive use is made in nuclear physics of a method of measuring the energy of a discrete heavy particle on the basis of the total ionization developed by that particle in an ionization chamber on deceleration. This method is based on the fact that a heavy particle, moving through matter, loses all of its energy to ionization. It is evident that a prerequisite for the applicability of the method is the fulfillment of the requirement that the mean free path of the particle in the ionization chamber be smaller than the dimensions of the chamber. With the entry of luminescent techniques, the sphere of applicability for this method has expanded.

Since the density of the scintillator is greater than the density of the gas by a factor of approximately 10^3, it has become possible on the basis of the total energy liberated in the material to measure the energy of particles exhibiting a total mean free path expressed in tens of $g \cdot cm^{-2}$, i.e., exhibiting energies of the order of 10^8 ev for heavy particles and of the order of 10^7 ev for electrons.

Can the applicability of this method be extended to include particles exhibiting energies of 10^{11}-10^{15} ev? To provide an answer to this question we must consider, even if only briefly, the mechanism by which such high-energy particles lose their e ergy

From the very outset, let us exclude the μ-mesons from our consideration, since these particles react somewhat too weakly with matter,

losing energy primarily only as a result of ionization. The high-energy μ-meson therefore exhibits such large mean free paths in dense material that it is virtually impossible to achieve its total deceleration in a layer of matter that is not overly thick. Electrons and γ-quanta, given a sufficiently thick layer of matter, expend all of their energy on the ionization of the atoms contained in the matter.

In connection with the high-energy nuclear-active particles, we know from a study of cosmic rays that a stream N(p) of such particles is strongly absorbed in the atmosphere, diminishing with depth in accordance with the law

$$N(p) = N_0 \exp(-p/L_p),$$

where p is the atmospheric pressure at the point of observation, L_p is the absorption mean free path equal to ~ 120 $g \cdot cm^{-2}$. Strong absorption of nuclear-active particles in the atmosphere indicates intensive loss of energy by these particles in their motions through matter. On which processes is the energy of a primary nuclear-active particle expended?

We know that the cross section of interaction for high-energy particles with atomic nuclei is close to the geometric cross section of the nucleus. As a result of the interaction, a fraction of the energy of a primary particle is expended on the formation of new particles (primarily π-mesons), and a small fraction of the energy is spent on the splitting of the nucleus with which the particle collided, the remaining energy staying with the primary particle which, on subsequent interaction, again will lose a fraction of its energy in analogous processes, etc. The generated charged π-mesons and similar particles with a lifespan of $\tau_0 \gtrsim 10^{-10}$ sec will interact in dense matter with nuclei, much like a primary particle, generating secondary particles and fissioning nuclei.

Since among the generated π-mesons there are generated π^0-mesons

with a lifespan of $\tau_0 \sim 10^{-16}$ sec, which decay into two γ-quanta virtually at the point of their formation, the energy of the primary nuclear-active particle is ultimately transformed into the energy of electromagnetic radiation (γ-quanta) and into the energy of strongly ionizing particles – the products of nuclear fission.

The energy of the γ-quanta in matter is totally expended on the ionization of the atoms of the material, just as was the case with the energy of the strongly ionizing particles.

If we take a block of dense material that is 1000 $g \cdot cm^{-2}$ thick and of sufficient lateral dimensions to prevent the passage of the secondary particles beyond the limits of the block through the side surfaces, a high-energy particle falling on such a block would be absorbed by it. In this case virtually all of the energy will be expended on the ionization and excitation of the atoms of the material and will, in the final analysis, be converted into heat.

If we were to measure the quantity of heat liberated in the block at the instant at which a nuclear-active particle impinged upon it, it would be possible to obtain an exact value of the energy E_0 of the incident particle. This method of measuring energy cannot be utilized for a great many reasons. However, it is possible to use the measurement of an intermediate quantity for the determination of E_0, this quantity being proportional to the final thermal effect, i.e., the total number of ion pairs I appearing in the block as a result of the incidence onto the block of a particle exhibiting an energy E_0.

It is clear that $E_0 = \lambda I$, where λ is the mean magnitude of the energy expended on the formation of a single ion pair in the material of the block.

Because of the similarity in the measurement of energy E_0 in a discrete particle on the basis of the total ionization effects and the

calorimetric measurements, the instrument used to accomplish this method of determining E_0 is known as an ionization calorimeter [1].

If we know the ionization distribution I(x) for the depth x in the absorbent of the ionization calorimeter we can determine the total ionization effect or the magnitude of I.

Indeed, if at the depth x (calculated from the upper base of the ionization calorimeter) I(x) dx ion pairs are formed in a layer dx thick at the instant at which a nuclear-active particle impinges on the ionization calorimeter, the quantity I may be defined from the obvious equality

$$I = \int_0^\infty I(x)dx$$

and, consequently,

$$E_0 = \lambda \int_0^\infty I(x)dx \simeq \lambda \int_0^{x_0} I(x)dx. \qquad (1)$$

For the correct determination of E_0 the thickness x_0 must be taken so that all secondary particles are decelerated in the ionization calorimeter, i.e., we must have $I(x_0) \simeq 0$.

Equation (1) superficially coincides with Eq. (2)

$$\overline{N \cdot E} = \lambda \int_0^\infty I(x)dx, \qquad (2)$$

which has been extensively used for a long time both in the study of cosmic radiation and in the dosimetry of penetrating radiation to measure the flow of energy contained in a large number of particles. However, Expressions (1) and (2) are basically different from one another, and this controls the experimental facilities available for the measurement of I(x) in the first and second cases.

If we want to measure the energy of a discrete particle, we must measure the ionization I(x) at all points of the absorbent at the iden-

tical instant of time corresponding to the incidence of the discrete particle onto the absorbent. In other words, the instantaneous distribution of ionization over the entire volume of the absorbent must be measured. Consequently, pulse ionization detectors should be used as the measurement instruments (the permissible measurement durations of the instantaneous distribution of ionization I(x) will be considered below).

When measuring a flow of energy carried by a large number of particles statistically distributed in time, the time-average of the quantity I(x) at each point of the absorbent must be measured. A prolonged measurement of I(x) is required for this purpose (the duration of the measurement will be defined by the assumed accuracy which will be a function only of the statistical fluctuations in the stream of particles at the point $\underline{x}$), i.e., in principle, the ionization detector must be of the integral type. If the stream of particles is steady, the measurements at the various points of the absorbent may be conducted at various times.

2. Parameters of the Ionization Calorimeter

In considering the ionization-calorimeter parameters which govern its operation, we will proceed under the assumption that this instrument is used for work with cosmic-ray particles.

The basic parameters of the ionization calorimeter are defined:

a) by physical processes as a result of which the energy of the primary particle, in the final analysis, is expended on ionization within the material making up the absorbent;

b) by the technical facilities of ionization measurement within the absorbent;

c) by the possible factors leading to "nondetection" of a fraction of the energy or by random coincidences;

d) by the features inherent in the cosmic-ray particles of high energy.

Let us consider the effect of the above-enumerated factors on the parameters of an ionization calorimeter. Let us begin with point (d).

In the design of the ionization calorimeter our main concern obviously involves, on the one hand, provision of the most complete transition of the energy of the primary particle to the energy of the electron-photon components and other ionization components within the confines of the ionization calorimeter and, on the other hand, to make certain that it is the energy liberated by only a single nuclear-active particle that is measured, eliminating or separating all other sources of energy liberation in the ionization calorimeter.

There are a number of features in work at sea level and on mountain peaks which must be taken into consideration in the design of the instrument.

One of these features involves the fact that nuclear-active high-energy particles frequently form groups in the atmosphere, i.e., several particles impinge simultaneously on the installation [2]. The greater the area of the installation, the greater the probability of the simultaneous incidence of two, three, and more nuclear-active particles. It is obvious that in the case of simultaneous incidence of several particles on the installation, the energy which is liberated in the ionization calorimeter will be equal to the sum of the energies of all particles impinging on the installation.

Evidently, to ensure reliable measurement of the energy of an individual nuclear-active particle, the design of the ionization calorimeter must make it possible to distinguish the incidence onto the calorimeter of a discrete nuclear-active particle, a group of nuclear-active particles, or the incidence of an extensive atmospheric shower,

etc. This condition may be satisfied if the ionization calorimeter is designed to consist of a large number of ionization detectors distributed appropriately (see below).

Let us now consider the process of the transition of energy from a nuclear-active particle to the electrons of an electron-photon cascade. In this case we will, for the time being, neglect the energy losses due to nuclear fission.

Let a nuclear-active particle experience an interaction at some point O_1 in the absorbent of the ionization calorimeter (Fig. 1). Consequently, at this point a fraction of the energy $(E_{\pi 0})_1$ was transferred to the π^0-mesons, while the remaining part, i.e., $E_0 - (E_{\pi 0})_1$ was retained by the nuclear-active particles: the nucleons, the $\pi^{\pm}$-mesons, etc. In view of the extremely small lifespan of the π^0-mesons, in practical terms it may be held that at the point O_1 there appeared γ-quanta distributed over some spectrum $dN_1(E_\gamma)/dE_\gamma$ with a total energy $(E_\gamma)_1 = \int_0^{E_0} E_\gamma \frac{dN_1}{dE_\gamma} dE_\gamma$. Each of the γ-quanta in the material of the calorimeter produces its own cascade shower of electrons. The entire aggregation of γ-quanta produced in the first interaction yields the total shower of $I_1(x)$.

Since for each "elementary" shower (produced by individual γ-quanta) the [following] equality is valid

$$E_\gamma = \lambda \int_0^\infty i(x)dx,$$

for the entire aggregation of γ-quanta the following equality is satisfied

$$(E_{\pi^0})_1 = \sum E_\gamma = \lambda \int_0^\infty \sum i(x)dx = \lambda \int_0^\infty I_1(x)dx. \qquad (3)$$

However, in addition to the π^0-mesons, secondary nuclear-active particles exhibiting a total energy $E_0 - (E_{\pi 0})_1$ will move from point O_1

in the direction of motion of the primary particle. One of the secondary particles may be subjected to interaction at some point O_2, as a result of which the energy $(E_{\pi 0})_2$ will be transferred to the π^0-mesons. The γ-quanta of these π^0-mesons will yield a certain total cascade shower of $I_2(x)$ for which Relationship (3) is valid

$$(E_{\pi^0})_2 = \lambda \int_0^\infty I_2(x)dx.$$

For any kth interaction accompanying the transfer of energy $(E_{\pi 0})_k$ to the π^0-mesons, the following equation will obviously be satisfied

$$(E_{\pi^0})_k = \lambda \int_0^\infty I_k(x)dx.$$

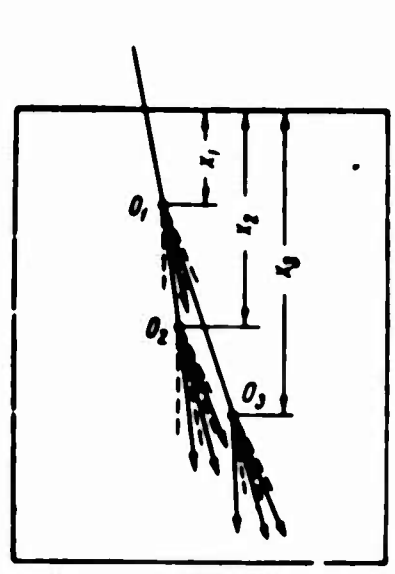


Fig. 1. Diagram clarifying transfer of energy E_0 from primary particle to the electron-photon component. O_1, O_2, O_3) Points of interaction for nuclear-active particles in ionization calorimeter. The dashed lines show the γ-quanta produced by the decay of the generated π^0-mesons, and the solid arrows show the secondary nuclear-active particles.

Fig. 2. Diagram showing development of partial electromagnetic cascades in an ionization calorimeter. The depth x of the calorimeter has been plotted along the axis of abscissas, and the ionization produced by the partial shower at the depth x is plotted along the axis of ordinates;

x_1, x_2, x_3) are the depths of points 0_1, 0_2, 0_3 (see Fig. 1).

With each interaction of secondary nuclear-active particles, a fraction of the energy will, first of all, go over into electromagnetic radiation (through the formation of π^0-mesons) and, secondly, degradation of the energy of discrete particles will occur (as a result of the generation of several nuclear-active particles with each interaction and the irreversible losses of energy to the formation of the π^0-mesons). Figure 2 shows this process schematically, the depths of the points 0_1, 0_2, ..., of nuclear-active particle interactions being plotted along the axis of abscissas, the magnitude of the ionization resulting from the cascade electrons due to the primary, secondary, etc., interactions being plotted along the axis of ordinates.

If the process of π^0-meson production were the only process by which the nuclear-active particles lost energy, the entire energy of the primary particle would eventually convert to the energy of the generated π^0-mesons. In this case,

$$E_0 = \sum_h (E_{\pi^0})_h = \lambda \int_0^\infty \sum I_h(x)\,dx.$$

Since

$$I(x) = \sum I_h(x),$$

$$E_0 = \lambda \int_0^\infty I(x)\,dx.$$

In order to evaluate the thickness x_0 of the absorbent used in the ionization calorimeter, in which the greatest fraction of the primary-particle energy converts to the energy of the π^0-mesons, we should undertake a more detailed examination of the manner in which, on the average, this process runs its course.

Let us assume that all of the nuclear-active particles, regardless

of their nature, exhibit an average mean free path L for interaction within the material making up the absorbent. Let us further assume that all of the nuclear-active particles, regardless of their nature, in each interaction on the average lose the identical fraction $\bar{k}$ of their energy to the formation of the π-mesons (for the time being, let us neglect losses in energy due to nuclear fission).

Let the spectrum of the nuclear-active particles in the shower formed by the primary particle exhibiting energy E_0 have the form $dE[\partial N(E, x)/\partial E]$ at a depth $\underline{x}$ calculated from the upper face of the ionization calorimeter. In the dx layer all of these particles will transfer energy to the π^0-mesons, equal to

$$dE_{\pi^0} = \frac{dx}{L}\int_0^{E_0} \frac{k}{3} E \frac{\partial N(E, x)}{\partial E} dE.$$

Here

$$\int_0^{E_0} E \frac{\partial N(E, x)}{\partial E} dE = E_{\text{n.a}}(x),$$

where $E_{ya.a.}(x)$ is the total energy of all nuclear-active particles at the depth $\underline{x}$.

Since the dissipation of the energy of the nuclear-active particles occurs only as a result of the transfer of energy to the π^0-mesons (within the framework of the adopted assumptions), $dE_{\pi^0} = -dE_{\text{n.a}}(x)$.

Consequently,

$$dE_{\text{n.a}}(x) = -\frac{k}{3} E_{\text{n.a}}(x) \frac{dx}{L}, \qquad E_{\text{n.a}}(x) = A \exp\left(-\frac{kx}{3L}\right), \quad (4)$$
$$E_{\text{n.a}}(0) = E_0 \text{ for } x = 0,$$

thus,

$$E_{\text{n.a}}(x) = E_0 \exp\left(-\frac{kx}{3L}\right). \qquad (5)$$

In a layer of finite thickness $\underline{x}$, the energy transferred to the π^0-mesons will be equal to

$$E_{\pi^0}(x) = E_0 - E_{\text{я.а.}}(x) = E_0\left[1 - \exp\left(-\frac{kx}{3L}\right)\right]. \qquad (6)$$

In order for 90% of the primary-particle energy, on the average, to be transferred to the π^0-mesons throughout the entire ionization calorimeter with a thickness x_0, the quantity x_0 must satisfy the condition

$$\frac{E_{\pi^0}(x_0)}{E_0} = 1 - \exp\left(-\frac{kx}{3L}\right) = 0{,}9,$$

i.e.,

$$x_0 = 7L/k. \qquad (7)$$

In the case of interactions with heavy nuclei $k \simeq 1$ [3]. Therefore, the minimum quantity of material exhibiting large Z in the ionization calorimeter must be equal to $x_{0\min} \simeq 7L$.

In materials exhibiting small Z the inelasticity of the nucleon interactions is small. For these $k = 0.3$-0.6 [4, 5]. Although there are no direct experimental data in the literature on the coefficient of inelasticity for the $\pi^{\pm}$-mesons of high energy, on the basis of certain indirect data we can assume that for these $k \simeq 1$. This means that with the interaction of a $\pi^{\pm}$-meson with a nucleus, on the average 1/3 of the energy of the $\pi^{\pm}$-meson impinging against the nucleus is transferred to the π^0-mesons.

If we neglect the energy losses due to nuclear fission, the average energy transfer to the π^0-mesons in the $\underline{x}$ layer of light material by a nucleon having an energy E_0 will be equal to

$$E_{\pi^0}(x) = E_0 - E_{\text{я.а.}}(x).$$

Here $E_{\text{я.а.}}(x)$ is the total energy of all nuclear-active particles at the depth $\underline{x}$ of the ionization calorimeter. Since in the given case the coefficients of inelasticity for the primary particle (the nucleon) and for the secondary particles (the $\pi^{\pm}$-mesons) vary, Eq. (4)-(7) in

the given case will not be valid. A fraction of the energy of all of the nuclear-active particles will be carried by the $\pi^{\pm}$-mesons ($E_{\pi}(x)$) and another fraction of the energy will be carried by the secondary nucleons ($E_n(x)$), i.e.,

$$E_{n.a}(x) = E_n(x) + E_n(x).$$

If we assume that no energy is transferred from the $\pi^{\pm}$-mesons to the nucleon component of the nuclear cascade, the change in the total energy of the nucleon component will be described by the equality

$$dE_n(x) = -kE_n(x)\frac{dx}{L}.$$

Since a "primary" nucleon with an energy of E_0 impinges on the boundary of the calorimeter, the energy E_n will be equal to

$$E_n(x) = E_0 \exp\left(-\frac{kx}{L}\right). \qquad (8)$$

The total energy of all $\pi^{\pm}$-mesons at the depth x will change in accordance with two processes:

First of all, with the interaction in the dx layer a fraction of the energy of the $\pi^{\pm}$-mesons will be transferred to the newly generated π^0-mesons,

$$dE_n' = -\frac{1}{3}E_n(x)\frac{dx}{L};$$

secondly, as a result of the interaction of the nucleon component of the shower in the dx layer, a fraction of its energy will be transferred to the $\pi^{\pm}$-mesons,

$$dE_n'' = \frac{2}{3}kE_n(x)\frac{dx}{L}.$$

Thus, the total change in the energy flow of all of the $\pi^{\pm}$-mesons in the dx layer will be equal to

$$dE_n(x) = -\frac{1}{3}E_n(x)\frac{dx}{L} + \frac{2}{3}kE_0\exp\left(-\frac{kx}{L}\right). \qquad (9)$$

The solution of this equation with $k \neq 1/3$ is given by

$$E_n(x) = \frac{2kE_0}{1-3k}\left[\exp\left(-\frac{kx}{L}\right) - \exp\left(-\frac{x}{3L}\right)\right] \qquad (10)$$

while with $k = 1/3$, by

$$E_n(x) = 2kE_0\frac{x}{3L}\exp\left(-\frac{x}{3L}\right). \qquad (11)$$

If $k = 1/3$,

$$E_{n.e.}(x) = E_0 \exp\left(\frac{x}{3L}\right)\left(1 + \frac{2}{3}\frac{x}{3L}\right),$$

and in order for $E_{n.e.}(x)/E_0 = 0{,}1$, we must have $x_{0min}/3L = 3{,}5$, i.e.,

$$x_0 \simeq 10{,}5\,L. \qquad (12)$$

However, if $k = 0.5$, $x_{0min} = 9L$. (13)

Until now we have neglected energy losses due to nuclear fission. However, with each interaction of a nuclear-active particle and the nucleus some fraction of its energy is expended on the splitting of a target nucleus. As a result of the fission, charged heavy particles (protons, α-particles) and neutrons are formed. These charged particles expend their kinetic energy in the calorimeter on ionization whereas the neutrons, if they are sufficiently fast, produce new nuclear fission. A certain fraction of the energy evades being recorded during the process of nuclear fission: a) the energy binding the nucleons in the nucleus; b) the energy of the low-energy neutrons leading to nuclear reactions with subsequent β-decay, in which the energy is liberated after a considerable lapse of time following the incidence of the primary particle onto the calorimeter, this particle producing the entire nuclear shower. Therefore, it may be assumed that virtually all of the energy expended on nuclear fission will be liberated for ionization in the calorimeter.

The fact of the occurrence of a great number of instances of nuclear fission in the calorimeter alone does not affect the accuracy of

determining the average magnitude of the primary-particle energy. However, in an individual measurement the nuclear fission may significantly affect the determination of particle energy.

The problem of recording strongly ionizing particles will be considered in greater detail below.

Up to the present time we have considered the physical processes which may lead to a reduction in energy (the energy binding the products of nuclear fission, low-energy neutrons). An opposite process, however, is possible, i.e., an elevation of the primary-particle energy; this elevation is possible because of the finite duration of the measurement of the ionization I(x). To evaluate this effect let us hold that the detectors being employed do not have a "threshold" of registration, i.e., they can measure as small a value of ionization as we please.

If the time required for the measurement of ionization is equal to τ, the energy liberated during this period of time by all of the cosmic particles in the ionization calorimeter is equal to

$$\Delta E = \bar{E}_{\text{к.л.}} S\tau,$$

where $\bar{E}_{\text{к.л.}}$ is the flow of energy of penetrating cosmic-radiation particles passing through 1 cm^2 per second. On mountain tops $\bar{E}_{\text{к.л.}} \simeq 5 \cdot 10^7$ $ev \cdot cm^{-2} \cdot sec^{-1}$.

If the energy of a discrete particle impinging on the calorimeter is equal to E_0, in order not to reduce the magnitude of the measured energy E_0 as a result of a finite period of measurement time, the following condition must be satisfied

$$\Delta E / E_0 \ll 1. \qquad (14)$$

The dimensions of the ionization calorimeter (its area S) are uniquely defined by the energies of those particles with which the investigator proposes to work. Thus, for example, in order to obtain accep-

table statistics involving particles exhibiting energies of 10^{11} ev, the area of the ionization calorimeter set up at an altitude of about 3000 m above sea level must be $\simeq 10^4$ cm^2. Hence $\tau \ll (1/5)$ sec.

If we assume an average elevation of particle energy by 10%,

$$\tau \leqslant 1/_{10} \text{ sec.}$$

It follows from these evaluations that the ionization detectors used in an ionization calorimeter must be rather fast-acting (proportional counters, pulse ionization chambers, luminescent detectors).

Summarizing all of the above, we can formulate the basic technical requirements which must be satisfied by an ionization calorimeter for the measurement of the energy of a discrete nuclear-active cosmic-radiation particle in the following manner.

1. With minimum extent of the ionization calorimeter (along the line of motion for the primary particle) the maximum transformation of primary-particle energy into π^0-mesons must be ensured in the absorbent.

2. The position of the ionization detectors must ensure separation of the physically various instances of the given energy liberation in the ionization calorimeter (a discrete particle, a group of particles, and an atmospheric shower).

3. The ionization detectors must be sufficiently fast-acting, introduce no distortion into the energy and angular distribution of the secondary particles in the absorbent of the calorimeter, and they must make it possible to achieve more complete recording of the ionization produced by the strongly ionizing particles.

3. Selection of the Absorbent Material

Earlier we estimated the thickness x_0 for absorbents of materials exhibiting various atomic numbers; the thickness, on the average, ensures the transfer of 90% of the primary-particle energy to the π^0-mesons. For substances with large Z we obtained: $x_0 = 7L$. For substances

with small Z, x_0 = (9-10.5) L. As we can see, the thickness x_0, expressed in mean free paths of interaction, is a weak function of Z. However, certain physical properties of the substances with various Z (and we have reference here primarily to density) compel preference for the heavy elements.

As a matter of fact, if the density of the substance is denoted by ρ, the extent of the ionization calorimeter in height will equal: $h = x_0/\rho$ (if we neglect the dimensions of the ionization detectors).

For purposes of evaluating x_0, having assumed the values of the mean free paths of interaction known for particles with energies of $\simeq 10^{10}$ ev, we obtain:

for carbon x_0 = 810-950 g·cm^{-2}, h_C = (810-950)/1.6 = 510-590 cm;

for iron x_0 = 900 g·cm^{-2}, h_{Fe} = 900/7.8 = 115 cm;

for lead x_0 = 1400 g·cm^{-2}, h_{Pb} = 1400/11.3 = 125 cm.

The ionization calorimeter must exhibit a certain "speed," this requirement being imposed by the rather frequent observation of cosmic particles of specific energy.

The "speed" of the ionization calorimeter is based on its geometric dimensions. If S_1 and S_2 denote the areas of the upper and lower bases of the calorimeter, and if $\underline{h}$ denotes the distance between these bases, the "speed" of the ionization calorimeter will approximately equal S_1S_2/h^2. We can see from this expression that with a change in $\underline{h}$ S_1 and S_2 must be altered in proportion to $\underline{h}$ in order to retain the constant "speed." Consequently, the use of substances with small Z as absorbents in the ionization calorimeter will cause the dimensions of such a calorimeter to be larger by a factor of approximately 4 than the dimensions of a calorimeter employing an absorbent of heavy material.

When working with an ionization calorimeter at high altitudes, where its weight P is of decisive significance, it is also most exped-

ient to use substances of great density.

Indeed, the ionization calorimeter should exhibit a definite speed Γ based on the conditions of the physical experiment. With $S_1 = S_2$, the speed

$$\Gamma = \frac{S^2}{h^2} = Sh\frac{S}{h^3} = \frac{V\sqrt{\Gamma}}{h^3}.$$

If ρ denotes the density of the calorimeter substance, $h = x_0/\rho$. Consequently,

$$\sqrt{\Gamma} = V\rho^3/x_0^3 = P\rho/x_0^3,$$

i.e., finally

$$P = \sqrt{\Gamma}x_0^3/\rho.$$

From this expression we can see that with a given speed Γ the weight of the ionization calorimeter is approximately inversely proportional to the density of the substance of which the calorimeter is made.

As will become evident from the following, the selection of the substance for the absorbent of the ionization calorimeter is governed in significant measure by the conditions of the development of electron-photon cascades and the methods by which these are recorded. For this reason let us again return to the problem of selecting the absorbent material after having considered the possible methods of recording the total ionization.

4. Methods of Recording Ionization

We can proceed along two basically diverse paths to measure the total ionization produced in a dense absorbent.

1. We can attempt the measurement of ionization at all points of the absorbent.

Such measurements may be carried out by using a scintillator as the absorbent, and a large number of photomultipliers [FEU] as the detector, making it possible to measure the total quantity of light lib-

erated at all points of the scintillator at the instant at which a nuclear-active particle impinges on it. (In this method, of all available procedures, the best recording results would be produced by the energy lost on nuclear fission). There is the possibility of a version (proposed by A.Ye. Chudakov) in which Cerenkov radiation is used in the place of the luminescence due to the charged particles (in this case the strongly ionizing particles are not recorded).

The majority of available scintillators and transparent substances exhibits low density for the utilization of the Cerenkov radiation, ~1 $g \cdot cm^{-3}$, and for this reason the total gathering of light must be accomplished with a volume of several tens of cubic meters. Technically, this is quite difficult. What is even more important is the fact that the integral gathering of light, although yielding a quantity proportional to E_0, will not permit satisfying the important Condition 2 (see Section 2). It therefore seems to us that this method shows little promise in connection with the measurement of the energy of a discrete cosmic-ray particle in the lower part of the atmosphere.

2. It is possible to measure the distribution of ionization (or a quantity proportional to it) through the depth of the absorbent; the individual detectors in this case are positioned at fixed depths of the absorbent (Fig. 3).

Let us consider this possibility in greater detail.

Let the entire absorbent x_0 thick be divided into n layers, each exhibiting a thickness of $x_1 = x_2 = \ldots = x_0/n$. Ionization detectors will be positioned beneath each layer. If these detectors do not introduce any changes into the energy and angular distribution of the shower particles, the ionization measured by the detectors situated at a depth x is uniquely associated with the ionization produced in the absorbent of the calorimeter at the depth x by all of the shower particles, i.e.,

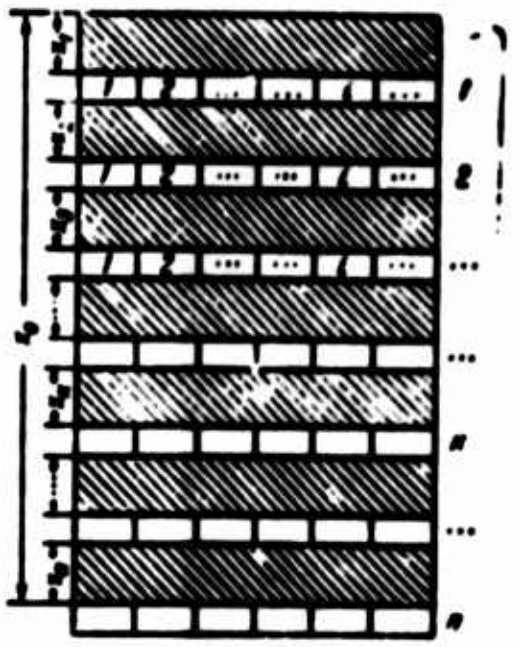


Fig. 3. Diagram showing positions of ionization detectors in ionization calorimeter. 1, 2, 3, ..., n) The number of rows of ionization detectors; x_1, x_2, ..., x_n } layers of absorbent; x_0 } over-all thickness of absorbent in ionization calorimeter.

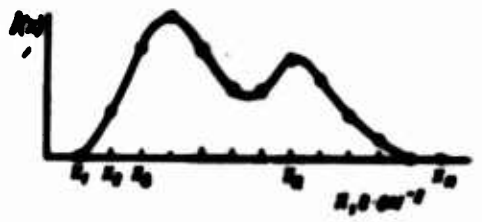


Fig. 4. An example of the distribution of ionization I(x) through the depth of the ionization calorimeter, by instantaneous measurement of this ionization at a finite number of cross sections x_1, x_2, ..., x_k, ..., x_n of the ionization calorimeter.

$$I_k(x) = \frac{\lambda_{det}}{\lambda_{absorb}} \cdot \frac{\beta_{absorb}}{\beta_{det}} \sum_i I^{ik}_{det}(x).$$

where I^{ik}_{det} is the ionization measured by the ith detector situated beneath the kth layer of the absorbent; $I_k(s)$ is the ionization in the calorimeter substance at the depth x; λ is the mean energy expended on the formation of a single ion pair; β is the average loss of energy due to ionization in 1 $g \cdot cm^{-2}$ of material.

Thus, the measurement of ionization by means of a finite number of detectors makes it possible to construct an I(x) curve yielding the instantaneous distribution of ionization over the entire thickness of the ionization calorimeter at the instant of incidence by a high-energy particle (Fig. 4). On the basis of the familiar Relationship (1), the energy of the incident particle is subsequently determined.

It is obvious that the thinner each of the absorbent layers (the greater the number of layers), the more reliable the identification of the average I(x) curve with the true distribution of ionization over all points of the absorbent. However, an increase in n leads to an increase in the number of ionization detectors. For this reason n may not be increased without limit and must be set in accordance with the potentials of existing engineering tech-

niques. The minimum number of rows can be derived from the following consideration.

The main part of the primary-particle energy in the ionization calorimeter is expended on the ionization by electrons of the cascade showers produced by γ-quanta from the disintegration of the π^0-mesons. The halfwidth of the cascade curve for a wide range of γ-quanta energies is of the order of (1.4-1.0) t_{max}, where t_{max} is the thickness of the substance layer in which the electron-photon shower develops to the maximum number of particles. Therefore, in order to avoid great errors in the derivation of the mean ionization-distribution curve through the depth of the ionization calorimeter, the thicknesses of the individual layers should not exceed the halfwidth of the cascade curve, i.e., they should be of the order of t_{max}.

Consequently,

$$x_h = \frac{x_0}{n} \cong t_{max} = \ln \frac{\bar{E}_\gamma}{E_c}.$$

where $\bar{E}_\gamma$ is some assumed energy of the γ-quanta (appearing as a result of the decay of the π^0-mesons), which should be measured by the ionization calorimeter with sufficient accuracy; E_c is the critical energy for the absorbent substance.

Hence the number of layers

$$n = \frac{x_0}{t_{max}} = \frac{x_0}{\ln \bar{E}_\gamma / E_c}. \qquad (15)$$

If the absorbent is made of a light substance, the average loss of energy by the nucleon amounts to (0.3-0.6) E_0 and with each interaction 10-20% of the nucleon energy is transferred to the π^0-mesons. The average multiplicity of produced $\pi^\pm$-mesons at nucleon energies of 10^{11}-10^{12} ev is equal to 5-10 (with $E_0 = 2 \cdot 10^{11}$ ev, $\bar{n}_s = 7$ [6]), and for this reason the average number of γ-quanta formed in each such interac-

tion must also be assumed to be equal to 5-10. Consequently, the average energy of the γ-quanta will be of the order of 1-3% of E_0.

However, if the absorbent is made of a heavy substance (with large Z), the energy losses by the nucleon will be greater, but there exists the possibility that the average multiplicity of generated π-mesons will also be greater. A quantity of the order of (0.01-0.03) E_0 should therefore be taken as the energy of the γ-quanta. (No particular refinement of this quantity is required, since it appears in the logarithm).

TABLE 1

Вещество 1	Показатель ослабления 2	E_c, Мэв 3
Углерод 4	40	72
Железо 5	12.5	18.4
Свинец 6	5.2	6.4

1) Substance; 2) shower unit, $g \cdot cm^{-2}$; 3) E_c, in Mev; 4) carbon; 5) iron; 6) lead.

For the sake of determinacy let us assume $E_\gamma = 0.01\ E_0$ and $E_0 = 10^{12}$ ev. In this case, $E_\gamma = 10^{10}$ ev. The magnitude of the critical energy E_c, just as the shower unit, is expressed in $g \cdot cm^{-2}$ and is a significant function of the atomic number Z of the substance. The values of E_c and the shower unit are presented in Table 1 [7].

To determine the number of rows n of detectors, let us assume the earlier-derived values of x_0. In this case we will have for carbon, iron, and lead, respectively, $n_C = 4$ rows, $n_{Fe} = 11$ rows, and $n_{Pb} = 37$ rows.

In order to be able to distinguish the process by which a given liberation of energy in the ionization calorimeter came about, each row in which the ionization measurement is being carried out must contain a number of ionization detectors. The geometric shape of the detectors and their positions in the rows must be selected so that on the basis of the ionization-distribution pattern in the detectors it becomes possible to reproduce the pattern of particle motion (or the motion of a group of particles) by means of the ionization calorimeter.

The greatest volume of information can be obtained if the detectors are prepared in the form of parallelepipeds of minimum thickness, square in shape, tightly filling the entire area of the calorimeter cross section (Fig. 5). In this case the position and the direction of motion for the nuclear-active particle through the calorimeter is determined in the simplest and most uniquely defined fashion (in Fig. 5 cross-hatching is employed to denote the detectors through which the particles passed and in which the ionization was recorded). Let us take a look at how in the case of such detectors their number will be a function of the sequential number of the calorimeter substance.

