AEROSPACE RESEARCH IN BULGARIA

Volume 24, Sofia, 2012
Space Research and Technology Institute
Bulgarian Academy of Sciences

Address
AEROSPACE RESEARCH IN BULGARIA
Space Research and Technology Institute
bl. 1, *Acad. G. Bonchev* St., Sofia 1113, Bulgaria

e-mail: journal@space.bas.bg

Pre-Publication Processing
Tsveta Srebrova

ISSN 1313 – 0927

Bulgarian Academy of Sciences **ISSN 1313 – 0927**

Aerospace Research in Bulgaria

24

Sofia, 2012

C o n t e n t s

С Ъ Д Ъ Р Ж А Н И Е

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 24, 2012, Sofia

THIN VISCOUS ELLIPTICAL ACCRETION DISCS WITH ORBITS SHARING A COMMON LONGITUDE OF PERIASTRON. V. LINEAR RELATIONS BETWEEN AZIMUTHAL-ANGLE AVERAGED FACTORS IN THE DYNAMICAL EQUATION

Dimitar Dimitrov

Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: dim@mail.space.bas.bg

Abstract

*We consider a model of elliptical **stationary** accretion discs developed by Lyubarskij et al. [4], which have derived a second order ordinary differential equation, describing the spatial structure of these objects. This dynamical equation contains seven integrals, arising from the azimuthal averaging along the elliptical disc particle orbits. They are functions on the **unknown** eccentricity distribution **e(u)**, its derivative **ė(u)** ≡ **de(u)/du** and the power **n** in the viscosity low **η = β Σ ^{n}**, where **u** ≡ **ln p**, **p** is the focal parameter of the concrete elliptical particle orbit. In the present paper, we derive linear relations between these **unknown** integrals, which may be useful to eliminate **three** of these quantities. It is also possible to eliminate even one more integral, but proving of this statement will be postponed in a forthcoming paper. The considered approach is maintained with a view to split the dynamical equation into a system of more simple differential equations.*

1. Introduction

The accretion phenomena have many impacts on the structure and evolution of large variety of astrophysical objects. Such processes may include both spherical accretion and/or accretion via discs. In the later case, the disc accretion mechanism is caused by the large angular momentum of the material, surrounding the compact body, and falling onto it as a final

result. In the present investigation we shall concentrate our attention over accreting compact objects having stellar masses. As a general rule, the matter, composing the accretion flow, is supplied by another star (the so-called donor star), orbiting around the accreting compact component of the binary stellar system. In some cases, the material of the disc may be available due to a disruption by the tidal forces of a very close orbiting body. But in spite of this possible (in some sense, more "exotic") situation, the mass of the accretion disc will disappear very soon because of the exhausting processes. These may be accumulation of mass over the surface of central star, jets and winds from the two surfaces of the disc (like the stellar winds in ordinary stars) or any other outflows removing the matter from the vicinity of the disc. Consequently, we would expect that such accretion discs may be treated as ***stationary*** objects for a time scales shorter than the corresponding time intervals for the discs existing in the close binary systems.

It is well known that the balance between the heating and cooling processes strongly determines the spatial structure and the time evolution of the accretion flows. A great variety of accretion disc models illustrates that the motions of the disc particles may essentially differ from the *Keplerian* one. This circumstance is able to change the flow so considerably, that in some parts of the disc the radial motion of matter *is not inward* (accretion), but is directed *outward* (excretion). This is the case for hot, advection-dominated accretion flows, which are usually optically thin in the radial direction. Therefore, the photons, produced at given radii, can travel long distances without being absorbed. Compton scattering of these photons heats or cools electrons at other radii of the considered accretion disc model. It may turn out to be, that at a certain radius, the Compton cooling rate is larger than the ***local*** viscous heating rate, i.e. the cooling effect is important in this situation. As pointed out by Yuan et al. [1], it is possible to obtain a self-consistent solution for the activity of an accretion disc around a black hole only when the luminosity of the disc L is less than 0,01 $L_{Eddington}$. Above this critical accretion rate, the *equilibrium* temperature of the electrons at the outer *radius* of the disc r_{out} is higher than the *virial* temperature, due to the strong Compton heating. As a result, the accretion is suppressed. Consequently, in this model, the activity of the black hole (more precisely, of its accretion flow) is expected to oscillate between an active and an inactive phase. The oscillations have time scales of the radiative time scale gas order at the outer radius r_{out} .

Another problem, associated with the description of the accretion disc structure around the black holes, is the following. The inner edges of these discs may have variable properties, if the matter inside the marginally stable orbit is magnetically connected to the disc. Then a non-zero torque is exerted on the inner disc edge and the accretion efficiency can be much higher than in the standard accretion disc model of Shakura & Sunyaev [2]. In the later case, this quantity is supposed to be equal to zero. The non-zero torque implies that, in the case of variable torque, transitions of the flow between different accretion types may be triggered [3].

In the present paper, we consider the problems related to the attempts to solve analytically the dynamical equation, governing the structure of elliptical accretion discs, rotating around a stellar mass objects. More specifically, we are dealing with the model of Lyubarskij et al. [4], which is a generalization of the work of Shakura & Sunyaev [2] to the case of elliptical accretion discs with orbits sharing a common longitude of the periastron. A very important property of the models [2] and [4] is that the trajectories of the disc particles are *Keplerian* ones. Consequently, our further conclusions cannot be applied to the above mentioned situations [1] and [3], i.e. our considerations shall avoid the cases of discs around black holes, and, especially, the disc regions too close to the central star. Such a limitation enables us also to escape the complications, related to the necessity to use general relativity for the description of disc dynamics. But these are not the only troubles, concerning the realistic treatment of the accretion flows by means of the Lyubarskij et al. model [4]. For example, angular momentum transport within young massive protoplanetary discs may be dominated by the self-gravity at the radii, where the disc is too weakly ionized to allow the development of the magnetorotational instability [5]. One important way to overcome the different problems, occurred in the theory of accretion discs, is to develop computer codes in order to perform numerical simulations of the processes in the accretion flows. Of course, such an approach may be applied for time-sequences of solutions, giving the evolution of the investigated objects. The difficulties, which arise in these searches, are very often caused by the vast volume of the needed computer capabilities. Numerical simulations of radiative processes in magnetized hot accretion discs (like these around black holes) are complicated, because the energy distributions of the particles and the photons span many orders of magnitude. The distributions may strongly depend on each other. Also, the radiative interactions behave significantly

differently, depending on the energy regime. Many complications in the computational procedures are due to the enormous difference in the time-scales of the processes [6].

There are many observational evidences that the accretion discs have complex spatial structure. Photometric and spectral studies in the near infrared region of the electromagnetic spectrum have led to the identification of a new class of accretion discs, whose members have an inner optically *thick* part, separated from an outer *also* optically *thick* part by an optically *thin* gap. This is in contrast to the discs that have inner disc holes. The authors of the paper [7] Espaillat et al. take for granted that the excess of the near infrared emission above the photosphere of the star LkCa 15 is a blackbody continuum, that can only be due to the optically *thick* material in an inner disc around the star. If this result is combined with the estimation of the radius of the *inner edge* of the outer disc, it reveals a gapped structure of the accretion disc. Espaillat et al. assume that the most likely mechanism for clearing the detected gap in the evolving disc of the star LkCa 15 is the forming of planets.

Returning to the theme of numerical simulations of the accretion flows, it is worthy to note that the *two*-dimensional hydrodynamical discs are nonlinearly unstable to the formation of vortices. Once formed, these vortices survive forever. But in *three* dimensions, numerical experiments show that *only* vortices in short boxes form and survive just as in *two* dimensions. The vortices in tall boxes are unstable and are destroyed. As pointed out by Lithwick [8], the unstable vortices decay into transient turbulent-like states, that transport angular momentum *outward* at a nearly constant rate for hundreds of orbital times. In the paper [8] was derived the criterion for the vortices to survive in *three* dimensions as they do in *two* dimensions. Namely, the azimuthal extend of the considered vortex must be larger than the local scale height of the accretion disc. When this condition is violated, the vortex is unstable and decays. Lithwick [8] concludes that a vortex with a given radial extend will survive in a *three*-dimensional disc if it is sufficiently weak (vortices are longer in azimuthal than in radial extend). The weak vortices behave *two*-dimensionally even if their width is much less than their height, because they are stabilized by rotation and behave as Taylor-Proudman columns [8]. It is also important to underline that the decaying of strong vortices might be responsible for the *outward* transport of angular momentum – a condition that is required for accretion discs to accrete. Obviously, the two-dimensional analytical model of

Lyubarskij et al. [4], in which the dynamical equation is a subject of our further considerations, does not include the vortices phenomena at all. That is why, this limitation must be kept in mind when the compatibility of this model to the real accretion discs is discussed. Nevertheless, we hope that at least some of its characteristics are realistic description of the elliptical discs in the nature and there is a reason to seek for analytical solutions of this model [4]. It must be stressed that the accretion disc theory itself contains certain unresolved problems and ambiguities. In particular, turbulent viscosity is frequently used in this theory to replace the microphysical viscosity in order to accommodate the observational need in discs, that leads to enhanced transport of energy and angular momentum. In paper [9] it is shown that the mean-field approach leads not to one, but to two transport coefficients that govern the mass and angular momentum transport. The authors of the above investigation conclude that the conventional approach suffers from an inconsistent neglect of the turbulent diffusion in the surface density equation. They constrain these two new transport coefficients for the specific cases of inward, outward and zero net mass transport. Hubbard and Blackman also find that one of the new transport terms can lead to oscillations in the mean surface density, which then requires a constant or small inverse Rossby numbers for accretion discs, to maintain a monotonic power-law density [9].

The above sketched difficulties and also many other complex problems of the accretion flows theory (cited in the references of the listed below papers), unambiguously imply that we must consider models with reasonable simplifications. What assumptions we shall made depends, of course, on the accretion disc features, which we want to describe. In a series of papers [10], [11] and [12], we have investigated ***stationary*** accretion discs with ***elliptical*** shape under the assumed viscosity law $\eta = \beta \Sigma^n$, where η is the viscosity coefficient, Σ is the surface density of the disc and β is a constant. The ***ellipticity*** is the dominant property, which is assumed to characterize all the considered cases. The power n is chosen to be a free parameter, which physically reasonable values lie in the range from about –1 to about +3. The cases when n is an *integer* are already treated in the papers [11] and [13], where the dynamical equation is expressed in an analytical form. In what follows, we shall attempt to simplify this equation for *noninteger* powers n. This division of the values of the parameter n into *integer*/*noninteger* meanings has purely mathematical origin, due to our

ability to solve analytically some integrals, entering into the dynamical equation. It has not physical foundations.

2. Dynamical equation for the elliptical accretion disc model

For brevity, we shall not write here in an explicit form the dynamical equation, governing the structure of the *stationary elliptical* discs *with orbits sharing a common longitude of periastron*. We only note that this matter is already studied and discussed in earlier papers [4], [10], [11], [12] and [13], and we refer the reader to these investigations. We shall remind only some definitions and assumptions, made in these publications, in order to be enough clear in the further exposition. We use the notations p and $u \equiv ln\ p$ for the focal parameter of each particle trajectory and its logarithm, respectively. We shall consider the power n in the viscosity law $\eta = \beta\, \Sigma^n$ as a constant parameter for each concrete considered model. This assumption means that n is the same constant throughout the disc, i.e. its derivative with respect to p (or $u \equiv ln\ p$) is equal to zero. We also assume that n may be either integer or noninteger, ranging between ≈ -1 and $\approx +3$, depending on the considered accretion disc model (but remaining as a constant in the framework of the model!). By $e \equiv e(u)$ we denote the eccentricity of the elliptical orbit of the particle, and by $\dot{e} \equiv \dot{e}(u) \equiv de/du \equiv de/dln\ p$ we understand the corresponding ordinary derivative. As it is already proved in [13], the dynamical equation, governing the structure of the accretion flow, is a second order ***homogeneous*** ordinary differential equation. Consequently, our problem is to simplify the coefficients, entering as multipliers into the two terms containing $\ddot{e}(u)$ and $\dot{e}(u)$ separately. In the paper [13] it is suggested that the procedure of the simplification may probably involve finding of linear relations between the following seven integrals $\mathbf{I_{0-}}$, $\mathbf{I_{0+}}$, $\mathbf{I_0}$, $\mathbf{I_1}$, $\mathbf{I_2}$, $\mathbf{I_3}$ and $\mathbf{I_4}$:

$$(1) \qquad \mathbf{I}_{0-}(e,\dot{e},n) \equiv \int_0^{2\pi} (1 + e\cos\varphi)^{n-3} [1 + (e - \dot{e})\cos\varphi]^{-(n+1)}\, d\varphi\ ,$$

$$(2) \qquad \mathbf{I}_{0+}(e,\dot{e},n) \equiv \int_0^{2\pi} (1 + e\cos\varphi)^{n-2} [1 + (e - \dot{e})\cos\varphi]^{-(n+2)}\, d\varphi\ ,$$

$$(3) \qquad \mathbf{I_j}(e,\dot{e},n) \equiv \int_0^{2\pi} (\cos\varphi)^{j} (1 + e\cos\varphi)^{n-2} [1 + (e - \dot{e})\cos\varphi]^{-(n+1)}\, d\varphi\ ;\ \ \mathbf{j} = 0, 1, 2, 3, 4,$$

where φ is the azimuthal angle over which the averaging is performed ([4], [10]). Using the above notations, we can write the dynamical equation of the

elliptical disc in the following form ([10], [13]):

$$(4) \qquad \sum_{i,k} \mathbf{A}_{ik}(e,\dot{e},n)\, \mathbf{I}_i(e,\dot{e},n)\, \mathbf{I}_k(e,\dot{e},n)\, \ddot{e} + \sum_{l,m} \mathbf{B}_{lm}(e,\dot{e},n)\, \mathbf{I}_l(e,\dot{e},n)\, \mathbf{I}_m(e,\dot{e},n)\, \dot{e} = 0 \,,$$

where the indices i, k, l and m independently take meanings 0-, 0+, 0, 1, 2, 3 and 4. Our base line in the present paper is to obtain ***linear*** relations between the integrals $\mathbf{I_i}(e,\dot{e},n)$, (i = 0-, 0_+, 0, 1, 2, 3, 4), which will allow us to reduce the number of these integrals in the *homogeneous* ordinary differential equation (4). As already mentioned in [13], this is another approach to perform simplifications of the considered dynamical equation (4). In the forthcoming calculations we suppose at first that, ***by hypothesis***, n, $n - 1$ and $n - 2$ are not equal to zero. Consequently, if these quantities appear as factors in the denominators of the derived intermediate and final expressions, they (by themself) cannot cause divergences of the results. After that, we include considerations of these particular cases, in order to ensure the completeness of the task solution. In the next paragraph we shall deduce expressions which will enable us to eliminate three of the above seven integrals, namely: $\mathbf{I_4}$, $\mathbf{I_2}$ and $\mathbf{I_1}$. In following papers, we shall also remove the integral $\mathbf{I_0}$ and shall discuss the linear independence of the remaining three integrals $\mathbf{I_3}$, $\mathbf{I_{0-}}$ and $\mathbf{I_{0+}}$. We stress that all the integrals are considered to be functions on $e(u)$, $\dot{e}(u)$ and n. The later quantity n has the same value for the entire area of the elliptical accretion disc, i.e. n does not depend on the focal parameter p ($u \equiv ln\ p$). Of course, for other *concrete* models n may have different (but also constant) values. As a final result, the integrals $\mathbf{I}_i(e,\dot{e},n)$ (i = 0-, 0+, 0, 1, 2, 3, 4) will depend on u, but in the following calculations we shall consider e and $\dot{e}$ as ***independent*** variables, having however in mind, that $\dot{e}(u) \equiv de(u)/du$. The later circumstance must be taken into account when a differentiation of the integrals (1) – (3) should be performed.

3. Linear relations between integrals

During the next calculations we shall use the well-known trivial relation $cos^2\varphi + sin^2\varphi = 1$ (or, equivalently: $cos^2\varphi = 1 - sin^2\varphi$), valid for all $0 \le \varphi \le 2\pi$. We also remember, that according to the original work of Lyubarskij et al. [4], we must limit our investigation to the cases when not only $|e(u)| < 1$, but also the conditions $|\dot{e}(u)| < 1$ and $|e(u) - \dot{e}(u)| < 1$ are fulfilled (see the denominators of the expressions in *Appendix A* of [4]). Such restrictions probably preserve us from the much more complicated situation, when

shock waves induced/generated by the singularities must be taken into account in the considered accretion disc model.

