AEROSPACE RESEARCH IN BULGARIA

Volume 22, 2008, Sofia
Space Research Institute
Bulgarian Academy of Sciences

Address
AEROSPACE RESEARCH IN BULGARIA
Space Research Institute
6, Moskovska St., Sofia 1000, Bulgaria

e-mail: office@space.bas.bg

Language Editor
Lubomira Kraleva

Pre-Publication Processing
Tsveta Srebrova

ISSN 1313 - 0927

Bulgarian Academy of Sciences ISSN 1313 – 0927

Aerospace Research in Bulgaria

22
Sofia, 2008

Contents

СЪДЪРЖАНИЕ

Bulgarian Academy of Sciences. Space Research Institute.
Aerospace Research in Bulgaria. 22, 2008, Sofia

THIN VISCOUS ELLIPTICAL ACCRETION DISCS WITH ORBITS SHARING A COMMON LONGITUDE OF PERIASTRON III. NUMERICAL EVALUATIONS OF THE VALIDITY DOMAIN OF THE SOLUTIONS

Dimitar Dimitrov

Space Research Institute - Bulgarian Academy of Sciences

Abstract

We have investigated the validity domain of the dynamical equation which defines the structure of a two-dimensional elliptical accretion disc model of Lyubarskij et al. [9]. Only cases with integer powers in the viscosity law $\eta \sim \Sigma^n$ are considered, namely n = -1, 0, +1, +2 and +3 (η is the viscosity coefficient, Σ is the disc surface density). This approach is adopted in view of the fact that the analytical expressions for the dynamical equation for these particular values of n are already derived in an earlier paper [7]. As a mathematical problem, we have to solve a second order ordinary differential equation with initial conditions – two arbitrary constants e_0 (the value of the eccentricity) and its derivative $\dot{e}_0$ for a given fixed value of the focal parameter p_0 of a selected elliptical trajectory. In the present paper we have chosen the following grid of values: e_0 = 0.00, +0.20 and +0.50 ; $\dot{e}_0$ varies by step 0.01 accordingly from –1.00 to +1.00, from –0.80 to +1.20 and from –0.50 to +1.50. The independent variable u in the dynamical equation is defined as a logarithm of the focal parameter p of the elliptical particle trajectories, i.e., $u \equiv \ln p$. Respectively, $e = e(u ; e_0, \dot{e}_0, n)$ and $\dot{e} = \dot{e}(u ; e_0, \dot{e}_0, n)$. By the definition of the problem, each eccentricity e must be a real function and from physical reasons the inequalities $| e(u) | < 1$, $|\dot{e}(u)| < 1$ and $| e(u) - \dot{e}(u) | < 1$ must be satisfied for every 3-tuple of parameters (e_0, $\dot{e}_0$, n). Then the dynamical equation is solved by means of numerical methods and the range of variation of u where the above restrictions are satisfied is found out. For each of the 15 combinations (n, e_0) the permitted range of variation of u as a function of $\dot{e}_0$ is presented graphically.

Introduction

Accretion flows in astrophysics are widely occurring phenomena. They are established for certain to exist around protostars, accreting compact objects in binary stellar systems, and also around supermassive black holes at the central regions of galaxies. In other words, configurations taking the form of flattened discs are commonly observed in astronomy. When accretion discs are highly flattened, the material in these objects typically orbits in circles, such that the centrifugal force to a first approximation balances the gravitational attracting force. The latter is often due to some external gravitational potential into which the matter has fallen, but can also (in part or in whole) be caused by the self-gravity of the material itself. Standard accretion disc theory describes the viscous evolution of flat circular discs around a single object (ordinary star, white dwarf, neutron star or black hole). However, accretion discs are often distorted in various ways. A disc around a single star may be composed of non-coplanar or elliptical orbits of its particles, resulting in a warped or eccentric disc, respectively. In addition, discs in binary and protoplanetary systems are tidally distorted by the presence of the orbiting companion. To describe of these phenomena, during the recent decade semi-analytical methods are developed and applied . They bring us nearer to the solution of these essentially three-dimensional, nonlinear problems [1]. Consequently, the simplest disc models are those whose disc elements follow circular orbits around a central potential in the same plane. But due to interactions with external forces, the discs are often neither coplanar nor circular.

It should be noted as well that a certain amount of work has also been done on eccentric and non-planar particle discs around planets. In this connection, it may be noted that the situation with regard to fluid discs is much less clear. Narrow planetary rings are often observed to be non-circular, exhibiting sometimes $m = 1$ (eccentric ring) or $m = 2$ distortions. This phenomenon can occur because the action of viscosity through the shear can destabilize non-axisymmetric density waves (viscous overstability). Another possible cause is parametric instability caused by a tidal interaction with a companion satellite. Such situations may also occur in larger in larger scale discs. There, viscosity is again capable to excite eccentricity rather than to cause it to decay. It makes sense to postulate that a combination of tides and parametric instability is causing disc eccentricity in dwarf novae stars, inspite of the circumstance that the conditions required for these instabilities (they depend on the viscosity law) are not understood sufficiently well at present. It is worth to note that an opportunity for cross-comparison between planetary ring dynamics and gaseous accretion discs is a promising perspective for better understanding of the

conditions required for viscosity-driven non-axisymmetric distortions to occur in discs. The preliminary results obtained at this stage are not essentially new but serve as an important test for disc models. This stimulated the collaboration between astronomers whose aim is studying the properties of eccentric discs in binary stellar systems.

Discs can become non-planar and/or eccentric due to interaction with the external constraints. For example, in the protostar field, most stars have binary companions and they form in crowded regions. Thus, protostellar discs are likely to be subject to interactions with companions or even with passing interlopers. It is also possible that when two stars come together to form a binary star, there can be an exterior protostellar disc surrounding the binary system as a whole. Dynamically, the impact of the stars affects the discs, but there is also a response reaction of the disc on the orbits of the stars themselves. It was shown, however, that even for the most compact star-forming environments known, disc penetrating encounters leading to capture of the star would be rare, with the probability of such an encounter being approximately 10 % over a disc's lifetime. But if the mean free-path in the star-forming core is sufficiently small so as to allow close encounters between young stellar objects on a time scale that is shorter than the lifetime of a protostellar disc, then the tidal interaction between a secondary object and the disc of the primary may nonetheless be a dynamically significant event in the disc's history.

S. Goodchild and G. Ogilvie have applied an eccentric accretion disc theory in a simplified form to the case of superhumps in SU Ursae Majoris cataclysmic variables and other similar systems [2]. In their model the disc contains 3:1 Lindblad resonance and it is believed that this resonance leads to growth of eccentricity and a modulation in the light curves due to the interaction of precessing eccentric discs with tidal viscous damping of eccentricity. These authors first worked out the theory in the simple case of a narrow ring and they arrived at the conclusion that the eccentricity distribution is locally suppressed by the presence of the resonance, creating a dip in the eccentricity at the resonant radius. Application of this theory to the case of superhumps in close binary stars confirms observationally this conclusion. It also gives possibilities for more accurate expressions for the precession rates of the discs than has been previously accomplished by means of simple dynamical estimates.

Time-resolved spectroscopy of the nova-like variable UU Aquarii is analyzed with eclipse mapping techniques to obtain spatially resolved spectra of its accretion disc and gas stream as a function of the distance from the disc centre in the spectral range $3600 \div 6900$ Å. The interpretation of the eclipse maps suggests that the

asymmetric structure in the outer disc, which was previously identified as a bright spot, may be a signature of an elliptical disc similar to those which probably are present in SU Ursae Majoris stars during outbursts [3]. The possibility that protoplanetary gaseous discs are dynamically unstable to axisymmetric and non-axisymmetric gravity perturbations is studied by E. Griv [4]. This investigation includes also perturbations that are produced by spontaneous disturbances. The analytical treatment of these phenomena (with characteristic scales larger than the vertical scale-height) is realized by using a local Wentzel-Kramers-Brillouin (WKB) approach. This paper reveals an interesting connection between non-axisymmetric accretion discs and their clumpy structures, which are gravitationally bound. The latter can collapse to become giant planets. There is an important feature of the planetary formation process. Namely, gravitationally unstable non-axisymmetric (e.g., spiral) perturbations can effectively transport both the angular momentum and the mass in a spatially inhomogeneous disc. Another investigation of the action of gravitational perturbers in thin cold astrophysical discs is carried out in paper [5]. Two types of density structures are found, depending on the mass of the perturbing body and on the amount of momentum transport in the disc. A gap around the whole circumference of the disc is opened if the perturber is more massive than a certain threshold. In the opposite case, a local S-shaped density modulation is generated (the so called "propeller"). It was found that the large-scale appearance of the "propellers" does not depend on the details of the scattering process but mainly on the effective strength of gravity perturbations. The solutions of the problem show that the characteristic spatial extentions of the structures depend on the mass of the perturber and the viscosity of the disc. This conclusion is in agreement with the already known result of the theory that the tidal torques exerted by a binary system on an accompanying disc may limit its spatial scope. Examples of such occurrences are an accreting protoplanet that is embedded in the nascent disc and which may clear an annulus about its orbital path, or a massive binary system which may maintain a cleared cavity within a circumbinary disc. The disc response may be studied as a function of binary mass ratio, separation between stars, inclination of the disc with respect to the orbital plane and the disc thickness. Theoretical considerations demonstrated that differentially rotating discs can precess approximately as rigid bodies provided the precessional timescale is sufficiently long when compared with the sound-crossing timescale for the disc. For self-gravitating disc models it was established that the usual condition for instability to non-axisymmetric perturbations is insufficient to ensure fragmentation and a higher level of instability, than is usually considered, appears to be required.

Numerical and analytical studies of accretion discs are motivated through reference to recent observational and theoretical results. In particular, in this respect the considerations connected with the new observational investigations on binarity among young stars should be mentioned. The analysis of nearby star-forming regions has revealed that the young stellar objects have a binary frequency in excess of that found for field stars. Roughly half of the current sample of these binary stars are associated with optically thick, geometrically thin, circumstellar accretion discs. It was found that the peak in frequency distribution over binary separation occurs at about 30 a.u., which is less than the typical disc sizes, which are about 100 a.u. These conclusions are based on the observational studies of nearby star-forming regions. Rotation of the accretion disc matter provides centrifugal support within a preferred plane which gives a geometrically thin circumbinary/circumstellar disc that is pressure supported perpendicular to that plane. In an early epoch the surrounding envelope of matter is cleared through the action of the stellar wind, giving exposed young stars/star and circumbinary/circumstellar disc, respectively.

Non-planar (or warped) accretion discs might be set up by the tidal forcing due to an inclined-orbit binary. Evidence of the non-planarity of binary orbits and disc planes is provided by the low inclinations of gas giants in the Solar system. Further, a small asymmetry in the inner regions of the dusty disc associated with β Pictoris has been interpreted as a warp generated by possibly a Jupiter-mass body on an inclined orbit embedded into the interior to the disc. Asymmetries found on the scale of the disc in this binary system might also have been caused by a stellar encounter.

Physical restrictions and features of the considered accretion disc model

In binary stellar systems the disc matter is continuously replenished by the donating companion through the mass transfer stream, which strikes the outer parts of the disc and causes a "hot spot" event. The latter is not expected to significantly affect the accretion disc dynamics. But if the rate of mass transfer is insufficient to maintain the disc temperature above the ionization temperature of hydrogen (i.e., 6000 K), then the disc can switch from an optically thin state to an optically thick state. This leads to thermal instability and outburst behaviour when the accretion rate may increase by two orders of magnitude and the disc may expand to fill the tidal radius of the compact object on time scales of weeks to months. In our further considerations we fully exclude such an obvious non-stationarity and concentrate only on the stationary accretion discs. It is also preferable to exclude explicitly

the detailed treatment of the thermal processes because the emission processes are always found to be specific to particular astronomical objects, and give little opportunity, if any, for cross-field fertilization. This remark is essential to be made, because our model of the accretion disc is not so detailed to be compared immediately with the observations. However, attention must be paid to these parts of the disc where the energy is dissipated as this can have feedback into the disc structure and thus affect the angular momentum transport.

Although astrophysical discs come in a variety of forms and have a wide range of length scales changing from planetary to galactic sizes, there can be a lot of common ground between them because of similar underlying dynamical properties. For this reason, research problems in the area are grouped together according to similarity of dynamical properties, rather then under astronomical field. Consequently, the investigation efforts may include researches in the general areas of hydrodynamics, magnetohydrodynamics, stability of self-gravitating differentially rotating systems and the theory of turbulent dissipative fluid systems as they relate to astrophysical discs. Both analytic approaches and numerical simulations are to be involved. In the present study we shall use analytical results about elliptical accretion discs obtained for particular samples of the viscosity law. These are already published in an earlier series of papers [6], [7], and [8], where the expressions of the dynamical equation (governing the structure of the disc) are given in explicit form. Here we continue the investigation of their properties. More specially, we concentrate our attention on the determination of the domain where the solutions are not only mathematically well defined, but also have reasonable physical meanings. Because of the complexity of the analytical expressions, the methods for solving the equations are purely numerical.

Another simplification of the accretion disc model considered by us should be mentioned. In fact, this is the elliptical disc first investigated by Lyubarskij et al. [9], where the internal torque is derived consistently from the basic fluid-dynamical equations within the context of the α-theory. The latter means that the assumption that effective dynamic viscosity is proportional to local pressure is valid for the entire disc. But large-scale electromagnetic fields are not included into consideration. This situation may not always correspond to the real objects in nature. Gaseous discs are observed around compact components in interacting binary stellar systems and around very young stars, which are still in the process of forming. They also occur around massive black holes in active galactic nuclei. Such discs consist of highly ionized plasma, therefore in them magnetic, as well as purely hydrodynamic processes, have to be taken into account. The effect of the magnetic field of surface

strength $B \approx 1$ T (over the surface of the secondary star) on the accretion disc structure is examined by Pearson [10]. The theory of precessing discs and resonant orbits is generalized to encompass resonances of higher order than 3:2 and is shown to retain consistency with the new mass ratio of the star AM Canum Venaticorum. Such investigations show that neglecting the magnetic field in the accretion disc model of Lyubarskij et al. [9] considered in the present paper may have additional difficulties if we try directly to compare it to the observational data. As mentioned also by Shalybkov and Rüdiger [11], the presence of an imposed vertical magnetic field may drastically influence the structure of thin accretion discs. If the field is sufficiently strong, the rotation law can depart from the Keplerian one. Of course, this situation is very different from that which is described by the dynamical equation derived in [9].

Neglecting the three-dimensional structure of the accretion flows may have important consequences. When the latter cannot cool via emission of radiation, they become vertically thick and nearly spherical. Thus, they are intrinsically multidimensional. Let H be the half-thickness of the disc at distance R from the center of the compact star. If the accretion disc is thick, so that $H \approx R$ and the Mach number $M \approx 1$, then all the thin disc approximations break down simultaneously. In particular, we may no longer neglect the radial pressure gradient, nor may we assume that the temperature gradients $\partial T/\partial z >> \partial T/\partial R$, where z is a coordinate perpendicular to the central disc plane. We also cannot neglect the heat flow in the radial direction. Thus, the accretion disc is no longer Keplerian, and the splitting of the solution of the disc structure into a radial part and a part perpendicular to the disc plane is no longer valid. The above remarks concerning the three-dimensional nature of accretion flows are valid not only for gaseous matter but also for dust composed discs. Such accretion discs exist around protostellar objects and are almost certainly self-gravitating during the early part of their lives, although at later times the central star dominates the gravitational potential. Albeit a dust shell may be located at quite a distance from the central binary star (up to about 1000 a.u.), it can be viewed as a remnant of a tidally-disrupted accretion disc. This seems to indicate that the disc and the orbits of the dust particles are not necessarily coplanar in the pre-main sequence stellar systems. Similar difficulties in simulations of accretion flows arise when clumping of the matter has to be taken into account. The formation of collapsed objects during the clumping of the material in a circumstellar discs is very hard to model numerically because of the huge range of spatial scales which are involved in the problem. Numerical experiments also demonstrate that failure to resolve the vertical structure of discs leads to greate errors in the numerical

solution for the evolution of the entire physical system. While the two-dimensional simulations allow much higher linear resolution in the two dimensions which they actually model, they do so at the cost of requiring that some assumptions must be made about the character of the system and its behavior in the third dimension. In the elliptical disc model of Lyubarskij et al. [9] investigated by us, where in fact an analytical description is used, the two-dimensional approach ensures that the final dynamical equation is a second order ordinary differential equation. This simplification of the problem is accompanied by the assumption of existence of local hydrostatic equilibrium at each point of the disc in the vertical z-direction.

But even when the third dimension is "eliminated" by means of appropriate assumptions about the vertical structure of the disc (i.e., it is not included explicitly in the dynamical equation), there are, of course, other inhomogeneities in the two-dimensional disc plane which complicate the analytical and numerical analysis. In the elliptical disc of Lyubarskij et al. [9] one such global inhomogeneity, namely the disc ellipticity, is incorporated into the description of disc equations. In that model global structures like spirals not taken into account. For example, the latter can arise due to shock waves, but they are not included into the model [9] because the quantities (e.g., the determinant of the metric, the introduced auxiliary functions, etc.) which are used, become divergent in that case. Consequently, the derived formulae in [9] are not appropriate to describe these complicated situations despite their closer similarity to the reality. This circumstance is, unfortunately, an additional reason obstructing the comparison "theoretical model – observations". The same is the case with the conclusion that the spirals contribute by about 16 and 30 per cents to the total line flux, respectively, for the He and CN4650 emission lines [12]. It is interesting to note that the comparison of the Doppler and eclipse maps for IP Pegasi reveals that the Keplerian velocities derived from the radial position of the spiral shocks are systematically larger than those inferred from the Doppler tomography. This indicates that the gas in the spiral shocks has sub-Keplerian velocities – another distinction from the model considered in the present paper.

The difficulties concerning the description of accretion disc structure become even more trouble some when the fragmentation into local density blobs must be included. These problems are by necessity solved by means of numerical methods. As computers become more powerful, the studies previously limited only to local patches of the accretion flows may be now expanded into global studies that encompass the entire disc. But in this area of investigations much must be done to overcome various obstacles. In particular, from the numerical simulations it follows that an error in the cooling rate as large as a factor of two will be sufficient to

suppress or enhance the fragmentation in a simulation, which would otherwise follow a very different evolutionary path [13]. An error of factor of two in the cooling rate will be equivalent to an error of ≈ 20 % in the photosphere temperature, due to the $T_{\rm eff}^{\ 4}$ dependence of the cooling on the temperature.

It would be stressed that a description of a real physical system (more specially, when it must be compared with the observational data) requires not only analytically or/and numerically valid simulations, but also relevant initial conditions and a correct and complete physical model. As mentioned by A. Nelson: "...Numerically valid simulations will be interesting only to the extent that they model real systems with real physics. The important point to take from these discrepant results is that not only do questions of numerical validity remain to be addressed, but also the physical models." [13]. Definitely, such conclusions remain valid also for analytical treatment of the problems, which concern the dynamics and structure of accretion discs.

Concrete values of the parameters and numerical evaluations of the validity domains for disc dynamical equation solutions

Briefly summarizing the results already obtained in the previous papers [6] and [7] about the accretion disc model worked out by Lyubarskij et al. [9], we shall remind that the viscosity coefficient η is related to the disc surface density Σ via the relation $\eta = \beta \Sigma^n$. Here β is a constant and the power n may, generally speaking, vary continuously within an interval with upper and lower boundaries determined on the basis of physical reasons, corresponding to the accepted accretion disc model. It is worth to note that the description of the disc given by Lyubarskij et al. [9] is an extension of the standard α - disc theory of Shakura and Sunyaev [14] to the case of accretion flows with elliptical orbits of the particles around the central star. In view of the fact that we have obtained earlier explicit forms of the dynamical equation only for five values of the power n [7], we limit our results in this paper also only to these particular values of n. Namely, our investigation of the validity domains is for the following fixed values:

n = -1, 0, +1, +2 and +3.

The dynamical equation determining the structure of the considered stationary accretion disc [9] is a second order differential equation. Hence, its solutions depend on two arbitrary integration constants. It is appropriate to choose as such constants the values of the eccentricity e and its derivative $\dot{e}$ at a given value u_0

of the argument. Each ellipse is described by its eccentricity e and focal parameter p. According to Lyubarskij et al. [9], it is plausible to make a transition to a new variable $u \equiv ln\, p$. Having in mind that for a circular orbit p is equal to the radius r of the orbit, the above substitution simply means that we have introduced a logarithmic scale to characterize the spatial size of the elliptical orbits of disc particles. Lyubarskij et al. [9] allow variations of u in the interval from $u = -3.0$ (innermost orbits of the disc) to $u = +3.0$ (outermost orbits of the disc). We shall keep these limits in the present study. The intermediate value $u = 0.0$ corresponds to $p = +1.0$. We have accepted this value as a point where the initial conditions are fixed also. Therefore, we introduce the following notations: $p_0 = +1.0$, $u_0 \equiv ln\, p_0 = 0.0$, $e_0 \equiv e(u_0)$ and $\dot{e}_0 \equiv \partial e(u)/\partial u \,|_{u=u0}$. In the next we have selected to investigate the solutions of the dynamical equation for three different values of e_0 for every fixed value of n; namely, $e_0 = 0.0$, $+0.20$ and $+0.50$. This choice ensures 15 combinations corresponding to the above representative discrete sample of values for n and e_0. In connection with this lattice of values, an important remark must be made on the signs of the eccentricity $e(u)$ and its derivative $\dot{e}(u) \equiv \partial e(u)/\partial u$ (in particular, this concerns e_0 and $\dot{e}_0$). By its way of deriving and by the formulation of the task itself (all particles' trajectories are postulated to be ellipses which apse lines are in line with each other), the dynamical equation has a symmetry in respect to the change of the sign of the eccentricity $e(u)$. The latter operation automatically does so with the signs of the derivatives $\dot{e}(u)$ and $\ddot{e}(u)$. In other words, if the periastron of an ellipse lies on the positive part of the abscissa, then its mirror ellipse with a periastron on the negative part is also a solution of the dynamical equation. This symmetry property can easily be seen from the explicit forms of the considered equation given for integer values of the power n [7]. The just above-mentioned characteristic must not be misleadingly identified with the possibility that some ellipses may have positive eccentricities (to say, in the inner parts of the accretion disc) and negative eccentricities (to say, in the outer parts of the same disc) and vice versa. Of course, if the accretion disc undergoes a mirror transformation as a whole (i.e., all ellipses change the signs of the eccentricities and their derivatives), formally the geometry of the accretion flow changes but physically the situation is not different. For this reason, we shall not include in our examples negative values of the constant of integration e_0. Negative initial values e_0, combined with negative/positive initial values of $\dot{e}_0$ will correspond to positive initial values of e_0 (with the same absolute values) combined with positive/negative values of $\dot{e}_0$ (again with the same absolute values as the former $\dot{e}_0$). Concluding this remark upon the signs of the eccentricities, it is worth to note that the notion "negative eccentricity" was introduced by Lyubarskij et al. [9] simply to

express analytically (and graphically) in a more compact form the global changes of eccentricity through the entire disc. "Negative eccentricity" simply means that the periastron of the ellipse (which really has a positive eccentricity with the same absolute value) lies on the negative part of the abscissa. Such representation is useful when there is a transition through the disc from one type of ellipses (with $e(u) > 0$) to another type of ellipses (with $e(u) < 0$) and vice versa. Only for generality, we mention that it is also helpful to use eccentricities which have complex values [15]. But this notion makes sense when more general models of accretion discs (where the simplification that the orbits share a common longitude of periastron is not already valid) are investigated. For our purposes such a complication is only an unnecessary "exotic".

Selecting the values of the parameters n and e_0, we are ready for each couple (n,e_0) to vary the second integration constant namely $\dot{e}_0$. This variation is performed with a much denser grid of points. The step of increasing of $\dot{e}_0$ is chosen to be 0.01 and the grid covers uniformly an interval of length 2.00. The initial (lower) and, correspondingly, the final (upper) boundaries of the interval are determined on the basis of the following reasons. A priori (by the definition of the problem) the eccentricity of the orbit $e(u)$ may change continuously from –1.0 (degenerate ellipse whose periastron has negative abscissa) to +1.0 (degenerate ellipse whose periastron has positive abscissa). The intermediate case $e(u) = 0$ corresponds to a circular orbit. In the original paper of Lyubarskij et al. [9] three auxiliary functions $Y(e,\dot{e}, n;\varphi)$, $Z(e,\dot{e},n;\varphi)$ and $W(e,\dot{e},n;\varphi)$ are introduced. They are averaged over the azimuthal angle φ by integrating over φ from 0 to 2π. Then the coefficients of the dynamical equation of the elliptical accretion disc are expressed through these functions and their first partial derivatives with respect to e and $\dot{e}$. Note that in contrast to paper [9], we use the overdot symbol (˙) to designate differentiation with respect to the variable u instead of differentiation with respect to time t. Such a change is not confusing because we consider only stationary accretion flows ([6], [7] and [8]). As pointed out in [6], an explicit computing of the angle averaged coefficients of the dynamical equation in the case of integer n ($n = -1, 0, +1, +2, +3$) requires analytical evaluations of seven integrals over the angle φ, which we denote as $\mathbf{I}_0(e,\dot{e},n),\ldots,$ $\mathbf{I}_0(e,\dot{e},n)$. Their integrands contain in the denominators multipliers of the following type: $[1+(e-\dot{e})\cos\varphi]^{n+1}$ and $[1+(e-\dot{e})\cos\varphi]^{n+2}$ (see equations (5) ÷ (7) in paper [6]). When $[e(u)-\dot{e}(u)]\cos\varphi = -1$, singularities arise which cannot be cancelled out by means of the other multipliers in the integrands. To avoid this peculiar situation, it is necessary to impose the condition $|e(u)-\dot{e}(u)| < 1$, which ensures that the integration over the azimuthal angle φ will not lead to divergences of the integrals

$\mathbf{I}_0(e,\dot{e},n),\ldots,\mathbf{I}_0(e,\dot{e},n)$. If we do so, the same inequality must hold also for the particular values of the eccentricity $e_0 \equiv e(u_0)$ and its derivative $\dot{e}_0 \equiv \dot{e}(u_0)$ $|e_0 - \dot{e}_0| < 1$, i.e., $-1 < e_0 - \dot{e}_0 < +1$. Having in mind that the initial value e_0 ($e_0 = 0.00, +0.20$ and $+0.50$ in our computations) is already fixed when we begin to vary the second initial value $\dot{e}_0$, we can write: $-1 - e_0 < -\dot{e} < 1 - e_0$, or multiplying by -1, we have: $-1 + e_0 < \dot{e} < 1 + e_0$. Consequently, for our particular choices of e_0, we obtain the following intervals for variation of the parameter $\dot{e}_0$:

1. $e_0 = 0.00$; then $-1.00 < \dot{e}_0 < +1.00$;
2. $e_0 = +0.20$; then $-0.80 < \dot{e}_0 < +1.20$;
3. $e_0 = +0.50$; then $-0.50 < \dot{e}_0 < +1.50$.

We emphasize that the above restrictions on $\dot{e}_0$ ensure that the divergences generated by the averaging over φ (i.e., averaging over the full elliptical trajectory length, corresponding to an arbitrary fixed value of the focal parameter p (or $u \equiv \ln p$) are avoided. But the elimination of the angular coordinate φ from the following sequence of mathematical manipulations and, possibly, other physical simplifications, does not guarantee that new divergences would not arise. For example, we may have expressions whose denominators include quantities like the multipliers $e(u)$, $\dot{e}(u)$, $1 - [e(u) - \dot{e}(u)]$, etc., taking zero values for some values of the independent variable $u \equiv \ln p$. Fortunately, these complications, in fact, do not generate unresolvable difficulties. By means of the L'Hospital's theorem (namely, that which resolves indeterminations of the type 0/0), it may be seen that analytically the divergences, caused by the nullifications of the denominators, cancel out by the nullifications of the corresponding nominators. As a result, the expressions are not divergent. However, troubles may arise at these singular points, when the expressions are evaluated with the help of numerical methods, because of the finite step of the lattice. This circumstance has to be taken into account when the numerical approach is used.

After the choice of the three parameters n, e_0 and $\dot{e}_0$ is already done, we are in a position to vary the parameter u (i.e., the focal parameter p; $u \equiv \ln p$). Confining ourselves to the directing paper of Lyubarskij et al. [9], we limit the range of variation of u from -3.0 to $+3.0$. In this paper we have varied u by step 0.001, which enables us to obtain more than enough dense grid of solutions of the dynamical equation. The latter is a second order ordinary differential equation with already fixed initial conditions e_0 and $\dot{e}_0$, and we have solved it numerically using the standard techniques for numerical solving of such equations. As discussed above, from physical point of view we have to select only these solutions for the eccentricity $e(u)$ which satisfy simultaneously the following three inequalities $|e(u)| < 1$, $|\dot{e}(u)| < 1$ and

$|e(u)-\dot{e}(u)| < 1$. Having a grid of 6000 solutions for different values of u (reminding that n, e_0 and $\dot{e}_0$ are fixed earlier), it is not difficult to establish the boundaries within which the above conditions are satisfied. Using a linear interpolation, the precision of the determinations is about $0.0001 \div 0.00001$. This accuracy is enough, all the more we express the results graphically on Figures $1 \div 5$.

Summarizing all the parametrization of the task, which we have applied, we have chosen 15 ordered tuples (n, e_0) ($n = -1, 0, +1, +2, +3$; $e_0 = 0.00, + 0.20, + 0.50$). For each of them we have drawn two-dimensional plots in the plane $(\dot{e}_0, u)$, giving descriptive representations of the domains of validity of the dynamical equation. The regions shaded by vertical lines on figures $1 \div 5$ denote the physically meaningless areas of solutions where the above mentioned three inequalities $|e(u)| < 1$, $|\dot{e}(u)| < 1$ and $|e(u) - \dot{e}(u)| < 1$ are not satisfied. The remaining unshaded ("white") regions are the domains of validity we are seeking for in the present paper. The boarders between the "permitted" and "forbidden" zones are delimited by thick curved lines.

Discussion and concluding remarks

Each of the graphics represented in Figures $1 \div 5$ (for preliminary fixed values of n and e_0) gives a visual illustration about the set of all possible solutions for the entire (global) accretion disc. For a given viscosity law $\eta = \beta \Sigma^n$ (n is fixed ; the value of the constant β is meaningless for our consideration) and selected initial conditions e_0 and $\dot{e}_0$, the focal parameter of the elliptical trajectories of disc particles must obey the restrictions imposed according to Figures $1 \div 5$. This ensures that the eccentricities $e(u)$ and their derivatives $\dot{e}(u)$ will have physically reasonable values. Evidently, the freedom remains that between trajectories samples with different eccentricities may be present, varying from the inner to the middle and to the outer parts of the concrete accretion disc model. In principle, even if the dynamical equation is not explicitly solved, graphics like these in Figures $1 \div 5$ give to some extent an idea of what properties of the solutions $e(u)$ may be generally expected. In turn, this information will be useful for more suitable choices of the initial conditions e_0 and $\dot{e}_0$ which guarantee the uniqueness of the solution to the accretion flow.

It is evident from Figures 1a ÷ 5a that when we choose the initial conditions to be given on a circle (i.e., $e_0 = 0$; $u = 0$; $p = 1$), then the forbidden zones for u are preferably situated in the upper half of the plane $(\dot{e}_0, u)$. An exception of this property is the case $n = 0$, corresponding to the absence of functional dependence between the viscosity coefficient η and the disc surface density Σ. In order to represent more clearly the descriptive results from Figures $1 \div 5$, let us return to

the definitions of the quantities u, focal parameter p and the eccentricity $e(u) \equiv e(p)$ [9]. If a and b are the major and minor semiaxes, respectively, of a single elliptical trajectory, then its eccentricity is given by $e = (1 - a^2/b^2)^{1/2}$. Taking into account that the focal parameter p of an ellipse is expressed through a and b as $p \equiv b^2/a$, we can write $e = (1 - a/p)^{1/2}$. Accepting the new variable $u \equiv \ln p$ [9], this inversely means $p \equiv \exp u$. There is a subtle detail. The argument of the logarithmic function must be dimensionless. So we have to express the focal parameter $p \equiv b^2/a$ as a product of two multipliers. One of them has constant value equal to unity and dimension of length – let us denote it by 1. The other multiplier is dimensionless and its value is equal to the value of b^2/a. When the definition $u \equiv \ln p$ is introduced, just the second (dimensionless) multiplier, denoted again by the symbol p because we are interested in the quantitative nature of the focal parameter is implicitly subtended. Strictly speaking, we have to write the "dimensionally right" focal parameter as $p \equiv 1 \times \exp u$, and hence, the quotient $a/(1 \times \exp u)$ is thus dimensionless. Making the above note, which is caused by the omission of the dimensional multiplier 1, we are sure that this simplification of notations does not lead to any erroneous conclusions.

We also emphasize the fact that when we consider accretion discs with elliptical trajectories of the particles, both semiaxes a and b may vary for different regions of the flow. Consequently, in the general case a and b are functions of u: $a = a(u)$ and $b = b(u)$. When we write the expression for the eccentricity $e = [1 - a(u) / (1 \times \exp u)]^{1/2}$ this does not mean that we have solved the problem at all. The semiaxis $a(u)$ is an unknown function of u. Hence, the solution of the dynamical equation is an unavoidable necessity in order to find the two-dimensional structure of the accretion disc. Having in mind that both semiaxes $a(u)$ and $b(u)$ are limited by the above functions (the disc has a finite size), it is possible to use the relation $e(u) = [1 - a(u)/(\exp u)]^{1/2}$ (we omit the multiplier 1) to make some conclusions when we consider boundary constraints. Table 1 contains some values of $\exp u$ when u varies within the interval [-3.0, +3.0].

Table 1

u	-3,0	-2.0	-1.0	0.0	+1.0	+2.0	+3.0
$\exp u$	0.0498	0.1353	0.3679	1.0000	2.7183	7.3891	20.0855

It is evident from these data that $exp\ u$ changes up to 400 times from its minimal to its maximal value in that interval. Roughly speaking, when u is "small" (i.e., u has negative values close to –3.0), the quotient $a(u)/exp\ u$ is larger in comparison to the case when u is "large" i.e., u has positive values close to +3.0). Correspondingly, in the first case the difference $1 - a(u)/exp\ u$ will be smaller then in the second case. In other words, it may be said to some naive extent, that the positive values of u suggest more elongated elliptical trajectories and the negative values of u are associated with orbits closer to circles.

Returning to Figures 1a ÷ 5a, only in the above sense we are able to conclude that except the case $n = 0$, the other cases $n = -1, +1, +2$ and $+3$ are more suggestive for a set of orbits which are not very elongated. Relative to the systems of boundary conditions given on "weakly" elongated ellipses ($e_0 = +0.20$; Figures 1b ÷ 5b) and on "strongly" elongated ellipses ($e_0 = +0.50$; Figures 1c ÷ 5c), this deduction does not alter essentially when $n = +1, +2$ and $+3$. Some differences appear for $n = -1$ and 0. There are practically not upper limits for u in the considered ranges of variation of $\dot{e}_0$. Consequently, such accretion discs probably contain also a larger number of high eccentricity orbital trajectories of their particles.

Another interpretation of the graphics in Figures 1 ÷ 5 may be given. By the definition of the task, the eccentricity $e(u\ ; e_0, \dot{e}_0, n)$ must be a real function of u for all allowed values of the parameters e_0, $\dot{e}_0$ and n, This requirement was already taken into account when the dynamical equation was solved numerically in order to determine its validity range. Therefore, the permitted (unshaded) regions in Figures 1 ÷ 5 already take into consideration this restriction. This means that $1 - a(u)/p \geq 0$, or $a(u) \leq exp\ u$. Taking a logarithm from the latter inequality, we obtain $ln\ a(u) \leq u$. Consequently, it is possible to give a new interpretation of the results presented in figures 1 ÷ 5. They may be considered as restrictions on the major semiaxes of the particle trajectories. Such limitations are imposed under definite values of the parameters n, e_0 and $\dot{e}_0$.

In conclusion, it should be mentioned again that despite the use of the analytical forms of the dynamical equation, it was solved by means of numerical methods. For this reason we do not give denser grid with respect to the parameter e_0. Moreover, we have not available analytical expressions for that equation in the cases where the power n is not integer. Nevertheless, Figures 1 ÷ 5 provide a general picture of the validity range of the dynamical equation.

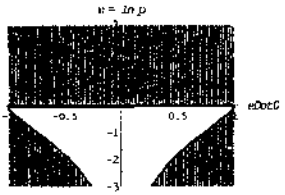


Fig. 1a. Case n = -1 ; e_0 = 0.00

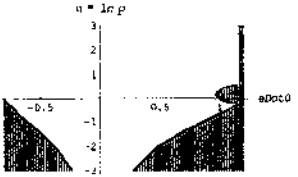


Fig. 1b. Case n = -1 ; e_0 = + 0.20

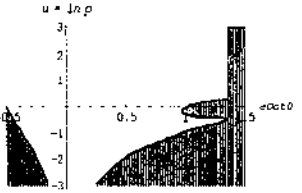


Fig. 1c. Case n = -1 ; e_0 = +0.50

Fig. 1. Permitted range of variation of $u \equiv \ln p$ (p is the focal parameter of the elliptical trajectory of a single disc particle) as a function of e_0. The vertically shaded regions are forbidden zones, which cannot be occupied by the focal parameters of the disc orbits. The power in the viscosity law $\eta \sim \Sigma^n$ is $n = -1$; e_0 is an initial value of the eccentricity (i.e., an arbitrary integration constant), which we choose to take three fixed values.

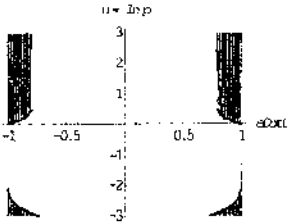


Fig. 2a. Case n = 0 ; e_0 = 0.00

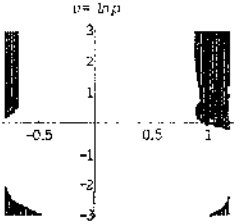


Fig. 2b. Case n = 0 ; e_0 = +0.20

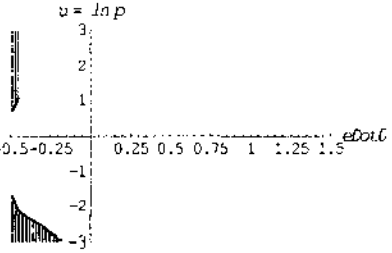


Fig. 2c. Case n = 0 ; e_0 = +0.50

Fig. 2. The same as in Fig. 1., but the power in the viscosity law $\eta \sim \Sigma^n$ is equal to $n = 0$.

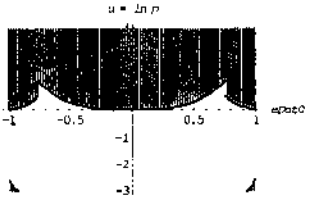


Fig. 3a. Case n = + 1 ; e_0=0.00

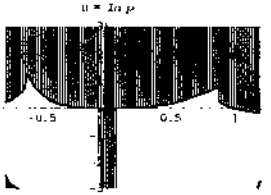


Fig. 3b. Case n = +1 ; e_0=+0.20

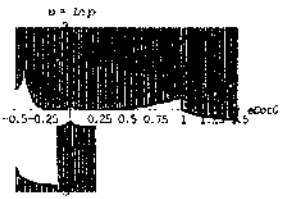


Fig. 3c. Case n = +1 ; e_0=+0.50

Fig. 3. The same as in Fig. 1., but the power in the viscosity law $\eta \sim \Sigma^n$ is equal to $n = +1$.

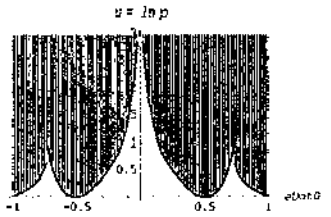


Fig. 4a. Case n = +2 ; $e_0 = 0.00$

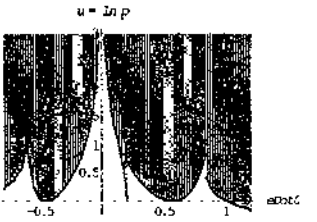


Fig. 4b. Case n = +2 ; $e_0 = +0.20$

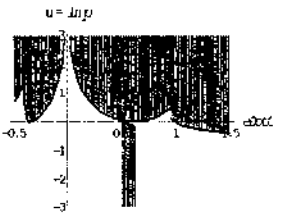


Fig. 4c. Case n = +2 ; $e_0 = +0.50$

Fig. 4. The same as in Fig. 1., but the power in the viscosity law $\eta \sim \Sigma^n$ is equal to n = +2 .

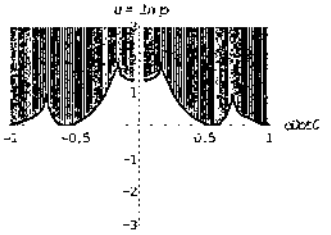


Fig. 5a. Case n = +3 ; $e_0 = 0.00$

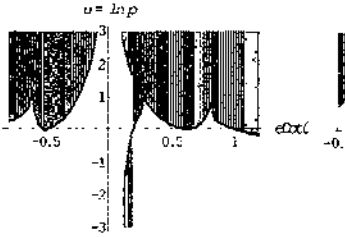


Fig. 5b. Case n = +3 ; $e_0 = +0.20$

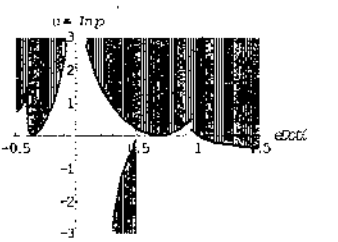


Fig. 5c. Case n = +3 ; $e_0 = +0.50$

Fig. 5. The same as in Fig. 1., but the power in the viscosity law $\eta \sim \Sigma^n$ is equal to n = +3 .