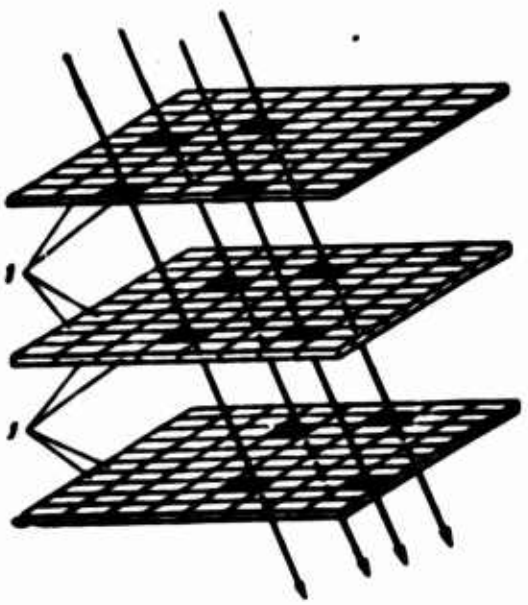


Fig. 5. An example of the possible distribution of ionization (1) detectors in an ionization calorimeter. The cross-hatched squares show the detectors through which passed the shower particles produced by nuclear-active particles.

Let us assume that σ denotes the area of each of the detectors, and that the cross-sectional area of the calorimeter is denoted by S (let us assume that the upper and lower bases of the calorimeter are identical). In this case, in each row we will have $S/\sigma = m$ detectors, while the entire calorimeter will have $m \cdot n$ detectors. Earlier we saw that in order to retain a constant "speed" for the calorimeters made of substances of various density, the cross-sectional area S of the calor-

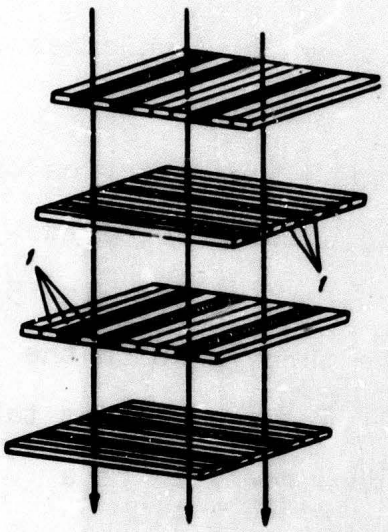


Fig. 6. An example of the possible distribution of ionization (1) detectors, making it possible to reconstruct spatial pattern of passage for a nuclear-active particle through an ionization calorimeter. The cross-hatched strips show the detectors through which passed the shower particles produced by nuclear-active particles.

imeter must be proportional to its height, i.e., S ~ h. Since m ~ S, calorimeters with various substances but of identical "speed" have m·n ~ h·n. Hence

$$\frac{(m\cdot n)_C}{(m\cdot n)_{Fe}} = \frac{h_C\cdot n_C}{h_{Fe}\cdot n_{Fe}} = (4\div 5)\cdot\frac{4}{11} \cong 1{,}5-1{,}8,$$

$$\frac{(m\cdot n)_{Pb}}{(m\cdot n)_{Fe}} = \frac{h_{Pb}\cdot n_{Pb}}{h_{Fe}\cdot n_{Fe}} = \frac{125}{115}\cdot\frac{37}{11} = 3{,}7.$$

Consequently, for the system of ionization detectors under consideration the use of light substances as absorbents yields no advantages in the number of detectors. The minimum number of detectors corresponds to substances exhibiting great density and average atomic weights (substances such as iron and copper).

Although the utilization of ionization detectors similar to those shown in Fig. 5 is quite attractive, at the contemporary level of engineering this is, apparently, difficult to achieve. As a matter of fact, if we accept as permissible the spatial "resolution" of the detectors at 10 cm, the area of such a detector will be 10·10 = 100 cm^2.

If S = 1 m^2, according to the above-cited calculations 1100 ionization detectors would be required for an ionization calorimeter made of iron, while 1600 detectors would be required for a calorimeter made of graphite.

The required minimum number of detectors can be reduced by an order of magnitude without impairing the resolving power of the instrument by using long and thin detectors, distributed in various rows mu-

tually perpendicular to one another (Fig. 6).

In this case the total number of detectors as a function of ionization-calorimeter dimensions will take the form: $m \cdot n \sim \sqrt{h} \cdot n$. If we assume that for a calorimeter with an iron absorbent $\sqrt{h} \cdot n = 1$, for a calorimeter with a carbon absorbent $\sqrt{h_C} \cdot n_C = 0{,}7 \div 0{,}8$ and if the absorbent is of lead $\sqrt{h_{Pb}} \cdot n_{Pb} = 3{,}3$. To evaluate the required number of ionization detectors in a calorimeter with an iron absorbent we assume that S = 1 m^2, with the width of each detector 10 cm. In this case m = 10 and n = 11 (see the calculations for the row number n). In other words, 110 detectors will be required. Given a calorimeter of the same "speed" with an absorbent made of graphite, ~80 detectors will be required, while for a calorimeter with a lead absorbent 360 detectors will be needed.

As can be seen from the calculations which have been carried out, the utilization of substances with low atomic numbers as absorbents (in ionization calorimeters with the given "speed" and the given spatial detector resolution) yields no significant advantage with regard to the number of required detectors.

Light substances used as absorbents in ionization calorimeters exhibit a number of negative properties which we will subsequently consider. In order to isolate cases of simultaneous incidence of several particles by means of an ionization calorimeter, the distances between the particles must not be smaller than some magnitude which defines the "resolving power" of the given calorimeter. It is obvious that the resolving power is a function of detector dimensions and of the lateral dimensions of the secondary-particle shower produced by one of the nuclear-active primary particles.

The reduction of the lateral detector dimensions leads to an increase in the over-all number of detectors in the calorimeter and these dimensions are therefore limited by the difficulties encountered in

fabrication and in operation with a large number of independent ionization detectors.

However, setting the problems of a technical nature aside for a moment, we can consider the problem of the basically attainable spatial resolution of an ionization calorimeter with various absorbents. The basic energy (and, consequently, ionization) is isolated in the calorimeter by electrons of a cascade shower. If all of the secondary, tertiary, and subsequent generations of nuclear-active particles formed in the ionization calorimeter were to move strictly in the direction of the primary nuclear-active particle, in this limit case ensuring minimum lateral dimensions for the entire shower, these dimensions would be finite. However, the electrons of a cascade shower, moving through matter, experience repeated scattering and depart from the initial direction, i.e., the axis of the shower, spreading out in some specific manner about the axis of the shower. The law governing the distribution of density for the electron stream at various distances from the axis of the shower is described by the function f(r, x) whose form, as demonstrated by calculation, is a function of the extent to which the shower has developed.

At the maximum development of the cascade shower (where the ionization effect is at its greatest) the mean square radius of the shower $\sqrt{\overline{R^2}}$ defining its lateral dimensions, is given by the formula [7]:

$$\sqrt{\overline{R^2}} = \frac{0{,}91E_s}{E_c},$$

where $\sqrt{\overline{R^2}}$ is expressed in shower units.

Approximately 3/4 of all particles are contained in a circle of radius $\sqrt{\overline{R^2}}$. It can therefore be maintained that the lateral dimensions of the shower (at the maximum of its development) are equal to $\sim\sqrt{2\overline{R^2}}$.

It is obvious that the maximum possible spatial resolution of the calorimeter is defined by the lateral dimensions of the cascade shower, i.e., by the quantity $2\sqrt{R^2}$. The values of the quantity $2\sqrt{R^2}$ in shower units and centimeters are presented in Table 2.

TABLE 2

1 Вещество поглотителя	Критическая энергия, 2 Мэв	$2\sqrt{R^2}$	
		3 в длинах единицах	4 в см
Углерод 5	72	0,53	13
Железо 6	18	2,1	3,4
Свинец 7	6,4	6,0	3,0

1) Absorbent substance; 2) critical energy, in Mev; 3) in shower units; 4) in cm; 5) carbon; 6) iron; 7) lead.

As can be seen from this table, the maximum attainable spatial resolution in calorimeters with light substances is smaller by a factor of 4 than in calorimeters with heavy substances.

In the construction of an ionization calorimeter some method must be employed to resolve the problem of measuring the ionization arising in the ionization detectors of the calorimeter. Thus given the same primary-particle, the magnitude of the ionization produced by that particle and its "descendants" in the absorbent is a function of the atomic number of the absorbent substance, and in the selection of the absorbent substance this function must be taken into consideration. The ionization produced in the detector by the electrons of the shower is proportional to the number of electrons in the shower. With a given primary-particle energy E_0, the number of electrons in the cascade shower is determined primarily by the critical energy E_c for the given substance, i.e., $I \sim 1/E_c$. This assumes that in substances with small and high atomic numbers the process of transferring energy to the electron-photon component of the shower proceeds with identical intensity. If, however, we take into consideration that the coefficients of inelasticity for the interaction of primary nucleons in substances with great Z are greater than in substances with low Z, the relationship between ionization and the atomic number of the absorbent substance will be even more pronounced than follows from the relation-

ship I ~ $1/E_c$. (Here we have completely set aside consideration of absorbents made of a mixture of substances: light substances for the generation of π^0-mesons and heavy substances for the subsequent development of electron-photon cascades.)

As we can see, and from the standpoint of the magnitude of the produced ionization, preference should be given to absorbent substances with large atomic numbers.

In summarizing, we can state that it is most expedient to use substances with medium atomic numbers exhibiting great density (iron, brass, copper) as the absorbent in an ionization calorimeter. Such substanceshave nuclear mean free paths of the order of several shower units (as a result, a small number of ionization-detector rows is obtained). The great density ensures, with a given over-all absorbent thickness ~1000 $g \cdot cm^{-2}$, comparatively small calorimeter dimensions and, consequently, for a given "speed," a small total number of ionization detectors. The comparatively low value of the critical energy E_c ensures a considerably greater magnitude of ionization for the given primary-particle energy. Moreover, the maximum attainable resolution in substances such as iron and copper is considerably greater than in light substances.

5. <u>The Role of Nuclear Fission in Energy Losses and the Accuracy of Measuring the Energy of a Discrete Particle</u>

In the passage of a shower of nuclear-active particles through the dense substance of an ionization calorimeter, the absorption of particles is governed exclusively by the nuclear interactions and the ionization losses of energy. (The decay mean free paths for all the known unstable particles the decay of which produces muons and neutrinos are considerably greater than the mean free paths for the nuclear interaction.)

In the event of a collision between an atomic nucleus and a particle of sufficiently high energy there occurs not only the generation of new particles, but the splitting of the target nucleus, this being accompanied by the escape from this nucleus of heavy particles, i.e., neutrons, protons, deuterons, and other heavier particles. The energy of these products of nuclear fission is drawn from the energy of the "primary" particle responsible for the fission.

In the overwhelming majority of cases, the energy of the heavy particles is not great, i.e., of the order of several tens of Mev. Therefore, the charged products of nuclear fission exhibit great specific ionizing capacity and, consequently, short mean free paths in material.

Neutrons with energies in tens of Mev and higher expend their energy on the splitting of atomic nuclei, i.e., on the formation of strongly ionizing particles. Neutrons with energies of ~Mev and lower will cause nuclear reactions as a result of which γ-quanta, electrons, and neutrinos will be emitted. In these processes a fraction of the energy will escape measurement, i.e., the energy introduced by the neutrino and the energy liberated as a result of the radioactive decay which, in the majority of cases, will take place upon conclusion of the measurement of the energy liberated by the primary particle in the ionization calorimeter.

It has been estimated that the energy provided by all of the neutrons exhibiting energies of ~10 Mev amounts to 3-5% of the energy of the primary particle.

Therefore, having considered the role of nuclear fission in the over-all process of energy loss, we will subsequently neglect the energy of the low-energy neutrons.

In nuclear fission a small fraction of energy is expended on rip-

ping the particles out of the nucleus (the binding energy). The charged particles (protons, α-particles, etc.), having been decelerated in the substance of the ionization calorimeter, expend all of their kinetic energy on ionization. After coming to a stop they no longer produce any nuclear reactions and, consequently, they do not return the binding energy expended on their removal from the split nucleus to the over-all energy balance. Neutrons however, which conclude their "life" in some nucleus, return the binding energy to the nucleus by which they are absorbed and a greater fraction of this energy is somehow liberated by the nucleus. The over-all quantity of binding energy expended in all nuclear fissions in the formation of charged heavy particles is of the same order of magnitude as the energy transferred to the low-energy neutrons, i.e., it amounts to several percent of E_0 and we will also neglect this.

Consequently, we can write $E_0 = \Sigma E_{\pi^0} + \Sigma E_{n.p.}$ where ΣE_{π^0} is the energy transferred to the π^0-mesons in all of the interactions, while $\Sigma E_{ya.r.}$ is the energy expended on nuclear fission in all of the interactions.

Let us consider how the quantity $\Sigma E_{ya.r.}$ is determined, what fraction of E_0 it represents in terms of order of magnitude, and how this fraction depends on E_0.

With each interaction between a nuclear-active particle and an atomic nucleus, fission of the atomic nucleus takes place with an average loss of energy $\bar{\varepsilon}_{ya.r.}$ expended on the fission; it has been demonstrated experimentally that the quantity $\bar{\varepsilon}_{ya.r.}$ is virtually independent of the energy of the impinging particle, if it exceeds ~1 Bev. During the development of the nuclear cascade in the substance of the ionization calorimeter, the energy of the secondary nuclear-active particles diminishes to some E_{krit} at which the basic energy loss in the interaction is a result not of the generation of new particles, but of

nuclear fissions. On the basis of experimental data $E_{krit} \approx 10^9$ ev.

In order to evaluate the fraction η of the primary-particle energy expended on nuclear fission, and in order to clarify the nature of the relationship between η and E_0, let us consider the following simplest scheme of energy losses:

a) in each interaction event involving any nuclear-particle exhibiting an energy $E > E_{krit}$, $\underline{m}$ charged particles of identical energy are generated, receiving a total energy $(1 - k_{\pi^0}) E_0$, where k_{π^0} is the fraction of the energy transferred to the π^0-mesons in the interaction;

b) with each interaction with a nucleus an energy of $\bar{\varepsilon}_{ya.r.}$ is expended on the nuclear fission;

c) with a secondary-particle energy of $E \leq E_{krit}$ there is no generation of secondary particles.

In this simplified examination the dissipation of energy in each interaction event, so long as secondary particles are being generated, takes place primarily as a result of the transfer of energy to the π^0-mesons.

This can be seen from the fact that the average energy of the secondary particles of the $\underline{i}$th generation $E > E_{krit}$ (otherwise there would be no generation of secondary particles).

The energy transferred to the generation particles of the $(i + 1)\underline{th}$ generation is equal to $(1 - k_{\pi^0}) E_i - \bar{\varepsilon}_{a.p.}$. The average energy of these particles

$$E_{i+1} = \frac{(1 - k_{\pi^0}) E_i - \bar{\varepsilon}_{a.p.}}{m} > E_{krit} > \bar{\varepsilon}_{a.p.}$$

i.e.,

$$\frac{(1 - k_{\pi^0}) E_i}{m + 1} > \bar{\varepsilon}_{a.p.}$$

Since $m \gg 1$,

$$\frac{\bar{\varepsilon}_{n.p.}}{(1-k_{\pi^0})E_i} \ll 1.$$

It is obvious that this inequality is all the more pronounced, the smaller the number of generations i, since $\bar{\varepsilon}_{ya.r.} = \text{const}$, while E_i increases with diminishing i. Therefore, we can maintain that

$$E_i = \frac{(1-k_{\pi^0})E_{i-1}}{m} = \frac{(1-k_{\pi^0})^i}{m^i}E_0.$$

Within the framework of the adopted simplifications, the process of cascade "multiplication" of nuclear-active particles will proceed until some kth generation whose particle energy will become equal to E_{krit}, i.e.,

$$E_{krit} = \frac{(1-k_{\pi^0})^k}{m^k}E_0.$$

Hence

$$k = \frac{\lg E_0/E_{krit}}{\lg m/(1-k_{\pi^0})}. \qquad (16)$$

All particles of the kth generation, and their number is equal to m^k, expend all of their energy, equal to E_{krit}, on nuclear fission. In other words, the energy $m^k \cdot E_{krit}$ will be spent on nuclear fission.

With each interaction involving nuclear-active particles (and of such interactions there were $\sum_{i=0}^{k-1} m^i$), an energy of $\bar{\varepsilon}_{ya.r.}$ is expended on nuclear fission. Therefore, the total loss of energy due to nuclear fission will be equal to:

$$\sum E_{n.p.} = m^k E_{krit} + \bar{\varepsilon}_{n.p.} \sum_{i=0}^{k-1} m^i.$$

Since $m \gg 1$,

$$\sum_{i=0}^{k-1} m^i \approx m^{k-1}, \quad \sum E_{n.p.} = E_{krit} \cdot m^k + \varepsilon_{n.p.} \cdot m^{k-1} =$$

$$= m^k E_{krit}\left(1 + \frac{\varepsilon_{n.p.}}{mE_{krit}}\right).$$

However, $\varepsilon_{n.p.} \leqslant E_{krit.}$ and for that reason

$$\frac{\bar{\varepsilon}_{n.p.}}{mE_{krit}} \ll 1, \sum E_{n.p.} \approx m^k E_{krit} = (1 - k_{\pi^0})^k E_0, \qquad (17)$$

where

$$k = \frac{\lg E_0 / E_{krit}}{\lg m / (1 - k_{\pi^0})}.$$

If in order to evaluate the quantity $\eta(E_0)$ for various primary-particle energies E_0 we assume $E_{krit} = 10^9$ ev, m = 10, and $k_{\pi^0} = (1/3)$, we obtain the result presented below

E_0, ev	10^{11},	10^{12},	10^{13}
$\eta(E_0)$	0,50,	0,36,	0,26

These evaluations show that the fraction η of the energy expended on nuclear fission in the ionization calorimeter represents a considerable part of the entire primary-particle energy and that η is an extremely weak function of the primary-particle energy E_0. Thus, with a change in E_0 by a factor of 100 (from 10^{11} to 10^{13} ev) η diminishes only by a factor of 2.

This nature of the relationship between η and E_0 is not associated with the fact that crude simplifications were employed in the calculation.

The calculations were carried out in the assumption that:

1. In the interaction of nuclear-active particles with an energy E secondary particles with an energy $E' = E/n$ are generated, with $n = aE^{1/4}$ (if E is expressed in 10^9 ev, a = 2.1).

2. All secondary particles exhibit an average coefficient of inelasticity K = 1, with a 1/3 fraction of the energy transferred to the π^0-mesons.

3. In each interaction an energy of $\varepsilon_{ya.r.}$ is lost on nuclear fission if $E > E_{krit}$. With $E < E_{krit}$ the entire energy of the particle is

expended on nuclear fission.

4. The coefficient of inelasticity of the primary particle (the nucleon) is k_n.

TABLE 3*

E_0, ev 1	$k_n=1$		$k_n=0.75$		$k_n=0.5$	
	$\bar{\varepsilon}_{n.p.}=0.6$	$\bar{\varepsilon}_{n.p.}=0.7$	$\bar{\varepsilon}_{n.p.}=0.6$	$\bar{\varepsilon}_{n.p.}=0.7$	$\bar{\varepsilon}_{n.p.}=0.6$	$\bar{\varepsilon}_{n.p.}=0.7$
10^{10}	0,59	0,62	0,67	0,69	0,74	0,76
10^{11}	0,35	0,38	0,43	0,45	0,50	0,52
10^{12}	0,25	0,27	0,29	0,31	0,33	0,35
10^{13}	0,19	0,21	0,21	0,24	0,23	0,26

* The values of $\bar{\varepsilon}_{ya.r.}$ are expressed in Bev.

1) E_0, in ev.

The results obtained from the calculations of η for various values of E_0 and k_n, carried out under the indicated assumptions, are presented in Table 3.

As we can see from Table 3 the weak relationship between η and E_0 is retained, even if we take into consideration the change in the multiplicity with a reduction in energy and secondary particles.

Electromagnetic cascades exhibit rather great mean free paths. Therefore, in the case of absorbent layers separating adjacent rows of ionization detectors into several shower units, $\Sigma E_{\pi 0}$ will be measured with sufficient accuracy, i.e., the measured value of $\Sigma E_{\pi 0}$ will differ little from the true value. The situation is quite different in the case of the measurement of $\Sigma E_{ya.r.}$ The products of nuclear fission exhibit short mean free paths, on the average ~1 $g \cdot cm^{-2}$. Consequently, their energy is liberated in the form of ionization in an absorbent layer ~1 $g \cdot cm^{-2}$ thick. Therefore, a greater part of the strongly ionizing particles are absorbed in the filters of the ionization calorimeter, without attaining the ionization detectors, thus leading to consider-

able fluctuations in the measured value of the quantity $\Sigma E_{ya.r.}$.

Nevertheless, on the average, the energy $\Sigma E_{ya.r.}$ will be measured correctly by the ionization calorimeter.

In order to evaluate the nature of the fluctuations in the measured quantity $\Sigma E_{ya.r.}$ and the influence of these fluctuations on the accuracy with which the energy E_0 of the primary particle is measured, let us consider a somewhat simplified diagram of energy liberation by strongly ionizing particles in the ionization calorimeter.

We will assume that as a result of a single nuclear-fission event occurring in the ionization calorimeter the strongly ionizing particles receive a total energy E_1. If the energy $\Sigma E_{n.p.} = \eta(E_0) \cdot E_0$ is expended throughout the entire ionization calorimeter on nuclear fission, this energy is liberated as a result of $\overline{N}_{ya.r.}$ standard events in each of which E_1 of energy was liberated. Consequently,

$$N_{n.p.} = \sum E_{n.p.} / E_1.$$

The determination of the primary-particle energy in the ionization calorimeter is carried out by measuring the ionization I_k in each row of ionization detectors and by subsequently summing the quantities I_k, i.e.,

$$E_0 = \lambda \Sigma I_k x_k.$$

For the sake of simplicity let us assume that all absorbent layers x_k have the identical thickness $\underline{x}$. In this case $E_0 = \lambda x \Sigma I_k$.

The ionization I_k in some $\underline{k}$th row of ionization detectors is produced by both the electrons of the electromagnetic cascades – I_k^e – and by strongly ionizing particles – $I_k^{s.i.}$ – i.e.,

$$I_k = I_k^e + I_k^{s.i.}.$$

Consequently

$$E_0 = \lambda x \left(\sum I_k^e + \sum I_k^{s.i.} \right).$$

However,

$$\lambda I_k^{e} = \Delta E_k^{e} / \Delta x, \qquad \lambda I_k^{c.n.} = \Delta E_k^{c.n.} / \Delta x,$$

where ΔE_k^{e} and $\Delta E_k^{s.i.}$ are, respectively, energies spent on ionization in the detectors of the kth row by electrons and strongly ionizing particles. Δx is thickness expressed in $g \cdot cm^{-2}$ of the working substance in the ionization detector.

Thus,

$$E_0 = (x/\Delta x)\left(\sum \Delta E_k^{e} + \sum \Delta E_k^{c.n.} \right) . \qquad (18)$$

This relationship is exact for average values of the energy E_0, i.e., for the average result of the repeated measurement of the energy of discrete particles exhibiting identical energies E_0. In other words,

$$\overline{E_0} = E_0 = (x/\Delta x)\left(\overline{\sum \Delta E_k^{e}} + \overline{\sum \Delta E_k^{c.n.}} \right).$$

Let us take note of the fact that

$$\overline{(x/\Delta x) \sum \Delta E_k^{e}} = \sum E_{n.p.}, \quad \overline{(x/\Delta x) \sum \Delta E_k^{c.n.}} = \sum E_{n.p.}$$

Let ε denote the energy liberated, on the average, for ionization in the ionization detector by strongly ionizing particles appearing as a result of a single nuclear fission. In this case the average value of the energy liberated in all ionization detectors by strongly ionizing particles may be written as follows:

$$\overline{\sum \Delta E_k^{c.n.}} = w \overline{N_{n.p}}\, \varepsilon = \frac{\Delta x}{x} \sum E_{n.p.},$$

where w is the probability of recording a single nuclear fission.

The energy liberated for ionization in the ionization detectors by the electrons of the cascade showers

$$\overline{\sum \Delta E_k^{e}} = (\Delta x / x) \sum E_{n^e}.$$

The quantity $w\bar{N}_{ya.r.} = \bar{n}$ is the average number of nuclear fissions yielding the basic contribution to the ionization due to the products of the nuclear fission. In our simplified calculation $\bar{n}$ is the average number of recorded nuclear fissions.

For each individual case of incidence of a particle having an energy E_0 onto the ionization calorimeter some number of nuclear fissions may be recorded. If $\underline{n}$ nuclear fissions are recorded, an ionization energy of $E(n)^{s.i.} = n\varepsilon$ is liberated in the ionization detectors. In accordance with this energy, on the basis of Expression (18), the primary particle energy will be determined

$$E_0' = (x/\Delta x)\left[\sum \Delta E_k^e + E^{s.i.}(n)\right] = (x/\Delta x)\left(\sum \Delta E_k^e + n\varepsilon\right). \quad (19)$$

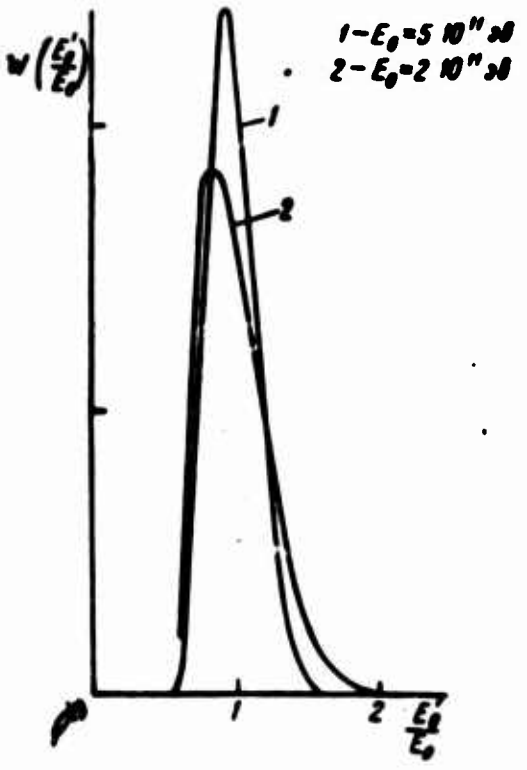


Fig. 7. Probability distribution for the measurement of energy E_0', for a primary-particle energy E_0, in the case in which the fluctuations in E_0' are a result exclusively of nuclear fission. $w(E_0'/E_0)$ is standardized so that $\int_0^\infty w\left(\frac{E_0'}{E_0}\right) d\left(\frac{E_0'}{E_0}\right) = 1$

Each of the $\underline{n}$ nuclear fissions is recorded independently of the other and for this reason the likelihood of recording $\underline{n}$ nuclear fis-

sions with an average number $\bar{n}$ of recorded fissions will be subject to the Poisson law, i.e.,

$$P(n) = [(\bar{n})^n / n!] e^{-\bar{n}}.$$

The probability distribution for the measurement of the given value of the ratio E_0'/E_0 for the above-considered values of the primary-particle energies is shown in Fig. 7. All curves in this figure have been standardized so that

$$\int_0^1 w\left(\frac{E'}{E_0}\right)\frac{dE'}{E_0} = 1.$$

We can see from Fig. 7 that with an increase in the energy of the primary particles there is a reduction in the fluctuations in the measured energy, these fluctuations having been brought about by the nuclear fission.

The distributions shown in Fig. 7 fail to take into consideration the fluctuations in the fraction of energy transmitted in each interaction event to the π^0-mesons, the fluctuations in the multiplicity of generated second particles, the energy loss as a result of nuclear fission, etc.

6. Selection of Ionization Detectors

The primary function of ionization detectors is to measure the ionization produced in an absorbent by shower particles. Ionization detectors should therefore not introduce distortions into the angular and energy distribution of shower electrons. This condition can be satisfied best of all if the ionization detectors are made of substances for which the Z are close to the atomic number of the absorbent substance. For example, if the absorbent is made of iron, the detectors may be made of ionization chambers (proportional counters) made of steel or brass (copper) and filled with argon or a mixture of argon and krypton. If the detectors are made of dense substances with the Z significantly

other than the atomic number of the absorbent, in order to satisfy the above-indicated condition the thicknesses of such detectors must be kept small. In selecting the detector thickness it should be borne in mind that the majority of the particles in an electron-photon shower exhibit an energy $E < E_c$. To prevent the ionization detector from introducing significant changes into the energy and, consequently, into the angular particle distribution, the energy losses due to ionization by individual shower particles in the detector must be significantly lower than the energy of the particles themselves. If the detector thickness, expressed in shower units of the substance of which the detector is made, is denoted x_d, if we denote the critical energy of the electrons for the absorbent substance by $(E_c)_p$, and if the critical energy for the detector substance is denoted by $(E_c)_d$, the condition of low energy losses in the detector is written as follows:

$$x_d(E_c)_d \ll (E_c)_p.$$

Hence it follows that if the ionization detectors are devices made of substances exhibiting low atomic numbers (for example, plastic scintillators, Cerenkov counters), their thicknesses must be of the order of 1 $g \cdot cm^{-2}$ and lower if the absorbent is of a substance having a large atomic number.

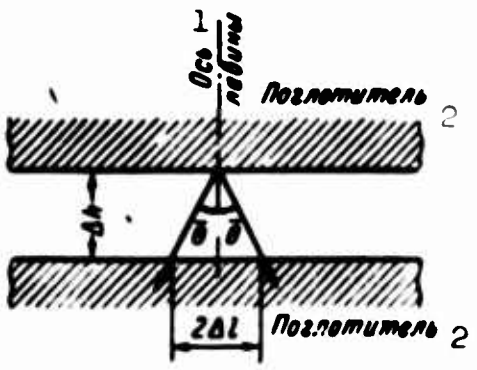


Fig. 8. Illustration of ionization-detector resolving power as a function of detector height h. $\bar{\theta}$) The average electron-scattering angle in the shower; 1) Shower axis; 2) absorbent.

In the selection of the ionization detectors it should be borne in mind that the practical spatial resolution of the detector depends not only on its width, but on its height. Let us turn to Fig. 8. The multiple scattering of the cascade electrons causes the electrons to assume various directions of motion. The higher the Z of the substance in which

the shower is developing, the greater the value of θ, i.e., the average angle between the axis of the shower (the direction of motion for the primary γ-quantum) and the direction of motion for the entire aggregation of electrons in the given shower cross section. Just as the width of the spatial distribution of electrons, the angle $\bar{\theta}$ is a function of the extent to which the cascade has developed, i.e., it is a function of its "age." The cascade electrons impinging on the ionization detector escape from the absorbent at an average angle $\bar{\theta}$ (with respect to the direction of primary-particle motion). If the height of the detector is given Δh, the entire electron-shower will spread out to a width of

$\Delta l \approx 2\bar{\theta}\Delta h$ on passage through the detector.

At the point of maximum development for a cascade shower, in a light substance $\bar{\theta} \sim 0.25$ and $\Delta l = 0.5\ \Delta h$. In a heavy substance (lead) electrons are scattered virtually isotropically at the point of maximum shower development, and $\bar{\theta} \approx 1$, i.e., $\Delta l \approx 2\Delta h$.

We can see from this examination that in order to attain the maximum possible spatial "resolution" in the ionization calorimeter using an absorbent made of a substance having a large Z (for example, with an absorbent made of lead) the height of the ionization detectors should not exceed their width. However, the width, in turn, should not exceed $2\sqrt{N}$.

7. Parameters of the Ionization Calorimeter for Work in the Upper Part of the Atmosphere and Beyond its Limits

The operating conditions for an ionization calorimeter in the upper part of the atmosphere, and particularly beyond the limits of the atmosphere, differ significantly from the operation coniditions at sea level and on mountain peaks.

At a great altitude where the residual air pressure amounts to $\sim 10\ g \cdot cm^{-2}$, the overwhelming portion of the primary protons and a sig-

nificant part of the primary α-particles will reach the ionization calorimeter without having had any interactions in the atmosphere and for this reason the primary particles will not be accompanied, at high altitudes, by aerial showers and secondary nuclear-active particles.

In that small percentage of cases (~10-15%) in which the high-energy primary proton nevertheless succeeds in interacting in the air above the instrument, because of the limited density of the atmosphere (lower by a factor of 50-100 than on mountain peaks or at sea level) the secondary nuclear-active particles will be separated from one another on the average by a distance 50-100 times that at mountain altitudes or sea level. Therefore, a single nuclear active particle will virtually always impinge on an ionization calorimeter, if there is an insufficient quantiy of substance close to the ionization calorimeter within the limits of its angular aperture on the side on which the primary particles impinged.

These operational features are all the more pronounced if the ionization calorimeter is situated beyond the limits of the atmosphere.

Under these operation conditions there is no need tc distinguish between incidence on the ionization calorimeter of aerial atmospheric showers, groups of nuclear-active particles, and the incidence of a single particle; there is therefore no need for a large number of ionization detectors in each row.

Since

$$E_0 = \lambda \int_0^{x_0} I(x)dx = \lambda \sum I_k x_k,$$

where I_k is the ionization value recorded in the kth row of detectors and x_k is the thickness of the deceleration medium of the ionization calorimeter between the kth and (k - 1)th rows, if we set all of the layers of the medium [substance] equal to the thickness ($x_1 = x_2 = \ldots = x$),

we will obtain

$$E_0 = \lambda s \sum I_k.$$

From this expression we see that in the case of ionization calorimeter operating at a great altitude or beyond the limits of the atmosphere we can do without one of the measuring elements, connecting all of the detectors in parallel, e.g., the ionization chambers. We can take particular advantage of this circumstance in the selection of the ionization detector.

The basic equality which describes the operating principle of an ionization calorimeter may be written somewhat differently, if we note that

$$\lambda I(s) = \frac{dE(s)}{ds} ds,$$

i.e.,

$$E_0 = \int_0^{s_n} \frac{dE(s)}{ds} ds = \sum \frac{dE_k}{ds} s_k,$$

where dE(x)/dx is the loss of energy resulting from the ionization by all of the charged particles at the depth $\underline{x}$ in a layer of 1 $g \cdot cm^{-2}$ of ionization-calorimeter substance. If all $x_k = x$,

$$E_0 = s \sum dE_k(s)/ds.$$

If a scintillator is chosen as a dE_k/dx detector, the quantity of light S in the scintillator is proportional to dE_k/dx, i.e., $S_k \sim dE_k/dx$ and, consequently,

$$E_0 = as \sum S_k = asS_0, \qquad (20)$$

where S_0 is the total quantity of light appearing in all scintillation detectors of the ionization calorimeter.

The recording of the total light flash may be accomplished in principle by a single FEU [photomultiplier].

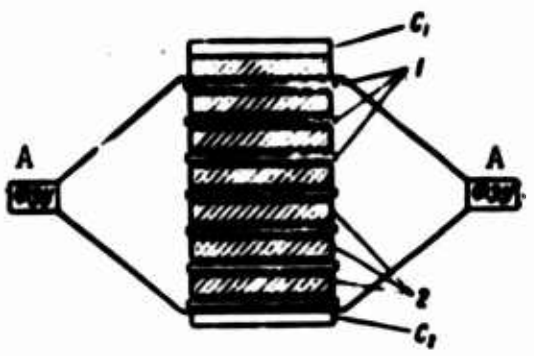


Fig. 9. Schematic representation of the simplest version of an ionization calorimeter with scintillators for operation in the upper part of the atmosphere. C_1 and C_2 are counters of charged particles, controling the operation of the ionization calorimeter. 1) Scintillator; 2) absorbent layers of the ionization calorimeter; A) FEU [Photomultiplier].

The simplest version of an ionization calorimeter for high-altitude measurements will have the form shown in Fig. 9.