3.1. Elimination of the integral $\mathbf{I_4}(e,\dot{e},n)$

According to the definition (3), we have that:

$$(5)\qquad \mathbf{I_4}(e,\dot{e},n) \equiv \int_0^{2\pi} (cos\varphi)^4\,(1+ecos\varphi)^{n-2}[1+(e-\dot{e})cos\varphi]^{-(n+1)}\,d\varphi =$$

$$= e^{-1}\int_0^{2\pi} cos^3\varphi[(1+ecos\varphi)-1](1+ecos\varphi)^{n-2}[1+(e-\dot{e})cos\varphi]^{-(n+1)}\,d\varphi =$$

$$= e^{-1}\int_0^{2\pi} cos^3\varphi(1+ecos\varphi)^{n-1}[1+(e-\dot{e})cos\varphi]^{-(n+1)}\,d\varphi -$$

$$- e^{-1}\int_0^{2\pi} cos^3\varphi(1+ecos\varphi)^{n-2}[1+(e-\dot{e})cos\varphi]^{-(n+1)}\,d\varphi\ .$$

The second integral is equal to $\mathbf{I_3}(e,\dot{e},n)$ (see definition (3)). Applying the relation $cos^2\varphi = 1 - sin^2\varphi$, we obtain:

$$(6)\qquad \mathbf{I_4}(e,\dot{e},n) = -e^{-1}\mathbf{I_3}(e,\dot{e},n) + e^{-1}\int_0^{2\pi} cos\varphi(1-sin^2\varphi)(1+ecos\varphi)^{n-1}\times$$

$$\times[1+(e-\dot{e})cos\varphi]^{-(n+1)}\,d\varphi = -e^{-1}\mathbf{I_3}(e,\dot{e},n) + e^{-1}\int_0^{2\pi} cos\varphi(1+ecos\varphi)^{n-1}\times$$

$$\times[1+(e-\dot{e})cos\varphi]^{-(n+1)}\,d\varphi + [e(e-\dot{e})]^{-1}\int_0^{2\pi} cos\varphi\ sin\varphi(1+ecos\varphi)^{n-1}\times$$

$$\times[1+(e-\dot{e})cos\varphi]^{-(n+1)}\,d[1+(e-\dot{e})cos\varphi]\ .$$

The second integral in (6) can be immediately expressed through the integrals $\mathbf{I_1}(e,\dot{e},n)$ and $\mathbf{I_2}(e,\dot{e},n)$, again using the definitions (3):

$$(7)\qquad \int_0^{2\pi} cos\varphi(1+ecos\varphi)^{n-1}[1+(e-\dot{e})cos\varphi]^{-(n+1)}\,d\varphi = \mathbf{I_1}(e,\dot{e},n) + e\mathbf{I_2}(e,\dot{e},n)\ .$$

In deriving of the third summand in the relation (6), we have used that the difference $(e - \dot{e})$ does not depend on the azimuthal angle φ. Consequently:

$$(8)\qquad \mathbf{I_4}(e,\dot{e},n) = e^{-1}\mathbf{I_1}(e,\dot{e},n) + \mathbf{I_2}(e,\dot{e},n) - e^{-1}\mathbf{I_3}(e,\dot{e},n) - [ne(e-\dot{e})]^{-1}\int_0^{2\pi} cos\varphi\ sin\varphi\times$$

$$\times(1+ecos\varphi)^{n-1}\,d\{[1+(e-\dot{e})cos\varphi]^{-n}\}\ .$$

Let us consider now the forth term in the above equality. Integrating by parts, we obtain:

(9) $$-[ne(e-\dot{e})]^{-1}\int_0^{2\pi} cos\varphi\ sin\varphi(1+ecos\varphi)^{n-1}\,d\{\,[1+(e-\dot{e})cos\varphi]^{-n}\,\} = -[ne(e-\dot{e})]^{-1}\times$$

$$\times\,\{\,cos\varphi\ sin\varphi(1+ecos\varphi)^{n-1}[1+(e-\dot{e})cos\varphi]^{-n}\,\Big|_0^{2\pi} -$$

$$-\int_0^{2\pi}[1+(e-\dot{e})cos\varphi]^{-n}\,d\{\,[\,cos\varphi\ sin\varphi(1+ecos\varphi)^{n-1}\}\,\} =$$

$$= [ne(e-\dot{e})]^{-1}\{-\int_0^{2\pi} sin^2\varphi(1+ecos\varphi)^{n-1}[1+(e-\dot{e})cos\varphi]^{-n}\,d\varphi\ +$$

$$+\int_0^{2\pi} cos^2\varphi(1+ecos\varphi)^{n-1}[1+(e-\dot{e})cos\varphi]^{-n}\,d\varphi\ -$$

$$-(n-1)e\int_0^{2\pi} cos\varphi\ sin^2\varphi(1+ecos\varphi)^{n-2}[1+(e-\dot{e})cos\varphi]^{-n}\,d\varphi\,\} =$$

$$= [ne(e-\dot{e})]^{-1}\{\,2\int_0^{2\pi} cos^2\varphi(1+ecos\varphi)^{n-1}[1+(e-\dot{e})cos\varphi]^{-n}\,d\varphi\ -$$

$$-\int_0^{2\pi}(1+ecos\varphi)^{n-1}[1+(e-\dot{e})cos\varphi]^{-n}\,d\varphi -$$

$$-(n-1)e\int_0^{2\pi} cos\varphi(1+ecos\varphi)^{n-2}[1+(e-\dot{e})cos\varphi]^{-n}\,d\varphi\ +$$

$$+(n-1)e\int_0^{2\pi} cos^3\varphi(1+ecos\varphi)^{n-2}[1+(e-\dot{e})cos\varphi]^{-n}\,d\varphi\,\}\,.$$

In the above derivation we have used the relation $cos^2\varphi + sin^2\varphi = 1$ and trivially following from it simple equality:

(10) $$-sin^2\varphi + cos^2\varphi = 2\,cos^2\varphi - 1\,.$$

Like the expression (7), we shall preliminary compute several auxiliary relations, which will help us further simplify the expression (9):

(11) $$\int_0^{2\pi}(1+ecos\varphi)^{n-2}[1+(e-\dot{e})cos\varphi]^{-n}\,d\varphi\ =$$

$$=\int_0^{2\pi}[1+(e-\dot{e})cos\varphi](1+ecos\varphi)^{n-2}[1+(e-\dot{e})cos\varphi]^{-(n+1)}\,d\varphi\ =$$

$$=\mathbf{I_0}(e,\dot{e},n)+(e-\dot{e})\mathbf{I_1}(e,\dot{e},n)\,.$$

By analogy with the above computation, we multiply ***both*** the nominator and the denominator of the integrals by $[1 + (e - \dot{e})cos\varphi]$. The

assumed condition $|e(u) - \dot{e}(u)| < 1$ for all u imply that this expression is never equal to zero. By a fully similar way, we evaluate the following integrals:

$$(12)\qquad \int_0^{2\pi} cos\varphi(1 + ecos\varphi)^{n-2}[1 + (e - \dot{e})cos\varphi]^{-n}\, d\varphi = \mathbf{I_1}(e,\dot{e},n) + (e - \dot{e})\mathbf{I_2}(e,\dot{e},n)\,,$$

$$(13)\qquad \int_0^{2\pi} cos^2\varphi(1 + ecos\varphi)^{n-2}[1 + (e - \dot{e})cos\varphi]^{-n}\, d\varphi = \mathbf{I_2}(e,\dot{e},n) + (e - \dot{e})\mathbf{I_3}(e,\dot{e},n)\,,$$

$$(14)\qquad \int_0^{2\pi} cos^3\varphi(1 + ecos\varphi)^{n-2}[1 + (e - \dot{e})cos\varphi]^{-n}\, d\varphi = \mathbf{I_3}(e,\dot{e},n) + (e - \dot{e})\mathbf{I_4}(e,\dot{e},n)\,,$$

where, of course, we have used definitions (3). Then, we continue the transformation of the right-hand side of (9):

$$(15)\qquad -[ne(e - \dot{e})]^{-1}\int_0^{2\pi} cos\varphi\, sin\varphi(1 + ecos\varphi)^{n-1}\, d\{[1 + (e - \dot{e})cos\varphi]^{-n}\} =$$

$$= [ne(e - \dot{e})]^{-1}\{\, 2\int_0^{2\pi} cos^2\varphi(1 + ecos\varphi)^{n-2}[1 + (e - \dot{e})cos\varphi]^{-n}\, d\varphi \; +$$

$$+ 2e\int_0^{2\pi} cos^3\varphi(1 + ecos\varphi)^{n-2}[1 + (e - \dot{e})cos\varphi]^{-n}\, d\varphi \; -$$

$$-\int_0^{2\pi}(1 + ecos\varphi)^{n-2}[1 + (e - \dot{e})cos\varphi]^{-n}\, d\varphi \; -$$

$$- e\int_0^{2\pi} cos\varphi(1 + ecos\varphi)^{n-2}[1 + (e - \dot{e})cos\varphi]^{-n}\, d\varphi \; +$$

$$+ (-ne + e)\int_0^{2\pi} cos\varphi(1 + ecos\varphi)^{n-2}[1 + (e - \dot{e})cos\varphi]^{-n}\, d\varphi \; +$$

$$+ (ne - e)\int_0^{2\pi} cos^3\varphi(1 + ecos\varphi)^{n-2}[1 + (e - \dot{e})cos\varphi]^{-n}\, d\varphi\,\} =$$

$$= [ne(e - \dot{e})]^{-1}\{\, 2\mathbf{I_2}(e,\dot{e},n) + 2(e - \dot{e})\mathbf{I_3}(e,\dot{e},n) + 2e\mathbf{I_3}(e,\dot{e},n) + 2e(e - \dot{e})\mathbf{I_4}(e,\dot{e},n) -$$
$$- \mathbf{I_0}(e,\dot{e},n) - (e - \dot{e})\mathbf{I_1}(e,\dot{e},n) - e\mathbf{I_1}(e,\dot{e},n) - e(e - \dot{e})\mathbf{I_2}(e,\dot{e},n) + (-ne + e)\mathbf{I_1}(e,\dot{e},n) +$$
$$+ (-ne + e)(e - \dot{e})\mathbf{I_2}(e,\dot{e},n) + (ne - e)\mathbf{I_3}(e,\dot{e},n) + (ne - e)(e - \dot{e})\,\mathbf{I_4}(e,\dot{e},n)\,\}\,.$$

Substituting this result into (8) and multiplying by $ne(e - \dot{e})$, we obtain the following expression for the integral $\mathbf{I_4}(e,\dot{e},n)$:

(16) $e(e-\dot{e})\mathbf{I_4}(e,\dot{e},n) = \mathbf{I_0}(e,\dot{e},n) + [e + (n-1)\dot{e}]\mathbf{I_1}(e,\dot{e},n) - 2\mathbf{I_2}(e,\dot{e},n) - [3e + (n-2)\dot{e}]\mathbf{I_3}(e,\dot{e},n).$

In the above derivation we have supposed that $n \neq 0$, $e(u) \neq 0$ and $[e(u) - \dot{e}(u)] \neq 0$. But the linear relation makes sense even if some of these conditions are not fulfilled. We shall now check the validity of (16) for these particular cases. In the next, we suppose that u is a certain value of the logarithm of the focal parameter p, for which we have $e(u) = 0$, or $[e(u) - \dot{e}(u)] = 0$, or both equalities $e(u) = 0$ and $[e(u) - \dot{e}(u)] = 0$ hold. In other words, the cases *integer*/*noninteger* n, *zero*/*nonzero* $e(u)$ and *zero*/*nonzero* $[e(u) - \dot{e}(u)]$ give $2^3 = 8$ combinations. The relation (16) is until now proved only for one case, namely $n \neq 0$, $e(u) \neq 0$ and $[e(u) - \dot{e}(u)] \neq 0$ ***simultaneously***. We shall now prove (16) for the rest ***seven*** cases, which may be considered (in some sense) as certain particular exceptional situations.

3.1.1. Case $n \neq 0$, $e(u) = 0$, $e(u) - \dot{e}(u) = 0$ => $e(u) = \dot{e}(u) = 0$.

The linear relation (16) can be written as:

(17) $0 = \mathbf{I_0}(0,0,n) - 2\,\mathbf{I_2}(0,0,n).$

We immediately compute that $\mathbf{I_0}(0,0,n) = \int_0^{2\pi} d\varphi = 2\pi$ and $\mathbf{I_2}(0,0,n) = \int_0^{2\pi} \cos^2\varphi \, d\varphi = \pi$.

Obviously, (17) is fulfilled.

3.1.2.1. Case $n \neq 0$, $n \neq 1$, $e(u) = 0$, $e(u) - \dot{e}(u) \neq 0$ => $\dot{e}(u) \neq 0$.

The relation (16) now becomes:

(18) $0 = \mathbf{I_0}(0,\dot{e},n) + (n-1)\dot{e}\mathbf{I_1}(0,\dot{e},n) - 2\mathbf{I_2}(0,\dot{e},n) - (n-2)\dot{e}\mathbf{I_3}(0,\dot{e},n)$.

We compute directly that:

(19) $$\mathbf{I_3}(0,\dot{e},n) = \int_0^{2\pi} \cos^3\varphi (1 - \dot{e}\cos\varphi)^{-(n+1)} \, d\varphi = -\dot{e}^{-1} \int_0^{2\pi} \cos^2\varphi [(1 - \dot{e}\cos\varphi) - 1] \times$$

$$\times (1 - \dot{e}\cos\varphi)^{-(n+1)} \, d\varphi = \dot{e}^{-1} \int_0^{2\pi} \cos^2\varphi (1 - \dot{e}\cos\varphi)^{-(n+1)} \, d\varphi \; -$$

$$- \dot{e}^{-1} \int_0^{2\pi} \cos^2\varphi (1 - \dot{e}\cos\varphi)^{-n} \, d\varphi = \dot{e}^{-1}\mathbf{I_2}(0,\dot{e},n) - \dot{e}^{-1} \int_0^{2\pi} (1 - \sin^2\varphi)(1 - \dot{e}\cos\varphi)^{-n} \, d\varphi \; =$$

$$= \dot{e}^{-1}\mathbf{I_2}(0,\dot{e},n) - \dot{e}^{-1} \int_0^{2\pi} (1 - \dot{e}\cos\varphi)^{-n} \, d\varphi + \dot{e}^{-2} \int_0^{2\pi} \sin\varphi (1 - \dot{e}\cos\varphi)^{-n} \, d(1 - \dot{e}\cos\varphi) \; =$$

$$= \dot{e}^{-1}\mathbf{I}_2(0,\dot{e},n) - \dot{e}^{-1}\mathbf{I}_0(0,\dot{e},n) + \mathbf{I}_1(0,\dot{e},n) + [(-n+1)\dot{e}^2]^{-1}\int_0^{2\pi} \sin\varphi d[(1-\dot{e}\cos\varphi)^{-(n-1)}] =$$

$$= -\dot{e}^{-1}\mathbf{I}_0(0,\dot{e},n) + \mathbf{I}_1(0,\dot{e},n) + \dot{e}^{-1}\mathbf{I}_2(0,\dot{e},n) - [(n-1)\dot{e}^2]^{-1} \times$$

$$\times \{ \sin\varphi[(1-\dot{e}\cos\varphi)^{-(n-1)}] \Big|_0^{2\pi} - \int_0^{2\pi} \cos\varphi[(1-\dot{e}\cos\varphi)^{-(n-1)}]\, d\varphi \} =$$

$$= -\dot{e}^{-1}\mathbf{I}_0(0,\dot{e},n) + \mathbf{I}_1(0,\dot{e},n) + \dot{e}^{-1}\mathbf{I}_2(0,\dot{e},n) + [(n-1)\dot{e}^2]^{-1}\int_0^{2\pi} \cos\varphi(1 - 2\dot{e}\cos\varphi +$$

$$+ \dot{e}^2\cos^2\varphi)(1-\dot{e}\cos\varphi)^{-(n+1)}\, d\varphi = -\dot{e}^{-1}\mathbf{I}_0(0,\dot{e},n) + \mathbf{I}_1(0,\dot{e},n) + \dot{e}^{-1}\mathbf{I}_2(0,\dot{e},n) +$$

$$+[(n-1)\dot{e}^2]^{-1}\mathbf{I}_1(0,\dot{e},n) - 2[(n-1)\dot{e}]^{-1}\mathbf{I}_2(0,\dot{e},n) + (n-1)^{-1}\mathbf{I}_3(0,\dot{e},n)\,.$$

Consequently, for $n \neq 1$ (by supposition), after multiplying the both sides of (19) by $\dot{e}(n-1)$, we have:

(20) $$-(n-1)\mathbf{I}_0(0,\dot{e},n) + [\dot{e}^{-1} + (n-1)\dot{e}]\mathbf{I}_1(0,\dot{e},n) + (n-3)\mathbf{I}_2(0,\dot{e},n) - (n-2)\dot{e}\mathbf{I}_3(0,\dot{e},n) = 0\,.$$

By direct computation we also get:

(21) $$\mathbf{I}_2(0,\dot{e},n) = \int_0^{2\pi} \cos^2\varphi(1-\dot{e}\cos\varphi)^{-(n+1)}\, d\varphi = -\dot{e}^{-1}\int_0^{2\pi} \cos\varphi[(1-\dot{e}\cos\varphi) - 1] \times$$

$$\times (1-\dot{e}\cos\varphi)^{-(n+1)}\, d\varphi = \dot{e}^{-1}\int_0^{2\pi} \cos\varphi(1-\dot{e}\cos\varphi)^{-(n+1)}\, d\varphi -$$

$$-\dot{e}^{-1}\int_0^{2\pi} \cos\varphi(1-\dot{e}\cos\varphi)^{-n}\, d\varphi = \dot{e}^{-1}\mathbf{I}_1(0,\dot{e},n) - \dot{e}^{-1}\int_0^{2\pi} (1-\dot{e}\cos\varphi)^{-n}\, d\sin\varphi =$$

$$= \dot{e}^{-1}\mathbf{I}_1(0,\dot{e},n) - \dot{e}^{-1}\{ \sin\varphi[(1-\dot{e}\cos\varphi)^{-n}] \Big|_0^{2\pi} - \int_0^{2\pi} \sin\varphi\, d\{ [(1-\dot{e}\cos\varphi)^{-n}] \} \} =$$

$$= \dot{e}^{-1}\mathbf{I}_1(0,\dot{e},n) - n\int_0^{2\pi} (1-\cos^2\varphi)[(1-\dot{e}\cos\varphi)^{-(n+1)}]\, d\varphi =$$

$$= \dot{e}^{-1}\mathbf{I}_1(0,e,n) - n[\mathbf{I}_0(0,e,n) - \mathbf{I}_2(0,e,n)]\,.$$

Therefore:

(22) $$(n-1)\mathbf{I}_2(0,e,n) = n\mathbf{I}_0(0,e,n) - \dot{e}^{-1}\mathbf{I}_1(0,e,n)\,,$$ or

(23) $$(n-3)\mathbf{I}_2(0,e,n) = n\mathbf{I}_0(0,e,n) - \dot{e}^{-1}\mathbf{I}_1(0,e,n) - 2\mathbf{I}_2(0,e,n)\,.$$

Substituting this result for $(n-3)\mathbf{I}_2(0,e,n)$ into the relation (20), we shall obtain the sought equality (18).