References

1. Ogilvie G. I. The dynamics of distorted discs. – In: The Physics of Cataclysmic Variables and Related Objects, ASP Conference Proceedings, vol. 261, Edited by B. T. Gänsicke, K. Beuermann, and K. Reinsch; San Francisco: Astronomical Society of the Pacific, 2002, p. 517.
2. Goodchild S., G. Ogilvie. The dynamic of eccentric accretion discs in superhump systems. - Monthly Not. Royal Astron. Soc., **368**, 2006, № 3, p. 1123.
3. Baptista R., C. Silveira, J. E. Steiner, K. Horne. Spatially resolved spectra of the accretion disc of the nova-like variable UU Aquarii. - Monthly Not. Royal Astron. Soc., **314**, 2000, № 4, p. 713.
4. Griv E. Gravitationally unstable protoplanetary discs. - Monthly Not. Royal Astron. Soc., **365**, 2006, № 3, p. 1007.
5. Sremcevic M., F. Spahn, W. J. Duschl. Density structures in perturbed thin cold discs. - Monthly Not. Royal Astron. Soc., **337**, 2002, № 3, p. 1139.

6. D i m i t r o v D. One possible simplification of the dynamical equation governing the evolution of elliptical accretion discs. – Aerospace Research in Bulgaria, 17, 2001, p.17.
7. D i m i t r o v D. Thin viscous elliptical accretion discs with orbits sharing a common longitude of periastron. I.Dynamical equation for integer values of the powers in the viscosity law. – Aerospace Research in Bulgaria, 19, 2005, p. 16.
8. D i m i t r o v D. Thin viscous elliptical accretion discs with orbits sharing a common longitude of periastron. II.Polynomial solutions to the dynamical equation for integer values of the powers in the viscosity law. - Aerospace research in Bulgaria, **21**, 2007, p. 7.
9. L y u b a r s k i j Yu. E., K. A. P o s t n o v, M. E. P r o k h o r o v. Eccentric accretion discs. - Monthly Not. Royal Astron. Soc., **266**, 1994, № 2, p. 583.
10. P e a r s o n K. J. Superhumps, magnetic fields and the mass ratio in AM Canum Venaticorum. - Monthly Not. Royal Astron. Soc., **346**, 2003, № 2, p. L21.
11. S h a l y b k o v D., G. R ü d i g e r. Magnetic field dragging and the vertical structure of thin accretion discs. - Monthly Not. Royal Astron. Soc., **315**, 2000, № 4, p. 762.
12. B a p t i s t a R., E. T. H a r l a f i s, D. S t e e g h s. Eclipse maps of spiral shocks in the accretion discs of IP Pegasi in outburst. - Monthly Not. Royal Astron. Soc., **314**, 2000, № 4, p. 727.
13. N e l s o n A. F. Numerical requirements for simulations of self gravitating and non-self gravitating discs. - Monthly Not. Royal Astron. Soc., **373**, 2006, № 4, p. 1039.
14. S h a k u r a N. I., R. A. S u n y a e v. Black holes in binary systems. Observational appearance. Astron. & Astrophys., **24**, 1973, № 3, p. 337.
15. O g i l v i e G. I. Non-linear Dynamics of eccentric discs. - Monthly Not. Royal Astron. Soc., **325**, 2001, № 1, p. 231.

ТЪНКИ ВИСКОЗНИ ЕЛИПТИЧНИ АКРЕЦИОННИ ДИСКОВЕ С ОРБИТИ ИМАЩИ ОБЩА ДЪЛЖИНА НА ПЕРИАСТРОНА. III. ЧИСЛЕНИ ОЦЕНКИ НА ОБЛАСТТА НА ВАЛИДНОСТ НА РЕШЕНИЯТА

Д. Димитров

Резюме

Ние сме изследвали областта на валидност на динамичното уравнение което определя структурата на един двумерен модел на елиптичен акреционен диск на Любарски и др. [9]. Разгледани са само

само случаи на целочислени степени в закона за вискозитета $\eta \sim \Sigma^n$, а именно n = -1, 0, +1, +2 и +3 (η е коефициентът на вискозитета, Σ е повърхностната плътност на диска). Този подход е възприет с оглед на факта, че аналитичните изрази за динамичното уравнение за тези частни стойности на n са вече получени в една по-ранна статия [7]. Като една математическа задача, ние трябва да решим едно обикновено диференциално уравнение от втори ред с начални условия – две произволни константи e_0 (стойността на ексцентрицитета) и неговата производна $\dot{e}_0$ за една дадена фиксирана стойност на фокалния параметър p_0 на една избрана елиптична траектория. В настоящата работа ние сме избрали следната мрежа от стойности: e_0 = 0.00, +0.20 и +0.50 ; $\dot{e}_0$ варира със стъпка 0.01 съответно от –1.00 до +1.00, от –0.80 до +1.20 и от –0.50 до +1.50. Независимата променлива u в динамичното уравнение е дефинирана като логаритъм от фокалния параметър p на елиптичните траектории на частиците, т. е. $u \equiv ln\, p$. Съответно, $e = e(u\,;\, e_0,\ \dot{e}_0\,,\, n)$ и $\dot{e} = \dot{e}(u;\ e_0,\ \dot{e}_0,\ n)$. Съгласно дефинирането на задачата, всеки един ексцентрицитет e трябва да бъде реална функция и въз основа на физически съображения трябва да бъдат удовлетворени неравенствата $|e(u)| < 1$, $|\dot{e}(u)| < 1$ and $|e(u) - \dot{e}(u)| < 1$ за всяка една подредена тройка от параметри $(e_0, \dot{e}_0, n)$. Тогава динамичното уравнение е решавано с помощта на числени методи и е намирана областта на изменение на u където са удовлетворени горните ограничения. За всяка една от 15 комбинации (n, e_0) допустимата област на изменение на u като функция на $\dot{e}_0$ е представена графически.

Bulgarian Academy of Sciences. Space Research Institute.
Aerospace Research in Bulgaria. 22, 2008, Sofia

ATMOSPHERIC CUT-OFFS IN THE GENERALIZED MODEL OF IONIZATION PROFILES DUE TO THE COSMIC RAY CHARGED PARTICLES IN PLANETARY IONOSPHERES AND ATMOSPHERES WITH 5 ENERGY INTERVAL APPROXIMATION OF THE IONIZATION LOSSES FUNCTION

Peter Velinov[1], Lachezar Mateev[1], Hanns Ruder [2]

[1]Central Solar-Terrestrial Influences Laboratory
Bulgarian Academy of Sciences
[2]Institut für Astronomie und Astrophysik
Eberhard-Karls-Universität Tübingen
e-mail: pvelinov@bas.bg; lnmateev@bas.bg

Abstract

An analytical and numerical approach for penetration and ionization of cosmic ray nuclei with charge Z in planetary ionospheres and atmospheres is considered in this paper. The electron production rates are calculated using new formulas, which couple the five main energy intervals in the ionization losses function (dE/dh). This is a five interval function, which performs better approximation of the measurements and experimental data in comparison with previous results for four interval ionization losses function. The geomagnetic cut-off rigidities and the energy decrease laws for the different intervals are used to create an intermediate transition energy region, which performs the coupling of the five main intervals in the ionization losses function. A new sixth energy interval for charge decrease in lower energies is taken into account. The case of vertical cosmic ray penetration is considered. The atmospheric cut-offs are calculated for 6 basic cases of atmospheric depth values.

Key words: *cosmic rays, ionization model, planetary ionospheres and atmospheres, space weather*

PACS Numbers: *94.10.-s, 94.20.-y, 96.40.-z*

This work is supported by DFG (Deutsche Forschungsgemeinschaft - German Scientific Research Society). The first two authors have Research Fellowship from the DFG (2005-2006).

Introduction

Cosmic rays (CR) are an essential factor in Earth's environment affecting physics and chemistry of planetary atmospheres [1, 2]. They are the main ionizing agent for the lower and middle atmosphere. That is why the study of the influence of cosmic ray penetration in the Earth's atmosphere and ionosphere [3, 4, 5] is important for understanding the solar-terrestrial relationships and space weather.

CR ionisation, excitation and dissociation are basic processes in the atmospheres and ionospheres of the other planets, too. For example detailed calculations of ionization profiles are made for Venus [6, 7], Jupiter [8], and Saturn [9]. It takes place also in the atmospheres of their satellites, for instance Titan [10], which possesses a dense atmosphere, similar to the terrestrial atmosphere. There are numerous CR effects in planetary atmospheres. A few examples are shown here for such essential CR influences.

Cosmic rays and electric fields and currents

It is already proved that CR influence atmospheric electric fields and thunderstorms. The high energy cosmic particles determine the conductivities in the terrestrial environment. CR have also an effect on Earth's global charge and global electric circuit [2]. In [2] it is shown that main causes of thunderstorm discharges are EAS (external atmospheric showers) which are generated by primary CR particles with energy more than 10^{14} eV. In [2] are described EAS discharge mechanisms for descending lightning (from cloud to ground) and ascending lightning (from ground to cloud).

Cosmic rays, clouds and climate formation

Many papers concerning relationships between CR penetration in the atmosphere, cloudiness and climate formation are presented in the monograph [2]. According to works [11, 12], CR ionization is the basis of clouds formation. CR determine the low cloud properties [2, 13]. The latitudinal dependence of low cloud amount on cosmic ray induced ionization is established in [14].

Several possible causes of global climate change have been discussed in scientific literature [2, 15, 16, 17, 18]: 1) internal variability in the climate system, e.g., changes of atmospheric and ocean circulation; 2) large volcanic eruptions, which are known to cause a sudden cooling lasting several years; 3) change of concentration of greenhouse gases (the increase of atmospheric CO_2 concentration during the last 100 years by about 30% has led to an increase of approximately 0.7°C in the global surface temperature); 4) orbital change of the Earth's motion

around the Sun; 5) changes in solar activity and CR intensity. The last possible cause is investigated in details in [2].

Some relationships between galactic cosmic rays and El Niño-Southern Oscillation trends are found in the International Satellite Cloud Climatology Project D2. These relationships are investigated in details in [19]. These relationships between cosmic rays in the atmosphere, cloud formation and their influence on climate are also treated in the review papers [20, 21, 22].

Some observations tending to investigate the nature of our Sun, in order to find causes or symptoms of its variable emission of light and heat are presented in the pioneer work of the famous astronomer William Herschel [23] yet in 1801, i.e. more than two hundred years ago. A strong coherence between solar variability and the monsoon phenomenon is established [24]. The altitude variations of cosmic ray induced production of aerosols and some implications for global cloudiness and climate are treated in [25].

A relationship exists between CR penetration and paleoclimate formation. A cosmic ray diffusion from the galactic spiral arms is observed in [26]. CR propagation reveals the spiral structure of the Milky Way and characterizes the ice age epochs on Earth [27]. Cosmic ray influence is also in a possible relationship with the celestial driver of phanerozoic climate [28]. A persistent solar influence on North Atlantic climate during the Holocene is found also in [29].

Analytical models for CR ionization in planetary ionospheres and atmospheres

All above mentioned processes and effects require the development of quantitative models for CR influence on the atmosphere and the ionosphere. For this purpose, a model of ionization profiles due to cosmic ray particles with charge Z in planetary ionospheres and atmospheres with multi-step energy interval approximation of the ionization losses function is created [2, 30, 31]. The energy decrease laws for different energy intervals of the ionization losses function are formulated. The corresponding electron production rate formulas are derived [32, 33]. It is assumed that the calculated energies are positive for the given atmospheric depth $\widetilde{h}$ and the projections (energy transformations) of the interval boundaries are located in the chosen energy interval. This model is developed for CR protons, Helium nuclei (alpha particles) [34] and nuclei with charge $Z > 2$. Two interval [35, 36], three interval [37, 38] and four interval [39, 40] approximations of

the ionization losses function dE/dh [1, 2, 40, 41] are taken into account. The corresponding electron production rate profiles for p, α - particles (He nuclei), L, M, H, VH groups of nuclei in the middle atmosphere and the lower ionosphere are calculated [42, 43].

The energy transformation laws were derived for the corresponding 5 interval approximation in the case of proton penetration [44, 45]. An intermediate transition region between neighbouring energy intervals was introduced [40]. But for more detailed calculations, a better approximation with 5 energy intervals for cosmic ray particles with charge Z can be applied. This new generalized model will be described in the present work. The model will contribute to the better accuracy of the problem solution [46, 47]. This 5 interval approximation is very close to the basic formula of Bohr-Bethe-Bloch for the ionization losses. A 6-th charge decrease energy interval is taken into account [2, 40]. The lower energy part of the ionization losses function is included in the new approximation [2, 40]. Vertical penetration of cosmic rays will be considered. It is the base for the full 3D electron production rate model.

Ionization Losses Formula (approximation of the Bohr-Bethe-Bloch function) for 5 energy interval model and nuclei with charge Z

The Bohr-Bethe-Bloch function is approximated in 5 energy intervals [1, 2]. One additional interval of charge decrease for cosmic ray (CR) particles from Z to 1 is included in (1) as interval 2.

(1)

$$-\frac{1}{\rho}\frac{dE}{dh}=\begin{cases}2.57\times10^{3}E^{0.5} & \text{if } kT\le E\le 0.15\ \text{MeV/n} \quad \text{, interval 1}\\ 1540E^{0.23} & \text{if } 0.15\le E\le E_{\alpha}=0.15Z^{2}\ \text{MeV/n} \quad \text{, interval 2}\\ 231\times Z^{2}E^{-0.77} & \text{if } E_{\alpha}\le E\le 200\ \text{MeV/n} \quad \text{, interval 3}\\ 68\times Z^{2}E^{-0.53} & \text{if } 200\le E\le 850\ \text{MeV/n} \quad \text{, interval 4}\\ 1.91\times Z^{2} & \text{if } 850\le E\le 5\times10^{3}\ \text{MeV/n} \quad \text{, interval 5}\\ 0.66\times Z^{2}E^{0.123} & \text{if } 5\times10^{3}\le E\le 5\times10^{6}\ \text{MeV/n} \quad \text{, interval 6}\end{cases}$$

E is the kinetic energy of charged particles, $\rho(h)$ is the neutral density of atmosphere, h is the altitude of the point, which is considered. The low energy part is taken into account. The ionization losses value is proportional to the square of the charge Z^2. They increase in interval 1 with increase of the kinetic energy E,

in interval 2 they increase with increase of the particle charge. That means, the particles decrease their energy below the value $E_a = 0.15 \times Z^2 < 200$ MeV during their penetration through the atmosphere. During this energy decrease the effective charge Z^* [2, 40] decreases and at the energy 0.15 MeV the particles become singly charged [2]. This is the energy at which the cosmic ray (CR) proton velocity equals the orbital electron velocity in hydrogen atom [2]. The charge decrease of CR nuclei which begins at energy E_a is due to the electron capturing in the nuclei shells. In intervals 3 and 4 there is a steep decrease of the ionization losses with the increase of kinetic energy. In interval 5 they are constant and in interval 6 they increase slowly with the increase of kinetic energy. This complex structure [1, 2, 46] of the ionization losses curve forms the basis of approximation (1).

Atmospheric cut-offs for 5 interval ionization losses function

The atmospheric cut-offs will be derived for those values of the travelling substance path, which correspond to the actual energy interval of the ionization losses function. For this purpose, the reverse value of the ionization losses function is integrated over the energy variable in the respective energy intervals [1]. The obtained value corresponds to the lowest energy of the particles which contribute to atmosphere ionization (with the assumption that the geomagnetic cut-off rigidity and the electric threshold for the actual height have smaller values [1, 2]).

Energy interval 1. We assume that the value of the travelling substance path $\tilde{h}$ following the CIRA model [48] is located in interval 1 from (1):

$$(2) \qquad \tilde{h} = \int_{kT}^{E_{A1}(h)} \frac{A dE}{2570 \times E^{0.5}} = \frac{A}{0.5 \times 2570} E^{0.5} \Big|_{kT}^{E_{A1}(h)} =$$

$$\frac{A}{1285}\left(E_{A1}^{0.5}(h) - (kT)^{0.5}\right)$$

The solution of equation (2) towards $E_{A1}(h)$ is the atmospheric threshold in interval 1. A is the atomic weight of particles. The integration is performed down to the energy of the thermal plasma of the solar wind.

(3) $$E_{A1}(h)=\left[\frac{1285}{A}\tilde{h}+(kT)^{0.5}\right]^2$$

The particle charge in this interval is equal to 1.

Energy interval 2 (charge decrease energy interval): The atmospheric cut-off in the second interval is calculated by means of the corresponding ionization path. In this case it is assumed that $E_{A1}(h) > 0.15$ MeV/n in (3). It is composed of two terms. In the first term the charge $Z = 1$. The charge in the second term decreases from Z^* to 1 with the energy decrease [40]. $\tilde{h}_i, i = 1,2$ is the travelling substance path in the corresponsing energy interval.

(4) $$\tilde{h} = \tilde{h}_1 + \tilde{h}_2 = \int_{kT}^{0.15} \frac{AdE}{2570E^{0.5}} + \int_{0.15}^{E_{A2}(h)} \frac{AdE}{1540E^{0.23}} =$$

$$\frac{A}{1285}\left(0.15^{0.5} - (kT)^{0.5}\right) + \frac{A}{1540\times 0.77}\left(E_{A2}^{0.77}(h) - 0.15^{0.77}\right)$$

The atmospheric cut-off is obtained as solution of equation (4) towards $E_{A2}(h)$:

(5) $$E_{A2}(h)=\left[0.15^{0.77} - \frac{1540\times 0.77}{1285}\left(0.15^{0.5} - (kT)^{0.5}\right) + \right.$$

$$\left. + \frac{1540\times 0.77}{A}\tilde{h}\right]^{1/0.77}$$

Energy interval 3: If the value $E_{A2}(h) > E_a$, then the ionization path corresponding to interval 3 is composed of 3 terms with respect to the first 3 energy intervals of the ionization losses function. The charge in interval 1 $Z = 1$. The charge in interval 2 decreases from Z to 1. The charge value in energy interval 3 is equal to Z. A is the atomic weight of cosmic ray particles.

(6) $$\tilde{h} = \tilde{h}_1 + \tilde{h}_2 + \tilde{h}_3 = \int_{kT}^{0.15} \frac{A}{2570E^{0.5}} + \int_{0.15}^{E_a} \frac{AdE}{1540E^{0.23}} +$$

$$+ \int_{E_a}^{E_{A3}(h)} \frac{dE}{231\frac{Z^2}{A}E^{-0.77}} =$$

$$\frac{A}{1285}\left(0.15^{0.5} - (kT)^{0.5}\right) + \frac{A}{0.77\times 1540}\left(E_a^{0.77} - 0.15^{0.77}\right) +$$

$$+ \frac{A}{1.77\times 231\times Z^2}\left(E_{A3}^{1.77}(h) - E_a^{1.77}\right)$$

After integration of equation (6), $E_{A3}(h)$ becomes its solution towards the unknown variable:

(7) $$E_{A3}(h) = \left[E_a^{1.77} - \frac{1.77\times 231\times Z^2}{1285}\left(0.15^{0.5} - (kT)^{0.5}\right) - \right.$$

$$- \frac{1.77\times 231\times Z^2}{0.77\times 1540}\left(E_a^{0.77} - 0.15^{0.77}\right) +$$

$$\left. + 231\times 1.77\frac{Z^2}{A}\tilde{h} \right]^{1/1.77}$$

Energy interval 4: If the condition $E_{A3}(h) > 200$ MeV/n is fulfilled, the ionization path is composed of the following 4 terms:

(8) $$\tilde{h} = \tilde{h}_1 + \tilde{h}_2 + \tilde{h}_3 + \tilde{h}_4 = \int_{kT}^{0.15} \frac{AdE}{2570E^{0.5}} + \int_{0.15}^{E_a} \frac{AdE}{1540E^{0.23}} +$$

$$\int_{E_a}^{200} \frac{dE}{231\frac{Z^2}{A}E^{-0.77}} + \int_{200}^{E_{A4}(h)} \frac{dE}{68\frac{Z^2}{A}E^{-0.53}} =$$

$$\frac{A}{1285}\left(0.15^{0.5}-(kT)^{0.5}\right)+\frac{A}{1540\times 0.77}\left(E_a^{0.77}-0.15^{0.77}\right)+$$

$$\frac{A}{231\times 1.77\times Z^2}\left(200^{1.77}-E_a^{1.77}\right)+\frac{A}{68\times 1.53\times Z^2}\left(E_{A4}^{1.53}(h)-200^{1.53}\right)$$

The charge Z values in the corresponding energy intervals 1 and 2 are formed as in the previous case of the atmospheric cut-off $E_{A3}(h)$. The solution of equation (8) yields the atmospheric cut-off $E_{A4}(h)$:

(9) $$E_{A4}(h)=\left[200^{1.53}+68\times 1.53\frac{Z^2}{A}\tilde{h}-\frac{68\times 1.53\times Z^2}{1285}\left(0.15^{0.5}-(kT)^{0.5}\right)-\right.$$

$$\left.-\frac{68\times 1.53\times Z^2}{1540\times 0.77}\left(E_a^{0.77}-0.15^{0.77}\right)-\frac{68\times 1.53}{231\times 1.77}\left(200^{1.77}-E_a^{1.77}\right)\right]^{1/1.53}$$

Energy interval 5: The condition $E_{A4}(h) > 850$ MeV/n determines the following ionization path calculation [1, 2]:

(10) $$\tilde{h}=\tilde{h}_1+\tilde{h}_2+\tilde{h}_3+\tilde{h}_4+\tilde{h}_5=\int_{kT}^{0.15}\frac{AdE}{2570E^{0.5}}+\int_{0.15}^{E_a}\frac{AdE}{1540E^{0.23}}+$$

$$\int_{E_a}^{200}\frac{dE}{231\frac{Z^2}{A}E^{-0.77}}+\int_{200}^{850}\frac{dE}{68\frac{Z^2}{A}E^{-0.53}}+\int_{850}^{E_{A5}(h)}\frac{dE}{1.91\frac{Z^2}{A}}=$$

$$\frac{A}{1285}\left(0.15^{0.5}-(kT)^{0.5}\right)+\frac{A}{1540\times 0.77}\left(E_a^{0.77}-0.15^{0.77}\right)+$$

$$\frac{A}{231\times 1.77\times Z^2}\left(200^{1.77}-E_a^{1.77}\right)+\frac{A}{68\times 1.53\times Z^2}\left(850^{1.53}-200^{1.53}\right)+$$

$$\frac{A}{1.91\times Z^2}\left(E_{A5}(h)-850\right)$$

The following transformation of equation (10) is done with the purpose to find the atmospheric cut-off $E_{A5}(h)$:

$$(11)\quad \frac{1.91\times Z^2}{A}\tilde{h}+850=E_{A5}(h)+\frac{1.91\times Z^2}{1285}\left(0.15^{0.5}-(kT)^{0.5}\right)+$$
$$\frac{1.91\times Z^2}{1540\times 0.77}\left(E_a^{0.77}-0.15^{0.77}\right)+\frac{1.91}{231\times 1.77}\left(200^{1.77}-E_a^{1.77}\right)+$$
$$\frac{1.91}{68\times 1.53}\left(850^{1.53}-200^{1.53}\right)$$

The solution of equation (11) towards $E_{A5}(h)$ is the following:

$$(12)\quad E_{A5}(h)=\frac{1.91\times Z^2}{A}\tilde{h}+850-\frac{1.91\times Z^2}{1285}\left(0.15^{0.5}-(kT)^{0.5}\right)-$$
$$\frac{1.91\times Z^2}{1540\times 0.77}\left(E_a^{0.77}-0.15^{0.77}\right)-\frac{1.91}{231\times 1.77}\left(200^{1.77}-E_a^{1.77}\right)-$$
$$\frac{1.91}{68\times 1.53}\left(850^{1.53}-200^{1.53}\right)$$

Energy interval 6: The condition $E_{A5}(h) > 5000$ MeV/n determines the calculation of the atmospheric cut-off $E_{A6}(h)$ from the following expression for the ionization path which is equal to the travelling substance path $\tilde{h}$ for kinetic energy decrease until absorption of the cosmic ray particles:

$$(13)\quad \tilde{h}=\tilde{h}_1+\tilde{h}_2+\tilde{h}_3+\tilde{h}_4+\tilde{h}_5+\tilde{h}_6=\int_{kT}^{0.15}\frac{AdE}{2570E^{0.5}}+\int_{0.15}^{E_a}\frac{AdE}{1540\times E^{0.23}}+$$
$$\int_{E_a}^{200}\frac{dE}{231\dfrac{Z^2}{A}E^{-0.77}}+\int_{200}^{850}\frac{dE}{68\dfrac{Z^2}{A}E^{-0.53}}+\int_{850}^{5000}\frac{dE}{1.91\dfrac{Z^2}{A}}+$$

$$\int_{5000}^{E_{A6}(h)} \frac{dE}{0.66\frac{Z^2}{A}E^{0.123}} = \frac{A}{1285}\left(0.15^{0.5} - (kT)^{0.5}\right) +$$

$$\frac{A}{1540\times 0.77}\left(E_a^{0.77} - 0.15^{0.77}\right) + \frac{A}{231\times Z^2\times 1.77}\left(200^{1.77} - E_a^{1.77}\right) +$$

$$\frac{A}{68\times Z^2\times 1.53}\left(850^{1.53} - 200^{1.53}\right) + \frac{A}{1.91\times Z^2}(5000-850) +$$

$$\frac{A}{0.66\times Z^2\times 0.877}\left(E_{A6}^{0.877}(h) - 5000^{0.877}\right)$$

The corresponding transformation of (13) is the following:

(14) $$0.66\times 0.877\frac{Z^2}{A}\tilde{h} + 5000^{0.877} = \frac{0.66\times Z^2\times 0.877}{1285}\left(0.15^{0.5} - (kT)^{0.5}\right) +$$

$$\frac{0.66\times Z^2\times 0.877}{1540\times 0.77}\left(E_a^{0.77} - 0.15^{0.77}\right) + \frac{0.66\times 0.877}{231\times 1.77}\left(200^{1.77} - E_a^{1.77}\right) +$$

$$\frac{0.66\times 0.877}{68\times 1.53}\left(850^{1.53} - 200^{1.53}\right) + \frac{0.66\times 0.877}{1.91}(5000-850) + E_{A6}^{0.877}(h)$$

The solution of (13) - (14) presents the atmospheric cut-off $E_{A6}(h)$ in energy interval 6:

(15) $$E_{A6}(h) = \left[5000^{0.877} + 0.66\times 0.877\frac{Z^2}{A}\tilde{h} - \frac{0.66\times 0.877\times Z^2}{1285}\times\right.$$

$$\left(0.15^{0.5} - (kT)^{0.5}\right) - \frac{0.66\times 0.877\times Z^2}{1540\times 0.77}\left(E_a^{0.77} - 0.15^{0.77}\right) -$$

$$\frac{0.66\times 0.877}{231\times 1.77}\left(200^{1.77} - E_a^{1.77}\right) - \frac{0.66\times 0.877}{68\times 1.53}\left(850^{1.53} - 200^{1.53}\right) -$$

$$\left.\frac{0.66\times 0.877}{1.91}(5000-850)\right]^{1/0.877}$$

The calculation of atmospheric cut-offs is important for evaluation of the lower boundary of integration in the corresponding ionization model [1, 2]. It determines the energy interval combinations between the initial cosmic ray kinetic energy intervals and the respective final ionization losses function energy intervals. These combinations create the different integral terms of the electron production rate model for the current valid values of altitude, zenith angle and azimuth angle.

References

1. Velinov P., G. Nestorov, L. Dorman. Cosmic Ray Influence on the Ionosphere and on Radio-Wave Propagation. Sofia, BAS Publ. House, 1974.
2. Dorman L. Cosmic Rays in the Earth's Atmosphere and Underground. Dordrecht, Kluwer Academic Publishers, 2004.
3. Velinov P. I. Y., L. Mateev, N. Kilifarska. Annalac Geophysicac, 2005, **23**, 9, 3043 - 3046.
4. Velinov P. I. Y., M. Buchvarova, L. Mateev, H. Ruder. Adv. Space Res., 2001, **27**, 11, 1901 - 1908.
5. Velinov P. I. Y., H. Ruder, L. N. Mateev. Cosmic Ray and Solar Energetic Particle Influences on the Planetary Ionospheres: Improved Analytical Approach. http://www.stil.bas.bg/11conf/AGENDA/PIM1
Solar - Terrestrial Influences, Proceedings of the Eleventh International Scientific Conference, Dedicated to the Year of Physics 2005, Sofia, 23 - 25 November 2005, Edited by S. Panchev, CSTIL BAS, Publishing House of Bulgarian Academy of Sciences, 2005, PIM 1, pp. 3 - 6.
6. Upadhyay H. O., R. P. Singh, R. N. Singh. Earth, Moon, and Planets. **65**, 1994, No 1, 89-94.
7. Upadhyay H. O., R. N. Singh. Adv. Space Res., **15**, 1995, no. 4, 99-108.
8. Capone L. A., J. Dubach, R. C. Whitten, S. S. Prasad. Icarus, **39**, 1979, p. 433-449.
9. Velinov P. I. Y., H. Ruder, L. Mateev, and M. Buchvarova, V. Kostov. Adv. Space Res., **33**, 2004, 2, 232 - 239.
10. Molina-Cuberos G.J., J. J. Lopez-Moreno, R. Rodrigo, L. M. Lara, K. O'Brien. Planet. Space Sci., **47**, 1999, 10, pp. 1347-1354(8)
11. Palle E., C.J. Butler, K. O'Brien. J. Atm. Sol.-Terr. Phys., **66**, 2004, 1779–1790
12. Usoskin I.G., G.A. Kovaltsov, Link between cosmic rays and clouds on different time scales, in: "Advances in Geosciences", v.2 (eds. W.-H. Ip, M. Duldig) Singapore, World Scientific, 2006, pp.321-330.
13. Marsh N., H. Svensmark. Phys. Rev. Lett. **85**, 2000, 5004.
14. Usoskin I. et al. Geoph. Res. Lett. **31**, 2004, L16109.
15. Witze A., H. Svensmark, Influence of Cosmic Rays on Earth's Climate, http://www.tmgnow.com/repository/global/CREC.html
16. Svensmark H., Influence of Cosmic Rays on Earth's Climate, Phys. Rev. Lett. **81**, 1998, 5027, http://www.dsri.dk/~hsv/prlresup2.pdf.
17. Svensmark H. Space Sci. Rev. **93**, 2000, 175 - 185.

18. Shaviv N. J. Cosmic Rays and Climate: in PhysicaPlus, 5: http://www.sciencebits.com/CosmicRaysClimate
19. Marsh N., H. Svensmark. J. Geoph. Res., **108**, 2003, D6, 6.
20. Carslaw K. S., R. G. Harrison, J. Kirkby. Cosmic Rays, Clouds, and Climate. Review: atmospheric science. http://www.seas.harvard.edu/climate/pdf/carslaw-2002.pdf
21. Svensmark H. Influence of Cosmic Rays on the Earth's Climate - http://www.junkscience.com/Greenhouse/influence-of-cosmic-rays-on-the-earth.pdf
22. Marsh N., H. Svensmark. Cosmic rays, clouds, and climate. http://www.dsri.dk/~hsv/SSR_Paper.pdf
23. William Herschel. Phil. Trans. Roy. Soc. London, **91**, 1801, 265.
24. Neff U. et al. Nature, **411**, 2001, 290.
25. Yu Fangqun. J. Geoph. Res., **107**, 2002, A7, 10.1029/2001JA000248.
26. Shaviv N. Phys. Rev. Lett. **89**, 2002, 051102.
27. Shaviv N. New Astronomy **8**, 2003, 39.
28. Shaviv N., J. Veizer. GSA Today, **13**, 2003, No. 7, 4.
29. Bond G. et al. Science, **294**, 2001, 2130-2136.
30. Velinov P. I. Y., L. N. Mateev. Compt. rend. Acad. bulg. Sci., **58**, 2005, 5, 511 - 516.
31. Velinov P. I. Y., H. Ruder, L. N. Mateev. Compt. rend. Acad. bulg. Sci., **58**, 2005, 8, 897 - 902.
32. Velinov P. I. Y., H. Ruder, L. Mateev. Analytical Model for Galactic and Solar Cosmic Ray Ionization in the Planetary Ionospheres and Atmospheres. The Second European Space Weather Week, 14 - 18 November 2005, European Space Research and Technology Centre (ESTEC), Noordwijk, The Netherlands, Poster Session 2, Progr. & Abstr., p. 93, European Space Agency (ESA), 2005. http://www.esa-spaceweather.net/spweather/esa_initiatives/
33. Velinov P. I. Y., L. N. Mateev. Compt. rend. Acad. bulg. Sci., **58**, 2005, 12, 1399 - 1404.
34. Velinov P. I. Y., H. Ruder, L. N. Mateev. Compt. rend. Acad. bulg. Sci., **58**, 2005, 9, 1033 - 1038.
35. Ruder H., P. I. Y. Velinov, L. N. Mateev. Compt. rend. Acad. bulg. Sci., **59**, 2006, 7, 717 - 722.
36. Velinov P. I. Y., L. Mateev. Analytical Approach to Cosmic Ray Ionization by Nuclei with Charge Z in the Middle Atmosphere - Distribution of Galactic CR Effects. Report C2.1-0003-06 on the XXXVI-th General Assembly of Committee of Space Research (COSPAR), Beijing, China, 16 - 23 July 2006. Adv. Space Res., 2007, **39**, (COSPAR) submitted
37. Velinov P. I. Y., H. Ruder, L. N. Mateev. Compt. rend. Acad. bulg. Sci., **59**, 2006, 8, 847 - 854.
38. Velinov P. I. Y., L. N. Mateev. Compt. rend. Acad. bulg. Sci., **59**, 2006, 10, 1001 - 1008.
39. Velinov P. I. Y., L. N. Mateev. Compt. rend. Acad. bulg. Sci., **60**, 2007, 1, 37 - 44.
40. Velinov P. I. Y., L. Mateev. Improved Cosmic Ray Ionization Model in Ionosphere and Atmosphere for Particles with Charge Z. Calculation of Electron Production Rate Profiles. Report on the COST 724 Management Committee Meeting and Scientific Event "Developing the Basis for Monitoring, Modelling and Predicting Space Weather", WG2 - The Radiation Environment of the Earth, Sofia, Bulgaria, 21 - 25 May 2007, pp. 1 - 11.

41. Al'pert Ya. L. Radio Wave Propagation and the Ionosphere. New York, Consultants Bureau Enterprises, Inc., 1963.
42. Velinov P. I. Y., L. Mateev. Ann. Geophys. 2007, **25**, submitted
43. Velinov P. I. Y., L. N. Mateev. Cosmic Ray Influence on the System Ionosphere - Atmosphere through Ionization Processes. Modeling Profiles of Electron Production Rates. Reports on the International Symposium on Recent Observations and Simulations of the Sun-Earth System (ISROSES), Varna, 2006, Heron Press Ltd., Sofia, 2006, p. 73, 102, 103; J. Atmos. Solar-Terr. Phys., 70, 2008, pp. 574 - 582.
44. Velinov P. I. Y., L. Mateev. Compt. rend. Acad. bulg. Sci., **60**, 2007, 6, 613 - 618.
45. Velinov P. I. Y., L. N. Mateev. Compt. rend. Acad. bulg. Sci., **60**, 2007, 8, 839 - 844.
46. Starodubzev S., A. Romanov. Penetration of Charged Particles through Substance. Tashkent, Proc. Acad. Sci. Uzbek. SSR, 1962.
47. Brasseur G., S. Solomon. Aeronomy of the Middle Atmosphere. Dordrecht, D. Reidel Publ. Comp., 1984.
48. CIRA 1986, Cospar International Reference Atmosphere. Amsterdam, North-Holland, 1986.

АТМОСФЕРНИ ПРАГОВЕ НА ОТРЯЗВАНЕ В ОБОБЩЕНИЯ МОДЕЛ НА ЙОНИЗАЦИОННИ ПРОФИЛИ ОТ ЗАРЕДЕНИТЕ ЧАСТИЦИ НА КОСМИЧЕСКИТЕ ЛЪЧИ В ПЛАНЕТНИТЕ ЙОНОСФЕРИ И АТМОСФЕРИ С АПРОКСИМАЦИЯ НА ФУНКЦИЯТА НА ЙОНИЗАЦИОННИТЕ ЗАГУБИ ВЪРХУ 5 ЕНЕРГИЙНИ ИНТЕРВАЛА

П. Велинов, Л. Матеев, Х. Рудер

Резюме

В настоящата статия е разгледан един аналитично-числен метод за описание на проникването и йонизацията от ядра на космическите лъчи със заряд Z в планетните йоносфери и атмосфери. Скоростта на електронната продукция се изчислява с нови формули, които съчетават петте основни енергийни интервала на функцията на йонизационните загуби (dE/dh). Тя е 5-интервална функция, която осъществява по-добра апроксимация на измерванията и експерименталните данни в сравнение с предходни резултати за 4–интервална функция на йонизационните загуби. Геомагнитните прагове на отрязване и законите за намаление на енергията за различните интервали се използуват за създаване на междинна преходна област на енергията, която осъществява съчетаване на петте основни интервала на функцията на йонизационните загуби. Въвежда се един нов шести енергиен интервал за намаление на заряда в ниските енергии. Разглежда се вертикално проникване на космическите лъчи. Атмосферните прагове на отрязване са изчислени за 6 основни случая на стойностите на дълбочината на атмосферата.

Bulgarian Academy of Sciences. Space Research Institute.
Aerospace Research in Bulgaria. 22, 2008, Sofia

ENERGY DECREASE LAWS AND ELECTRON PRODUCTION RATES IN THE GENERALIZED MODEL OF IONIZATION PROFILES DUE TO THE COSMIC RAY CHARGED PARTICLES IN PLANETARY IONOSPHERES AND ATMOSPHERES WITH 5 ENERGY INTERVAL APPROXIMATION OF THE IONIZATION LOSSES FUNCTION

Peter Velinov[1], Hanns Ruder[2], Lachezar Mateev

[1]Central Solar-Terrestrial Influences Laboratory
Bulgarian Academy of Sciences
[2]Institut für Astronomie und Astrophysik
Eberhard-Karls-Universität Tübingen
e-mail: pvelinov@bas.bg; lnmateev@bas.bg

Abstract

An analytical and numerical approach for penetration and ionization of cosmic ray nuclei with charge Z in the planetary ionospheres and atmospheres is considered in this paper. The electron production rates are calculated using new formulas, which couple the five main energy intervals in the ionization losses function (dE/dh). This is a five interval function, which performs better approximation of the measurements and experimental data in comparison with previous results for four interval ionization losses function. The geomagnetic cut-off rigidities and the energy decrease laws for the different intervals are used for creation of an intermediate transition energy region, which performs the coupling of the five main intervals in the ionization losses function. A new sixth energy interval for charge decrease in lower energies is taken into account. The case of vertical cosmic ray penetration is considered. The corresponding energy decrease laws and initial energy interval boundary values are evaluated.

Key words: *cosmic rays, ionization model, planetary ionospheres and atmospheres, space weather*

PACS Numbers: *94.10.-s, 94.20.-y, 96.40.-z*

This work is supported by DFG (Deutsche Forschungsgemeinschaft - German Scientific Research Society). The first two authors have Research Fellowship from the DFG (2005-2006).

Energy decrease laws and boundary crossings

The calculation of energy decrease laws over the boundaries of the ionization losses function energy intervals is very important for the creation of improved electron production rate model. The formulas which describe this kind of cosmic ray particles kinetic energy decrease will be derived as follows.