At great altitudes the intensity of the cosmic-radiation particles is considerably higher than at sea level or in mountainous areas and in accordance with this fact the total flow of cosmic-radiation energy incident on 1 cm^2 of surface per 1 second is greater by approximately 2 orders of magnitude than on mountain peaks and is equal to $\sim 5 \cdot 10^9$ ev $cm^{-2} \cdot sec^{-1}$.

On the one hand, a large flow of cosmic-radiation particle energy requires a smaller time interval τ within which the energy E_0 of the primary particle is measured in order for the error in E_0 due to random coincidences to fall within the given limits. On the other hand, great intensity on the part of high-energy particles makes possible a reduction in the dimensions of the ionization calorimeter (at least, the lateral dimensions) and thus makes it possible to reduce the likelihood of recording two or more particles simultaneously.

If for the evaluation of τ we again assume $S = 10^4$ cm^2 and that the permissible error $\Delta E_0 / E_0 \leqslant 0{,}1$, we obtain

$$\Delta E_0 = 5 \cdot 10^9 \cdot 10^4 \tau = 5 \cdot 10^{13} \text{ ev}.$$

If $E_0 = 10^{11}$ ev, then $\Delta E_0 / E_0 < 0{,}1$ when $\tau < 2 \cdot 10^{-3}$ sec.

8. Recording of Ionization Pulses from a Large Number of Detectors

In working with an ionization calorimeter the individual conducting the experiment is confronted with the problem of measuring the in-

stantaneous distribution of ionization in a large number of detectors (of the order of 100-1000). It is obvious that in order to solve this problem each detector must be connected to a separate amplitude analyzer. The collection of separate analyzers, together with the device for the "recording" of the measured pulses, forms a measurement and recording system for pulse amplitudes. Such systems can satisfy certain requirements.

1. Amplitude-measurement systems must be sufficiently fast-acting in order to conduct the measurement of ionization within so brief a period of time as to make the appearance of random pulses unlikely. However, the system for the recording of the measured pulses may function rather slowly, since in work with cosmic-radiation particles the frequency of observation for high-energy particles is low. Therefore the time required for the recording of the measured pulses may be of the order of tens of seconds.

2. The distribution of ionization between the various detectors is extremely nonuniform and subject to intense fluctuations due to the following factors. First of all, the recorded phenomena may be brought about by particles of various energies. As a result 10-fold fluctuations in the detector ionization can be anticipated. (If the ionization calorimeter is constructed so that it records particles with an energy of E_{min} with sufficient frequency, particles with energies of 10 E_{min} will be recorded at a frequency lower by a factor of 50-100.) Secondly, the distribution of ionization along the calorimeter may be quite varied. As a result, given this primary-particle energy the magnitude of ionization in the various detectors may vary by factors from 20 to 50. Thus the measurement and recording system must exhibit a great dynamic range and make possible measurements and recordings of pulse amplitudes varying over a 500-1000-fold range.

3. The pulse measurement and recording system must be interference-free. This requirement is associated with the fact that, as a rule, an ionization calorimeter is called upon to function in conjunction with some device for the observation of interaction processes, most frequently with a Wilson chamber, a spark chamber, counters, etc. The operation of the Wilson chamber (the flashing of lights, the operation of valves, exhaust) may produce significant interference of both an electromagnetic and acoustical nature (the microphone effect).

4. The amplitude-analyzer installation and the recording system must be quite simple and function reliably. Otherwise the installation, containing 100-1000 amplitude analyzers, will virtually be incapable of operation.

II. POSSIBLE APPLICATIONS OF THE IONIZATION CALORIMETER

At the initial stage of ionization-calorimeter development the authors thought that this instrument, operating in conjunction with a Wilson chamber, would function only in the role of a detector of primary-particle energy, while the Wilson chamber is to be used to study the elementary processes of interaction.

Some experience has already been accumulated in working with ionization calorimeters on mountain peaks. The projects carried out at the Cosmic-Ray Laboratory of the Scientific-Research Institute of Nuclear Physics at Moscow State University, which provided the impetus for the application of this new method, demonstrated the possibility of employing the ionization calorimeter as an independent physical instrument to study the characteristics of nuclear interactions between high-energy particles [3]. The work carried out by this laboratory in conjunction with the FIAN [The P.N. Lebedev Institute of Physics of the USSR Academy of Sciences] showed the exceptional effectiveness of combining an ionization calorimeter with a Wilson chamber [6].

At the present time all of the major physics institutes of the nation (the Physics Institute of the USSR Academy of Sciences, the Institute of Nuclear Physics of the Kazakhstan SSR Academy of Sciences, the Physics Institute of the Georgian SSR Academy of Sciences, the Scientific Research Institute of Nuclear Physics of Moscow State University), engaged in a study of cosmic radiation, have developed or are completing development of large installations constructed on the basis of combining an ionization calorimeter with a Wilson chamber. In a number of cases, for purposes of analyzing secondary particles, the Wilson chamber is placed inside a magnetic field. Ionization calorimeters are finding some application in foreign laboratories [8]:

However, the ionization calorimeter has considerably greater potentials than those being realized at the present time.

Let us consider those of these potentials which may be used for the solution of the problems formulated at the beginning of this article.

1. The Study of the Chemical Composition of Primary Cosmic Radiation in the Region of High and Very High Particle Energies

As follows from the operational principle of the ionization calorimeter, the latter is an instrument designed to measure the energy of a nuclear-active particle, and it is equally well suited for the measurement of the energy of a proton and the multiply charged particles of primary cosmic radiation. Moreover, in the motion through matter of particles exhibiting very high energies the latter rapidly dissipate their energy (the transfer of energy to the π^0-mesons), so that the ionization calorimeter is equally well suited for the measurement of particle energies of 10^{11} ev and of 10^{15} ev (for very-high particle energies the energy will be measured even better, since there will be smaller fluctuations in the energy expended on nuclear fission).

The ionization calorimeter may therefore be used for the direct measurement of the energy spectrum of primary cosmic-radiation particles.

If we give rein to our imagination for a moment and assume (and apparently the near future will show these assumptions to have been quite modest) that we have an ionization calorimeter with a geometric factor $\Gamma = 10^4$ cm^2·sterad aboard an artificial satellite of the earth, the operation of such an ionization calorimeter for a period of a year will make it possible to measure the absolute intensity and to study the energy spectrum of the primary particles all the way to energies of 10^{15} ev.

If we place an instrument over the ionization calorimeter to permit measurement of the charge of the primary particle (a Cerenkov counter or large proportional counters), it will become possible to study the chemical composition of the primary cosmic particles to very high energies of $\sim 10^{15}$ ev.

It is curious to note that the study of the chemical composition of primary cosmic radiation in the high-energy regions by means of an ionization calorimeter and a detector of charge Z is less subject to the effect of the various methodological errors than the study of the chemical composition with the methods presently in use in the region of low energies $<10^{10}$ ev/nucleon.

The situation here involves the fact that in measuring a flow of primary particles of unlike charge at a given geomagnetic latitude, particles exhibiting identical magnetic rigidity are recorded, rather than particles exhibiting identical energies. However, the ionization calorimeter can be used to determine the chemical composition of the primary particles exhibiting equal energies.

Let us assume that the energy spectra of the primary particles

with the given charge Z and the mass number A follow the power law

$$N_Z(E)dE = a_Z(cdE/E^\gamma),$$

where E is the energy for a single nucleon.

If the spectrum of the primary protons is written in the form

$$N_{Z=1}(E)dE = (cdE/E^\gamma),$$

the coefficient a_Z yields the fraction of the nuclei having the charge Z from the stream of protons with identical energies per nucleon. For example, on the basis of geomagnetic measurements we have

$a_2 = 0.07\text{-}0.08$ for $Z = 2$, $\quad a_{3\text{-}5} = 1.5\cdot10^{-3}$ for $Z = 3\text{-}5$,

$a_{6\text{-}8} = 5\cdot10^{-3}$ for $Z = 6\text{-}8$, $\quad a_{\geq10} = 2.5\cdot10^{-3}$ for $Z \geq 10$.

If we measure the flow of particles with the given energy $E_0 = AE$,

$$N_Z(E_0)dE_0 = a_Z A^{\gamma-1}\frac{cdE_0}{E_0^\gamma} = a_Z A^{\gamma-1}N_{Z=1}(E_0)dE_0. \qquad (21)$$

i.e.,

$$\frac{N_Z(E_0)}{N_{Z=1}(E_0)} = a_Z A^{\gamma-1} \approx a_Z(2Z)^{\gamma-1} = B_Z.$$

With $\gamma = 2.7$ we will obtain

$Z = 2$	$B_2 = (0{,}07 + 0{,}08)\cdot 10 = 0{,}7 \div 0{,}8$
$Z = 3 \div 5$	$B_{3\div5} = 1{,}5\cdot10^{-3}$
$Z = 6 \div 8$	$B_{6\div8} = 90\cdot5\cdot10^{-3} = 0{,}45$
$Z \geq 10$	$B_{\geq10} = 160\cdot2{,}5\cdot10^{-3} = 0{,}4_0$

i.e., in relation to the flow of protons the α-particles make up 70-80%, the group of medium nuclei, 45%, and the group of heavy nuclei $\gtrsim$40%, whereas with measurements at the given geomagnetic latitude (in a low-energy region) the streams of these nuclei with respect to the protons yield

$$N_Z/N_{Z=1} = 2^{\gamma-1}a_Z$$

(in the low-energy region $\gamma \simeq 2.4$).

Z	2	$3 \div 5$	$6 \div 8$	≥ 10
$N_Z/N_{Z=1}$, %	$18 \div 21$	0,4	1,3	0,6

In measuring the charge of the primary particles by means of electron methods (in particular, in the region of low values of $Z \sim 3$-6), the nuclear interactions between protons and α-particles in the charge detector may be the principal factors responsible for errors. In such interactions showers of charged particles, simulating particles with $Z > 1$, are generated.

The situation is considerably improved in the measurement of the chemical composition of high-energy primary cosmic particles recorded by means of an ionization calorimeter, since there is a significant increase in nuclei with charge $Z \geq 2$ relative to the protons.

2. A Study of the Characteristics of Nuclear Interactions Involving Primary High-Energy Cosmic Particles

The study of the relationship between the interaction cross sections of elementary particles and the energies of the latter is of great significance for the contemporary theory of elementary particles.

In particular, at the present stage it is extremely important to trace the course of the cross section with energies all the way to 10^{12}-10^{13} ev for inelastic proton-proton interactions σ_{pp}.

At first glance it might seem that this problem could be solved with a large ionization calorimeter having an area of 10-20 m^2 set up at the top of a mountain, together with a system of filters containing hydrogen (for example, on the basis of the paraffin-carbon difference effect), including a detector of inelastic interactions.

As a matter of fact, we know that at an altitude of 3200 m above sea level a flow of nuclear-active particles with energies of $\geq 10^{12}$ ev is equal to $1.0 \cdot 10^{-8}$ $cm^{-2} \cdot sec^{-1} \cdot ster^{-1}$ [9]. Consequently, approximately 100 such particles will fall each day on an ionization calorimeter of large area, let us say 10 m^2. This great intensity would be enough in order to measure the cross section of an inelastic interaction between

nuclear-active particles exhibiting an energy $\geq 10^{12}$ ev and protons with a statistical accuracy of several percent. However, the actual situation is quite different.

First of all, in the lower part of the atmosphere the overwhelming part of the nuclear-active high-energy particles is accompanied by aerial atmospheric showers. Such particles are not suitable for the measurement of the interaction cross section. As a result the intensity of the flow of particles with $E \geq 10^{12}$ ev, suitable for the measurement of the interaction cross section, diminishes approximately by a factor of 10.

Secondly, among the nuclear-active high-energy particles ($E \geq 10^{12}$ ev) in the lower part of the atmosphere the $\pi^{\pm}$-mesons apparently make up the greatest fraction (up to 40%) [10].

Therefore, in order to study the nucleon-nucleon interactions, we must take only the neutral nuclear-active particles (the neutrons), whose flow is smaller by a factor of 2 than the entire flow of the nucleon component.

Because of these two factors, the actual flow of nucleons with $E \geq 10^{12}$ ev, suitable for the measurement of the cross section, at the tops of mountains amounts approximately to 5% of the flow of all of the nuclear-active particles of the indicated energy. This flow is too small even with an installation having an area of 10 m^2 to measure the cross section of inelastic nucleon interactions with nucleons.

The problem of studying the relationship between σ_{pp} and the energy of the primary protons can, in principle, be successfully resolved by using an ionization calorimeter in measurements conducted beyond the limits of the atmosphere.

The advantages here are the following:

1) the proton intensity is greater by a factor of $\sim 10^3$ than on

mountain peaks;

2) the admixture of $\pi^{\pm}$-mesons and other secondary singly charged particles is absent.

If polyethylene and carbon filters are alternated above the ionization calorimeter, a charge detector can be employed to distinguish cases of the incidence onto the filter of a singly charged primary particle (of a proton of the given energy) and this detector can also be used to record cases of the passage of a proton through the filter without interaction, and it thus becomes possible to use the difference method in order to obtain the sought relationship between σ_{pp} and E_0.

In order to illustrate the potentials opening up in this direction, let us use the example of an ionization calorimeter which, together with a charge detector, has a geometric factor $\Gamma = 10^4$ cm^2·sterad. In this case, during the course of the year-long measurements σ_{pp} may be measured with a statistical accuracy of 2% for protons exhibiting an energy of $E_0 \geq 10^{12}$ ev, while for protons with an energy of $\geq 10^{13}$ ev, the accuracy is 15% (if we assume that in the indicated energy interval the exponent of the primary-particle integral spectrum $\gamma - 1 = 1.7$).

If we take a substance with the given atomic weight A as the filter, the relationship between the cross section of the inelastic interaction σ_{pA} of the primary protons and the atomic weight of the target nucleus and the energy E_0 of the primary proton can be investigated.

In this case, for the above example with $\Gamma = 10^4$ cm^2·sterad, the statistical-expectation errors in the determination of σ_{pA} amount to ~0.4% for $E_0 \geq 10^{12}$ ev and 3% for $E_0 \geq 10^{13}$ ev.

The measurements of the cross sections of a nucleus-nucleus interaction with atomic weights A_1 and A_2 in the high-energy region are of great interest.

From the data presented in the previous section it follows that

with a fixed energy for the recorded particle in the primary cosmic radiation, e flows of heavy nuclei are commensurate with proton flows (if the chemical composition of the cosmic rays in the region of high and very high energies is the same as in the low-energy region). As a result, there arises a true possibility of studying a number of characteristics of the interactions of nuclei with nuclei in the energy region 10^{12}-10^{14} ev with high statistical accuracy.

The development of a hydrodynamic theory of high- and very high-energy particle interaction by Landau [11], Milekhin [12], and a number of other authors made possible the establishment of a relationship between the equation of state for a superheated vacuum $p = C^2\varepsilon$ (p is pressure, C is the speed of sound, and ε is the energy density) and the experimentally observed characteristic of interaction.

On the basis of the hydrodynamic theory, there exists a relationship of the following form between the average multiplicity of the $\bar{n}$ secondary particles generated through collision and the energy E_0 of the primary particle:

$$\bar{n} = aE_0^{\nu}\left(\nu = \frac{1}{2}\frac{1-C^2}{1+C^2}\right).$$

Therefore, if the coefficient ν is to be determined by experiment, it is easy also to determine C^2, i.e., to derive the equation of state of that nuclear material which is produced at the instant of collision between high-energy particles.

Despite the urgency of this problem, the question of the type of relationship between $\bar{n}$ and E_0 has not yet been clarified, although it was formulated more than 10 years ago.

It follows from what was said earlier that the ionization calorimeter can be employed to solve this problem as well.

If in the experiments under consideration we make use of a detec-

tor which permits the recording of the number of secondary charged particles generated through interaction, the nature of the relationship between $\bar{n}$ and E_0 in the energy region at least to 10^{13} ev can be clarified in the experiments to measure the cross section of an inelastic pp-interaction on the basis of the difference effect and here it will also be possible to measure ν with an accuracy of ± 0.04.

3. The Photoemulsion Method in the Study of Elementary Nuclear Processes

The range of problems considered in Sections 1 and 2 can be resolved, in principle, by using electronic sensors in conjunction with radio-engineering facilities, the information subsequently being transmitted to the earth by radio.

However, this range of problems by no means covers the entire range of interest in high and very-high energy particles.

Of particularly significant importance is the clarification of the mechanism of particle interaction in the energy region 10^{13}-10^{15} ev, depending on the energy and the nature of the impinging particle, and on the atomic weight of the target nucleus.

Of particular interest in this case is the clarification of the degree of difference between the elementary interaction event in the case of very high energies and the event with which we are familiar for particles exhibiting energies of 10^{11}-10^{12} ev.

At the present time, nuclear photoemulsion is the only method available for a detailed study of the characteristics of the elementary event at energies $\geq 10^{12}$ ev.

At high primary-particle energies the secondary particles in the elementary processes escape at such small angles that over a path involving tens of centimeters they diverge from one another by a fraction of a millimeter. Under these conditions neither the Wilson chamber nor

a spark chamber can "resolve" particles moving so close to one another and these devices record them as a single particle.

The resolving power of photoemulsions is hundreds of times greater easily permitting the separate detection of two parallel tracks produced by two particles separated by a distance of 1 μ.

Moreover, in recent times there have come into extensive use electron-sensitive nuclear photoemulsions interlaid with thin layers of lead in which develops the electron cascade which is fixed by the emulsions in order to measure the energy of the generated π^0-mesons.

However, the emulsions exhibit the property of fixing all particles passing through them during the exposure. Therefore, in order to employ an ionization calorimeter to measure the energy of that primary particle which brought about the interaction observed in the nuclear emulsion, we must learn to define uniquely the interactions in the emulsion with the particles recorded by he ionization calorimeter.

In the Cosmic-Radiation Laboratory of the Institute of Nuclear Physics at Moscow State University various versions for a combination of nuclear photoemulsions with an ionization calorimeter have been worked out. This method, known as the method of controlled photoemulsions, was proposed in 1956 by N.L. Grigorov and worked out in cooperation with V.Ya. Shestoperov, V.A. Sobinyakov, A.V. Podgurska [13], coworkers at the Cosmic Radiation Laboratory. This method is intended for the study of the energy spectrum of the generated π^0-mesons in interactions involving particles of known energy and atomic nuclei.

The essence of this method can be gathered from Fig. 10. Plates coated with the nuclear photoemulsion are positioned beneath thin lead filters over two rows of cylindrical ionization chambers whose axes are mutually perpendicular to one another. An ionization calorimeter is placed beneath these chambers.

If the primary particle interacts in the generator made of a substance with the given atomic weight, the generated π^0-mesons, on decaying, yield high-energy γ-quanta. The γ-quanta, impinging on the lead, produce powerful electron-photon showers in the lead and these penetrate the nuclear photoemulsions and enter the two upper rows of ionization chambers. Extensive ionization is recorded in the chambers through which the shower passed. Therefore, the intersection of the two rows of chambers in which the greatest ionization was recorded can serve as a rough determination of the position of the nuclear plate through which passed the given shower recorded by the chambers of the two upper rows.

The nuclear-active particles ($\pi^{\pm}$-mesons, nucleons), left over from the first interaction, impinge on the ionization calorimeter and liberate their energy $\Sigma E_{ya.a.}$ there.

The upper rows of chambers beneath the lead filters measure the total energy transferred to all of the π^0-mesons in the first interaction, i.e., ΣE_{π^0}. Consequently, the primary-particle energy $E_0 = \Sigma E_{\pi^0} + \Sigma E_{n.a.}$. On the basis of the readings provided by the ionization chambers of the upper rows, the photosensitive plate through which the shower passed is removed from the installation, chemically processed, and the shower is then found with the aid of a microscope. The processing of this shower yields the energy of the γ-quanta responsible for the shower.

Because the point of shower passage is checked by the ionization chambers, the authors referred to this method as the method of controlled nuclear photoemulsions.

In the form described above, this method is used for investigations on mountain peaks [14] and has already yielded interesting results, illuminating a number of phenomena relating to cosmic-ray parti le ener-

gies of 10^{12}-10^{13} ev observed in the atmosphere [15].

However, the described technique is not suitable for investigations at high altitudes with photoemulsion stacks, since because of the significantly greater intensity of primary cosmic high-energy particles than is encountered on mountain peaks on the area of a single plate covered by the surface of two intersecting chambers, so many showers occurring at various times will be recorded during the course of a single exposure that it will be impossible to identify each shower with the readings of the ionization calorimeter.

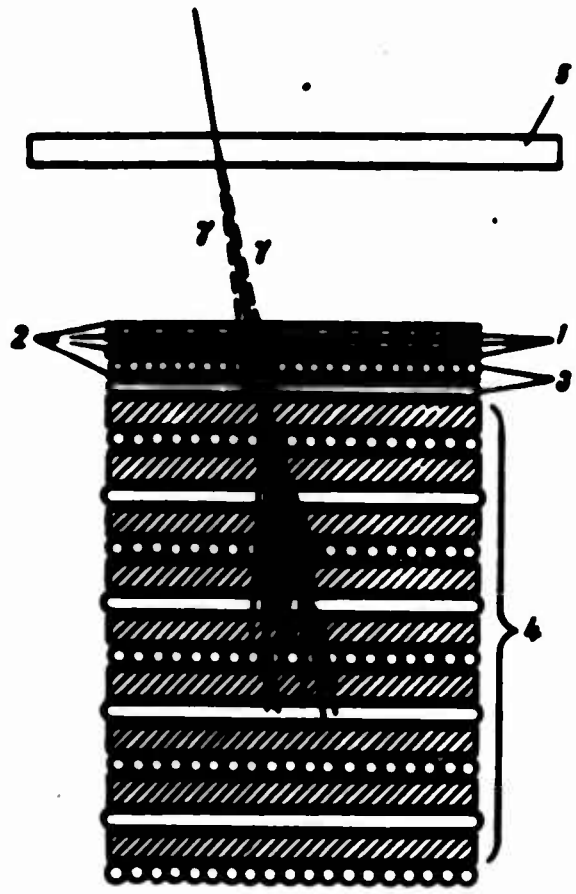


Fig. 10. Schematic representation of installation to study processes of π^0-meson generation by cosmic-ray particles with energies of 10^{12}-10^{13} ev by the method of controlled nuclear photoemulsions. 1) Nuclear photoemulsions; 2) layers of lead; 3) ionization chambers; 4) ionization calorimeter; 5) π^0-meson generator.

In order to make the "resolving" power of the ionization chambers suitable for the conditions of the high intensity with which the events are recorded, it must be considerably improved. Accordingly, chambers with a somewhat undersize diameter should be employed, but this makes

it difficult to use the method.

The development of the method of controlled photoemulsions for conditions of high-altitude investigations proceeded along two lines.

I.D. Rapoport [16] proposes the utilization of ZnW-type phosphor scintillators in contact with high-sensitivity photographic films for the determination of the point of electron-photon shower passage through the photographic plate, one of the scintillators being fixed in position, and the other moving with the recording of the subsequent shower.

Another version of an ionization calorimeter for operations with emulsion stacks of large volume was proposed by N.L. Grigorov and I.D. Rapoport [17]. This version contains an ionization calorimeter made of an assembly of thin scintillation plates positioned between lead plates each 1-2 cm thick. The scintillations are photographed stereoscopically by means of an image converter. In order to achieve better refinement of the coordinates of the electron-nuclear shower, shower indicators made of ZnW phosphor scintillators and photographic films or thick-layered x-ray film of the RT-6 type are positioned between the lead plates of the calorimeter.

What are the prospects offered by the method of controlled photoemulsions in the study of nuclear interactions when applied to the study of high and very-high energy primary cosmic radiation?

Let us assume that there is an installation with a geometric factor $\Gamma = 10^4$ $cm^2 \cdot$ sterad and nuclear photoemulsions controlled by an ionization calorimeter aboard a heavy artificial satellite. Such an installation will record approximately four interactions between particles exhibiting energies of $\sim 10^{14}$ ev per day. If we employ special measures to reduce the background in the nuclear emulsions (setting a small angle of inclination for the satellite's orbital plane relative to the

plane of the equator and selecting a circular orbit ranging between 300-400 km in altitude), the emulsion under such conditions can be kept for a period of 20-30 days beyond the limits of the atmosphere. In other words, approximately 100 interactions involving particles with energies of ~10^{14} ev can be recorded by the photoemulsions during the exposure. For these interactions it will be possible to undertake a detailed study of the basic characteristics of the elementary event in the case of very high energies, i.e., the energy spectrum of the π^0-mesons, the coefficient of inelasticity and the distribution of the interactions on the basis of the inelasticity coefficients, the multiplicity of generated particles, the relationship between these characteristics and the atomic weight of the target nucleus and the mass of the incident particle.

These investigations of the characteristics of very-high energy particle interactions, in addition to providing the solution to fundamental problems of a theoretical nature, also have the significance that they make possible a clarification of the nature of the processes on which the formation of extensive atmospheric showers is based, and thus they provide the scientific basis for the study of particles of even greater energies (~10^{16}-10^{17} ev).

4. <u>The Study of High-Energy Electrons and Photons in Primary Cosmic Rays</u>

We are thoroughly aware of the great significance ascribed at the present time to the study of the electron component of primary cosmic radiation and the study of high-energy γ-quanta. There are three basic problems which must be resolved in the investigations of the electron component.

1. The reliable detection of high-energy electrons (positrons) in the primary radiation and the proof of their galactic origin.

2. The study of the energy spectrum of the electrons (positrons).

3. The determination of the relationship between the electrons and the positrons.

Measurements carried out at an altitude of 30-35 km [18] showed that the intensity of the electron component in primary cosmic radiation amounts approximately to 1-2%.

The difficulty in studying the electron component of primary cosmic radiation is based on the fact that its intensity is quite low, and also on the fact that there are few characteristic indications on the basis of which it would be possible to distinguish the electron from the primary proton. And since the flow of protons in the primary cosmic radiation is greater than the measured flow of electrons by a factor of ~100, even the small probability (~1%) of simulating an electron with a photon leads to tremendous errors in the determination of the electron flow.

The basic difference between the electron and the proton - the great difference in their resting masses - leads to a situation in which the passage of the electron through a substance exhibiting a a large atomic number Z (for example lead) is accompanied by an intense loss of energy due to bremsstrahlung and the subsequent formation of an electron-photon cascade shower of charged particles. The entire charged-particle shower, having expended its energy on the ionization of the substance, is absorbed with comparative rapidity. As a result the mean free path of the high-energy electron (together with all of the secondary particles) in the lead proves to be relatively small.

The high-energy protons, having passed through the substance, experience virtually no loss of energy due to bremsstrahlung and produce no cascade showers. However, as a result of nuclear interactions they generate π^0-mesons whose decay yields γ-quanta which produce cascade showers.

Therefore, if the electrons are detected only on the basis of an indication that a particle with a single charge produces cascade showers in a substance with a large atomic number Z, in a large percentage of cases the proton of the primary cosmic rays will simulate electrons.

It seems to us that in order to achieve a more reliable isolation of electrons, the particles should be distinguished on the basis of mass and mean free path.

As a matter of fact, with the given total proton and electron energy, because of the great differences in their masses, their velocities will vary. For example, with $E_p = E_e = 10^{10}$ ev the quantity $\beta_e = v_e/c$ (v_e is the velocity of the electron, $\underline{c}$ is the speed of light in a vacuum) will be other than 1 by less than $4.5 \cdot 10^{-4}$. The same quantity $\beta_p = v_p/c$ for the proton will differ from 1 by $8 \cdot 10^{-3}$. This difference in electron and proton velocities can be employed to distinguish between them, if a Cerenkov gas-discharge counter is employed. This counter represents a vessel filled with a nonluminescent gas which is kept at the appropriate pressure. An electron, on passing through this gas, because of its greater velocity, will produce Cerenkov radiation which can be recorded by a photomultiplier (FEU), while the proton, because of its lower velocity, will not produce this type of radiation. For example, in freon at normal temperature under a pressure of 10 atm protons with an energy $\gtrsim 10^{10}$ ev and electrons with an energy $\gtrsim 5 \cdot 10^{6}$ ev will produce Cerenkov radiation.

Consequently, if an ionization calorimeter situated beneath the Cerenkov gas-discharge counter is to be used to isolate particles with energies less than 10^{10} ev, the particles with a single charge, producing the Cerenkov radiation, cannot be protons.

In order to determine the passage of the electrons recorded in this manner, their energies must be compared (this energy having been

measured by the ionization calorimeter) with the value of the critical energy $E_{\kappa}(\psi, \theta, \varphi)$, which corresponds to a geomagnetic latitude ψ at which the electron was recorded (θ and φ are, respectively, the polar angle and azimuth determining the direction of electron influx). If $E_e \geq E_k$, the electron is not of terrestrial origin but came to the earth from without. If $E_e < E_k$, the electron is of terrestrial origin and originated either in the atmosphere or in the recording instrument under the action of the primary cosmic rays.

To study the energy spectrum of electrons we naturally require an ionization calorimeter (if we are to work within the energy region 10^9-10^{10} ev).

Finally, the relationship between the electrons and the positrons in the primary cosmic rays will be measured on the basis of the azimuthal effect of cosmic rays.

The investigation of the electron component of cosmic rays by means of heavy artificial satellites of the earth also offers another two advantages which make possible a significant reduction in the likelihood of protons simulating electrons.

The utilization of a rather thick layer of ionization-calorimeter substance together with an auxiliary filter and the appropriate apparatus to record the charged particles makes possible the absorption of the entire shower of particles generated by the primary electron. At the same time, a proton of the same energy (with $E_p > 10^9$ ev) will not, as a rule, be absorbed by such a filter. This circumstance provides additional possibilities of distinguishing between electrons and protons.

Finally, if we are to study the electron spectrum in the energy region $E_e \lesssim$ 10-15 Bev, i.e., electrons subjected to the influence of the terrestrial magnetic field, with the motion of a satellite from the equator to the pole electrons of ever-decreasing energy will be toler-

ated in the recording instrument, as a result of which their intensity will increase with increasing geomagnetic latitude. At the same time, the flow of protons producing the Cerenkov radiation will remain constant, since the energy threshold of proton recording is set by the features of the instrument and must be rather high ($E_{porog} \geq 10^{10}$ ev). As a result, the simulation of electrons by high-energy protons $E_p \geq \geq 10^{10}$ev will not be a function of the position of the instrument and will remain a constant quantity, producing some background which may be measured (or at least the upper limit of the background can be evaluated on the basis of measurement results obtained in the vicinity of the equator).

One of the possible versions of an instrument to record electrons in the primary cosmic radiation in which the above-enumerated principles are employed is shown in Fig. 11.

The study of the high-energy γ-quanta in the primary cosmic radiation at this stage can, by agreement, be subdivided into two basic problems:

1) The over-all γ-radiation of the Galaxy and, possibly, the Metagalaxy.

2) The local sources of γ-quanta in our Galaxy and beyond its limits.

The solution of the first problem requires the measurement of the stream of γ-quanta from various regions of the celestial sphere, the measurement of the energy spectrum of the γ-quanta, and the relationship between that spectrum and the direction of γ-quanta arrival.

The second problem calls for the solution of the question as to whether or not there exist local sources and, if such are detected, a study of the spectral characteristics of the source-emitted γ-quanta.

The evaluation of the over-all γ-radiation of the Galaxy leads to

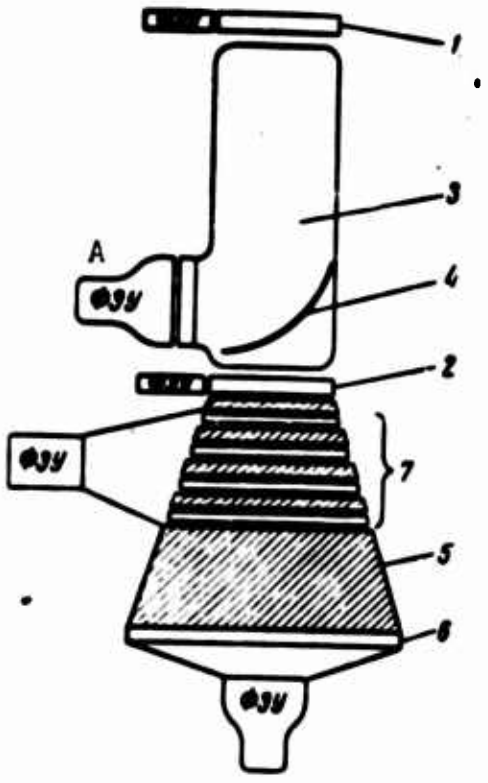


Fig. 11. Schematic representation of the instrument to record high-energy electrons in primary cosmic rays. 1) Charge detector and "telescope" counter; 2) "telescope" counter; 3) Cerenkov gas-discharge counter; 4) spherical mirror; 5) thick lead absorbent of charged particles; 6) charged-particle counter; 7) ionization calorimeter; A) FEU [Photomultiplier].

a situation in which it is possible to anticipate streams of γ-quanta amounting to 10^{-4}-10^{-5} of the stream of charged particles (given equal energies for the γ-quanta and the charged particles). With respect to the local sources, it is quite difficult to arrive at any, even partly valid, estimates of the anticipated intensity. It is quite likely that these streams [flows] are considerably smaller than the flows produced by nuclear interactions of cosmic rays and the interstellar medium in our Galaxy.

The instruments intended for the solution of the above-indicated range of problems must exhibit specific features making possible the detection of γ-quanta against the background of a tremendous stream of charged particles and they should also provide for the possibility of determining the direction of γ-quanta influx. It is obvious that in order to seek out local sources of γ-quanta the accuracy with which the direction of γ-quanta influx is localized must be a fraction of a degree for a large total recording-instrument aperture. Moreover, in all cases the instrument must yield an energy spectrum of the recorded γ-quanta.

It seems to us that the solution of the entire complex of problems can be obtained by means of an instrument whose operational principle is easily understood from Fig. 12. This instrument is based on a track-

ing two-section spark chamber and an ionization calorimeter.

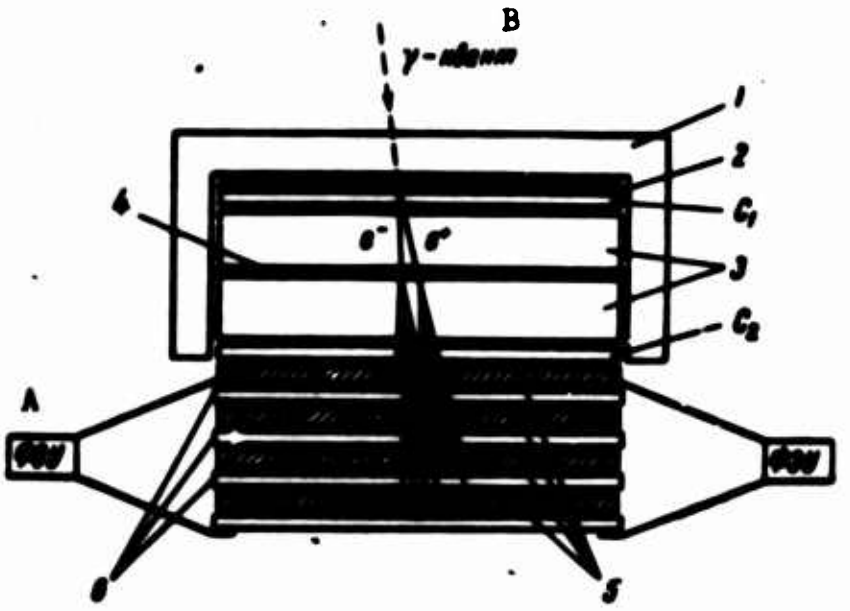


Fig. 12. Schematic representation of an instrument to study the energy spectrum of primary γ-rays and to seek out local sources of γ-quanta. 1) Bell-shaped cover of scintillation plastic; 2) electron-positron pair generator; C_1, C_2) Charged-particle counters; 3) two-section spark chamber; 4) barrier made of a substance with a high atomic number; 5) ionization-calorimeter absorbent; 6) scintillator; A) FEU [Photomultiplier]; B) γ-quantum.