3.1.2.2. Case $n = 1$, $e(u) = 0$, $e(u) - \dot{e}(u) \neq 0 \Rightarrow \dot{e}(u) \neq 0$.

The relation (16) becomes:

(24) $$0 = \mathbf{I_0}(0,\dot{e},1) - 2\mathbf{I_2}(0,\dot{e},1) + \dot{e}\mathbf{I_3}(0,\dot{e},1) .$$

We compute directly that:

(25) $$\mathbf{I_3}(0,\dot{e},1) = \int_0^{2\pi} cos^3\varphi(1 - \dot{e}cos\varphi)^{-2}\, d\varphi = \dot{e}^{-3}\int_0^{2\pi}(-\dot{e}^3cos^3\varphi)(1 - \dot{e}cos\varphi)^{-2}\, d\varphi =$$

$$= -\dot{e}^{-3}\int_0^{2\pi}(1 - 3\dot{e}cos\varphi + 3\dot{e}^2cos^2\varphi - \dot{e}^3cos^3\varphi)(1 - \dot{e}cos\varphi)^{-2}\, d\varphi +$$

$$+ \dot{e}^{-3}\int_0^{2\pi}(1 - 3\dot{e}cos\varphi + 3\dot{e}^2cos^2\varphi)(1 - \dot{e}cos\varphi)^{-2}\, d\varphi = -\dot{e}^{-3}\int_0^{2\pi}(1 - \dot{e}cos\varphi)\, d\varphi +$$

$$+ \dot{e}^{-3}\mathbf{I_0}(0,\dot{e},1) - (3\dot{e}/\dot{e}^3)\mathbf{I_1}(0,\dot{e},1) + (3\dot{e}^2/\dot{e}^3)\mathbf{I_2}(0,\dot{e},1) =$$

$$= -2\pi\dot{e}^{-3} + 0 + \dot{e}^{-3}\mathbf{I_0}(0,\dot{e},1) - 3\dot{e}^{-2}\mathbf{I_1}(0,\dot{e},1) + 3\dot{e}^{-1}\mathbf{I_2}(0,\dot{e},1) .$$

Let us evaluate the third nonzero term in the right-hand side:

(26) $$-3\dot{e}^{-2}\mathbf{I_1}(0,\dot{e},1) = -3(2\dot{e}^3)^{-1}\int_0^{2\pi}2\dot{e}cos\varphi(1 - \dot{e}cos\varphi)^{-2}\, d\varphi =$$

$$= 3(2\dot{e}^3)^{-1}\int_0^{2\pi}(1 - 2\dot{e}cos\varphi + \dot{e}^2cos^2\varphi)(1 - \dot{e}cos\varphi)^{-2}\, d\varphi -$$

$$- 3(2\dot{e}^3)^{-1}\int_0^{2\pi}(1 + \dot{e}^2cos^2\varphi)(1 - \dot{e}cos\varphi)^{-2}\, d\varphi =$$

$$= 3(2\dot{e}^3)^{-1}2\pi - 3(2\dot{e}^3)^{-1}\mathbf{I_0}(0,\dot{e},1) - 3(2\dot{e})^{-1}\mathbf{I_2}(0,\dot{e},1) .$$

Substituting this result into equation (25), we obtain:

(27) $$\mathbf{I_3}(0,\dot{e},1) = [-\dot{e}^{-3} + 3(2\dot{e}^3)^{-1}]2\pi + [\dot{e}^{-3} - 3(2\dot{e}^3)^{-1}]\mathbf{I_0}(0,\dot{e},1) +$$
$$+ [3\dot{e}^{-1} - 3(2\dot{e})^{-1}\mathbf{I_2}(0,\dot{e},1) , \quad \text{or}$$

(28) $$\mathbf{I_3}(0,\dot{e},1) = 2\pi\,(2\dot{e}^{-3}) - (2\dot{e}^{-3})\mathbf{I_0}(0,\dot{e},1) + 3(2\dot{e})^{-1}\mathbf{I_2}(0,\dot{e},1) .$$

In straightforward way we find that:

(29) $$3\dot{e}^{-1}\mathbf{I_2}(0,\dot{e},1) = 3\dot{e}^{-1}\int_0^{2\pi}cos^2\varphi(1 - \dot{e}cos\varphi)^{-2}\, d\varphi = 3\dot{e}^{-3}\int_0^{2\pi}\dot{e}^2cos^2\varphi(1 - \dot{e}cos\varphi)^{-2}\, d\varphi =$$

$$= 3\dot{e}^{-3}\int_0^{2\pi}(1 - 2\dot{e}cos\varphi + \dot{e}^2cos^2\varphi)(1 - \dot{e}cos\varphi)^{-2}\, d\varphi - 3\dot{e}^{-3}\int_0^{2\pi}(1 - 2\dot{e}cos\varphi) \times$$

$$\times (1 - \dot{e}cos\varphi)^{-2}\, d\varphi = 3\dot{e}^{-3}2\pi - 3\dot{e}^{-3}\mathbf{I_0}(0,\dot{e},1) + 6\dot{e}^{-2}\mathbf{I_1}(0,\dot{e},1) .$$

Dividing the both sides of this equality by 6, we have:

(30) $$(2\dot{e}^3)^{-1}2\pi = (2\dot{e}^3)^{-1}\mathbf{I_0}(0,\dot{e},1) - \dot{e}^{-2}\mathbf{I_1}(0,\dot{e},1) + (2\dot{e})^{-1}\mathbf{I_2}(0,\dot{e},1) .$$

Substituting this into (28) and multiplying by $\dot{e}$, we shall finally obtain the following *intermediate result*:

(31) $0 = \dot{e}^{-1}\mathbf{I_1}(0,\dot{e},1) - 2\mathbf{I_2}(0,\dot{e},1) + \dot{e}\mathbf{I_3}(0,\dot{e},1)$.

Our further computations include explicit analytical evaluations of the integrals $\mathbf{I_0}(0,\dot{e},1)$ and $\mathbf{I_3}(0,\dot{e},1)$:

(32) $$\mathbf{I_0}(0,\dot{e},1) = \int_0^{2\pi}(1 - \dot{e}cos\varphi)^{-2}\, d\varphi = 2\pi\,(1 - \dot{e}^2)^{-3/2}\,,$$

according to formula **858.535** from [14].

(33) $$\mathbf{I_1}(0,\dot{e},1) = \int_0^{2\pi} cos\varphi(1 - \dot{e}cos\varphi)^{-2}\, d\varphi = -\,\dot{e}^{-1}\int_0^{2\pi}[(1 - \dot{e}cos\varphi) - 1](1 - \dot{e}cos\varphi)^{-2}\, d\varphi\ =$$

$$= \dot{e}^{-1}\int_0^{2\pi}(1 - \dot{e}cos\varphi)^{-2}\, d\varphi - \dot{e}^{-1}\int_0^{2\pi}(1 - \dot{e}cos\varphi)^{-1}\, d\varphi\ .$$

From Dwight [14], formula **858.525**, we find that:

(34) $$\int_0^{2\pi}(1 - \dot{e}cos\varphi)^{-1}\, d\varphi = 2\pi\,(1 - \dot{e}^2)^{-1/2}\,.$$

Combining evaluations (32) and (34) into (33), the result is:

(35) $$\mathbf{I_1}(0,\dot{e},1) = 2\pi\,\dot{e}^{-1}(1 - \dot{e}^2)^{-1}(1 - \dot{e}^2)^{-1/2} - 2\pi\,\dot{e}^{-1}(1 - \dot{e}^2)^{-1/2} =$$
$$= 2\pi\,\dot{e}(1 - \dot{e}^2)^{-3/2} = \dot{e}\mathbf{I_0}(0,\dot{e},1)\,.$$

Substituting the above result into (31), we finally obtain the necessary relation (24).

3.1.3. Case $n \neq 0$, $e(u) \neq 0$, $e(u) - \dot{e}(u) = 0 \Rightarrow \dot{e}(u) = e(u) \neq 0$.

The relation (16) can be written as:

(36) $0 = \mathbf{I_0}(e,\dot{e} = e,n) + n\dot{e}\mathbf{I_1}(e,\dot{e} = e,n) - 2\mathbf{I_2}(e,\dot{e} = e,n) - (n + 1)\dot{e}\mathbf{I_3}(e,\dot{e} = e,n)$.

We compute directly that:

(37) $$\mathbf{I_3}(e,\dot{e} = e,n) = \int_0^{2\pi} cos^3\varphi(1 + ecos\varphi)^{n-2}\, d\varphi = e^{-1}\int_0^{2\pi}[(1 + ecos\varphi) - 1](1 + ecos\varphi)^{n-2}\ \times$$

$$\times\, cos^2\varphi\, d\varphi = e^{-1}\int_0^{2\pi} cos^2\varphi(1 + ecos\varphi)^{n-1} d\varphi\ - e^{-1}\int_0^{2\pi} cos^2\varphi(1 + ecos\varphi)^{n-2}\, d\varphi =$$

$$= -\,e^{-1}\mathbf{I_2}(e,\dot{e} = e,n) + e^{-1}\int_0^{2\pi} cos\varphi(1 + ecos\varphi)^{n-1} d\, sin\varphi = -\,e^{-1}\mathbf{I_2}(e,\dot{e} = e,n) +$$

$$+ e^{-1}\{\, sin\varphi\, cos\varphi(1 + ecos\varphi)^{n-1}\Big|_0^{2\pi} - \int_0^{2\pi} sin\varphi\, d[cos\varphi(1 + ecos\varphi)^{n-1}]\,\} =$$

$$= -e^{-1}\mathbf{I}_2(e,\dot{e}=e,n) + e^{-1}\int_0^{2\pi} sin^2\varphi(1+ecos\varphi)^{n-1}d\varphi + (n-1)e^{-1}\int_0^{2\pi} sin\varphi\, cos\varphi \;\times$$

$$\times (1+ecos\varphi)^{n-2}esin\varphi\, d\varphi = -e^{-1}\mathbf{I}_2(e,\dot{e}=e,n) + e^{-1}\int_0^{2\pi}(1-cos^2\varphi)(1+ecos\varphi)^{n-1}d\varphi +$$

$$+ (n-1)\int_0^{2\pi} cos\varphi\, sin^2\varphi(1+ecos\varphi)^{n-2}d\varphi = -e^{-1}\mathbf{I}_2(e,\dot{e}=e,n) + e^{-1}\int_0^{2\pi}(1+ecos\varphi)^{n-2}d\varphi +$$

$$+ \int_0^{2\pi} cos\varphi(1+ecos\varphi)^{n-2}d\varphi - e^{-1}\int_0^{2\pi} cos^2\varphi(1+ecos\varphi)^{n-2}d\varphi - \int_0^{2\pi} cos^3\varphi(1+ecos\varphi)^{n-2}d\varphi +$$

$$+ (n-1)\int_0^{2\pi} cos\varphi(1+ecos\varphi)^{n-2}d\varphi - (n-1)\int_0^{2\pi} cos^3\varphi\,(1+ecos\varphi)^{n-2}d\varphi \;=$$

$$= -e^{-1}\mathbf{I}_2(e,\dot{e}=e,n) + e^{-1}\mathbf{I}_0(e,\dot{e}=e,n) + \mathbf{I}_1(e,\dot{e}=e,n) - e^{-1}\mathbf{I}_2(e,\dot{e}=e,n) -$$
$$- \mathbf{I}_3(e,\dot{e}=e,n) + (n-1)\mathbf{I}_1(e,\dot{e}=e,n) - (n-1)\mathbf{I}_3(e,\dot{e}=e,n)\,.$$

Therefore:

(38) $$0 = e^{-1}\mathbf{I}_0(e,\dot{e}=e,n) + n\mathbf{I}_1(e,\dot{e}=e,n) - 2e^{-1}\mathbf{I}_2(e,\dot{e}=e,n) - (n+1)\mathbf{I}_3(e,\dot{e}=e,n)\,.$$

Multiplying (38) by e and taking into account that for the considered value of u $e(u) = \dot{e}(u)$, we complete the proof of the linear relation (36).

3.1.4. Case $n = 0$, $e(u) = 0$, $e(u) - \dot{e}(u) = 0 \Rightarrow \dot{e}(u) = 0$.

The relation (16) can be written as:

(39) $$0 = \mathbf{I}_0(0,0,0) - 2\mathbf{I}_2(0,0,0)\,.$$

In this case $\mathbf{I}_0(0,0,0) = \int_0^{2\pi} d\varphi = 2\pi$ and $\mathbf{I}_2(0,0,0) = \int_0^{2\pi} cos^2\varphi\, d\varphi = \pi$. Then (39) immediately follows.

3.1.5. Case $n = 0$, $e(u) = 0$, $e(u) - \dot{e}(u) \neq 0 \Rightarrow \dot{e}(u) \neq 0$.

The relation (16) becomes:

(40) $$0 = \mathbf{I}_0(0,\dot{e},0) - \dot{e}\mathbf{I}_1(0,\dot{e},0) - 2\mathbf{I}_2(0,\dot{e},0) + 2\dot{e}\mathbf{I}_3(0,\dot{e},0)\,.$$

The direct computation gives:

(41) $$\mathbf{I}_3(0,\dot{e},0) = \int_0^{2\pi} cos^3\varphi(1-\dot{e}cos\varphi)^{-1}d\varphi = -\dot{e}^{-1}\int_0^{2\pi} cos^2\varphi[(1-\dot{e}cos\varphi)-1](1-\dot{e}cos\varphi)^{-1}d\varphi =$$

$$= \dot{e}^{-1}\int_0^{2\pi} cos^2\varphi(1-\dot{e}cos\varphi)^{-1}d\varphi - \dot{e}^{-1}\int_0^{2\pi} cos^2\varphi\, d\varphi = \dot{e}^{-1}\mathbf{I}_2(0,\dot{e},0) - \pi\,\dot{e}^{-1}\,.$$

Multiplying by $2\dot{e}$, we shall obtain:

(42) $2\dot{e}\mathbf{I}_3(0,\dot{e},0) - 2\mathbf{I}_2(0,\dot{e},0) = -2\pi$.