Boundary crossing between energy intervals 1 and 2. We assume that the initial kinetic energy of cosmic ray particles before penetration in the atmosphere is $0.15<E_k<E_a$. Boundary crossing occurs if $\tilde{h}_2 < \tilde{h}(h)$. The following condition is fulfilled: $\tilde{h}(h) < \tilde{h}_1(E_{21}(h) = kT) + \tilde{h}_2$ The final energy, corresponding to E_k and $\tilde{h}(h)$ is $E_{21}(h)$.

$$(1) \qquad \tilde{h} = \tilde{h}_1 + \tilde{h}_2 = \int_{E_{21}(h)}^{0.15} \frac{AdE}{2570 \times E^{0.5}} + \int_{0.15}^{E_k} \frac{AdE}{1540 \times E^{0.23}} =$$

$$\frac{A}{1285}\left(0.15^{0.5} - E_{21}^{0.5}(h)\right) + \frac{A}{1540 \times 0.77}\left(E_k^{0.77} - 0.15^{0.77}\right)$$

The transformed equation which contains the unknown variable $E_{21}(h)$ is the following:

$$(2) \qquad \frac{1285}{A}\tilde{h} = 0.15^{0.5} - E_{21}^{0.5}(h) + \frac{1285}{1540 \times 0.77}\left(E_k^{0.77} - 0.15^{0.77}\right)$$

The final energy after the CR particles penetration in the atmosphere at altitude h is:

$$(3) \qquad E_{21}(h) = \left[0.15^{0.5} + \frac{1285}{1540 \times 0.77}\left(E_k^{0.77} - 0.15^{0.77}\right) - \frac{1285}{A}\tilde{h}\right]^2$$

Boundary crossing between energy intervals 2 and 3. The condition $\tilde{h}_3 < \tilde{h}(h)$ for fixed $E_a < E_k < 200$ MeV/n causes interval boundary crossing of the energy E_a. The final energy $E_{32}(h)$ after penetration of CR particles at height h for traveling substance path $\tilde{h}(h) < \tilde{h}_2(E_{32}(h) = 0.15) + \tilde{h}_3$ is calculated as follows:

(4) $$\tilde{h} = \tilde{h}_2 + \tilde{h}_3 = \int_{E_{32}(h)}^{E_a} \frac{A dE}{1540 \times E^{0.23}} + \int_{E_a}^{E_k} \frac{dE}{231 \frac{Z^2}{A} E^{-0.77}} =$$

$$\frac{A}{1540 \times 0.77}\left(E_a^{0.77} - E_{32}^{0.77}(h)\right) + \frac{A}{231 \times 1.77 \times Z^2}\left(E_k^{1.77} - E_a^{1.77}\right)$$

Equation (4) is transformed with the purpose to be solved towards the unknown variable $E_{32}(h)$:

(5) $$\frac{1540 \times 0.77}{A}\tilde{h} = E_a^{0.77} - E_{32}^{0.77}(h) + \frac{1540 \times 0.77}{231 \times 1.77 \times Z^2}\left(E_k^{1.77} - E_a^{1.77}\right)$$

The final energy $E_{32}(h)$ after CR penetration at height h is therefore defined with the following expression:

(6) $$E_{32}(h) = \left[E_a^{0.77} + \frac{1540 \times 0.77}{231 \times 1.77 \times Z^2}\left(E_k^{1.77} - E_a^{1.77}\right) - \frac{1540 \times 0.77}{A}\tilde{h}\right]^{1/0.77}$$

Boundary crossing between energy intervals 3 and 4. The condition $\tilde{h}_4 < \tilde{h}(h)$ causes the interval boundary crossing of the energy 200 MeV/n for initial kinetic energy of particles $200 < E_k < 850$ MeV/n. From the relation $\tilde{h}(h) < \tilde{h}_3(E_{43}(h) = E_a) + \tilde{h}_4$ it follows that

(7) $$\tilde{h} = \tilde{h}_3 + \tilde{h}_4 = \int_{E_{43}(h)}^{200} \frac{dE}{231 \frac{Z^2}{A} E^{-0.77}} + \int_{200}^{E_k} \frac{dE}{68 \frac{Z^2}{A} E^{-0.53}} =$$

$$\frac{A}{231 \times Z^2 \times 1.77}\left(200^{1.77} - E_{43}^{1.77}(h)\right) + \frac{A}{68 \times Z^2 \times 1.53}\left(E_k^{1.53} - 200^{1.53}\right)$$

The transformed equation (7) has the form:

$$(8)\qquad \frac{231\times 1.77\times Z^2}{A}\tilde{h}=200^{1.77}-E_{43}^{1.77}(h)+\frac{231\times 1.77}{68\times 1.53}\left(E_k^{1.53}-200^{1.53}\right)$$

The corresponding final kinetic energy $E_{43}(h)$ which characterizes the CR penetration in the atmosphere for intervals 3 and 4 from (1) in [1] is:

$$(9)\qquad E_{43}(h)=\left[200^{1.77}+\frac{231\times 1.77}{68\times 1.53}\left(E_k^{1.53}-200^{1.53}\right)-\frac{231\times 1.77\times Z^2}{A}\tilde{h}\right]^{1/1.77}$$

Boundary crossing between energy intervals 4 and 5. This case follows from the condition $\tilde{h}_5<\tilde{h}(h)$. It is valid for $850<E_k<5000$ MeV/n with the relationship $\tilde{h}(h)<\tilde{h}_4\left(E_{54}(h)=200\right)+\tilde{h}_5$. The penetrating CR particles cross the boundary 850 MeV/n:

$$(10)\qquad \tilde{h}=\tilde{h}_4+\tilde{h}_5=\int_{E_{54}(h)}^{850}\frac{dE}{68\dfrac{Z^2}{A}E^{-0.53}}+\int_{850}^{E_k}\frac{dE}{1.91\dfrac{Z^2}{A}}$$

Equation (10) is solved towards the unknown variable $E_{54}(h)$ which presents the final kinctic energy for this case, namely

$$(11)\qquad E_{54}(h)=\left[850^{1.53}+\frac{68\times 1.53}{1.91}(E_k-850)-68\times 1.53\frac{Z^2}{A}\tilde{h}\right]^{1/1.53}$$

Boundary crossing between energy intervals 5 and 6. The next energetic interval boundary in (1) from [1], which is crossed, has the value 5000 MeV. This case takes place, when $\tilde{h}_6<\tilde{h}(h)$ and $5000<E_k<5\times 10^6$ MeV/n. The condition $\tilde{h}(h)<\tilde{h}_5\left(E_{65}(h)=850\right)+\tilde{h}_6$ is also valid:

$$(12)\qquad \tilde{h}=\tilde{h}_5+\tilde{h}_6=\int_{E_{65}(h)}^{5000}\frac{dE}{1.91\dfrac{Z^2}{A}}+\int_{5000}^{E_k}\frac{dE}{0.66\dfrac{Z^2}{A}E^{0.123}}$$

The final kinetic energy $E_{65}(h)$ is calculated as follows:

$$(13) \qquad E_{65}(h) = 5000 - 1.91\frac{Z^2}{A}\tilde{h} + \frac{1.91\left(E_k^{0.877} - 5000^{0.877}\right)}{0.66 \times 0.877}$$

Initial energies for interval boundaries

The initial kinetic energies of the interval boundaries are needed for formation of intermediate transition regions between energetic intervals in the improved cosmic ray ionization model.

***Boundary between energy intervals 1 and 2*: 0.15 MeV/n.** It is assumed that $\tilde{h}$ generates the condition $E_{0.15;2}(h) < E_a$ in the following expression:

$$(14) \qquad \tilde{h} = \int_{0.15}^{E_{0.15;2}(h)} \frac{AdE}{1540E^{0.23}} = \frac{A}{1540 \times 0.77} E^{0.77}\Big|_{0.15}^{E_{0.15;2}(h)} =$$

$$\frac{A}{1540 \times 0.77}\left(E_{0.15;2}^{0.77}(h) - 0.15^{0.77}\right)$$

The unknown variable for the initial energy $E_{0.15;2}(h)$ is calculated from (14):

$$(15) \qquad E_{0.15;2}(h) = \left[0.15^{0.77} + \frac{1540 \times 0.77}{A}\tilde{h}\right]^{1/0.77}$$

It creates the upper boundary of the intermediate region between energy intervals 1 and 2 in (1) from [1].

***Boundary between energy intervals 2 and 3*: E_a.** It is assumed that the value of $\tilde{h}$ generates the condition $E_{E_a,3}(h) < 200$ MeV/n in the next equation:

$$(16) \qquad \tilde{h} = \int_{E_a}^{E_{E_a,3}(h)} \frac{dE}{231\frac{Z^2}{A} \cdot E^{-0.77}} = \frac{E_{E_a,3}^{1.77}(h) - E_a^{1.77}}{231\frac{Z^2}{A} \times 1.77}$$

The initial value of E_a in interval 3 is obtained from equation (16):

$$(17) \qquad E_{E_a,3}(h)=\left[E_a^{1.77}+231\frac{Z^2}{A}\times 1.77\times\tilde{h}\right]^{1/1.77}$$

***Boundary between energy intervals 3 and 4*: 200 MeV/n.** It is assumed that the value of $\tilde{h}$ generates the condition $E_{200,4}(h)<850$ MeV/n in the next equation:

$$(18) \qquad \tilde{h}=\int_{200}^{E_{200,4}(h)}\frac{dE}{68\frac{Z^2}{A}E^{-0.53}}=\frac{E_{200,4}^{1.53}(h)-200^{1.53}}{68\frac{Z^2}{A}\times 1.53}$$

The corresponding initial energy $E_{200,4}(h)$ is calculated from equation (18):

$$(19) \qquad E_{200,4}(h)=\left[200^{1.53}+68\times 1.53\times\frac{Z^2}{A}\times\tilde{h}\right]^{1/1.53}$$

***Boundary between energy intervals 4 and 5*: 850 MeV/n.** It is assumed that the value of $\tilde{h}$ generates the condition $E_{850,5}(h)<5000$ MeV/n in the next equation:

$$(20) \qquad \tilde{h}=\int_{850}^{E_{850,5}(h)}\frac{dE}{1.91\frac{Z^2}{A}}=\frac{E_{850,5}(h)-850}{1.91\frac{Z^2}{A}}$$

The corresponding initial energy value $E_{850,5}(h)$ is calculated from equation (20):

$$(21) \qquad E_{850,5}(h)=850+1.91\frac{Z^2}{A}\tilde{h}$$

***Boundary between energy intervals 5 and 6*: 5000 MeV/n.** It is assumed that the value of $\tilde{h}$ generates the condition $E_{5000,6}(h)<5\times 10^6$ MeV/n in the next equation:

$$(22)\qquad \tilde{h} = \int_{5000}^{E_{5000,6}(h)} \frac{dE}{0.66\dfrac{Z^2}{A}E^{0.123}} = \frac{E_{5000,6}^{0.877}(h) - 5000^{0.877}}{0.66\times\dfrac{Z^2}{A}\times 0.877}$$

The corresponding initial energy value $E_{5000,6}(h)$ is calculated from equation (22):

$$(23)\qquad E_{5000,6}(h) = \left[5000^{0.877} + 0.66\times\frac{Z^2}{A}\times 0.877\times\tilde{h}\right]^{1/0.877}$$

Energy decrease laws in internal regions, corresponding to energy intervals 1-6

These energy decrease laws are used for calculation of the ionization losses and the electron production rate in the internal regions of all 6 energy intervals in (1) from [1]. It means, that both the initial energy E_k and the final energy $E_i(h)$ belong to the corresponding energy interval i, $i\in\{1,\ldots,6\}$ from (1) in [1].

Energy interval region 1. The condition $\{E_k, E_1(h)\}\in[kT, 0.15]$ MeV/n is fulfilled. E_k is the initial energy and $E_1(h)$ is the final energy for this case of energy decrease law.

$$(24)\qquad \tilde{h} = \int_{E_1(h)}^{E_k} \frac{AdE}{2570\times E^{0.5}} = \frac{A\left(E_k^{0.5} - E_1^{0.5}\right)}{1285}$$

The final energy $E_1(h)$ is calculated from equation (24) as follows:

$$(25)\qquad E_1(h) = \left[E_k^{0.5} - \frac{1285}{A}\tilde{h}\right]^2$$

Energy interval region 2. The condition $\{E_k, E_2(h)\}\in[0.15, E_a]$ MeV/n is fulfilled. E_k is the initial energy and $E_2(h)$ is the final energy for this case of energy decrease law.

$$(26)\qquad \tilde{h} = \int_{E_2(h)}^{E_k} \frac{AdE}{1540\times E^{0.23}} = \frac{A\left(E_k^{0.77} - E_2^{0.77}(h)\right)}{1540\times 0.77}$$

The final energy $E_2(h)$ is calculated from equation (26) as follows:

$$(27)\qquad E_2(h) = \left[E_k^{0.77} - \frac{1540}{A}\times 0.77\times\tilde{h}\right]^{1/0.77}$$

Energy interval region 3. The condition $\{E_k, E_3(h)\} \in [E_a, 200]$ MeV/n is fulfilled. E_k is the initial energy and $E_3(h)$ is the final energy for this case of energy decrease law.

$$\tilde{h} = \int_{E_3(h)}^{E_k} \frac{dE}{231\frac{Z^2}{A}E^{-0.77}} = \frac{E_k^{1.77} - E_3^{1.77}(h)}{231\frac{Z^2}{A}\times 1.77} \tag{28}$$

The final energy $E_3(h)$ is calculated from equation (28) as follows:

$$E_3(h) = \left[E_k^{1.77} - 231\frac{Z^2}{A}\times 1.77\times \tilde{h}\right]^{1/1.77} \tag{29}$$

Energy interval region 4. The condition $\{E_k, E_4(h)\} \in [200, 850]$ MeV/n is fulfilled. E_k is the initial energy and $E_4(h)$ is the final energy for this case of energy decrease law.

$$\tilde{h} = \int_{E_4(h)}^{E_k} \frac{dE}{68\frac{Z^2}{A}E^{-0.53}} = \frac{E_k^{1.53} - E_4^{1.53}(h)}{68\times\frac{Z^2}{A}\times 1.53} \tag{30}$$

The final energy $E_4(h)$ is calculated from equation (30) as follows:

$$E_4(h) = \left[E_k^{1.53} - 68\times 1.53\times\frac{Z^2}{A}\times\tilde{h}\right]^{1/1.53} \tag{31}$$

Energy interval region 5. The condition $\{E_k, E_5(h)\} \in [850, 5000]$ MeV/n is fulfilled. E_k is the initial energy and $E_5(h)$ is the final energy for this case of energy decrease law.

$$\tilde{h} = \int_{E_5(h)}^{E_k} \frac{dE}{1.91\times\frac{Z^2}{A}} = \frac{E_k - E_5(h)}{1.91\times\frac{Z^2}{A}} \tag{32}$$

The final energy $E_5(h)$ is calculated from equation (32) as follows:

$$E_5(h) = E_k - 1.91\times\frac{Z^2}{A}\times\tilde{h} \tag{33}$$

Energy interval region 6. The condition $\{E_k,\ E_6(h)\} \in [5000,\ 5\times10^6]$ MeV/n is fulfilled. E_k is the initial energy and $E_6(h)$ is the final energy for this case of energy decrease law.

$$\widetilde{h} = \int_{E_6(h)}^{E_k} \frac{dE}{0.66\times\frac{Z^2}{A}E^{0.123}} = \frac{E_k^{0.877} - E_6^{0.877}(h)}{0.66\times\frac{Z^2}{A}\times 0.877} \tag{34}$$

The final energy $E_6(h)$ is calculated from equation (34) as follows:

$$E_6(h) = \left[E_k^{0.877} - 0.66\times 0.877\times\frac{Z^2}{A}\times\widetilde{h} \right]^{1/0.877} \tag{35}$$

Electron production rate in 5 energy intervals with charge decrease in the ionization losses function

The improved CR ionization model includes the electron production rate terms in 6 energy intervals of the ionization losses function and 5 intermediate transition region terms between the basic intervals. The lower boundary of integration E_{min} is chosen as the maximum of the atmospheric cut-off and the geomagnetic cut-off rigidity [1, 2]. The case of vertical penetration of cosmic rays is considered. This improved model can be extended to the 3-dimensional case in the Earth environment with introduction of the Chapman function [1, 2], which takes into account the Earth sphericity. Then all possible combinations of initial and final energy intervals of CR penetration must be included.

Lower boundary of integration E_{min}: The following case of lower integration boundary is assumed [1, 2]:

$$kT \le E_{A1}(h) \le E_{min} \le 0.15 < E_a \text{ MeV/n} \tag{36}$$

The next equation presents the corresponding electron production rate. $\rho(h)$ is the neutral density [1], Q is the energy for formation of one electron-ion pair [2]. Formula (1) from [1] is taken into account.

$$q(h) = \frac{\rho(h)}{Q}\left\{ 2570\int_{E_{min}}^{0.15} D(E)\left[E_1(h)\right]^{0.5} dE + 2570\int_{0.15}^{E_{0.15,2}(h)} D(E)\left[E_{21}(h)\right]^{0.5} dE + \right. \tag{37}$$

$$1540 \int_{E_{0.15.2}(h)}^{E_a} D(E)\left[E_2(h)\right]^{0.23} dE + 1540 \int_{E_a}^{E_{E_{32}}(h)} D(E)\left[E_{32}(h)\right]^{0.23} dE +$$

$$231 \times Z^2 \int_{E_{E_{43}}(h)}^{200} D(E)\left[E_3(h)\right]^{-0.77} dE + 231 \times Z^2 \int_{200}^{E_{200.4}(h)} D(E)\left[E_{43}(h)\right]^{-0.77} dE +$$

$$68 \times Z^2 \int_{E_{200.4}(h)}^{850} D(E)\left[E_4(h)\right]^{-0.53} dE + 68 \times Z^2 \int_{850}^{E_{850.5}(h)} D(E)\left[E_{54}(h)\right]^{-0.53} dE +$$

$$\int_{E_{850.5}(h)}^{5000} D(E)\frac{dE}{dh}\left[E_5(h)\right] dE + \int_{5000}^{E_{5000.6}(h)} D(E)\frac{dE}{dh}\left[E_{65}(h)\right] dE +$$

$$0.66 \times Z^2 \int_{E_{5000.6}(h)}^{\infty} D(E)\left[E_6(h)\right]^{0.123} dE$$

The ionization losses function in interval 5 with final energies $E_5(h)$ from (33) and $E_{65}(h)$ from (13) is calculated with the following formula:

$$(38) \qquad -\frac{1}{\rho}\frac{dE}{dh}\left[E_5(h)\right] = -\frac{1}{\rho}\frac{dE}{dh}\left[E_{65}(h)\right] = 1.91 \times Z^2$$

Conclusion

For many practical purposes of space weather, e.g., for the impact of cosmic rays on the ozone layer and formation of clouds in the troposphere, it is important to know precisely the cosmic ray induced ionization, its distributions and its variations with location, time, solar and geomagnetic activity. For this goal mainly two types of models are created: 1) analytical and 2) numerical.

In the present work a generalized *analytical* CR ionization model is proposed. A model of this type can be formed for every altitude, azimuth and zenith angle in the middle atmosphere and the lower ionosphere. Some combinations of energy intervals will arise which will cause the derivation of corresponding new formulas for energy integration in the model of the electron production rate with multi-interval ionization losses function, which was presented in this paper. The full 3D integration with introduction of the Chapman function [1, 2] can then be done which will provide

higher accuracy of the electron production rate model. The coupling of the energy intervals of the ionization losses function are done.

The computational procedure will choose logically the respective lower boundary and mathematical expression for third energy integration which corresponds to the ionization losses function energetic interval, the atmospheric cut-off or geomagnetic cut-off rigidity.

The integration over the full altitude-zenith-azimuth surface will be performed with account to the local Chapman function value which corresponds to the travelling substance path for the point which is calculated. These values differ from one another which will increase the number of interval combinations and the complexity of the electron production rate model.

The improved model can be realized on PC [1, 3]. It will possess higher accuracy [1, 4]. It can be applied for practical calculations [4-6]. The cosmic ray flux measurements [7, 8] are taken into account in the improved ionization model. This improved model includes a new 5 energy interval approximation of the ionization losses function, which is more adequate to experimental data [1, 9]. The charge decrease Z of particles in the additional second energy interval (interval 2 in (1) from [1]) is taken into account.

The analytical and numerical results in this work are important for the study of solar-terrestrial processes and space weather. Our results can explain quantitatively a lot of interesting phenomena in the terrestrial environment, as for example the appearance of abnormal ionization in the Earth's atmosphere and the lower ionosphere associated with solar cosmic ray flux enhancement on 23 February 1956 [10]. This is the most powerful and the greatest solar proton event which is observed up to now. One first corresponding evaluation of this event is given in the monograph [2].

The second type of CR ionisation models are the *numerical* models and codes. A simulation of extensive air showers with application of the CORSIKA (Cosmic Ray Simulations for KASKADE) programming system is made in [11]. The variation of atmospheric depth profile on different time scales is shown in [12]. The formation of the Pfotzer maximum of the electron production rate from cosmic rays is presented in [13-15]. But these investigations could not explain satisfactorily the first experimental data obtained yet more than 70 years ago.

The cosmic ray intensities in the stratosphere were measured for the first time in 1933 by the group of the Nobel price winner (1923) Robert Millikan [15]. The cosmic ray ionization at high altitudes was determined in 1934 by the group of another Nobel price winner (1927) Arthur Compton [16].

For better agreement between the experiments and the models our group began to model the ionization profiles in the whole atmosphere [17, 18]. Cosmic ray induced ionization rates q produced by galactic cosmic rays (GCR) in the Earth's atmosphere are obtained on the basis of a recent model using Monte Carlo simulations and parameterization of the primary spectrum. The simulations are carried out with CORSIKA 6.52 code using FLUKA 2006 and QGSJET II hadronic interaction subroutines. The energy deposit of GCR proton induced air showers is estimated. The ionization profiles for minimum and maximum of solar activity are calculated on the basis of the previously defined and obtained cosmic ray induced ionization yield function Y and parameterization of the cosmic ray spectrum [19, 20].

For our simulations we used the recent version CORSIKA 6.52 code [21] with corresponding hadronic interaction models FLUKA [22] and Quark Gluon String with JETs QGSJET [23]. The FLUKA 2006 code is used for simulation of hadronic interaction below 80 GeV/nucleon and QGSJET for hadronic interaction above 80 GeV/nucleon, respectively. The choice to use FLUKA hadronic interaction model is based on the recommendation of CORSIKA authors [24]. The hadronic event generator FLUKA is used only for the description of inelastic interactions below energy of several 100 GeV [25]. Within FLUKA these collisions are handled by different hadronic interaction models above, around and below the nuclear resonance energy range.

All these investigations require the efforts of big international collectives and collaborations [26]. One such collaboration is the Working Group 2 of Action 724 to European Co-operation in the field of Scientific and Technical Research - COST 724 "Developing the Scientific Basis for Monitoring, Modelling and Predicting Space Weather".

References

1. Velinov P. I. Y., L. N. Mateev, H. Ruder. Aer. Space Res. in Bulgaria (in this issue)
2. Velinov P., G. Nestorov, L. Dorman. Cosmic Ray Influence on the Ionosphere and on Radio-Wave Propagation. Sofia, BAS Publ. House, 1974.
3. Press W. H., B. P. Flannery, S. A. Teukolsky, W. T. Vetterling. Numerical Recipes in C++ - The Art of Scientific Computing. Cambridge, Cambridge University Press, 2002.
4. Usoskin I., K. Alanko-Huotari, G. Kovaltsov, K. Mursula. J. Geophys. Res., **101**, 2005, A12108.
5. Johnson R. E. Energetic Charged-Particle Interactions with Atmospheres and Surfaces, Berlin, Springer-Verlag, 1990.
6. Usoskin I. G., G. A. Kovaltsov. J. Geophys. Res., **111**, 2006, D21206.
7. Neher H. V. J. Geophys. Res., **76**, 1971, 1637-1651.

8. Van Allen J. A. Ch. 14 in "Physics and Medicine of the Upper Atmosphere". Univ. New Mexico Press, 1952.
9. Sternheimer R. Interaction of Radiation with Substance. In: Fundamental Principles and Methods of Particle Detection. Methods of Experimental Physics, Vol. V, A. Nuclear Physics, New York-London, Academic Press, 1961.
10. Bailey D. K. Proc. IRE, **47**, 1959, 2, 255-266.
11. Keilhauer B., J. Bluemer, R. Engel, H. O. Klages, M. Risse. Astroparticle Physics, **22**, 2004, 249-261.
12. Wilczynska B., D. Gora, P. Homola, J. Pekala, M. Risse, H. Wilczynski. Astroparticle Physics, **25**, 2006, 106–117.
13. Stozhkov Y. I. The role of cosmic rays in the atmospheric processes. J. Phys. G: Nucl. Part. Phys. **29**, 2003, 913-923 doi:10.1088/0954-3899/29/5/312, http://www.iop.org/EJ/abstract/0954-3899/29/5/312
14. Ziegler J. F. IBM J. Res. Develop. **42**, 1998, 1, 117.
15. Bowen I. S., R. A. Millikan. Cosmic-ray intensities in the stratosphere. Phys. Rev., **43**, 1933, 9, pp. 695-700. http://authors.library.caltech.edu/5451/
16. Compton A. H., R. J. Stephenson. Cosmic-Ray Ionization at High Altitudes, Phys. Rev., **45**, 1934, 7, pp. 441-450, http://adsabs.harvard.edu/abs/1934PhRv...45..441C
17. Velinov P. I. Y., A. Mishev. Compt. rend. Acad. bulg. Sci., **60**, 2007, 5, 495 - 502.
18. Mishev A., P. I. Y. Velinov. Compt. rend. Acad. bulg. Sci., **60**, 2007, 5, 513 - 518.
19. Mishev A., P. I. Y. Velinov. Compt. rend. Acad. bulg. Sci., **60**, 2007, 7, 725 - 734.
20. Velinov P. I. Y., A. Mishev. Compt. rend. Acad. bulg. Sci., **60**, 2007, 9, 947 - 956.
21. Heck D. et al., CORSIKA: A Monte Carlo Code to Simulate Extensive Air Showers. Report FZKA 6019 Forschungszentrum Karlsruhe, 1997.
22. Fasso A. et al., 2003. The physics models of FLUKA: status and recent developments, Computing in High Energy and Nuclear Physics 2003 Conference (CHEP2003), La Jolla, CA, USA, March 24-28, 2003, (paper MOMT005), eConf C0303241.
23. Kalmykov N, S. Ostapchenko, A. Pavlov. Nuclear Physics B - Proc. Suppl **52**, 1997, 17.
24. Heck D. et al., 2003. Influence of Low-Energy Hadronic Interaction Programs on Air Shower Simulations with CORSIKA. Proc. 28th Int. Cosmic Ray Conference, Tsukuba (Japan), 279-282.
25. Heck D. Low-Energy Hadronic Interaction Models. Proc. XIIIth Int. Symp. on Very High Energy Cosmic Ray Interactions, Pylos, Greece (2004) Nucl. Phys. B (Proc. Suppl.), 151, 2006.
26. Usoskin I., L. Desorgher, P. I. Y. Velinov, M. Storini, E. Flueckiger, R. Buetikofer, G. A. Kovalstov. Solar and Galactic Cosmic Rays in the Earth's Atmosphere. Progress Report of Working Group 2 of COST 724 Project. Acta Geophys., **66**, 2007 (in press)

ЗАКОНИ ЗА НАМАЛЕНИЕ НА ЕНЕРГИЯТА И СКОРОСТ НА ЕЛЕКТРОННАТА ПРОДУКЦИЯ В ОБОБЩЕНИЯ МОДЕЛ НА ЙОНИЗАЦИОННИ ПРОФИЛИ ОТ ЗАРЕДЕНИТЕ ЧАСТИЦИ НА КОСМИЧЕСКИТЕ ЛЪЧИ В ПЛАНЕТНИТЕ ЙОНОСФЕРИ И АТМОСФЕРИ С АПРОКСИМАЦИЯ НА ФУНКЦИЯТА НА ЙОНИЗАЦИОННИТЕ ЗАГУБИ ВЪРХУ 5 ЕНЕРГИЙНИ ИНТЕРВАЛА

П. Велинов, Х. Рудер, Л. Матеев

Резюме

В настоящата статия е разгледан един аналитично-числен метод за описание на проникването и йонизацията от ядра на космическите лъчи със заряд Z в планетните йоносфери и атмосфери. Скоростта на електронната продукция се изчислява с нови формули, които съчетават петте основни енергийни интервала на функцията на йонизационните загуби (dE/dh). Тя е 5 - интервална функция, която осъществява по-добра апроксимация на измерванията и експерименталните данни в сравнение с предходни резултати за 4 - интервална функция на йонизационните загуби. Геомагнитните прагове на отрязване и законите за намаление на енергията за различните интервали се използуват за създаване на междинна преходна област на енергията, която осъществява съчетаване на петте основни интервала на функцията на йонизационните загуби. Въвежда се един нов шести енергиен интервал за намаление на заряда в ниските енергии. Разглежда се вертикално проникване на космическите лъчи. Изчислени са съответните закони за намаление на енергията и началните стойности на границите на енергийните интервали.

Bulgarian Academy of Sciences. Space Research Institute.
Aerospace Research in Bulgaria. 22, 2008, Sofia

ENHANCING THE EFFICIENCY IN CHECKING CONSTRAINTS SATISFACTION WHEN PLANNING GROUND-BASED AND SPACE EXPERIMENTS, USING AN ALTERNATIVE PROBLEM

Atanas Atanassov

Solar-Terrestrial Influences Laboratory, Department in Stara Zagora
e-mail: At_M_Atanassov@yahoo.com

Abstract

Situational analysis lies in the basis of space and ground-based experiment planning. It is connected with the use of complex computational models of the environment and with verification of the restricting conditions, due to the character of the conducted experiments and the solved scientific tasks.

The present work proposes a formulation of the situational analysis on the basis of the finite abstract automata theory. On this basis, optimization of the situational analysis is suggested by formal schemes for adaptation to the conditions of the model environment. The efficiency enhancement is illustrated by results from the application of the proposed real-time optimization for photometric system control.

Keywords: *Situational analysis; planning and scheduling; constraints satisfaction; Mealy's automata; satellites experiments*

Introduction

To solve scientific and practical problems, connected with ground-based and satellite experiments and measurements, special activity is required – planning, connected with checking the satisfaction of a multitude of geometrical, physical and other constraints [1]. The planning will involve the application of an adequate mathematical model for analysis of a number of constraints and for their formulation, this representing the so-called situational analysis (SA). It allows to determine suitable

time intervals to conduct experiments and measurements and to optimize complex and expensive scientific programmes. The practice of the situational analysis involves check-up of different constraints, such as:

- A satellite is within the Earth's radiation belt zone, within the zone of the shock wave or of the magnetopause of the Earth's magnetosphere [2];
- The angle between the observed object and another bright object on the celestial sphere (Sun, Moon, bright starts) is less than a definite value [3];
- A satellite passes over a territory of the Earth's surface [4];
- A satellite is within the visibility zone of an observation point on the Earth's surface;
- The Earth's radiation background is within admissible limits with a view to the conducted measurements [3];
- A satellite is properly orientated towards the force lines of the Earth's magnetic field;
- The magnetic force line where the measuring device is located, pierces the polar oval region;
- The visibility axis of the optical instrument falls within a region of the celestial sphere or the Earth's surface, which is interesting from the point of view of the conducted experiment.

The appropriate orientation of the model of the space platform with scientific equipment on board in respect to the simulation of experiments and measurements in the objective space will be called situation. Each situation can be presented in the general case by a predicate function S:

(1) $$S = S(\vec{R}, A, t) = \overline{0,1}$$

In (1) $\{\vec{R}\} = \langle \vec{r}_1(t), \vec{r}_2(t), ..., \vec{r}_n(t) \rangle$ - the multitude of the radii-vectors of the objects in the model space, $\{A\} = \langle \alpha_1(t), \alpha_2(t), ..., \alpha_m(t) \rangle$ - the multitude of vector or scalar fields, describing certain properties of the model space – and t – the time. Due to the complex character of the processes and phenomena, occurring in space, S may have a complex analytical representation.

Actually, we can have a combination of several constraints restricting conditions. In addition, we shall examine such conditions that are independent on one another (none of them is represented by the others). Thus, the multitude of conditions $\{\gamma_i\}$ can be juxtaposed to the multitude of predicate functions $\{s_i\}$.

The implementation of situation S will require the fulfillment of the following identity:

(2) $$S = s_1 \wedge s_2 \wedge ... \wedge s_n = 1$$

Presenting the planning process by finite automata formalism

It is obvious that the check-up of the identity of (2) requires calculation of the predicate functions s_i, which is connected with considerable calculation (expenses). According to (2) we can consider $\{s_i\}$ as being arranged in the sense of consequences computational consequence of calculation of s_i:

(3) $$S = (...(s_1 \wedge s_2) \wedge ... \wedge s_{n-1}) \wedge s_n$$

In (3), every predicate function s_i is calculated if those before it in $\{s_i\}$ have valuc of one. A measure of the ineffectiveness of scheme (3) can be the expenses for calculation of s_i for i<k, when $s_k = 0$.

Every discrete-deterministic model may be treated as abstract finite automata [6] and presented by 6-tuple:

(4) $$F = < Z, X, Y, \varphi, \psi, z_0 >,$$

functioning in discrete automata time. In (4), Z is the set of internal states, X is the set of incoming (input) signals, Y is the set of exit (output) signals, $z(t + \Delta t) = \varphi(z(t), x(t))$ – transition function and z_0 - initial state. For the classical Mealy's automata we have:

(5) $$\begin{cases} z(t + \Delta t) = \varphi(z(t), x(t)) \\ y(t) = \phi(z(t), x(t)) \end{cases}$$

We can apply this formal approach to the situational analysis. For moment t, the set of input signals $X(t)$ is the total of all values determining the discrete model. In many cases, these are vectors which determine the measurement place or the observation direction. The internal states of the situational analysis automata (SAA) can be characterized by information entropy, which is proportional to the quantity of processed information for one cycle, related to moment t. These states are related with the implementation of different model calculations, corresponding to the different constraints (situational conditions) γ_i whose satisfaction is

verified for every cycle of automata. We can reason that the state $z(t)$ will depend on the order of situational conditions γ_i in $\{\gamma_i\}$, and on the evolution of the mathematical model. With applying (3), each state of the automata will depend on the input signal x(t) only and will not depend on the previous state of the automata:

(6) $$z(t+\Delta t)=\varphi(x(t))$$

Automata without storage (such as the SAA) have only one stable state, which coincides with the initial one. Here we have an example of such automata which may be excited in terms of one cycle consecutively to different internal states. According to (5), for the SAA to function, no storage is necessary. This leads to simplicity and easy accomplishment. However, the lack of a storage and an appropriate transition function makes SAA ineffective for space research applications.

It is important to have in mind that each constraint γ_i is executed in a time interval T_i and is not executed in the next one T_i^*. The only practically imposed requirements for T_i and T_i^* that we shall take into account are that they are finite and $T_i, T_i^* >> \Delta t$.

A priori information about the fasibility execution of conditions γ_i with time is laching. The check-up of γ_i can be verified by tracing the evolution of the deterministic numerical model, underlying (1).

Adaptation and Optimization Strategy

Problem A will be called inverse to B, if the objectives of A and B are opposite [7]. In our case, the problem for specifying the temporal interval, in which the conditions $\{\gamma_i\}$ are satisfied, is inverse to the problem for specifying its adjacent in which they are not satisfied or, according to (3) at least one of them is not satisfied. Obviously, this problem is more efficient since it is sufficient to check whether only one condition γ_i is not satisfied.

(7) $$\overline{S} = \overline{s_1 \wedge s_2 \wedge \ldots \wedge s_k \wedge \ldots \wedge s_n} = \overline{s_1} \vee \overline{s_2} \vee \ldots \vee \overline{s_k} \vee \ldots \vee \overline{s_n}$$

In fact, according to the above-said, $T_i >> \Delta t$. If, for discretization moment t we have $s_k = 0$, then we can presume that for $t+\Delta t$ this same condition will be satisfied again.

The strategy for implementing situational analysis is reduced to a consecutive verification of the conditions. If all conditions are satisfied, then we have

a situation of the searched type. In the case when for moment t a condition is found out which is not satisfied, there is not a situation. When for t' this condition is dissatisfied, once again the next condition in the set $\{\gamma_i\}$ is being verified.

Here, we shall examine two approaches, related with the organisation of the conditions' verifications. The first approach is related with replacement of the first encountered condition in the first place of the set of conditions [8]. Instead of (7), as a result of the disjunction commutation, we can write down:

$$(8) \qquad \overline{S} = \bar{s}_k \vee \bar{s}_1 \vee \bar{s}_2 \vee \ldots \vee \bar{s}_{k-1} \vee \bar{s}_{k+1} \ldots \bar{s}_n$$

The verifying algorithm for (3) always begins from the first element. The replacement of each dissatisfied condition when applying the alternative strategy is equivalent to adaptation of (3) to the conditions of the model medium. The latter is a function of the transition $\psi(z(t), x(t))$.

The second approach requires to treat the set of conditions as a ring-shaped structure [9]. This means that the last element of $\{\gamma_i\}$ is followed by the first one.

$$(9) \qquad \overline{S} = \bar{s}_k \vee \bar{s}_{k+1} \vee \ldots \vee \bar{s}_n \vee \bar{s}_1 \vee \ldots \vee \bar{s}_{k-1}$$

This allows to apply the alternative strategy without rearranging the set of conditions only by moving a pointer along this ring-shaped structure until finding a dissatisfied condition. The shifting of the pointer along the ring-shaped structure until finding dissatisfied condition plays the role of a transition function. A situation is found out when a full circle along the ring-shaped structure is completed.

Analysis and Effectiveness Estimation

In order to estimate the effectiveness, the following two cases are of interest:

a) The predicate functions s_i on the basis of which identify (2) is verified, which are characterized by equal or almost equal computational expenses $\xi_i \cong const$;

b) The predicate functions s_i on the basis of which identify (2) is verified, which are characterized by different computational expenses.

In the first case, for each step of SAA performance based on the offered strategy, $(k-1).\xi$ computational units will be saved in contrast to the direct application of (3) where the dissatisfied condition is in the k^{th} position. The transition

function will not be of any significance since the different predicate functions are characterised by equal computational expenses.

As an illustration we will give an analysis, related with the transition of a couple of satellites over the visibility zone of a ground-based observation station, proper mutual configuration and distance between them, lack of factors hampering the observation of satellites from the Earth (Moon), lack of factors preventing the observation of the optical instruments from one of the satellites (Moon, Sun). This situation considers the problem for a possible synchronous observation form the Earth's surface and from one of the satellites of an emission, caused by neutral particles injection for studying the Earth's ionosphere (The *Aktiven* experiment).

In the second case where the computational expenses for the separate conditions are different, for the saved computational expenses we can write down $\sum_{i=1}^{k-1}\xi_i$. Without going into detail, it is worth noting that the arrangement of the conditions in $\{\gamma_i\}$ (3) and the transition function $z(t)$ are significant. Despite the quasicausal choice of an unfeasible condition in applying the alternative strategy, there are certain peculiarities. The first algorithm may carry a condition with large computational expenses to the first position. If this condition remains in the beginning of the set of conditions $\{\gamma_i\}$, the effectiveness of the analysis may be reduced. With the second approach such a problem does not exist as a result of the constant arrangement of the conditions and their circulation.

We have an example of a situational problem of the second type when including in the analysis the conditions, related with the parameters of the medium α_i . The condition for the angle between the satellite velocity vector and the magnetic field vector $\vec{B}(\vec{r},t)$ to have an appropriate value is important for the analysis of active experiments. (The *Aktiven* experiment). The radiation background is important in astrophysical observations of x-ray sources [3]. When phenomena connected with particles dissipation in the polar oval region are investigated, it is important to measure the magnetic field in the region of the force lines along which the dissipation occurs [6]. In all presented cases, the verified conditions are related with complex model computations.

Example for application

For effective real time control of a zenith photometer [11] it is required to check conditions,related with the Sun's and Moon's position for elimination

of their influence on the measurements. The restriction for the Sun is to be at an angle φ' below the horizon and for the Moon - to close with direction to the zenith an angle bigger than φ''

Two check-up conditions contain two versions for arranging $\{\gamma_i\}$. In the first one, the verification of the condition for the Sun is in the first place and for the Moon - in the second; in the other version, the places are reversed. To perform the analysis for one year period with a time step of one minute without optimization, 412,264 checks for the Sun's position and 525,600 checks for the Moon's position are made, respectively, by the first arrangement version, 525,600 and 171,025 checks are performed by the second version. The application of the optimization approach leads to solving the same task with 468,161 and 186,659 checks only.

Conclusion

This paper examines the situational analysis of space experiments on the basis of the finite automata theory. It allows formalization of the computing processes and search of possibilities for optimization of the computing algorithms at a higher abstraction level.

The advantages of the suggested approach are obvious even in the examined simple application case. The properties of the examined optimization can be revealed best in the analysis with participation of a larger number of check-up conditions, which, on the other hand, are connected with considerable computational expenses. During real-time experiment planning and, especially during their control, the analysis' acceleration can be of major importance.

Besides, a strategy is examined for optimization of the situational analysis, based on the solution of the inverse problem. This, on the other hand, allows to apply algorithms for adaptation to the conditions of the model environment. Two algorithms are suggested whose application can considerably enhance the analysis' efficiency.

References

1. Boddy Mark S., Bonnie H. Bennett, Brian A. Isle, Rainer A. Isle. NASA Planning and Scheduling Applications: Emerging Technologies and Mission Trends., Adventium Labs Tech Report, May 28, 2004.
http://www.adventiumlabs.org/Publications/PlanningGrantFinal28 May2004.pdf
2. Прохоренко В. И. Ситуационный анализ орбит хвостового и аврорального зондов в проекте "ИНТЕРБОЛ", Пр. 1037, ИКИ АН СССР, 1985 г.

3. Гайдаров П., А. Суханов, Ат. Атанасов, В. Рязанова. Планирование астрофизических экспериментов в проекте "Шипка", Пр.-1534, ИКИ АН СССР, 1989 г.
4. Бажинов И.К. и соавт., Навигационное обеспечение полета орбитальнго комплекса "Солют-6"-"Союз"-"Прогресс", М.: Наука, 1985. 376 с.
5. Kuzmin A.K. et al. UV-spectrometer in INTERBALL Project to map ionospheric characteristics in the magnetic field line footprint from the satellite Auroral Probe, in "Interball mission and payload", 401p.,SRI , 1996.
6. Советов Б.Я., С.А. Яковлев. Моделирование систем, М.: "Высшая школа", 1985. 271 с.
7. Гладун В.П. Планирование решений. Киев , "Наукова думка", 1987.
8. Атанасов А.М. Адаптивен алгоритъм за ситуационен анализ при следене на обекти в силно изменяща се обстановка. Депон. ЦИНТИ - Hg 1399/89
9. Atanassov At. Adaptive Approach for Situational Analysis of Space Experiments. Balkan Physics Letters, (Proc. Suppl.) BPU-4 (2000), 577-580.
10. http://www.izmiran.ru/projects/APEX/
11. Atanassov At. Development of a Computer Model to Control a Scanning Photometric System for Atmospheric Investigations, Balkan Physics Letters, (Proc. Suppl.) BPU-4 (2000)

ПОВИШАВАНЕ НА ЕФЕКТИВНОСТТА НА СИТУАЦИОННИЯ АНАЛИЗ ПРИ ПЛАНИРАНЕ НА НАЗЕМНИ И КОСМИЧЕСКИ ЕКСПЕРИМЕНТИ С РЕШАВАНЕ НА "АЛТЕРНАТИВНА ЗАДАЧА"

А. Атанасов

Резюме

Ситуационният анализ стои в основата на планирането на космчески и наземни експерименти. Той е свързан с използването на сложни изчислителни модели на средата и проверка на ограничителни условия произтичащи от същността на провежданите експерименти и решаваните научни задачи.