The spark chamber is shielded from the top and at the sides by means of scintillator 1 which prevents the recording of those charged particles which have entered the spark chamber from without.

An emitter [radiator] 2 (a plate of lead or tungsten) is positioned above spark chamber 3 and beneath the emitter we have the charged-particle counter C_1 which is connected so as to coincide with the second counter C_2 and the ionization calorimeter 5 and 6. The γ-quantum striking the radiator is converted in the latter to an electron and a positron. These two particles, given sufficiently great γ-quantum energy, pass through the counters, through the spark chamber blocked off by means of plate 4 made of a substance having a large Z(Pb or W), and are then recorded by the ionization calorimeter. An electrical pulse controlling the operation of the spark chamber is produced as a result. The pattern in the spark chamber is photographed. The celestial sphere or the dial of an extremely precise clock is photographed at the pre-

cise instant of time, on the same piece of film.

In the region of low energies ~10^8 ev such a system yields reliable identification of the γ-quanta on the basis of the spark-chamber photograph, since in this case two particles converging at a single point - the radiator - can be seen in the upper section.

At energies of ~10^9 ev and higher, when the angle between the electron and the positron is so small that both streamers merge into one, and only a single track is visible in the photograph of the upper spark-chamber section, the developing shower will be observed in the lower section, i.e., a number of tracks converging in a small volume in plate 4. The direction of this shower will indicate the direction of the particle producing the observed pattern.

What are the potentials of such a γ-telescope?

These potentials are governed by two parameters: the accuracy with which the direction of the γ-quantum arrival is measured and the "speed" of the "control-counter - spark-chamber - ionization-calorimeter" system.

The direction of motion for the γ-quantum is defined by the position of the bisectrix of the angle formed by the components of the pair. Consequently, the smaller the angle between the electron and the positron, the more exact the definition of the direction of γ-quantum influx to the telescope. This angle is basically governed by the multiple scattering of the pair particles in the radiator and in the wall of the spark chamber.

As a matter of fact, the scattering angle for e^+ and e^- at the instant the pair appears is equal to $\theta_p \approx m_0c^2/E_\gamma$, where m_0c^2 is the energy of the electron resting mass, equal to 0.51 Mev, and E_γ is the energy of the γ-quantum.

The divergence angle for the pair components due to multiple scat-

tering is equal to

$$\theta_{pas} = \sqrt{\left(\frac{E_c}{E^+}\right)^2 t + \left(\frac{E_c}{E^-}\right)^2 t} = E_c \sqrt{t} \sqrt{\frac{1}{(E^+)^2} + \frac{1}{(E^-)^2}},$$

where E_c = 21 Mev, $\underline{t}$ is the thickness of the layer of substance through which the pair particles have passed, this quantity expressed in shower units , and E^+ and E^- are, respectively, the energy of the positron and the electron.

If we write that

$$E^+ = \alpha E_\gamma, \quad E^- = (1-\alpha) E_\gamma$$

$$\theta_{pas} = \frac{21 \text{Mev} \sqrt{t}}{E_\gamma} \sqrt{\frac{1}{\alpha^2} + \frac{1}{(1-\alpha)^2}}.$$

As an estimate we can assume $\alpha = 1/4$, $1 - \alpha = 3/4$ (or, conversely, $\alpha = 3/4$, $1 - \alpha = 1/2$). In this case,

$$\bar{\theta}_{pas} = \frac{17 \text{Mev} \sqrt{t}}{E_\gamma}$$

The comparison of the two expressions for θ_p and $\bar{\theta}_{ras}$ shows that when $t \geq 1/300$ of a shower unit $\bar{\theta}_{ras} > \theta_p$.

Since for effective recording of the γ-quanta $\underline{t}$ must be of the order of 1, $\theta_{ras} \gg \theta_p$.

Let us assume $t = 0.5$, $E_\gamma = 10^9$ ev, in which case

$$\bar{\theta}_{pas} \approx 0.7^\circ$$

Consequently, by means of the spark chamber it is possible to determine the direction of the primary γ-quantum at an energy $E_\gamma \geq 10^9$ ev with an accuracy no lower than 0.7°. At an energy $E_\gamma = 10^{10}$ ev this accuracy of $\sim 0.1^\circ$ approaches the accuracy of the determination of direction for a single particle.

If we assume that the area of the spark chamber is $S = 10^4$ cm^2 and that the geometric factor Γ for the entire device - for the chamber and the ionization calorimeter - is of the order of 10^4 cm^2·sterad, it will

be possible by means of such a telescope to:

1) record the flow of γ-quanta with energies of $\geq 10^9$ ev at an accuracy of 1% during the measurement day, if the intensity I_γ of the γ-quanta amounts to 10^{-4} of the proton intensity I_p; the accuracy will be 3-4% if $I_\gamma/I_p = 10^{-5}$;

2) detect local sources of γ-quanta with an energy of $E_\gamma \geq 10^9$ ev if the flow of γ-quanta from these in the vicinity of the earth is of the order of $2 \cdot 10^{-10}$ cm$^{-2} \cdot$sec^{-1} (with $I_\gamma/I_p = 10^{-4}$) and $\sim 6 \cdot 10^{-11}$ cm$^{-2} \cdot \cdot$sec^{-1}, if $I_\gamma/I_p = 10^{-5}$, i.e., such a γ-telescope, in principle, makes possible the recording of the direction of a stream of energy contained in the γ-quanta that is smaller by a factor of 10^{10} than the flow of energy produced by all of the charged cosmic-ray particles

5. Conclusion

This brief review of the new methods of working with high-energy cosmic-ray particles and some of the potentials for their application to investigations in the physics of outer space shows what fundamental strides in contemporary knowledge regarding primary cosmic radiation and nuclear processes involving high and very high energies can be expected in the near future, when the vigorously expanding conquest of outer space will provide physicists with the possibility of applying an established method to the investigation of cosmic radiation.

Despite the initial apparent paradox, we believe that the near future of the physics of high and very-high energy particles (10^{12}-10^{15} ev) is inseparably associated with the successful conquest of outer space. There are prospects along this line that were completely inaccessible to acceleration techniques of the past decade, to the method of studying cosmic radiation in the lower part of the atmosphere, and to the indirect methods carried out on the basis of extensive atmospher-

ic showers.

Received
9 June 1964

REFERENCES

1. N.L. Grigorov, V.S. Murzin, I.D. Rapoport, Zh. eksperim. i teor. fiz. [Journal of Experimental and Theoretical Physics], 34, 506, 1958.
2. N.L. Grigorov, A.V. Podgurskaya, V.A. Sobinyakov, V.Ya. Shestoperov, Zh. eksperim. i teor. fiz. 33, 1099, 1957.
3. N.L. Grigorov, V.S. Murzin, I.D. Rapoport. Zh. eksperim. i teor. fiz., 36, 1068, 1959.
4. V.V. Guseva, N.A. Dobrotin, N.G. Zelevinskaya, K.A. Kotel'nikov, A.M. Lebedev, S.A. Slavatinskiy, Izv. AN SSSR, ser. fiz. [Bulletin of the USSR Academy of Sciences, Physics Series], 26, 549, 1962.
5. B.V. Tolkachev. Zh. eksperim. i teor. fiz., 46, 43, 1964.
6. N.L. Grigorov, V.V. Guseva, N.A. Dobrotin, A.M. Lebedev, K.A. Kotel'nikov, V.S. Murzin, I.D. Rapoport, S.V. Ryabikov, S.A. Slavatinskiy, Tr. Mezhdunarodnoy konf. po kosmich. lucham., Izd-vo AN SSSR, [Transactions of the International Conference on Cosmic Radiation. Academy of Sciences USSR Press], 1, 140, 1960.
7. S.Z. Belen'kiy, Lavinnyye protsessy v kosmicheskikh luchakh. [Shower Processes in Cosmic Rays], Gostekhizdat, [State Power Engineering Press], 1948.
8. Kh.P. Babayan, Ya.S. Babetski, N.G. Boyadzhyan, Z.A. Buya, N.L. Grigorov, Ye.S. Loskevich, E.A. Mamidzhanyan, Ye.I. Massal'skiy, A.A. Oles', Ch.A. Tret'yakova, V.Ya. Shestoperov, Izv. AN SSSR, ser. fiz. ,26, 558, 1962.
9. N.L. Grigorov, Zh. eksperim. i teor. fiz., 45, 1919, 1963.

10. K.I. Alekseyeva, N.L. Grigorov, Dokl. AN SSSR [Proc. Acad. Sci. USSR], 117, 593, 1957.

11. L.D. Landau. Izv. AN SSSR, ser. fiz., 17, 15, 1953.

12. G.A. Milekhin, Izv. AN SSSR, ser. fiz., 26, 635, 1962.

13. N.L. Grigorov, V.Ya. Shestoperov, V.A. Sobinyakov, A.V. Polgurskaya. Materialy soveshchaniya po metodike tolstosloynykh fotoemul'siy [Materials From a Conference on the Methods of Thick-Layered Photoemulsions], OIYAI [United Nuclear Research Institute], 1957, page 168.

14. Kh.P. Babayan, S.I. Brikker, N.L. Grigorov, A.V. Podgurskaya, A.N. Savel'yeva, V.Va. Shestoperov, Zh. eksperim. i teor. fiz., 47, 379, 1964.

15. Kh.P. Babayan, N.G. Boyadzhyan, N.L. Grigorov, E.A. Mamidzhanyan, Ch.A. Tret'yakova, V.Ya. Shestoperov, Zh. eksperim. i teor. fiz. 46, 1525, 1964.

16. I.D. Rapoport, PTE [Journal of Experimental Equipment and Techniques], No 5, 130, 1960.

17. N.L. Grigorov, I.D. Rapoport, Avt. svid. No 145779. Vydano 3 maya 1962 g.[Patent No. 145779. Issued 3 May 1962].

18. P. Meyer and R. Vogt. Phys. Rev. Letters, 6, 193, 1961.

Manuscript Page No.	[Transliterated Symbols]
100	эв = ev = elektron-vol't = electron-volt
105	п = p = pogloshcheniye = absorption
113	я. а. = ya.a. = yaderno-aktivnyye [chastitsy] = nuclear-active [particles]
115	н = n = nuklon = nucleon
117	к.л. = k.l. = kosmicheskiy luch = cosmic ray
122	дет = det = detektor = detector
122	погл = pogl = poglotitel' = absorbent
132	я. р. = ya.r. = yadernyye rasshchepleniya = nuclear fissions
133	крит = krit = kriticheskiy = critical
136	Бэв = Bev = billion electron-vol't = billion electron-volt
137	с. и. = s.i. = sil'no ionizuyushchiy = strongly ionizing
141	д = d = detektor = detector
141	п = p = poglotitel' = absorbent
162	ФЭУ = FEU = fotoelektronnyy umnozhitel' = photomultiplier
164	порог = porog = porog = threshold
167	п = p = para = pair
168	рас = ras = raskhozhdeniye = divergence
168	Мэв = Mev = megaelektron-vol't = mega-electron-volt

UDK 523.16:523.16.031.6

RECORDING THE EFFECTS OF THE HIGH-ALTITUDE THERMONUCLEAR EXPLOSION OF 9 JULY 1962 WITH THE "KOSMOS-5" SATELLITE

Yu.I. Gal'perin and A.D. Bolyunova

The radiation effects of the American high-altitude thermonuclear explosion of 9 July 1962 over Johnston Island were recorded aboard the "Kosmos-5" satellite. A flash of hard radiation was detected at the instant of the explosion, far beyond the limits of its zone of direct visibility. This flash, known as a γ-dawn, apparently represents the recording of γ-radiation associated with the explosion. The explosion occurred at $12^h00^m09^s.8 \pm 1^s$ Moscow time. In the first minutes after the explosion positively charged particles – protons, α-particles, fission fragments, and positrons drifting toward the west – were detected approaching the oncoming "Kosmos-5" satellite. After 10 minutes electrons with energies of several Mev became dominant. In the region magnetically conjugate to Johnston Island, at altitudes of about 500 km and close to the Brazilian magnetic anomaly at altitudes of 200-300 km, comparatively soft electrons were detected and their absorption in the atmosphere evidently caused an aurora polaris over the Pacific Ocean. The maximum recorded intensity for 9 July, one hour after the explosion, amounted to $\sim 2 \cdot 10^9$ electrons $\cdot$ $cm^{-2} \cdot sec^{-1}$ over the South Atlantic. The maximum of the resulting radiation belt was situated over the magnetic equator, with an altitude of ~1350 km over Johnston Island, changing with longitude. The drop in intensity was most significant during the first twenty-four hours of the existence of the belt, and then gradually diminished. The intensity in the center of the belt diminished by approximately an order of magnitude over a period of four months. An increase in the radiation background was observed significantly below the region in which the stable radiation belts existed. The rate of disappearance for this excess over the background produced by cosmic rays is close to the rate at which the intensity diminished at the maximum of

artificial belt. The nature of the recorded phenomena is discussed here.

On 28 May 1962, the satellite "Kosmos-5" was launched into an orbit inclined to the equator at an angle of 49°, with the apogee at 1600 km and the perigee at an altitude of 204 km. By 9 July the altitude of the apogee had diminished to 1512 km, while the apogee point had shifted to the northern hemisphere to a latitude of 49° N.

The apparatus aboard the "Kosmos-5" satellite had been intended for geophysical investigations in connection with geoactive corpuscular radiation of low energies and was described in detail in [1, 2]. The analysis of the consequences resulting from the high-altitude thermonuclear explosion of 9 July 1962 was carried out primarily on the basis of measurements provided by the El electron indicator and the Geiger counter. Let us briefly recall their basic characteristics. The El indicator consisted of a fluorescent $Sr_3(PO_4)_2$[Eu] screen 1.4 $mg \cdot cm^{-2}$ thick, covered by a piece of aluminum foil 0.4 $mg \cdot cm^{-2}$ thick inside a solid angle about 1/12 steradian (about 10 steradians were covered with from 1 to 3 $g \cdot cm^{-2}$ and the rest with $\gtrsim 10$ $g \cdot cm^{-2}$). The glow of the screen was recorded by means of a photomultiplier whose signal was amplified and transmitted to a radio-telemetry system with a memory unit. The sensitivity of the El indicator was periodically controlled during the flight by means of a beam of electrons with a tritium target. The indicator is virtually insensitive to x-ray emission. Electrons with an energy $\lesssim 1$ Mev were picked up only through the indicator window. The rotation of the satellite made it possible to distinguish this radiation on the basis of the characteristic modulation resulting from the anisotropy of charged-particle distribution in the trap formed by the magnetic field of the earth. Electrons with energies $\gtrsim 2$ Mev penetrated the indicator from virtually all sides and for this reason their anisotropy led virtually to no modulation with rotation.

The Geiger counter mounted on the inside of the satellite frame was a standard STS-5 halogen counter shielded with 3.4 $g \cdot cm^{-2}$ of lead. The elements of the construction provided additional shielding of about 0.8 $g \cdot cm^{-2}$ of aluminum in a solid angle of about 2π steradians and up to ~25 $g \cdot cm^{-2}$ in the remaining direction. The counter area was 4.3 cm^2. With the high intensity of the particles in the artificial belt, the counter frequently changed over to the descending branch of its counting characteristic. This characteristic for particularly high levels of intensity was determined directly from telemetry-measurement data obtained in a number of orbits inside the artificial belt by comparing the counting rates with the electron-indicator readings which did not exceed the scale.

The sensitivity of the apparatus to electrons was determined on the basis of results obtained from laboratory calibration prior to the launch and on the basis of the calibration of a completely analogous complex of apparatus in a far wider range of energies. On the basis of these data we can evaluate the sensitivity of the El indicator and the Geiger counter for the recording of β-radiation from fission products captured by the terrestrial magnetic field. The energy spectrum of the electrons in the artificial radiation belt may be assumed to conform with laboratory measurements [3] on the basis of the pattern of electron-drift rates obtained for 9 July 1962 [4]. In this case 1 μa of photomultiplier current in the El indicator, with a pitch angle of 90°, corresponded to $2 \cdot 10^7$ particles $\cdot$ $cm^{-2} \cdot sec^{-1}$, while 1 pulse $\cdot$ sec^{-1} with the Geiger counter corresponded to $2 \cdot 10^3$ particles $\cdot$ $cm^{-2} \cdot sec^{-1}$. Prior to the explosion of 9 July 1962, the counter had primarily recorded protons with an energy of >50 Mev and the background produced by cosmic rays [5].

The first effect of the explosion on 9 July 1962 was recorded at

the instant $12^h00^m09^s.8 \pm 1^t$ in the form of a sharp jump in the counting rate of the Geiger counter, approximately by three orders. (The measurement ambiguity at the initial instant of time is explained by the fact that we cannot exclude the possibility here of the counter having transferred to the descending branch of the counting characteristic). The electron sensor signal in this case changed by no more than 0.01 μa, which indicates the great hardness of the radiation. At that instant of time the "Kosmos-5" satellite was situated at a point whose coordinates were $\varphi = 44^\circ.9$ N, $\lambda = 115^\circ.7$E and h = 1442 km at a distance of about 7500 km from the region of the explosion (along an arc of the great circle of the terrestrial sphere). At this distance from the satellite the zone of direct visibility was higher than 1200 km and, consequently, the point of the explosion (h = 400 km [6]) was deep beneath the horizon, even at the altitude of the satellite. This burst, lasting about 2 minutes, is shown in Fig. 1.

Let us consider the possible factors responsible for the described burst in the counting rate beyond the limits of the direct-visibility zone of the point of explosion.

The electrons injected by the explosion and captured in the trap formed by the magnetic field of the earth move toward east as a result of drift and, having gone around the earth, overtake the satellite approaching the point of the explosion from the west. However, even for an energy of 10 Mev at the geomagnetic latitude of the satellite the drift time for the electrons from the point of the explosion to the satellite is more than 5 minutes and, consequently, the possibility of their being recorded at the instant $12^h00^m09^s.8$ must be excluded.

The positively charged particles (protons, α-particles, fission fragments, or positrons) which appear in the trap formed by the magnetic field drift in a westerly direction toward the satellite. However,

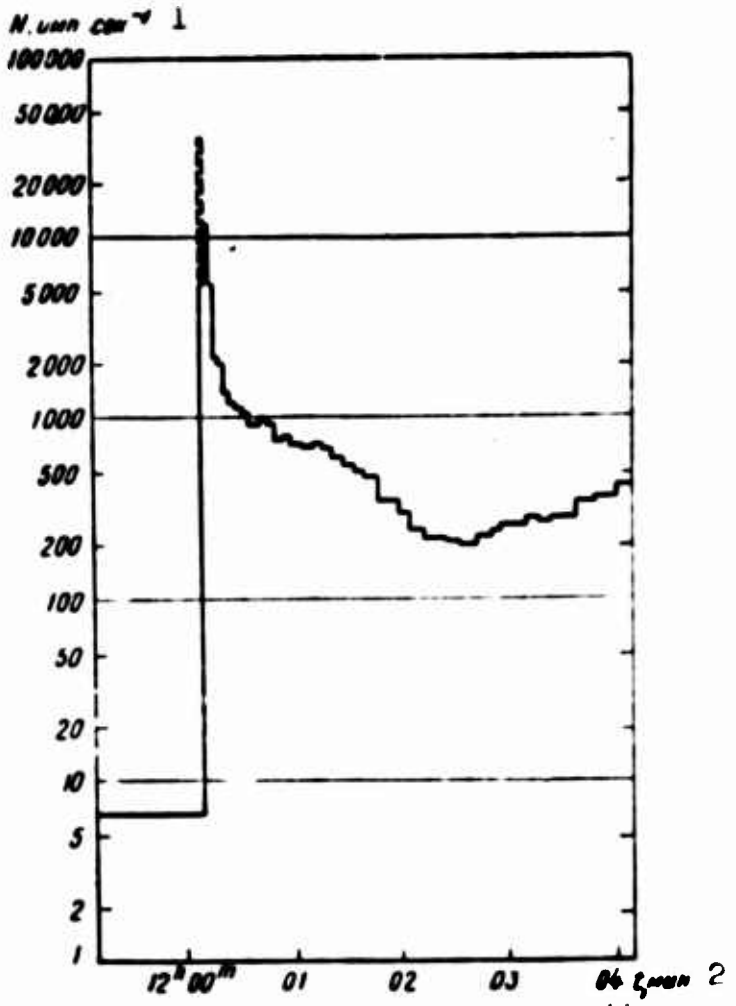


Fig. 1. Radiation surge on 9 July 1952 at $12^h00^m09^s.8$. The counting rate of the shielded Geiger counter is plotted along the axis of ordinates in logarithmic scale (3.4 $g \cdot cm^{-2}$ Pb + 0.8 $g \cdot cm^{-2}$ Al) in pulses per second, corrected for "dead time." 1) N, pulses$\cdot sec^{-1}$; 2) t, in minutes.

the particle energies required for longitudinal drift at the high speeds required to accomplish the shift from the point of the explosion to the longitude of the satellite within several seconds are, apparently, somewhat too great (>500 Mev).

Taking into consideration the exceptionally low efficiency of the halogen counter with respect to neutrons, we might assume that the recording of these directly is also quite unlikely. On the other hand, the neutrons formed in the explosion, leaving the vicinity of the explosion without hindrance, decay throughout the entire region of the terrestrial geomagnetic field which is within the zone of the direct visibility of the point of explosion. Their β-disintegration could result in the ap-

pearance of electrons with small pitch angles within a fraction of a second after the explosion on the same magnetic line of force as the one on which the satellite was situated at the moment [7]. However, even without a more detailed quantitative calculation it is obvious that if the described surge in the counting rate of the shielded counter were a result of electrons with energies of several hundred kev, appearing with the β-disintegration of the neutrons, it would have been recorded by the electron indicators aboard the satellite.

There is, however, another means of recording the neutron waves being propagated from the region of the explosion, i.e., through the γ-radiation appearing as a result of the (n, γ) reactions within the frame of the satellite and in the atmosphere in the case of neutron irradiation. This hypothesis must be subjected to a detailed examination with consideration of the multiple scattering and deceleration of the atmospheric neutrons. It would be nice if we could explain the shape of the peak by the passage of the satellite through the neutron wave, with the more decelerated neutrons arriving later, so that at the 1st, 10th, and 100th seconds we have the influx of neutrons with energies, respectively, of 350 kev, 3.5 kev, and 30 ev, causing the γ-radiation as a result of the (n, γ) reactions.

The appearance of γ-radiation outside of the zone of direct point-of-explosion visibility within the first two minutes after the explosion is possible also as a result of the multiple scattering of the delayed γ-radiation from the region of the explosion (the instantaneous γ-radiation is somewhat too brief in duration in order to produce this effect), and also in the event of the ejection of γ-radioactive fission products to altitudes in excess of 1200 km. Holding the average kinetic energy of the fragments to amount to about 1 Mev/nucleon [8], we find that these, under favorable conditions, are capable of attaining the zone of

direct visibility in approximately 0.1 second. At the end of this interval of time the γ-radiation of the fission products is still quite significant (about 1 Mev·sec^{-1} per 1 fission, according to [9]) and, apparently, can be recorded. However, the rapid diminution of the counting rate in the interval $12^h00^m12^s$-$12^h00^m24^s$ does not agree with the data on the rate of descent for γ-radioactivity of fission products cited in [9]. On the other hand, the tremendous extent of the subsequently formed artificial radiation belt (see below) indicates that the products of the explosion were actually carried to significant altitudes.

We may therefore draw the conclusion that the described surge of hard radiation, detected by the Geiger counter aboard the "Kosmos-5" satellite at the instant of the high-altitude thermonuclear explosion beyond the limits of direct visibility, most likely represents the recording of γ-radiation produced by the explosion. We will refer to this phenomenon in the future as the "γ-dawn."

It follows from all of the above that the explosion of 9 July 1962 occurred at $12^h00^m09^s.8 \pm 1^s$ Moscow time.

We are not familiar with any other data obtained by means of satellites at the time of the explosion on 9 July 1962 over the Pacific Ocean, with the exception of the counter measurements taken with the British-American "Ariel" satellite [10]. At the instant of the explosion this satellite was distant from the point of the explosion by approximately the same distance (7400 km) as the "Kosmos-5," and on the geomagnetic meridian passing close to Johnston Island, with $L \approx 5$, but its altitude was 819 km and, consequently, the region of direct visibility of the point of the explosion lay above 2000 km. A surge in the counting rate was also recorded aboard the "Ariel" satellite, this surge lasting 2 minutes; however, the low interrogation rate of the memory system

(1 interrogation in 92 seconds) and the lack of precision in time coordination (± 10 sec) enabled the authors [10] to establish only that the surge occurred "after no more than 20 ± 10 seconds after the explosion."* Moreover, the data of the single counter aboard the satellite made it impossible to evaluate the hardness of radiation. The authors [10], attempting to interpret this surge, examined mechanisms for which the time interval between the explosion and the surge involved tens of seconds, and they came to the conclusion that the most valid explanation for the surge "is the scattering from the trap of outer-belt electrons whose pitch angles were altered under the action of the magnetohydrodynamic perturbation associated with the explosion." In any event, such an explanation is not applicable to the surge in hard radiation recorded by the "Kosmos-5" satellite, without considering the possibility of a simultaneous acceleration of electrons (or even protons) in the natural radiation belt. However, the virtual simultaneity of the instant of explosion and the surge in radiation aboard the satellite "Kosmos-5" (with an accuracy of $\pm 1^s$), apparently, precludes the possibility of such an interpretation.

On the other hand, we can assume that the radiation surge aboard the "Ariel" satellite was indeed recorded immediately after the explosion, as was the case aboard the "Kosmos-5." This is borne out by the results from the measurements of radio-wave absorption in Alaska [11] along virtually the same geomagnetic meridian as Johnston Island and "Ariel" satellite at the instant of time under consideration, but in the higher geomagnetic latitudes ($L \approx 5.5$). On the basis of the data in [11], the ionization of the ionosphere over Alaska increased sharply within no more than after 2 seconds after the explosion at $L \gtrsim 5$. For D-layer region over Alaska under consideration, where the ionization took place (at a distance of 5600 km from Johnston Island), and for the

"Ariel" satellite, the boundary of the zone of direct visibility of Johnston Island was situated at approximately the same altitude, and the geometric conditions for the scattering of radiation from the region of explosion are also analogous. Since the geomagnetic latitudes (or the values of L) were also virtually identical, any of the mechanisms responsible for the surge along the geomagnetic longitude of the explosion region, i.e., the propagation of charged particles in the direction of increasing L, the multiple scattering from the explosion region of γ-radiation or neutrons (with subsequent (n, γ)-reactions), or the ejection of γ-radioactive products to altitudes >2000 km, i.e., in the zone of direct visibility, apparently should have led to the simultaneous appearance of ionizing radiation over Alaska and aboard the "Ariel" satellite. It may therefore be assumed that the described surge of the γ-dawn at the instant of the explosion in actual fact was recorded by the counter aboard the "Ariel" satellite and, possibly, also served as the direct cause of the jump in absorption in the ionosphere over Alaska. However, is not out of the question that the ejection of the products of explosion to great altitudes plays a significant role in the γ-dawn phenomenon, i.e., ejection into the zone of direct visibility from the satellite. In this case the surge of the γ-dawn aboard the "Ariel" satellite may have started significantly later than aboard the "Kosmos-5," depending on the rate of propagation for the products of the explosion from an altitude of ~1200 km (required for the "Kosmos-5") to an altitude of ~2000 km (required for the "Ariel"), while the ionospheric absorption over Alaska was brought about by certain other factors. Unfortunately, neither the time of the surge nor its intensity in these satellites could be compared, since the counter aboard the "Ariel" was off scale throughout virtually the entire 2 minutes of the γ-dawn surge.

After the first surge the intensity of the hard radiation sharply increased throughout the entire region of the space intersected by the orbit of the "Kosmos-5" satellite. Starting as of 12^h02^m the Geiger counter recorded hard radiation with increasing intensity. It is possible that γ-radiation was also present in the expanding cloud of products of the explosion, but on the basis of subsequent observations it is possible to draw the conclusion that these might have been positrons or radioactive fission fragments with energies of no less than several Mev, drifting toward the west, in the direction toward the satellite. Within 8 minutes after the explosion, when the "Kosmos-5" was located to the east of Japan (the natural coordinates were L = 1.4 and B = 0.25), there also began a sharp increase in the electron-indicator signals recording comparatively soft radiation. At 12^h10^m (L = 1.3; and B = 0.24) the signal increased by more than an order of magnitude. It is significant that the energy of electrons capable of attaining the longitude of the satellite, on having completed a turn around the terrestrial globe, even at 12^h08^m and 12^h10^m would be no less than 7.7 Mev and 6.5 Mev, respectively, while there are virtually no such electrons in the β-spectrum of the fission products [4]. Moreover, the trajectories of the mirror points of the particles recorded from 12^h08^m to 12^h10^m descend to the surface of the earth in the region of the Brazilian magnetic anomaly, which must be passed by the electrons prior to reaching the "Kosmos-5." Thus it is highly unlikely that the indicator signals recorded at that time were caused by electrons, and the indicator is not sensitive to γ-radiation. Apparently, the "Kosmos-5" satellite was capable of recording positively charged particles appearing in conjunction with the thermonuclear explosion: protons, α-particles, fission fragments, or positrons.

The intensity maximum recorded at 12^h13^m at at altitude of 1100 km

and virtually at the latitude of Johnston Island, but to the west ($L \approx$ ≈ 1.2 and $B = 0.23$) is obviously associated with the flight of the satellite through the resulting radiation belt, since in the days following extensions of this new artificial belt were situated in this region. One day later, the intensity in this region diminished by a factor of ~30. During the pass over the geomagnetic equator at 12^h20^m at an altitude of about 800 km the electron indicators recorded a soft-radiation intensity with deep modulation due to anisotropy. At the same time the Geiger counter was recording extremely hard radiation ($\gtrsim 10^3$ pulses·sec^{-1}), possibly also associated with the γ-radioactivity of the products of the explosion being propagated through the atmosphere.

Approximately at 12^h24^m-12^h28^m the "Kosmos-5" satellite, at an altitude of 500-600 km, passed through the region magnetically conjugate to the region of the explosion. A new hard-radiation maximum was recorded at this time, as well as a soft component. Then with approach to the perigee, the intensity of soft radiation diminished sharply; however, the counting rate of the counter continued to remain high. After Δt = 93 minutes after the explosion, in the southern hemisphere ($\varphi = 49^\circ$ S; $\lambda = 250^\circ$ E) at an altitude of about 210 km, the counting rate of the Geiger counter amounted to approximately 1000 pulses·sec^{-1}, which is greater by a factor of approximately 150 than prior to the explosion at this same point. An extremely interesting phenomenon was detected by the electron indicators at 12^h40^m-12^h50^m during the approach to the Brazilian magnetic anomaly. Apparently, electrons with energies of 50-100 kev were recorded, and their intensity increased sharply at altitudes of 200-300 km ($B > 0.23$) with diminishing magnetic field strength B along the trajectory. Apparently, the "Kosmos-5" satellite at this time "overtook" the electrons with energies of 50-100 kev completing

their first drift revolution about the earth that had been injected as a result of the explosion. The higher-energy electrons injected into the same shells and exhibiting greater drifting rates, had by that time already been absorbed in the atmosphere above the Brazilian magnetic anomaly and it therefore became possible to re[illegible] only the soft component. The soft particles, drifting slowly toward the Brazilian anomaly, gradually descended into the dense atmosphere and their absorption, apparently, produced an aurora polaris over the Pacific Ocean.

Particularly great intensity in the high-energy particles was observed in the vicinity of the Brazilian magnetic anomaly where the magnetic field strength had diminished significantly in comparison with the other longitudes. Within an hour after the explosion, over an extensive region around the geomagnetic equator ($\lambda \sim 325\text{-}340^{\circ}E$) at altitudes of 600-850 km the omnidirectional intensity of electrons with energies of $\lesssim 1$ Mev reached $2 \cdot 10^9$ particles $\cdot$ cm$^{-2} \cdot$ sec^{-1} according to the data of the electron indicator. The readings of the Geiger counter aboard the "Kosmos-5" satellite for the line L = 1.25, B = 0.20 gauss over the South Atlantic within Δt = 60 minutes after the explosion can be compared with the published measurement data for the "Injun-1" satellite which carried a virtually identically shielded SpB counter which passed through this same line over South Africa at the instant of time Δt = 45 min [12]. Because of the various counter areas and the inaccurate calibration, it is more convenient to compare the measured counting rates in pulses $\cdot$ cm$^{-2} \cdot$ sec^{-1}, i.e., $7 \cdot 10^4$ for the "Injun-1" and $1.2 \cdot 10^5$ for the "Kosmos-5." Bearing in mind that at these very early stages of the existence of the artificial radiation belt, the intensities may actually differ somewhat and it may therefore be concluded that the agreement is satisfactory. The greatest values of L which were attained by the "Kosmos-5" satellite during the period in which the artificial radi-

ation belt appeared reached L = 2.3. The counting rate increased significantly at these latitudes as well. Thus, for example, within 87 minutes of the explosion at an altitude of 1450 km (L = 2.0) over the territory of the Soviet Union the counting rate of the Geiger counter amounted to approximately 1000 pulses·sec^{-1} which is greater by a factor of 30 than for the same region a day prior to the explosion. This indicates that the distribution of the intensity recorded by the single shielded counter changed significantly throughout a tremendous region of the space surrounding the earth, all the way to distances of several earth radii.

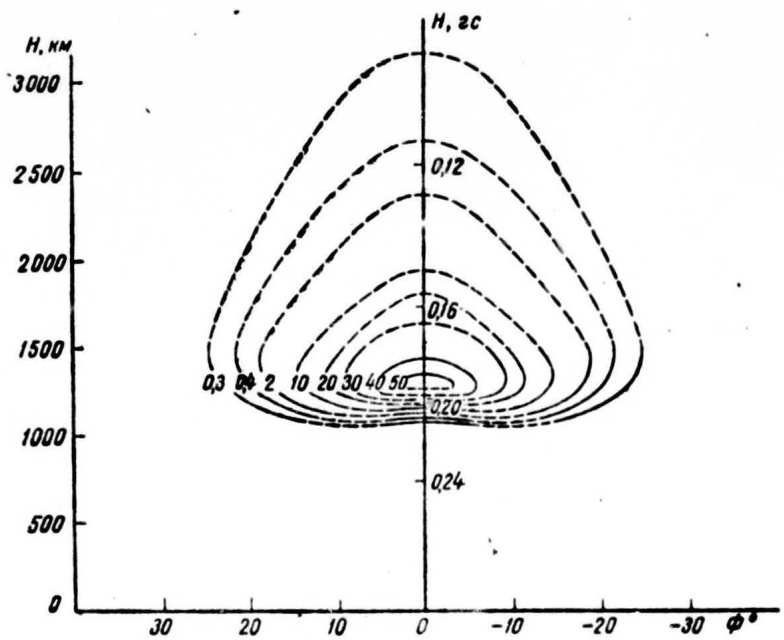


Fig. 2. Distribution of intensity in artificial radiation belt over Johnston Island 10 days after the explosion. The geomagnetic latitude is plotted along the axis of abscissas. The isolines indicate the current magnitudes of the El sensing element in μa (for conversion to an omnidirectional stream of electrons expressed in units of electron·cm^{-2}·sec^{-1} it is necessary to multiply by $2 \cdot 10^7$). The values of terrestrial magnetic field strength are indicated in the equatorial plane. In the region above the apogee of the "Kosmos-5" satellite the isolines are extended (by means of a dashed line) approximately along the geomagnetic lines of force.