Further we also evaluate that:

(43) $\mathbf{I}_1(0,\dot{e},0) = \int_0^{2\pi} cos\varphi(1 - \dot{e}cos\varphi)^{-1}d\varphi = -\dot{e}^{-1}\int_0^{2\pi}[(1 - \dot{e}cos\varphi) - 1](1 - \dot{e}cos\varphi)^{-1}d\varphi =$

$= -2\pi\dot{e}^{-1} + \dot{e}^{-1}\int_0^{2\pi}(1 - \dot{e}cos\varphi)^{-1}d\varphi$.

Consequently:

(44) $\dot{e}\mathbf{I}_1(0,\dot{e},0) - \mathbf{I}_0(0,\dot{e},0) = -2\pi$.

Combining (42) and (44), we attain to the relation (40).

3.1.6. Case $n = 0$, $e(u) \neq 0$, $e(u) - \dot{e}(u) = 0 \Rightarrow \dot{e}(u) = e(u) \neq 0$.

The linear relation (16) can be written as:

(45) $0 = \mathbf{I}_0(e,\dot{e} = e,0) - 2\mathbf{I}_2(e,\dot{e} = e,0) - 3\dot{e}\mathbf{I}_3(e,\dot{e} = e,0)$.

To prove the above statement, we must perform evaluation of the integrals $\mathbf{I}_0(e,\dot{e} = e,0) = \int_0^{2\pi}(1 - \dot{e}cos\varphi)^{-2}d\varphi$, …, $\mathbf{I}_3(e,\dot{e} = e,0) = \int_0^{2\pi}cos^3\varphi(1 - \dot{e}cos\varphi)^{-2}d\varphi$. Clearly, this is fully analogous to the estimation of the integrals in the case **3.1.2.2.** We must only replace $-\dot{e}(u)$ in the denominators of the integrals by $e(u)$ and proceed by the same way, when we were proving the relation (24). We shall not write out these clumsy calculations again, in order to prove validity of the relation (45).

3.1.7. Case $n = 0$, $e(u) \neq 0$, $e(u) - \dot{e}(u) \neq 0$.

The linear relation (16) now becomes:

(46) $e(e - \dot{e})\mathbf{I}_4(e,\dot{e},0) = \mathbf{I}_0(e,\dot{e},0) + (e - \dot{e})\mathbf{I}_1(e,\dot{e},0) - 2\mathbf{I}_2(e,\dot{e},0) - (3e - 2\dot{e})\mathbf{I}_3(e,\dot{e},0)$.

In an earlier paper [11] (formulas (3a) – (3d)) we have already derived in explicit form ***analytical*** expressions for the integrals, entering in (46). We shall now rewrite in a little more compact form these results. Let us denote by $\boldsymbol{A}(e,\dot{e})$ the multiplier:

(47) $\boldsymbol{A}(e,\dot{e}) = 2\pi\dot{e}^{-2}(1 - e^2)^{-3/2}[1 - (e - \dot{e})^2]^{-1/2}$.

Then, according to [11] (formulas (3a) – (3d)), with this simplification of the notations, we have:

(48) $\mathbf{I}_0(e,\dot{e},0) = \boldsymbol{A}(e,\dot{e})\{ e\dot{e}[1 - (e - \dot{e})^2]^{1/2} - e(e - \dot{e})(1 - e^2)[1 - (e - \dot{e})^2]^{1/2} +$
$+ (e - \dot{e})^2(1 - e^2)^{3/2}\}$,

(49) $\mathbf{I}_1(e,\dot{e},0) = \boldsymbol{A}(e,\dot{e})\{\,(e-\dot{e}-e^3)[1-(e-\dot{e})^2]^{1/2} - (e-\dot{e})(1-e^2)^{3/2}\,\}$,

(50) $\mathbf{I}_2(e,\dot{e},0) = \boldsymbol{A}(e,\dot{e})\{\,(-1+e^2+e\dot{e})[1-(e-\dot{e})^2]^{1/2} + (1-e^2)^{3/2}\,\}$,

(51) $\mathbf{I}_3(e,\dot{e},0) = \boldsymbol{A}(e,\dot{e})e^{-2}(e-\dot{e})^{-1}\{-e^2(1-e^2)^{3/2} + [e^2 - e^4 - e^3\dot{e} - \dot{e}^2 + 2e^2\dot{e}^2 + \dot{e}^2(1-e^2)^{3/2}][1-(e-\dot{e})^2]^{1/2}\}$,

(52) $e(e-\dot{e})\mathbf{I}_4(e,\dot{e},0) = \boldsymbol{A}(e,\dot{e})e^{-2}(e-\dot{e})^{-1}\{\,e^3(1-e^2)^{3/2} + (-e^3+e^5+e^4\dot{e}+3e\dot{e}^2 - 5e^3\dot{e}^2 - 2\dot{e}^3 + 3e^2\dot{e}^3)[1-(e-\dot{e})^2]^{1/2} + (-3e\dot{e}^2+2\dot{e}^3)(1-e^2)^{3/2}[1-(e-\dot{e})^2]^{1/2}\,\}$.

Let us now compute the right-hand side of the equality (46):

(53) $\mathbf{I}_0(e,\dot{e},0) + (e-\dot{e})\mathbf{I}_1(e,\dot{e},0) - 2\mathbf{I}_2(e,\dot{e},0) - (3e-2\dot{e})\mathbf{I}_3(e,\dot{e},0) = \boldsymbol{A}(e,\dot{e})e^{-2}(e-\dot{e})^{-1} \times$
$\times\{\,e^2(e-\dot{e})e\dot{e}[1-(e-\dot{e})^2]^{1/2} - e^3(e-\dot{e})^2(1-e^2)[1-(e-\dot{e})^2]^{1/2} +$
$+ e^2(e-\dot{e})^3(1-e^2)^{3/2} + e^2(e-\dot{e})^2(e-\dot{e}-e^3)[1-(e-\dot{e})^2]^{1/2} - e^2(e-\dot{e})^2(1-e^2)^{3/2} +$
$+ (2-2e^2-2e\dot{e})e^2(e-\dot{e})[1-(e-\dot{e})^2]^{1/2} - 2e^2(e-\dot{e})(1-e^2)^{3/2} +$
$+ (3e-2\dot{e})e^2(1-e^2)^{3/2} + (-3e+2\dot{e})(e^2-e^4-e^3\dot{e}-\dot{e}^2+2e^2\dot{e}^2 +$
$+ \dot{e}^2(1-e^2)^{3/2}[1-(e-\dot{e})^2]^{1/2}\,\} =$
$= \boldsymbol{A}(e,\dot{e})e^{-2}(e-\dot{e})^{-1}\{(-e^3+e^5+e^4\dot{e}+3e\dot{e}^2 - 5e^3\dot{e}^2 - 2\dot{e}^3 + 3e^2\dot{e}^3)[1-(e-\dot{e})^2]^{1/2} +$
$+ e^3(1-e^2)^{3/2} + (-3e\dot{e}^2+2\dot{e}^3)(1-e^2)^{3/2}[1-(e-\dot{e})^2]^{1/2}\,\} = e(e-\dot{e})\mathbf{I}_4(e,\dot{e},0)$.

Hence, the linear relation (46) is proved. With this, we have also completed the validity of relation (16) in the general case. That is, for *integer*/*noninteger* powers n, *zero*/*nonzero* values of $e(u)$ (for $|e(u)| < 1$) and $\dot{e}(u)$ (for $|\dot{e}(u)| < 1$), and also for *zero*/*nonzero* values of $[e(u) - \dot{e}(u)]$ (for $|e(u) - \dot{e}(u)| < 1$).

3.2. Elimination of the integral $\mathbf{I}_2(e,\dot{e},n)$

Generally speaking, the approach in the computing of $\mathbf{I}_4(e,\dot{e},n)$ is the following: we perform a series of evaluations of $\mathbf{I}_4(e,\dot{e},n)$, decomposing its integrand into such, containing into their nominators powers of $cos\varphi$ equal or less than 4, and the same denominators $[1 + (e-\dot{e})cos\varphi]^{n+1}$. After that, we transfer the repeatedly appeared integrals $\mathbf{I}_4(e,\dot{e},n)$ in the *right-hand* side into the *left-hand* side, in order to combine all $\mathbf{I}_4(e,\dot{e},n)$. Unfortunately, such a procedure ***does not work at all*** when we try to apply it for elimination of the integral $\mathbf{I}_3(e,\dot{e},n)$ (we suppose that the linear relation (16) is already used for removing of the integral $\mathbf{I}_4(e,\dot{e},n)$). The reason for this unsuccessful attempt is that the multiplier before $\mathbf{I}_3(e,\dot{e},n)$ equals to zero for all values of the variable u. It is suspected that this impossibility is in relation to the linear independence of the considered seven integrals (1), (2) and (3). We shall not deal with this problem in the present paper and continue to the evaluation of the integral $\mathbf{I}_2(e,\dot{e},n)$.

According to the definition (3), we can write:

(54) $\mathbf{I}_3(e,\dot{e},n) = \int_0^{2\pi} cos^3\varphi(1+ecos\varphi)^{n-2}[1+(e-\dot{e})cos\varphi]^{-(n+1)}\,d\varphi =$

$$= e^{-1}\int_0^{2\pi} cos^2\varphi[(1+ecos\varphi)-1](1+ecos\varphi)^{n-2}[1+(e-\dot{e})cos\varphi]^{-(n+1)}\,d\varphi \;=$$

$$= e^{-1}\int_0^{2\pi} cos^2\varphi(1+ecos\varphi)^{n-1}[1+(e-\dot{e})cos\varphi]^{-(n+1)}\,d\varphi \;-$$

$$- e^{-1}\int_0^{2\pi} cos^2\varphi(1+ecos\varphi)^{n-2}[1+(e-\dot{e})cos\varphi]^{-(n+1)}\,d\varphi \;=$$

$$= -e^{-1}\mathbf{I_2}(e,\dot{e},n) + e^{-1}\int_0^{2\pi}(1-sin^2\varphi)(1+ecos\varphi)^{n-1}[1+(e-\dot{e})cos\varphi]^{-(n+1)}\,d\varphi \;=$$

$$= -e^{-1}\mathbf{I_2}(e,\dot{e},n) + e^{-1}\int_0^{2\pi}(1+ecos\varphi)^{n-1}[1+(e-\dot{e})cos\varphi]^{-(n+1)}\,d\varphi \;+$$

$$+ [e(e-\dot{e})]^{-1}\int_0^{2\pi} sin\varphi(1+ecos\varphi)^{n-1}[1+(e-\dot{e})cos\varphi]^{-(n+1)}\,d[1+(e-\dot{e})cos\varphi]\,.$$

Applying a relation analogous to (7), we obtain:

(55) $$\mathbf{I_3}(e,\dot{e},n) = -e^{-1}\mathbf{I_2}(e,\dot{e},n) + e^{-1}\mathbf{I_0}(e,\dot{e},n) + \mathbf{I_1}(e,\dot{e},n) \;-$$

$$- [ne(e-\dot{e})]^{-1}\int_0^{2\pi} sin\varphi(1+ecos\varphi)^{n-1}d\{\,[1+(e-\dot{e})cos\varphi]^{-n}\,\} \;=$$

$$= e^{-1}\mathbf{I_0}(e,\dot{e},n) + \mathbf{I_1}(e,\dot{e},n) - e^{-1}\mathbf{I_2}(e,\dot{e},n) - [ne(e-\dot{e})]^{-1}\{\, sin\varphi(1+ecos\varphi)^{n-1} \times$$

$$\times [1+(e-\dot{e})cos\varphi]^{-n}\Big|_0^{2\pi}\} + [ne(e-\dot{e})]^{-1}\{\int_0^{2\pi} cos\varphi(1+ecos\varphi)^{n-1}[1+(e-\dot{e})cos\varphi]^{-n}d\varphi -$$

$$- e(n-1)\int_0^{2\pi} sin^2\varphi(1+ecos\varphi)^{n-2}[1+(e-\dot{e})cos\varphi]^{-n}d\varphi\,\} \;=$$

$$= e^{-1}\mathbf{I_0}(e,\dot{e},n) + \mathbf{I_1}(e,\dot{e},n) - e^{-1}\mathbf{I_2}(e,\dot{e},n) + [ne(e-\dot{e})]^{-1}\{\int_0^{2\pi} cos\varphi[1+(e-\dot{e})cos\varphi] \times$$

$$\times(1+ecos\varphi)^{n-1}[1+(e-\dot{e})cos\varphi]^{-(n+1)}d\varphi - (n-1)e\{\int_0^{2\pi}(1+ecos\varphi)^{n-2} \;\times$$

$$\times [1+(e-\dot{e})cos\varphi]^{-(n+1)}d\varphi + (e-\dot{e})\int_0^{2\pi} cos\varphi(1+ecos\varphi)^{n-2}[1+(e-\dot{e})cos\varphi]^{-(n+1)}d\varphi -$$

$$-\int_0^{2\pi} cos^2\varphi(1+ecos\varphi)^{n-2}[1+(e-\dot{e})cos\varphi]^{-(n+1)}d\varphi -$$

$$- (e - \dot{e}) \int_0^{2\pi} \cos^3\varphi (1 + e\cos\varphi)^{n-2} [1 + (e - \dot{e})\cos\varphi]^{-(n+1)} d\varphi \,\} \,\} =$$

$$= e^{-1}\mathbf{I_0}(e,\dot{e},n) + \mathbf{I_1}(e,\dot{e},n) - e^{-1}\mathbf{I_2}(e,\dot{e},n) + [ne(e - \dot{e})]^{-1}[\mathbf{I_1}(e,\dot{e},n) + e\mathbf{I_2}(e,\dot{e},n) +$$
$$+ (e - \dot{e})\mathbf{I_2}(e,\dot{e},n) + e(e - \dot{e})\mathbf{I_3}(e,\dot{e},n) - (n-1)e\mathbf{I_0}(e,\dot{e},n) - (n-1)e(e - \dot{e})\mathbf{I_1}(e,\dot{e},n) +$$
$$+ (n-1)e\mathbf{I_2}(e,\dot{e},n) + (n-1)e(e - \dot{e})\mathbf{I_3}(e,\dot{e},n)] .$$

Consequently, we have the following expression:

(56) $$\mathbf{I_3}(e,\dot{e},n) = \{e^{-1} - (n-1)[n(e - \dot{e})]^{-1}\}\mathbf{I_0}(e,\dot{e},n) +$$
$$+ \{1 + [ne(e - \dot{e})]^{-1} - (n-1)n^{-1}\}\mathbf{I_1}(e,\dot{e},n) + \{- e^{-1} + [n(e - \dot{e})]^{-1} + (ne)^{-1} +$$
$$+ (n-1)[n(e - \dot{e})]^{-1}\}\mathbf{I_2}(e,\dot{e},n) + [n^{-1} + (n-1)n^{-1}]\mathbf{I_3}(e,\dot{e},n) .$$

It is evident that the integrals $\mathbf{I_3}(e,\dot{e},n)$ from the both sides of the above equality cancel out, and it is impossible to determine any linear relation between $\mathbf{I_3}(e,\dot{e},n)$, $\mathbf{I_0}(e,\dot{e},n)$, $\mathbf{I_1}(e,\dot{e},n)$, and $\mathbf{I_2}(e,\dot{e},n)$, ***as already mentioned above***. Nevertheless, the result (56) may be used to eliminate the integral $\mathbf{I_2}(e,\dot{e},n)$. Multiplying (56) by $ne(e - \dot{e})$, we obtain:

(57) $$[e + (n-1)\dot{e}]\mathbf{I_2}(e,\dot{e},n) = (- e + n\dot{e})\mathbf{I_0}(e,\dot{e},n) - [1 + e(e - \dot{e})]\mathbf{I_1}(e,\dot{e},n) .$$

In the above derivation, we again have supposed that simultaneously are fulfilled the following three conditions: $n \neq 0$, $e(u) \neq 0$, $e(u) - \dot{e}(u) \neq 0$, for every considered value of the independent variable $u \equiv ln\ p$ (p is the focal parameter of the particle orbit). We note that (57) makes sense even if some (or even all) of these restrictions are violated.

3.2.1. Case $n \neq 0$, $e(u) = 0$, $e(u) - \dot{e}(u) = 0$ => $e(u) = \dot{e}(u) = 0$.

The relation (57) is obviously satisfied, because $\mathbf{I_1}(0,0,n) = \int_0^{2\pi} \cos\varphi \, d\varphi = 0$.