В настоящата работа се предлага формулировка на ситуационния анализ на основата на крайните абстрактни автомати. На тази основа се прави оптимизиране на ситуационния анализ чрез формални схеми за адаптация към условията на моделната среда. Повишаването на ефективността се илюстрира с резултати от прилагане на предлаганата оптимизация в реално време за управление на фотометрична система.

Bulgarian Academy of Sciences. Space Research Institute.
Aerospace Research in Bulgaria. 22, 2008, Sofia

AN ADAPTIVE PARALLEL INTEGRATOR OF ORDINARY DIFFERENTIAL EQUATIONS SYSTEM FOR SPACE EXPERIMENT SIMULATION

Atanas Atanassov

Solar-Terrestrial Influences Laboratory - Bulgarian Academy of Sciences
e-mail: At_M_Atanassov@yahoo.com

Abstract

Different possible sources are discussed for enhancement of the calculation time when solving ordinary differential equations systems to forecast space objects' motion. This paper presents an approach for building an integrator of ordinary differential equations systems for simultaneous solving of motion equations of multiple objects. A parallelization of calculation on the base of threads is offered. A method for synchronization is presented. The technological advance and the invasion of multi-core processors make actual the examined approach for developing an integrator of ordinary differential equations systems.

Keywords and Phrases: *ordinary differential equations integrator/solver; adaptive algorithms; multi threading, thread parallelism;*

1. Introduction

Computer simulation has long ago become a tool for quality improvement of space experiments, which is applied in their design, planning and control. The various dynamical models are based on ordinary differential equations systems (ODEsS). The intensive use of such models (for example, of the motion of different bodies in the near Earth's or other planets' space) may be related with increase of the consumed processor time in the case of computer accomplishment. When tackling this issue (if it exists), the focus is on creating more effective numerical methods and computer programs for ODEsS integration.

The delay in the course of processor frequency increase is expected to be

successfully compensated with increase of their core number. Presently, the AMD and Intel available double core processors of on the market have reasonable prices. Quad core processors are manufactured, too. However, soon expected the manufacturing of single-core processors and the development of 8 and 16 core processors is expected to end soon. Multi-core processors are intended at disposing two or more processors on one crystal in order to work at lower voltage and with lower frequency. The latter aims to reduce released heat, which is a major issue in increasing operational frequency. This, however, does not ensure automatical raising of the calculation power of multi-core processors, equivalent to one-core processors. It is also necessary to transform serial algorithms into parallel ones, taking into account the specifics of the processor's architecture. The perspective for personal computers to compete with the expensive computer clusters, specialized in parallel calculations, will require a change in algorithmic thinking and development of new program tools on the basis of the calculation processes' parallelization.

The integration of ODEsS as an initial value problem (IVP) can be written down in the following way:

$$\dot{y}_i = \varphi_i(t, y_i), \quad y_i(t_0) = y_i^0 \tag{1}$$

The solution of (1) as IVP is obtained as a series of values of $y(t_m)$ for $t_m = t_0 + m.\Delta t$.

The keystones in parallel calculations theory within the context of IVP are formulated in a series of articles [1, 2, 3]. Similar issues are examined in [4] with application in the field of orbital dynamics. In [5, 6] the prognostication of the orbital motion of a large number ($N \sim 10^3$) of objects is examined on the basis of analytical methods. [7, 8] expose other approaches to solving the same problem, however, with different numerical methods.

The development of ODEsS integrators (ODEsSI) and their application for integration of space bodies' motion equations is one of the fields, which has always been attractive for its final objective - achieving better precision and effectiveness.

In this work, we will focus on an approach for building a parallel ODEsSI (PODEsSI), based on explicit Runge-Kutta-Fehlberg (RKF) schemes [8,10] and on its application for satellite orbits' integration. Preliminary results are obtained, using double-core processors. The calculation model's parallelization is based on threads.

2. Space objects motion equations

We can examine the solution of a satellite's motion equations (material objects, N>>1) in the gravitational field of a central body as IVP:

$$\ddot{\vec{r}}_n = \vec{f}_n(t, \vec{r}_n, \dot{\vec{r}}_n), \vec{r}_n(t_0) = \vec{r}_n^0, \dot{\vec{r}}_n(t_0) = \vec{v}_n^0. \quad (2)$$

Index n is used to denote the object: $1 \le n \le N$. System (2) is of the second order and can be reduced to a first order system:

$$\begin{cases} \dot{\vec{r}}_n = v_n \\ \dot{\vec{v}}_n = \vec{f}_n(t, \vec{r}_n, \dot{\vec{r}}_n), \quad \vec{r}_n(t_0) = \vec{r}_0, \quad \dot{\vec{r}}_n(t_0) = \vec{v}_0 \end{cases} \quad (3)$$

The right - hand side of (2) can include in addition to the acceleration from a central gravitational force $\vec{f}_n^{grav}$ different kinds of perturbations [11, 12]:

$$\dot{f}_n(t, \vec{r}_n, \dot{\vec{r}}_n) = \vec{f}_n^{grav} + \vec{f}_n^{atm} + \vec{f}_n^{Sun} + \vec{f}_n^{Moon} + \vec{f}_n^{light} + \vec{f}_n^{eldin} \quad (4)$$

By analogy with [1] we can point to several reasons as a result of which the solution of one IVP can be related with the use of much calculation time:

- complex calculation models are used for the calculation of different disturbances in (4);
- the mathematical model describes the motion of numerous objects;
- the integration interval $[t_0, t_{end}]$ is large;
- multiple solutions of IVP to determination of the simulation model's parameters.

The computer modelling of multi-satellite space experiments, involving a large set of instruments, intended to solve multiple scientific problems, aimed at producing numerous parameters, contains the above-mentioned reasons. Appropriate orbital elements are determined during the experiment design stage, which, within the expected instrument operation period, are expected to have optimum conditions for their implementation. The simulation of different aspects of the instruments' operation in model conditions allows specifying some of their parameters (optical, mechanical, electrical, informational, etc.) in order to obtain statistically reliable results.

3. Possibilities for parallelization

Gear [1] points to two approaches for calculation parallelization by numerical integration of ODEsS:

- Parallelism across the method;

– Parallelism across the system.

The phrase „parallelism across the method" expresses the possibility for different calculation stages within the framework of one method to be executed independently and simultaneously on different processors. This paper will be focused mainly on the explicit classical RKF schemes, which are used to build the integrator of ODEsS (OSEsSI). The used schemes feature different precision order and are based on the calculation of functions $\vec{f}_n(t_k, \vec{r}_n, \dot{\vec{r}}_n)$ for $t_{in} < t_i < t_{in-1}$. Each of the stages, related with calculation of $g_{i,0}$ and $g_{i,k}$ for moments t_k is based on the previous one: $\vec{f}_n(t_{k-1}, \vec{r}_{n,k-1}, \dot{\vec{r}}_{n,k-1})$ and $g_{i,k-1}$. As a result, the possibilities for parallel calculations of coefficients $g_{i,k}$ are restricted with regards to the separate system equations. Each of the six $g_{i,k}$ can be calculated on a separate processor; with two processors, each of them can be computed by three coefficients $g_{i,k}$. The implementation of such kind of parallel calculations, however, requires special compilers.

The phrase "parallelism across the system" means that one equation or a group of equations, part of a large ODEsS, can be solved on a separate processor. This kind of parallelism reflects to some extent the character of the solved problem. It is very suitable for application in simultaneous integration of a large number of equations of type (2).

The parallelization of adaptive integrator of ODEsS, based on "parallelism across the system" is examined in this paper.

4. Multi body ODEsS integrator

Fig. 1 shows the functional diagram of the integrator. The basic subroutine ***rkfasd*** controls the choice of the integration scheme. The classical schemes of Feldberg [9,10] are used – subroutines ***prkf0a, prkf2a, prkf4a, prkf6a*** and ***prkf8a***. Subroutine ***kalkgr*** serves to evaluate the error and to verify whether the current scheme is suitable or, another one should be selected.

Integration with a variable step within interval Δt is selected, if the scheme with maximal precision is not sufficient. Subroutine ***pertur*** calculates the right - hand side of (3). The serial version of the integrator is directly called on the basis of subroutine ***rkfasd***. It is designed as storage automata and can integrate simultaneously many ODEsS with individual scheme choice for each system or step-size control, if the precision of the highest-order scheme is not sufficient.

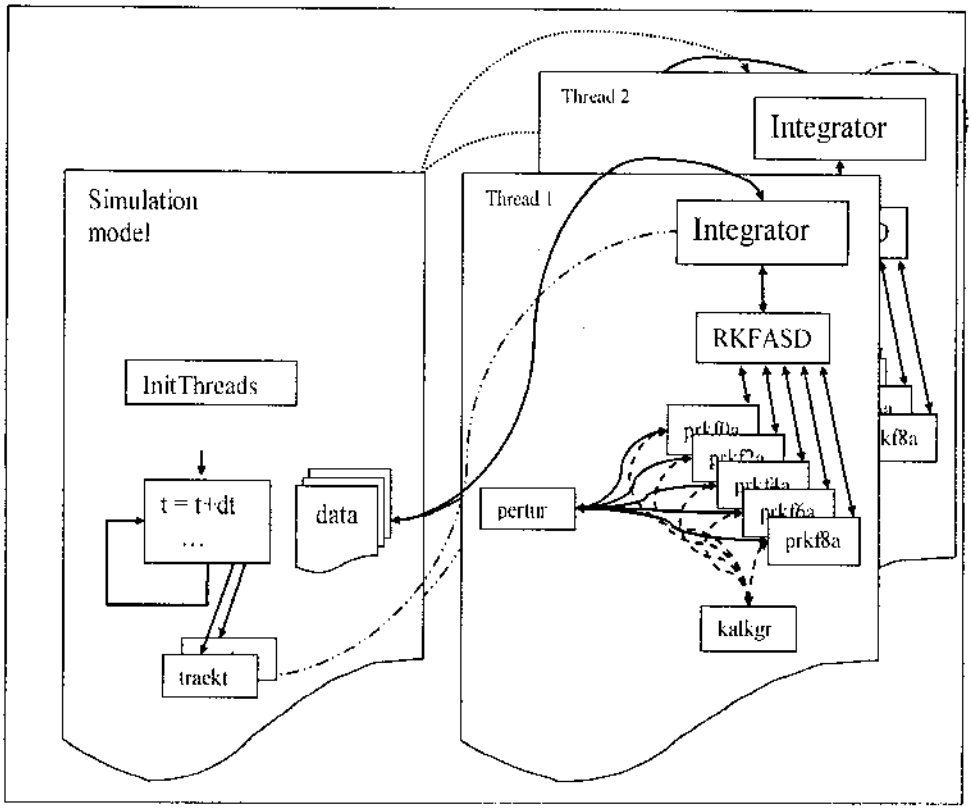


Fig. 1. General functional diagram of the integrator.

The parallel version of ODEsSI (PODEsSI) is initialized before the first integration step. A certain amount of integrator threads (Ith) are started according to the number of processor cores. These threads are left in suspended state up to the first addressing PODEsSI. The ITh perform the selection of different ODEsS in competitive conditions on the basis of a synchronizing code, shown below and proceed until reaching the maximum number. The choice is protected from doubling of ODEsS and solving one system more than once by different IThs on the basis of a special event. While one ITh is performing the choice of ODEsS, the remaining threads are waiting for the choice to be completed. When one ITh finishes the integration of the respective system for the current moment t and step Δt, it tries to select the next system, if there are any remaining.

5. Threads' synchronization

All calculations, related with the integration of ODEsS are portioned and each portion is related with one specific time step. This requires synchronization of the main and secondary threads ITh. The respective code in the basic parent's thread is, as follows:

```
  SUBROUTINE traekt(num_sat,dt)
                ...
 glb_counter = 0
a:DO i=1,num_th
    k= SetEvent(thread_par(i)%ha_beg) !  Events for releasing threads
 END DO a
   k= WaitForMultipleObjects(num_th, ha_end,WaitAll,Wait_infinite)
b:DO i=1,num_th
    k= ResetEvent(thread_par(i)%ha_end) ! Events for starting threads
 END DO b
                ...
END SUBROUTINE traekt
```

The above fragment of subroutine **traekt** illustrates the control of the PODEsSI. The counter ***glb_counter*** which counts the number of ODEsS is zeroed before starting the integration. In the first **a:DO** construct, the event objects ***ha_beg*** for starting the threads' operation are set out in a signalled state. After this, subroutine **traekt** passes to a waiting state until all treads ascertain that all systems are solved. The second **b:DO** construct makes non-signalled the event objects ***ha_end*** which point to the end of each ITh. The latter ensures the waiting state in the integration step.

Instead of the traditional call of the integrator as a subroutine, by analogy with the serial version without parallel calculations (connection by address) we have now synchronization of the parent's thread and the IThs. Instead of transmitting data by actual and formal arguments, now global data are used for transferring the address of allocatable user-defined type data which contain the generic data - coordinates of radius vectors, velocities, tolerance, etc.

The threads IThs perform two kinds of synchronizations. The first one is with the basic thread and should perform the step integration of ODEsS. For this purpose, event objects with handlers ***ha_beg*** are used to start the choice of ODEsS and their solving; other event objects with handlers ***ha_end*** (by one for each ITh), signal the end of the integration. These event objects are the same as the above described in connection with subroutine **traekt**. The second communication is

between IThs only and is devoted to concurrent ODEsS. As a result of this synchronization it is possible that while one thread is choosing a subsequent ODEsS, all others are waiting. In this way, the solution of one ODEsS with different threads is avoided. The allocation of the systems among IThs is achieved within the frames of "**b: DO WHILE()**" construction, shown in the next fragment:

```
SUBROUTINE     Integrator (th_id_num)
   ...
a: DO WHILE(.true.)
     k=  WaitForSingleObject(ha_beg,WAIT_INFINITE) ! Event for threads starting
  b: DO WHILE(glb_counter.LT. numsat)              ! ODEsS concurrent distribution
     k= WaitForSingleObject(ha_1,WAIT_INFINITE)    ! prohibition of ODEsS distribution
                glb_counter= glb_counter + 1;       ! serial number of ODEsS
                loc_counter= glb_counter;           ! remembering in local thread's
                                                     storage
        k= SetEvent(ha_1)                           ! release of ODEsS distribution
         IF(loc_counter.GT.numsat) EXIT
      CALL  rkfasd(loc_counter,m,transfer_data(loc_counter)%t , ... )
    END DO b                                        ! end of concurrent distribution
   k= ResetEvent(thread_par(th_id_num)%ha_beg)
   k=    SetEvent(thread_par(th_id_num)%ha_end)
 END DO a
   ...
END SUBROUTINE  Integrator
```

6. Analysis of the effectiveness

The entire PODEsSI code is developed on Compaq Visual Fortran Professional Edition v.6.6 under WindowsXP, using QuickWin library. Tests were conducted on one core Intel and AMD processors for checking the correctness. Experiments were conducted on a double core Athlon to establish the effectiveness of the integration of motion equations of six satellites with different orbits. The experiments have shown that, with double core processors, the calculation time is not possible to be decreased. The serial code execution time which is transformed into a thread increases about twice. Besides, time is necessary for the threads' synchronization between parent's and children's IThs. As a result, the execution time of the parallel version exceeds the time of the serial one by about 30%.

7. Conclusion and future work

The basic part of the theoretic and experimental investigations which are related with different aspects of solving ODEsS-IVP has private character and

restricted practical application. The increase of the effectiveness and the investigation of ODEsS integration have major significance in solving complex problems related with difficult simulation models of physical reality. These models can describe both moving satellites and different satellite subsystems. The whole simulation model can include models of instruments and mechanisms, of the environment where the measurements are conducted, as well as of investigation method and investigated object. Some of the private models can be deterministic and others - stochastic. In such conditions, both the IThs and other threads can be initialized in completely competitive conditions - for example, for calculation of the geophysical parameters by orbits of satellites included in the model, the geomagnetic field, the directions of the instrument optical axis, for measurements imitation and analysis of the statistical reliability. Further analysis is necessary to explain the behaviour and effectiveness of the adaptive parallel integrator in the framework of such a global model.

We can consider normal the above results obtained using double core processors. It is necessary to conduct experiments with quad core processors. We can expect decrease of the execution time of a parallel code compared to the execution time of a serial code.

References

1. Gear C. W. The Potential for Parallelism in Ordinary Differential Equations, in Computational Mathematics, ed. Simeon Fatunla, Boole Press, Dublin, pp33-48, 1987.
2. Jackson K. R. A Survey of Parallel Numerical Methods for Initial Value Problems for Ordinary Differential Equations, IEEE Transactions on Magnetics, 27 (1991), pp. 3792-3797.
3. Stone L. C., S. B. Shukla, B. Neta. Parallel Satellite Orbit Prediction Using a Workstation Cluster, *International J. Computer and Mathematics with Applications*, **28**, (1994), 1–8.
4. Neta B. Parallel Solution of Initial Value Problems, *Proc. Fourth International Colloquium on Differential Equations*, D. Bainov, V. Covachev, A. Dishliev (eds), Plovdiv, Bulgaria, 18-23 August 1993, **2**, (1993), 19–42.
5. Phipps W. E., B. Neta, D. A. Danielson. Parallelization of the Naval Space Surveillance Satellite Motion Model, *J. Astronautical Sciences*, **41**, (1993), 207–216.
6. Neta B., D. A. Danielson, S. Ostrom, S. K. Brewer. Performance of Analytic Orbit Propagators on a Hypercube and a Workstation Cluster.
7. Atanassov At. Integration of the Equation of the Artificial Earth's Satellites Motion with Sclection of Runge-Kutta-Fehlberg schemes of Optimum Precision Order., Aerospace Research in Bulgaria, 2007, 21, 24-34.
8. Kolyuka Yu.F., T.I. Afanasieva, T.A. Gridchin. Precise Long-Term Prediction of

the Space Debris Object Motion, Proceedings of the 4th European Conference on Space Debris, 18-20 April 2005, ESA/ESOC, Darmstadt, Germany.

9. F e h l b e r g E. Klassische Runge-Kutta formeln funfter und siebenter Ordnung mit Schrittweitenkontrolle., Computing, v. 4, p.93-106, 2969.

10. F e h l b e r g E. Klassische Runge-Kutta formeln vierter und nidrigerer Ordnung mit Schrittweitenkontrolle und ihre Anwendung auf Warmeleitungsprobleme., Computing, v. 6, p. 61-71, 1970.

11. П р о х о р е н к о В. И.. Описание универсальной программы расчета навигационной информации о положении искусственного спутника Земли. Пр. ИКИ АН СССР №263, 1976, с.80.

12. П о п о в и ч П. Р., Б. Ц. С к р е б у ш е в с к и й. Баллистическое проектирование космических систем., Москва, Машиностроение, 1987, 239с.

ПАРАЛЕЛЕН ИНТЕГРАТОР НА СИСТЕМИ ОТ ОБИКНОВЕНИ ДИФЕРЕНЦИАЛНИ УРАВНЕНИЯ ЗА СИМУЛИРАНЕ НА КОСМИЧЕСКИ ЕКСПЕРИМЕНТИ

Ат. Атанасов

Резюме

Обсъждат се различни възможни източници за нарастване на изчислителното време при решаване на системи от обикновени диференциални уравнения за прогнозиране на движението на космически обекти. В настоящата работа се разглежда подход за построяване на интегратор на системи от обикновени диференциални уравнения предназначен за едновременно решаване на уравнения на движение на много обекти. Предлага се паралелизация на изчисленията на основата на нишки. Разглежда се начин за синхронизация.

Bulgarian Academy of Sciences. Space Research Institute.
Aerospace Research in Bulgaria. 22, 2008, Sofia

NEUTRINO AND GRAVITON MASS ESTIMATIONS BY A PHENOMENOLOGICAL APPROACH

Dimitar Valev

Stara Zagora Department, STIL - Bulgarian Academy of Sciences
e-mail: valev@gbg.bg

Abstract

The ratio between the proton and electron masses is shown to be close to the ratio between the strong and electromagnetic interaction coupling constants at Extremely Low Energy (ELE). Based on the experimental data, this relation has been extended for the weak and gravitational interactions, too. Thus, a mass relation has been found, according to which the rest mass of the Lightest Free Massive Stable Particle (LFMSP), acted upon by a particular interaction, is proportional to the coupling constant of the respective interaction at ELE. On the basis of this mass relation, the electron neutrino and graviton masses have been approximately estimated to 2.1×10^{-4} *eV/c*^{2} *and* 2.3×10^{-34} *eV/c*^{2}*, respectively. The last value is of the order of the magnitude of* $\frac{\hbar H}{c^2}$*, where H is the Hubble constant. It is worth noting that this value has been obtained by fundamental constants only, without consideration of any cosmological models.*

1. Introduction

Although the neutrino and the graviton belong to different particle kinds (neutral lepton and carrier of gravity, respectively), they have some similar properties. Both particles are not acted upon by the strong and the electromagnetic interactions, which makes their detection and investigation exceptionally difficult. Besides, both have masses that are many orders of magnitude lighter than the masses of the rest particles and they are generally accepted to be massless.

Decades after the experimental detection of the neutrino by Reines and

Cowan [1], it was generally accepted that the neutrino mass $m_{0\nu}$ is rigorously zero. In the Fermi's theory of β-decay [2] as well as in the electroweak theory [3, 4] and hence, in the Standard Model (*SM*), the neutrinos have been accepted massless. Despite this, attempts to determine the neutrino mass have been made as early as it was found. The recent experiments bound $m_{0\nu}$ on the top and its upper limit has decreased millions of times in the latest experiments, as compared to the initial estimations of Pauli [5].

The first experiment, hinting that the neutrino probably possesses a mass, was carried out by Davis et al. [6]. The total flux of neutrinos from the Sun was found about three times lower than the one, predicted by theoretical solar models, thereby creating the problem for the solar neutrino deficit. This discrepancy can be explained if some of the electron neutrinos transform into another neutrino flavour. Within the frame of the *SM*, however, there is no place for massive neutrinos and neutrino oscillations. As a result, the detection of neutrino oscillations appears crucial for the *SM* and it requires its extension in direction to the *GUT*, *SUSY*, Superstring/M-theory and others.

Later, the experimental observations showed that the ratio between the atmospheric ν_μ and ν_e fluxes is less than the theoretical predictions [7, 8]. This discrepancy became known as the atmospheric neutrino anomaly. Again it could be explained by the neutrino oscillations. The crucial experiments with the 50 *kton* neutrino detector Super-Kamiokande found strong evidence for oscillations (and hence - mass) in the atmospheric neutrinos [9].

The direct neutrino measurements allow to bound the neutrino mass. The upper limit for the mass of the lightest neutrino flavour ν_e was obtained from experiments for measurement of the high-energy part of the tritium β-spectrum and recent experiments yield ~ 2 eV/c^2 upper limit [10, 11]. As a result of the recent experiments, the upper mass limits of ν_μ and ν_τ reach 170 keV/c^2 [12] and 18.2 MeV/c^2 [13], respectively. The Solar and atmospheric neutrino experiments allow to find the neutrino mass squared differences $\Delta m_{12}^2 = m_2^2 - m_1^2$ and $\Delta m_{23}^2 = m_3^2 - m_2^2$, but not the absolute value of the neutrino masses. The astrophysical constraint of the neutrino mass is $\sum m_n < 2.2\ eV/c^2$ [14]. The recent extensions of *SM* lead to non-zero neutrino masses, which are within the large range from 10^{-6} eV/c^2 to 10 eV/c^2.

Although the graviton hasn't been experimentally detected yet, most of the quantum gravity models posit a neutral spin-2 particle, appearing carrier (gauge boson) of gravity. Similarly to the case with the neutrino before 1998, the prevailing current opinion is that the graviton is massless. This opinion is connected with Einstein's theory of general relativity, where gravity is described by a massless field of spin 2 in a generally covariance manner. The nonzero graviton mass produces a finite range of gravity $r_g \sim \lambda\!\!\!^{-}_g = \frac{\hbar}{m_g c}$. There are two kinds of astrophysical methods for estimation of the upper limit of the graviton mass (or low limit of $\lambda\!\!\!^{-}_g$) – static and dynamic methods. The static methods are based on the search of difference between Yukawa potential for massive graviton and Newton potential for massless graviton. The Solar system measurements infer $\lambda_g > 2.8 \times 10^{15}\ m$, that is equivalent to $m_g < 4.4 \times 10^{-22}\ eV/c^2$ [15]. Rich galactic clusters allow to estimate $m_g < 2 \times 10^{-29}\ \ h^{-1}\ eV/c^2$ [16], where $h \approx 0.65$ is a dimensionless Hubble constant. This is the lowest limit of the graviton mass and therefore this value is used in the present paper. The dynamic methods are based on the differences of the emission and propagation of the gravitational waves from binary stellar systems in cases of massless or massive graviton. The possibilities of the astrophysical measurements to limit the graviton mass, including Laser Interferometer Space Antenna (*LISA*), are still of the order of the static tests magnitude in the Solar system [17]. Will and Yunes [18] suggest considerable improvement of these results for extra-Galactic massive binaries.

The theoretical estimations of the graviton mass are most often based on the assumption that the Compton wavelength of the graviton $\lambda\!\!\!^{-}_g$ is close to the Hubble distance $c/H \sim 1.4 \times 10^{26}\ m$, which produces a value of the graviton mass $m_g \sim \frac{\hbar H}{c^2} \sim$ $10^{-33}\ eV/c^2$, where $H \approx 65\ km\ s^{-1} Mpc^{-1}$ is the Hubble constant [19, 20]. Woodward et al. [21] obtain a value of the graviton mass 4.3×10^{-34} eV/c^2 for an infinite stationary universe, but the expansion of the Universe is a fact, long ago established.

2. Determination of the coupling constants at extremely low energy

Four fundamental interactions are known - strong, electromagnetic, weak and gravitational, whose strength decreases in the direction from strong to gravitational interaction of dozens orders of magnitude. A measure for the interaction strength is a dimensionless quantity, namely the coupling constant of the interaction (α_i), which is determined from the cross section of the respective processes. It is known that the coupling constants of the interactions are running [22, 23]. With increase of the processes energy, the coupling constant of the strong interaction decreases and reaches $\alpha_s \approx 0.11$ at 189 [24] *GeV*, and the rest coupling constants increase. Since modern experiments are performed with energy of hundreds *GeV*, a value of the weak coupling constant, close on the electroweak scale, approaches the coupling constant of the electromagnetic interaction. The electromagnetic coupling constant increases exceptionally slow and remains $\alpha \approx 7.81 \times 10^{-3}$ even at energy $E \sim m_W c^2 \approx 80.4\ GeV$ [25]. The coupling constant would fulfill the role of unique property of the particular interaction when the energy is fixed. Recent unified theories predict unification of the four interactions on the Planck scale ($E \sim 10^{19}\ GeV$). Close to such energy, the four interactions and coupling constants merge. Thus, with the energy increase, the interactions (and the coupling constants) converge and become more hardly differentiated from each other. A similar situation occurs with the particles since their rest masses become a negligible part of the full masses. The purpose of this paper is to relate the rest masses of the lightest stable particles, acted upon by the respective interactions with the respective coupling constants. That is why it is necessary to determine the coupling constants under conditions when the interactions (and the particles) are differentiated as much as possible and, this is the case when the energy of the processes reaches extremely low value. Thus, each coupling constant obtains a unique asymptotical value at Extremely Low Energy (*ELE*), which will be designated by $\alpha(0)$.

The coupling constant of the electromagnetic interaction α_e is known the fine structure constant $\alpha_e(0) = \dfrac{e^2}{\hbar c} \equiv \alpha \approx 7.30 \times 10^{-3}$, where e is the elementary electrical charge, $\hbar$ - Planck constant, c – the light velocity.

The coupling constant of the weak interaction α_w is determined by the expression:

(1) $$\alpha_w = \frac{G_F m^2 c}{\hbar^3}$$

where G_F is Fermi coupling constant, m – the interacting particle mass.

As mentioned above, it is necessary the coupling constants to be determined at *ELE* of the processes. The lowest-energy process, involving the weak interaction, is the neutron β - decay. The process is running extremely slowly ($\tau \sim 880\ s$), and a minimum quantity of energy is released $\Delta E \approx 0.78\ MeV \sim m_e c^2 = 0.511\ MeV$. Therefore, m in expression (1) is substituted by the electron mass m_e and for the weak coupling constant at *ELE*, the following value is obtained:

(2) $$\alpha_w(0) \sim \frac{G_F m_e^2 c}{\hbar^3} \approx 3.00 \times 10^{-12}$$

This value of the weak coupling constant is close to the one accepted in [26], where $\frac{\alpha_w}{\alpha} \sim 10^{-10}$. The value of the weak coupling constant, obtained from (2) is minimal as a result of the minimal energy of the neutron β - decay. The rest decays involve the weak interaction, occurring with considerably higher energy and, as a result, the typical values of α_w are in orders of magnitude higher than such energy. For this reason α_w often determines by (1) replacing m with the pion or proton mass.

It is known [27] that in cases of slow nucleons scattering (without angular momentum) the strong coupling constant reaches a maximum value $\alpha_s \sim 10 \div 17$, in [28] α_s is accepted ~ 15 and in [29] $\alpha_s \sim 14$. As a result, the strong coupling constant at *ELE* is accepted $\alpha_s(0) \sim 14 \pm 1$ in this paper. The strong coupling constant is often accepted a unit but this value corresponds to high energy of experiments within the *GeV* sector.

The coupling constant of the gravitational interaction is determined by the expression:

(3) $$\alpha_g = \frac{G_N m_e m_p}{\hbar c} \approx 3.21 \times 10^{-42}$$

where G_N is the universal gravitational constant and m_e and m_p are the electron and proton masses. This estimation is a conventional medial value between

$$\alpha_g = \frac{G_N m_e^2}{\hbar c} \text{ and } \alpha_g = \frac{G_N m_p^2}{\hbar c}.$$

Thus the coupling constant of the strong, electromagnetic, weak and gravitational interactions at *ELE* are 14, 7.30×10^{-3}, 3.0×10^{-12} and 3.21×10^{-42}, respectively.

3. Phenomenological mass relation for free massive stable particles

Among the hundreds observed particles, only several free particles are notable, which are stable or at least their lifetime is longer than the age of the Universe. These particles are the proton (p), electron (e), photon (γ), neutrinos (ν_i) and hypothetical graviton (g). These particles play a key role since together with the quasi-stable neutron they build the known matter in the Universe. Although according to the quantum chromodynamics the proton is composed by quarks, it has displayed itself as an undivided particle in the recent experiments. Only *free massive stable* particles are examined in this paper. The quarks and gluons are bound in the hadrons by confinement and they cannot be immediately detected in the experiments, and the photon is massless. Therefore, these particles are not the subject of this paper.

The ratio between the proton and electron masses is $\frac{m_p}{m_e}\approx 1836$. On the other side, the ratio between the strong and electromagnetic coupling constants at *ELE* is $\frac{\alpha_s(0)}{\alpha}\approx 1918$. The two ratios differ by less than 5%, which is less than the error of $\alpha_s(0)$. Consequentially, the found relation is hardly an incidental coincidence, therefore would be written:

$$\frac{m_p}{m_e}\approx\frac{\alpha_s(0)}{\alpha} \tag{4}$$

Thus, it is shown that the ratio between the proton and electron masses is close to the ratio between the strong and electromagnetic coupling constants at ELE. The proton and electron are the Lightest Free Massive Stable Particles (*LFMSP*), acted upon by the strong and electromagnetic interactions, respectively. This relation is important since it connects the masses of *LFMSP*, acted upon by the strong and electromagnetic interactions and the respective coupling constant at *ELE*. The relation (4) suggests that the mass of *LFMSP*, acted upon by the strong (or the electromagnetic) interaction is proportional to the respective coupling constant at *ELE*, i.e. $m_p\approx k\alpha_s(0)$ and $m_e\approx k\alpha$, where k is a constant. Therefore, it is

interesting to examine whether this rule is valid both for the weak interaction, which is several orders of magnitude weaker than the electromagnetic interaction and for the gravity, which is dozens orders of magnitude weaker than the weak interaction. *LFMSP* acted upon by the weak interaction is the electron neutrino and *LFMSP* acted upon by the gravity most probably appears the hypothetical graviton. Although the rest masses of the two particles are yet unknown, the direct neutrino mass experiments and the theoretical models suggest the ν_e mass between $10^{-6}\ eV/c^2$ and $2\ eV/c^2$, i.e. is several orders of magnitude lighter than the electron. Again, the astrophysical constraints allow to find the upper limits of the graviton mass and according to these constraints, if the graviton really exists its mass would be less than $3.1 \times 10^{-29}\ eV/c^2$, i.e. dozens orders of magnitude lighter than ν_e. Table I presents the obtained in Section 2 coupling constants of the interactions at *ELE*, as well as the masses of *LFMSP*, acted upon by the respective interactions. The experimental upper limits of the electron neutrino and graviton masses are presented, too.

Table I shows that with increasing the interaction strength (coupling constant), the mass m_{mon} of *LFMSP* acted upon by the respective interaction also increases. The data from Table I are presented in a double-logarithmic scale in Fig.1, which shows that the trend is clearly expressed.

Table I: Coupling constants of interactions at ELE and the masses of LFMSP acted upon by the respective interaction.

Fundamental Interaction	Coupling Constant	Particle	Experimental Mass (eV/c^2)	Calculated Mass (eV/c^2)
Strong	14	p	9.38×10^8	9.8×10^8
Electromagnetic	7.30×10^{-3}	e	5.11×10^5	-
Weak	3.00×10^{-12}	ν_e	$0 < m < 2$	2.1×10^{-4}
Gravitational	3.21×10^{-42}	g	$< 3.1 \times 10^{-29}$	2.3×10^{-34}

The points in Fig. 1, corresponding to the electron and proton masses and to the upper limit masses of the electron neutrino and graviton, are approximated by the least squares with a power law:

(5) $Logm_{min} = 0.880 Log\alpha(0) + 8.493$

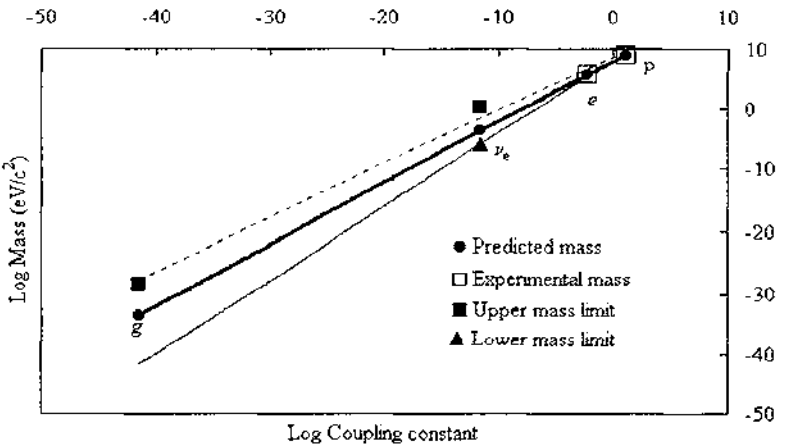


Fig. 1. Dependence of the mass of LFMSP acted upon by a particular interaction on the coupling constant of the respective interaction at ELE. The dashed line presents the approximation (5) of e, p and the upper mass limits of g and ν_e. The thin solid line presents the approximation (6) of e, p and the lowest limit mass of ν_e taken from the models. The thick solid line presents the strict linear approximation (S=1).

Although this approximation is only on 4 points, the correlation found is close and the correlation coefficient reaches $r = 0.997$, which supports the power law. The slope (S) is a little smaller than a unit, but it should be reminded that instead of the electron neutrino and graviton masses, their upper limit values are used, which produce a certain underestimation of the S value. Therefore, it can be said that the regression is close to a linear one. Similarly, the points, corresponding to the electron and proton masses and to the low limit of the electron neutrino mass from the theoretical models are approximated with a power law, too:

(6) $Logm_{min} = 1.196 Log\alpha(0) + 7.883$

This approximation is similar to (5) and the correlation is close again (r = 0.999). Now the slope is a little bigger than a unit. The certain slope overestimation is due to the fact that in this case instead of the still unknown value of the electron neutrino mass, the lowest neutrino mass limit is used from the models. Approximations (5) and (6) produce at $\alpha_g(0) = 3.21 \times 10^{-42}$ a value for the graviton mass in an

interval from 1.8×10^{-42} eV/c^2 to 3.1×10^{-29} eV/c^2, i.e. it suggests that the graviton possesses a negligible, yet a *nonzero mass*. These approximations show that the mass of *LFMSP*, acted upon by a particular interaction, increases with the increase of the respective coupling constant $\alpha(0)$ by a power law with $S \sim 1$, i.e. close to the linear one. The coincidence (4) of the ratio between the proton and electron masses with the ratio between the strong and electromagnetic coupling constants at *ELE* also supports a linear dependence (without intercept) between m_{min} and $\alpha(0)$. Hence, the experimental data suggest a linear dependence ($S = 1$) between the mass of *LFMSP* and the coupling constants:

(7) $\quad Log m_{min} = Log\alpha(0) + k_0$

where k_0 is a constant.

Expression (7) transforms into $m_{min} = 10^{k_0}\alpha(0) = k\alpha(0)$. In this way, the experimental data and constraints suggest that the mass of *LFMSP*, acted upon by a particular interaction, is proportional to the coupling constant of the respective interaction at *ELE*:

(8) $\quad m_{i\,min} = k\alpha_i(0)$

where k is a constant, i =1, 2, 3, 4 - index for each interaction and *LFMSP* acted upon by a respective interaction.

Constant k can be determined by the fine structure constant ($\alpha_1 \equiv \alpha$) and the electron mass ($m_{1min} \equiv m_e$) since both are measured with very high precision $k = \frac{m_e}{\alpha} \approx 7.00 \times 10^7 eV/c^2$. The substitution of this value in (8) yields the mass relation:

(9) $\quad m_{i\,min} = \frac{m_e}{\alpha}\alpha_i(0)$

Therefore, *LFMSP* corresponds to each fundamental interaction and its mass is proportional to the strength of the respective interaction.

4. Neutrino and graviton mass estimations and discussions

The found mass relation (9) can be examined by the strong interaction since the proton mass is measured with high precision. The application of the mass relation on the strong interaction yields the lightest stable hadron mass $m_p \approx 9.80 \times 10^8$ eV/c^2. Therefore, the proton mass value obtained by the mass relation would be

only 4.5% higher than the experimental value of m_p. This result shows that the found mass relation possesses heuristic power. The application of mass relation (9) on the weak interaction allows to evaluate the mass of the electron neutrino $m_{\nu e} \approx 2.1 \times 10^{-4}\, eV/c^2$. This value is close to the prediction of the simple *SO(10)* model for the lightest neutrino mass $m_{\nu e} = 2.4 \times 10^{-4}\, eV/c^2$ [30].

Finally, the application of mass relation (9) on the gravitational interaction produces an estimation of the graviton mass $m_g = \frac{m_e}{\alpha}\alpha_g = \frac{G_N m_e^2 m_p}{e^2} \approx 2.3 \times 10^{-34}\, eV/c^2$, which is several orders of magnitude smaller than the upper limit of the graviton mass, obtained by astrophysical constraints. Thus, the graviton mass has been estimated by *fundamental constants* only - universal gravitational constant, electron and proton masses and elementary electrical charge. It is worth noting that this value is of the order of the magnitude of $\frac{\hbar H}{c^2}$, but by the suggested approach it has been obtained without consideration of any cosmological models.

The calculated masses of the four free stable particles are presented in the last column of Table I. It shows that the fitting of the calculated values and the experimental data is satisfactory. In this way, the found mass relation is remarkable since it unifies the masses of four stable particles of completely different kinds (p, e, ν_e and g) and covers an extremely wide range of values, exceeding 40 orders of magnitude. The found mass relation allows approximate estimation of the neutrino and graviton masses, which affords the opportunity for its verification in the nearest years when the direct neutrino mass measurements are expected to reach the necessary threshold of sensitivity.

The obtained value $m_{\nu e} \approx 2.1 \times 10^{-4}\, eV/c^2$ and the results from the solar and atmospheric neutrino experiments allow to estimate the masses of the heavier neutrino flavours - ν_μ and ν_τ. The results from the Super Kamiokande experiment lead to the neutrino mass squared difference $\Delta m_{23}^2 \sim 2.7 \times 10^{-3}\, eV^2$ [31]. Recent results on solar neutrinos provide hints that the Large Mixing Angle (*LMA*) Mikheyev-Smirnov-Wolfenstein (*MSW*) solution is more probable than the Small Mixing Angle (*SMA*) *MSW* [32]. The *LMA* leads to $\Delta m_{12}^2 \sim 7 \times 10^{-5}\, eV^2$ [33] and the *SMA* leads to $\Delta m_{12}^2 \sim 6 \times 10^{-6}\, eV^2$ [34]. In this way, both solutions yield $m_{\nu\tau} \approx 0.05\, eV/c^2$. The most appropriate *LMA* yields $m_{\nu\mu} \approx 8.4 \times 10^{-3}\, eV/c^2$, and *SMA* yields $m_{\nu\mu} \approx 2.5 \times 10^{-3}\, eV/c^2$. Thus, the obtained values of the neutrino masses support the normal

hierarchy of neutrino masses.