The nature of the change in particle intensity as a function of the natural coordinates B, L approximately within an hour after the explo-

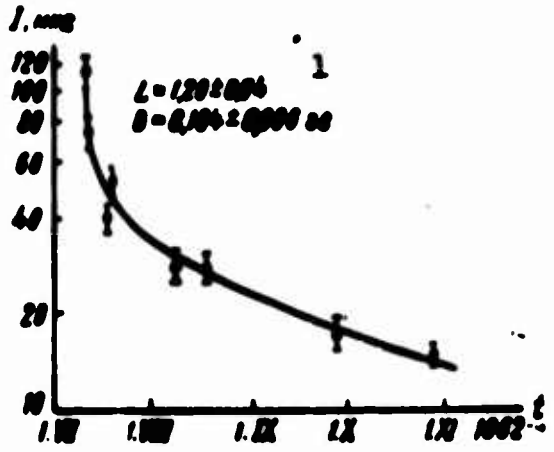


Fig. 3. Time reduction in maximum intensity of electrons with energies of $\lesssim 1$ Mev in the region of the center of the artificial belt. The current of the El electron indicator is plotted along the axis of ordinates in logarithmic scale and is given in v. 1) Gauss.

sion indicates that an artificial radiation belt had been formed by that time. The maximum intensity of the belt was found in the vicinity of the geomagnetic equator (or in the vicinity of the cosmic-ray equator), its altitude being a function of longitude, since it is associated with the local intensity of the real terrestrial magnetic field. The altitude of the intensity maximum over Johnston Island was approximately 1350 km (Fig. 2), while over the Atlantic Ocean it was 700-900 km. With the passage of time the intensity of the particles at the maximum of the artificial radiation belt gradually diminished. The rate of this diminution was at its maximum during the first days of the existence of the artificial belt and then gradually diminished (Fig. 3). The figure shows that over a period of four months the intensity at the center of the artificial radiation belt diminished by appoximately an order of magnitude. At the same altitudes, but in higher geomagnetic latitudes, the rate of intensity reduction during the first 10 days was significantly greater and continued to increase with increasing latitude. The rate of diminution during the first days was also quite high at lower altitudes.

The "Kosmos-5" detected a new phenomenon at even lower altitudes, i.e., a rise in the radiation background noise recorded by the counter in the low latitudes beneath the region of the radiation belt. Figure 4 shows the reduction in the counting-rate excess for $L = 1.10$ and $B \geq 0.24$ over the background noise produced by the cosmic rays (~2.5

pulses·sec^{-1}). In comparison with the level for 10 July the 10-fold reduction occurred within approximately $5 \cdot 10^6$ seconds. A more-or-less analogous increase in background noise beneath the radiation belt also occurred in the higher latitudes. Let us consider the possible factors responsible for this new phenomenon. We might suppose that this was the γ-radiation of the fission products that had spread out uniformly throughout the atmosphere at medium altitudes (50-100 km) over a period of several days. However, the measured rate of decrease in intensity is lower than the rate of γ-radioactivity reduction according to [9], although as a result of the downward diffusion of heavy fragments we should expect an even more rapid reduction than on the basis of [9]. This explanation is therefore not valid. Another possibility is the recording of protons with energies >50 Mev whose distribution by pitch angles and (possibly by energies) was disrupted as a result of the phenomena in the magnetosphere accompanying the nuclear explosion (an analogous assumption was made in [10] with respect to electrons to explain γ-dawn surge). In this case the rate of intensity reduction is also somewhat too low, since the average density of the atmosphere along the drift trajectory for such particles is tremendous. For purposes of comparison let us indicate that the excess intensity of protons with energies ~55 Mev, detected inside the radiation belt, i.e., for $B < 0.22$ gauss (with minimum mirror-point altitudes of ~250-700 km in the Brazilian anomaly) with the aid of nuclear photoemulsions aboard the satellites [13], disappeared at approximately the same rate. However, in the latter case the average density of the atmosphere along the drift trajectory must be lower by several orders of magnitude than for the region below the radiation belt described above. It may be assumed that this radiation at low altitudes represents the high-energy electrons pouring out of the region of the artificial radiation belt as a result

of the random reduction of the pitch angle on collision with an atmospheric atom or electron. However, in this case the excess at L = 1.20-1.25, where the belt maximum is located, apparently would exceed the excess at L = 1.10. The experimental data do not confirm this. Finally, for the time being we cannot exclude the possibility that this slowly diminishing excess in the counting rate is brought about by the radioactivity induced in the hull of the satellite as a result of the irradiation after the explosion. In this case this possibly might simply be an extension of the γ-dawn phenomenon. The study of this effect is continuing.

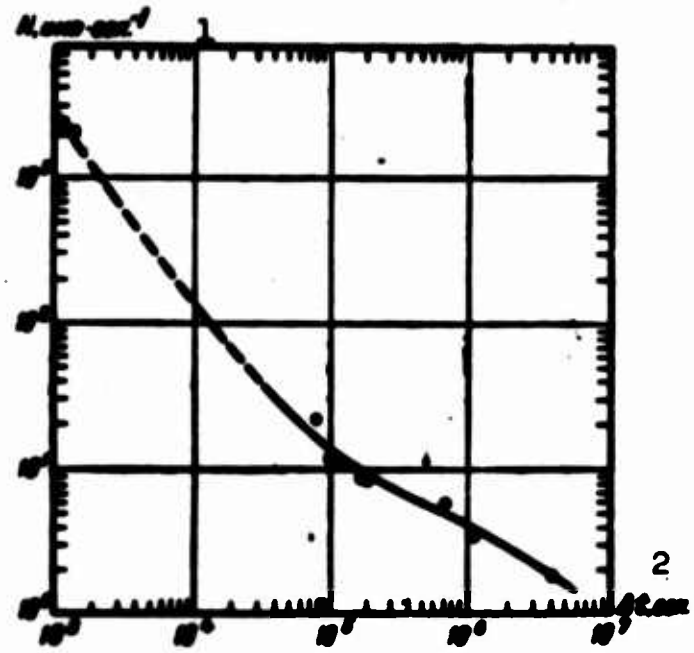


Fig. 4. Time reduction in the counting-rate excess over the background of cosmic rays near the fission equator (L = 1.10). The time interval (sec) from the instant of the explosion is plotted along the axis of abscissas in logarithmic scale; the squares show the first observations after the explosion which, at least partially, may be associated with the recording of the delayed γ-radiation of the fission products. 1) N, pulses·sec^{-1}; 2) Δt, in sec.

The above-described measurements indicate the extremely great life span of the artificially introduced hard radiation, particularly in the equatorial regions. This made considerably more complex the geophysical investigations of natural corpuscular radiation in the upper atmosphere, even after a lapse of several months after the explosion.

Earlier [14] we had already taken note of the fact that the measurements of the hard component in the natural radiation belts, these measurements having been carried out after an intense series of nuclear tests in the summer of 1958 (which included high-altitude nuclear explosions), could have been complicated by the recording of strongly penetrating radiation due to fission products. Apparently, available data of the measurements carried out at that time could advantageously be reanalyzed from this standpoint.

In conclusion, we consider it our duty to express our deep-felt gratitude to V.I. Krasovskiy for his advice and constant attention to this project, as well as to the co-workers at the Department of the Physics of the Upper Atmosphere of the Institute of the Physics of the Atmosphere, USSR Academy of Sciences, for their great assistance in the execution of the project and their useful consideration of the result, as well as to the many co-workers in other agencies and institutions who helped us in this work.

Received

11 March 1964

REFERENCES

1. V.I. Krasovskiy, Yu.I. Gal'perin, V.V. Temnyy, T.M. Mulyarchik, N.V. Dzhordzhio, M.Ya. Marov, A.D. Bolyunova, B.P. Potapov, O.L. Vaysberg and M.L. Bragin, Geomagnetizm i aeronomiya [Geomagnetism and Aeronomy], 3, No. 3, 401, 1963.
2. Yu.I. Gal'perin and V.I. Krasovskiy, Kosmich. issled. [Cosmic Research], 1, No. 1, 126, 1963.
3. R.E. Carter, F. Reaines, J.J. Wagner and M.E. Wyman, Phys. Rev., 113, 280, 1959.
4. G.R. Ochs, D.T. Farley, Jr. and K.L. Bowles, J. Geophys. Res., 68,

701, 1963.

5. V.V. Temnyy, Kosmich. issled., 1, No. 1, 139, 1963.
6. W.L. Brown, W.H. Hess and J.A. Van Allen, J. Geophys. Res., 68, 605, 1963.
7. C.M. Crain and P. Tamarkin, J. Geophys. Res., 66, 35, 1961.
8. Experimental nuclear physics, edited by E. Segre, II, Foreign Literature Press, 1955, page 121.
9. K. Way and E.R. Wigner, Phys. Rev., 73, 1318, 1948.
10. A.C. Durney, H. Elliot, R.J. Hynds and J.J. Quenby, Nature, 195, No. 4848, 1245, 1962.
11. R.P. Basler, R.B. Dyce and H. Leinbach, J. Geophys. Res., 68, 741, 1963.
12. B.J. O'Brien, C.D. Laughlin and J.A. Van Allen, Nature, 195, 939, 1962.
13. R. Filz, E. Holeman and H. Yagoda, Trans. A.G.U., 44, 78, 1963.
14. I.S. Shklovskiy, V.I. Krasovskiy and Yu.I. Gal'perin, Izv. AN SSSR, ser. geofiz. [Bulletin of the USSR Acad. Sci. Geophysics Series], No. 12, 1799, 1959.

Manuscript Page No.

[Footnote]

180 "The time of the explosion in [10] is assumed to be $12^h00^m15^s$, i.e., 5.2 seconds later than the time recorded aboard the "Kosmos-5" satellite, and later by 6 seconds than the time of $12^h00^m09^s$ indicated in [6].

UDK 538.691

THE POSSIBILITY OF CAPTURING CHARGED PARTICLES IN THE FIELD OF A MAGNETIC DIPOLE IF ENERGY IS LOST AS A RESULT OF EMISSION*

V.M. Vakhnin and I.N. Shvachunov

The investigation of the possibility of capturing charged particles in the field of a magnetic dipole, carried out earlier for plane (two-dimensional) motion, is extended in this article to the case of an arbitrary three-dimensional particle motion.

Use is made of a method of phase trajectories in a four-dimensional phase space. It is demonstrated that the "critical trajectories" and the possibility of capture exist in three-dimensional motion as well.

The possibility of capturing charged particles in the field of a magnetic dipole as a result of the existence of "critical" trajectories for which infinitely small losses of kinetic energy on the part of a particle lead to the transition from an "uncaptured" to a "captured" trajectory were considered in References [1, 2].

The analysis was carried out by the method of phase trajectories, with the specific study of the trajectories limited to the case of the motion of a particle in the equatorial plane of the dipole. The phase space in this case was described by 2 variables, i.e., it had the form of a two-dimensional phase plane. It was pointed out in Reference [1] that in order to analyze the arbitrary motion of the particle it must be represented in the somewhat more complex four-dimensional phase space. The present work deals with an investigation of the existence and nature of the "critical" trajectories beyond the equatorial plane which are described by the four-dimensional phase space having the coordinates

$$w = \frac{r}{a}, \quad u = \frac{dr}{d\varphi}\frac{1}{a}, \quad \theta, \quad \psi = \frac{d\theta}{d\varphi}$$

(r is the distance between the particle and the center of the dipole; a is the Stoermer length unit; and φ, θ are the angular spherical coordinates).

DIFFERENTIAL EQUATIONS FOR THE "PHASE TRAJECTORIES" OF CHARGED-PARTICLE MOTION

The equation of motion for a charged particle in a permanent magnetic field $\vec{H}$ can be presented in vector form:

$$m\frac{dv}{dt} = \frac{e}{c}[vH], \qquad (1)$$

where $\vec{v}$ is the particle velocity, e is the electrical charge of the particle, which may be either positive or negative; c is the electrodynamic constant, $m = m_0/(1 - \beta^2)^{1/2}$ is the relativistic mass of the particle (m_0 is the resting mass, $\beta = v/c$).

The differential equation of motion for a charged particle in the field of a magnetic dipole is conveniently considered in a spherical system of coordinates r, θ, φ (the angle θ is counted from the equatorial plane of the dipole, with $-\pi/2 < \theta < \pi/2$). If the coordinate origin is made to coincide with the center of the dipole, and if its magnetic moment $\vec{M}$ were to be directed along the negative polar axis, then it is well known [3] that the intensity $\vec{H}$ of the dipole magnetic field will be expressed by

$$H = \frac{3(Mr)r}{r^5} - \frac{M}{r^3}. \qquad (2)$$

The components of this field can be presented in spherical coordinates as follows:

$$H_r = -\frac{2M\sin\theta}{r^3}, \qquad (3)$$

$$H_\theta = \frac{M\cos\theta}{r^3}, \qquad (4)$$

$$H_\varphi = 0. \qquad (5)$$

Having expressed $\vec{v}$ and $\vec{H}$ in terms of their components in spherical coordinates in the vectorial equation (1) and considering (5), we obtain the following differential equations:

$$m = (\ddot{r} - r\dot{\theta}^2 - r\cos^2\theta\dot{\varphi}^2) = -\frac{e}{c} r\cos\theta\dot{\varphi}H_\theta, \tag{6}$$

$$\frac{m}{r}\left[\frac{d}{dt}(r^2\dot{\theta}) + r^2\sin\theta\cos\theta\dot{\varphi}^2\right] = -\frac{e}{c} r\cos\theta\dot{\varphi}H_r, \tag{7}$$

$$\frac{m}{r\cos\theta}\frac{d}{dt}(r^2\cos^2\theta\dot{\varphi}) = \frac{e}{c}(\dot{r}H_\theta + r\dot{\theta}H_r). \tag{8}$$

Having introduced the new variable ds = vdt (s is the length of the arc measured along the trajectory of the particle) and having substituted the value of H_r and H_θ from (3) and (4), respectively, we obtain

$$r'' = r\theta'^2 + r\cos^2\theta\varphi'^2 - \frac{a^2\cos^2\theta\varphi'}{r^3}, \tag{9}$$

$$\theta'' = \frac{2a^2\sin\theta\cos\theta\varphi'}{r^3} - 2\frac{r'\theta'}{r} - \sin\theta\cos\theta\varphi'^2, \tag{10}$$

$$\varphi'' = 2\varphi'\theta'\operatorname{tg}\theta - 2\frac{r'\varphi'}{r} - \frac{2a^2\theta'\operatorname{tg}\theta}{r} + \frac{a^2r'}{r^4}, \tag{11}$$

where $a = \sqrt{eM/mvc}$ is the characteristic Stoermer coefficient, while the prime denotes differentiation with respect to the new variable s.

In order to derive the equations for the phase trajectories of a charged particle in phase space, let us assume the polar angle φ as the independent variable. In this case, using the equality [4]

$$\frac{d^2r}{d\varphi^2} = \frac{r''\varphi' - \varphi''r'}{\varphi'^3}$$

from Eqs. (9) and (11) we will obtain

$$\frac{d^2r}{d\varphi^2} = r\left(\frac{d\theta}{d\varphi}\right)^2 + r\cos^2\theta + \frac{2}{r}\left(\frac{dr}{d\varphi}\right)^2 - \frac{a^2}{r}\left(\frac{ds}{d\varphi}\right)\left[\frac{\cos^2\theta}{r} + \right.$$

$$\left. + \frac{1}{r^3}\left(\frac{dr}{d\varphi}\right)^2 - \frac{2}{r^2}\left(\frac{dr}{d\varphi}\right)\left(\frac{d\theta}{d\varphi}\right)\operatorname{tg}\theta\right]. \tag{12}$$

Analogously, taking into consideration that

$$\frac{d^2\theta}{d\varphi^2} = \frac{\theta''\varphi' - \varphi''\theta'}{\varphi'^3},$$

from Eqs. (10) and (11) we will have

$$\frac{d\psi}{d\varphi} = -\sin\theta\cos\theta - 2\left(\frac{d\theta}{d\varphi}\right)^2 \operatorname{tg}\theta - \frac{a^2}{r^2}\left(\frac{ds}{d\varphi}\right)\left[\frac{1}{r}\left(\frac{dr}{d\varphi}\right)\left(\frac{d\theta}{d\varphi}\right) - \right.$$

$$\left. - 2\sin\theta\cos\theta - 2\left(\frac{d\theta}{d\varphi}\right)^2 \operatorname{tg}\theta\right]. \qquad (13)$$

Having substituted into (12) and (13) the following expression familiar to us from differential geometry

$$\frac{ds}{d\varphi} = r\left[\cos^2\theta + \frac{1}{r^2}\left(\frac{dr}{d\varphi}\right)^2 + \left(\frac{d\theta}{d\varphi}\right)^2\right]^{1/2}$$

and making the transition in (12) and (13) to phase coordinates

$$w = \frac{r}{a}, \quad u = \frac{dr}{d\varphi}\frac{1}{a}, \quad \theta, \quad \psi = \frac{d\theta}{d\varphi},$$

we will obtain, respectively,

$$\frac{du}{dw} = \frac{w}{u}\psi^2 + \frac{w}{u}\cos^2\theta + 2\frac{u}{w} - 2u\psi\operatorname{tg}\theta \mp \frac{1}{wu}\left[\cos^2\theta + \left(\frac{u}{w}\right)^2 - \right.$$

$$\left. - 2\frac{u}{w}\operatorname{tg}\theta\psi\right]\left[\cos^2\theta + \left(\frac{u}{w}\right)^2 + \psi^2\right]^{1/2}. \qquad (14)$$

$$\frac{d\psi}{dw} = -\frac{\sin\theta\cos\theta + 2\psi^2\operatorname{tg}\theta}{u} + \frac{1}{uw^2}\left[\frac{u}{w}\psi - 2\sin\theta\cos\theta - \right.$$

$$\left. - 2\psi^2\operatorname{tg}\theta\right]\left[\cos^2\theta + \left(\frac{u}{w}\right)^2 + \psi^2\right]^{1/2}. \qquad (15)$$

The minus sign in front of the square brackets in (14) corresponds to trajectory segments with positive curvature, i.e., convexity averted from the coordinate origin. The plus sign corresponds to trajectory segments with negative curvature.

Moreover,

$$\frac{d\theta}{dw} = \frac{d\theta}{d\varphi}\frac{d\varphi}{dw} = \frac{\psi}{u}. \qquad (16)$$

The system of differential equations (14)-(16) describes the phase trajectories of a charged particle in four-dimensional phase space.

"THE CRITICAL TRAJECTORIES" OF CHARGED-PARTICLE MOTION

As was demonstrated in [1, 2] and in accordance with the familiar statements of the theory of oscillations [5], the "critical [particle]

trajectories" separating the captured and uncaptured in the case of two-dimensional motion (in the equatorial plane of the dipole) are defined by a singular point of the phase plane and the limiting curves leading to it. An unstable circular trajectory corresponds to the singular point in physical space, while spiral trajectories asymptotically winding onto and off the circular trajectory correspond to the limiting curves [separatrices]. In the case under consideration, where the motion takes place outside the equatorial plane and is described by phase trajectories in four-dimensional space (w, u, θ, ψ), the pattern of the critical trajectories retains the same qualitative indicators, i.e., there exist unstable trajectories surrounding the magnetic dipole, and there are the asymptotically converging spiral trajectories. The difference from the two-dimensional case consists only in that the trajectories oscillate in the direction of the meridional angle θ, rising above the equatorial plane, intersecting it, descending beneath the plane, etc. In phase space this form of motion corresponds to:

a) the trajectories surrounding the dipole (circular type), i.e., the limit cycles forming the two-dimensional set (surface);

b) spiral-shaped trajectories, screwlike phase curves asymptotically approaching the limit cycles.

The phase curves corresponding to the asymptotic spiral trajectories form a dense three-dimensional set which, in four-dimensional phase space, separates the regions corresponding to the various trajectory types. This three-dimensional set can therefore carry the designation of "three-dimensional separatrix hypersurface." Let us now prove the above-discussed statements.

Let us consider the projections of the trajectories in the four-dimensional phase space onto the plane (w, u) for rather low values of the variables θ and ψ. From Eq. (14), neglecting the infinitesimals of

higher order in θ and ψ, we obtain

$$\frac{du}{dw} = \frac{w}{u} + 2\frac{u}{w} \mp \frac{[1+(u/w)^2]^{1/2}}{uw}. \qquad (17)$$

However, this same equation defines the phase trajectories corresponding to the two-dimensional motion of particles (in the equatorial plane of the dipole), which was considered in [1, 2]. Consequently, the projections of the three-dimensional separatrix hypersurfaces onto the plane (w, u) in a region of rather low θ and ψ coincide with the corresponding separatrices in two-dimensional motion. This provides a basis for the contention that in the three-dimensional case there also exist trajectories of the "enveloping," "loop uncaptured," and "loop captured" types, analogous to the two-dimensional cases. But because the spatial phase trajectories, in addition to the coordinates $\underline{w}$ and $\underline{u}$, are also functions of the coordinates θ and ψ, the charged-particle trajectories in three-dimensional space will differ from the corresponding types of two-dimensional trajectories in that the motion along the coordinate θ is also superposed on the enveloping or loop motion.

The equation for the projections of the trajectories in four-dimensional space onto the plane (θ, ψ) can be derived from (15) and (16):

$$\frac{d\psi}{d\theta} = 2\psi \operatorname{tg}\theta - \frac{\sin\theta\cos\theta}{\psi} + \frac{1}{w^2\psi}\left[\frac{u}{w}\psi - 2\sin\theta\cos\theta - 2\psi^2\operatorname{tg}\theta\right] \times$$
$$\times\left[\cos^2\theta + \left(\frac{u}{w}\right)^2 + \psi^2\right]^{1/2}. \qquad (18)$$

where w = const, u = const.

At the singular point w = 1, u = 0 given sufficiently small values of θ and ψ, having neglected the infinitesimals of higher orders in θ and ψ, we can present Eq. (18) in the form

$$\frac{d\psi}{d\theta} = -\frac{3\theta}{\psi}. \qquad (19)$$

Having integrated (19), we obtain

$$3\theta^2 + \psi^2 = C^2. \qquad (20)$$

where C is an arbitrary constant of integration. Equation (20) describes phase trajectories of the limit-cycle type. Continuously changing the value of C in (20), we can easily see that the limit cycles form a continuous two-dimensional surface which in the region of rather low values θ and ψ coincides with the plane (θ, ψ). With finite θ and ψ the surface of the limit cycles becomes, general speaking, curvilinear, but has no discontinuities. The continuity follows from the fact that the system of differential equations (14)-(16), describing the particle phase trajectories in four-dimensional phase space, has no singular points when $\theta \neq 0$ and $\psi \neq 0$.

As a matter of fact, the singular points of System (14)-(16) are defined, as usual, by the fact [6] that at these points the derivatives du/dw, $d\psi/dw$, and $d\theta/dw$ will simultaneously change into a indeterminacy of the 0/0 type.

From (14)-(16) we see that a necessary (but insufficient) condition for this will be the condition u = 0 and ψ = 0.

If u → 0 and u/w → 0, ψ → 0 and tan θ is finite, i.e., $\theta \neq \pi/2$, from (14) and (15) we obtain, respectively,

$$\frac{du}{dw} = \frac{\cos^2\theta}{u}\left(w \mp \frac{\cos\theta}{w}\right). \tag{21}$$

$$\frac{d\psi}{dw} = -\frac{\sin\theta\cos\theta}{u}\left(1 + \frac{2\cos\theta}{w^2}\right). \tag{22}$$

A sufficient condition for the existence of the singular points of the system of equations (14)-(16), other than the requirement that u = 0 and ψ = 0, is the condition of the simultaneous equality to zero on the part of the numerators in the right-hand parts of Eqs. (21) and (22). For this let us initially consider the system of equations

$$\cos^2\theta\left(w+\frac{\cos\theta}{w}\right)=0,$$

$$\sin\theta\cos\theta\left(1+\frac{2\cos\theta}{w^2}\right)=0 \qquad (23)$$

for the case K < 0.

This system of equations has no real positive solutions and, consequently, with K < 0 the system of equations (14)-(16) has no singular point. The solution $\theta = \pi/2$ when $\psi = 0$ corresponds to the motion of a particle along a dipole axis. This case of motion is of no interest for the present problem and is not considered in this work.

Let us now examine the system of equations (21) and (22) for the case K > 0. In this case we have

$$\cos^2\theta\left(w-\frac{\cos\theta}{w}\right)=0,$$

$$\sin\theta\cos\theta\left(1+\frac{2\cos\theta}{w^2}\right)=0. \qquad (24)$$

The only real solutions for System (24) will be $\theta = 0$ and $w = 1$. Consequently, the only singular point of the system of differential equations (14)-(16) is the point ($w = 1$, $u = 0$, $\theta = 0$, $\psi = 0$).

It follows from the theorem of phase-trajectory set density that the surface of limit cycles must correspond to a dense three-dimensional set of phase trajectories asymptotically approaching that surface [the surface of limit cycles]. This dense three-dimensional set also forms three-dimensional separatrix hypersurfaces.

Consequently, the mapping point characterizing the motion of a charged particle along a trajectory close to the critical, due to the loss of energy on the part of the particle as a result of emission, may intersect the three-dimensional separatrix hypersurfaces and transfer from an "uncaptured" trajectory to a "captured" trajectory. The separatrix intersection process is similar to the two-dimensional case described in [1, 2], i.e., a charged particle losing its kinetic energy as a result of radiation initially transfers from the enveloping trajec-

tory to an "uncaptured" loop trajectory, and then to a loop "captured" trajectory.

Received
14 March 1964

REFERENCES

1. V.M. Vakhnin and G.A. Skuridin, Dokl. AN SSSR [Proc. Acad. Sci. USSR], 135, 1354, 1960.
2. V.M. Vakhnin, G.A. Skuridin and I.N. Shvachunov, Kosmich. issled. [Cosmic Research], 1, No. 3, 414, 1963.
3. I.Ye. Tamm, Osnovy teorii elektrichestva [Fundamentals of the Theory of Electricity], Gostekhteoretizdat [State Publishing House of Theoretical and Technical Literature], Moscow, 1954.
4. G.M. Fikhtengol'ts, Kurs differentsial'nogo i integral'nogo ischisleniya [A Course in Differential and Integral Calculus], Vol. 1, Gostekhteoretizdat, 1951.
5. A.A. Andronov, A.A. Vitt and S.E. Khaykin, Teoriya kolebaniy [The Theory of Oscillations], Fizmatgiz [Physics-Mathematics Press], 1959.
6. V.V. Nemytskiy and V.V. Stepanov, Kachestvennaya teoriya differentsial'nykh uravneniy [A Qualitative Theory of Differential Equations], Gostekhteoretizdat, 1947.

Manuscript Page No.

[Footnote]

191 *The results of this project were reported at the 4th All-Union Conference on Magnetohydrodynamics (Riga, June, 1964).

UDK 539.121.64:539.165

SPECTRUM OF VERY HIGH-ENERGY ELECTRONS ARISING IN THE β-DECAY OF THE ALBEDO NEUTRONS

A.I. Yershkovich

The equilibrium integral spectrum of β-electrons in the case of energies of the order of several Mev is calculated. It is demonstrated that this spectrum agrees with the experimental results obtained in the area of the outer radiation zone.

Electrons exhibiting an energy E of the order of several Mev were detected in the outer zone of radiation during the flights of the first and second soviet spacecraft [1, 2]. It might be assumed that these electrons originate in the β-decay of albedo neutrons exhibiting energies of hundreds and thousands of Mev [3, 4]. Nakada [3] constructed a differential spectrum of β-electrons in the assumption that the spectrum of fast neutrons at a geocentric distance $\underline{r}$ measured in earth radii is equal to $2E_n^{-1.8}/r^2$ and E_n^{-2}/r^2 $cm^{-2} \cdot sec^{-1} \cdot Mev^{-1}$, where E_n is the kinetic energy of the neutrons in Mev. As a matter of fact, the albedo neutron spectrum far from the earth diminishes with increasing E_n considerably faster, since with increasing E_n there is a reduction in the solid angle from which the neutrons emanate to the point under consideration. Since this effect was not taken into consideration,* the flow of β-electrons with energies $E \geq 5$ Mev, obtained in [3], proved to be unjustifiably large (see [2]).

The Lenchek and Singer [5] calculations show that the directed differential intensity of albedo neutrons with energies of the order of several hundreds of Mev, leaving the atmosphere of the earth, with res-

pect to time, on the average, are equal to*

$$j_n(E_n) = 1{,}5\,E_n^{-1.8} \quad [\text{cm}^{-2}\cdot\text{sec}^{-1}\cdot\text{steradian}^{-1}\cdot\text{Mev}^{-1}]. \quad (1)$$

With extremely high neutron energies $j_n(E_n)$ must diminish more rapidly than $\sim E_n^{-1.8}$. Indeed, the energy spectrum of the secondary particles above ~500 Mev must repeat the spectrum of the primary cosmic rays [7]. We know that in the hardness interval $R = Pc/Ze$ from 1 to 10 Bev (the corresponding interval of proton kinetic energy ranges from 0.4 to 9 Bev) the differential hardness spectrum is proportional to $R^{-2.5}$ [8]. Hence it is not difficult to find that

$$j_n(E_n) = 6{,}25\cdot 10^2 \frac{E_n + Mc^2}{[E_n(E_n + 2Mc^2)]^{1.75}} [\text{cm}^{-2}\cdot\text{sec}^{-1}\cdot\text{steradian}^{-1}\cdot\text{Mev}^{-1}], \quad (2)$$

where Mc^2 is the energy of the resting neutron, equal to 940 Mev. In the regions of energies E_n of the order of several hundreds of Mev, the neutron intensity defined by Relationship (2), $j_n(E_n) \sim E_n^{-1.8}$; with $E_n \gg Mc^2$ $j_n \sim E_n^{-2.5}$. The proportionality factor in (2) is defined from the condition that Expressions (1) and (2) sew together with $E_n = 400$ Mev.

Let us define the intensity of the albedo neutrons at the point S situated at a distance $\underline{r}$ from the center of the earth. Let θ be the angle between the directions of the primary and secondary particles. In this case $\sin\theta = P_\perp/P$, where $P_\perp$ and P are, respectively, the lateral and total impulses. We know that the lateral impulse $P_\perp$ of the secondary particles changes but slightly over a wide interval of primary-particle energies [9]. The average value of $P_\perp \approx 3\cdot 10^2$ Mev/c [10, 11]. The least value of the zenith angle ξ at which fast neutrons with impulse P escape from the atmosphere is obviously equal to $\xi_{min} = \pi/2 - \theta$. Thus,

$$\cos\xi_{min} = (3\cdot 10^2)/Pc, \quad (3)$$

where Pc is measured in Mev. Let us assume that neutrons with impulse P

escape from the atmosphere isotropically with $\xi_{min} \leq \xi \leq \pi/2$. Since $j_n \approx$ const along the neutron trajectory, the differential intensity at the point S is equal to

$$I_n(E_n, r) = j_n(E_n)\Omega(E_n, r)[\text{cm}^{-2}\cdot\text{sec}^{-1}\cdot\text{Mev}^{-1}], \qquad (4)$$

where $\Omega(E_n, r)$ is the solid angle from which the neutrons with energies of E_n come to point S. In a polar system of coordinates with its center at point S and an axis passing through the center of the earth, $d\Omega = 2\pi \sin\alpha d\alpha$, where α is the polar angle. Since $\sin\xi = r\sin\alpha$,

$$\Omega = \frac{2\pi}{r^2}\int_{\xi_{min}}^{\pi/2} \frac{\sin\xi\cos\xi\, d\xi}{[1-(\sin^2\xi/r^2)]^{1/2}} = 2\pi\left[\left(1-\frac{\sin^2\xi_{min}}{r^2}\right)^{1/2} - \left(1-\frac{1}{r^2}\right)^{1/2}\right]. \qquad (5)$$

With $r^2 \gg 1$, having used (3), we find

$$\Omega = (\pi/r^2)\cos^2\xi_{min} = [(2{,}8\cdot 10^3)/r^2](Pc)^{-2}. \qquad (6)$$

According to (2), (4), and (6), with $r^2 \gg 1$ and $E_n \gg Mc^2$

$$I_n(E_n, r) \sim E_n^{-4.3}/r^2.$$

The integral intensity of β-decay electrons with a kinetic energy $E \geq E_1$ at the equator can be evaluated on the basis of the formula

$$I(E \geq E_1, r) = nvT\,[\text{cm}^{-2}\cdot\text{sec}^{-1}], \qquad (7)$$

where $\underline{n}$ is the number of electrons exhibiting the energy $E \geq E_1$, produced per unit volume per unit time, and $\underline{v}$ is the velocity of these electrons, T representing the lifespan of the electrons whose mirror points are situated close to the plane of the equator. Obviously,

$$n = \frac{1}{t_n}\int_{E_{n\,min}(E_1)}^{\infty} \frac{I_n(E_n, r)}{v_n\gamma_n} W(E_n, E_1)\,dE_n\,[\text{cm}^{-3}\cdot\text{sec}^{-1}], \qquad (8)$$

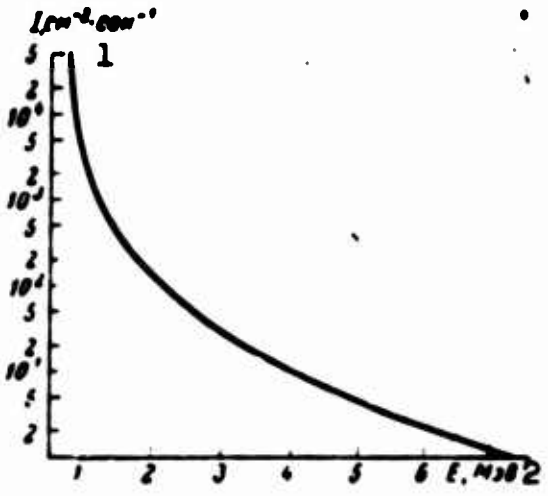
where $t_n \approx 1100$ sec is the average lifespan of the neutron, v_n is its velocity, and $\gamma_n = (1-v_n^2/c^2)^{-1/2}$. $\underline{W}(E_n, E_1)$ is the probability that on the decay of a neutron having an energy E_n there will be produced an electron having a kinetic energy $E \geq E_1$ in a laboratory system (the function $W(E_n, E_1)$ was defined earlier [4]), $E_{n\,min}(E_1)$ is the minimum neutron

energy beginning from which W $\neq$ 0. According to [4]

$$E_{n\min}(E_1) = Mc^2(2{,}53E_0 - (2{,}53^2 - 1)^{1/2}(E_0^2 - 1)^{1/2}] - Mc^2, \qquad (9)$$

where

$$E_0 = E_1 / mc^2 + 1.$$



1) I, in $cm^{-2} \cdot sec^{-1}$; 2) E, in Mev.