3.2.2. Case $n \neq 0$, $e(u) = 0$, $e(u) - \dot{e}(u) \neq 0$ => $\dot{e}(u) \neq 0$.

The relation (57) takes the form:

(58) $$(n-1)\dot{e}\mathbf{I_2}(0,\dot{e},n) = n\dot{e}\mathbf{I_0}(0,\dot{e},n) - \mathbf{I_1}(0,\dot{e},n) .$$

We compute that:

(59) $$\mathbf{I_2}(0,\dot{e},n) = \int_0^{2\pi} \cos^2\varphi (1 - \dot{e}\cos\varphi)^{-(n+1)} d\varphi = \int_0^{2\pi} (1 - \dot{e}\cos\varphi)^{-(n+1)} d\varphi -$$

$$- \int_0^{2\pi} \sin^2\varphi (1 - \dot{e}\cos\varphi)^{-(n+1)} d\varphi = \mathbf{I_0}(0,\dot{e},n) - \dot{e}^{-1} \int_0^{2\pi} \sin\varphi (1 - \dot{e}\cos\varphi)^{-(n+1)} d(1 - \dot{e}\cos\varphi) =$$

$$= \mathbf{I_0}(0,\dot{e},n) + (n\dot{e})^{-1} \int_0^{2\pi} \sin\varphi \, d[(1 - \dot{e}\cos\varphi)^{-n}] = \mathbf{I_0}(0,\dot{e},n) + (n\dot{e})^{-1} \{\sin\varphi (1 - \dot{e}\cos\varphi)^{-n} \Big|_0^{2\pi} -$$

$$- \int_0^{2\pi} \cos\varphi (1 - \dot{e}\cos\varphi)^{-n} d\varphi\} = \mathbf{I_0}(0,\dot{e},n) - (n\dot{e})^{-1} \int_0^{2\pi} \cos\varphi (1 - \dot{e}\cos\varphi)(1 - \dot{e}\cos\varphi)^{-(n+1)} d\varphi =$$

$= \mathbf{I_0}(0,\dot{e},n) - (n\dot{e})^{-1}\mathbf{I_1}(0,\dot{e},n) + n^{-1}\mathbf{I_2}(0,\dot{e},n)$.

From this equality follows:

(60) $(1 - n^{-1})\mathbf{I_2}(0,\dot{e},n) \equiv (n-1)\dot{e}(n\dot{e})^{-1}\mathbf{I_2}(0,\dot{e},n) = \mathbf{I_0}(0,\dot{e},n) - (n\dot{e})^{-1}\mathbf{I_1}(0,\dot{e},n)$.

Multiplying by $n\dot{e} \neq 0$, we obtain the seeked equality (58).

3.2.3. Case $n \neq 0$, $e(u) \neq 0$, $e(u) - \dot{e}(u) = 0 \Rightarrow \dot{e}(u) = e(u) \neq 0$.

The relation (57) in this case becomes:

(61) $[e + (n-1)\dot{e}]\mathbf{I_2}(e,\dot{e}=e,n) \equiv n\dot{e}\mathbf{I_2}(e,\dot{e}=e,n) = (n-1)\dot{e}\mathbf{I_0}(e,\dot{e}=e,n) - \mathbf{I_1}(e,\dot{e}=e,n)$.

We directly compute that:

(62) $\mathbf{I_2}(e,\dot{e}=e,n) = \int_0^{2\pi} cos^2\varphi(1 + ecos\varphi)^{n-2}d\varphi = e^{-1}\int_0^{2\pi} cos\varphi[(1 + ecos\varphi) - 1] \times$

$\times(1 + ecos\varphi)^{n-2}d\varphi = -e^{-1}\int_0^{2\pi} cos\varphi(1 + ecos\varphi)^{n-2}d\varphi + e^{-1}\int_0^{2\pi}(1 + ecos\varphi)^{n-1}d\,sin\varphi =$

$= -e^{-1}\mathbf{I_1}(e,\dot{e}=e,n) + e^{-1}\{ sin\varphi(1 + ecos\varphi)^{n-1}\Big|_0^{2\pi} - \int_0^{2\pi} sin\varphi\, d[(1 + ecos\varphi)^{n-1}]\} =$

$= -e^{-1}\mathbf{I_1}(e,\dot{e}=e,n) + (n-1)\int_0^{2\pi}(1 - cos^2\varphi)(1 + ecos\varphi)^{n-2}d\varphi =$

$= -e^{-1}\mathbf{I_1}(e,\dot{e}=e,n) + (n-1)\mathbf{I_0}(e,\dot{e}=e,n) - (n-1)\mathbf{I_2}(e,\dot{e}=e,n)$.

Multiplication of the both sides by $\dot{e}(u) = e(u) \neq 0$ gives the result:

(63) $n\dot{e}\mathbf{I_2}(e,\dot{e}=e,n) = (n-1)\dot{e}\mathbf{I_0}(e,\dot{e}=e,n) - \mathbf{I_1}(e,\dot{e}=e,n)$,

that proves (61).

3.2.4. Case $n = 0$, $e(u) = 0$, $e(u) - \dot{e}(u) = 0 \Rightarrow \dot{e}(u) = 0$.

The relation (57) is obviously true, because $\mathbf{I_1}(0,0,0) = \int_0^{2\pi} cos\varphi\, d\varphi = 0$.

3.2.5. Case $n = 0$, $e(u) = 0$, $e(u) - \dot{e}(u) \neq 0 \Rightarrow \dot{e}(u) \neq 0$.

The relation (57) now becomes:

(64) $-\dot{e}\mathbf{I_2}(0,\dot{e},0) = -\mathbf{I_1}(0,\dot{e},0)$.

It is evident that:

(65) $\mathbf{I_2}(0,\dot{e},0) = \int_0^{2\pi} cos^2\varphi(1 - \dot{e}cos\varphi)^{-1}d\varphi = -\dot{e}^{-1}\int_0^{2\pi} cos\varphi[(1 - \dot{e}cos\varphi) - 1](1 - \dot{e}cos\varphi)^{-1}d\varphi =$

$= -\dot{e}^{-1}\int_0^{2\pi} cos\varphi\, d\varphi + \dot{e}^{-1}\int_0^{2\pi} cos\varphi(1 - \dot{e}cos\varphi)^{-1}d\varphi = \dot{e}^{-1}\mathbf{I_1}(0,\dot{e},0)$.

Multiplication of the above equality by $-\dot{e}(u) \neq 0$ gives (64).

3.2.6. Case $n = 0$, $e(u) \neq 0$, $e(u) - \dot{e}(u) = 0 \Rightarrow e(u) = \dot{e}(u) \neq 0$.

The relation (57) can be written in the following form:

(66) $$0 = -e\mathbf{I_0}(e,\dot{e}=e,0) - \mathbf{I_1}(e,\dot{e}=e,0) .$$

We directly compute that:

(67) $$\mathbf{I_1}(e,\dot{e}=e,0) = \int_0^{2\pi} cos\varphi(1+ecos\varphi)^{-2}d\varphi = e^{-1}\int_0^{2\pi}[(1+ecos\varphi)-1](1+ecos\varphi)^{-2}d\varphi =$$

$$= -e^{-1}\int_0^{2\pi}(1+ecos\varphi)^{-2}d\varphi + e^{-1}\int_0^{2\pi}(1+ecos\varphi)d\varphi =$$

$$= -e^{-1}\mathbf{I_0}(e,\dot{e}=e,0) + 2\pi e^{-1}(1-e^2)^{-1/2},$$

where we have used formula **858.525** from [14]. But according to formula **858.535** from the same source [14]:

(68) $$\mathbf{I_0}(e,\dot{e}=e,0) = \int_0^{2\pi}(1+ecos\varphi)^{-2}d\varphi = 2\pi(1-e^2)^{-3/2} = (1-e^2)^{-1}2\pi(1-e^2)^{-1/2},$$

which means that:

(69) $$2\pi(1-e^2)^{-1/2} = (1-e^2)\mathbf{I_0}(e,\dot{e}=e,0) .$$

Substituting this result into (67), we have that:

(70) $$\mathbf{I_1}(e,\dot{e}=e,0) = -e^{-1}\mathbf{I_0}(e,\dot{e}=e,0) + e^{-1}(1-e^2)\mathbf{I_0}(e,\dot{e}=e,0) = -e\,\mathbf{I_0}(e,\dot{e}=e,0) .$$

Hence, (66) is proved.

3.2.7. Case $n = 0$, $e(u) \neq 0$, $e(u) - \dot{e}(u) \neq 0$.

The linear relation (57) can be written as:

(71) $$(e-\dot{e})\mathbf{I_2}(e,\dot{e},0) = -e\mathbf{I_0}(e,\dot{e},0) - [1+e(e-\dot{e})]\mathbf{I_1}(e,\dot{e},0) .$$

Let us calculate at first, using as before, the explicit expressions for the integrals from paper [11] (formulas (3a) – (3c)). In the present paper we have written them as the expressions (48), (49) and (50) for $\mathbf{I_0}(e,\dot{e},0)$, $\mathbf{I_1}(e,\dot{e},0)$, and $\mathbf{I_2}(e,\dot{e},0)$, respectively. From (50) we obtain:

(72) $$(e-\dot{e})\mathbf{I_2}(e,\dot{e},0) = \boldsymbol{A}(e,\dot{e})\{(e-\dot{e})(-1+e^2+e\dot{e})[1-(e-\dot{e})^2]^{1/2} + (e-\dot{e})(1-e^2)^{3/2}\}.$$

Further we evaluate the right-hand side of (71):

(73) $$-e\,\mathbf{I_0}(e,\dot{e},0) - [1+e(e-\dot{e})]\mathbf{I_1}(e,\dot{e},0) = \boldsymbol{A}(e,\dot{e})\{-e^2\dot{e}[1-(e-\dot{e})^2]^{1/2} +$$
$$+ e^2(e-\dot{e})(1-e^2)[1-(e-\dot{e})^2]^{1/2} - e(e-\dot{e})^2(1-e^2)^{3/2} - [1+e(e-\dot{e})](e-\dot{e}-e^3)\times$$
$$\times[1-(e-\dot{e})^2]^{1/2} + [1+e(e-\dot{e})](e-\dot{e})(1-e^2)^{3/2}\} =$$
$$= \boldsymbol{A}(e,\dot{e})\{(-e+e^3+\dot{e}-e\dot{e}^2)[1-(e-\dot{e})^2]^{1/2} + (e-\dot{e})(1-e^2)^{3/2}\} .$$

But $(e-\dot{e})(-1+e^2+e\dot{e}) = -e+e^3+\dot{e}-e\dot{e}^2$ and hence:

(74) $$-e\,\mathbf{I_0}(e,\dot{e},0) - [1+e(e-\dot{e})]\mathbf{I_1}(e,\dot{e},0) =$$
$$= \boldsymbol{A}(e,\dot{e})\{(e-\dot{e})(-1+e^2+e\dot{e})[1-(e-\dot{e})^2]^{1/2} + (e-\dot{e})(1-e^2)^{3/2}\}.$$

The right-hand sides of (72) and (74) coincide and this proves the relation (71). This also completes the validity of (57)in the general case of *integer*/*noninteger* powers *n*, *zero*/*nonzero* values of $e(u)$ (for $|e(u)| < 1$) and *zero*/*nonzero* values of $[e(u) - \dot{e}(u)]$ (for $|e(u) - \dot{e}(u)| < 1$).

3.3. Elimination of the integral $\mathbf{I_1}(e,\dot{e},n)$

In the next derivation we shall use not only definitions (3), but also definition (1) and (2). According to the later and the identity $cos\varphi^2 + sin\varphi^2 = 1$, we have:

$$(75)\qquad \mathbf{I_{0+}}(e,\dot{e},n) = \int_0^{2\pi}(1 + e\cos\varphi)^{n-2}[1 + (e - \dot{e})\cos\varphi]^{-(n+2)}\, d\varphi \;=$$

$$= \int_0^{2\pi}(\cos\varphi^2 + \sin\varphi^2)(1 + e\cos\varphi)^{n-2}[1 + (e - \dot{e})\cos\varphi]^{-(n+2)}\, d\varphi \;=$$

$$= (e - \dot{e})^{-1}\int_0^{2\pi}\cos\varphi[1 + (e - \dot{e})\cos\varphi - 1](1 + e\cos\varphi)^{n-2}[1 + (e - \dot{e})\cos\varphi]^{-(n+2)}\, d\varphi \;-$$

$$-\,(e - \dot{e})^{-1}\int_0^{2\pi}\sin\varphi(1 + e\cos\varphi)^{n-2}[1 + (e - \dot{e})\cos\varphi]^{-(n+2)}\, d[1 + (e - \dot{e})\cos\varphi] \;=$$

$$= \;(e - \dot{e})^{-1}\int_0^{2\pi}\cos\varphi(1 + e\cos\varphi)^{n-2}[1 + (e - \dot{e})\cos\varphi]^{-(n+1)}\, d\varphi \;-$$

$$= -\,(e - \dot{e})^{-1}\int_0^{2\pi}\cos\varphi(1 + e\cos\varphi)^{n-2}[1 + (e - \dot{e})\cos\varphi]^{-(n+2)}\, d\varphi \;+$$

$$+\,[(n + 1)(e - \dot{e})]^{-1}\int_0^{2\pi}\sin\varphi(1 + e\cos\varphi)^{n-2}\, d\{\,[1 + (e - \dot{e})\cos\varphi]^{-(n+1)}\,\} \;=$$

$$= (e - \dot{e})^{-1}\mathbf{I_1}(e,\dot{e},n) - (e - \dot{e})^{-2}\int_0^{2\pi}\{[1 + (e - \dot{e})\cos\varphi] - 1\}(1 + e\cos\varphi)^{n-2} \;\times$$

$$\times\,[1 + (e - \dot{e})\cos\varphi]^{-(n+2)}\, d\varphi = [(n + 1)(e - \dot{e})]^{-1}\{\,\sin\varphi(1 + e\cos\varphi)^{n-2} \;\times$$

$$\times\,[1 + (e - \dot{e})\cos\varphi]^{-(n+1)}\Big|_0^{2\pi} - \int_0^{2\pi}[1 + (e - \dot{e})\cos\varphi]^{-(n+1)}\, d\,[\sin\varphi(1 + e\cos\varphi)^{n-2}]\,\} \;=$$

$$= (e - \dot{e})^{-1}\mathbf{I_1}(e,\dot{e},n) - (e - \dot{e})^{-2}\int_0^{2\pi}(1 + e\cos\varphi)^{n-2}[1 + (e - \dot{e})\cos\varphi]^{-(n+1)}\, d\varphi \;+$$

$$+\,(e - \dot{e})^{-2}\int_0^{2\pi}(1 + e\cos\varphi)^{n-2}[1 + (e - \dot{e})\cos\varphi]^{-(n+2)}\, d\varphi \;-$$

$$- [(n+1)(e-\dot{e})]^{-1}\int_0^{2\pi} cos\varphi(1+ecos\varphi)^{n-2}[1+(e-\dot{e})cos\varphi]^{-(n+1)}\,d\varphi \;+$$

$$+ (n-2)e[(n+1)(e-\dot{e})]^{-1}\int_0^{2\pi}(1-cos^2\varphi)(1+ecos\varphi)^{n-3}[1+(e-\dot{e})cos\varphi]^{-(n+1)}\,d\varphi =$$

$$= (e-\dot{e})^{-1}\mathbf{I_1}(e,\dot{e},n) - (e-\dot{e})^{-2}\mathbf{I_0}(e,\dot{e},n) + (e-\dot{e})^{-2}\mathbf{I_{0+}}(e,\dot{e},n) -$$

$$- [(n+1)(e-\dot{e})]^{-1}\mathbf{I_1}(e,\dot{e},n) + (n-2)e[(n+1)(e-\dot{e})]^{-1}\int_0^{2\pi}(1+ecos\varphi)^{n-3}\times$$

$$\times [1+(e-\dot{e})cos\varphi]^{-(n+1)}\,d\varphi \;-$$

$$- (n-2)e[(n+1)(e-\dot{e})]^{-1}\int_0^{2\pi} cos^2\varphi(1+ecos\varphi)^{n-3}[1+(e-\dot{e})cos\varphi]^{-(n+1)}\,d\varphi\,.$$

The later integral in the right-hand side in the above equality can easily be computed:

(76) $$\int_0^{2\pi} cos^2\varphi(1+ecos\varphi)^{n-3}[1+(e-\dot{e})cos\varphi]^{-(n+1)}\,d\varphi \;=$$

$$= e^{-1}\int_0^{2\pi} cos\varphi[(1+ecos\varphi)-1](1+ecos\varphi)^{n-3}[1+(e-\dot{e})cos\varphi]^{-(n+1)}\,d\varphi \;=$$

$$= e^{-1}\mathbf{I_1}(e,\dot{e},n) - e^{-2}\int_0^{2\pi}[(1+ecos\varphi)-1](1+ecos\varphi)^{n-3}[1+(e-\dot{e})cos\varphi]^{-(n+1)}\,d\varphi \;=$$

$$= e^{-1}\mathbf{I_1}(e,\dot{e},n) - e^{-2}\mathbf{I_0}(e,\dot{e},n) + e^{-2}\mathbf{I_{0-}}(e,\dot{e},n)\,,$$

where we have taken into account the definition (1). Substituting (76) into (75), we arrive at the following linear dependence between the four integrals $\mathbf{I_{0-}}(e,\dot{e},n)$, $\mathbf{I_{0+}}(e,\dot{e},n)$, $\mathbf{I_0}(e,\dot{e},n)$ and $\mathbf{I_1}(e,\dot{e},n)$:

(77) $$0 = [(e-\dot{e})^{-2}-1]\mathbf{I_{0+}}(e,\dot{e},n) + (n-2)[(n+1)(e-\dot{e})]^{-1}(e-e^{-1})\mathbf{I_{0-}}(e,\dot{e},n) +$$
$$+ \{\,(n-2)[(n+1)e(e-\dot{e})]^{-1} - (e-\dot{e})^{-2}\,\}\mathbf{I_0}(e,\dot{e},n) + \{\,(e-\dot{e})^{-1} -$$
$$- [(n+1)(e-\dot{e})]^{-1} - (n-2)[(n+1)(e-\dot{e})]^{-1}\,\}\mathbf{I_1}(e,\dot{e},n)\,.$$

We choose to estimate from the above relation the integral $\mathbf{I_1}(e,\dot{e},n)$. After the multiplication of the both sides of (77) by $(e-\dot{e})$, the result is:

(78) $$2(n+1)^{-1}\mathbf{I_1}(e,\dot{e},n) = (n-2)(1-e^2)[(n+1)e]^{-1}\mathbf{I_{0-}}(e,\dot{e},n) +$$
$$+ [(e-\dot{e})^2-1](e-\dot{e})^{-1}\mathbf{I_{0+}}(e,\dot{e},n) + \{\,(e-\dot{e})^{-1} - (n-2)[(n+1)e]^{-1}\,\}\mathbf{I_0}(e,\dot{e},n)\,.$$

Another multiplication by $(n+1)e(e-\dot{e})$ leads to the next evaluation:

(79) $$2e(e-\dot{e})\,\mathbf{I_1}(e,\dot{e},n) = (n-2)(e-\dot{e})(1-e^2)\,\mathbf{I_{0-}}(e,\dot{e},n) +$$
$$+ (n+1)e[(e-\dot{e})^2-1]\,\mathbf{I_{0+}}(e,\dot{e},n) + [3e+(n-2)\dot{e}]\,\mathbf{I_0}(e,\dot{e},n)\,.$$

We note that (79) is derived under the assumptions $n \neq -1$, $e(u) \neq 0$ (for $|e(u)| < 1$) and $[e(u) - \dot{e}(u)] \neq 0$ (for $|e(u) - \dot{e}(u)| < 1$), which guarantees that the denominators into the expressions (75) – (78) will also be different

from zero. As in the previous cases, concerning the integrals $\mathbf{I_4}(e,\dot{e},n)$ and $\mathbf{I_2}(e,\dot{e},n)$, the derived equality (79) makes sense even if these limitations are not fulfilled for a certain value $u \equiv ln\ p$. We now shall check this statement.

3.3.1. Case $n \neq -1$, $e(u) = 0$, $e(u) - \dot{e}(u) = 0$ $\Rightarrow$ $e(u) = \dot{e}(u) = 0$.

The equality (79) is obviously true, because the both sides are equal to zero.

3.3.2.1. Case $n = 2$, $e(u) = 0$, $e(u) - \dot{e}(u) \neq 0$ $\Rightarrow$ $\dot{e}(u) \neq 0$.

Again (79) has both sides equal to zero.

3.3.2.2. Case $n \neq -1$, $n \neq 2$, $e(u) = 0$, $e(u) - \dot{e}(u) \neq 0$ $\Rightarrow$ $\dot{e}(u) \neq 0$.

The relation (79) in this case can be written as:

(80) $$0 = -(n-2)\dot{e}\mathbf{I_{0-}}(0,\dot{e},n) + (n-2)\dot{e}\mathbf{I_0}(0,\dot{e},n)\ .$$

Because $(n-2)\dot{e} \neq 0$, we may divide by this quantity, to obtain:

(81) $$\mathbf{I_{0-}}(0,\dot{e},n) - \mathbf{I_0}(0,\dot{e},n) = 0\ .$$

Using definitions (1) and (3), direct computation shows that:

(82) $$\mathbf{I_{0-}}(0,\dot{e},n) = \int_0^{2\pi} (1 - \dot{e}cos\varphi)^{-(n+1)}\, d\varphi = \mathbf{I_0}(0,\dot{e},n)\ .$$

Hence, equality (81) is trivially proved and the same is true for (79).

3.3.3. Case $n \neq -1$, $e(u) \neq 0$, $e(u) - \dot{e}(u) = 0$, $\Rightarrow$ $\dot{e}(u) = e(u) \neq 0$.

The linear relation (79) can be written in the following way:

(83) $$0 = -(n+1)e\mathbf{I_{0+}}(e,\dot{e}=e,n) + (n+1)\dot{e}\mathbf{I_0}(e,\dot{e}=e,n)\ .$$

Because $\dot{e}(u) = e(u) \neq 0$ and $(n+1) \neq 0$,we can divide the both sides by $(n+1)e$:

(84) $$\mathbf{I_0}(e,\dot{e}=e,n) - \mathbf{I_{0+}}(e,\dot{e}=e,n) = 0\ .$$

From definitions (2) and (3) we directly compute that:

(85) $$\mathbf{I_0}(e,\dot{e}=e,n) = \int_0^{2\pi} (1 + ecos\varphi)^{n-2}\, d\varphi = \mathbf{I_{0+}}(e,\dot{e}=e,n)\ .$$

That is, equality (84) and, hence, the linear relation (79) are also true in that case.

3.3.4. Case $n = -1$, $e(u) = 0$, $e(u) - \dot{e}(u) = 0$ $\Rightarrow$ $\dot{e}(u) = 0$.

The equality (79) now becomes $0 = 0$ and it is trivially fulfilled.

3.3.5. Case $n = -1$, $e(u) = 0$, $e(u) - \dot{e}(u) \neq 0 \Rightarrow \dot{e}(u) \neq 0$.

The relation (79) now becomes:

(86) $0 = 3\dot{e}\mathbf{I_0}_{\text{-}}(0,\dot{e},-1) - 3\dot{e}\mathbf{I_0}(0,\dot{e},-1)$, or $\mathbf{I_0}_{\text{-}}(0,\dot{e},-1) = \mathbf{I_0}(0,\dot{e},-1)$.

We have $\mathbf{I_0}(0,\dot{e},-1) = \int_0^{2\pi} d\varphi = 2\pi$ and $\mathbf{I_0}_{\text{-}}(0,\dot{e},-1) = \int_0^{2\pi} d\varphi = 2\pi$. Therefore, (86) and correspondingly (79) are satisfied.

3.3.6. Case $n = -1$, $e(u) \neq 0$, $e(u) - \dot{e}(u) = 0 \Rightarrow \dot{e}(u) = e(u) \neq 0$.

In this particular case, the both sides of the equality (79) are equal to zero.

3.3.7. Case $n = -1$, $e(u) \neq 0$, $e(u) - \dot{e}(u) \neq 0$.

The relation (79) now can be written as:

(87) $2e(e - \dot{e})\mathbf{I_1}(e,\dot{e},-1) = -3(e - \dot{e})(1 - e^2)\mathbf{I_0}_{\text{-}}(e,\dot{e},-1) + 3(e - \dot{e})\mathbf{I_0}(e,\dot{e},-1)$.

Because $(e - \dot{e}) \neq 0$, we may divide (87) by $(e - \dot{e})$ and check the validity of the equality:

(88) $2e\mathbf{I_1}(e,\dot{e},-1) = -3(1 - e^2)\mathbf{I_0}_{\text{-}}(e,\dot{e},-1) + 3\mathbf{I_0}(e,\dot{e},-1)$.

In an earlier paper [11], we have already computed for $n = -1$ in an explicit form the following analytical expressions for the integrals $\mathbf{I_0}(e,\dot{e},-1)$, $\mathbf{I_1}(e,\dot{e},-1)$ and $\mathbf{I_0}_{\text{-}}(e,\dot{e},-1)$ (see in [11] formulas (2a), (2b) and (2h), respectively):

(89) $\mathbf{I_0}(e,\dot{e},-1) = \pi(2 + e^2)(1 - e^2)^{-5/2}$,

(90) $\mathbf{I_1}(e,\dot{e},-1) = -3\pi e(1 - e^2)^{-5/2}$,

(91) $\mathbf{I_0}_{\text{-}}(e,\dot{e},-1) = \pi(2 + 3e^2)(1 - e^2)^{-7/2}$.

Consequently:

(92) $2e\mathbf{I_1}(e,\dot{e},-1) = -6\pi e^2(1 - e^2)(1 - e^2)^{-7/2}$,

(93) $-3(1 - e^2)\mathbf{I_0}_{\text{-}}(e,\dot{e},-1) + 3\mathbf{I_0}(e,\dot{e},-1) = -3\pi(2 + 3e^2)(1 - e^2)(1 - e^2)^{-7/2} + 3\pi(2 + e^2)(1 - e^2)(1 - e^2)^{-7/2} = -6\pi e^2(1 - e^2)(1 - e^2)^{-7/2}$.

The right-hand sides of (92) and (93) are equal, and, hence, the linear relation (87) is proved. Thus, the reliability of the linear relation (79) is shown to remain valid in the general case of *integer*/*noninteger* powers n, $e(u)$ *equal* or *not equal* to zero (for $|e(u)| < 1$) and $[e(u) - \dot{e}(u)]$ *equal* or *not equal* to zero (for $|e(u) - \dot{e}(u)| < 1$) for *arbitrary* values of $u \equiv ln\ p$. These

conditions may be separately or simultaneously encountered, and the equality (79) may be used without specifying any restrictions, like the above considered.

4. Conclusions

In a series of papers ([10] – [13]), we have investigated the dynamical equation (4), governing the structure of the ***stationary elliptical*** discs in the model developed by Lyubarskij et al. [4]. Our main goal is to reveal the properties of this second order ordinary differential equation in a fully (as far as possible) analytical manner, without introducing any additional simplifications into the model, except these which already exist in the original development [4]. The first successful step in this direction was the establishment that the dynamical equation is (in the most general case) a homogeneous differential equation [12]. The next step in the simplification of the equation was to eliminate ***four*** among the ***seven*** integrals, entering as functions of $e(u)$, $\dot{e}(u)$ and n into equation (4). *In the **present** paper we pointed out how to do so with **three** of them, namely*: $\mathbf{I_4}(e,\dot{e},n)$, $\mathbf{I_2}(e,\dot{e},n)$ and $\mathbf{I_1}(e,\dot{e},n)$. *The elimination of the **fourth** integral* $\mathbf{I_0}(e,\dot{e},n)$ *will be considered in a **forthcoming** paper*. As a final result, they may be represented, by means of ***linear*** relations, through two integrals, namely: $\mathbf{I_{0-}}(e,\dot{e},n)$ and $\mathbf{I_{0+}}(e,\dot{e},n)$. The later two integrals may be shown to be linearly independent functions on $e(u)$ and $\dot{e}(u)$ for every fixed (physically reasonable) value of the power n in the viscosity law $\eta = \beta\,\Sigma^n$. This statement *will also be proved in a **forthcoming** paper*. The problem with the integral $\mathbf{I_3}(e,\dot{e},n)$ ***still remains unresolved***. It is unclear are the three integrals (***considered together***) $\mathbf{I_{0-}}(e,\dot{e},n)$, $\mathbf{I_{0+}}(e,\dot{e},n)$ and $\mathbf{I_3}(e,\dot{e},n)$ linearly independent, or, opposite, the later integral can also be expressed as a linear combination of $\mathbf{I_{0-}}(e,\dot{e},n)$ and $\mathbf{I_{0+}}(e,\dot{e},n)$. This matter relates to the main aim of our investigations. Namely, to express the dynamical equation (4) as a sum of several terms, each factorized as a product of one of these ***two*** (or, may be, ***three***) ***linearly independent*** integrals and coefficients, which are functions on $e(u)$, $\dot{e}(u)$ and n. The linear independence would imply nullification of the coefficients. This leads to splitting of the equation (4) into a system of *probably* more simple differential equations about the ***unknown*** function $e(u)$.

References

1. J u a n, F., F. X i e, J. P. O s t r i k e r. Global compton heating and cooling in hot accretion flows., Astrophys. J., **691**, 2009, № 1, part 1, pp. 98–104.
2. 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, pp. 337–355.
3. C a o, X., J.-D. X u. Radiation properties of an accretion disk with a non-zero torque on its inner edge., Publ. Astron. Soc. Japan, **55**, 2003, pp. 149–154.
4. L y u b a r s k i j, Y u. 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, pp. 583–596.
5. R i c e, W. K. M., P. J. A r m i t a g e. Time-dependent models of the structure and stability of self-gravitating protoplanetary discs., Monthly Not. Royal Astron. Soc., **396**, 2009, № 4, pp. 2228–2236.
6. V u r m, I., J. P o u t a n e n. Time-dependent modeling of radiative processes in hot magnetized plasmas., Astrophys. J., **698**, 2009, № 1, part 1, pp. 293–316.
7. E s p a i l l a t, C., N. C a l v e t, K. L. L u h m a n, P. D'A l e s s i o. Confirmation of a gapped primordial disk around LkCa 15., Astrophys. J. Letters, **682**, 2008, № 2, part 2, pp. L125–L128.
8. L i t h w i c k, J. Formation, survival, and destruction of vortices in accretion disks., Astrophys. J., **693**, 2009, № 1, part 1, pp. 85–96.
9. H u b b a r d, A., E. G. B l a c k m a n. Identifying deficiencies of standard accretion disc theory., Monthly Not. Royal Astron. Soc., **390**, 2008, № 1, pp. 331–335.
10. D i m i t r o v, D. V. One possible simplification of the dynamical equation governing the evolution of elliptical accretion discs., Aerospace Research in Bulgaria, **17**, 2003, p.17–22.
11. D i m i t r o v, D. V. 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**, 2006, pp.16–28.
12. D i m i t r o v, D. V. Thin viscous elliptical accretion discs with orbits sharing a common longitude of periastron. IV. Proof of the homogeneity of the dynamical equation, governing of the disc structure, for arbitrary powers n in the viscosity law $\eta = \beta \Sigma^n$., Aerospace Research in Bulgaria, **24**, 2010.
13. D i m i t r o v, D. V. 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, pp. 7–23.
14. D w i g h t, H. B. Tables of integrals and other mathematical data., Fourth edition, New York, MacMillan company, 1961.

ТЪНКИ ВИСКОЗНИ ЕЛИПТИЧНИ АКРЕЦИОННИ ДИСКОВЕ С ОРБИТИ, ИМАЩИ ОБЩА ДЪЛЖИНА НА ПЕРИАСТРОНА. V. ЛИНЕЙНИ ЗАВИСИМОСТИ МЕЖДУ УСРЕДНЕНИТЕ ПО АЗИМУТАЛНИЯ ЪГЪЛ МНОЖИТЕЛИ В ДИНАМИЧНОТО УРАВНЕНИЕ

Д. Димитров

Резюме

Ние разглеждаме модел на ***стационарни*** елиптични акреционни дискове, разработен от Любарски и др. [4], които са получили обикновено диференциално уравнение от втори ред, описващо пространствената структура на тези обекти. Това динамично уравнение съдържа седем интеграли, възникващи при азимуталното усредняване по протежение на елиптичните орбити на частиците от диска. Те са функции на ***неизвестното*** разпределение на ексцентрицитета $e(u)$, неговата производна $\dot{e}(u) \equiv de(u)/du$ и степенния показател n в закона за вискозитета $\eta = \beta \Sigma^n$, където $u \equiv \ln p$, p е фокалният параметър на конкретната елиптична орбита на частицата. В настоящата статия, ние извеждаме линейни зависмости между тези ***неизвестни*** интеграли, които могат да бъдат полезни за елиминирането на ***три*** от тези величини. Възможно е да бъде елимиран допълнително още един интеграл, но доказването на това твърдение ще бъде отложено в една предстояща статия. Разгледаният подход е поддържан с цел да се разцепи динамичното уравнение на една система от по-прости диференциални уравнения.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 24, 2012, Sofia

THIN VISCOUS ELLIPTICAL ACCRETION DISCS WITH ORBITS SHARING A COMMON LONGITUDE OF PERIASTRON. VI. SIMPLIFICATION OF THE DYNAMICAL EQUATION

Dimitar Dimitrov

Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: dim@mail.space.bas.bg

Abstract

We continue the series of papers, devoted to the investigation and simplification of the dynamical equation, governing the structure of the ***stationary*** *elliptical accretion discs. These studies are in the frameworks, specified according to the model of Lyubarskij et al. [7]. In addition to the previous examinations, we find one more linear relation between the coefficients of this second order ordinary differential equation, which enables us to eliminate effectively at least four of them. This is in the course of our approach to reduce the number of these functions, depending on the eccentricity, its derivative and the power* $\boldsymbol{n}$ *in the viscosity law* $\boldsymbol{\eta = \beta \Sigma^{n}}$*. They appear in the equation during the process of averaging (i.e. integrating) over the azimuthal angle of the elliptical orbits. At the present stage of the investigations, there still remain three integrals of the indicated type. Except the case of integer values of* $\boldsymbol{n}$*, their analytical solutions are not known. In connection with the linear dependence or independence of these functions (this is a subject of our future studies), the dynamical equation of the elliptical accretion discs may be split into a system of corresponding number of more simple equations about the unknown eccentricities of the particle orbits. Such an approach is in accordance with our base line, carried out through the referred series of papers, to make a progress, as much as possible, into the solving of the task by means of* ***purely analytical*** *methods. And only when the further advance in this way (if the final solution is not already attained) is so complicated, that it is impasse, to use numerical simulations.*

1. Introduction

There are numerous observational examples of accretion discs, orbiting around one of the components in a binary stellar system. The observational evidences, proving the existence of such astronomical objects, are based mainly on the accretion processes on to the compact object, located at the center of the disc. It may happen that the supply of mass to the central object can vary for different physical and geometrical reasons. This is possible to occur during the state transitions of the accretion flow or due to the variable interactions in the eccentric binary system. Such an activity, in principle, is a subject of damping by the viscosity of the disc matter. The mass supply on to the primary star (black hole, neutron star, white dwarf) is, in the end, determined by the mass supply at the outer disc edge. The authors of the investigation [1] compare this physical process for two accretion disc models: such as with finite and with infinite sizes. They find significant differences between these two cases. Namely, the infinite disc solution overestimates the viscous damping. They conclude that, generally speaking, the damping becomes very strong when the viscous time at the outer edge of the disc turns longer than the modulation time scale [1]. We consider the above example, in order to underline how important may be the right evaluation even of a *single* parameter, when we describe the accretion flow events.