Now it is not clear yet what the cause for the relationship between the masses of *LFMSP* and the interaction coupling constants is but its existence is confirmed by the recent experimental data and constraints. Most probably the found mass relation represents an expression of a universal symmetry, including free stable particles of most diverse kinds (hadron, charged lepton, neutral lepton and carrier of gravity).

It is interesting that mass relation (9) is very similar to equation (10) derived in [35] by a totally different approach, namely by strong gravity model and astrophysical constraints:

$$(10) \qquad m = \left(\frac{\hbar^2}{G_N R_0} \right)^{1/3} \frac{q^2}{\hbar c}$$

where $R_0 = c/H$ is the Hubble distance and q is the "main charge".

The mass dimension coefficient of proportionality $m_0 = \left(\frac{\hbar^2}{G_N R_0} \right)^{1/3} \sim$

$1.05 \times 10^{-28}\ kg$ is close to $k = \frac{m_e}{\alpha} \approx 1.25 \times 10^{-28}\ kg$.

The presence of an exceptionally small, yet nonzero mass of the graviton, involves a finite range of the gravity $r_g \sim \lambda\!\!\!^{-}_g$ and Yukawa potential of the gravitational field $\phi(r) = -\frac{G_N M}{r} \exp(-r / \lambda\!\!\!^{-}_g)$, where $\lambda\!\!\!^{-}_g$ is Compton wavelength of the graviton $\lambda\!\!\!^{-}_g = \frac{\hbar}{m_g c} \sim c/H$. Since is in order of magnitude of the Hubble distance, the deviation of Newton potential from Yukawa potential is manifested very weakly at a distance $r << c/H$. As a result, the experimental determination of the graviton mass will be a serious challenge. Yet, it can be expected that appropriate astrophysical or laboratory experiments will be found for experimental determination of the graviton mass. The massive graviton might turn of considerable importance for the description of the processes in the nuclei of the active galaxies and quasars, the gravitational collapse as well as for the improvement of the cosmological models.

The massive graviton places other challenges before the modern unified theories. Among them are the famous van Dam-Veltman-Zakharov (*vDVZ*)

discontinuity [36, 37] and the violation of the gauge invariance and the general covariance. Yet, there are already encouraging attempts to solve *vDVZ* discontinuity in anti de Sitter background [38].

5. Conclusions

Among the huge multitude of observed particles, a small group of several free particles is noticeable, which are stables or, at least their lifetimes are longer than the age of the Universe. These particles are very important since together with the quasi-stable neutron they build the known matter in the Universe.

The interaction coupling constants are determined at *ELE* close to $m_e c^2$ when the interactions are most different in their strength and are differentiated clearly from one another. It is shown that the ratio between the proton and electron masses is close to the ratio between the strong and electromagnetic coupling constants at *ELE*. Further on, based on the experimental data, a mass relation has been found, according to which the rest mass of the Lightest Free Massive Stable Particle (LFMSP), acted upon by a particular interaction, is proportional to the coupling constant of the respective interaction at ELE. This mass relation is remarkable, since it connects the masses of the particles, most important for the stability of the Universe, i.e. the stable particles, with the most substantial property of the fundamental interactions, i.e. their coupling constants. The applied approach suggests that the graviton mass is nonzero. The electron neutrino and graviton masses are approximately estimated to $2.1 \times 10^{-4}\, eV/c^2$ and $2.3 \times 10^{-34}\, eV/c^2$, respectively. The obtained value of the graviton mass is of the order of the magnitude of $\frac{\hbar H}{c^2}$ and this value has been found by fundamental constants only, without consideration of any cosmological models. The masses of the heavier neutrinos ν_μ and ν_τ are estimated by the results of the solar and atmospheric neutrino experiments. The mass of ν_τ is close to $0.05\, eV/c^2$ while the estimation of m_{ν_μ} depends to a certain extent on the *MSW* solution. Yet, both solutions support the normal hierarchy of the neutrino masses.

Acknowledgements

I would like to thank Prof. I. G. Koprinkov for his useful discussions.

References

1. Reines F., C. L. Cowan. Detection of the free neutrino, Phys. Rev., 92, 1953, 830-831.
2. Fermi E. Versuch einer theorie der beta-strahlen. I., Zeit. Phys., 88, 1934, 161-177.
3. Weinberg S. A model of leptons, Phys. Rev. Lett., 19, 1967, 1264-1266.
4. Salam A. Elementary particle theory: Relativistic groups and analyticity, In Nobel symposium 8, Stockholm: Almqvist and Wiksell, 1968, 367-377.
5. Pauli W. Les theories quantique du magnetisme. L'electron magnetique. In Proc. 6th Solvay congr. on phys., Bruxelles, 1930, 175-176.
6. Davis R. et al. Search for neutrinos from the Sun, Phys. Rev. Lett., 20, 1968, 1205-1209.
7. Hirata K. S. et al. Experimental study of the atmospheric neutrino flux, Phys. Lett. B, 205, 1988, 416-420.
8. Casper D. et al. Measurement of atmospheric neutrino composition with the IMB-3 detector, Phys. Rev. Lett., 66, 1991, 2561-2564.
9. Fukuda Y. et al. Evidence for oscillation of atmospheric neutrinos, Phys. Rev. Lett., 81, 1998, 1562-1567.
10. Weinheimer C. et al. High precision measurement of the tritium beta spectrum near its endpoint and upper limit on the neutrino mass, Phys. Lett. B, 460, 1999, 219-226.
11. Lobashev V. M. et al. Direct search for mass of neutrino and anomaly in the tritium beta-spectrum, Phys. Lett. B, 460, 1999, 227-235.
12. Assamagan K. et al. Upper limit of the muon-neutrino mass and charged-pion mass from momentum analysis of a surface muon beam, Phys. Rev. D, 53, 1996, 6065-6077.
13. Barate R. et al. An upper limit on the tau neutrino mass from three- and five-prong tau decays, Eur. Phys. J. C, 2, 1998, 395-406.
14. Bahcall J. N. Solar neutrinos: Where we are, where we are going. Astrophys. J., 467, 1996, 475-484.
15. Talmadge C. et al. Model-independent constraints on possible modifications of Newtonian gravity, Phys. Rev. Let., 61, 1988, 1159-1162.
16. Goldhaber A. S., M. M. Nietto. Mass of the graviton, Phys. Rev. D, 9, 1974, 1119-1121.
17. Finn L.S., P.J. Sutton. Bounding the mass of the graviton using binary pulsar observations, Phys. Rev. D, 65, 2002, 044022-044028.
18. Will C.M., N. Yunes. Testing alternative theories of gravity using LISA, Class. Quant. Grav., 21, 2004, 4367-4381.
19. Freund P.G. et al. Finite-range gravitation, Asrophys. J., 157, 1969, 857-867.
20. Gershtein S.S. et al. Evolution of the Universe and the graviton mass, Phys. Atom. Nucl., 61, 1998, 1420-1429.
21. Woodward J.F. et al. Mach's principle and the rest mass of the graviton, Phys. Rev. D, 11, 1975, 1371-1374.
22. Bethke S. et al. Experimental investigation of the energy dependence of the strong coupling strength, Phys. Lett. B, 213, 1988, 235-241.

23. Hinchliffe I. Quantum chromodynamics, Eur. Phys. J. C, 15, 2000, 85-94.
24. Acciarri M. et al., Study of the structure of hadronic events and determination of α_s at $\sqrt{S}$ =130 GeV and 136 GeV, Phys. Lett. B, 371, 1996, 137-148.
25. Mohr P.J., B.N. Taylor. Physical constants, Eur. Phys. J. C, 15, 2000, 73-74.
26. Dictionary of science, Oxford, Oxford University Press,1999.
27. Davidov A. Atoms, nuclei and particles, Kiev, Naukova Dumka Press, 1971.
28. Gibson W., B. Pollard. Symmetry principals in elementary particle physics, Cambridge, Cambridge University Press, 1976.
29. Prokhorov A. Physical encyclopedical dictionary, Moscow, Sovyetskaya Encyclopediya Press, 1984.
30. Dermisek R. Neutrino masses and mixing, quark-lepton symmetry, and strong right-handed neutrino hierarchy, Phys. Rev. D, 70, 2004, 073016-073022.
31. Gonzalez-Garcia M.C., C. Pena-Garay. Three-neutrino mixing after the first results from K2K and KamLAND, Phys. Rev. D, 68, 2003, 093003-093012.
32. Krastev P.I., A.Y. Smirnov. Global analysis with SNO: Toward the solution of the solar neutrino problem, Phys. Rev. D, 65, 2002, 073022-073036.
33. Maltoni M. et al. Status of three-neutrino oscillations after the SNO-salt data, Phys. Rev. D, 68, 2003, 113010-113017.
34. Albright C.H. Overview of grand unified models and their predictions for neutrino oscillations, Nucl. Instr. Meth. Phys. Res. A, 472, 2001, 359-363.
35. Massa C. Is there a fundamental acceleration, Nuovo Cimento, 109, 1994, 25-33.
36. van Dam H., M.J. Veltman. Massive and mass-less Yang-Mills and gravitational fields, Nucl. Phys. B, 22, 1970, 397-411.
37. Zakharov V.I. Linearized gravitation theory and the graviton mass, JETP Lett., 12, 1970, 312-320.
38. Kogan I.I. et al. The m→0 limit for massive graviton in dS_4 and AdS_4 — How to circumvent the van Dam–Veltman–Zakharov discontinuity, Phys. Lett. B, 503, 2001, 173-180.

ОЦЕНКИ НА МАСИТЕ НА НЕУТРИНОТО И ГРАВИТОНА ПОСРЕДСТВОМ ЕДИН ФЕНОМЕНОЛОГИЧЕН ПОДХОД

Д. Вълев

Резюме

Показано е, че отношението между масите на протона и електрона е близко до отношението между константите на свързване за силното и електромагнитното взаимодействие при екстремално ниска енергия (*EHE*). На базата на експерименталните данни тази връзка бе разширена и за слабото и гравитационното взаимодействие. По такъв начин бе намерена масова връзка, съгласно която масата на покой на най-леката свободна масивна стабилна частица, участвуваща в дадено взаимодействие, е пропорционална на константата на свързване при *EHE* за съответното взаимодействие. На базата на тази масова връзка масите на електронното неутрино и гравитона бяха оценени съответно на $2.1 \times 10^{-4}\ eV/c^2$ и $2.3 \times 10^{-34}\ eV/c^2$. Последната стойност е от порядъка на $\frac{\hbar H}{c^2}$, където H е константата на Хъбъл. Заслужава си да се отбележи, че тази стойност е получена само посредством фундаментални константи, без разглеждане на никакъв космологичен модел.

Bulgarian Academy of Sciences. Space Research Institute.
Aerospace Research in Bulgaria. 22, 2008, Sofia

TEXTURE EDGE DETECTION TECHNIQUES USING TEXTURE FEATURES AND GRADIENT-BASED EDGE DETECTORS

Mariana Tsaneva

Solar-Terrestrial Influences Laboratory – Bulgarian Academy of Sciences
e-mail: mtsaneva@hotmail.com

Abstract

Texture edges are the locations where there is an abrupt change in image texture properties. Two novel texture edge detection techniques, which used texture feature extraction and segmentation before conventional gradient-based or zero crossing edge detection algorithms (Roberts, Sobel, Prewitt, Laplacian, Kirsch, Robinson and Canny) are proposed and tested in this paper. The most common image segmentation methods are used in experiments for texture segmentation: fuzzy c-means, gray level quantization, histogram thresholding, median cut and principal components transformation/median cut. Results from applying these texture edge detection techniques to satellite images are presented and analyzed.

1. Introduction

Spatial, spectral and texture properties have been used for many years in many remote sensing applications such as mapping and analysis of ground cover. Texture analysis methods and techniques grew in popularity during the late 1990s as the resolution of satellite images increased. Texture analysis has been studied in computer vision literature. Several definitions of texture have been formulated by researchers and one early definition is due to Haralick: "The image texture we consider is non figurative and cellular... An image texture is described by the number and types of its (tonal) primitives and the spatial organization or layout of its (tonal) primitives... A fundamental characteristic of texture: It cannot be analyzed without a frame of reference of tonal primitive being stated or implied. For any smooth gray-tone surface, there exists a scale such that when the surface is examined, it has no

texture. Then as resolution increases, it takes on a fine texture and then a coarse texture" [1].

The visual interpretation of images has relied on image spatial properties. Spectral features are also efficiently used for discrimination and identification of the objects in interpreting images besides spatial features. Remote sensing in the visible and near infrared ranges of the electromagnetic spectrum finds wide application as a method for discrimination of natural objects, their states and texture peculiarities, as well as for monitoring of ecosystems. In [2], results are presented for recognition of natural objects along a trace on the ground surface of the territory of Bulgaria using data for the solar radiation reflected from these objects. Data were collected by means of a trace multichannel spectrometer onboard the MIR spatial station.

The role of image texture and the incorporation of texture features in image classification procedures is very important with the ever-increasing spatial resolution of remotely sensed data. Augustejin [3] uses a neural network classifier for ground cover identification in satellite images based on texture measures and compares a range of texture features (co-occurence matrices, gray-level differences, texture-tone analysis, features derived from Fourier transform, and Gabor filters). Raw pixel based classification is used to provide a baseline. The results show that the best feature set depends on the data to be classified.

Unser [4] developed local linear transforms for texture measurement. He proposed using four 2 x 2 Hadamard masks: the first measures the magnitude, the other three masks approximate the derivatives in horizontal, vertical and diagonal directions. After convolution with masks, the computation of the local variance using a moving window in the image is applied.

In [5], the authors introduce a novel texture description scheme. Texture representation uses a combination of edge and region statistics. The results show that edge-based representation is the best.

Feature-based texture analysis derives texture measures directly from the image by local operators, statistical attributes and examination of images in the frequency domain. The main focus of remote sensing scientists has involved the use of second-order statistics from the gray-level co-occurrence matrix. Many different methods are used to extract textural information from images: structural, statistical, model-based and transform-based. Recent research has focused on texture measures derived from moving a fixed-size, odd-numbered window through the image and calculating a variety of different pixel relationships [6]. Relatively few studies have focused on the use of lower-order statistical properties of images [7].

An edge in a nontextured image is a sequence of pixel location where there is an abrupt change in gray level (intensity) values on either side of the pixel locations. A texture edge is associated with changes in the texture characteristics. Conventional gradient–based edge detection algorithms can detect only local changes in intensity values and they are suitable only for nontextured images. When applied to textured images, gradient–based edge detection methods detect intensity edges, which are micro-edges [8].

Edge detection techniques often use a mask (2 x 2, 3 x 3.pixels) that is convolved with the pixel in the window (Fig. 1). An edge detector can be [9]:

1. Directional which can be used to detect orientations at 90° intervals (the Roberts, Sobel and Prewitt operators).
2. Directional which can be used to detect orientation at 45° intervals (the Kirsch and Robinson operators).
3. Non-directional (the Laplacian).

The canny edge detection algorithm [10] is based on the zero crossing of the image function second derivative. It is known to many as the optimal edge detector.

Kazuhiro [11] proposed a method which can be applied to colour and texture edge extraction. In this method, an edge is defined not as a point where there is a large change in intensity, but as a region boundary based on the separability. which depends only on the shape of an edge. This characteristic enables easy selection of the optimum threshold value for the extraction of an edge.

Texture edge detection techniques can be used for textured object classification and object recognition. They are able to work with boundary and contour detection techniques on stages of object boundary detection or region of interest detection.

The goal of this research is to detect the macro edge separating the two or more texture regions on satellite images. The objective of this work is to develop techniques to extract texture edges in and between texture regions.

2. Texture edge detection

Edge detection is a research field within image processing and computer vision, in particular within the area of feature extraction. Edge detection consists of creating a binary image from gray scale image by finding a high intensity gradient or the second derivative zero crossings. The edges identified by edge detection are often disconnected. Discontinuities are bridged, if the distance between the two edges is within some predetermined threshold. Region boundaries and edges are closely related, since there is often sharp adjustment in intensity at the region

boundaries.

Texture edges can be micro-edges and macro-edges. Micro-edges can be detected using small-distance operators. Macro-edges need large-size edge detectors. There are many texture edge detection techniques using different texture features and different algorithms including edge detectors, image enhancement, edge models, etc. In this research, conventional edge detection techniques are used after novel proposed preprocessing scheme for texture feature extraction. Micro-edges are obtained using conventional edge detectors: the Roberts, Sobel, Prewitt, Laplacian, Kirsch, Robinson and Canny edge detection algorithms. Macro-edges are obtained using texture segmentation before edge detection. The most common segmentation algorithms are used: fuzzy c-means, gray level quantization, histogram thresholding, median cut and principal components transformation/median cut.

3. Edge detectors

Edges are used to find region boundaries. Boundaries and their parts (edges) are perpendicular to the direction of the gradient. Gradient operators can be divided into:

1. Operators approximating the first or second derivatives of the image function using differences. Operators which are rotationally invariant (Laplacian) need one convolution mask only (Fig. 1d). Operators which are able to detect edge direction (Roberts, Sobel, Prewitt, Kirsch and Robinson) are represented by a collection of masks, each corresponding to a certain direction (Fig. 1a, b, c, e, f).
2. Operators based on the zero crossings of the image function second derivative (Canny edge detector).

The Roberts operator approximates the gradient of the image function. The Sobel, Prewitt, Kirsch and Robinson operators approximate the first derivative. The Laplace operator approximates the second derivative [12].

The Canny edge detection algorithm is known as the optimal edge detector based on three criteria:

1. Detection criterion – important edges should not be missed, there should be no spurious responses.
2. Localization criterion – distance between the actual and located position of the edge should be minimal.
3. One response criterion – minimizes multiple responses to a single edge (partly covered by the first criterion).

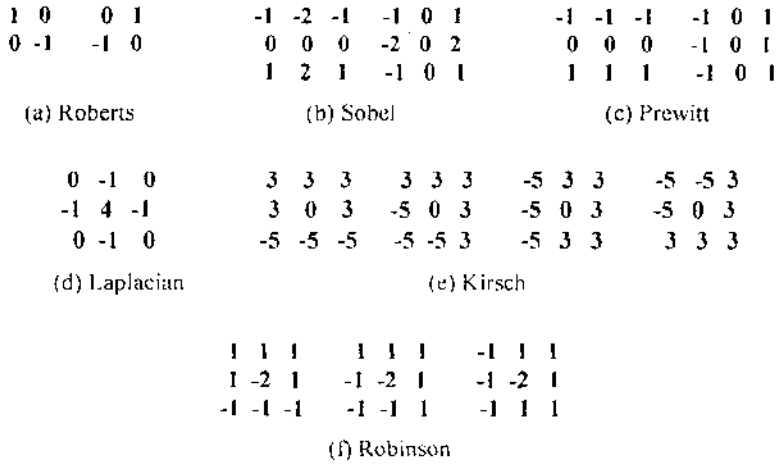


Fig. 1. Edge detectors masks

The Canny edge detector works as follows:

1. Smoothes the image to eliminate noise.
2. Finds the image gradient to highlight regions with high spatial derivatives.
3. Suppresses any pixel that is not at the maximum (nonmaximal suppression).
4. Reduces the gradient array by hysteresis. It uses two thresholds. If the magnitude is below the first threshold, it is set to zero (made a nonedge). If the magnitude is above the high threshold, it is made an edge. If the magnitude is between the two thresholds, then it is set to zero.

Each edge detection operator has its advantages and disadvantages (Table1). All gradient-based edge detection operators (Roberts, Sobel, Prewitt, Kirsch and Robinson) are able to detect edge direction, but they depend on the size of objects and on noise. The Roberts, Laplace and Canny edge detectors have their own advantages and disadvantages.

4. Texture features

Every pixel is described by texture feature, which must be homogenous inside a texture region. In this research, the two novel scalar texture features are used: texture feature 1 and texture feature 2 from [13]. For every pixel in the image, a 5 x 5 pixel window is considered, where P(x,y) is the gray level value of pixel (x,y). For every 5 x 5 pixel mask, the average intensity value M of pixels is calculated.

Table 1. Advantages and disadvantages of edge detection operators

Operator	Advantages	Disadvantages
The Roberts operator	The simplest and fastest operator..	It has high sensitivity to noise.
The Sobel operator	It is able to determine gradient direction.	It depends on the size of objects and has sensitivity to noise.
The Prewitt operator	It is able to determine gradient direction.	It depends on the size of objects and has sensitivity to noise.
The Laplace operator	It gives gradient magnitude only and is rotationally invariant.	It responds twice to some edges in the image.
The Kirsch operator	It is able to determine gradient direction.	It depends on the size of objects and has sensitivity to noise.
The Robinson operator	It is able to determine gradient direction.	It depends on the size of objects and has sensitivity to noise.
The Canny edge detector	The optimal edge detector based on three criteria: detection criterion, localization criterion and one response criterion. Optimal for step edges corrupted by white noise.	Smoothes the shape too much and tends to create closed loops of edges.

Texture features (TF) are defined as a difference of intensity value of the central pixel and the average intensity value in the feature masks:

$$M_1 = (P(x+2,y-2)+P(x+2,y)+P(x+2,y+2)+P(x+1,y-1)+ P(x+1,y+1)+P(x,y-2)+P(x,y)+P(x,y+2)+P(x-1,y-1)+ P(x-1,y+1)+P(x-2,y-2)+P(x-2,y)+P(x-2,y+2))/13 \quad (1)$$

$$TF_1 = P(x,y)-M_1 \quad (2)$$

$$M_2 = (P(x+2,y-1)+P(x+2,y+1)+P(x+1,y-2)+P(x+1,y)+ P(x+1,y+2)+P(x,y-1)+P(x,y+1)+P(x-1,y-2)+P(x-1,y)+ P(x-1,,y+2)+P(x-2,y-1)+P(x-2,y+1))/12 \quad (3)$$

$$TF_2 = P(x,y)-M_2 \quad (4)$$

These texture features are implemented in image processing and analysis software ImageJ [14] as macros. These macros display two resultant texture feature images (Figs. 3 and 4).

5. Texture edge detection techniques

Two novel edge detection techniques are proposed. The first technique is based on texture feature edge detection. The second is based on texture feature image segmentation and edge detection. Texture segmentation has been done by five image segmentation algorithms in CVIPtools [15]: fuzzy c-means, gray level quantization, histogram thresholding, median cut and principal components transformation/median cut.

The first technique works as follow:

1. Texture feature extraction by texture feature 1 and 2 [13, 14].
2. Applying conventional edge detection operators on resultant texture feature images [15].

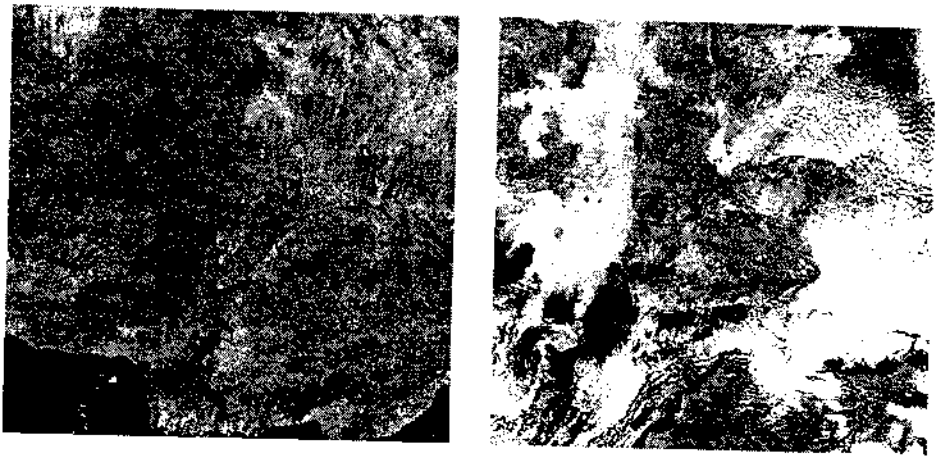
The second technique works as follow:

1. Texture feature extraction by texture feature 1 and 2 [13, 14].
2. Resultant image segmentation by using the image segmentation algorithms in CVIPtools [15].
3. Applying conventional edge detection operators on resultant texture segmentation images [15].

6. Experiments

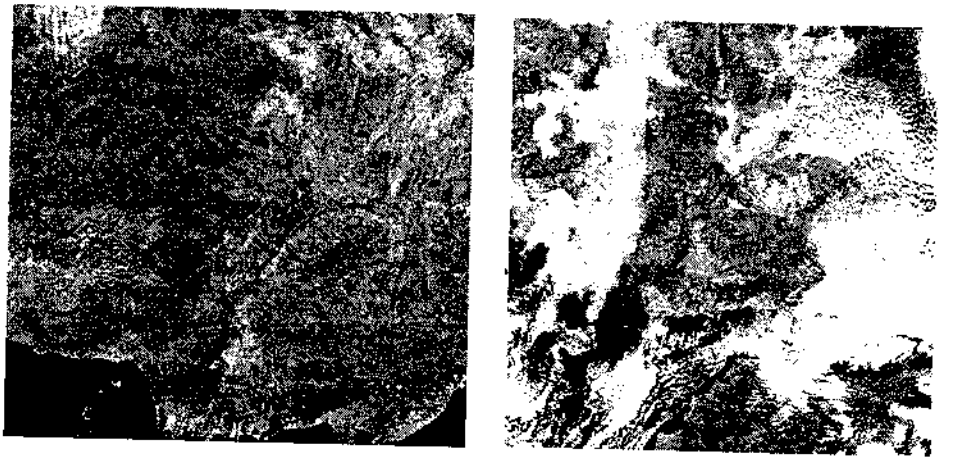
The MODIS flying on NASA's Terra satellite captured the image (Fig. 2a) of floods in the border region of Western Turkey, Eastern Greece, and SouthEastern Bulgaria on November 22. The second image (Fig. 2b) shows the region on November 9, under normal conditions. Both images were made with a combination of visible and infrared light [16]. The Evros River, called the Meriç River in Turkey, flows from Bulgaria south along the border between Turkey and Greece and into the Aegean Sea. Flowing into the Evros from the east, the Ergene River is similarly flooded, as are a number of other tributaries, including the Ardas.

These satellite images are used to obtain texture feature images (Figs. 3 and 4). The implemented experiments apply the proposed texture edge detection techniques. The obtained results are presented and analyzed.



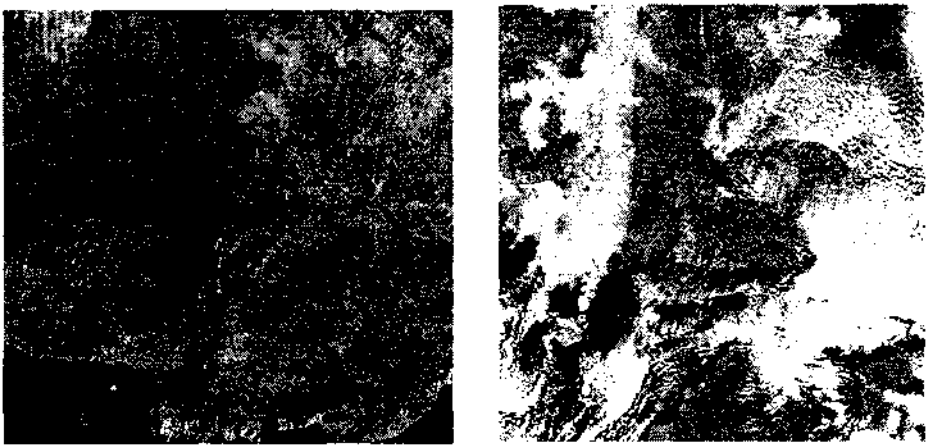
(a) (b)

Fig. 2. The MODIS satellite images obtained on 9.11.2007 (a) and 22.11.2007 (b)



(a) (b)

Fig. 3. The texture feature 1 images: 9.11.2007 (a) and 22.11.2007 (b)



(a) (b)

Fig. 4. The texture feature 2 images: 9.11.2007 (a) and 22.11.2007 (b)

7. Results and discussion

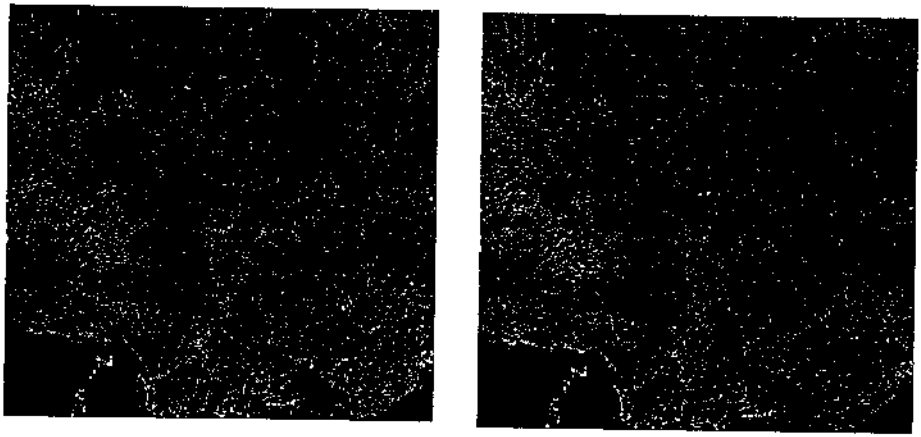
To evaluate the techniques, it is necessary to have some comparison basis for. In this research, the basis is another technique that has been widely accepted by the scientists: the Gray-Tone Difference Matrix (GTDM) and the "flat texture image" (the image minus the median filtered image).

The GTDM has been proposed in an attempt to define texture measures correlated with human perception of textures. In this work, five different features are used to quantitatively describe such perceptual texture properties as coarseness, contrast, busyness, complexity, and texture strength [17].

The "flat texture images" (Fig. 5) represent a texture because the operation the image minus the median filtered image removes the effect of the overall intensity level.

The edge detection operators are applied on the resultant images. The results show that proposed texture edge detection techniques have been successfully applied on satellite images with obtained rate commensurable to edge detection on these images. In the statistical approach, there is no reliable distinction between the results based on the comparison technique and the proposed techniques.

Figures 6, 7 and 8 show the complied results for the proposed technique.

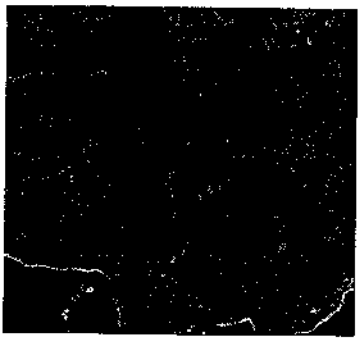


(a) (b)

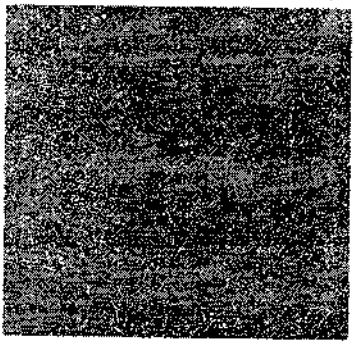
Fig. 5. The "flat texture images": 9.11.2007 (a) and 22.11.2007 (b)

(a) *Roberts*

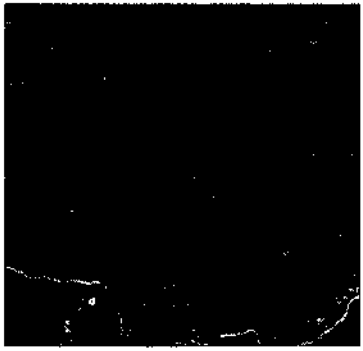
(b) *Sobel*



(c) *Prewitt*

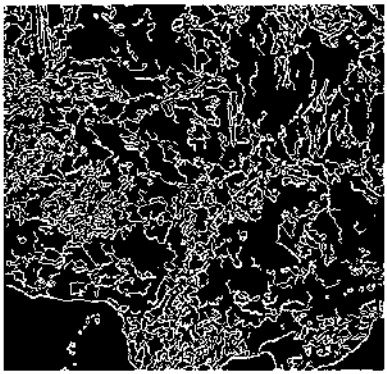


(d) *Laplacian*



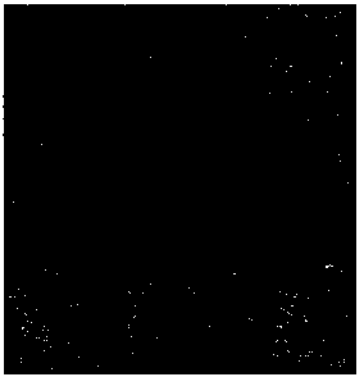
(e) *Kirsch*

(f) *Robinson*



(g) *Canny*

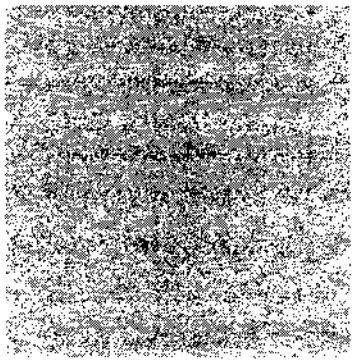
Fig. 6. The results from applying the first proposed texture edge detection technique to the satellite image obtained on 9.11.2007



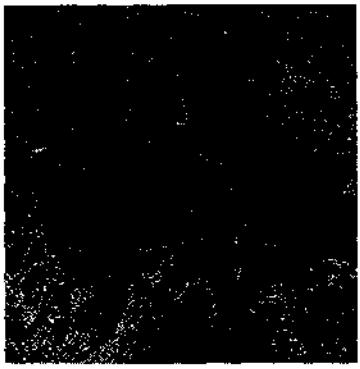
(a) *Roberts*

(b) *Sobel*

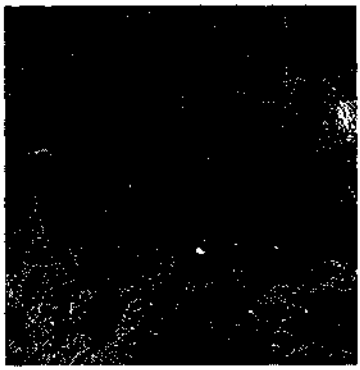
(c) *Prewitt*



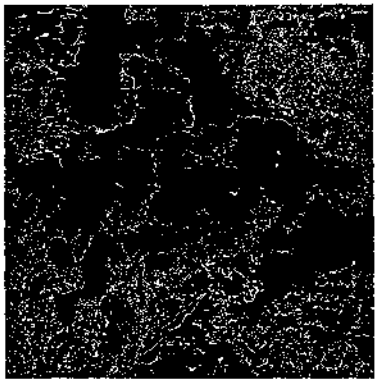
(d) *Laplacian*



(e) *Kirsch*

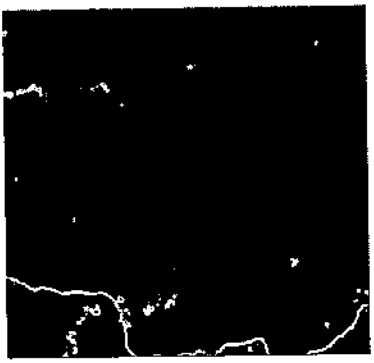
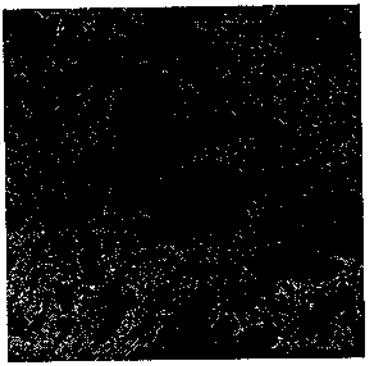


(f) *Robinson*

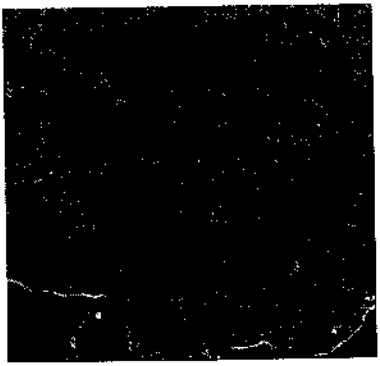
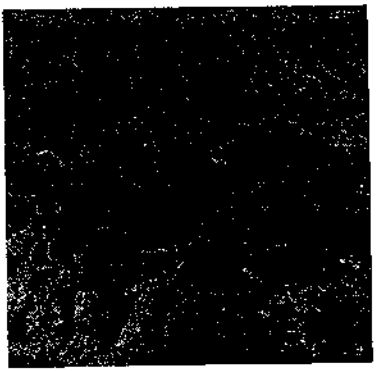


(g) *Canny*

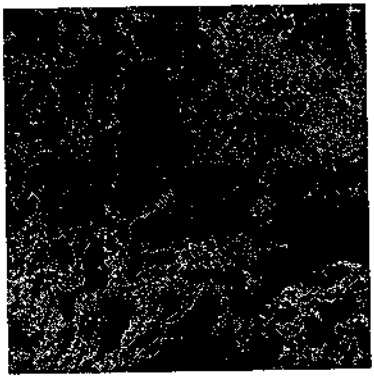
Fig. 7. The results from applying the first proposed texture edge detection technique to the satellite image obtained on 22.11.2007

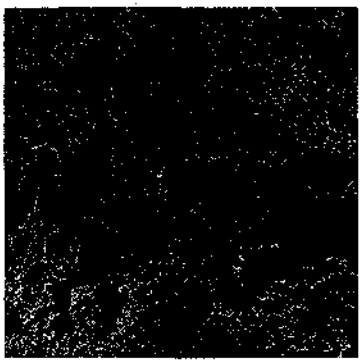
The fuzzy c-means segmentation and the Kirsch edge detection

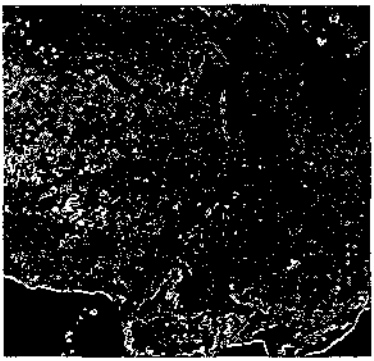
The gray level quantization segmentation and the Kirsch edge detection

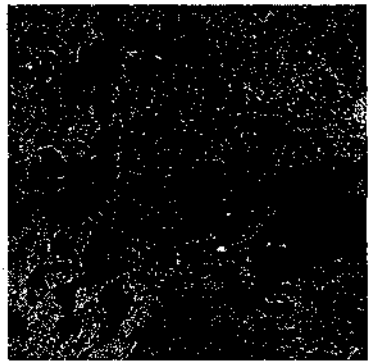



The histogram thresholding segmentation and the Kirsch edge detection



The median cut segmentation and the Kirsch edge detection





The principal component transformation/median cut segmentation and the Kirsch edge detection

(a) (b)

Fig. 8. The results from applying the second proposed texture edge detection technique on the stage of the Kirsch operator to the satellite images obtained on 9.11.2007 (a) and 22.11.2007 (b)

8. Conclusion

Two novel texture edge detection techniques are developed and tested. They have been applied to MODIS satellite images of flooded regions of Turkey, Greece and Bulgaria on 9.11.2007 (normal conditions) and 22.11.2007 (floods). The results evidence of the successful application of the proposed techniques for texture edge detection.

The feature work is addressed to identifying texture boundaries, which is crucial for higher–level processing of contours. Since textures consist of contours, a contour processing stage will process all the contours within the texture, which is unnecessary, as these contours do not represent boundaries. Therefore, it is beneficial to inhibit texture regions before higher–level processing, such as contour extraction occurs.

References

1. Haralick R. M. Statistical and structural approaches to texture, Proceedings of the IEEE, 1979, 786-804.
2. Krezhova D. Recognition of natural objects along a trace of the Earth's surface by spectral reflectance characteristics and photoimages, Annual of the Jubilee International Scientific Session "50 years University of Mining and Geology", 46, 2003, 375-379.
3. Augustejin M. F. Performance evaluation of texture measures for ground cover identification in satellite images by means of a neural-network classifier, IEEE Transactions on Geoscience and Remote Sensing, 33, 1995, 616-625.
4. Unser M. Local linear transforms for texture measurements, Signal Processing, 2, 1986, 61-79.
5. Kuan J. K. P., P. H. Lewis. Complex texture classification with edge information, Proceedings on Second International Conference on Visual Information System, San Diego, 1997.
6. Lillesand T. M., R. W. Kiefer. Remote Sensing and Image Interpretation, 4th edition, New York, Wiley, 2000.
7. Weszka J. S., C. R. Dyer, A. Rosenfeld. A comparative study of texture measures for terrian classification, IEEE Transaction on Systems, Man, and Cybernetics, 6, 1976, 269-285.
8. Pavan Kumar G. Edge detection in textured images, MS Theses, Department of Computer Science and Engineering, Indian Institute of Technology, Madras, August 1998, http://speech.cs.iitm.ernet.in/Main/publications/MSTheses/PavanThesis.ps.gz.
9. http://www.icaen.uiowa.edu/~dip/LECTURE/PreProcessing3.html#edge.
10. Canny J. A computational approach to edge detection, IEEE Transactions on Pattern Analysis and Machine Intelligence, 8, 1986, 679-714.
11. Kazuhiro F. Edge extraction method based on separability of image features, IEICE Transactions on Information and Systems, E78-D, No 12, 1995, 1533-1538.
12. http://www4.gu.edu.au:8080/adt-root/uploads/approved/adt-QGU20050810.132625/public/07Chapter4.pdf
13. Tsaneva M. Texture features for segmentation of satellite images, Cybernetics and Information Technologies, 8, No 3, 2008 (in press).
14. Rasband W. S. ImageJ, U. S. NIH, Bethesda, Maryland, USA, 1997-2007, http://rsb.info.nih.gov/ij/.