It was assumed that the primary process determining the lifespan T of the electrons is the loss of energy resulting from Coulomb interactions with particles of the exosphere. If N is the concentration of cold plasma in the exosphere, the ionization losses of the ultrarelativistic electrons ($E \gg mc^2$) are equal to [6]

$$-(dE/dt)_i = 7{,}62 \cdot 10^{-9} N \cdot \{\ln E / mc^2 - \ln N + 73{,}4\} [ev \cdot sec^{-1}]. \qquad (10)$$

It follows from Relationships (4), (6)-(8), (10) that the equatorial value of the flow of captured β-decay electrons when $r^2 \gg 1$ is proportional to $1/r^2N(r)$.

The figure shows the integral spectrum of β-electrons in the region of the outer radiation-zone maximum (r = 4), calculated in accordance with (7). The value of N is set equal to $1.5 \cdot 10^2$ cm^{-3} [12]. With $E \gtrsim 1.5$ Mev this spectrum is approximately proportional to $E^{-3.7}$, which corresponds to the results of the measurements that were carried out aboard the "Explorer XII" satellite (the differential spectrum $\sim E^{-5}$) [13].* The flow ratio $I(E \gtrsim 1.6 \text{ Mev})/I(E \gtrsim 5 \text{ Mev})$ in the outer zone amounts approximately to 10^2 with a change in the geocentric distance r from 15000 to 55000 km ([13], Table 3). This value is close to the

one calculated (see Figure). On the basis of the estimate derived from the second spacecraft [2], at the maximum of the outer zone $I(E \geq 5$ Mev) < 1 cm$^{-2}\cdot$sec^{-1}. If we take into consideration the known indeterminacy of the spectrum of fast albedo neutrons and the concentration N(r), we must acknowledge that this estimate satisfactorily agrees with the calculation results. The experimental data [2, 13], apparently, favor the hypothesis that electrons of extremely high energy appear in the terrestrial radiation zones on the β-decay of the albedo neutrons. The effect of the betatron mechanism may lead to significant variations in the intensity of these electrons, without altering the form of their energy spectrum.

Received

21 February 1964

REFERENCES

1. S.N. Vernov, A.Ye. Chudakov, P.V. Vakulov, Yu.I. Logachev, Dokl. AN SSSR, [Proc. Acad. Sci. USSR], 125, No. 2, 304, 1959.
2. S.N. Vernov, A.Ye. Chudakov, P.V. Vakulov, Yu.I. Logachev and A.G. Nikolayev, Sb. "Iskusstvennyye sputniki Zemli," [Collection entitled "Artificial Satellites of the Earth,"], vyp. 5. Izd-vo AN SSSR [Issue 5. Acad. Sci. USSR Press] 1960, page 24.
3. M.P. Nakada. J. Geophys. Res., 68, No. 1, 47, 1963.
4. A.I. Yershkovich, Izv. AN SSSR, ser. geofiz., [Bulletin of the USSR Academy of Sciences, Geophysics Series], No. 9, 1457, 1963.
5. A.M. Lenchek and S.F. Singer, Planet. Space Sci., 11, No. 10, 1151, 1963.
6. V.L. Ginsburg, S.I. Syrovatskiy, Proiskhozhdeniye kosmicheskikh luchey [Origin of Cosmic Rays], Izd-vo AN SSSR, 1963.
7. G. Messel, Collection entitled, "Fizika kosmicheskikh luchey,"

["The Physics of Cosmic Rays"], 2, pod. red. Dzh. Vil'sona. IL, M., [2, edited by Dzh. Wil'son [sic], Foreign Literature Press, Moscow,], 1956, page 106.

8. E.P. Ney, Sb. "Solnechnaya aktivnost' i kosmicheskiye luchi," [Collection entitled "Solar Activity and Cosmic Rays], per. pod ped. V. P. Shabanskogo, IL. M., [Translation edited by V.P. Shabanskiy, Foreign Literature Press, Moscow,] 1962, page 21.

9. D.H. Perkins, Progress in elementary particle and cosmic ray physics, ed. J.G. Wilson and S.A. Wouthuysen, Amsterdam, 5, 259, 1960.

10. V.B. Fretter, L.F. Khansen, Tr. Mezhdunar. konf. po kosmicheskim lucham [Transactions of the International Conference on Cosmic Rays], 1, Izd-vo AN SSSR, M., 1960, page 134.

11. R.R. Daniel, N. Kameswara Rao, P.K. Malhotra and Y. Tsuzuki, Nuovo cimento [10], 16, No. 1, 1, 1960.

12. A.M. Lenchek and S.F. Singer. J. Geophys. Res., 67, No. 4, 1263, 1962.

13. W.G. V. Rosser, B.J. O'Brien, J.A. Van Allen, L.A. Frank and C.D. Laughlin, J. Geophys. Res., 67, No. 12, 4533, 1962.

14. I.S. Shklovskiy, V.I. Krasovskiy, Yu.I. Gal'perin, Izv. AN SSSR, ser. geofiz., No. 12, 1799, 1959.

Manuscript Page No.

[Footnotes]

200 B.A. Tverskoy drew our attention to the possible significance of this effect.

201 Reference [5] shows the derivation $j_n = 0.5\ E_n^{-1.8}$ cm^{-2}·sec^{-1}·steradian^{-1}·Mev^{-1}. The factor 2 was dropped in the derivation of this expression. Moreover, the authors [5] assumed a cosmic-ray intensity at the earth of $J_0 = 0.1$ cm^{-2}·sec^{-1}·steradian^{-1}, which corresponds to the maximum of solar ac-

tivity. Since J_0 close to the activity minimum increases by a factor of two (see, for example [6]), it is expedient to assume an average value for J_0 equal to 0.15 $cm^{-2} \cdot sec^{-1} \cdot steradian^{-1}$. As a result, we obtain Expression (1).

203 It must be borne in mind that a certain fraction of the captured electrons exhibiting energies of several Mev could be a result of the high-altitude nuclear explosions carried out since the summer of 1958 [14].

UDK 629.198.61

PROBLEMS OF ENGINEERING PSYCHOLOGY IN COSMONAUTICS AND CERTAIN INVESTIGATIONAL RESULTS

V.G. Denisov, Ye.S. Zav'yalov, A.P. Kuz'minov, M.M. Sil'vestrov, V.D. Yazdovskiy

The article presents a definition of the subject of engineering psychology and its problems in the development of spacecraft control systems and the training of cosmonauts [astronauts] in their professional activities. A complex method is proposed for the evaluation of the closed "operator-vehicle" system in which, in conjunction with the application of the methods of cybernetics and information theory, providing engineering estimates for the technical elements of the system, extensive use is also made of physiological recording of the biopotentials of various human functional systems characterizing the level of tension and ability of the operator to function. There is a discussion of some of the results obtained in the effort to develop efficient systems of indication, signaling, and manual control. Certain of the problems involved in the development of training facilities are touched upon. The use of a system of training facilities consisting of complex, specialized, and functional training facilities is recommended for the solution of problems relating to the training and maintenance of skills on the parts of the cosmonauts. The development of on-board training facilities is recommended for the maintenance of skills in prolonged space flights.

Engineering psychology is a new trend in science, arising out of the need for a scientific approach to technical control installations and the coordination of their characteristics with the psycho-physiological potentials of the operator in a single closed cybernetic "man-machine" system.

Engineering psychology and research in connection with the latter assume particular significance in cosmonautics because a piloted space vehicle together with its equipment represents a complex multicircuited control system, while the conditions of flight aboard the vehicle are quite specific and frequently unusual. Thus the cosmonaut will encounter difficulties in controlling the systems of the vehicle, particularly in the event that the technical facilities of these systems are developed without consideration of the cosmonaut's potentials for control.

At least three basic problems of engineering psychology can be formulated in connection with cosmonautics, i.e., the investigation of the psycho-physiological potentials of the operator to control a space vehicle and its systems, and the change in these potentials under the influence of space-flight conditions; the substantiation of requirements imposed on vehicle control systems in order to coordinate their characteristics with the psycho-physiological potentials of the operator; and the development of methods and facilities to train cosmonauts for their professional activity, these methods and facilities to provide for the acquisition of skills and habits on the part of the operators to correspond to the characteristics of the control systems.

The action of the operator, who represents the controlling link in the closed "cosmonaut-vehicle" systems may be characterized by the following three stages: the acquisition and sensing of information from various instruments, signaling devices, and the external medium; the evaluation of the acquired information and the working out of control responses for vehicle control, i.e., evaluation of the sensed parameters of the control process, the ambient medium (surrounding conditions), disturbing influences, generalization, interpretation, and comparison of these [parameters] with the parameters of the given control process and the taking of decisions with regard to the nature, magnitude, and

purpose of the control response; the application of control responses to vehicle control units or the issuance of some other control commands

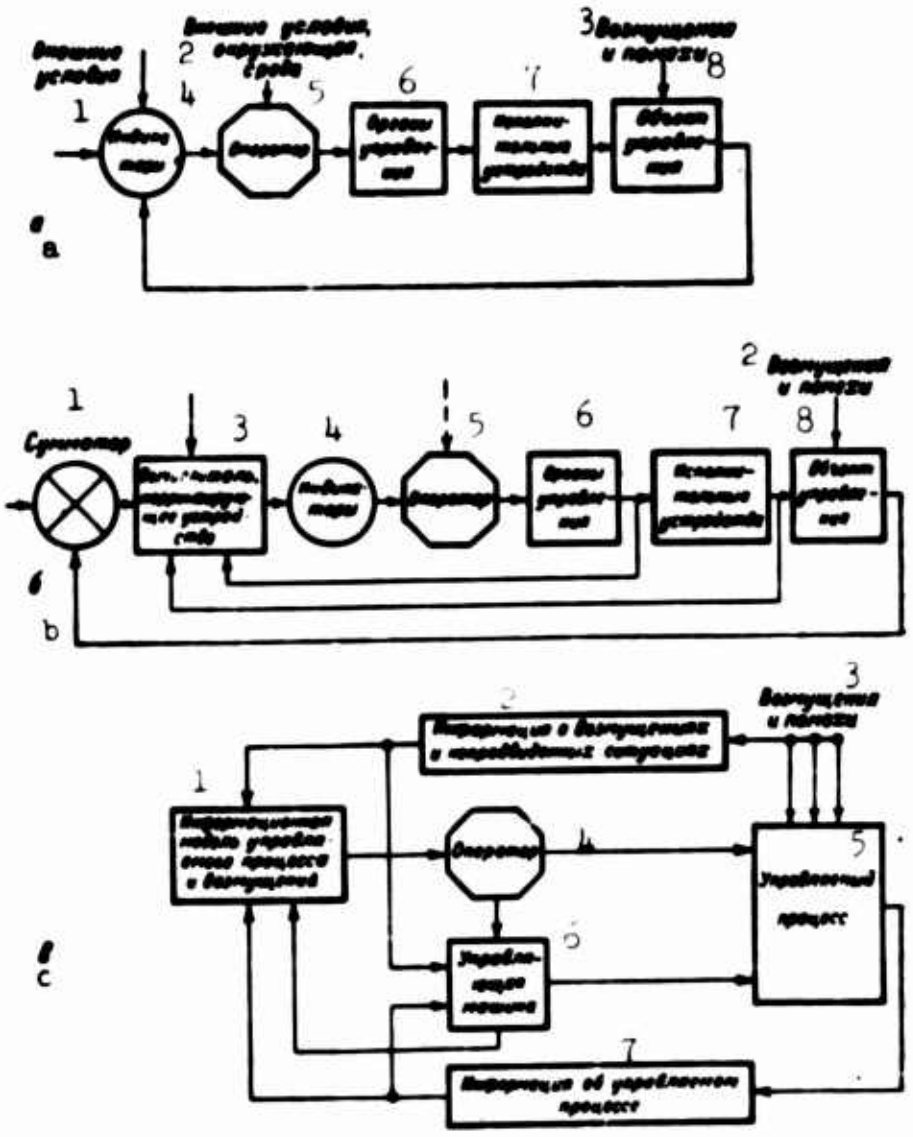


Fig. 1. Typical control systems aboard a space vehicle, requiring the participation of an operator. a) Manual control system; b) semi-automatic (directorial) control system; c) training control system with participation of operator; a1) External conditions; a2) external conditions, ambient medium; a3) disturbances and interference; a4) indicators; a5) operator; a6) control unit; a7) servomechanisms; a8) controlled member; b1) adder; b2) disturbances and interference; b3) computer, correction device; b4) indicators; b5)operator; b6) control unit; b7) servomechanisms; b8) controlled member; c1) information model of control process and disturbances; c2) information on disturbances and unanticipated situations; c3) disturbances and interference; c4) operator; c5) controlled process; c6) servomechanism; c7) information on controlled process.

Typical diagrams of control systems involving the participation of an operator aboard a space vehicle are shown in Fig. 1.

The systems of vehicle orientation and descent, the system for

transfer from orbit to orbit, and other control systems can be reduced to these diagrams. To ensure the appropriate accuracy and efficiency of the manual-control process the information systems must be adapted for rapid stimulation of the sense organs and the information-evaluation systems of the operator when evaluating the status of the controlled member and in carrying out control assignments. This can be ensured only through the issuance to the operator not only of quantitative but of qualitative information, which in totality determines the value of the information for the operator [1]. The value of the information can be defined as an increment to the probability of accomplishing the tasks for which the information is being collected.

If less stress is laid on information quality, the memory capacity of the operator must be increased and, consequently, longer operator training is required in order to achieve at least a certain minimum level of qualitative information required to elevate the stimulation of the sense organs and the information-evaluating systems of the operator as he controls the member.

It follows from the above that specifically human activity in work with any machine involves comprehension of signals from the machine and in assignment to these signals of a certain significance in order to interpret the conceptual content of the incoming information and on this basis to select the required action.

The emergency warning device which functions on the binary principle and, for example, indicates a fire in the element, the probability of the occurrence of such a fire being given by $P_1 = 1/1000$, provides the operator with continuous reports as to the condition of the element, the average value of information quantity being determined by the expression

$$I_{cp} = -P_1 \log_2 P_1 - P_2 \log_2 P_2,$$

where $P_2 = 1 - P_1$ is the probability of their being no fire in the element. Having substituted the values of P_1 and P_2, we obtain $I_{sr} = 0.01$ binary units. The average value of the information quantity characterizing entropy of the system and the measure of the indeterminacy of its condition is small, since the usual natural state of the element is characterized by an absence of fire. The low indeterminacy of the state of the system (the absence of fire) is a virtually indisputable fact. However, the continuous information being provided by this warning device is extremely important for the human being. This flow of information is not evaluated in the human mind in the same manner as is information used in engineering.

It has been established that man during a process of activity uniquely accounts for the stress of a situation and in order to avoid functional excesses "cleanses" the incoming information of secondary elements.

Proceeding from the above, in our opinion the quantity of information used by an operator during the control process cannot be determined exclusively by the conventional probability method used in engineering.

There is a need to introduce certain coefficients characterizing the value of the incoming information as well as the ability of the operator to perceive this information into the expression for the determination of the quantity of information flowing into the operator.

The average quantity of information in the transmission to and receipt by the operator of information may be expressed by the formula

$$I = -\sum_{j=1}^{m} A_j P_j(x) \log P_j(x) - \sum_{j=1}^{m} P_j(y) \left[\sum_{i=1}^{n} B_i P_{yi}(x) \log P_{yi}(x) \right],$$

where $\underline{x}$ is the input quantity which is being reported, $\underline{y}$ is the received quantity, $P_j(x)$ is the probability of the quantity $\underline{x}$ assuming

the value of y prior to the arrival of the information, $P_j(y)$ is the probability that the received quantity y will acquire the value y_j if the incoming information is not certain, $P_{yi}(x)$ is the probability of the input quantity after the receipt of the information, A_j is the coefficient expressing the value of the information on the quantity x_1, and B_1 is the coefficient expressing the level of perception of the information on x_1.

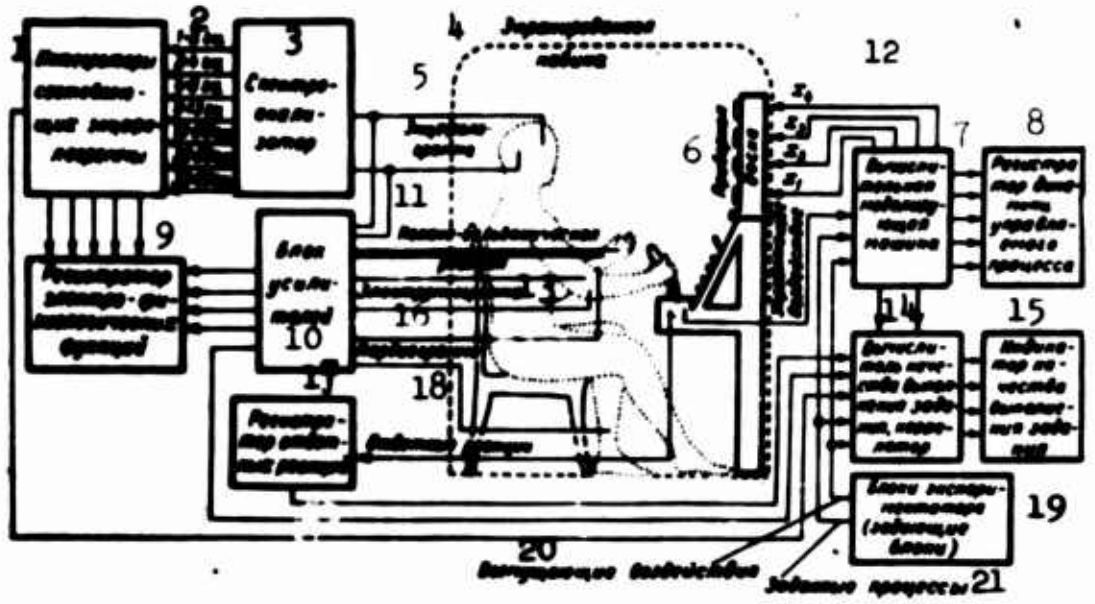


Fig. 2. Diagram showing the modeling of control processes with integrated recording of control parameters and electrophysiological indicators for the operator. 1) Encephalogram component integrators; 2) cps; 3) spectral analyzer; 4) shielded cabin; 5) encephalogram; 6) instrument panel; 7) computer-simulator; 8) control-process dynamics recorder; 9) electrophysiological function recorder; 10) amplifier block; 11) dermogalvanic reaction; 12) control action; 13) electromyogram; 14) correlator, computer to determine quality of assignment execution; 15) indicator showing quality of assignment execution; 16) cardiogram; 17) response-reaction recorder; 18) response reactions; 19) experimentor units (assignment unit); 20) disturbing actions; 21) assigned processes.

The average quantity of information in transmission to the operator and his receipt of said information is equal to the difference between the entropy of input-quantity distribution (with consideration of the value of this information) and the averaged entropy of probability distribution for this quantity after receipt of the signal (with consideration of the level of perception for the incoming information).

The magnitude of the value coefficient A_j and the perception coef-

ficient B_1 is determined by a study of the informational processes, with consideration of the selected sequence of information input, its value, the efficiency of the indicators and signaling devices, and the distribution of information among the units and analyzer systems.

A study of the interaction between the operator and the machine in a closed system of space vehicle control is carried out primarily by experiment. Depending on the problems that have been posed, uniquely cybernetic methods are employed (the simulation of a closed control system, the application of information theory), as well as physiological, hygienic, and experimental-psychological methods, i.e., a wide range of methods is used under experimental conditions to permit study of the functional potentials and limitations of a human being in his interactions with the machine [vehicle] during the control process.

A method of simulating the control system of a space vehicle was employed in the present article which deals with the engineering-psychological investigations associated with space flight. Modeling made possible a study of the time characteristics involved in the work of a cosmonaut with instrument equipment and control levers; these characteristics are extremely important for the efficient design of the control systems of future space vehicles and to resolve problems associated with the proper assembly of instrumentation and the problems dealing with efficient indication and warning systems. Modeling [simulation] also made possible the studying and testing of individual elements included in the systems of indication, warning [signaling], and manual control under dynamic operating conditions, within the structure of that activity which the cosmonaut carries out under actual conditions [2].

Figure 2 shows the general view of a control-process simulation diagram with integrated recording of control-process and operator elec-

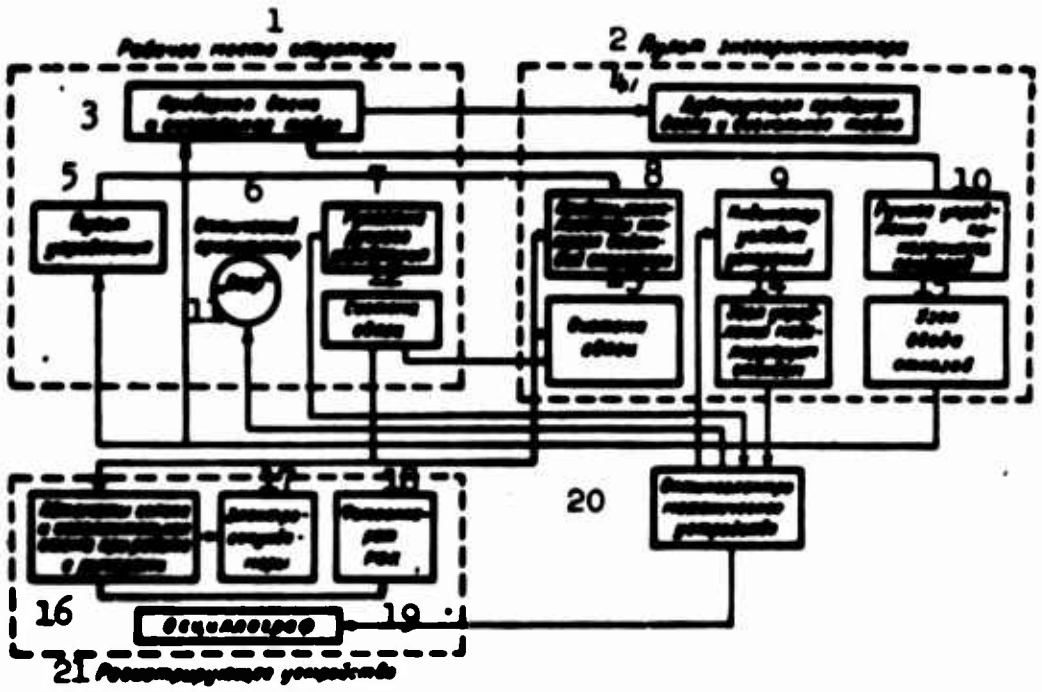


Fig. 3. Diagram of modeling installation. 1) Operator's work position; 2) experimentor console; 3) instrument panel and signal board; 4) back-up instrument panel and signaling board; 5) control panel; 6) optical orientator; 7) manual-control lever; 8) instruments and warning devices to monitor operator actions; 9) angular-acceleration indicator; 10) manual control of instrument readings; 11) view; 12) communications system; 13) communications system;14) control unit for modeling installation; 15) nonoperation input unit; 16) automatic voice units and final control scheme for lever operation; 17) electronic timer; 18) RFK camera; 19) oscillograph; 20) optical-electrical mechanical device; 21) recording device.

trophysiological-indicator parameters. We can see from the diagram that the control process is modeled by means of a computer simulator from which the signals characterizing the condition of the controlled process are transmitted to the signaling and indication system. The operator receives the information regarding the state of the controlled process, works out control signals, and at the same time alters the controlled process and its coordinates.

A recording device is connected to the modeling unit for purposes of recording the dynamics of control. The motor and speech response reactions are recorded by means of the response-reaction recorder. The circuit [diagram] includes a device for the recording of biological currents from various sections of the functional systems of the human operator. For purposes of analyzing the encephalogram, a spectrum analy-

zer is used to isolate from the general encephalogram curves exhibiting various frequencies (rhythms). Integrators provide for the determination of the average level of each frequency component of the encephalogram for the selected interval of time. The diagram also shows a quality-control computer which is used to determine the transient responses on the basis of integral limits.

In investigating the psychophysiological potentials of the operator to control the vehicle"Vostok" we found that the most important elements of the modeling installation used in our work were: the operator's cabin, the console of the experimentor, the earth-and-sky apparent-image simulator, computers, and programming devices (Fig. 3). Use was made of a mock-up of the cabin of the "Vostok" vehicle whose equipment differed neither with respect to external appearance nor with respect to design from the corresponding equipment in the actual cabin. The dynamics of the angular shift occurring in vehicle control were simulated by means of electronic models (MNB-1 and IPT-5) and special units. The motion-picture film projection system provided for the simulation at one of the portholes of the relative shifting of the image of the earth and the stellar sky, as would be observed by a cosmonaut in an actual flight. The device simulating the manual control of orientation makes possible the introduction of changes in the angular positions of the vehicle by means of a control stick.

The operational elements of the automatic orientation system, the on-board programming installation, the control system, the vehicle communications system, and the operation of the instrumentation are simulated in the modeling [simulation] installation.

With the aid of the recording installation (Fig. 4), on the basis of the photographic records showing the readings of the electronic timer prior to the action of the stimulation and after the response of the

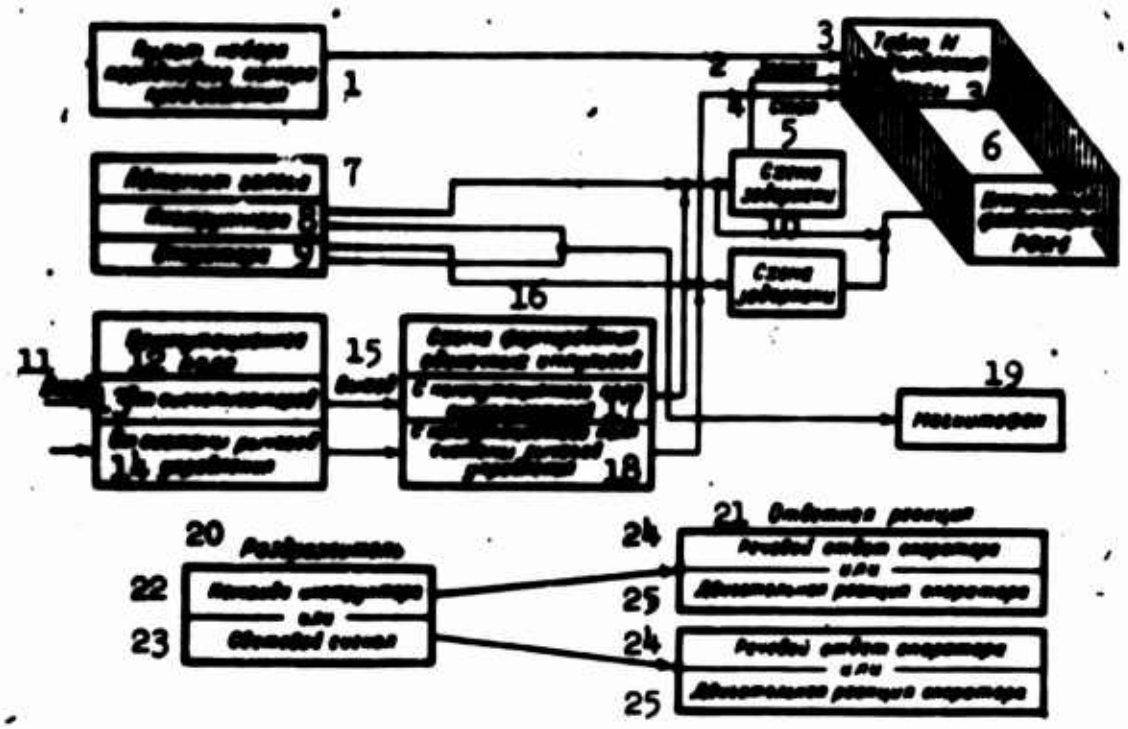


Fig. 4. Diagram of recording installation. 1) Panel for recording of sequential presentation number; 2) start; 3) N presentation table; 3a) clock; 4) stop; 5) delay circuit; 6) RFK-5 impulse camera; 7) automatic voice unit; 8) of instructor; 9) of operator; 10) delay circuit; 11) input; 12) switching field; 13) from warning devices; 14) from control-lever system; 15) output; 16) circuit for the formation of discrete pulses; 17) from warning-device switching field; 18) from control-lever system switching field; 19) magnetic-tape recorder; 20) stimulation; 21) response reaction; 22) order from instructor; 23) or light signal; 24) speech response from operator; 25) or motor reaction on the part of the operator.

operator, the readings of the instruments, the active signals and the positions of the control levers, as well as on the basis of the speech interrogation of the test subject as recorded on magnetic tape, provides a possibility for rendering judgement regarding the time and accuracy with which the motor and sensory actions of the operator have been carried out. The movement of the control handle is recorded on the oscillograph [3].

The investigation of the closed "human being-machine" system must be carried out not only on the basis of quantitative and qualitative characteristics of the control process, but also on the basis of the electrophysiological indicators of the operator [4]. The electrophysiological indicators of the operator recorded during the control process

make it possible to establish the difference between the systems which, on the basis of initial external indications, seem to be identical. Thus with identical output results for quality control from two systems of the same designation, the magnitude of the bioelectrical activity of high-frequency rhythms in an encephalogram recorded during the control process as the operator was working with a system that is more "difficult" for a human being increases in comparison with the corresponding bioelectrical activity of the encephalogram for ar operator working with an "easier" system. Thus, in order to evaluate the manual and semi-automatic control systems, we must use the indicators of the control-process dynamics (the transient functions, their nature, the time of the transient response, the integral limits of the transient responses), errors in steady induced regimes, the expenditure of energy on control, the quantity of incorrect actions on the part of the operator, the time characteristics of his operational activity (the latent period of response reactions, the time required for the completion of given operations), emotional stress, determined on the basis of operator reactions and the data of the integrated recording of electrophysiological indicators (the electroencephalogram, dermogalvanic reactions, from the electromyogram, the electrocardiogram, the respiration rate), as well as the formation of a habit.

There arises the problem of working out a generalized criterion for the evaluation of the various control systems, said criterion to account both for the indicators of control dynamics and the psychophysiological stress of the cosmonaut as he executes his control actions. The total entropy of the physiological indicators of the operator and the deviations of the controlled quantities from the given values can be used as such a generalized criterion, since some degree of indeterminacy is inherent in the closed "human being-machine" system, constant-

ly under the influence of disturbing factors.

For a quantitative evaluation of the systems in accordance with this generalized criterion, based on total entropy, a large quantity of information must be processed. Therefore, the application of this criterion calls for extensive utilization of digital computer techniques.

The development of space-vehicle control systems including an operator must be carried out with consideration of the psychophysiological potentials and functional limitations of the human being. The characteristics of the indication installations must be coordinates with the features in the operation of the perception apparatus, while the characteristics of the control units must be coordinates with the features involved in the motor apparatus of the human being.

The execution of the applied problems in space flight depends in great measure on the proper consideration of the functional potentials and capacities of the human being and the efficient [rational] distribution of functions among the members of the crew and the space-vehicle control systems, since a decisive factor limiting the potentials of the human being is the pronounced divergence between the rate of human reaction and the dynamic characteristics of an automatic system.

In the development of space-vehicle control systems including a human link, it is absolutely necessary to recall, on the one hand, the great plasticity of the nervous system, permitting adaptation to changing conditions of the medium, and on the other hand, to recall that there are limitations to individual psychophysiological human functions. For the sensory functions these limitations are defined by the thresholds of absolute and relative sensitivity, as well as by the maximum degree of attentiveness (the simultaneous perception of 5-7 isolated objects). It should be borne in mind in this case that subliminal unconscious signals affect the perception of information by the operator

[4-6].

Our work demonstrates that in various cases the time of information perception from a given instrument and its evaluation amounts to 0.28-0.8 second. The maximum reaction time, however, to signals appearing in rapid succession comes to about 0.25 second in the case of intervals of 0.57 to 2.91 second between signals, with the limit interval between signals within which an adequate reaction is possible coming to no more than 0.5 second. A number of experimental projects have been carried out in connection with the potentials of a cosmonaut to execute various types of activity during flight. Thus during the process of determining the potentials of visual control in the meeting of space vehicles, it was determined that the acuity of visual determination of angular magnitudes by the operator in calculating the speed and range of the approaching object came to 3.5', and the resolving power with respect to angular velocity in the case of an observation period of 10 seconds amounted to 0.35 angular minutes per second [7].

The efficiency with which a human being utilizes levers and tools under conditions in which there is no loss of energy and in which there is no friction changes radically in the case of weightlessness. This is borne out by experiments carried out with minimum friction (the friction coefficient 0.0015) for the support carrying the test subject. With insignificant friction it was virtually impossible to move a lever "toward" or "away" from oneself. Maximum pressure was required in order to shift the lever, simultaneously maintaining the test subject in a stable position. The effectiveness of the pressure applied by the operator in comparison with the effectiveness of the pressures applied by a human being working under conventional conditions changed sharply. Thus the effectiveness of the pressure was lower by a factor of 27 in the case of movement toward, and lower by a factor of 18 in the case of

movement away from oneself, and lower by a factor of 3 in the case of rotary movements. The magnitude of the pressure on compression and extension with both hands showed no significant change [8]. Under conditions of reduced gravity, the muscular strength of the hand is reduced in the execution of motor reactions. Thus at an initial level of 45-60 kg the strength of the hand in 87% of all cases is reduced by 4-16 kg, and there is an impairment in the indicators produced in a study of given muscular stresses, which could lead to an impairment in the precision with which movement is accomplished. In working with control levers designed for a force of 750g, the accuracy with which the operator's muscular exertions are determined can be significantly impaired. To protect the cosmonaut against unfavorable effects of weightlessness, we can employ artificial gravity. However, in setting up artificial gravity it is necessary to take into consideration the negative effects exerted on a human being by the Coriolis acceleration and the change in the magnitude of the artificial gravity.

To determine the possibility of controlling a spacecraft, we studied the retention and precision of movement in various joints at accelerations from 4 to 15 units [9]. It has been established that with accelerations greater than 6 units, mobility in the large joints of the arms and legs is impossible. The hand and fingers retain sufficient mobility at accelerations of 8 units. At accelerations below 15 units, the test subject may execute certain movements in the wrist and finger joints, but mobility in the joints is significantly reduced. The potentials of the operator with regard to control in the case of acceleration may be improved by means of regulated training procedures, the application of various facilities to elevate the ability of the organism to withstand the effect of acceleration, and these improvements can also be achieved through efficient design of the manual-control circuit.

Results obtained during the flights of our cosmonauts have shown that under space-flight conditions lasting up to 5 days the motor patterns and ability to function experience no significant change. Experimental investigations directed at a study of individual aspects of this problem made possible the obtaining of the data discussed below.

In developing the problems of engineering psychology in cosmonautics during the first flights of vehicles of the "Vostok" type it was necessary to take into consideration the following features of cosmonaut work.

1) Total automation of vehicle control processes. The cosmonaut was responsible only for functions of control in connection with the radiocommunications system and the monitoring of the other systems; in individual cases, he can execute manual control of vehicle orientation.

2) The specificity of orientation-control dynamics, consisting in the fact that the reaction of the space vehicle to control actions is comparatively slow.