The variations of the properties of the accretion discs are subject not only to *internal* changes of their parameters, but also as consequences of *external* influences on the stellar system. The interactions between a “star + accretion disc” system and another star will perturb the disc, possibly resulting in significant modifications of the disc structure and its physical properties. It is suggested that such encounters are capable to trigger fragmentation of the disc, to form brown dwarfs or gas giant planets [2]. In the later paper are simulated star-star encounters, where the primary star has a self-gravitating, marginally stable protostellar disc, and the secondary star has not disc. The results of this investigation of the variations of the disc structure and its dynamics may be summarized in the following way. The stellar encounter is to prohibit the fragmentation, because compressive and shock heating stabilizes the disc and the radiative cooling is insufficient to trigger the gravitational instability. The conclusions from these simulations [2] show, that the encounter strips the outer regions of the disc. This can be realized either by tidal tails or by a capture of matter to form a disc around the secondary star. As a final result, the interaction triggers a readjustment if

the initially existing primary accretion disc turns to such with a steeper surface density profile. We conclude from such studies, that accretion discs have not only good possibilities to survive the encounters with the other stars, but also to preserve their relatively smooth spatial distribution of different physical parameters, characterizing the global accretion flow. It is worthy to note, that the *fractal* concepts have recently been introduced in the accretion disc theory as a new feature. As pointed out by Roy & Ray [3], due to the fractal nature of the flow, its continuity condition undergoes modifications. They show (through completely analytical solutions of the equilibrium point conditions), that their results give indications, that the *fractal properties enable the flow to behave like an effective continuum of lesser density*. The mass accretion rate exhibits a fractal scaling behavior, and the entire fractal accretion disc is stable under *linearized* dynamic perturbations.

As we know, the accretion discs in the binary stellar systems are frequently occurred objects and naturally arises the question about the relation of the shape of the disc and the eccentricity of the binary orbit. Marzari et al. [4] study the evolution of circumstellar massive discs around the primary star of a binary system. Especially, they concentrate on the computation of the disc eccentricity and its dependence on the binary orbit eccentricity. The conclusions are that the self-gravitation of the massive discs leads to discs that have (on average) low eccentricity. They establish that the orientation of the disc, computed with the standard dynamical method, always librates, instead of circulating as in the simulations *without* self-gravitation. The simulations show that the accretion disc eccentricity decreases with the binary eccentricity. This result is found also in models *without* self-gravitation. Generally speaking, the investigation [4] is in agreement with the statement that the disc self-gravitation appears to be an important factor in determining the evolution of the *massive* accretion discs in the binary systems. One additional complicating factor, which possibly affects the shape of the accretion flow, is its orientation with respect to the spin axis of the central body. Modelling of the overall shape of an accretion disc in a semidetached binary system is performed, for example, in the paper of Martin et al. [5]. In this investigation, the mass is transferred on to a spinning black hole, which spin axis is misaligned with the orbital rotation axis of the binary. It is assumed that the accretion disc around the black hole is in a *steady* state. It turns out, that its outer regions are subject to differential precession, caused by the tidal torques of the companion star.

The later tend to align the outer parts of the disc with the orbital plane. But the inner regions of the disc are subject to differential precession caused by the general relativity action (Lense-Thirring effect). Such an influence tends to align the rotation of the inner parts of the accretion disc with the spin of the black hole. There are many other examples, both theoretical and observational, illustrating that the disc midplane may be inclined relative to the binary orbital plane. Under suitable physical conditions, similar geometrical configuration is expected to induce warping and rigid body precession of the disc. Fragner & Nelson [6] find that thick discs (with aspect ratio = 0.5 and low viscosity parameter α) precess as rigid bodies with very little warping or twisting. They are observed to align with the binary orbital plane on the viscous evolution time. On the other hand, thinner discs with higher viscosity, in which warp communication is less efficient, develop significant twists, before to achieve a state of rigid body precession. Under the most extreme conditions considered in [6] (with aspect ratio = 0.01; $\alpha = 0.1$ and $\alpha = 0.005$), it is established that the accretion discs can become broken or disrupted by the strong differential precession. Discs that become highly twisted are observed to align with the binary orbital plane on time scales much shorter than the viscous time scale and, possibly, on the precession time [6]. These examples, concerning the complicated internal and external interactions in the accretion flows, demonstrate some of the difficulties, which may occur, when the shape of the disc must be established in a quantitative manner. In the present paper, we investigate a particular accretion disc model, having an elliptical shape. The trajectories of the disc particles are assumed *ad hoc* to be ellipses, sharing a common longitude of periastron. More precisely, the dynamical equation, with which we shall deal, describes and governs the structure of the disc in the model developed by Lyubarskij et al. [7], and which generalizes the standard α-disc model of Shakura & Sunyaev [8]. These two papers do not involve in their considerations strongly magnetized accretion flows. Taking into account such a circumstance, it is a matter of interest to mention the recent study of Murphy et al. [9], devoted to the *large-scale* magnetic fields in the viscous resistive accretion discs. According to the theory of the winds from cold steady-state discs, having near Keplerian velocity field, there is a necessity for a large-scale magnetic field at near equipartition strength to present. However, as mentioned in [9], this required minimum magnetization (for these disc models) has never been tested with time dependent simulations. In order to eliminate this omission, Murphy et

al. [9] investigate the time evolution of a Shakura-Sunyaev accretion disc [8], threaded by a weak magnetic field. Its strength is such that the disc magnetization falls off rapidly with the radius. The results lead to the conclusion that the *large-scale* magnetic field introduces only a small perturbation to the disc structure and the accretion remains driven by the dominant viscous torque. However, their numerical simulations reveal that a superfast magnetosonic jet is observed to be launched from the innermost regions of the disc and *continues to be a stationary event for a long time*. The main conclusions, following from these numerical simulations, are that the astrophysical discs with superheated surface layers could drive analogous outflows, even if their midplane magnetization is low. In order the accretion process to proceed, the turbulent viscosity must extract a sufficient angular momentum. The authors of the investigation [9] conclude that the magnetized outflows are no more than byproduct, rather than a fundamental driver of the accretion. Nevertheless, if the midplane magnetization increases towards the center of the accretion disc, a natural transition to an inner *jet-dominated* disc could eventually be achieved. We shall pick out a little more attention to the important process of the angular momentum transfer in the accretion flows.

2. Mechanisms of angular momentum transport

It is believed that the microphysical viscosity is too small to produce the observed protoplanetary accretion disc lifetimes. Instead of that, it is suggested a new approach, based on the turbulent transport. In the later case, the turbulent motion takes the place of thermal motion. Though the source of such turbulence remains a matter of discussion, this process can provide the correct order of magnitude of the observed accretion rates in these objects for reasonable surface densities [10]. It may happen that the main accretion mechanism is not the turbulent viscosity, as can be seen in the situation with the magnetorotational instability. According to the numerical simulations performed in [10], the requirement for energy conservation is a significant constraint on the accretion driving processes, such as the magnetorotational instability. The mechanism of angular momentum transport in accretion discs is debated for a long time. In this stream of investigations, it should be noted that although the magnetorotational instability appears to be a promising explanation of the accretion events, in the poorly ionized regions of accretion discs there may not be favorable conditions for this instability to operate. In the research [11] is revisited the

possibility of transporting angular momentum by thermal *convection*. There is shown that strongly turbulent convection can drive outward angular momentum transport at a rate that is (under certain conditions) compatible with the observed rate in the discs. The results of Lesur & Ogilvie [11] are indicative that *convection* might be another way to explain ***global*** disc evolution. Such a scenario will be realistic provided that a sufficiently unstable *vertical* temperature profile can be maintained.

Another recent research, devoted to the role of magnetic field in the angular momentum transport, is performed in the paper [12]. The physical situation (advection-dominated accretion flow with a toroidal magnetic field) and the geometric configuration (quasi-spherical accretion flow) are much different in comparison with the standard disc model of Shakura & Sunyaev [8], and also from the model Lyubarskij et al. [7]. Nevertheless, it is worthy to note, that the conclusions in [12] may have, to some extent, a reference to the former two models. In the research of Khesali & Faghei [12] it is assumed that (like in [7] and [8]), the angular momentum transport is due to the viscous turbulence and the α-prescription is used for the kinematic coefficient of viscosity. In this paper [12], a self-similar solution is used, in order to solve the equations that govern the dynamical behaviour of the accretion flow. According to the conclusions of Khesali & Faghei [12], their solution provides some insights into the dynamics of *quasi-spherical* accretion flows and avoids many of the strictures of *steady* self-similar solutions. The results in [12] show that the behaviour of the physical quantities in a ***dynamical*** advection-dominated accretion flow is different from that for a ***steady*** accretion flow or a disc using a polytropic approach. This model also implies that the flow has a differential rotation that is a sub-Keplerian at small radii and super-Keplerian at large radii. Such different results are also obtained if a polytropic accretion flow is used. Also, the behaviour of the advection-dominated flows in the presence of a large toroidal magnetic field implies that different results are obtained using *steady-state* self-similar models in contrast to the *dynamical* case. The above remarks have to be referred/related (in some conditional and restricted sense) to the ***steady-state*** case of the model of Lyubarskij et al. [7], which dynamical equation we are going to simplify further in the present paper. The same note is also significant in view of the fact that the classical models [7] and [8] do not involve considerations of large-scale magnetic fields. Restricting our attention only to the particular case of steady-state elliptical accretion discs with orbits sharing a common

longitude of periastron, we shall preliminary eliminate from our treatment the more unusual so called “mini-discs”. It is believed that they arise due to the accretion on to black holes in wind-fed binaries and collapsars. Formation of such small rotating discs is combined with some peculiar properties. They accrete on the free-fall time-scale, *without the help of the viscosity*, and, nevertheless, they can have a high radiative efficiency [13]. We shall not, of course, apply our results to these inviscid “mini-discs”. In principle, in the nature may exist even more “exotic” accretion discs. As pointed out by Zhang et al. [14], the temperature of the hot accretion flows around black holes is sufficiently high for the ignition of nuclear reactions. In the usual studies of the hot accretion flows, the viscous dissipation is considered as the only heating mechanism. In the same time, the heating caused by nuclear reactions is not considered at all. The calculations of Zhang et al. [14] indicate that the energy generation rate of nuclear reactions is *at most* one per cent of the viscous heating. Consequently, they are rather not important and the dynamics of the accretion flow can be calculated in the usual way, without the need to consider the heating due to the nuclear reactions.

3. Definitions and notations

With a view to be more explicit in our further exposition, we shall rewrite briefly some of the definitions and notations introduced and used in our earlier papers, dealing with the same problem. For more detailed descriptions and comments on this theme, the reader is directed to the paper ([15], section 2) and the references therein. We introduce the independent variable $u \equiv ln\ p$, where p is the focal parameter of the ellipse, approximating the orbit of the considered disc particle. The eccentricity of the ellipse is denoted by $e \equiv e(u)$, understanding that the orbits of the particles, belonging to different regions of the accretion disc, may, generally speaking, have different shapes/elongations. Further we use the notation $\dot{e} \equiv \dot{e}(u) \equiv de(u)/du \equiv de/d(ln\ p)$ for the ordinary derivative of the eccentricity $e(u)$. We shall consider the viscosity law $\eta = \beta\ \Sigma^{n}$, with η – viscosity parameter, β is a constant, Σ is the surface density of the accretion disc. The power n is assumed to be a constant for every examined case. In paper [15] are established several linear relations between the following integrals, obtained after the averaging over the azimuthal angle φ:

(1) $$\mathbf{I_{0\text{-}}}(e,\dot{e},n) \equiv \int_0^{2\pi} (1 + e\cos\varphi)^{n-3}[1 + (e - \dot{e})\cos\varphi]^{-(n+1)}\, d\varphi\ ,$$

(2) $$\mathbf{I_{0+}}(e,\dot{e},n) \equiv \int_0^{2\pi} (1 + e\cos\varphi)^{n-2}[1 + (e - \dot{e})\cos\varphi]^{-(n+2)}\, d\varphi\ ,$$

(3) $$\mathbf{I_j}(e,\dot{e},n) \equiv \int_0^{2\pi} (\cos\varphi)^{j}(1 + e\cos\varphi)^{n-2}[1 + (e - \dot{e})\cos\varphi]^{-(n+1)}\, d\varphi\ ;\ \mathbf{j} = 0,$$

1, 2, 3, 4.

In the earlier paper [15], we have shown how the integrals $\mathbf{I_4}(e,\dot{e},n)$, $\mathbf{I_2}(e,\dot{e},n)$ and $\mathbf{I_1}(e,\dot{e},n)$ can be expressed through linear relations of the integrals $\mathbf{I_0}(e,\dot{e},n)$, $\mathbf{I_{0\text{-}}}(e,\dot{e},n)$, and $\mathbf{I_{0+}}(e,\dot{e},n)$. Our present aim is to extend this approach, finding a linear relation between the last three integrals, namely, to determine how $\mathbf{I_0}(e,\dot{e},n)$ may be written out as a linear combination of $\mathbf{I_{0\text{-}}}(e,\dot{e},n)$ and $\mathbf{I_{0+}}(e,\dot{e},n)$.