15. U m b a u g h S., Southern Illinois University Edwardsville. CVIPtools. http://www.ee.siue.edu/CVIPtools.
16. http://modis.gsfc.nasa.gov/gallery/individual.php?db_date=2007-11-30.
17. A m a d a s u n M., R. K i n g. Textural features corresponding to textural properties, IEEE Transactions on System, Man Cybernetics, 19, No 5, 1989, 1264-1274.

ТЕХНИКИ ЗА ОПРЕДЕЛЯНЕ НА ТЕКСТУРНИ ГРАНИЦИ, ИЗПОЛЗВАЩИ ТЕКСТУРНИ ПРИЗНАЦИ И ГРАДИЕНТНИ ДЕТЕКТОРИ

М. Цанева

Резюме

Текстурни граници се наричат местата в изображение, където има рязка промяна на текстурните характеристики. В работата са предложени и тествани две нови техники за определяне на текстурни граници, които използват извличане на текстурни признаци и сегментиране преди прилагането на конвенционалните градиентни алгоритми и алгоритми с пресичаща нулата втора производна за определяне на граници (на Roberts, Sobel, Prewitt, Laplacian, Kirsch, Robinson и Canny). В експериментите за текстурно сегментиране са използвани най-разпространените методи за сегментиране на изображения: fuzzy c-means, gray level quantization, histogram thresholding, median cut and principal components transformation/median cut. Представени са и са анализирани резултатите, които са получени от прилагането на тези техники за определяне на текстурни граници към спътникови изображения.

Bulgarian Academy of Sciences. Space Research Institute.
Aerospace Research in Bulgaria. 22, 2008, Sofia

SUBSTRATE MEDIA SELECTION FOR USE IN SPACE GREENHOUSES: LABORATORY APPARATUS FOR ASSESSMENT OF SEED GERMINATION RESPONSES TO GRADUATED SUBSTRATE SURFACE MOISTURE

Plamen Kostov [1], Ivan Stoyanov [2], Svetlana Sapunova [1]

[1] Space Research Institute - Bulgarian Academy of Sciences
[2] Institute of Plant Physiology - Bulgarian Academy of Sciences
e-mail: plamen_Kostov@space.bas.bg

Abstract

The study proposed is a part of research on the development of an algorithm for automatic control of plant growth environment in space greenhouses in order to maintain optimal conditions in chosen substrate at each plant development stage. In particular the dependence of seed germination on the substrate moisture level was investigated. For the purpose, a seed germination and plant growth-testing apparatus was developed to evaluate substrate agro-properties at different constant moisture, light intensity and temperature levels. The effect of Balkanine substrate moisture on the germination percentage and rate for seeds of a Lettuce crop was evaluated. Highest seed germination (82-84 %) at moisture levels between the water holding capacity and saturation and delay in passing to the activation phase at lowering water potentials was observed and discussed. A critical level of about 22% volumetric water content (VWC) for radical emergence was determined. Initial substrate moisture of 0.35-0.5 cm^3/cm^3 VWC in the sowing substrate layer is recommended for the experimental salad crop. Discussions about the number and heights of experimental dishes necessary to achieve the desired gradient of moisture levels on the substrate surface and about the possibilities provided by the apparatus to record experimentally profiles of the germination–environmental parameters relations for Balkanine substrate are drawn.

Introduction

The key to successful plant development in space is to provide suitable root environment for the growing plants. Systems for controlling root environment are required to provide steady substrate moistening, non-stop and balanced transfer of

nutrients to plant roots and good gas exchange.

A range of factors, conflicting and inadequate requirements have raised problems on the development of root modules for space application. Mass, volume and power constraints have been caused by the requirement to get maximum yield from minimum plant area, power and time. This imposes hard power and space restrictions on the equipment and reduces the root module volume which leads to higher density of components in the root area. The problems of water containment and liquid and gas phase separation in microgravity, discussed in the NASA Technical Memorandum, presented by Steinberg S.L. et al. [1], are of great importance for providing adequate water-air balance in small volume root environment. The separate-phase systems based on porous solids have established themselves for space utilization in most plant growth facilities. Substrates of solid particles are considered to be the most appropriate medium for plant growth in long term space experiments because of their longevity, repeated use and repeated crops in the same substrate.

Microgravity affects heat transfer, mass exchange processes, fluid behavior, gradients of nutritive concentration in the substrates, and capillary properties of artificial soils. The substrate media selected for plant growth research purposes should have defined physical and hydraulic characteristics. The problems with selection of a substrate capable to support suitable root environment in small volume containers on earth and in space are due to the impossibility to comply with all the requirements which are quite conflicting and sometimes mutually incompatible. In addition the process of selection involves a lot of long, hard and expensive analysis in specialized laboratories.

A lot of researches have been directed at study and control of fluid movement in a multi-porous medium. Another important issue is related to seed germination and plant growth responses to various conditions in such medium. There is a need of measurements of the substrate agro-properties that contribute to reveal its productivity. This corresponds with the research work of Kostov et al. [2] on the development of the Svet-3 SG equipment - a space greenhouse (SG) of third generation which allows evaluating plant status and optimizing growth conditions during the experiment in order to provide most favorable conditions at each stage of plant development. Such adaptive control of the plant environmental parameters requires development of an algorithm for automatic control of substrate moisture so that optimal conditions are maintained in the rhizosphere. Therefore there is a need of experiments to show what different moisture levels are necessary to be maintained in the chosen substrate at different plant development stages - from sowing to harvest.

Seed germination is strongly dependent on a variety of factors such as substrate temperature and moisture, physical properties and chemical substrate composition, quantity and quality of light, depth of sowing, plant variety, preliminary seed treatment etc. Such an experiment requires available equipment to maintain different constant light intensity, temperature and substrate moisture levels. Soil moisture is critical as it affects how quickly water penetrates into the seed. Moisture in appropriate amount for fast, uniform and complete germination of the seeds is very important.

With these considerations in mind, a seed germination and plant growth-testing apparatus was developed to study some critical problems on providing adequate water and air supply to the plant roots. It allows testing a variety of substrate media in order to evaluate substrate agro-properties at different constant moisture levels. Adjusting light intensity and constant temperature maintained by inverter air conditioner allow recording profiles of the germination–environmental parameters relations.

Materials and Methods

SUBSTRATE MATERIAL

Balkanine™ (Stoilov, G., I. Petkov, D. Dimitrov, (1979), Bulgarian Patent # 40343) was used as a root medium during the experiments.

After determining the water holding capacities for four Particle Size Distributions (PSD) of Balkanine, the 1.5-2.0 mm fraction was selected to be used for seed experiments.

PLANT MATERIAL

Lettuce (*Lactuca Sativa* L.), "Yellow beauty" variety, seeds s/n 8711989*171 (SORTOVI SEMENA I RASTITELNA ZASHTITA LTD. Sofia, Bulgaria) were used for the experiments.

LABORATORY APPARATUS

An apparatus for assessment of the agro-physical properties of substrates was developed.

Description of the apparatus

An apparatus, designed for these experiments, was specially developed and produced. It is an advanced modification of the device described by Xiao-Chun Wan et al. [3]. The apparatus is composed of a plastic water container with storage capacity of about 25 l, twelve plastic pots of different heights filled with the substrate subjected to test and a water reservoir (Mariotte siphon arrangement) which represents a 20 l polyvinyl chloride (PVC) bottle (Fig. 1).

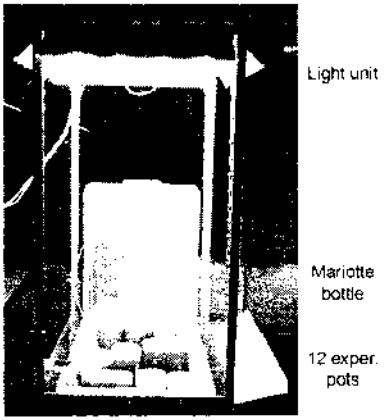


Fig. 1. A view of the experimental apparatus

The water reservoir and the container are connected by a polyethylene tubing (6 mm OD), which siphons water from the reservoir to the water container in which the pots are placed.

The water container is surrounded by a light aluminum mechanical structure, 1.2 m high, which holds fixed a movable light unit. The light unit can be positioned horizontally or at some angle and provides adjusting distance from the lighted container surface.

A horizontal shelf 60 cm long, 13 cm wide and with a slope of 15° of the narrow side is mounted on the aluminum structure. Five Petri dishes are placed there and used for a germination test experiment performed by standard method (Fig. 2).

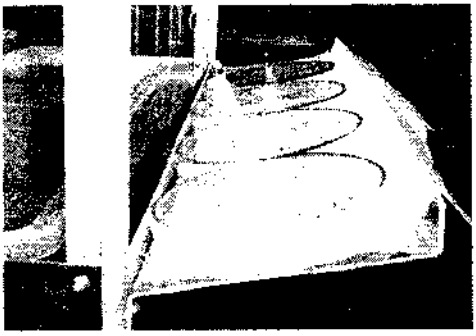


Fig. 2. Five Petri dishes are placed on the shelf

Mariotte siphon (bottle)

The water reservoir (20 liters PVC bottle) is equipped with a vertically inserted stainless steel pipe (6 mm OD) allowing ambient air to enter the reservoir through the water in the form of air bubbles. The water pressure in the bottom end of the pipe is equal to the atmospheric pressure and the water level in the container is maintained constant at the height of this end of the pipe. Another pipe is inserted below the air-entry pipe to siphon the water from the water reservoir through a flexible hose to the water container. The Mariotte bottle can be positioned at different heights to allow adjusting the desirable water table in the container. The Mariotte reservoir is filled with sufficient water amount to support a two-month plant experiment without need of recharging.

Water container and experimental pots (EP) for substrate

The water container has internal dimensions L=55 x D=35 x H=13 cm and allows disposing the 12 PVC pots filled with substrate. The container is placed on a levelled laboratory table next to the Mariotte bottle. The container has a firm bottom which bears the load of the 12 pots with substrate and the water without deformation providing good levelling and the same height of the water table in each pot with substrate.

Twelve pots, respectively 3.5, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18 and 20 cm high and with sowing area of about 86.5 cm^2 allow testing the agro-physical properties of particulate substrates at different substrate surface moisture levels and the same other environmental conditions. The pots are filled to the brim with substrate which has been previously packed to the desirable bulk density. 27 small holes, uniformly situated on the bottom pot wall, permit the water to enter the substrate from the container.

Light unit

The substrate surface is lighted by four 15 W fluorescent lamps, 60 cm long and with "Cool white" light characteristic. A 6 W fluorescent lamp of the same light characteristic is used to ensure the minimal lighting necessary for the operation of a photo camera during the "night". A timer controls the light unit operation and adjusts the photoperiod from 0 to 24 hours in steps of 15 min. A simplified hanging system provides a possibility to adjust the distance to the experimental surface at 5 to 100 cm in steps of 2 cm.

Petri controls

A shelf for the five Petri dishes (D=120 mm, H=20 mm) is mounted on the long side of the metal structure. The percentage germination of the seeds selected

for the main experiment is determined in a control experiment with 100 seeds sown in each Petri dish. Both experiments are simultaneously conducted. The Petri dishes are positioned at the mean height of the experimental pots with substrate (about 12 cm) so that all the seeds for both, the control and main experiment, are placed under comparable environmental conditions.

Observation system and data collection

Although the apparatus maintains automatically the water level in the substrate and the air-controlled system takes care of the environment, a set of sensors provide continuous monitoring of the changes in some environmental parameters. Four temperature sensors SMT160-30 (SMARTEC B.V, Delpratsingel 26, 4811 AP Breda, Netherlands) and a relative air humidity sensor SMT RH05 (SMARTEC) are located in places of interest. A photo camera Konica Q-M100 takes pictures automatically at previously set time intervals. Data obtained from the sensors and the photo camera are collected and saved in a computer. The computer operates in a mode of a Dial-Up Server which permits remote access to the data collected by a modem connection. A Web camera is connected to another computer operating as an IP Server and provides user access through the Web for visual observation in real time during the experiment. A data collection system allows inclusion of other sensors in order to expand the range of the environmental variables monitored.

Description of the experiments

The apparatus described above provides a possibility for testing all kinds of substrates (Balkanine in our case) at different moisture levels. Adjusting light intensity and constant temperature maintained by an inverter air conditioner allow recording the profiles of the germination–environmental parameters relations.

Experiment objective – evaluation of the effect of Balkanine moisture on the germination time and percentage for seeds of a Lettuce crop.

Initial conditions for the experiment:

- Balkanine substrate, 1.5-2.0 mm PSD;
- Twelve pots of different heights, filled and compacted with substrate and providing 12 different water tables above and below the sowing substrate surface: h_1 = +1.5 cm, h_2 = 0 cm, h_3 = -1 cm, h_4 = -2 cm, h_5 = -3 cm, h_6 = -4 cm, h_7 = -5 cm, h_8 = -7 cm; h_9 = -9 cm, h_{10} = -11 cm, h_{11} = -13 cm and h_{12} = -15 cm;
- Seeds of Lettuce (*Lactuca Sativa L.*), "Yellow beauty" variety; 12 pots with 100 seeds sown in each one;
- Water – from the water-main, purified by CFS-SOLVO®, coagulant for potable water treatment;

- Lighting period – 16 h day/8 h night;
- Air temperature in the laboratory - controlled at (25±2)°C - a typical ambient temperature level for the living cabin environment onboard space vehicles;
- Relative air humidity - (60-75) %, uncontrolled;
- Petri control experiment – 5 Petri dishes (D120 mm x H20 mm) with 100 seeds sown in each one;
- Duration of the experiment – 1+2 weeks.

The experiment was performed in two stages – (1) initial capillary substrate moistening in every EP till equilibrium between moistening and evaporation from the substrate surface was reached (in the first week), and (2) 100 seeds were placed among the particles of the surface substrate layer in each of the 12 EPs and 100 seeds in each of the Petri dishes (Petri control) for the next two weeks (see Fig. 3).

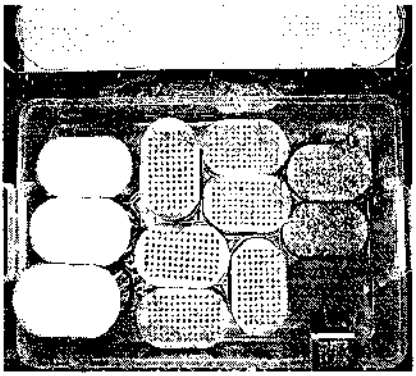


Fig. 3. Seeds are sown in the Petri controls and experimental pots using templates

The seeds in EP No. 1 were covered by a glass plate and the pot was dipped 1.5 cm below the water surface. A seed was counted germinated when the shoot reached a height of 10-15 mm. Then it was removed. As an exception to the procedure described above some plant shoots were left to the end of the experiment for visual evaluation of the germination uniformity in each EP.

A Petri control experiment was performed in 5 Petri dishes simultaneously with the main experiment in order to evaluate the seed germination for the used batch of seeds. Two sheets of filtering paper were placed in each dish and damped with 5 ml of purified water. 100 seeds were arranged on the paper using a sowing template. the dishes were covered and placed on the apparatus shelf. Both experiments (the main one and the Petri control) were performed under the same

environmental conditions. The shoots (10-15 mm) were counted and removed from the Petri dishes every day. Data about the seed germination in the Petri controls were statistically processed and the results about the germination time and percentage in the EPs were presented graphically.

A surface substrate layer of about 1 cm in height was removed in each EP and the absolute water content was determined. Data obtained were added to the data set obtained during measurements of the physical and hydro-dynamic properties of Balkanine in Russia, USA and Bulgaria (Zakharov [4]; Bingham at al., [5, 6]; Jones and Or, [7]; Ivanova at al. [9, 10]).

Substrate-water characteristic curve (SWCC) gives matric head (h) - volumetric water content (θ) relation. Fitted SWCC of Balkanine (1-2 mm PSD) is determined by Jones and Or [8] using the van Genuchten [11] nonlinear model, defined as -

$$\theta = \theta_r + (\theta_s - \theta_r)/[1 + (\alpha.|h|)^n]^m$$

where θ is current volumetric water content, θ_s is saturated volumetric water content, θ_r is residual volumetric water content, α, n, m are fitting parameters, and h is matric head (cm).

Total daily evaporation and evapo-transpiration during the experiment were determined by measuring the weight of the Mariotte reservoir with water.

The experimental apparatus ability to maintain relatively constant substrate moisture content and water gradient in each EP during the experiment was analyzed. Some suggestions for additional evaporation and transpiration measurements were made.

Results

The picture shown in Fig. 4 gives an idea of the reducing number of germinated seeds in the higher seedling pots.

Fig. 4. A view of the experiment

The germination percentage for the used batch of seeds selected for the main experiment, determined in the Petri control, is 81%±1.58 SD. The seed germination percentage distribution depending on water table in the 12 EPs is shown in Fig. 5.

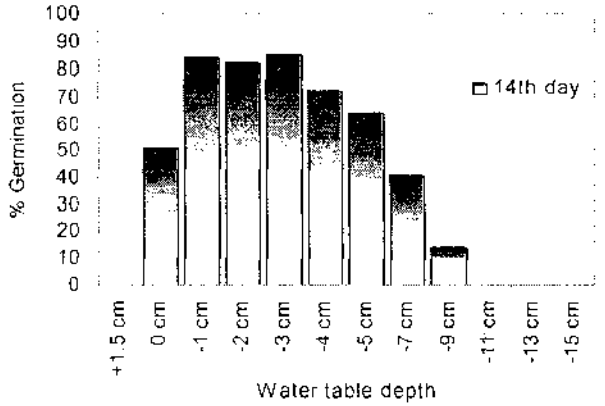


Fig. 5. Seed germination percentages in EP at the 14th DAP depending on the water table depth

Highest seed germination of (82-84) % is observed in the EPs No. 3, 4 and 5 where the substrate moisture in the sowing layer is between the water holding capacity and saturation. Water imhibitions in the lowering water potential from EP No. 6 towards EP No. 12 reduces the seed water content, postpones passing to the activation phase and delays entering the radical growth phase. Radical emergence and growth occurs when the water content exceeds a critical level. For our experiment this critical level is about 22% volumetric water content.

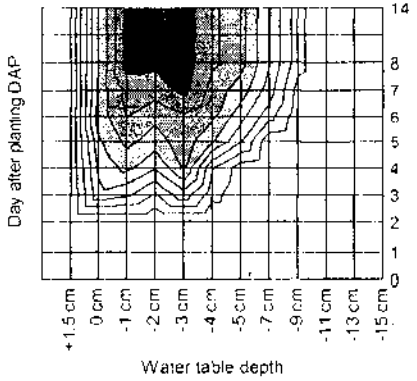


Fig. 6. A 3-dimensional plot of seed germination rates versus the water table depth in the 14-day germination period

From EP No. 6 to EP No. 9 the germination time is prolonged by 1 to 5 days respectively (Fig. 6).

Discussions

Soil moisture is critical as it affects the rate of water penetration into the seed. The germination experiment results suggest clearly a strong dependence of the seed germination percentage and rate on substrate moisture. Highest germination was observed at high moisture levels (0.35-0.5 cm^3/cm^3) between water holding capacity and saturation (Fig. 5). A few centimeters lower water table in the next EPs (matric suction at inner-particle moisture) reduces germination to zero. At the same time the germination time prolongs by 3÷8 days. Hence the automatic system for initial Balkanine wetting should provide higher moisture in the sowing substrate layer what was fixed to 0.35-0.5 cm^3/cm^3 volumetric water content or 60-85% of saturation for the experimental salad crop.

The water status of the surface substrate layers was determined for the ten EPs (No 3-12) (Fig. 7).

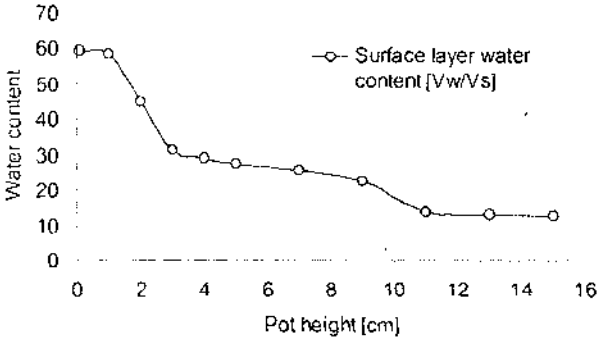


Fig. 7. Absolute volumetric water content in the surface substrate layer against the experimental pot height at the end of the experiment

The zeolite used for Balkanine production has double porosity - inner-particle and inter-particle. Zakharov [4], and Jones (in Bingham et al., [5, 6]) measure the basic hydro-physical characteristics and report that the sharp drop in the matric potential at 22% volumetric water content, due to full macro-pore water draining, leads to drastic decrease of water conductivity (Fig. 8).

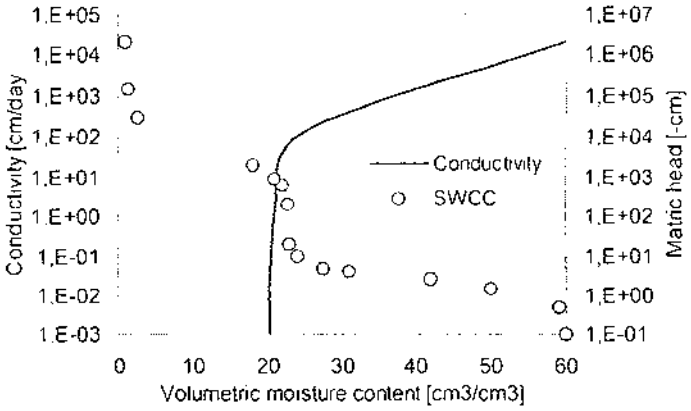


Fig. 8. Matric head and water conductivity of 1-2 mm Balkanine versus volumetric moisture content. (Jones and Or, 1999)

The total daily evaporation and evapo-transpiration during the experiment were determined by measuring the weight of the Mariotte reservoir with water. With some simplifications a mean daily evaporation of 0.15 cm^3/cm^2.day (without plants) was calculated. Shaidorov [12] reports about a mean daily evaporation of 0.24 cm^3/cm^2.day in plant experiments with Balkanine and at air flow rate of about 0.3 m/sec.

Measured and plotted values of surface layer water contents (Fig. 9) in EPs shows close congruence with SWCC plotted for a similar dish. Using the RETC program (van Genuchten et al., [13]) the retention and conductivity characteristics were modelled where θ_s=0.58 cm^3/cm^3 (saturation) and θ_r=0.23 cm^3/cm^3 (pot No. 9) were enclosed in model.

As evident from Fig. 9, the unsaturated hydraulic conductivity curve is considerably smoother than the one determined during hydro-dynamical measurements (Fig. 8) because of the evaporation. Evaporation from surface substrate layer was maintained relatively constant by controlling the ambient temperature and the streamlining air flow. Moistures of the surface substrate layers measured for EP No. 10, 11 and 12 were 0.13-0.14 cm^3/cm^3 and corresponded to the Balkanine maximum hygroscopicity.

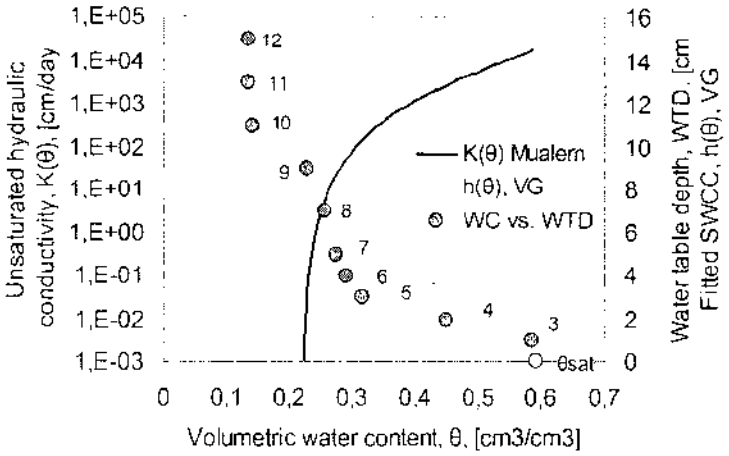


Fig. 9. Measured water contents in 1-cm surface layer for each EP (filled circlets with the EP number), fitted "SWCC" (VG model), and modelled unsaturated hydraulic conductivity versus volumetric water content (Mualem model)

Conclusions

Summarizing the results we can conclude the following: (1) using measurement data about the physical, hydrostatical and hydrodynamical properties of a particular substrate, the number and heights of experimental dishes necessary to achieve desired gradient of moisture levels on the substrate surface when placed in water container for capillary moistening can be determined; (2) using the apparatus described above the germination percentage and norm for seeds of experimental plants can be precisely determined; (3) series of plant experiments can be carried out at different temperature and light intensity gradients, substrate water control algorithm for optimal seed germination can be developed.

Acknowledgements

This research was supported in part by the Bulgarian Ministry of Education and Science under Contract Number KI-1-01/03.

The authors would like to thank technician Diana Antonova for the help and professional work on manufacturing the device and conducting laboratory tests.

References

1. Steinberg S. L., D. W. Ming, D. Henninger. Plant Production Systems for Microgravity: Critical Issues in Water, Air, and Solute Transport through Unsaturated Porous Media, NASA/TM-2002-210774.
2. Kostov P., T. Ivanova, I. Dandolov, S. Sapunova, I. Ilieva. Adaptive Environmental Control for Optimal Results during Plant Microgravity Experiments. Acta Astronautica (ISSN 0094-5765/02). Vol. 51, No. 1-9, 2002, pp. 213-220.
3. Wan X.-C., Guang-Xi Wang, Izumi Washitani. Seed germination responses of Monochoria korsakowii Regel et Maack, a threatened paddy weed, to temperature and soil moisture. Plant Species Biology 19, 2004, pp. 203-207.
4. Zakharov S.B. Some Agro-physical Properties of the Balkanine Substrate for the Svet Greenhouse Regarding the Ground Stage of the Studies, Proc. Second Micro-Symposium "Svet-90", Varna, Bulgaria, October 5-11, 1990, pp. 56-66.
5. Bingham G., S. Jones, I. Podolsky, B. S. Yendler. Porous Substrate Water Relations Observed During the Greenhouse-2 Flight Experiment, SAE Technical Paper Series 961547, 26th International Conference on Environmental Systems, Monterey, CA, July 8-11, 1996.
6. Bingham G. E., S. B. Jones, D. Or, I. G. Podolski, M. A. Levinskikh, V. N. Sytchov, T. Ivanova, P. Kostov, S. Sapunova, D. B. Bubenheim, G. Jahns. Microgravity effects on water supply and substrate properties in porous matrix support systems. Acta Astronautica, Vol. 47, No. 11, 2000, pp. 839-848.
7. Jones S. B., D. Or. Design of Porous Media for Optimal Gas and Liquid Fluxes to Plant Roots. Soil Sci. Soc. Am. J. 1998. 62:563–573.
8. Jones S. B., D. Or. Microgravity effects on water flow and distribution in unsaturated porous media: Analyses of flight experiments, WATER RESOURCES RESEARCH, Vol. 35, No. 4, 1999, pp. 929-942.
9. Ivanova T., I. Stoyanov, G. Stoilov, P. Kostov, S. Sapunova. Zeolite Gardens in Space, Proceedings of the Sofia Zeolite Meeting' 95, NATURAL ZEOLITES Sofia'95, 18-25 June 1995, (ISBN 954-642-015-8), G. Kirov, L. Filizova & O. Petrov (eds.), PENSOFT, 1997, pp. 3-10.
10. Ivanova T. N., P. T. Kostov, O. E. Petrov, I. I. Ilieva. Zeolite for Space Greenhouse Experiments on "MIR" Orbital Station, Microporous and Mesoporous Materials, Special issue "ZEOLITE'02", manuscript No. #MMM/MST/02/100TH, 2002.
11. van Genuchten, M. Th., A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci. Soc. Am. J. 1980, 44:892-898.
12. Shaydorov Yu. I., I. Stoyanov, R. O. Geodakian, V. M. Simonov. "Balkanine" substrate moisture influence on lettuce plants growth. Proc. XVII meeting of Soc. countries on Space Medicine and Biology, Gagra, USSR, May 27 – June 1, 1985, pp. 192-193, (in Russian).
13. van Genuchten M. Th., F. J. Leij, S. R. Yates. The RETC Code for Quantifying the Hydraulic Functions of Unsaturated Soils, EPA/600/2-91/065, U.S.EPA, Ada, Oklahoma, 74820.

ПОДБОР НА СУБСТРАТНИ СРЕДИ ЗА ПРИЛОЖЕНИЕ В КОСМИЧЕСКИ ОРАНЖЕРИИ: ЛАБОРАТОРЕН АПАРАТ ЗА ОЦЕНКА НА КЪЛНЯЕМОСТТА НА СЕМЕНА В УСЛОВИЯ НА ПОВЪРХНОСТНИ СУБСТРАТНИ СЛОЕВЕ С ГРАДИРАНИ ВЛАЖНОСТИ

П. Костов, И. Стоянов, С. Сапунова

Резюме

Предложеното изследване е част от научноизследователска работа по разработването на алгоритъм за автоматично управление на космически оранжерии с цел поддържане на оптимални условия във всеки етап от развитието на растенията. В частност беше изследвана зависимостта на кълняемостта на семена от почвената (субстратна) влажност. За целта беше разработен апарат за тестване на процеса на развитие на растенията, предназначен за оценка на качествата на субстрати при различни, постоянни нива на влажност, интензивност на светлината и температура. Беше оценено влиянието на влажността на субстрат "Балканин" върху процента на кълняемост и нормата на покълване на семена от салата (*Lactuca Sativa* L.). Най-висока кълняемост на семената (82-84 %) беше наблюдавана при влажност на субстрата между пределната почвена влагоемност и насищане. При намаляващи водни потенциали беше отчетено забавяне на преминаването към активационна фаза. Критична обемна влажност на субстрата от 22% беше установена за покълването на минимален процент семена. За конкретната експериментална култура беше препоръчано първоначално овлажняване на субстрата до 0.35-0.5 cm^3/cm^3 обемно влагосъдържание в посевния субстратен слой. Дискутиран е метод за определяне на броя и височината на експерименталните съдове, необходими за постигане на желан градиент на нивата на влажност в повърхностния слой на субстрата, както и възможностите, предоставяни от апарата, за експериментално снемане на профили на зависимостите кълняемост – параметри на средата за субстрат "Балканин".

Bulgarian Academy of Sciences. Space Research Institute.
Aerospace Research in Bulgaria. 22, 2008, Sofia

MODULES BASED ON NUTRITIONAL ADAPTOGENES FOR ASTRONAUTS AND GROUND CONTROL

Ljubka Georgieva, Iliana Nacheva, Tsvetan Tsvetkov

Institute of Cryobiology and Food Technology
e-mail: ljubka_georgieva@abv.bg

Abstract

The established reasonable causality and close relation between physical activities of astronauts and their nutritional regime provoke the necessity of creating appropriate biostimulating foods in the form of food additives complex.

On the basis of highly developed methods 2 specified modules are created, including 6 new types of lyophilized biostimulants, rich in nourishing proteins, essence aminoacids, poly unsaturated fatty acids, plant fibres, probiotic complexes of antioxidants, vitamins, micro and macroelements, lypotropic and other bioactive substances. Combined with suitable Bulgarian space foods, developed as the First Space Menu by the Institute of Cryobiology and Food Technology, the new type of biostimulating foods provide through alimentary way effective bioprotection of the body under extreme labour conditions and help to maintain high level of vitality, activity of the astronauts and the ground control during the preparation-training period as well.

Keywords: *lyophilized biostimulants, adaptogenes, space foods*

During the last years the multi-aspect investigations into space medicine and biology offer precious data revealing the influence of some factors over the human organism in space conditions. The most characteristic syndrome -"motion sickness", typical for humans during space flight, is due to increased acceleration. In response to the stress factors occurring during space flight, the main protective reaction of the organism is to mobilize the regulatory mechanisms of homeostasis and adaptive reorganization of metabolism. At the same time, such durable action of this complicated complex of factors over the human body leads to reduction of the adaptive abilities, working capacity, immune resistance and a series of

pathological metabolism disorders and the whole body as well. The total dynamic balance, physical and mental activity of humans, also in space condition, depends on three main groups of factors:

1) factors due to flight dynamics;
2) physico-hygienic factors;
3) psychological factors.

Nutrition is one of the main physico-hygienic factors of the working capacity. As an outer synchronize, the alimentary factor has a determinative role for maintaining high level of the astronaut's working capacity, adaptation and healthy status. Its role is to provide protection of the human organs and systems under extreme conditions of life during space flight and it helps to improve the adaptive ability of the organism and the overall working capacity. That is why, the organism under extreme conditions has to be provided with adequate quantitative and qualitative source of nutrition substances equivalent to the intensity of adaptive processes.

Main objective: "Creation of multi-aspect nutrition biostumulant modules for improving the working capacity, activity and adaptation of the organism under extreme conditions during space flight and training".

For the implementation of the main objective we apply the following approaches:

Scientific approach:

I. Formulating the composition of the new biostimulants:

• **Based on the pathogenic principle** - in order to obtain the necessary effect on the human organism under extreme conditions, the composition of each functional food is in accordance with the specifics of the metabolitic processes, energy consumption, substrate utilization and character of the adaptive changes depending on the duration and type of the physical exercises, nervous and physical exhaustion, training programme (by the ground control) etc. In order to achieve better efficiency and wide-spectrum influence on the above-mentioned pathological processes during the space flight and training exercises by the ground control, we have developed multicomponent biostumulating functional foods. Furthermore, the so-called "oxidative stress" of the organism and the professional disbiosis, occurring under extreme conditions, can be overcome due to the composition of the new food adaptogenes with biologically active substances with high antioxidant index in combination with probiotic, prebiotic and enzyme complexes.

• **Based on the preliminary dosed qualitative and quantitative composition** in respect to the characteristics of the final product, its organoleptic

properties, stable consistency and high utility, and microbial purity.

• **Based on the national and European health regulations for food safety and quality**

Developed in compliance with the above regulations, the new food adapatogenes can be classified as follows:

1) Food adaptogenes with stimulating effect on working capacity;

2) Adaptogenes with influence on tissue metabolism;

3) Adaptogenes – synbiotics with polyfunctional effect.

Based on their composition and functional effect, the new bioproducts can be combined into modules for individual food prophylaxis and recreation regimes with various pathological conditions and disorders of the organism under the unfavourable factors during space flight and nervous psychic loadings along with the physical fatigue as a result of the training exercises by the ground control.

II. Technological approach

The new generation of space foods are technologically developed in accordance with new combined methods for technological processing – multi stage cryobiotechnology with the following stages included:

- thermal treatment according to conventional methods →biotechnologi - cal processing
- enzyme immobilization + cell immobilization → fermentation processes → cryoprotection → freezing →freeze-drying.

III. Qualitative analysis of the new foods

Each product is tested in regard with the organoleptic, physicochemical, biochemical and microbiological parameters in compliance with standard methods for analysis.

Results and discussion:

Evaluation of the ingredient composition – from the medico-biological point of view, the ingredients of the new bioproducts from this module can be classified as follows:

- **structural ingredients** – cereals, nuts, fruits, fermented milk (sheep, goat, etc.);
- **energy ingredients** – fructose, honey, plant oils, polysaccharides;
- **ingredients with stimulating and regulatory function** – nutritive and

- curative plants – herbal extracts and substrates, oligosaccharides and microseaweeds – blue-green, red, brown etc; honey products-propolis, bee milk, enzymes, probiotic complexes of microbial cultures – lactic acid bacteria - **Lactobacillus bulgaricus, strain 3556, Lactobacillus acidophilus, strain 1379, Bifidobacterium bifidum, strain 1370, Streptococcus thermophilus, strain 1374, brewer's yeasts Saccharomyces cerevisiae, strains 577 and 1248** - have favourable effect in cases of **disbacterioses from various etiology, including professional bacterioses.**

These new bioproducts are 100 % natural foods without synthetic additives with plant extracts. Plants are substrates for oxidation as they maintain the acid balance, hormone balance, mediators, macro and microelements which improve the disintoxitaion function of the liver and the regeneration function of the nervous and endocrine systems due to the chemical composition [2,3].

The plant extracts are suitable for immobilization of useful microorganisms - probiotics of various origin and morphology. That's why the composition of the new functional foods contains oligo- and polysaccharides as a source of alimentary fibres. cryoprotectors and matrixes for immobilization of natural bioactive substances – probiotics and enzyme complexes of the new foods [4].

Organoleptic indicators – lyophilized products with fine consistency, pleasant aroma and taste properties specific for each product depending on its comnposition.

Chemical composition of the food lyophilized adaptogenes – series of "Synbiotics"(%/100g) : Proteins – from 8.67 to 15.30; Fats - from 17.00 to 28.00; Carbohydrates – from 60.73 to 63.50; Minerals from 3.45 to 4.90.

Active acidity (pH) of the lyophilized adaptogenes – series of "Synbiotics" varies from 5.0 to 5.7.

The residual moisture composition of all three bioproducts after lyophilization is low – from 1.98% to 3.45%. This allows to qualify these foods as high by concentrated foods.

The energy value (100 g) varies from 538.42 kcal (2252.76 kJ) to 583.48 kcal (2441.28kJ).

The mineral composition – average values of five repetitions from the obtained results are shown on Table 1.

As it is apparent from the Table 1 the new lyophilized functional foods from the series of "Synbiotics" are outlined with high composition of macroelements - phosphor, calcium, potassium, magnesium, sodium. This is due to the presence of

poly- and oligosaccharides – seaweed components, pectin, plant and vegetable extracts, brewer` s yeasts etc [1].

Table 1. Mineral composition of food adaptogenes – mg/100g

Product	P	Ca	Mg	K	Na	Fe	Mg
"Sb-Lyo-1"	443.90	522.30	51.80	780.00	247.00	2.66	0.06
"Sb-Lyo-2"	595.10	881.30	68.50	897.00	357.00	2.39	0.24
"Sb-Lyo-3"	1559.10	774.40	77.50	1913.00	628.00	7.59	0.21

Results from the biochemical analyses:

The new lyophilized food adaptogenes are characterized by high assimilation-95%. The total effect of the resultant proteolytic activity of the combined symbiotic starter cultures leads to a higher degree of coagulation of the whey and casein fraction of protein. In addition, the hydrolytic effect of enzyme proteases complex "attacks" mainly milk casein and at the same time increases its assimilation. These results were confirmed by the electrophoretic investigations conducted on the casein and whey proteins of the three types of synbiotics. The temperature treatment of the samples during lyophilization has no inluence on the quantity of fractions related to the loss of electrophoretic mobility and their denaturation.

The need of the human organism of essential aminoacids is scientifically proved and complies with the composition of the standard "ideal" protein according to FAO. The biological value of the proteins of the new bioproducts is significantly higher compared to the standart protein according to FAO/WHO. The content essential aminoacids - lysine, isoleucine, leucine, valine, threonine and tryptophan in percent dose is significantly higher than in the standart protein. Their values in standart protein vary from 1.0 to 7.0 g/100 g protein whereas the composition in the new bioproducts is significantly higher up to 15.0% and 60.00% in leucine.

The chemical score of all aminoacids from the spectrum of the non-essential ones is above 100 - an indicator of their balance and high biological value. Furthermore, it is extremely important for the organism under extreme conditions of life and work to dispose of optimal proportion of the essential aminoacids close to that of the proteins in the human body. The proportion non-essential/total quality of aminoacids of the new foods is 55.60% higher than that of the standart protein – 35.00% - an indicator of good aminoacid balance.

The biochemical analyses of the fatty acid composition of the new bioproducts have proved high content of free fatty acids. Free fatty acids are extremely important due to their rapid utilization during metabolism. Their formation provides half of the total amount of energy during metabolism [7]. In all three products, investigations have indicated high amount of free fatty acids and especially in the product "SB – Lyo-1" that proves the favourable effect of the double immobilization and particularly - the role of the hydrocolloid matrix, responsible for the cryoprotective properties for maintaining high lypolitic activity of the microbial lipase used during freeze-drying. Starter cultures have also lypolitic effect [6].

The results from the biochemical analysis of some of the main water- and fatty soluble vitamins in the new foods with adaptagene effect prove the relatively high content of provitamin A (β-carotene), vitamin E, watersoluble vitamins of group B and ascorbic acid – vitamin C that is due to the total effect of all ingredients of their composition. It is evident that the highest amount belongs to the vitamins with high antioxidant activity: ***vitamin E***, ascorbic acid (vitamin C), β-carotene (provitamin A). The healthy effect of the new type of foods is determined by the composition of water soluble vitamins of group B – B_1 (thiamin) and B_2 (riboflavin). B_1 favours the assimilation of carbohydrates, facilitates the normal function of the nervous system, muscles and cardiovascular activity whereas riboflavin, in combination with other biologically active substances, provides for the assimilation of carbohydrates, proteins and fats, stimulates growth and prevents inflammatory processes.

The high concentration of bioactive substances, including vitamins as well, of the products of the series "Synbiotics" is a factor for their antioxidant activity, spectrophotometrically determined, on the basis of the total antioxidant capacity (radical scavenging activity), the so-called "TROLOX Equivalent" [5]. This indicator provides information about the overal content of antioxidants in the product of various origin and the synergic relations between them.

The results presented on Table 2 show the relatively high antioxidant activity of the functional foods with adaptogene effect, measured by "TROLOX Equivalent", ant the value is close to that of the fresh fruits and some red wines – e.g. "Merlo".