The experimental study of operations carried out by a cosmonaut during work shows that during the normal flight of a vehicle of the "Vostok" type all actions of the cosmonaut involving the instrumentation are accomplished within the time available. As a matter of fact, there is even some "surplus" of time. In neither case does the time required to read off the instrument readings limit the execution of these operations. For purposes of determining geographic coordinates on the basis of materials provided by investigation no more than 7.7 seconds are required. To determine the numerical value of one of the parameters characterizing the state of the medium in the cabin (temperature, oxygen content, carbon-dioxide content, etc.) from 2.2 to 3.1 seconds are needed (depending on the type of instrument face and its position on the instrument panel). The time required for the perception

of light signals ranges from 0.8 to 1.2 seconds, depending on the color and light characteristics of the signal and its position on the signal panel. We know that the proper assembly of a gauge group improves conditions of instrument-reading perception. The instruments used aboard space vehicles must be assembled into groups having a common functional purpose. The problem of external distinction between instruments is closely associated with grouping, since external distinction is a signal indicator considerably facilitating the "search" for an instrument on the instrument panel and for the perception of its readings. This difference can be achieved by varying the shapes of the instruments, their dimensions, and primarily by altering the instrument faces (the shape and color of the dials, the position of the scales, the shape and color of the needles, etc.). From this standpoint it is quite advantageous for the instruments to differ with respect to one another in terms of read-off device. Indication conditions can be considerably improved by combining into a single group the indicators of tape and profile mnemonic systems, as well as those executed in the form of a counter.

Our investigations have shown that the conditions of perception are not identical with combined three-needle indicators. The situation is decidedly better in the perception of readings from scales positioned in the upper part of the instrument than if the scales are situated at the side. Therefore, in combined instruments it is best to position the indicators of the most important parameters at the top.

It has been shown experimentally that upon application of 5-6 signals to a single signal group, the attention effect of the other signals diminishes. The lighting of an emergency or particularly important signal against a background of many similar signals may produce a delayed reaction on the part of the cosmonaut. Therefore, in constructing

the signaling [warning] systems for space vehicles, where a great quantity of such signaling devices is required, it is best to position these signals in individual panels separated by function.

To improve the efficiency and reliability of the warning system, the attraction properties of these signals must be raised. In particular, to improve the attraction of emergency and warning signals, they must be executed so as to flash at the onset of their operation, with transition to continuous burning after the response reaction of the cosmonaut. All signals requiring immediate response should be backed up by means of sound signals. All critical parameters must be signaled [indicated] with light or sound signals.

We know that achromatic paint, offering the greatest contrast, provides the best conditions of legibility. Achromatic paint is widely used in aviation where it has proved itself, since flight proceeds under conditions of a considerable lack of time. The situation is different in space flight, where the operational duties are completed within the available time, but the duration of a space flight differs from that of a flight in an aircraft. A space flight may last from several days to several months and even years. In a prolonged flight the interior of the cabin assumes particular significance for a human being. The cabin under such conditions serves the cosmonaut not only as his work station and laboratory, but as his sleeping quarters as well. The cosmonaut is therefore not indifferent to the interior appointment of the cabin and all of the equipment within. A cabin that provides no contrast and instruments and control levers that are roughly made and not pleasing to the eye will not enhance the maintenance of high nervous-system tonus and even, quite the opposite, can produce a state of irritation in the cosmonaut and impair his ability to work. Space-vehicle cabins and their equipment must therefore be painted in a manner which

takes into consideration the conditions of the flight and the nature of the artificial and natural lighting. The instruments and levers must be fabricated so that they satisfy high artistic and aesthetic requirements. The instrument panels and the control consoles, as well as the instruments themselves, must be painted in light pleasing shades which harmonize with the color of the cabin. As an exception we can cite the instruments intended for indication of parameters of vital significance.

The faces of these instruments, intended to provide exact and rapid readings, are best painted in shades providing the greatest contrast (black-white).

In our examination of the problems associated with the assembly of a space-vehicle cabin we should take note of the fact that the aesthetic and artistic approach should be extended in equal measure to the control units, i.e., the buttons, switches, etc. From this standpoint, the most appropriate shape for the control levers is the keyboard which in recent times has gained widespread acceptance in the assembly of control positions. These are not only pleasing in shape, but exhibit a number of advantages associated with convenience of use. The keyboard provides for the more compact distribution of the control units, and it is convenient to identify them; the "on-off" positions are easily visible on the console and are convenient when working with gloves.

In the development of space-vehicle instrumentation the instrument making experience accumulated by engineering psychology was used to some extent in various branches of industry and particularly in the aviation industry. However, the specific conditions of flight in space vehicles impose unique requirements on the monitoring-measuring instruments. The means of adapting the information systems to effective stimulation by means of an analyzer, to the rapid and exact evaluation of information and, consequently, to the execution of assignments with re-

gard to control aboard an aircraft and a space vehicle are both similar and different. A space flight in terms of its characteristics is sharply delineated into phases (stages). In this connection, unlike the flight in an aircraft, the individual phases (stages) such as, for example, an orbital flight, may proceed over an extremely prolonged period of time, while entry into and descent from an orbit to the earth takes but a comparatively short period of time. The nature of space flight is such that there is no need for the cosmonaut to receive information from all or from most of his instruments simultaneously; to the contrary, in each stage he uses but a limited number of instruments. During the orbital-flight stage he has no need to refer to those instruments designed for either launch or landing. Only on rare occasions will he have to refer to those instrument complexes required for the linking of space vehicles with space stations. As a result, in order to facilitate the work of the cosmonaut, information must be supplied in stages. This principle obviously is most easily realized if the space-vehicle indication systems make extensive use of indicators operating with electron-beam tubes.

Let us dwell briefly on the results of comparative investigations of a number of manual vehicle-orientation control systems. The principles employed in the evaluation of systems have been described above and were used in an investigation of the systems of manual angular vehicle-motion control with a relay control law, where the angular velocity of object motion attained a certain (constant) velocity upon appearance of a control signal. The specificity of the transient response to control of the object's angular motion included its duration and the absence of any natural damping. A three-channel contact control key and a special keyboard console were experimentally compared.

The experiments showed that prolonged depression of a key at a

certain angle calls for concentrated attention on the part of the operator to maintain the key in the required position. Prolonged concentration of attention on the maintenance of the key hinders the accomplishment of the required monitoring of instrument information and warning signals. As a result of fatigue in the hand and a reduction in the attention devoted to the maintenance of the key in the required position leads to involuntary shutting off and erroneous switching on of control channels. The angular-motion control process has been changed radically through resort to the keyboard console. The average values of the parameters characterizing the work of an operator handling these control units are presented in the table.

Наименование параметров 1	2 Управление рукояткой	3 Управление клавишным пультом
4 Среднее время ориентации в относительных единицах (по сравнению с идеальным управлением)	1,35	1,17
5 Среднее значение обобщенной ошибки углового положения, в градусах	5.87	3.37
6 Средний расход энергии, в условных единицах	15,4	11,9
7 Среднее количество ошибочных действий по управлению в процессе одного цикла ориентации	3,00	0,05
8 Среднее время считывания одного индикатора в процессе управления, в сек.	3,14	3,05

1) Parameter designation; 2) key control; 3) keyboard-console control; 4) average time for orientation in relative units (in comparison with ideal control); 5) average value of generalized error in angular position, in degrees; 6) average expenditure of energy, in conventional units; 7) average number of erroneous actions with respect to control during the process of a single orientation cycle; 8) average time required to read off a single indicator during the control process, in sec.

From the table we can see the advantages of controlling orientation by means of a keyboard console over control accomplished by means of a three-channel contact key. In the first control case the average time for orientation under identical initial conditions is shorter, the expenditure of energy on the execution of a given control problem is reduced, there is a significant reduction in the average number of erroneous orientation actions, and it is easier to train the skill of

controlling the system. Moreover, the use of a keyboard console makes it possible to raise the operator's range of scanning, to distribute his attention more effectively, and at the same time to reduce the average time required to perceive instrument readings during the process of executing a control action, and it also makes it possible for the operator to be distracted for a short time (in the case of need) for some other activity (for example, to operate the radio key, etc.).

Data from the integrated recording of electrophysiological indicators also show that operators executing control by means of a keyboard console worked under more favorable conditions.

S.G. Mel'nik and Yu.A. Pozanov worked together with us on an experimental project on the possibility of determining the direction of "movement" of objects on the ground through a porthole relative to the heading line when these objects were moving at low velocities.

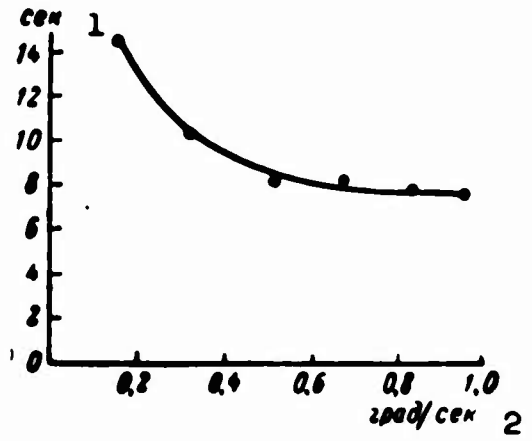


Fig. 5. The time required to perceive a deviation with respect to the angle of yaw as a function of the angular velocity of object "movement." 1) sec; 2) deg/sec.

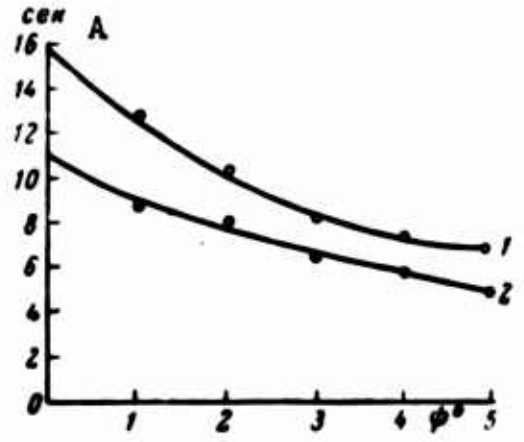


Fig. 6. The time required to perceive deviations with respect to the angle of yaw as a function of its magnitude. 1) $\omega = 0.32$ deg/sec; 2) $\omega = 0.95$ deg/sec; A) sec.

The time required to perceive deviations with respect to yaw as a function of the angular velocity of object motion is presented in Fig. 5.

The curve shows that with an increase in the velocity of motion

on the part of points of reference on the ground, the time required to perceive the direction of motion is reduced.

Figure 6 shows the time required to perceive a deviation with respect to the angle of yaw as a function of the magnitude of this deviation at various angular velocities of object movement within the instrument's field of view.

The curve shows that with an increase in the angle of yaw the perception time is reduced, since with an increase in the angle between the direction of object movement and the longitudinal line of the grid, given the same velocity of motion, the velocity component perpendicular to the heading line increases and movement distinguishable by eye is accumulated within a shorter period of time. Experiments also showed that without limitation of the observation time the accuracy of perceiving the direction of object motion parallel to the heading line is virtually independent of the velocity of object motion. The mean error in the determination of the direction of object movement relative to the heading line does not exceed 1°. The error in the determination of the angle of yaw with a reduction in the velocity of motion at the surface of the earth will increase if not enough time is available.

After the initial manned space flights we can assume that a trained cosmonaut is capable of functioning adequately during a flight in unanticipated and emergency situations, to control the space vehicle, to repair the space-vehicle systems (in the case of an emergency), to carry out various measurements and computations, including those not covered in the program. It is true, and this should be stressed, that all of this is still in need of thorough study and refinement during the course of the forthcoming manned flights into outer space.

For the successful completion of the assignments confronting a cosmonaut in flight, thorough training of the vehicle's crew on the

ground is required.

The development of the required skills calls for the utilization of training systems. These training facilities vary greatly in terms of designation and their significance in the system of cosmonaut training and structural execution. In this connection, it would be a good idea

to present an approximate description of the training-facility system required for the training of space-vehicle crews.

Depending on the type and nature of cosmonaut training, all training facilities can be divided into two basic groups: trainers and training facilities designed to condition the organism of the cosmonaut to the conditions of space flight; training facilities to develop the skills of the cosmonauts with respect to vehicle and vehicle-system control.

In addition, we can have combined training facilities in which the cosmonauts are trained to perform their activities aboard the space vehicle with skill, while at the same time conditioning the cosmonauts' organisms to the extreme conditions of cosmic flight. The first group might include trainers and facilities for physical and special training, as well as training devices designed to simulate the ambient medium and conditions of space flight. Below we will dwell in greater detail only on the second group which includes the training facilities designed to develop the professional skills required for vehicle and vehicle-system control. Depending on the effect on the human organism exerted by the factors of motion, these training facilities may be of the dynamic or of the static variety. Dynamic training facilities involve devices in which the cabin or work position of the operator being trained moves in conjunction with the modeled dynamics of the vehicle. Thus in dynamic training devices the human organism is acted upon by angular velocities and angular and linear acceleration. In static training installations

the cabin is in a fixed position.

Depending on the volume of the systems being modeled, and also depending on the stages of the flight and the problems being simulated, the training facilities designed to develop professional skills on the part of the cosmonauts may be designated as universal, integrated, specialized, or functional.

The universal trainer is a unified multipurpose facility designed to train crews for various types of space vehicles. It is an extremely complex installation whose characteristics can be altered and adjusted in accordance with the characteristics of various projected or existing space vehicles.

The integrated trainer is designed to provide for the training of all members of a crew in activities aboard a specific type of space vehicle during all stages of flight.

Proceeding from contemporary concepts of possible types of space vehicles and the need for the training of cosmonauts as thoroughly as possible, provision should be made for the following types of integrated trainers: for the crew of an interplanetary vehicle, for the crew of a space station, for the crews of satellite vehicles, and for the crews of astroplanes.

The specialized trainer is one intended for the development of skills in a wide range of specific operations required to carry out special assignments in a forthcoming space flight. In the specialized trainers the members of the crew must develop the skills to use certain of the systems designed to solve one or more flight problems.

The functional trainer is intended for the development of skills in connection with operations of a single type of activity or to train in such individual human functions as concentration, alertness, etc. This trainer requires no precise modeling of any specific system with

which the operator is called upon to work. The trainer must simulate only that which is required to elevate the functional capacity of the human being for the given type of activity.

In connection with the fact that without corresponding reinforcement, certain of the operator acquired skills may be lost, in addition to the above-indicated training facilities, specialized on-board trainers are required and these should be designed to maintain the required skills of the cosmonauts to carry out vital operations in the case of a prolonged space flight. These trainers may include, for example, a unit designed to train in landing on the earth or on planets, a unit to train in rendezvous with space stations, etc. The development and production of universal trainers at the present time is hardly expedient. The problem here lies in the fact that the assignments which must be carried out aboard space vehicles of various types, and the characteristics of the control element, are extremely varied. The reproduction of the characteristics of the various control elements, changing over a wide range, and the simulation of the operation of all space-vehicle systems would make the universal trainer cumbersome, expensive, and highly inefficient. Moreover, with the subsequent conquest of space, the number of various vehicles will increase and the training capacity of the universal trainer will not satisfy the requirements of training the crews for the various types of space vehicles.

Our analysis leads to the conclusion that trainers for the development of professional skills in the work performed by members of crews must be constructed in accordance with the designation of each space vehicle and in accordance with the features of cosmonaut activity aboard the corresponding type of vehicle.

For thorough cosmonaut training, as well as for purposes of scientific and research projects directed at the study of efficient methods

of indication and signaling, at the study of manual and semi-automatic control systems, and the assembly of space-vehicle equipment, we must have three types of trainers: integrated, specialized, and functional. This would provide for efficient training of an individual for professional activity under conditions of space flight.

In this work we have dwelt only on certain general problems of engineering psychology which, as the schedules for the conquest of outer space are expanded and as the control systems and the vehicle equipment with which a cosmonaut will have to work in flight become more complex, will be called upon to resolve an ever-increasing range of problems associated with the development of space-vehicle systems.

The first results in the area of engineering psychology in cosmonautics offer the hope that the methods of engineering psychology and the meaningful work of psychologists, doctors, engineers, designers, and the cosmonauts themselves will aid in attaining more effective solutions to the problems of interaction between operator and machine in such a complex biomechanical system of control as a piloted space vehicle.

Received

28 February 1964

REFERENCES

1. V.G. Denisov, Sb. "Problemy kosmicheskoy biologii," [Collection entitled "Problems of Space Biology,"] 2. Izd-vo AN SSSR, [2. Acad. Sci. USSR Press], 1962 page 54.
2. N.I. Boyko, S.G. Darevskiy, Ye.S. Zav'yalov et al. Tez. dokl. II s"yezda ob-va psikhologov, M., [Report at the 2nd Congress of the Society of Psychologists, Moscow], 1963, page 7.
3. Ye.S. Zav'yalov, A.P. Kuz'minov, V.I. Mankevich, Voprosy psikhol-

ogii, [Problems in Psychology], 3, 8, 1963.

4. A.P. Kuz'minov, V.F. Onishchenko, M.M. Sil'vestrov. Tez. dokl II s"yezda ob-va psikhologov, M., 1963, page 63.
5. L.B. Itel'son, Voprosy filosofii, [Problems in Philosophy], 4, 33, 1962.
6. B.F. Lomov, Chelovek i tekhnika, [Man and Engineering], Izd-vo LGU, L., [Leningrad State University Press, Leningrad], 1963.
7. R.F. Brissenden, E.C. Liheberry. Aerospace Eng., 21, No. 6, 64, 1963.
8. J.T. Cebutano, H.S. Alecsander, Amer. Rocket Soc., 145, 1961.
9. R. Grandpierre, F. Violedte, J. Physiol., 54, 1, 1962.

Manuscript Page No.	[Transliterated Symbol]
210	cp = sr = sredneye = average

UDK 629.198.61

REACTIVITY STATE OF ANIMAL ORGANISM AFTER SUBJECTION TO CERTAIN SPACE-FLIGHT FACTORS

V.V. Antipov, V.I. Davydov, E.F. Panchenkova,

P.P. Saksonov and G.A. Chernov

Data are given on the changes in the reactivity of animals to physical exertion following subjection to g-forces, the combined action of acceleration and x-irradiation, and irradiation by 120-Mev protons. The shifts detected indicate a persistent change in the stability to physical load following subjection to acceleration, ionizing radiation and combinations of these factors. The change in the tolerance of centrifuged animals to physical load correlates with ceruloplasmin shifts in the blood serum.

The problem of the reactivity of an organism under flight conditions has acquired great practical importance in connection with man's successful mastery of outer space.

In evaluating the reaction of an organism during flight to one or another factor, it is necessary to take its functional state into account, and this depends on many characteristics of the flight and primarily on its duration and the type of orbit. Thus, for example, as the flight time increases, the importance of the vibrations and accelerations that arise during launching of the vehicle decreases, and the importance of such factors as weightlessness, cosmic radiation, etc., increases.

It is quite obvious that under actual space-conditions, the widest imaginable variety of combinations in the action of the factors is pos-

sible – as regards force, duration and sequence – and that this makes it difficult to analyze the mechanisms and predict the over-all effect. On the other hand, knowledge of possible deviations in the organism's reactions under these conditions is necessary for elaboration and use of various methods for protecting it from the injurious effects of space flight.

Despite the fact that certain advances have already been scored in this field, many problems of reactivity in cosmic bioloby and medicine still require solution. This applies in particular to study of the organism's functional state, its tolerance for various factors in the external environment after being subject to transverse acceleration, corpuscular radiation, weightlessness and the like. Investigation of the organism's reactions when dynamic factors of space flight and ionizing radiation act in concert is of great practical interest.

A number of papers have been devoted to study of the organism's reactivity after subjection to acceleration [1, 2].

Another important factor in space flight is cosmic radiation – one of the principal obstacles to the conquest of outer space [3, 4]. To a considerable degree, the cosmic radiation consists of high-energy protons generated by solar chromospheric layers; these particles are also found in the composition of the outer and inner radiation belts. This enhances substantially the possibility of radiation injury in space flights of long duration.

Numerous studies indicate that the reactivity of animals after γ- and x-irradiation, during the phase of acute radiation sickness, varies quite substantially. Definite interest attaches to study of the organism's reactivity after the manifest clinical symptoms of acute radiation sickness have disappeared, with the object of obtaining a basis for evaluating the degree to which the disturbed functions have been re-

stored. It is known that proton irradiation produces more persistent irreversible changes in certain organs and tissues and results in a greater number of cases (as compared with γ- and x-irradiation) of malignant neoplasms as a remote aftereffect [5].

The present study has resolved some of the problems outlined above: the resistance of the organism was studied after subjection to accelerations, bombardment with high-energy protons and the combined action of acceleration and ionizing radiation. An attempt was also made to unveil some of the mechanisms behind the organism's tolerance changes as an effect of transverse accelerations. The functional state of the organism after acceleration and exposure to ionizing radiation was evaluated by applying a physical load: the animals were made to swim until they perished. Application of such an extraordinary load makes it possible to bring out more fully the state of the organism's compensatory mechanisms after subjection to the space-flight factors under study.

METHOD

Experiments were performed on 700 nonlineal mice and 80 rats.

A study of the animals' tolerance to physical exertion following acceleration was made in the first series of experiments. In one case the mice were investigated 15 minutes after subjection to acceleration, and then 1, 2 and 4 hours and 1, 3 and 7 days after a single dose of acceleration (8 g, 20 min in the ventrodorsal direction). In another case the animals were subjected to three acceleration treatments at weekly intervals and given the physical-exertion test 4 hours after the last centrifuging.

The experiments of the second series were performed on male rats following irradiation with 120 Mev protons in a dose of 700-850 or 1400-1770 rad at a dose rate of 60 ± 10 rad/min. The animals were proton-bombarded in the pulsed beam of the OIYaI synchrocyclotron (located at

Dubno) with a flux density of 10^8-10^9 protons per $cm^2 \cdot sec^{-1}$. There were 100 pulses per second, each lasting 200-400 μsec. The proton energies were reduced to 120 Mev by use of polyethylene blocks 1.5 m thick, supplemented by 6 cm of lead. The flux intensity and the proton dose were deter ined from the activity induced in carbon indicators.

Animals of this series were tested for tolerance to physical load 40 days after irradiation, and then the weights of the spleen and left suprarenal glands were determined.

In a third series of experiments, we studied the tolerance of the mice for physical exertion after combined subjection to acceleration and x-irradiation. The animals were subjected to a g-force (8 g for 15 min) 4 hours before irradiation in doses of 400 and 700 r. The irradiation conditions were as follows: 180 kv, 10 ma, Al-1.0 mm, Cu 0.5 mm, dose rate 13 r/min. Also determined for the mice that had been subjected to acceleration were the ceruloplasmin activity in the blood serum, using R. Henry's method [6], and the biological activity of whole blood after preliminary hemolysis of the erythrocytes with distilled water and precipitation of the blood proteins with hydrated zinc oxide ($NaOH - ZnSO_4$). The supernatant liquid was tested for serotonin creatinine sulfate on a section of rat large intestine.

The physical-exertion te t imposed on the animals was that of allowing them to swim until they perished. The water temperature in the swimming tub was 19-20°. In view of the seasonal differences in the tolerance of the animals to various environmental factors, each experimental group had its own control.

The experimental data obtained were processed by the Litchfield probit analysis method [7], with determination of the average effective swimming time (ET_{50}) of the experimental animals. The statistical processing given other indicators amounted to determination of the arithme-

tic mean and the root-mean-square error.

RESULTS AND DISCUSSION

The results of the first-series experiments are assembled in Tables 1 and 2. As will be seen from Table 1, the tolerance of the mice (males and females) to physical exertion varies as follows between various points in time after subjection to g-forces. At first (15 minutes after acceleration), the tolerance of the experimental animals shows no differences from that of the control; 1-2 hours after the acceleration, the ET_{50} increases ($P < 0.05$), but it has dropped sharply 4 hours after the acceleration ($P < 0.05$). On the 7th day after centrifuging, there was a definite increase in the tolerance of the mice to physical load. The data obtained represent testimony to the effect that the functional changes occurring during acceleration last for a very long time.

TABLE 1

Average Time before Death of 50% of Mice (ET_{50}) During Physical Exertion Following Acceleration

1 Дата опыта	2 Время обследования после действия ускорения	3 Количество животных	4 ET_{50}, мин	5 Доверительные границы (мин) при $P = 0.05$
24.01.62	Контроль 7	23	170	8 149÷194
6	15 мин. 9	24	160	142÷172
	4 час. 10	23	21	19÷ 23
25.01.62	11 24 час.	24	150	120÷187
29.01.62	Контроль 7	24	114	95÷137
	12 1 час.	24	160	147÷174
	2 часа 13	24	154	145÷164
	7 Контроль	24	280	240÷326
2.02.62	4 час. 10	24	20	17÷ 24
	7 Контроль	24	205	193÷218
6.02.62	2 часа 3-и сутки	24	200	184÷217
	Контроль 7	24	120	95÷172
	15 7-ми сутки	24	310	268÷372

1) Date of experiment; 2) time between acceleration and test; 3) number of animals; 4) ET_{50}, min; 5) confidence limits (min) with $P = 0.05$; 6) 24 January 1962; 7) control; 8) 149 to 194; 9) 15 min; 10) 4 hours; 11) 24 hours; 12) 1 hour; 13) 2 hours; 14) on 3rd day; 15) 7th day.

The changes observed following acceleration can be related to shifts in the activity of certain biological active compounds in the blood.

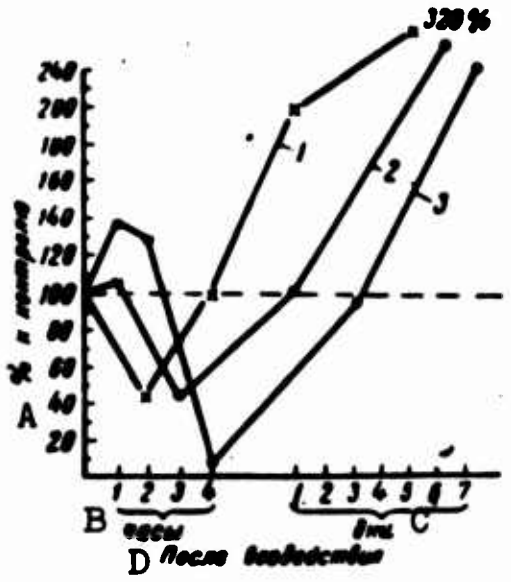


Fig. 1. Comparative data on variation of ceruloplasmin activity (1), biological activity of blood (2), and mean effective swimming time (3) after acceleration. A) % with reference to control; B) hours; C) days; D) after disturbance.

Juxtaposition of the curves in Fig. 1 indicates that the general direction taken by the changes is quite uniform. The correlation between these three indices is most distinctly manifest on the 5th to 7th days after acceleration: with increasing ceruloplasmin and biological activity of the blood, the tolerance of the centrifuged mice to physical exertion increases. The literature contains references to the effect that ceruloplasmin increases in the plasma under various stress-type disturbances and in many pathological states. The universal nature of this reaction has made it possible to conclude that ceruloplasmin is

TABLE 2

Seasonal Changes in Tolerances of Mice to Physical Exertion 4 Hours after Acceleration

1 Дата опыта	2 Группы	3 Количество животных	4 ET_{50}, мин	Доверительные границы (мин) при $P = 0.05$ 5
6 24.01.62	7 Контрольная	23	170	8 149÷194
	Опытная 9	23	21	19÷23
30.01.62	7 Контрольная	24	280	240÷325
	Опытная 9	24	20	17÷24
2.02.62	7 Контрольная	24	130	98÷172
6.02.62	7 Контрольная	24	205	193÷218
7.02.62	Контрольная	24	184	166÷204
	Опытная 9	24	199	179÷220
12.02.64	7 Контрольная	18	93	73÷119
		10	38	24÷59
	Опытная 9	10	71	41÷109
9.03.62	7 Контрольная	15	21	16÷28
	Опытная 9	15	41	28÷69
	7 Контрольная	20	34	23÷51
	Опытная (трехкратное ускорение) 10	20	64	48÷96
30.03.64	Контрольная 7	30	11	8÷14

1) Date of experiment; 2) group; 3) number of animals; 4) ET_{50}, min; 5) confidence limits (min) with P = 0.05; 6) 24 January 1962; 7) control; 8) 149 to 194; 9) experimental; 10) experimental (three accelerations).

an adaptation enzyme [8, 9]. It is also interesting to note that in humans in a state of heightened activity, awareness, excitement and sleeplessness, the copper content in the blood rises (the basic portion – 96% – of the copper in the blood is bound in ceruloplasmin) and, conversely, the blood copper content drops in states of drowsiness and relaxation [9]. It would appear that there is something in common in the mechanisms (most probably neuroendocrine) behind the rise in ceruloplasmin, the biological activity of the blood and the physical tolerance of the centrifuged mice. The nonuniform reactivity that we observed in the animals after acceleration in various months of the year (January, February, March) indicates that the neuroendocrine system is involved to some degree (Table 2).

Thus, we established that the tolerance of the control animals for physical exertion is quite substantially different for different times of the year: in March, the ET_{50} was reliably lower than the figures obtained in January and at the beginning of February (Table 2, Fig. 2). Figure 2 presents results that indicate changes in the physical tolerance of mice 4 hours after acceleration. Our attention is drawn to the opposite directions taken by the changes in the physical stability of the control and experimental mice in January and March. It is also known from literature sources that readjustment of the hypothalamo-hypophyseal-endocrine system takes place on seasonal changes in the environment [10]. Thus, we observe a perverted response of the hypothalamo-hypophyseal system in response to adequate environmental stimuli in animals that have been subjected to ionizing radiation [11, 12].

The results of the second series of experiments are presented in Table 3. In all groups in which the animals were subjected to proton irradiation, the swimming ET_{50} of the animals was shorter than that of the control group. For irradiation in a dose of 1400-1770 rad, the ET_{50} was

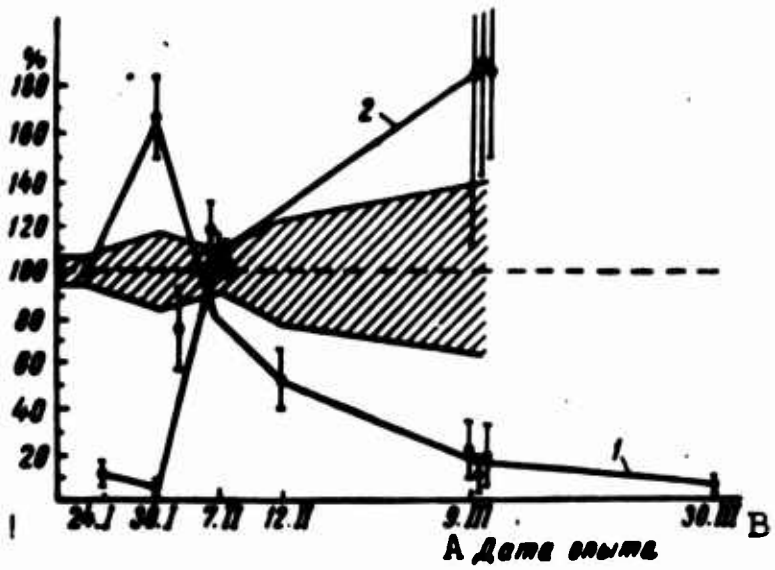


Fig. 2. Seasonal changes in tolerance for physical load 4 hours after acceleration. 1) Control of 24 January 1962 taken as 100%; 2) own control group taken as 100%; shaded zone indicates reliable control interval (%) with P = 0.05. A) Date of experiment; B) 30 March.

TABLE 3

Average Swimming Time and Weight of Spleen and Suprarenal Gland (Left) in Rats 40 Days after Irradiation with 120-Mev Protons

1 Группы	2 Количество животных	3 Средний вес животных, г	4 ВТ, мин	5 в % к контролю	6 Вес селезенки 7 мг	5 в % к контролю	8 Вес надпочечника 7 мг	в % к контролю
9 Контроль	10	225,5±7,2	22 (19÷26) 10		817,3±57,2		18,3±1,24	
11 1400—1700 рад	4	156,8±0,6	10(8÷13)	45	453,0±94,5	55	22,5±2,6	123
700—850 рад	19	208,8±8,5	18 (15÷21)	82	620,6±64,4	75	19,1±0,87	105
700—850 рад 12 (Ф + А)*	15	210,6±9,2	19 (15÷24)	86	613,6±14,9	75	19,7±1,2	108

Note: *F + A: phenatine (2 mg/kg) + ascorbic acid (200 mg/kg).

1) Group; 2) number of animals; 3) average weight of animals, g; 4) minutes; 5) in % of control; 6) weight of spleen; 7) mg; 8) weight of suprarenal gland; 9) control; 10) 19 to 26; 11) rad; 12) (F + A).*

45%, and with 700-850 rad it was 82% with reference to the control. In animals that had been irradiated with a dose of 1400-1770 rad, the weight of the spleen was 45% lower than in the control, while the weight of the suprarenals was 23% higher. The weight of all organs in rats

that had been irradiated with a dose of 700-850 rad differed little from the corresponding figures for the control rats. Thus, despite the absence of clinical symptoms of radiation injury in proton-irradiated rats, the tolerance for physical load was lower than in the control.

The group of animals that had been irradiated with a dose of 700-850 rad was given an intraabdominal injection of phenatine (2 mg/kg) and ascorbic acid (200 mg/kg) 30 min before subjection to the physical load with the purpose of raising the tolerance of the irradiated rats to this exertion. It is known that whole-body irradiation is accompanied by a considerable decrease in the ascorbic acid content in the suprarenals [13, 15]. Injection of this vitamin into the organism should obviously increase the tolerance of the irradiated rats for physical exertion, the more so because the synthesis of corticoids uses ascorbic acid. On the other hand, it has been established that phenatine (β-phenylisopropyl nicotinamide), which has a sympathicomimetic effect, increases ability to perform work.

The physical-exertion test indicated that the ET_{50} of swimming was somewhat higher for animals that had been given phenatine and ascorbic acid than in rats that had not been given the preparation. The swimming ET_{84} for irradiated and control animals approximated the value of the corresponding index for the control group: 29 and 30 min, respectively. However, the increase in swimming time in irradiated animals to which the preparations had been administered was statistically unreliable as compared with the irradiated control. The absence of a substantial increase in physical endurance in the irradiated animals after administration of phenatine to them is obviously related to functional insufficiency of the suprarenal glands, since it is known that the basic effect of preparations of this type is simply to raise the sensitivity of adrenoreactive systems to adrenaline and sympathetic impulses.

In the third series of experiments, an attempt was made to evaluate the tolerance to physical exertion after combined subjection to acceleration and irradiation, the most realistic simulation of spaceflight conditions. The experiments were performed in March. We expected that since the tolerance of the centrifuged mice for physical exertion was higher during this period, the resistance to the effects of ionizing radiation would also be higher. The experimental data secured are shown in Fig. 3. As will be seen from the figure, the mortality of the animals after combined subjection to acceleration and irradiation was lower than for x-irradiation taken alone. This relationship was observed for both of the doses used. Acceleration before irradiation increased the $DL_{50/30}$ by approximately 100 r as compared with the figure for the irradiation control. The animals that had survived irradiation were given the physical-endurance test. The swimming ET_{50} was 11 min for the control, 11.2 min for acceleration + irradiation with 400 r, 9 min for acceleration + irradiation with 700 r and 5 min for irradiation with 700 r. As will be seen from these data, the swimming ET_{50} for the animals that had survived the combined treatment was higher than in the corresponding irradiation control. The ET_{50} correlated with the survival rate on the 30th day after irradiation.