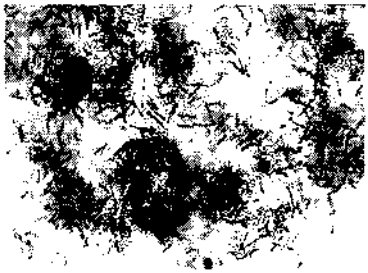


Table 2. Composition of some main antioxidants and radical scavenge activity of new adaptogenes – series of "Synbiotics"

Product	Ascorbic acid mg / %	Total carotenoides mg / %	Total Polyphenols mg/100g	Radical scavenge activity mmol TE/1000
"Sb-Lyo-1"	70.00	6.48	3700	81
"Sb-Lyo-2"	111.00	6.28	1300	115
"Sb-Lyo-3"	53.55	13.40	2000	169

Results from the microbiological investigations

The high antioxidant activity of the foods investigated with high concentration of useful microflora is 9.5×10^8 - 9.5×10^9 after lyophilization, i.e. approximately 80.00% active living cells. This fact can be explained by the effect of the hydrocolloids used - oligo and polysaccharide matrixes. They allow the stabilization of the enzyme activity of immobilized cell and increase cells survivability after lyophilization and during storage. In addition, the bioactive substances in the composition have a favourable effect on the fermentation process.

The new functional foods have been investigated in accordance with standard indicators to detect pathogenic microorganisms. The results are presented on Table 3.

Table 3. Microbiological status of food adaptogenes – series of "Synbiotics"

Type and amount of investigated microorganisms															
Total number of aerobic mesophilic microorganisms CFU/g Up to		Coliforms, CFU/g		Clostridium CFU/g		Salmonella sp. in 25g		Staphylococcus aureus in 1g		Moulds, CFU/g		Yeasts, CFU/g		Lactic acid microorganisms CFU/g	
Norm	Analysis	Norm	Analysis	Norm	Analysis	Norm	Analysis	Norm	Analysis	Norm	Analysis	Norm	Analysis	Norm	Analysis
1000	250-510	should not be determined	not determined	should not be determined	not determined	should not be determined	not determined	should not be determined	not determined	100	10-25	100	25-45	10	10

As it is apparent from Table 3, the new foods contain no pathogenic microflora and comply with the standard requirements for microbial purity. The absence of pathogenic microorganisms proves that the whole technological process is conducted in accordance with the sanitary norms and requirements. The technology - freeze-drying has a bactericidal effect and suppresses the process of pathogenic microflora growth.

The technological, biochemical, physicochemical and microbiological investigations conducted lead to the following conclusions:

Conclusions:

1. The new synbiotic bioproducts with adaptogene effect, obtained according to modern cryobiology methods, show the high biological value and favourable physiological effect for bioprophylaxis against pathological conditions occurring during space flight and strenuous training exercises.

2. After lyophilization, the nutrition and biological value of the new food adaptogenes, created for space crew and ground control, is maintained in respect to the composition of the main nutrients and biogenic components. This allows us to determine the process as accurately programmed and conducted under optimal conditions.

3. The new biostimulants can be combined with other Bulgarian lyophilized functional foods with proved effect against obesity, hyperlipoproteinemies, cardiovascular disorders, osteoporosis, intestinal disorders etc. The innovative effect of our scientific investigations is based on the principally new methods for developing nutrition modules and biostimulants for individual nutrition regimes and programmes for astronauts.

References

1. Allen L.H., R.D. Wood. (1994), Calcium and Phosphorus. In : Modern nutrition in health and disease. Shils ME, Olson JA,Shike M.(editors) – 8-th ed. Williams& Wilkins, vol. I:144-163.
2. Donald J., N.D. Brown. (1993), Probiotics and the intestinal ecosystem . Let's live, November. 45-47.

3. Gissen A.S. (1995). Of probiotics, prebiotics & synbiotics. Iss VRP Inc.Vit. Res. Prod. Newsletters, Sept., USA.
4. Luc de Vuyst (2000).Technology aspects related to the application of functional starter cultures. Food Technol.Biotechnol., 38(2)105-112.
5. Miller H.E. et all.(2000). Antioxidant Content of Wole Grain Breakfast Cereals. Fruits and Vegetables; Journal of the American College of Nutrition, Vol.19, № 3, 312 - 319.
6. Salminen S.at al.(1998) Demonstration of safety of probiotics – safety assured Scandinavian University Press.
7. Vaughan E.E., B. Mollet (1998) Nahrung, Probiotics in the new millennium, 43(3), 148 - 153.

МОДУЛИ ОТ ХРАНИТЕЛНИ АДАПТОГЕНИ ЗА АСТРОНАВТИ И НАЗЕМНИ СЪСТАВИ

Л. Георгиева, И. Начева, Цв. Цветков

Резюме

Установената причинна обусловеност и връзка между работоспособността и активността на космонавтите и тяхното хранене, обосновават необходимостта от създаване на подходящи биостимулиращи храни, под формата на комплекс от хранителни корегиращи добавки.

На базата на високотехнологични методи са създадени два вида специализирани модула, включващи 6 вида нови лиофилизирани биостимуланти, богати на пълноценни протеини, есенциални аминокиселини, полиненаситени мастни киселини, растителни фибри, пробиотични комплекси, на антиоксиданти, витамини, микро-и макроелементи, на липотропни и други биоактивни вещества. Съчетани с подходящи български космически храни, разработени в института като първо българско космическо меню, новите биостимулиращи храни осигуряват по алиментарен път, ефективната биозащита на организма, поставен в екстремални условия на труд и бит, допринасят за поддържане висока ниво на жизненост, работоспособност и адаптивност на космонавтите, както и на наземните екипи по време на подготвително-тренировъчния период.

Bulgarian Academy of Sciences. Space Research Institute.
Aerospace Research in Bulgaria. 22, 2008, Sofia

MODELING AND ANALYSIS OF ELECTRO-MECHANICAL FLIGHT ACTUATOR SYSTEM

Marek Hièár

Honeywell, spol. s r.o., Aerospace, Advanced Technology
Spielberk Office Center, Holandská 2/4, 639 00 Brno, Czech Republic
e-mail: Marek.Hicar@honeywell.com

Abstract

Electro-mechanical flight actuator (EMFA) systems improve efficiency, reliability, safety, and maintainability in More Electric Aircraft (MEA) programs intended for next generation commercial aircrafts, as compared to the traditional hydraulic systems. A power-by-wire system actuation approach, based on incorporating high torque density brushless motors featuring superior characteristics, is increasingly used in the design of both primary and auxiliary flight control surfaces, improving the system performance while reducing its weight. An EMFA system based on a permanent magnet three-phase synchronous machine fed by a Variable Frequency Current Hysterisis Controlled Inverter and additionally incorporating two outer loops, a position and a velocity loop, is proposed in this work. A MATLAB Simulink/SimPowerSystems software tool is used to model the system and analyze its functionality, accuracy, stability, disturbance rejection and acceleration capabilities, for specific cases of landing and takeoff.

Intoduction

More Electric Architectures have primarily involved the use of much larger number of electromechanical actuators for primary flight and secondary controlled surfaces. Power-by-wire technology payoffs include elimination of the central hydraulic system, reduction of maintenance support, increased survivability significant reliability improvement, more efficient use of secondary power, and improve ability to turn off or reconfigure a damaged or inoperative flight control surface.

Fly-by-wire or fly-by-light control of EMFAs is executed by Flight Control Systems and is part of flight-critical vehicle components. Advanced simulation engineering software tools are used for control design, development, testing and verifying of the simulated results [1, 2]. The elimination of hydraulic and pneumatic secondary power systems will improve maintainability, increase reliability, reduce life cycle costs, increase energy efficiency and improve flight readiness. This explains why, in recent years, there has been much interest in the 'all-electric' aircraft, and its supporters have emphasized the serious consequences of hydraulic fluid loss and the weight and space disadvantage of a centralized hydraulic power distribution system, especially where there are large distances between the primary power source and the actuators. Therefore, electric surface actuation systems will only show a significant weight saving advantage, if the hydraulic system is removed completely from the aircraft [3].

Electro-mechanical actuators involve conversion of rotary motion (from an electrically powered source) into linear or rotary displacement. There are many designs of modern linear actuators and every company that manufactures them tends to have its own proprietary methods and designs. In one approach, a rotor driver is mechanically connected to a long shaft so that the rotation of the electric motor will make the shaft rotate. The shaft has a continuous helical thread machined on its circumference running along the length. Threaded onto this shaft is a nut with corresponding helical threads. The nut is kept from being able to rotate with the shaft (this often involves an interlocking of the nut with a stationary part of the actuator). Therefore, when the shaft is rotated, the nut will be driven up or down the threads depending on the direction of rotation. By fashioning linkages to the nut, this can be converted into usable linear displacement. Most current actuators are built either for fast speed, high torque capacity, or a compromise between the two.

Typical landing flaps drive system architectures of rotary motion conversion into linear displacement for flaps mechanism is [4]:

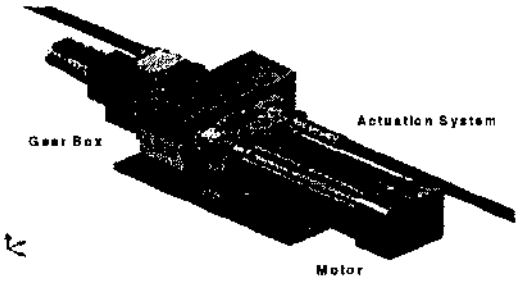


Fig.1. Landing Flap Drive System

Permanent Magnet Synchronous Machine SimPowerSystems Model Description

This paper deals with the Permanent Magnet Synchronous Machine (PMSM) model applied to drive the flap actuation system. The PMSM operates in motoring mode where positive mechanical torque is required. The electrical and mechanical parts of the machine are each represented by a second-order state-space model. The sinusoidal model assumes that the flux established by the permanent magnets in the stator is sinusoidal, which implies that the electromotive forces are sinusoidal. For the trapezoidal machine, the model assumes that the flux established by the permanent magnets is purely trapezoidal, which implies a trapezoidal electromotive forces waveform [5].

The sinusoidal model electrical system implements the following equations in $d-q$ transformed plane:

$$\frac{di_d}{dt} = \frac{1}{L_d} u_d - \frac{R}{L_d} i_d + \frac{L_q}{L_d} p\omega_r i_q,$$

$$\frac{di_q}{dt} = \frac{1}{L_q} u_q - \frac{R}{L_q} i_q + \frac{L_d}{L_q} p\omega_r i_d - \frac{\lambda p\omega_r}{L_q},$$

$$T_e = 1{,}5p\left(\lambda i_q + \left(L_d + L_q\right) i_d i_q\right), \quad (1)$$

where L_q, L_d - axis inductance, R - stator windings resistance, u_d, u_q - axis voltage, ω_r - rotor velocity, λ - amplitude of the flux induced by the permanent magnets of the rotor in the stator phases, p - number of pole pairs, T_e - electromagnetic torque.

Trapezoidal model electrical system in a, b, c frame:

$$\frac{di_a}{dt} = \frac{1}{3L_S}\left(2u_{ab} + u_{bc} - 3R_S i_a + \lambda p\omega_r\left(-2\phi_a' + \phi_b' + \phi_c'\right)\right),$$

$$\frac{di_b}{dt} = \frac{1}{3L_S}\left(-u_{ab} + u_{bc} - 3R_S i_b + \lambda p\omega_r\left(\phi_a' - 2\phi_b' + \phi_c'\right)\right),$$

$$\frac{di_c}{dt} = -\left(\frac{di_a}{dt} + \frac{di_b}{dt}\right), \quad T_e = p\lambda\left(\phi_a' i_a + \phi_b' i_b + \phi_c' i_c\right), \quad (2)$$

L_S - stator windings inductance, R - stator windings resistance, i_a, i_b, i_c - a, b, c phase currents, $\phi_a^{'}, \phi_b^{'}, \phi_c^{'}$ - a, b, c phase electromotive forces, u_{ab}, u_{bc} - a, b and b, c phase to phase voltages.

Mechanical system:

$$(3) \qquad \frac{d\omega_r}{dt} = \frac{1}{J}\left(T_e - F\omega_r - T_m\right), \quad \frac{d\Theta}{dt} = \omega_r,$$

J - inertia of rotor and load, F - friction of rotor and load, Θ - rotor angle position, T_m - load torque.

The sinusoidal machine is simulated in the $d-q$ rotor reference frame and the trapezoidal machine in the $a\,b\,c$ frame. Stator windings are connected in wye to an internal neutral point.

Variable Frequency Current Hysterisis (VFCH) Controlled Inverter SimPowerSystems Model

A DC / AC inverter is represented by current comparison blocks and full-bridge inverter. It uses ideal switches and antiparallel diodes and also demonstrates the VFCH Control provided by current comparison between reference and actual values. Six pulses are generated for a three-arm bridge.

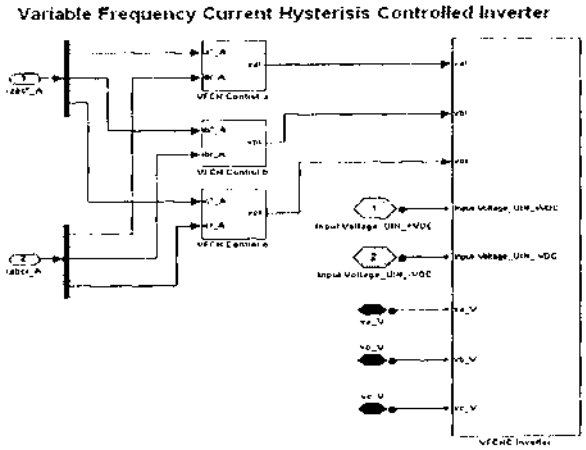


Fig.2. Variable Frequency Current Hysterisis Controlled Inverter - Layer 1

Comparion blocks include also low pass filters to eliminate high frequency noise. The hysteresis modulation is a feedback current control method where the motor current tracks the reference current within a hysteresis band. The following figure shows the operation principle of the hysteresis modulation. The controller generates the sinusoidal reference current of desired magnitude and frequency that is compared with the actual motor line current. If the current exceeds the upper limit of the hysteresis band, the upper switch of the inverter arm is turned off and the lower switch is turned on. As a result, the current starts to decay. If the current crosses the lower limit of the hysteresis band, the lower switch of the inverter arm is turned off and the upper switch is turned on. As a result, the current gets back into the hysteresis band. Hence, the actual current is forced to track the reference current within the hysteresis band.

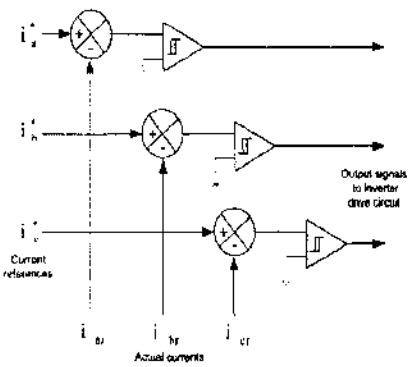


Fig. 3. Variable Frequency Current Hysterisis Control Scheme

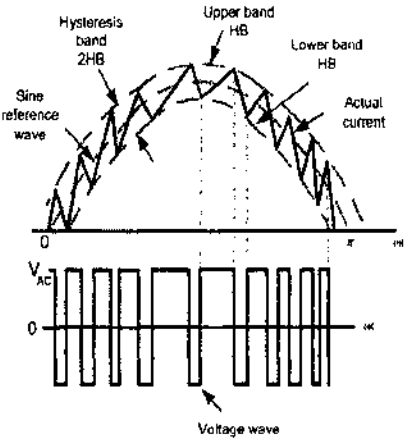


Fig. 4. Voltage Wave Control Input Signal Generation

The inverter is designed to accomodate Dead Zone T_d _s to eliminate simultaneously turning on the switches in *a, b* or *c* phase (if both switches in one phase are on, it produces short current in that phase, Fig.5). Each ideal switch consists of a diode to keep the current flow in the desired direction and also anti parallel diode (free wheeling diode) to ensure current flow for inductive loads.

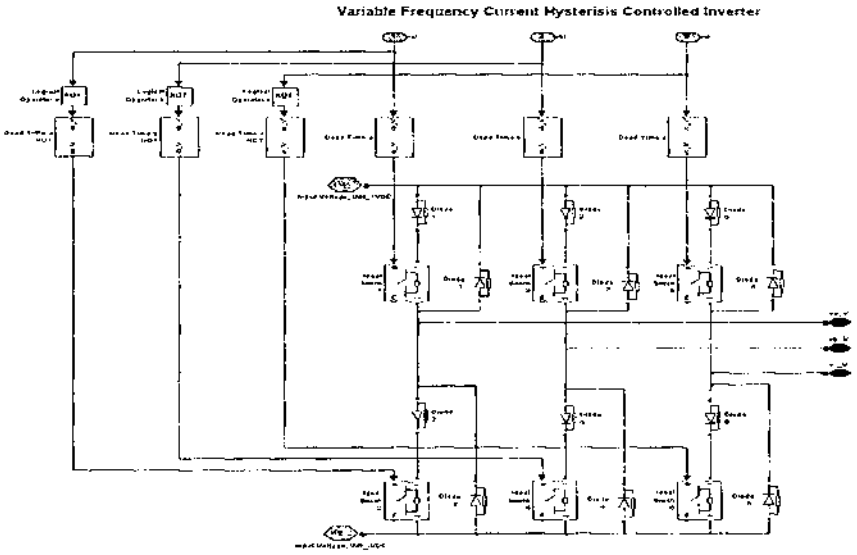


Fig. 5. Variable Frequency Current Hysterisis Controlled Inverter - Layer 2

Electro-Mechanical Flap Actuation SimPowerSystems Model

Flaps generally increase the camber of the wing and therefore increase the wing's lift. This allows a slower speed without stalling. Effects of the flap also permit steeper glide angle in the landing approach. The flap is a hinged surface attached to the trailing edge of the wing.

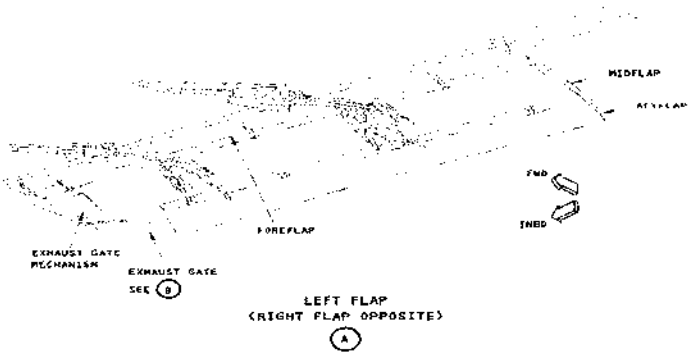


Fig.6 Trailing Edge Flaps Extracting

Flaps on both wings move in unison in response to the flight deck control input. When the flaps are not is use, they are stored as part of the trailing edge of the wing. The slotted flap is equipped with tracks, rollers, or hinges of a special design. During operation, the flap moves downward and rearward away from the position of the wing. The opened slot allows a flow of air over the upper surface of the flap to streamline the airflow and to improve the efficiency of the flap [6].

This paper deals with a high-lift device as an integrated model which simulates an Electro-Mechanical Flight Actuator System as a Landing Flap Drive System (Fig.7). A Synchronous Motor is fed by the VFVHC Inverter. Input parameter is Reference Flaps Position_F_grade which starts executing the retracting or extending flap mechanism. Position and Speed PI Controllers are used and final error signals enter the $dq2abc$ transformation block. Measured Rotor Angle_Θ_rad, Actual Flaps Position_F_c_grade and Rotor Speed_ω_m are the control loop feedbacks to Position and Speed Controller and $dq2abc$ calculation block:

$$
\begin{aligned}
i_a^* &= i_{qref\_A}\cos(\Theta_e) + i_{dref\_A}\sin(\Theta_e) + i_{oref\_A},\\
i_b^* &= i_{qref\_A}\cos\left(\Theta_e - 2\frac{\pi}{3}\right) + i_{dref\_A}\sin\left(\Theta_e + 2\frac{\pi}{3}\right) + i_{oref\_A},\\
i_c^* &= i_{qref\_A}\cos\left(\Theta_e + 2\frac{\pi}{3}\right) + i_{dref\_A}\sin\left(\Theta_e - 2\frac{\pi}{3}\right) + i_{oref\_A}.
\end{aligned}
\tag{4}
$$

Calculated i_a^*, i_b^*, i_c^* are compared with measured Stator Currents_i_{sa}, i_{sb}, i_{sc} and the result is processed by Variable Frequency Current Hysterisis Control (Fig.2, 3).

The Permanent Magnet Synchronous Machine has load torque input represented by air drag doing the extending or retracting process. Circular motion is transformed to linear by:

$$
F_c = \int \frac{\omega_m r}{p}, \quad r \text{ - radius of single transmission.} \tag{5}
$$

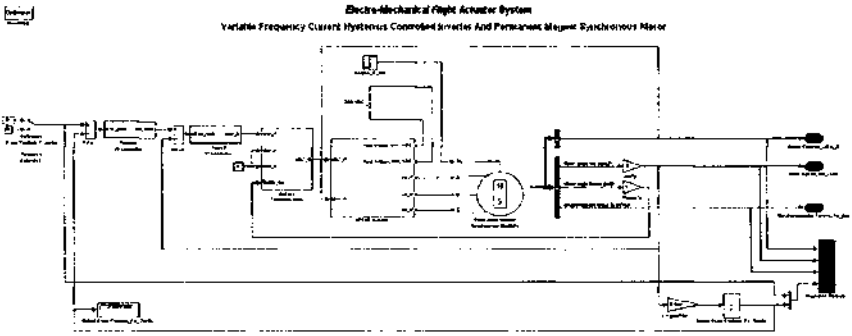


Fig. 7. Variable Frequency Current Hysterisis Controlled Inverter - Layer 2

Electro-Mechanical Flap Actuator Simulation Results

In this paper, forward and backward movement of the flap control lever is transferred directly by fly-by-wire signals as Reference Flaps Position. Landing flap functionality in extracting mode is displayed in Fig. 8. Reference Flap Position has been increased by step-step position function to reach the final extended position of 40 degrees. Stator currents, rotor speed and actual flaps position involving the load torque to the synchronous machine is shown. Full flaps extension is finished in 3,2 s.

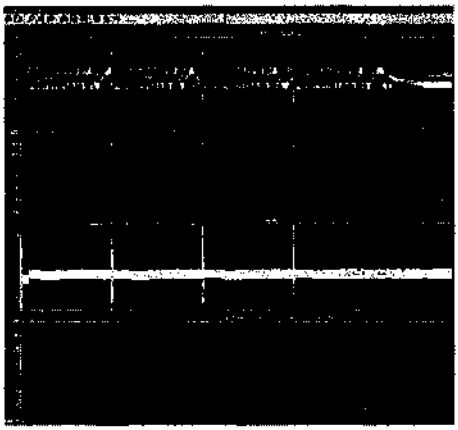


Fig. 8. Landing Flaps Extending (0 - 40 degrees)

Fig. 9 shows detail of landing flap extension from 20 to 30 degrees and the permanent magnet synchronous machine accelerating process: VFCHC Inverter stator currents and speed control. When 30 degrees position is reached, the synchronous machine is producing torque to oppose only the input load torque and remain in same position at zero rotational speed.

Backward movement of flaps - retracting is demonstrated in Fig. 10., the synchronous machine moves in opposite direction, with the same input torque profile and currents corresponding to acceleration profile.

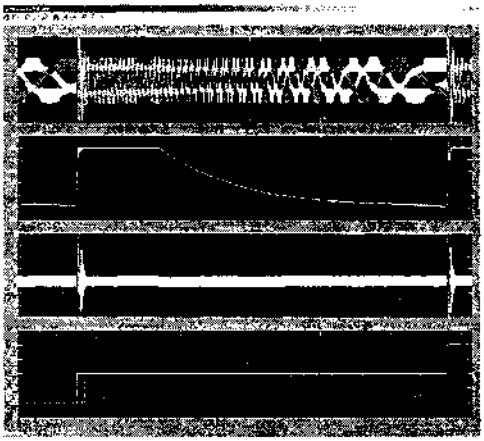


Fig. 9. Landing Flaps Extracting (20 - 30 degrees)

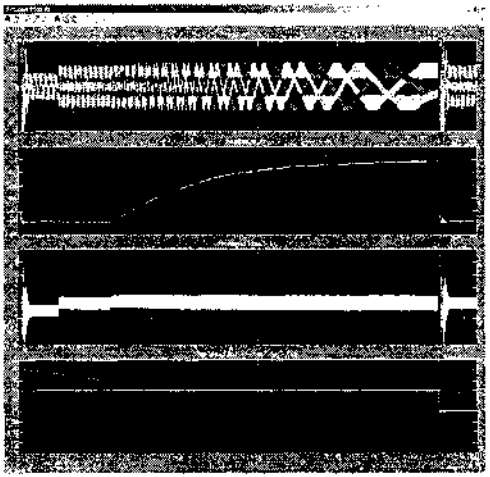


Fig. 10. Landing Flaps Retracting (40 - 30 degrees)

Conclusion

The advantages of using Electro-Mechanical Actuating Systems have been claimed for control surface applications. Electrical machines can execute accurately and quickly either multiple functions or single tasks. This paper has demonstrated the application of SimPowerSystems toolbox of MATLAB/Simulink for actuating systems, including control, analysis, integration, and verification. Extending and retracting operation modes of landing flap actuating system are demonstrated within a power optimized concept.

References

1. Dawson D. Air Force Research Laboratory's Air Vehicles Directorate, www.google.com
2. Lyshevski S., V. Skormin, R. Colgren. High-Torque Density Integrated Electro-Mechanical Flight Actuators, www.google.com
3. Annaz F. Fundamental design concepts in multi-lane smart electromechanical actuators, www.google.com
4. ROSNER-TDL, Supplier of Customised Actuation Systems, www.rosner-tdl.de
5. MATLAB Version 7.1.0.246 (R14) Service Pack 3, Simulink, SimPowerSystems.Permanent Magnet Synchronous Machine, (Louis-A. Dessaint and R. Champagne (Ecole de Technologie Superieure, Montreal)).
6. Lombardo D. Advanced Aircraft Systems, ISBN 0-07-038603-X.

МОДЕЛИРАНЕ И АНАЛИЗ НА ЕЛЕКТРОМЕХАНИЧНА СИСТЕМА ЗА АКТИВАЦИЯ НА ПОЛЕТА

М. Хичар

Резюме

Електромеханичните системи за активация на полета (EMFA) подобряват ефикасността, надеждността, безопасността и възможностите за поддръжка на програмите MEA, предназначени за търговски

въздухоплавателни средства от следващо поколение, в сравнение с традиционните хидравлични системи. Методът за активация чрез система с проводниково захранване, основаващ се на използването на безчеткови двигатели с висока степен на усукване, които притежават по-добри характеристики, се използва все по-често в разработката както на основни, така и на спомагателни повърхности за контрол на полета, подобрявайки к.п.д. на системата и намалявайки теглото й. В работата се предлага EMFA система, основаваща се на трифазен синхронен двигател с постоянен магнит, захранван от променливо-честотен инвертор с токово-хистерезисно управление, включващ и два външни контура – позиционен и скоростен. За моделиране на системата и анализ на нейната функционалност, точност, устойчивост, защита от смущения и възможности за ускоряване в конкретни случаи на кацане и отлитане е използвано програмното средство MATLAB Simulink/ SimPowerSystems.

Bulgarian Academy of Sciences. Space Research Institute.
Aerospace Research in Bulgaria. 22, 2008, Sofia

COHERENT FORMATION AND RECEPTION OF FREQUENCY HOPPING SPREAD SPECTRUM SIGNALS IN AERO-COSMIC RADIO LINES

Antonio Andonov[1], Galina Cherneva[1], Zoya Hubenova[2]

[1]Todor Kableshkov University of Transport, Sofia
[2]Space Research Institute – Bulgarian Academy of Sciences
e-mail: andonov@vtu.bg; zhubenova@space.bas.bg

Abstract

The creation of coherent radio lines is one of the ways to increase the noise-resistance and reach conditions making possible optimal signal reception. The paper presents a method and a device for the purpose mentioned, both protected by an application for a patent. It is based on the two American patents of Motorola Company known up to now and used to accomplish a coherent connection in USA Air Force. The approach suggested provides for simpler and more effective technical implementation to build coherent radio lines under the condition of fast signal fluctuations. This makes possible its use in both aviation and cosmic radio lines where it is necessary to transmit an increasing amount of telemetric information as coherence ensures considerably higher amplification with the received signals' processing in comparison with the ways used up to now. An algorithm of signal processing based on the theory of non-linear filtration has been suggested and grounded.

1. Introduction

A number of modern information transmission systems such as the systems of space radio connection and of mobile communications are characterized by common requirements for the used signals. These requirements are most completely satisfied by using frequency hopping spread spectrum signals (FHSS). The properties of those signals in combination with the optimal methods of their processing make it possible to ensure a high level of accuracy with measuring distance and speed, to

combine the transmission of information with trajectory measurements, to increase the efficiency of radio systems' operation with regard to electromagnetic compatibility and, under multi-beam conditions, to achieve energy and structural obscurity of emission, etc.

For space radio lines and mobile communications of limited energy resources, it is extremely urgent to develop systems of connection with coherent JLCCF, as they make it possible to ensure considerably greater signal amplification with coherent processing in the receiver. However, at the same time, the requirements concerning the system of synchronization and monitoring delays are considerably increasing. The purpose of this paper is to propose a method and a device for accurate independent and coherent reproduction of the shape of the transmitted FHSS in the transmitter and the receiver.

2. Systems of monitoring the delay with coherent receiving signals of JLCCF

Two main patents of coherent systems have been known up to now. They belong to the Motorola company and have been applied to implement a radio connection by using FHSS and fast-moving objects (Air Force airplanes), i.e. under the conditions of quick fluctuations of delay [1, 2]. A coherent system of synchronization with a phase synthesizer is described in [2].

The principal peculiarity of the system proposed consists of using a special device to synthesize the phase in the receiver as a local generator: a synthesizer performing the voltage-phase transformation. In this way, the controlled signal for the phase synthesizer should be equal to the value of the current phase of the frequency synthesized. The structural diagram for the delay observation system (DOS) is given in Fig.1.

The input signal in the receiver is presented by the equation:

(1) $$S_1(t) - \mathrm{Re}\left\{e^{j[\omega_p t + \varphi_1(t)]}\right\},$$

where ω_p is the frequency of the pilot signal.

The phase of the received coherent FHSS can be written in the form:

(2) $$\varphi_1(t) = \omega_0 \tau + \int_0^t \omega(t')dt',$$

where ω_0 is the central frequency of the signal, $\omega(t')$ is the amplitude of the random oscillation of rectangular shape that determines the frequency values of the signal of

JLCCF and has been distributed evenly between ω_p and $-\omega_p$.

To accomplish the synchronization between the received and the reference signals according to delay, it is necessary to reduce the phase error to zero, i.e.:

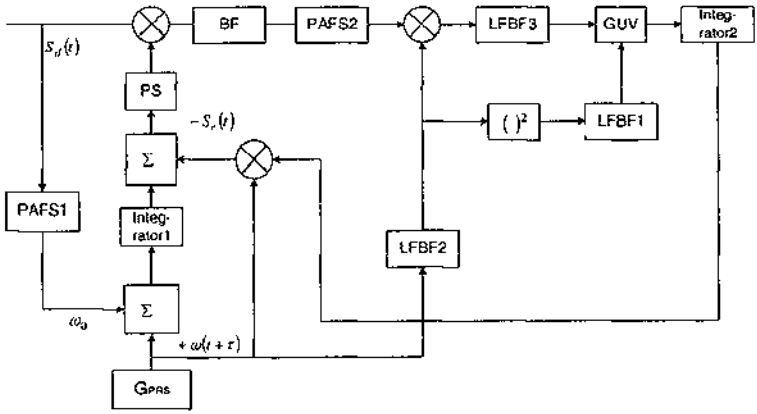


Fig.1

(3) $$\Delta\varphi = \varphi_2(t) - \varphi_1(t),$$

where $\varphi_2(t)$ is the integral phase of the supporting signal.

For a case when $\tau << \tau_c$ the approximate equation has been satisfied:

(4) $$\Delta\varphi \approx \tau \frac{d\varphi_1(t)}{dt},$$

or $$\frac{d\varphi_1(t)}{dt} \approx \omega_0 + \omega(t).$$

Hence, the assessment of the output signal phase can be presented in the form:

(5) $$\hat{\varphi}(t) = \varphi_2(t) - \tau \frac{d\varphi_1(t)}{dt}.$$

At that, the error of the obtained assessment can be evaluated by the expression:

(6) $$\Delta\omega(t) = \omega(t+\tau) - \omega(t),$$

where $|\Delta\omega(t)| \le |\omega(t)|$.

The method suggested can be implemented by the diagram in Fig. 1 in the following way. The signal determining the law of the phase change for the phase synthesizer (PS), is formed at the output of the first generator. It presents the integral sum of the phases of the central frequency obtained at the output of the circle of phase automatic frequency setting (PAFS 1) of the input radio signal and the signal entering from generator of pseudo-random sequence (G_{PRS}) determining the value of $\omega(t)$. The preliminary evaluation of the phase error formed on the basis of delay τ_3 is taken out from the obtained value of the phase in the extracting device and the result obtained is used as a controlling effect for the PS. In this way, the PS generates a signal of the form:

$$(7) \qquad S_2(t) = \mathrm{Re}\left\{\exp j\left[-\tau_3\omega(t) + \omega_0(t+\tau) + \int_0^{t+\tau}\omega(t')dt'\right]\right\}.$$

The synthesizer examined presents an address memory device (MD) storing *exp jx*, values where *x* is a number fed at the MD input. At that, a digital-to-analog transformer operating according to discrete $\mathrm{Re}\{\exp jx\}$ is used at the MD output.

The signal of the mixer output is of the form:

$$(8) \qquad S_m(t) = \mathrm{Re}\{\exp[(\tau-\tau_3) - \bar{\omega}(t)] - \tau_3\tau\dot{\omega}(t)\}.$$

Signal $S_m(t)$ passing through band filter (BF) enters at the input of the circle of PAFS 2. In the case examined, according to its structure and purpose the circle of PAFS 2 is analogous to the circle of PAFS 1 in DOS described in [1]. The signal at the output of the circle of PAFS 2 is fed to the first input of the multiplier of the circuit of delay evaluation, while at the other input of the circuit, the value of $\omega(t)$ averaged in low frequency band filter (LFBF1) enters, obtained by G_{PRS}. The time constant of LFBF1 corresponds to the time constant of BF. In this way, at the output of the multiplier, the following signal is obtained:

$$(9) \qquad S_4(t) = (\tau-\tau_3)\bar{\omega}^2(t).$$

Besides that, the averaged value of $\bar{\omega}(t)$ is fed to the square power device and after averaging in LFBF2 with a time constant bigger than the one of LFBF1. the signal obtained is in the form:

$$(10) \qquad S_5(t) = \bar{\omega}^2(t).$$

This signal is fed as a divisor at the input of a division diagram. In the capacity of dividend, a signal of the following type is used:

(11) $\overline{S}_{\partial}(t) = (\tau - \tau_з)\overline{\omega}^2(t)$,

Obtained from $S_{\partial}(t)$ by averaging in LFBF3. In this way, a new evaluation of the delay is formed:

(12) $\tau'_з = \frac{S_{\partial}(t)}{S_d(t)} = \tau - \tau_з$,

which at the integrator output is in the form:

(13) $\overline{\tau}'_з = \overline{\tau} - \tau_з$

The evaluation enters a multiplier where the consecutive signal is formed for the correction of the frequency of the DOS monitoring system.

3. Phase synthesizer

The aim is to develop a device for independent and accurate (coherent) reproduction of the phases in the transmitter and receiver of FHSS and on the base of it, a coherent device of synchronization serviceable with random fluctuations of delays [3].

The problem is solved by creating a frequency synthesizer including a source of a standard (bearing) signal, generator of primary signals and switching circuit. The series of primary signals with angular frequency corresponding to the bearing signal frequency is taken to the switching-over circuit output in the form of a phase-displaced line. The switching-over circuit is synchronized with one of the primary signals and switching-over can be controlled by choosing the number of primary signal cycles.

The advantage of the frequency synthesizer suggested is that the generator of primary signals generates them with the same spectrum frequency as the source of the bearing signal (S_{BS}) and hence, if the switching circuit has been synchronized with one of the primary signals, the signal chosen with the same phase is obtained every time with the synthesis' start which makes it possible to control the synthesized signal phase and to reproduce it independently in the transmitter and the receiver.

The performance of the suggested frequency synthesizer as an example is shown in Fig. 2 (block PS). According to [3], the frequency synthesizer consists of a source of a bearing signal, generator of primary simple signals, accomplished in the form of a multi-terminal delay line with N terminals, switching-over circuit, divisor

of frequencies synchronized with one of the simple signals of the generator and connected with the tact input of the switching-over circuit.

The synthesizer operates in the following way: The signal of the frequency standard generator of angular frequency ω_0 is fed to the frequency standard generator input (multi-terminal delay line (MTDL) with N terminals). The switching circuit is synchronized through the divisor of frequencies and switches over after every k cycles of the primary signal. If the time interval between two switchings over is T_s, the phase difference between the signal synthesized and the bearing one increases by phase θ for time T_s. As $\theta=2\pi/N$, the signal synthesized can be supported for a random period of time with multiple usage of N primary signals and the synthesized signal frequency ω_c is $\omega_c=\omega_0+k\theta/T_s$. By the control of the switching circuit (the choice of k), the output signal phase can be changed.

On the basis of the synthesizer suggested, a device for synchronization of coherent FHSS was obtained. The device consists of a transmitter that can be regulated, receiver, diagram of monitoring by delay, extrapolator of delay and the bearing signal source, generator of primary simple signals, switching circuit, divisor of frequencies.

The advantage of the device suggested is that the moment to start the synthesis of the desired signal in the receiver correlating unit can be determined on the basis of extrapolation of the delay evaluation as receiving the same phase as the one used in the transmitter.

The implementation of the device for coherent synchronization as an example according to [3] shown in Fig. 2. The device of synchronization consists of a transmitter that can be regulated, receiver, diagram of monitoring by delay, extrapolator of delay and source of a bearing signal, generator of primary simple signals, switching circuit, divisor of frequencies.

The device operates is the following way: The current value of delay is evaluated in the standard diagram of monitoring by delay. To compensate the delay, the extrapolator determines value $\tau^*(t)$ and from it, overtaking in time, i.e. $\tau^*(t)=t-\tau(t)$ of transmitting in transmitter that is implemented by feeding the signal at the input of the generator of bearing signals (multi-delay line).

Let us denote one of the n in number signal elements, which are approximately equal to zero out of the interval, by S[0,T]. Then the signal transmitted can be expressed as follows:

$$S(t)=\sum_{k=0}^{n} S_k(t-iT). \tag{14}$$

Considering the random delay, the obtained useful signal is:

$$S(t,\tau(t)) = S(t-\tau(t)) = \sum_{i=0}^{n} S(t-iT-\tau(t)). \quad (15)$$

Let us present the discrete parameter $\theta(t)$ and the random delay $\tau(t)$ in the form of known information function, i.e.:

$$S(t,\theta,\tau) = S[t-\tau(t),\theta(t-\tau(t))]. \quad (16)$$

The discrete parameter takes constant values of the tact intervals $\theta(t) = \theta_i, t \in [t_i, t_{i+1}]$. The values of the information parameter of the tact

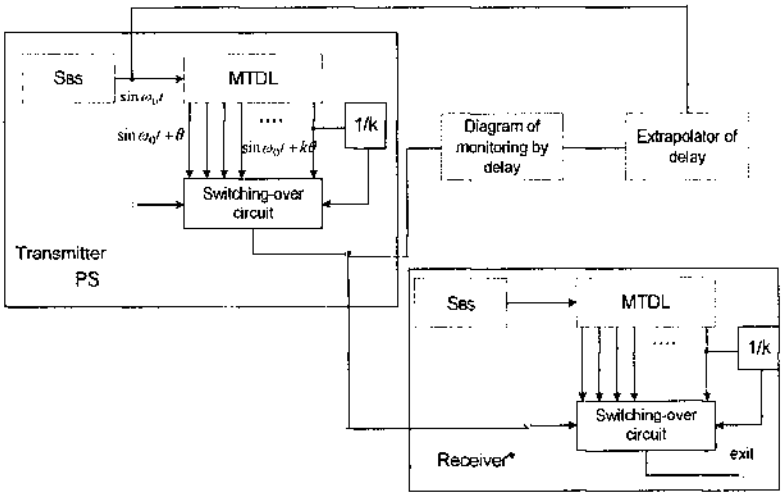


Fig.2

intervals form a normal Markov chain $\theta_i(t), i = 0,1,...n$ with n states and a known matrix of transmitting state i into state j $\Pi = \Pi_{ij}$ and a vector of initial states $\bar{p} = p_i$. The boundaries of the tact intervals are determined by the random delay $\tau(t)$, i.e. $t_i = t_i(\tau)$. With accomplishment of certain delay, the boundaries of the tact intervals are $t_i = iT + \tau(t_i)$.

In the tact interval t_i, t_{i+1} signal $S(t,\theta,\tau)$ coincides with the elementary

signal $S(t-iT-\tau(t))$ if $\theta(t)=\theta_i$. The process $\lambda(t)$ satisfies the system of stochastic differential equations:

$$\frac{\partial\lambda_i(t)}{\partial t}=f_i(t,\lambda)+n_i(t). \tag{17}$$

Here, $f_i(t,\lambda)$ are functions satisfying the condition of Lipschitz and $n_i(t)$ is Gauss's noise of intensity $b_{ij}(t,\lambda)$. The apriori probability characteristics of the process $\lambda(t)$ have been determined by the equation of Fokker-Plank-Kolmogorov:

$$\lambda(t)\frac{\partial W}{\partial t}=-\sum_{\alpha=1}^{n}\frac{\partial}{\partial\lambda}[a_\alpha(t,\alpha)W]+\frac{1}{2}\sum_{\alpha=1}^{n}\sum_{\gamma=1}^{n}\frac{\partial^2 b_{\alpha\gamma}(t,\alpha)W}{\partial\lambda_\alpha\lambda_\gamma}\equiv L[W], \tag{18}$$

where $W=W(t,\lambda)$ is the apriori the probability density of the process, $\lambda(t)$.

The observation of the signal $S(t,\theta,\tau)$ has been made against noise background, i.e. it has the form:

$$\xi(t)=S(t,\theta,\tau)+n(t), \tag{19}$$

where $n(t)$ is uncorrelated with $\theta(t)$ and $\tau(t)$ white noise with characteristics $m\{n(t)\}=0$.

4. Conclusions

The proposed device and algorithm of optimal FHSS receiving allow independent and accurate control on the phases in the transmitter and the receiver. The possibility of direct assessment of the random delay provides to obtain an algorithm including a wide range of problems. The device and the algorithm are protected by an application for a patent [3].

References

1. Patent 4023103 (USA).
2. Patent 4066964 (USA).
3. Application for a patent No 108437/12.12.03. Coherent device for synchronizing FHSS

КОХЕРЕНТНО ФОРМИРАНЕ И ПРИЕМАНЕ НА СИГНАЛИ СЪС СКОКООБРАЗНО ИЗМЕНЕНИЕ НА НОСЕЩАТА ЧЕСТОТА В АЕРОКОСМИЧЕСКИ РАДИОЛИНИИ

А. Андонов, Г. Чернева, З. Хубенова

Резюме

Създаването на кохерентни радиолинии е един от пътищата за повишаване на шумоустойчивостта и достигане до условия, реализиращи оптималното приемане на сигнали. В настоящата статия се разглежда метод и устройство за посочената цел, защитени чрез заявка за патент. Той се базира на досега известните два такива американски патента на фирмата Motorola, използвани при реализацията на кохерентна връзка във ВВС на САЩ. Предложеният подход дава възможност за по-проста и ефективна техническа реализация за изграждане на кохерентни радиолинии в условия на бързи флуктуации на сигнала. Това позволява неговото използване както в авиационните, така и в космическите радиолинии, при които е необходимо да се предават по-големи по обем данни с телеметрична информация, тъй като кохерентността осигурява значително по-голямо усилване при обработка на приеманите сигнали в сравнение с досега използваните. Предложен и обоснован е и алгоритъм за обработка на сигнала, основаващ се на теорията за нелинейна филтрация.

Bulgarian Academy of Sciences. Space Research Institute.
Aerospace Research in Bulgaria. 22, 2008, Sofia

ELECTROENCEPHALOGRAM ELECTROOCULOGRAPHIC ARTEFACTS POWER SPECTRUM ANALYSIS

Plamen Manoilov

A. Kanchev University of Rousse, Faculty of EEA, Department of KTT
e-mail: pmanoilov@ecs.ru.acad.bg

Abstract

Since noninvasive electroencephalogram (EEG) was discovered, started its wide use for medical diagnostics. Soon after that, attempts for using the EEG for devices' control were made. In both cases, electrooculographic (EOG) artefacts, especially subjects eye-blinks, noticeably contaminate EEG power spectrum and impede the analysis and classification of the signal. Artefacts are noises introduced to the EEG signal by not central nervous system (CNS) sources of electric fields inside and outside subjects body. In this paper, an analysis of the power spectrum of eye-blinking artefacts is described with a connection of using EEG for pattern recognition during mental tasks performance based brain-computer interface (BCI), working with α- and μ-rhythms (8-13 Hz) brain potentials. The goal of the study is to determine the influence of eye-blinks' power spectrum on EEG and choose a method for their handling.

Introduction

The ability to communicate is a typical and determinated characteristic of human beings and plays a vital role in their relationship. This communication is richer and more sophisticated than any other form of communication. Verbal and written messages are usually sent by the mouth or hands and received by the eyes or ears with the mediation of the brain. While the communication between humans has been extensively developed and studied, the communication between humans and machines is in its initial phase. The progress of neurology and computer science give the possibility to establish an immediate connection between human brain and the computer - Brain computer interface (BCI).

BCI will raise the quality of life of disabled people, offering a new communication channel. Brain control could be used by healthy people too, as additional possibility for control.

The Alternative Control Technology (ACT) program of the US Air Force Research Laboratory, Dayton, Ohio, among the variety of hands-free controls, using input from eyes, head, speech and electromyographic (EMG) includes the use of electroencephalographic (EEG) systems that allow communication with computers while the pilots' hands remain engaged in other activities [1].

Most of BCI studies use EEG, recorded from the human scalp. EEG is a noninvasive and easy to perform method, which does not require expensive and heavy equipment [2].

During EEG recording the subject moves and glances about, as it is expected of anyone asked to sit in a chair for a long time and engage in repetitive tasks, requiring mental effort. The movements introduce periods of electrical noise (artefacts) that are difficult to discriminate from neural activity.

Frontal muscles EMG can dominate the β- or μ-brain rhythms frequency range at frontal locations. Eye-blinks (known as electrooculographic (EOG) artefacts) can affect the θ- or even μ-rhythm range at frontal and central scalp locations [3, 4]. Thus, as it is possible the user to control output device by raising his eyebrows or blinking his eyes the mentioned EMG activity might obscure the user's actual EEG control.

Artefacts can dramatically alter the EEG recorded at the scalp [5], to bring to false results and conclusions during the investigation of EEG-based BCI. Studies, pretending neuroprosthesis control [6], show this risk. Later study [7] proves frontal muscles EMG influence over the control.

There are, however, other sources [**8**, 9], where the authors pretend increased data transfer rate and specially state that the investigations are done without any artefact removal, "...No (!) trials were rejected due to artifacts" [**8**].

As a result of this conflicting information, a decision was taken to study the subjects eye-blinks power spectrum in the context of particular use of EEG and after that choose a method for their handling.

In this paper an investigation of eye-blinks' influence over the EEG intended for *pattern recognition technique* based BCI, using α- and μ-rhythms in 8-13 Hz range is described. BCI uses the patterns issued during various tasks, evoked increasing brain activity, like visual presentation of four images, imaginary image rotation, imaginary human body limbs rotation, mathematical tasks, hyperventilation etc.

EOG artefacts analysis

In the study, an EEG database, recorded in the Technical University of Delft, the Netherlands, where in 2004 started a project on building an EEG-based BCI, is used.

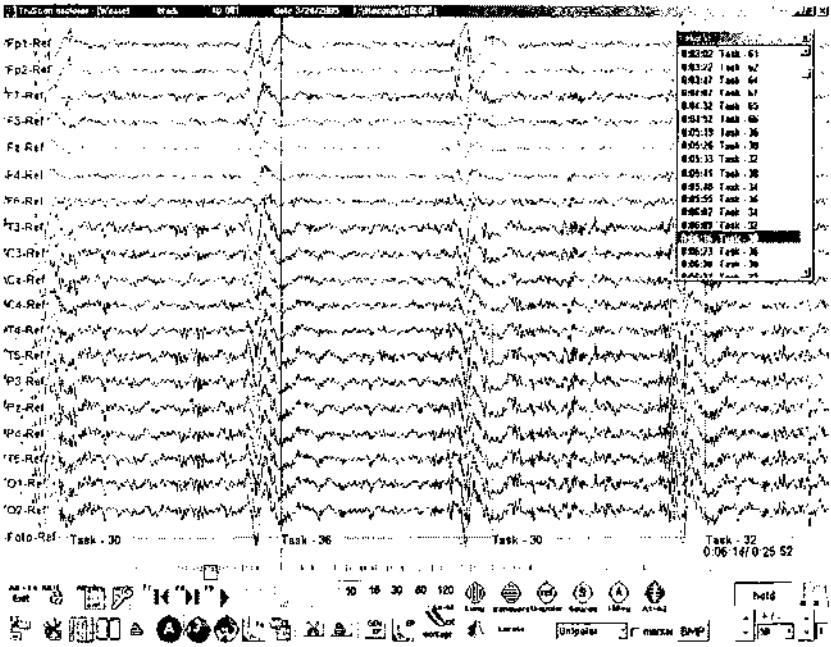


Fig. 1. Tasks'succession. The vertical lines mark every task's start. Eye-blinks stand at every tasks end out by their five times higher amplitude

To find out EOG artefacts influence over the working range power spectrum, an analysis of the planned blinks at every tasks end during the *visual presentation* tasks - having numbers 30, 32, 34, 36 - including *visual presentation* on the monitor screen respectively of *yellow triangle, green dot, red cross* and *blue lines* in one session, Fig. 1, is done. Every task runs 5 times in one session. The tasks follow each other in a pseudorandom order, to avoid the subjects learning. According to the experiment schedule, the first 5-seconds EEG are free from blinks. The next 3-seconds contain a planned blink. A model for the analysis is synthesized as (1), (2):

$$(1) \qquad P_{av}^{b}(k) = \frac{1}{RN} \sum_{m=1}^{R} \left[X_R^b(k) X_R^{b*}(k) \right]$$

(2) $$P_{av}^{nb}(k) = \frac{1}{RN}\sum_{m=1}^{R}\left[X_R^{nb}(k)X_R^{nb*}(k)\right]$$, where

$P_{av}^{b}(k)$, $P_{av}^{nb}(k)$ are the average powers for k frequency respectively for segments with blinks and without blinks.

f_s is the sampling rate, 256 Hz.

N is the number of input discretes for the analysis. The ratio N / f_s defines the frequency resolution.

X^b, X^{nb} are arrays which contain the frequency components, calculated after the Fourier analysis of the input segment which contains, X^b, and which does not contain a blink, X^{nb}. A N-point Hamming window $H(N)$, according to (3) is applied in advance:

(3) $$X(k) = \sum_{n=0}^{N-1} x(n)e^{-i2\pi kn/N}H(n)$$, where

X^{b*}, X^{nb*} are the corresponding complex conjugate arrays.

As a result of a preliminary study [10], the length of the eye-blinks is determinated as approximately 3 s (they depend on the particular subject and his psycho-physiological condition). The spontaneous and planned blinks' study [11] stated, that they have equal duration. The result allows to study the planned blinks during tasks' 3X performance. They are in predefined time segments and could be processed automatically. Averaging the power spectrum of blinks during different mental tasks' performance is possible, as a base for the comparison parts of EEG without blinks, recorded just before the blinks during the same tasks' performance are taken. A result of other preliminary study of mental tasks characteristics [14], states that tasks 3X have similar patterns.

Every blink is selected in three-second interval (768 discretes), enveloped parts from EEG before and after the visible maximum, Fig. 2. Other three-second segments just before the blinks are selected from the same tasks. The white noise level (neighbor neurons' activity [12]) in the averaged power spectrum is $\sqrt{20}$ times lower [13] in comparison with the white noise of a single segment.

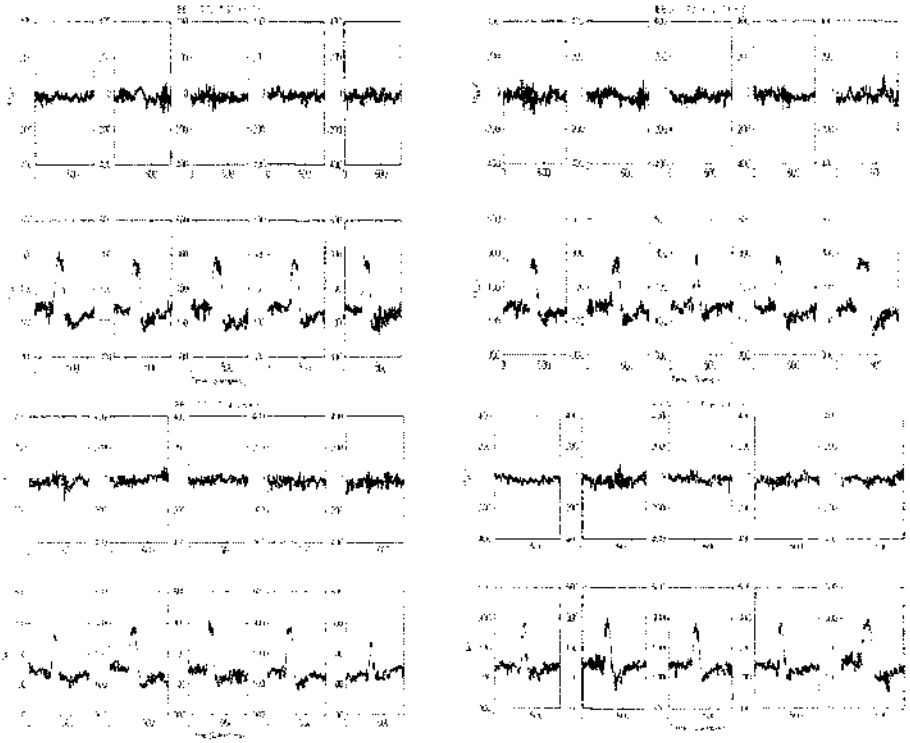


Fig. 2. EEG segments from one session selected for the analysis

Every blink is selected in three-second interval (768 discretes), enveloped parts from EEG before and after the visible maximum, Fig. 2. Other three-second segments just before the blinks are selected from the same tasks. The white noise level (neighbor neurons' activity [12]) in the averaged power spectrum is $\sqrt{20}$ times lower [13] in comparison with the white noise of a single segment.

Graphs of the averaged power spectra envelopes of the segments with and without blinks in all EEG channels are shown in Fig. 3.

The EEG amplitude without blinks is different for every channel. It is the lowest in frontal placed Fp1 and Fp2, Fig. 1. Channels Fp1 and Fp2 are the most suitable to discover eye-blinks by controlling the EEG amplitude in the time domain. When the threshold is properly set, errors probability could be minimal.

The blinks' power spectrum is concentrated in the 2-3 Hz frequency band. The blinks could be discovered by controlling the powers of 2-3 Hz frequency components after the Fourier analysis is done.

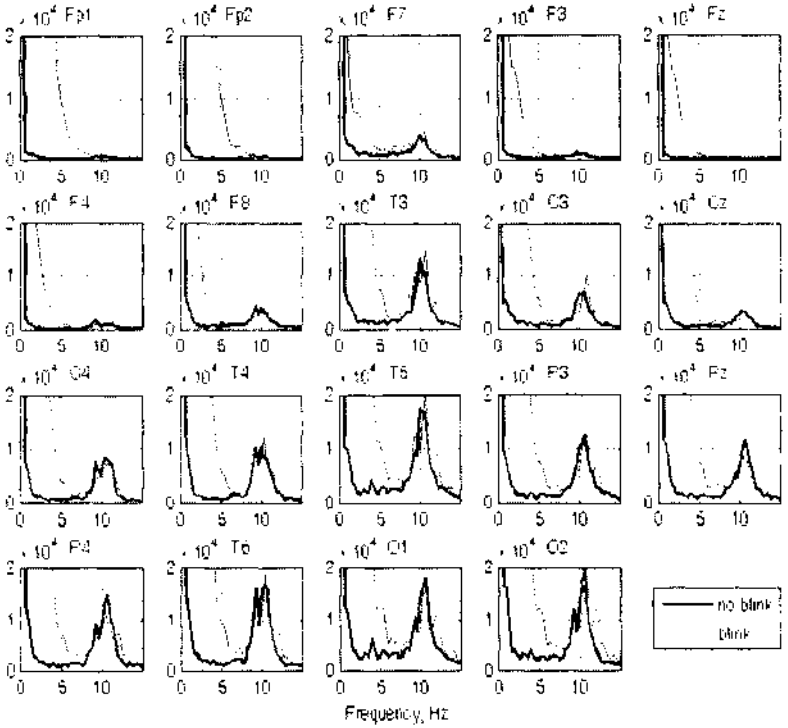


Fig. 3. Averaged power spectra in all channels

From the averaged power spectra man can see that the blinks influence is different in particular channels. In the segments with blinks, in frontal placed electrodes Fp1, Fp2, F7, F3, F4 and F8, averaged powers of working range frequency components dominate over the averaged powers of the segments without blinks. This obviously is a result of the eye-lids activation EMG potentials. In the rest of the channels C3, C4, P3, P4 etc there are frequencies with lower power in segments with blinks. Unlike sources, where EOG artefacts are defined as "no CNS artefacts", the author's suggestion is that the work range frequencies power reduction in segments with blinks is a result of brain activity. Consequently, it is wrong to define eye-blinks' influence in every channel as a result only of EMG artefacts' diffusion along and/or across the scalp accordingly to particular rules [15].

Besides low frequencies, powers of segments with blinks differ from blinks-clean segments more than 50% for frequencies in the range 8-13 Hz. This difference is commensurable and in some channels higher than the expected power changes as a result of mental tasks performance. Blinks presence definitely lowers the the probability for right mental tasks classification. Consequently, it is inadmissible to analyze segments that contain blinks without any preprocessing.

To have blinks free EEG segments for the study, a decision is taken, to reject the segments which contain subjects eye-blinks. Unlike EEGs used in the medical practice, where a short lasting part could contain very important information (for epilepsy diagnosis) and it is not allowed to loose it, mental tasks last a long time.The database is big enough and rejecting parts with eye-blinks will not lower the matter of the study.

Conclusions and future work

EOG artefacts identification, determination of their influence on the working range and its elimination from real data are necessary steps in EEG processing. They are inevitable when the data are intended to train the BCI classificator.

The power spectrum of the subject's eye blinks is concentrated in the range between 0,5-3 Hz. The power of the low frequency components is many times higher than the power of the EEG without blinks.

In the range 8-13 Hz, in part of the electrodes (T5, T6, O1, O2 etc), the power introduced by the blinks, is in times higher than the energy of the signal without blinks and much higher than the power difference, that is a result of the mental task's performance (which is 10-20% [14]). This correlation can bring considerable errors during the classification. In the study of BCI, working with brain rhythms in the range 8-13 Hz, eye-blinking artefacts should be eliminated.

The results of the study will be used to prepare the input vector for training the BCI classificator.

References

1. Middendorf M. et al. Brain–Computer Interfaces Based on the Steady-State Visual-Evoked Response, IEEE Transactions on Rehabilitation Engineering, Vol. 8, No. 2, 2000, pp. 211-214.
2. Wolpaw J. R. Brain-computer interface technology: A Review of the First International Meeting, IEEE Transactions on Rehabilitation Engineering, vol. 8, No. 2, 2000. pp. 164-173.

3. Goncharova I. I. et al. EEG-Based Brain–Computer Interface (BCI) Communication: Scalp Topography of EMG Contamination, Society of Neuroscience, vol. 26, 2000, p. vol. 8, No. 2, 2000, pp. 164–173.
4. McFarland D. J., A.T. Lefkowicz, J. R. Wolpaw. Design and Operation of an EEG-Based Brain–Computer Interface (BCI) with Digital Signal Processing Technology, Behavioral Research Methods and Instrumental Computing, vol. 29, 1997, pp. 337–345.
5. Allison B.Z., P3 or not P3: Toward a Better P300 BCI, Ph.D. Dissertation, University of California, San Diego, 324 pages, http://www.cis.gsu.edu/brainlab/PapersOtherWritings.htm.
6. Lauer R.T., P.H. Peckham, K.L. Kilgore. EEG-Based Control of a Hand Grasp Neuroprothesis, Neuro-Report, vol. 10, 1999, pp. 1767–1771.
7. Lauer R.T. et al. Applications of Cortical Signals to Neuroprosthetic Control: A Critical Review, IEEE Transactions on Rehabilitation Engineering, vol. 8, 2000, pp. 205 -208.
8. Blankertz B. et al. Single trial detection of EEG error potentials: A tool for increasing BCI transmission rates, *Artificial Neural Networks - ICANN 2002*, pp. 1137-1143, http://ida.first.fhg.de/publications/BlaSchDorCur02.pdf.
9. Schalk G. et al. EEG-Based Communication: Presence of an Error Potential, Clinical Neurophysiology, vol. 111, 2000, pp. 2138-2144.
10. Manoilov P.K. Electroencephalogram Electrooculographic Artefacts Analysis, Proceedings of the national conference with an international participation "ELECTRONICS 2006", Sofia, June 1-2 2006, pp. 156-161.
11. Manoilov P.K. et al. EEG Eye-Blinking Artefacts Duration Analysis, Proceedings of the Workshop "Services and Software Architectures, Infrastructures and Engineering for enchancing EU Citizen's Quality of Life", Bucharest, Romania, May 18 2007, pp. 140-147.
12. Malmivuo J., R. Plonsey. Bioelectromagnetism, Principles and Applications of Bioelectrical and Biomagnetic Fields, New York, Oxford, Oxford University Press, 1995, 482 pages, http://www.tut.fi/~malmivuo/bem/bembook/bembook.zip
13. Boyanov B.G. Digital Signal Processing, part I, Briagprint, Varna, 2003, 162 pages, (in Bulgarian).
14. Manoilov P.K. Classifying User Activities Based on EEG Analysis, Anniversary scientific session '2006, NMU "V. Levski", Dolna Mitropolia, 2006, pp. 180-187.
15. Boeijinga P.H., F.H. Lopes da Silva. A New Method to Estimate Time Delays Between EEG Signals Applied to Beta Activity of the Olfactory Cortical Areas, Electroencephalography and Clinical Neurophysiology, vol. 73, Sepòåmber, 1989, pp. 198-205.

АНАЛИЗ НА ЧЕСТОТНИЯ СПЕКТЪР НА ЕЛЕКТРООКУЛОГРАФСКИ АРТЕФАКТИ В ЕЛЕКТРОЕНЦЕФАЛОГРАМИ

Пл. Маноилов

Резюме

С откриването на неинвазивната електроенцефалограма (ЕЕГ), започва нейното широко прилагане в медицинската диагностика. Скоро след това се извършват експерименти по прилагане на ЕЕГ за управление на устройства. И в двата случая електроокулографските (ЕОГ) артефакти, по-точно премигванията на очите на субекта, забележимо повлияват енергийния спектър на ЕЕГ, с което затрудняват анализа и класифицирането на сигнала. Артефактите са шумове, внесени в ЕЕГ от непринадлежащи към централната нервна система (ЦНС) на субекта източници на електрически полета в и извън човешкото тяло. В статията се описва анализ на енергийния спектър на премигвания, във връзка с прилагане на ЕЕГ в мозък-компютър интерфейс (МКИ), базиран на метода разпознаване на образци, получени при изпълнение на мисловни задачи, работещ с α- и m-мозъчни ритми в диапазона 8-13 Hz. Изследването има за цел да определи влиянието на енергийния спектър на премигванията върху ЕЕГ, за да се избере подходящ метод за тяхната обработка.

Bulgarian Academy of Sciences. Space Research Institute.
Aerospace Research in Bulgaria. 22, 2008, Sofia

ON SOME RESULTS FROM THE OBSERVBATIONS OF THE TOTAL SOLAR ECLIPSES OF 11/08/1999 AND 29/03/2006

Milen Zamfirov, Petar Getsov

Space Research Institute - Bulgarian Academy of Sciences
e-mail: mzamfirov@space.bas.bg; director@space.bas.bg

Abstract

The paper reviews some results from the observations of the last two solar eclipses that took place on 11/08/1999 and 29/03/2006. The methods and means used during the photo sessions have been described. Pictures from different stages have been attached, as follows: partial eclipse, total eclipse, solar corona, Baily's Beads, crescents.

Introduction

The phenomenon of the solar eclipse can be observed when the lunar disk is projected over the visible part of the solar disk, and the lunar shadow moves over the Earth's surface.

The article reviews some of the results from the observations of two solar eclipses from 11/08/1999 and 29/03/2006. The methods and means used during the photo session are described.

The solar eclipses are very important to science. During a solar eclipse, the solar corona, the solar protuberances and the Baily's Beads can be observed. Solar eclipses also have great biological, geophysical, and meteorological influence on the Earth and its inhabitants [1], and they are one of the most attractive phenomena we can observe.

Tasks and objectives

During solar eclipses, photographic observations are the main research method for these phenomena. Our main goal was to take pictures of protuberances on the Sun's limb, solar corona and Baily's Beads; to observe "running shadows" and to take amateur pictures of the Sun with other objects from the landscape.

Solar eclipse – 11/08/1999.

The dark stripe spread over Northeast Bulgaria on 11th August 1999 for comparatively short time, and in its full phase it lasted approximately 2 hours. The solar eclipse of 1999 was comparatively short – 2 hrs and 27 sec. We made the observation from the Town of General Toshevo, district of Dobrich. The main shadow stripe entered Bulgaria at 2:09 pm local time near the Town of Silistra and left the country nearby Cape Shabla at 2:12.3 pm (Fig.1).

Solar eclipse – 29/03/2006.

The total solar eclipse was observed from the eastern most areas of South America, Central Africa, Middle Asia and South Siberia [2]. We observed the eclipse from the Town of Serik, South Turkey. In this latitude, the eclipse started at 12:38.11 pm and finished on 3:13.19 pm (all phases).

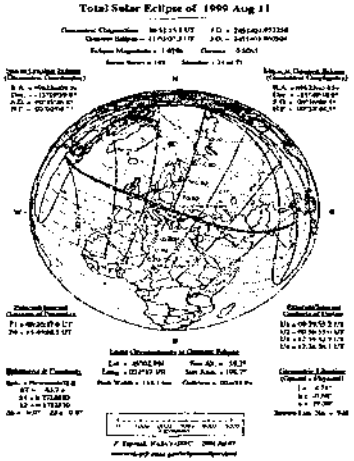


Fig. 1. Line of totality (11/08/1999) [3]

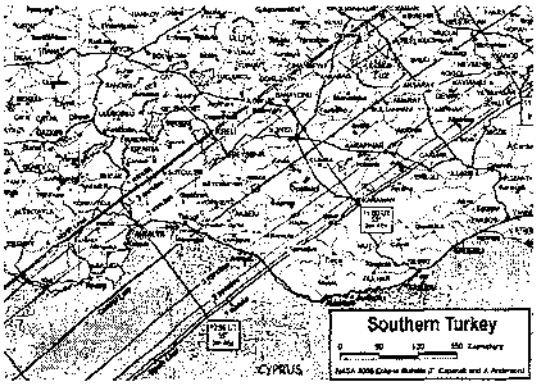


Fig. 2. Line of totality, crossing South Turkey (29/032006) [4]

Means and Methods

Telescope. Model: Konuspace 500, Telescope Newton, Reflector, F= 500 mm, D= 114 mm, f/ 4.3.

Filter. The filter used for the eclipses is Mylar, neutral type. In contrast to saturated SFO-filters, Mylar filters are more transparent and the Sun's color on the picture is not yellow-orange, but bluish [5].

Photographic camera, film and exposure time

36 pictures were shot with *Praktika* photo camera. The camera was attached to the telescope by a converter and installed in the telescope's focal plane. Various speeds were required, so that we could implement different tasks. We used standard Kodak 200 ASA color film, because the Sun is a very bright object and, despite the filters, the light that reached the negative, was sufficient.

The solar eclipses were shot at different speeds in order to receive several pictures or at least one good shot. This was necessary, because during the different phases of the solar eclipse, the solar crescent changes its size continuously.

Table 1. Speeds used to shoot the phases of the solar eclipses

Speed	Phase
1/500	First contact
1/250	1/2 of the Sun
1/125	1/4 of the Sun
1/60	1/10 of the Sun
1/30	Before totality
1/1000	Totality phase
1/500	Totality phase
1/250	Totality phase
1/125	Totality phase
1/60	Totality phase
1/30	Totality phase

Results

Gradually, the dark side increased and the Sun started to look more and more like a thinning crescent. The day light weakened. The temperature decreased by 6 °C.

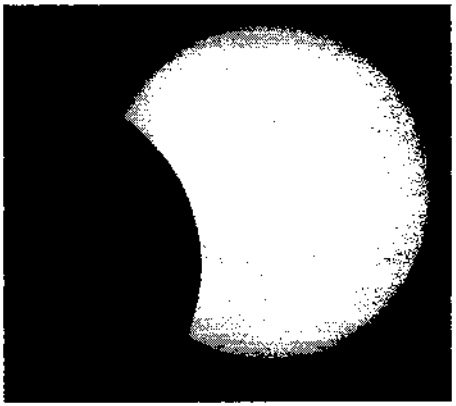


Fig. 3. Thinning solar crescent (11/08/1999).

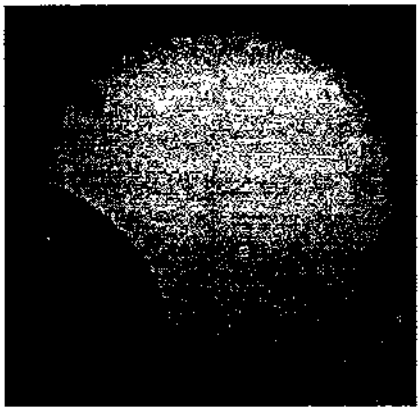


Fig. 4. Thinning solar crescent (29/03/2006).

Unlike the solar eclipse of 11/08/1999, when the Sun was in the maximum of its activity, now it was in its minimum and it was difficult to extract from the pictures areas where protuberances could be observed.

Solar corona. The solar corona is almost circular, due to the solar minimum, and is silver-white. When the Sun is in its maximum, the corona is greatly stretched-out along the solar equator and a separate ray can be viewed.

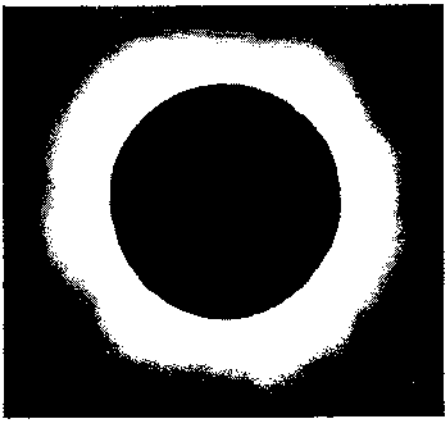


Fig. 5. The solar corona (11/08/1999).

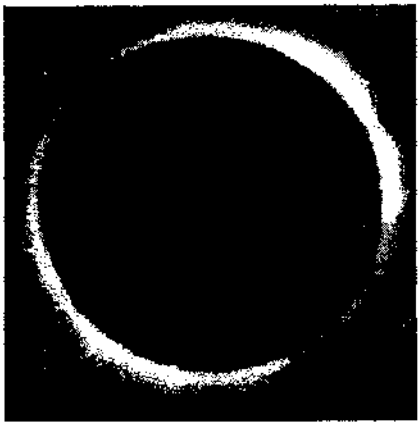


Fig. 6. The solar corona (29/03/2006)

Running shadows. Running shadows can be observed right before and immediately after the totality phase. They are considered to be an effect of the turbulent flows of the solidity in the Earth's atmosphere, which leads to focusing and unfocusing of the sun rays. It is extremely difficult to shoot the rays, but it was very easy to observe them.

Baily's Beads. They occurred before and after the totality phase. They appear because of the craters on the lunar limb.

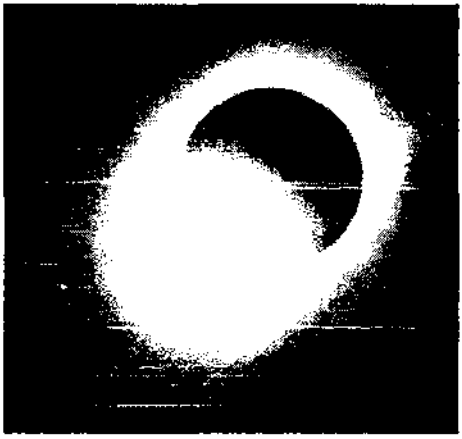


Fig. 7. Baily's Beads (11/08/1999)

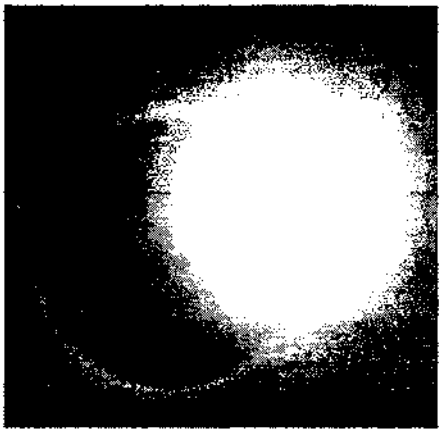


Fig. 8. Baily's Beads (29/03/2006)

Shadow of the earth objects. Since the luminous crescent gradually decreases, the shadows become sharp (without penumbras). The objects parallel to the solar crescent have extremely sharp shadows, while in other spatial dispositions, the shadow is asymmetric. This is due to the smaller angular sizes of the revealed parts of the Sun.

Fig. 9. Shadow of an Earth object (29/03/2006)

Crescents. Thousands of small crescents can be observed due to the small holes along and between the sheets, used as camera obscura.

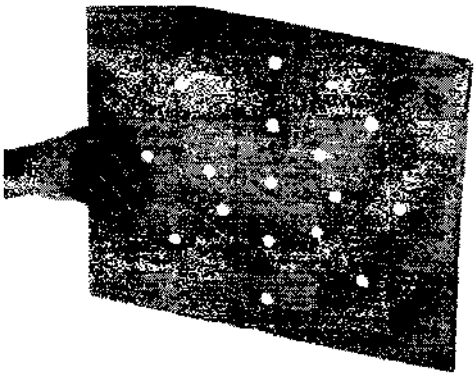


Fig. 10. A board with circular holes, allowing the sunlight during the partial phase (29/03/2006).

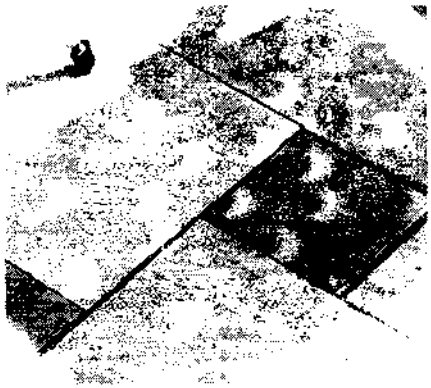


Fig. 11. The crescents observed from the board with circular holes, letting the sunlight (29/03/2006).

Celestial bodies. During the totality of the solar eclipse, the observed celestial bodies were Venus and Mercury.

Results

The pictures of the solar eclipses of 11/08/1999 and 29/03/2006 show that, due to the solar minimum, the solar corona from 29/03/2006 is almost circular and silver-white (Fig. 6). On 11/08/1999, the Sun was in its maximum and the solar corona was greatly stretched-out along the solar equator, so separate rays can be seen (Fig. 5).

Conclusion

We made 72 pictures with exposure time varying from 1/30 to 1/1000 seconds, shooting the phases of the two solar eclipses, the Baily's Beads and crescents, the running shadows and the planets visible during the totality phase of the solar eclipse and lots of amateur pictures.

References

1. bg.wikipedia.org/wiki/
2. bgastronomy.com/index.asp?ID=362&Cat=16
3. http://sunearth.gsfc.nasa.gov/eclipse/SEplot/SEplot1951/SE1999Aug11T.GIF
4. sunearth.gsfc.nasa.gov/eclipse/SEmono/TSE2006/TSE2006fig/TSE2006-fig13b.GIF
5. http://solar-eclipses.hit.bg/Urok.html

ОТНОСНО НЯКОИ РЕЗУЛТАТИ ОТ ПЪЛНИТЕ СЛЪНЧЕВИ ЗАТЪМНЕНИЯ НА 11/08/1999 И 29/03/2006

М. Замфиров, П. Гецов

Резюме

В статията се разглеждат някои резултати, свързани с двете слънчеви затъмнения на 11.08.1999 г. и 29.03.2006 г. Описани са методите и средствата при провеждането на снимките. Показани са снимки на частичната фаза на затъмнението, пълната фаза, слънчевата корона, броеница на Бейли, сърпчета.

A NEW BOOK ABOUT NATURAL HAZARDS

Millions of victims, much more injured and homeless, great amount of losses and heavy destructions – these could be just a part of the negative consequences due to the different natural hazards all over the world. Even today – at the beginning of 21 century, humanity still pays the great price to the terrible effects of natural hazards. Almost everywhere in the world it is proved that the victims and damages strongly depend on of the quality of constructions and the ability of the population to mitigate disasters. The effectiveness of emergency planning and execution, the fast and well organized reaction to the negative effects of the disaster are among the important factors influencing the mitigation of the different hazards. This is a great topic and now – during the EARTH'S YEAR – 2008 – everywhere in the world a lot of resources, educational programs and active measures are under taken to mitigate the effects of the different natural hazards – earthquakes, tsunami, global climate change and all other catastrophic events on the earth's surface. Even 4 years after the devastating tsunami in the Indian ocean (The Boxing day event) which took more than 300, 000 victims and left more than 1, 000, 000 homeless, people are not able to fully recognize what happened on this holiday. And when it will be repeated?...

It is much clearer now that humanity will experience an increasing number of disasters in the future due to the dynamic earth. The global climate change, the ozone hole development, the El Nino phenomenon and the fast geodynamic events, registered by the GPS in the solid earth, as well as the increased urbanization are the main factors which can augment the negative consequences of Nature for the everyday life of populations and societies.

Therefore, the new by published book *Natural Hazards and Ecological Catastrophes – Study, Prevention, Protection* is extremely useful and topical.

The author – Prof. Garo Mardirossian from the Space Research Institute of the Bulgarian Academy of Sciences needs no a special presentation. Well known for his popular scientific books, he is working hard on the problems of natural

hazards and their study by aerospace methods and techniques. This is his 4th book on the topic. All of them were accepted by the audience with great interest.

In this last book, at the beginning, some general problems related with natural hazards, their generation and development, are highlighted and the many physical properties underlying their ability to produce negative effects. Many terms like risk, vulnerability, multi risk and preliminary assessment of the consequences are among the initial themes developed in the book. Separately, a lot of space is deducated to ecological catastrophes, war consequences and geophysical weapons. Step by step, the study goes deeper discovering the nature and effects of such events as earthquakes, landslides, cyclones, volcanic eruptions, tsunami, avalanches, floods, forest fires, thunderstorms, tornadoes, etc. A lot of data about the observed cases, statistics about the victims and the damages, the most extreme cases etc. are under the focus of the author. Special attention is paid to the natural hazard on the territory of Bulgaria. The main parameters generating the negative effects to the population and the infrastructure are discussed. A separate chapter is dedicated to each individual disaster type. Special attention is paid to prevention and protection, as well as to individual and societal directions. The presentation of the potential of aerospace technologies for such investigations and observations and their high effectiveness is very useful. Thus, supporting the idea that such methodologies could be a very effective and useful tool against natural hazards, the book provides a lot of improvements in that direction.

The book is presented well to the readers and illustrated with many diagrams, figures and photos. It could be useful to many different specialists – scientists and researchers in the fields of geophysics and ecology, geography and meteorology, etc. as well as to the decision-makers and civil defence servants, engineers and land planners, as well as the students and pupils from schools and universities. The book contains 372 pages, 56 figures, 20 tables, etc. Useful annexes contain the measurement scales about the different natural hazards.

Due to its actual content and usefulness and because such books are rare not only for us but also in the world practice, the appearance of the book of Prof. Mardirossian is an important event for the scientific community.

Assoc. Prof. Boyko Ranguelov

Мардиросян, Г. „Природни бедствия и екокатастрофи – изучаване, превенция, защита". Академично издателство „Проф. Марин Дринов", София, 2007, 372 стр.

Technology Transfer Office

The Technology Transfer Office in the Space Research Institute at the Bulgarian Academy of Sciences was established in November 2007 as a department within the institute according to the project of the PHARE program: 'Establishing of Technology Transfer Offices at the Bulgarian Public Research Organizations'.

Fields of action

- Transfer of technologies, innovations and know-how
- Searching for resolving of concrete production problems
- Evaluation of technologies – technical, economical, financial and ecological
- Consultancy services related to technology transfer
- Preparing, consulting and coordination of projects
- International cooperation, organization of training and scientific events – workshops, seminars, presentations etc.

Research Areas

- Remote Sensing
- Telemedicine
- Space Materials Knowledge and Nanotechnologies
- Localisation of Moving Objects
- Defense and Security, with respect to their potential to be introduced for commercialisation in the economy

Target groups of the TTO's activities

Government and non-government organizations and institutions at national, regional and local level, and different sectors of the national economy – industry, agriculture, transport etc.

Managing Board: Chairman – Peter Getsov
Advisory board: Chairman – Gergin Gerginov

Project coordinator: Ognian Petrov
Director of the office: Christina Kovacheva

Main Partner
Institute of Geography at the Bulgarian Academy of Sciences

Associated partners
- New Bulgarian University - Center for Continuing Education;
- Bulgarian Astronautical Federation;
- KOSMOS NPP;
- Bulgarian Association of Aviation and Space Medicine and Psychology;
- GIS-Transfer Center;

For contacts
Tel/Fax: +359 2 873 67 33
E-mail: tto_sri@space.bas.bg
www.tto-sri.com

ПРОГРАМА ФАР 2005 - Проект BG 2005/017-353.10.06
Създаване на центрове за технологичен трансфер в българските висши училища и държавни научноизследователски организации - Грантова схема Номер на публикацията: BG 2005/ESC/G/TTO

PHARE 2005 Programme – Project BG 2005/017-353.10.06
Establishment of Technology Transfer Offices at Bulgarian Public Research Organisations Grant Scheme
Reference Number: BG 2005/ESC/G/TTO

The project is implemented with financial support of the European Union and Republic of Bulgaria
Проектът се изпълнява с финансовата подкрепа на Европейския съюз и Република България