An increase in the radioresistance of centrifuged animals was also noted by Zellmer [16]. At the present time, it is difficult to give an exhaustive explanation for this phenomenon. It can only be stated that tissue hypoxia (particularly in the hypothalamic region of the brain) due to the g-forces may result in an increase in the activity of the pituitary-adrenal cortex link in the neuroendocrine regulatory system. It is known that an increase in the functional activity of the adrenal cortex produced by administration of salicylates, AKTG [ACTH] or desoxycorticosterone acetate increases to a certain degree the resistance of

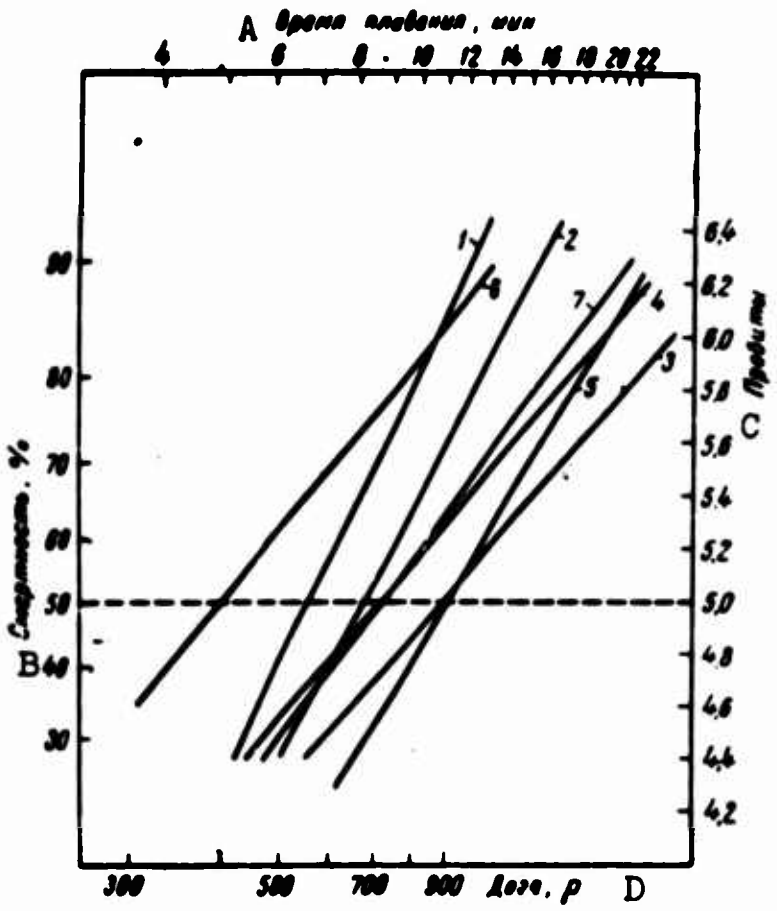


Fig. 3. Probit lines. 1) Mortality on 30th day after x-irradiation; 2) mortality on 30th day after acceleration + x-irradiation; 3) physical endurance of control animals; 4) after irradiation with 400 r; 5) after acceleration + irradiation with 400 r; 6) after irradiation with 700 r; 7) after acceleration + irradiation with 700 r. The graduation on the lower abscissa is for curves 1 and 2, and that on the upper abscissa refers to curves 3-7. A) Swimming time, min; B) mortality, %; C) probits; D) dose, r.

animals to whole-body irradiation [17]. This mechanism in the "protective" effect of radiation would obviously not be the only one. Tissue hypoxia may give rise to formation of other tissue metabolism products. It is interesting in this connection to examine literature data on the existence of a specialized hemopoietic factor resisting the effects of ionizing radiation and generated during hypoxia.

CONCLUSIONS

1. A statistically reliable decrease had taken place in the physical tolerance of the animals 4 hours after acceleration, and was followed by a rise in this index one week after subjection to the g-forces. Seasonal variations in physical endurance were noted in animals that had

been subjected to acceleration.

2. The changes in the reactivity of centrifuged animals to physical exertion correlate with ceruloplasmin shifts in the blood.

3. The tolerance of the animals for physical load was reduced 40 days after irradiation with 700-1770 rad of 120-Mev protons.

4. Preliminary centrifuging elevated the resistance of the animals to x-irradiation to some degree.

Received
7 May 1964

REFERENCES

1. V.Ye. Belay, P.V. Vasil'yev and S.P. Kolchin, Farmakol. i toksikologiya [Pharmacology and Toxicology], 26, No. 5, 559, 1963.
2. I.M. Khazen, Collection entitled "Aviats. i kosmich. meditsina" [Aviation and Space Medicine], Izd-vo AMN SSSR [Publishing House of the Academy of Medical Sciences USSR], Moscow, 1963, page 463.
3. V.V. Antipov, N.N. Dobrov and P.P. Saksonov, Collection entitled "Problemy kosmich. biologii" [Problems of Space Biology], 3. Izd-vo AN SSSR [Academy of Sciences USSR Press], 1963, page 113.
4. N.M. Sisakyan, V.V. Parin, V.V. Antipov, N.N. Dobrov and P.P. Saksonov, Izv. AN SSSR, seriya biol. [Bulletin of the Academy of Sciences USSR, Biology Series], 3, 341, 1964.
5. S.P. Yarmonenko, E.B. Kurlyandskaya and G.A. Avrunina, Collection entitled "Aviats. i kosmich. meditsina." Izd-vo AMN SSSR, Moscow, 1963, page 510.
6. R.J. Henry, N. Chiamory and J.D. Vacob, Proc. Soc. Exp. Biol. Med., 104, No. 4, 620, 1960.
7. J.T. Litchfild and E.J. Wilcoxon, J. Pharmacol. Exper. Therap., 96, 99, 1949.

8. S. Martens, Swenska Lakartidingen [Swedish Medical Journal], 12, 809, 1959.

9. B.M. Gekht, Collection entitled "Fiziologiya i patologiya dientsefal'noy oblasti golovnogo mozga" [Physiology and Pathology of the Diencephalon], Izd-vo AN SSSR, 1963, page 432.

10. G.I. Markelov, Zabolevaniya vegetativnoy sistemy [Diseases of the Vegetative System], Kiev, 1948.

11. A.V. Lebedinskiy, Z.N. Nakhil'nitskaya and N.P. Smirnova, Meditsinskaya radiol. [Medical Radiology], No. 7, 3, 1959.

12. B.I. Davydov, Radiobiologiya [Radiobiology], 1, 394, 1961.

13. L.C. Wyman, R. Whitney and D. Patt, Federat. Proc., 12, 161, 1953.

14. B.C. Wexler, R. Peucharz and S.F. Thomas, Amer. J. Physiol., 183, 71, 1955.

15. L.V. Bozhenko, Meditsinskaya radiol. [Medical Radiology], No. 2, 44, 1957.

16. R.W. Zellmer, G.J. Womack, R.C. McNee and R.G. Allen, Aerosp. Med., 33, No. 3, 356, 1962.

17. E.H. Betz, Contribution a L'etude du Syndrome endocrinen provoque par L'irradiation totale de L'organisme [Contribution to Study of the Endocrine Syndrome Produced by Whole-Body Irradiation of the Organism], George Thone, Liege, 1955.

Manuscript Page No.

[List of Transliterated Symbols]

236 ОИЯИ = OIYaI = Ob"yedinennyy institut yadernykh issledovaniy = United Nuclear Research Institute

243 АКТГ = AKTG = adrenokortikotropnyy gormon = adrenocorticotropic hormone (ACTH)

UDK 629.198.61

EXPERIMENTAL INVESTIGATIONS OF THE INFLUENCE OF LANDING IMPACT g-FORCES ON THE ANIMAL ORGANISM

S.A. Gozulov, G.P. Mirolyubov, N.N. Popov and N.I. Frolov

The results of experimental studies of the effect of landing-impact g-forces on animals are reported. The manner in which morphological changes in the internal organs and the degree of disturbance to the functions of the cardiovascular system depend on the magnitude of the disturbance is determined.

The impact of an airplane with the ground in a forced landing or the touchdown of separable parts of aerospace vehicles (cabins, escape capsules, etc.) on the surface of the ground or water is attended by impact overloads of great magnitude.

In recent years, interest in study of the effect on the organism of landing overloads has heightened in connection with the development of systems for landing space vehicles on the earth or other planets [1, 2].

Distinctive characteristics of the landing overloads include their short duration, high rise rates of the order of thousands of units per second, and their high peak magnitude.

The above overload parameters are determined by the character of the soil on which deceleration takes place and the speed at which the cabin is descending at the moment of impact. Increasing the descent velocity or reducing the deceleration path of the vehicle to be landed results in a sharp increase in the overload. For a given descent rate, an increase in the peak overload results from a shortening of its duration

– of the deceleration path.

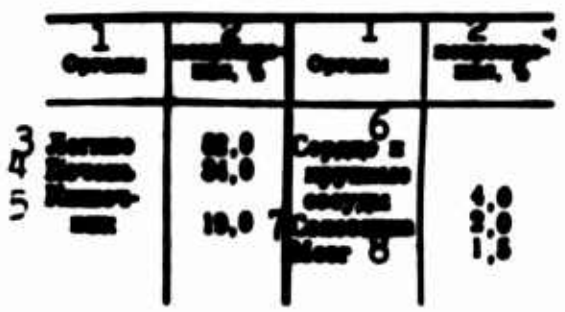
According to foreign authors, the optimum touchdown speed of a space-vehicle cabin from the standpoint of engineering calculations is of the order of 7-9 m/sec [3]. Depending on the design characteristics and landing engines of the vehicle, the most favorable descent speed may be increased to 20-30 m/sec. On the basis of personnel-safety considerations, it is recommended that the touchdown speed be below 12.2 m/sec, since failure of the retro-firing landing engines can be allowed for in this case [4].

In the present investigation, we studied the consequ nces for animals (rats and dogs) of impact overloads produced on touchdown at a speed of 4 to 13.6 m/sec. The highest values of the average overload reached 800-900 min, and the duration varied from 10 to 1 msec.

The experiments were run on an impact machine that made it possible to vary the falling height (speed) of the platform and the path (time) of its deceleration within the necessary limits. The animals were secured to the dropping platform in a special bed, placed in the horizontal position in such a way that the overload would act in the dorsoventral direction.

TABLE 1

Frequency of Injuries to Certain Organs in Rats (in %)



1) Organ; 2) injuries, %; 3) lungs; 4) liver; 5) intestine; 6) heart and major vessels; 7) spleen; 8) brain.

A total of 169 experiments were performed with a single fall (100 on rats and 69 on dogs), plus 40 with repeated (up to 5) applications of the impact overload (on 3 rats and 5 dogs). The electrocardiogram and respiratory frequency were registered for the rats, while for the dogs, in addition to the above, we traced the arterial pressure by the tachooscillator method, using an electromagnetic or piezoelectric

sensor. The arterial pressure was recorded from the carotid artery, which was diverted into a skin flap, or, in some of the experiments, from the animal's caudal artery by a supplementary technique.

The results of studies of the physiological functions indicated for the rats that the nature of the pulse-frequency variation after the disturbance depends on the magnitude of the overload sustained. For relatively small overloads, the pulse frequency either rises (43%) or drops (57%) by 20-30% immediately after touchdown. After a heavy g-force has been sustained, however, the pulse frequency decreases in all cases to 50% of the initial value, i.e., an absolute depression of the function is noted [5].

This qualitative change in the reaction as a function of the magnitude of the g-force sustained had a characteristic limit for each range of landing speed. Thus, for example, at speeds from 7.5 to 10 m/sec, a slackening of pulse frequency was observed after g-forces of 350-400 units, while for speeds of 11-13.2 m/sec, the corresponding phenomenon arose at overloads as low as 200-250 units. If, on the other hand, the touchdown speed did not exceed 6 m/sec, no decrease in pulse frequency was observed in the animals under the same experimental conditions even with g-forces of 500 units and more.

The data obtained show that for a given descent speed, the extent to which the pulse slackens increases with increasing operating overload or, for a given overload, with increasing touchdown speed. The same type of correspondence is observed in the extent of injury to the internal organs. In the absolute majority of cases, when the impact disturbance caused a slackening of the animal's pulse frequencies it was accompanied by injury to their internal organs (ruptures of the capsule or parenchyma of an organ, extravasation of varying degrees, and, in occasional cases, fat embolism). In 100 rats subject to the disturbance,

damage to one or more organs was observed in 82 cases, while in the other 18 only occasional subcutaneous or intramuscular extravasation was noted.

The frequency of injuries to the rats' internal organs was not uniform (Table 1). Most frequently injured were the lungs, liver, loops of the intestine and the mesentery; the heart, major vessels and brain were less frequently damaged. In occasional cases, we also noted extravasation in the perinephric cellular tissue and subcapsular extravasation in the spleen. No cases of bone injury were observed with g-forces of this magnitude.

Pathomorphological examinations showed that in addition to the numerous hemorrhages resulting from impact overloads of large magnitude, extensive rupturing of the lungs and liver occur, in relation, as are the injuries in other regions, to the direction of the impact and the situation of the organs (impingement on adjacent parts of the skeleton, as well as collisions of organs with other organs). Landing overloads of great magnitude may, in combination with primary injury to the organs and tissues of the organism, also produce secondary degenerative changes matching the morphological pattern of brain concussion, which is more distinct in the dog [6]. Following application of small g-forces, i.e., those that do not produce absolute bradycardia under the present experimental conditions, the changes registered in the electrocardiogram were insignificant and inconsistent.

As the operating g-force is increased, characteristic changes appear in the electrocardiograms; the S-wave deepens and the T-wave rises to amplitudes of 0.05-0.1 mv; the S-T segment also shifts to a position above the isoelectric line. In about 12-15% of cases, polytopic ventricular extrasystoles, auricular fibrillation and incomplete atrioventricular block are observed after the impact disturbance. The changes regis-

tered may indicate a functional disturbance that arises not only as a result of extracardial causes, but also directly from "transitory myocardial excitability block" due in this case to a mechanical impact disturbance of the tissues and organs [7].

TABLE 2

Change in Frequencies of Pulse and Respiration in Dogs on Application of Landing Overloads (up to 200 units)

	1 Кол-во опыт.	2 До приземления: 4 30 мин.	5 1 мин.	6 10 сек.	3 После приземления: 6 10 сек.	5 1 мин.	4 30 мин
7 Частота пульса	55	145* ±38	134 ±45	144 ±36	193 ±16	142 ±38	145 ±28
8 Частота дыхания	55	36 ±10	44 ±27	43 ±23	55 ±31	49 ±34	29 ±13

*Average ± root-mean-square deviation.

1) Number of experiments; 2) before landing; 3) after landing; 4) 30 minutes; 5) 1 minute; 6) 10 seconds; 7) pulse frequency; 8) respiratory frequency.

In all of the experiments on dogs, the same same 13-13.6-m/sec landing speed was used, and the average impact-overload value varied from 80 to 870 units. Landing with a g-force below 200 units (55 experiments) produced the maximum quickening of the respiratory and cardia frequencies immediately after impact, and in the first 10-15 sec. Then the pulse and respiration returned to the initial level within 1-2 min. The dynamics of the pulse and respiratory frequency changes are shown in Table 2.

Only in two cases do we observe a slackening of the pulse (primary brachycardia) and a brief suspension of respiration during the first 10 sec after application of a g-force smaller than 200 units; this is then supplanted at 20 seconds by a quickening. The arterial pressures prevailing before and after the disturbance were registered in 19 exper-

iments of this series (g-force smaller than 200 units). Following the variations of pulse frequency immediately after landing, the maximum and minimum arterial pressures rose by 10-30 mm Hg.

In two dogs, polytopic ventricular extrasystoles were observed to arise on the electrocardiogram against a background of sharp change in the pulse frequency at the 5th and 10th seconds after the disturbance. During this interval, the pulse varied in one dog from 264 to 156 (initial level) over 5-10 sec, and in the other from 172 to 126 beats/min (also down to the initial level). Usually, the function normalized in the dogs at the 3rd-5th minute after the g-force disturbance was applied.

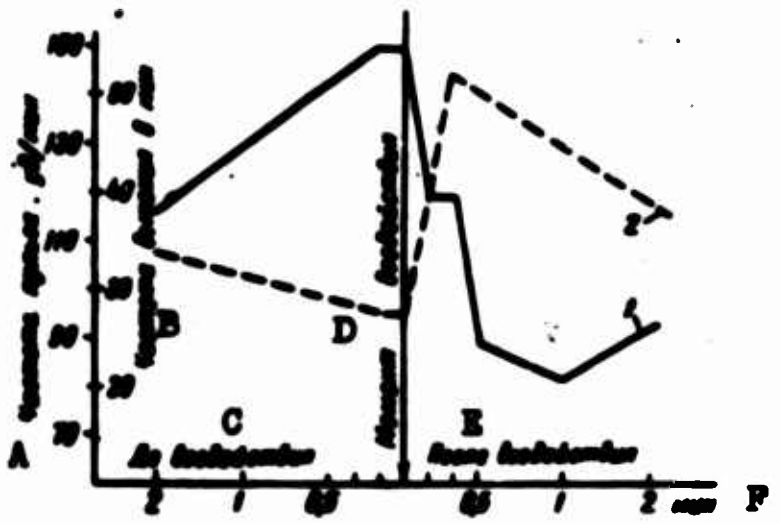


Fig. 1. Changes in pulse and respiratory frequencies in dog on application of landing g-force greater than 200 units. 1) Pulse; 2) respiration. A) Pulse frequency, beats/min; B) respiratory frequency, per minute; C) before disturbance; D) moment of disturbance; E) after disturbance; F) minutes.

In experiments with g-forces greater than 200 units (14 experiments), a slackening of the pulse frequency by 40-60% and of the respiratory frequency by 20-30%, or a 10-30-sec stoppage of respiration, was observed in the dogs immediately after the disturbance. Figure 1 shows typical curves of pulse-frequency and respiratory-frequency variation.

Qualitative reaction changes were also in evidence in the case of arterial pressure in this series of experiments. As a rule, we observed at the first minute after the overload a decrease in the maximum arter-

ial pressure by 30 mm Hg and a 20-mm-Hg drop in the minimum pressure, and it did not reach the initial level, i.e., there was only a relative depression of the function after the disturbance.

No pathological disturbances were observed on the electrocardiograms of the dogs after application of large g-forces. A characteristic observation (as for the smaller animals) was a deepening of the S-wave and an increase in the T-wave against a general background of tachycardia.

To evaluate the cumulative effects, we ran a series of experiments on white rats and dogs in which the impact g-force was applied repeatedly.

The rats were subjected to the disturbance three times, at 300-350 units with a 10-minute interval. The changes registered here in the respiratory frequency and pulse indicate that the changes observed in the physiological functions are intensified when the impact disturbances are repeated (Fig. 2). The disturbances to the cardiac rhythm (extrasystole, atrioventricular block) become particularly pronounced. As will be seen from Fig. 2, the first impact disturbance produced minor primary slackening (inhibition) of the function. Subsequent disturbances were accompanied by exacerbating disturbances to myocardial function, making their appearance immediately after the disturbance.

In this series, the dogs were subjected to 4-5 applications of moderate impact overloads (within the 200-unit range), the duration of which was 10-15 msec, with an interval of 2-5 days between disturbances. In the case of repeated disturbances of relatively small magnitude, no increase was manifest in the functional derangements (Fig. 3), as had been the case in the white rats subjected to large overloads. Moreover, the provision of a sufficiently long interval between disturbances resulted in a lesser degree of functional deviations in the repeated ex-

periments. As will be seen from Fig. 4, a decrease in the extent of the pulse-frequency deviations from the first to the fourth disturbance is observed with repetition of relatively small disturbances separated by adequate intervals of time. It is as though an adaptation to the extraordinary force of the stimulus has taken place. The sharpness of the reflex reactions diminishes, despite an increase in the effective overload (from 92 to 158 units), and the heart accomplishes its work at a more favorable physiological level. Thus, repeated application of landing shocks may produce different physiological effects depending on its magnitude and frequency of repetition.

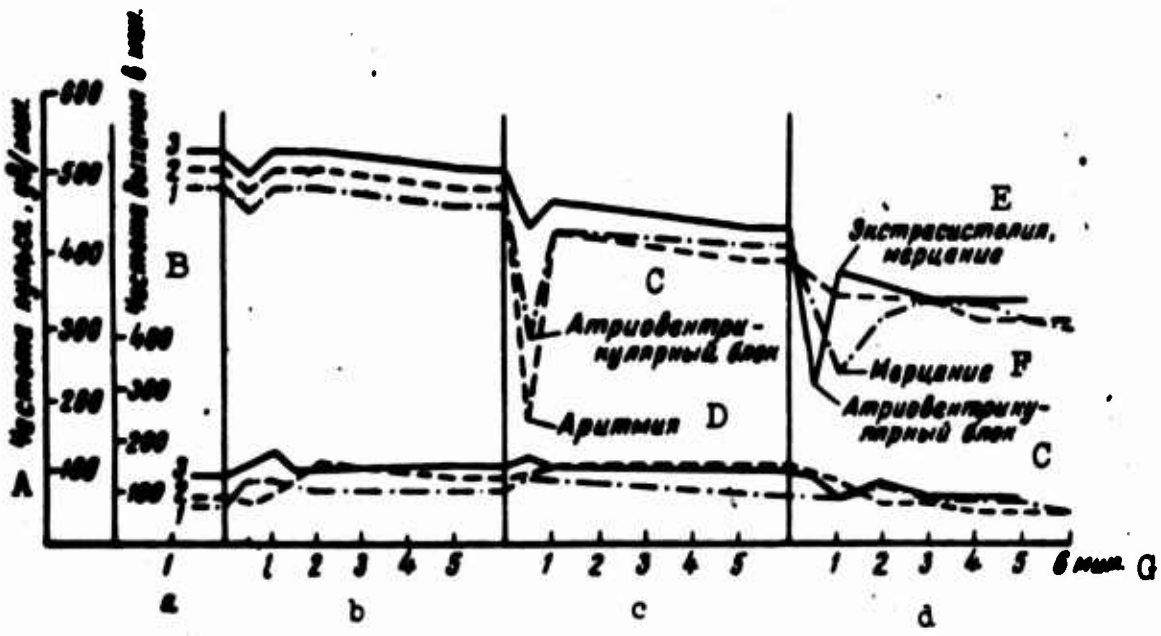


Fig. 2. Changes in pulse and respiratory frequencies of white rats on repeated application of landing shocks. a) Before disturbance; b,c,d) 1st, 2nd and 3rd disturbances, respectively; 1,2,3) Numbers assigned to animals. A) Pulse frequency, beats/min; B) respiratory frequency, per minute; C) atrioventricular block; D) arrhythmia; E) extrasystolia, fibrillation; F) fibrillation; G) minutes.

In addition to the functional disturbances, considerable damage to the internal organs occurs in dogs when they are subjected to landing overloads of great magnitude. As in the small animals, the lungs, liver, intestine, heart, etc., were most frequently injured in the dogs. As a rule, more serious damage was detected on autopsy after repeated disturbances, and an intensification of the destructive and dystrophic

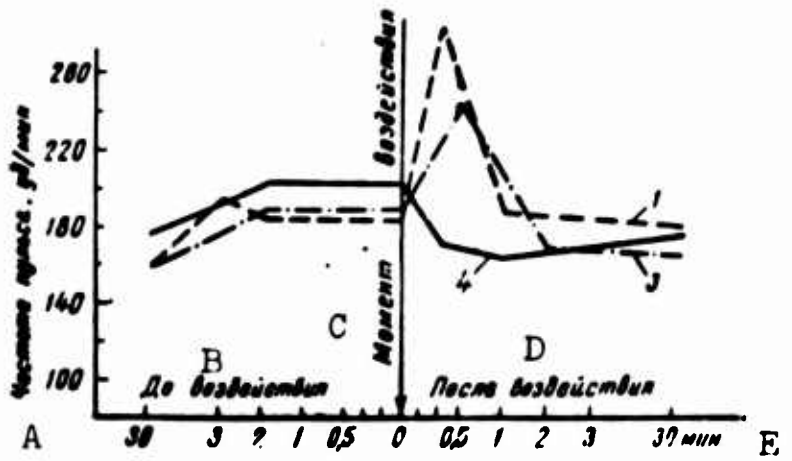


Fig. 3. Variation of pulse frequency in dogs on repeated application of landing-shock overload. 1,3,4) Ordinal numbers of disturbances. A) Pulse frequency, beats/min; B) before disturbance; C) time of disturbance; D) after disturbance; E) minutes.

changes in the brain matter was observed on histological examination.

Experiments performed on cats, rabbits and mice showed that damage to the lungs arises as a result of transmission of the impact pressure that arises in the abdominal cavity [8, 9]. Severe injuries to the internal organs are noted for g-forces greater than 300 units and descent rates of 6-8 m/sec. The hypothesis has also been advanced [10] that smaller animals perish after impact overload not as a result of hemorrhaging or injuries to the organs, but from brain concussion – a judgment based on the observation that at the moment of impact, the animals entered convulsions not related to the gravity of the macroscopic injuries observed on autopsy. Histological examinations also support the possibility of brain concussions having occurred as a result of the landing overload [6].

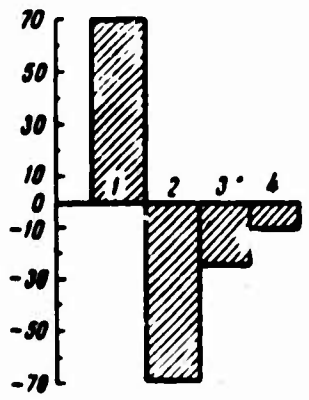


Fig. 4. Magnitudes of relative pulse-frequency changes during first 10 sec after repetition of landing-shock overloads, in dogs. The difference between the pulse frequencies before and after the disturbances is plotted against the axis of ordinates.

Of no lesser importance for the condition of the organism are the disturbances to the heart function (extrasystolic arrhythmia, atrioven-

tricular block, ventricular fibrillation) and the drop in arterial pressure below the physiological level. Since these phenomena arise concurrently with damage to the internal organs, it is not possible fully to exclude their secondary origin. Nevertheless, there is reason for ascribing all of these disturbances to the direct effect of the impact overload on the tissues of the organs of the body, including the heart muscle. Similar observations of changes in heart function after subjection to explosion waves and air blasts, in which there was no macroscopic injury done to the organs, were reported in Reference [11]. It would appear that the reflexogenic region of the heart and major vessels is of prime importance in the genesis of the derangements described here. To ascertain the importance of reflexes from the region of the heart during subjection to g-forces, we performed experiments on rats in which they received measured blows on the chest. Here, together with other deviations, the electrocardiogram showed brachycardia, atrioventricular block and polytopic ventricular extrasystoles, i.e., a pattern similar to that observed when the impact overload acts on the entire organism emerged.

The material cited above represents testimony to the complexity of the mechanisms by which the physiological functions are disturbed when landing-impact overloads are applied to the organism and the necessity of devoting penetrating study to them. Specific recommdations for ways and means to ensure the safety of landing humans in one or another space vehicle can be developed on this basis.

Received
3 February 1964

REFERENCES

1. M.D. Cassidi and P.I. Sullivan, Aerospace Engineering, 20, 3, 38, 1961.
2. R.R. Hesberg, Publ. Nat. Acad. Sci. Nat. Res. Council, 913, 66, 1961.
3. Y.A. Holcomb, Ballistic Missile and Space Technol., 1, 107, 1960.
4. Aviat. Week and Space Technol., 78, No. 21, 69, 1963.
5. S.A. Gozulov, Voyenno-meditsinskiy zhurnal [Journal of Military Medicine], No. 1, 62, 1962.
6. M.I. Kas'yanov and G.P. Mirolyubov, Collection entitled "Aviatsionnaya i kosmicheskaya meditsina" [Aviation and Space Medicine], Izd-vo AMN SSSR i Fiziologicheskogo ovshchestva [Publishing House of the Academy of Medical Sciences USSR and the Physiological Society], 1963, page 236.
7. B.M. Fedorov, Vliyaniye nervnoy regulyatsii na aritmii serdtsa [Influence of Nervous Regulation on Arrhythmia of the Heart], Izd-vo AMN SSSR [Publishing House of the Academy of Medical Sciences USSR], 1963.
8. R.F. Rushmer and Y.L. Green, J. Aviat. Med., 17, 84, 1946.
9. M.C. Donald, C. Kelly and R. Rey, J. Aviat. Med., 3, 92, 1948.
10. A.I. Gold, H.B. Hance, M. Kornhauser and R.W. Lauton, Aerospace Med., 33, No. 2, 204, 1962.
11. C.F. Clemendson, J. British Interplanet Sci., 17, No. 9, 279, 1960.

UDK 551.521.6

DISTRIBUTION CHARACTERISTICS OF ELECTRONS WITH ENERGIES AROUND 100 kev AT MODERATELY HIGH ALTITUDES ABOVE THE EARTH

V.V. Temnyy

A volume of data on the distribution of electrons with $E_e > 50$ kev and protons with $E_p > 50$ Mev trapped by the earth's magnetic field was obtained prior to 9 July 1962 in the region $L \leq 2$ by the "Kosmos-3" and "Kosmos-5" satellites [1-3]. Here the electron fluxes in this region were also observed at points at which the field strength B at fixed L exceeds the value of the field strength B_1 at an altitude of 100 km over the earth's surface in the South Atlantic anomaly at the same L. When the mean levels of captured-electron intensity were plotted in B, L-coordinates, the scattering of the experimental points with respect to these levels was considerably higher than the corresponding scatter obtained for the high-energy protons [4]. This indicates that the scatter obtained for the electron fluxes is not a result of inaccuracy of the B, L-coordinates employed [5]. If we map levels of equal intensity in a three-dimensional coordinate system, taking the geographic longitude λ into account (B, L and λ), it is found that the scatter of the experimental values decreases substantially as compared with that obtained in the two-dimensional representation (B and L without λ). Strictly speaking, the third coordinate should have been λ_m – the geomagnetic longitude. For the case under consideration, however, the discrepancy between λ_m and λ does not change the picture substantially.

Let us examine the positions of the experimental surfaces of equal electron intensity that we have obtained in the (B, L, λ) system. For intensities lower than $2 \cdot 10^7$ particles $\cdot$ cm$^{-2} \cdot$ sec^{-1}, we observe the following pattern over the entire region $L \leq 2$: as we move toward the east in the direction of electron drift from $\lambda \sim 0^\circ$ to $\lambda \sim 120^\circ$ of western longitude, the field strength on surfaces of equal intensity increases progressively as we move toward the east in the direction of electron

drift from a minimum at $\lambda \sim 0^\circ$ (B_{min}) to a maximum at $\lambda \sim 120^\circ$ of western longitude. As we move further in the same direction, the field strength diminishes progressively from the maximum at $\lambda \sim 120^\circ$ of western longitude to a minimum at $\lambda \sim 0^\circ$. No substantial dependence of the resulting shift on local time could be detected. Given adiabatic electron drift, the field strength B should be constant on surfaces of equal intensity. The picture obtained can be accounted for, for example, by the presence of a constant electric field that shifts the captive-electron mirror points into a region with higher field strength in the longitude interval from 0° to 120° of western longitude and into a region with lower strength in the interval from 120° western longitude to 0°. Here, the fluxes registered with $B > B_{min}$ may remain captive after transit of the South Atlantic anomaly.

Another possible cause of the variation of the positions of equal-intensity surfaces with longitude may be a mechanism that lowers the captive-electron mirror points in a region with high field strengths. In our motion toward the east from $\lambda \sim 0^\circ$ to $\lambda \sim 120^\circ$ of western longitude, the shift of the surface toward larger B can be accounted for in terms of predominance of the electron-influx process over the absorption process in these regions, while the shift toward smaller B for λ from $\sim 120^\circ$ of western longitude to 0° would be due to preferential absorption of electrons in this region.

The particle flux escaping the capture region in drift around the earth can be evaluated from data on the longitude variation of the intensity with L and B fixed. It follows from the distributions obtained for L = 1.6, the flux intensity at the bottom of the line of force has diminished by about $8 \cdot 10^6$ particles $\cdot$ $cm^{-2} \cdot sec^{-1}$ after one drift revolution. If it is assumed that most of the electrons registered have energies of 100-150 kev, as has been confirmed experimentally in many cases [6, 7] and is consistent with the spectrum of the electrons in the inner belt according to data up to 9 July 1962 as given in [8], then the estimates given indicate that for electrons of these energies the average flux escaping from the capture region to the maximum-loss region (on meridians from 120° of western longitude to 0°) comes to approximately $\bar{I}_t \sim 1.5 \cdot 10^3$ particles $\cdot$ $cm^{-2} \cdot sec^{-1}$. The rate of this "precipitation" varies substantially with longitude.

From a known number of escaping electrons, we can determine the quantity τ – the "capture region replenishment time." For the given $\bar{I}_t$

and a known flux intensity in the equatorial region, the quantity τ amounts to approximately $3 \cdot 10^6$ sec. For other L in the region $L \leq 2$, the values of I_t and τ are of the same order as for L = 1.6.

Thus, it can be affirmed that at low altitudes above the surface of the earth in the region $L \leq 2$ there exists a substantial longitude variation in the positions of surfaces of equal intensity with respect to the field strength B.

The distribution described differs from that obtained for adiabatic electron motion at high altitude.

In conclusion, the author expresses his gratitude to V.I. Krasovskiy and Yu.I. Gal'perin for their discussion of the results obtained.

Received

20 July 1964

REFERENCES

1. Yu.I. Gal'perin and V.I. Krasovskiy, Kosmich. issled [Cosmic Research], 1, No. 1, 126, 1963.
2. V.I. Krasovskiy, Yu.I. Gal'perin, N.V. Dzhordzhio, T.M. Mulyarchik and A.D. Bolyunova, Kosmich. issled., 1, No. 1, 132, 1963.
3. V.V. Temnyy, Kosmich. issled., 1, No. 1, 139, 1963.
4. V.V. Temny, Proc. V Intern. Space Sci. Symp. North-Holland Publ. Compl. Amsterdam (in preparation).
5. C.E. McIlwain, J. Geophys. Res., 66, 3681, 1961.
6. T.M. Mularchik and O.L. Vaisberg, Proc. V Intern. Space Sci. Symp. North-Holland Publ. Comp. Amsterdam (in preparation).
7. V.I. Krasovskiy, Yu.I. Gal'perin, V.V. Temnyy, T.M. Mulyarchik, N.V. Dzhordzhio, M.Ya. Marov and A.D. Bolyunova, Geomagnetizm i aeronomiya [Geomagetism and Aeronomy], 3, No. 3, 408, 1963.
8. W.L. Imhof, R.W. Smith and P.C. Fisher, Proc. III Intern. Space Sci. Symp. North-Holland Publ. Comp. Amsterdam, 1963, page 438.

Manuscript Page No.	[Transliterated Symbol]
259	т = t = teryayemyy = escaping, being lost

DISTRIBUTION LIST

DEPARTMENT OF DEFENSE	Nr. Copies
HEADQUARTERS USAF	
AFNINDE	1
ARL (ARB)	1
AFCIN-4	1
OTHER AGENCIES	
AEC	2
ARMY (FSTC)	3
ATD	2
CIA	1
DIA	4
NAFEC	1
NASA (ATSS-T)	73
NAVY	3
NSA	6
OAR	1
OTS	2
PWS	1
PGE (Steensen)	1
PGE	15
RAND	1
AFCRL (CRXLR)	1

MAJOR AIR COMMANDS	Nr. Copies
CONAC (CIPO-D)	2
DDC	20
AFSC	
SCFDD	1
TDBDP	10
TDBTL	5
TDGS	1
SSD (SSFAR)	2
TDBXT	1
TDC	1
TDF	3
TDEWG	2
TDEWT	1
TDFS	5
TDEEC (MINSHALL)	2
AEDC (AEY)	1
AFETR (MTW)	1
AFFTC (FTF)	1
AFMDC (MDF)	1
AFWL (WLF)	1
AMD (AMRF)	1
APGC (PGF)	1
ASD (ASFA)	25
BSD (BSF)	1
ESD (ESY)	1
RADC (RAY)	1
TDB	1