4. Elimination of the integral $\mathbf{I_0}(e,\dot{e},n)$

We have already obtained the following relation between the integrals $\mathbf{I_2}(e,\dot{e},n)$, $\mathbf{I_0}(e,\dot{e},n)$ and $\mathbf{I_1}(e,\dot{e},n)$ ([15], formula (57)):

(4) $$[e + (n-1)\dot{e}]\mathbf{I_2}(e,\dot{e},n) = (-e + n\dot{e})\mathbf{I_0}(e,\dot{e},n) - [1 + e(e - \dot{e})]\mathbf{I_1}(e,\dot{e},n)\ .$$

To resolve the present problem, we begin with an another derivation of the integral $\mathbf{I_2}(e,\dot{e},n)$. As before, we suppose that $ne(u)[e(u) - \dot{e}(u)] \neq 0$ for the considered value of $u \equiv \ln p$. The particular cases, when this condition is violated, will be considered separately. According to definition (3):

(5) $$\mathbf{I_2}(e,\dot{e},n) = \int_0^{2\pi} \cos^2\varphi(1 + e\cos\varphi)^{n-2}[1 + (e - \dot{e})\cos\varphi]^{-(n+1)}\, d\varphi =$$

$$= \int_0^{2\pi} (1 - \sin^2\varphi)(1 + e\cos\varphi)^{n-2}[1 + (e - \dot{e})\cos\varphi]^{-(n+1)}\, d\varphi =$$

$$= \int_0^{2\pi} (1 + e\cos\varphi)^{n-2}[1 + (e - \dot{e})\cos\varphi]^{-(n+1)}\, d\varphi\ +$$

$$+ (e - \dot{e})^{-1} \int_0^{2\pi} \sin\varphi(1 + e\cos\varphi)^{n-2}[1 + (e - \dot{e})\cos\varphi]^{-(n+1)}\, d[1 + (e - \dot{e})\cos\varphi] =$$

$$= \mathbf{I_0}(e,\dot{e},n) - [n(e-\dot{e})]^{-1}\int_0^{2\pi} sin\varphi(1+ecos\varphi)^{n-2} d\{[1+(e-\dot{e})cos\varphi]^{-n}\} = \mathbf{I_0}(e,\dot{e},n) -$$

$$- [n(e-\dot{e})]^{-1}\{\, sin\varphi(1+ecos\varphi)^{n-2}[1+(e-\dot{e})cos\varphi]^{-n}\Big|_0^{2\pi} -$$

$$- \int_0^{2\pi} cos\varphi(1+ecos\varphi)^{n-2}[1+(e-\dot{e})cos\varphi]^{-n} d\varphi +$$

$$+ (n-2)e\int_0^{2\pi} sin^2\varphi(1+ecos\varphi)^{n-3}[1+(e-\dot{e})cos\varphi]^{-n}\, d\varphi\,\} = \mathbf{I_0}(e,\dot{e},n) +$$

$$+ [n(e-\dot{e})]^{-1}\int_0^{2\pi}[1+(e-\dot{e})cos\varphi]\, cos\varphi(1+ecos\varphi)^{n-2}[1+(e-\dot{e})cos\varphi]^{-(n+1)} d\varphi \; -$$

$$- (n-2)e[n(e-\dot{e})]^{-1}\int_0^{2\pi}[1+(e-\dot{e})cos\varphi](1-cos^2\varphi)(\,1+ecos\varphi)^{n-3} \times$$

$$\times [1+(e-\dot{e})cos\varphi]^{-(n+1)} d\varphi \; = \mathbf{I_0}(e,\dot{e},n) + [n(e-\dot{e})]^{-1}\mathbf{I_1}(e,\dot{e},n) + n^{-1}\mathbf{I_2}(e,\dot{e},n) \; -$$

$$- (n-2)e[n(e-\dot{e})]^{-1}\int_0^{2\pi}(1-cos^2\varphi)(1+ecos\varphi)^{n-3}[1+(e-\dot{e})cos\varphi]^{-(n+1)} d\varphi \; -$$

$$- (n-2)en^{-1}\int_0^{2\pi}(1-cos^2\varphi)cos\varphi(1+ecos\varphi)^{n-3}[1+(e-\dot{e})cos\varphi]^{-(n+1)} d\varphi \; =$$

$$= \mathbf{I_0}(e,\dot{e},n) + [n(e-\dot{e})]^{-1}\mathbf{I_1}(e,\dot{e},n) + n^{-1}\mathbf{I_2}(e,\dot{e},n) \; -$$

$$- (n-2)e[n(e-\dot{e})]^{-1}\int_0^{2\pi}(1+ecos\varphi)^{n-3}[1+(e-\dot{e})cos\varphi]^{-(n+1)} d\varphi \; +$$

$$+ (n-2)e[n(e-\dot{e})]^{-1}\int_0^{2\pi} cos^2\varphi(1+ecos\varphi)^{n-3}[1+(e-\dot{e})cos\varphi]^{-(n+1)} d\varphi \; -$$

$$- (n-2)en^{-1}\int_0^{2\pi} cos\varphi(1+ecos\varphi)^{n-3}[1+(e-\dot{e})cos\varphi]^{-(n+1)} d\varphi \; +$$

$$+ (n-2)en^{-1}\int_0^{2\pi} cos^3\varphi(1+ecos\varphi)^{n-3}[1+(e-\dot{e})cos\varphi]^{-(n+1)} d\varphi \; =$$

$$= \mathbf{I_0}(e,\dot{e},n) + [n(e-\dot{e})]^{-1}\mathbf{I_1}(e,\dot{e},n) + n^{-1}\mathbf{I_2}(e,\dot{e},n) - (n-2)e[n(e-\dot{e})]^{-1}\mathbf{I_{0\text{-}}}(e,\dot{e},n) +$$
$$+ (n-2)e[n(e-\dot{e})]^{-1}[e^{-1}\mathbf{I_1}(e,\dot{e},n) - e^{-2}\mathbf{I_0}(e,\dot{e},n) + e^{-2}\mathbf{I_{0\text{-}}}(e,\dot{e},n)] \; -$$
$$- \; (n-2)en^{-1}[e^{-1}\mathbf{I_0}(e,\dot{e},n) - e^{-1}\mathbf{I_{0\text{-}}}(e,\dot{e},n)] +$$
$$+ (n-2)en^{-1}[e^{-1}\mathbf{I_2}(e,\dot{e},n) - e^{-2}\mathbf{I_1}(e,\dot{e},n) + e^{-3}\mathbf{I_0}(e,\dot{e},n) - e^{-3}\mathbf{I_{0\text{-}}}(e,\dot{e},n)] \,.$$

Computing of the last three integrals in the right-hand-side of the above equation is expressed in the next three equalities:

$$(6)\qquad \int_0^{2\pi} cos\varphi(1+ecos\varphi)^{n-3}[1+(e-\dot{e})cos\varphi]^{-(n+1)}d\varphi \;=$$

$$= e^{-1}\int_0^{2\pi}(1+ecos\varphi-1)(1+ecos\varphi)^{n-3}[1+(e-\dot{e})cos\varphi]^{-(n+1)}d\varphi \;=$$

$$= e^{-1}\mathbf{I_0}(e,\dot{e},n) - e^{-1}\mathbf{I_{0\text{-}}}(e,\dot{e},n)\,,$$

$$(7)\qquad \int_0^{2\pi} cos^2\varphi(1+ecos\varphi)^{n-3}[1+(e-\dot{e})cos\varphi]^{-(n+1)}d\varphi \;=$$

$$= e^{-1}\int_0^{2\pi} cos\varphi(1+ecos\varphi-1)(1+ecos\varphi)^{n-3}[1+(e-\dot{e})cos\varphi]^{-(n+1)}d\varphi \;=$$

$$= e^{-1}\mathbf{I_1}(e,\dot{e},n) - e^{-2}\mathbf{I_0}(e,\dot{e},n) + e^{-2}\mathbf{I_{0\text{-}}}(e,\dot{e},n)\,,$$

$$(8)\qquad \int_0^{2\pi} cos^3\varphi(1+ecos\varphi)^{n-3}[1+(e-\dot{e})cos\varphi]^{-(n+1)}d\varphi \;=$$

$$= e^{-1}\int_0^{2\pi} cos^2\varphi(1+ecos\varphi-1)(1+ecos\varphi)^{n-3}[1+(e-\dot{e})cos\varphi]^{-(n+1)}d\varphi \;=$$

$$= e^{-1}\mathbf{I_2}(e,\dot{e},n) - e^{-2}\mathbf{I_1}(e,\dot{e},n) + e^{-2}\mathbf{I_0}(e,\dot{e},n) - e^{-3}\mathbf{I_{0\text{-}}}(e,\dot{e},n)\,.$$

After some algebra, the expression (5) gives the following result for the integral $\mathbf{I_2}(e,\dot{e},n)$:

$$(9)\qquad n^{-1}\mathbf{I_2}(e,\dot{e},n) = (n-2)\dot{e}(1-e^2)[ne^2(e-\dot{e})]^{-1}\mathbf{I_{0\text{-}}}(e,\dot{e},n) +$$
$$+ n^{-1}\{\,2-(n-2)\dot{e}[e^2(e-\dot{e})]^{-1}\,\}\,\mathbf{I_0}(e,\dot{e},n) +$$
$$+ n^{-1}\{\,(e-\dot{e})^{-1}+(n-2)\dot{e}[e(e-\dot{e})]^{-1}\,\}\,\mathbf{I_1}(e,\dot{e},n)\,.$$

Multiplying this equality by $ne^2(e-\dot{e})$, we shall obtain:

$$(10)\qquad e^2(e-\dot{e})\mathbf{I_2}(e,\dot{e},n) = (n-2)(1-e^2)\dot{e}\mathbf{I_{0\text{-}}}(e,\dot{e},n) + [2e^2(e-\dot{e}) - (n-2)\dot{e}]$$
$$\mathbf{I_0}(e,\dot{e},n) + [e^2 + (n-2)e\dot{e}]\,\mathbf{I_1}(e,\dot{e},n)\,.$$

This is another linear relation, which enables us (like the equality (4)) to eliminate the integral $\mathbf{I_2}(e,\dot{e},n)$. We shall now check the validity of (10) in the cases when the power n, the eccentricity $e(u)$ and its derivative $\dot{e}(u)$ may vanish separately or simultaneously for some value $u \equiv ln\ p$ – a situation that was preliminary excluded in deriving of (10).

4.1. Case $n \neq 0$, $e(u) = 0$, $e(u) - \dot{e}(u) = 0 \Rightarrow e(u) = \dot{e}(u) = 0$.

The relation (10) can be written as the equality 0 = 0, i.e., it is right.

4.2. Case $n \neq 0$, $e(u) = 0$, $e(u) - \dot{e}(u) \neq 0 \Rightarrow \dot{e}(u) \neq 0$.

The equality (10) becomes:

(11) $0 = (n-2)\dot{e}\mathbf{I_0}_{\text{-}}(0,\dot{e},n) - (n-2)\dot{e}\mathbf{I_0}(0,\dot{e},n)$.

If $n = 2$, (11) is trivially fulfilled. If $n \neq 2$, after the division of both sides by $(n-2)\dot{e}$, we must prove that $\mathbf{I_0}_{\text{-}}(0,\dot{e},n) = \mathbf{I_1}(0,\dot{e},n)$. This was already done earlier: see equalities (81) and (82) from paper [15].

4.3. Case $n \neq 0$, $e(u) \neq 0$, $e(u) - \dot{e}(u) = 0 \Rightarrow \dot{e}(u) = e(u) \neq 0$.

We can write (10) in the following way:

(12) $0 = (n-2)(1-e^2)\dot{e}\mathbf{I_0}_{\text{-}}(e,\dot{e}=e,n) - (n-2)\dot{e}\mathbf{I_0}(e,\dot{e}=e,n) + (n-1)e\dot{e}\mathbf{I_1}(e,\dot{e}=e,n)$.

Dividing by $\dot{e} \neq 0$, we arrive at the next formula, which we must to prove, in order to verify (10) in this particular case:

(13) $(n-1)e\,\mathbf{I_1}(e,\dot{e}=e,n) = (n-2)\mathbf{I_0}(e,\dot{e}=e,n) - (n-2)(1-e^2)\mathbf{I_0}_{\text{-}}(e,\dot{e}=e,n)$.

We directly compute that:

(14) $$\mathbf{I_1}(e,\dot{e}=e,n) = \int_0^{2\pi} \cos\varphi(1+e\cos\varphi)^{n-2}\,d\varphi = \int_0^{2\pi}(1+e\cos\varphi)^{n-2}\,d(\sin\varphi) =$$

$$= \sin\varphi(1+e\cos\varphi)^{n-2}\Big|_0^{2\pi} + (n-2)e\int_0^{2\pi}\sin^2\varphi(1+e\cos\varphi)^{n-3}\,d\varphi =$$

$$= (n-2)e\int_0^{2\pi}(1-\cos^2\varphi)(1+e\cos\varphi)^{n-3}\,d\varphi = (n-2)e\int_0^{2\pi}(1+e\cos\varphi)^{n-3}\,d\varphi -$$

$$- (n-2)\int_0^{2\pi}\cos\varphi(1+e\cos\varphi-1)(1+e\cos\varphi)^{n-3}\,d\varphi = (n-2)e\mathbf{I_0}_{\text{-}}(e,\dot{e}=e,n) -$$

$$- (n-2)\int_0^{2\pi}\cos\varphi(1+e\cos\varphi)^{n-2}\,d\varphi + (n-2)\int_0^{2\pi}\cos\varphi(1+e\cos\varphi)^{n-3}\,d\varphi =$$

$$= (n-2)e\mathbf{I_0}_{\text{-}}(e,\dot{e}=e,n) - (n-2)\mathbf{I_1}(e,\dot{e}=e,n) +$$

$$+ (n-2)e^{-1}\int_0^{2\pi}(1+e\cos\varphi-1)(1+e\cos\varphi)^{n-3}\,d\varphi =$$

$$= (n-2)e\mathbf{I_0}_{\text{-}}(e,\dot{e}=e,n) - (n-2)\mathbf{I_1}(e,\dot{e}=e,n) +$$

$$+ (n-2)e^{-1}\int_0^{2\pi}(1 + e\cos\varphi)^{n-2}\,d\varphi - (n-2)e^{-1}\int_0^{2\pi}(1 + e\cos\varphi)^{n-3}\,d\varphi =$$

$$= (n-2)e\mathbf{I_0}_{-}(e,\dot{e}=e,n) - (n-2)\mathbf{I_1}(e,\dot{e}=e,n) + (n-2)e^{-1}\mathbf{I_0}(e,\dot{e}=e,n) -$$
$$- (n-2)e^{-1}\mathbf{I_0}_{-}(e,\dot{e}=e,n) .$$

Multiplying the both sides by $e(u) \neq 0$, we obtain the equality (13), and, correspondingly, the linear relation (10) is proved.

4.4. Case $n = 0$, $e(u) = 0$, $e(u) - \dot{e}(u) = 0 \Rightarrow e(u) = \dot{e}(u) = 0$.

The equality (10) may be written as 0 = 0, and it is obviously fulfilled.

4.5. Case $n = 0$, $e(u) = 0$, $e(u) - \dot{e}(u) \neq 0 \Rightarrow \dot{e}(u) \neq 0$.

Now (10) becomes:

(15) $$0 = -2\dot{e}\mathbf{I_0}_{-}(0,\dot{e},0) + 2\dot{e}\,\mathbf{I_0}(0,\dot{e},0) , \text{ or } \mathbf{I_0}_{-}(0,\dot{e},0) = \mathbf{I_0}(0,\dot{e},0) .$$

This is true, because $\mathbf{I_0}(0,\dot{e},0) = \int_0^{2\pi}(1 - \dot{e}\cos\varphi)^{-1}d\varphi = \mathbf{I_0}_{-}(0,\dot{e},0)$.

4.6. Case $n = 0$, $e(u) \neq 0$, $e(u) - \dot{e}(u) = 0 \Rightarrow \dot{e}(u) = e(u) \neq 0$.

The linear relation(10), which must be proved, in this case becomes:

(16) $$0 = -2(1-e^2)\dot{e}\mathbf{I_0}_{-}(e,\dot{e}=e,0) + 2\dot{e}\mathbf{I_0}(e,\dot{e}=e,0) - e^2\mathbf{I_1}(e,\dot{e}=e,0) .$$

We have already computed (formula (70) from [15], multiplied by $-e^2(u) \neq 0$), that in this case we have:

(17) $$-e^2\mathbf{I_1}(e,\dot{e}=e,0) = e^3\mathbf{I_0}(e,\dot{e}=e,0) .$$

Taking into account that $e(u) = \dot{e}(u)$, we also compute (substituting (17) into (16)):

(18) $$-2(1-e^2)e\mathbf{I_0}_{-}(e,\dot{e}=e,0) + 2e\mathbf{I_0}(e,\dot{e}=e,0) + e^3\mathbf{I_0}(e,\dot{e}=e,0) =$$
$$= -2(1-e^2)e\mathbf{I_0}_{-}(e,\dot{e}=e,0) + e(2+e^2)\mathbf{I_0}(e,\dot{e}=e,0) .$$

For later purposes, we evaluate in an explicit form the integral $\mathbf{I_0}(e,\dot{e}=e,0)$, using the formula **858.538** from Dwight [16]:

(19) $$\mathbf{I_0}(e,\dot{e}=e,0) = \int_0^{2\pi}(1 + e\cos\varphi)^{-2}\,d\varphi = 2\pi[(1-e^2)(1-e^2)^{1/2}]^{-1} .$$

In the considered case, because $\dot{e}(u) = e(u)$, the integral $\mathbf{I_0}(e,\dot{e}=e,0)$ is a function only on one independent variable, namely e. We shall differentiate this integral with respect to e. Having in mind, that differentiation and

integration are linear operations, which sequence of carrying out may be interchanged, we write:

(20) $d\,\mathbf{I_0}(e,\dot{e}=e,0)/de = (d/de)[2\pi(1-e^2)^{-3/2}] = 3e(1-e^2)^{-1}2\pi(1-e^2)^{-3/2} = 3e(1-e^2)^{-1}\mathbf{I_0}(e,\dot{e}=e,0)$,

where we have applied the result (19). From the other hand:

(21) $$d\,\mathbf{I_0}(e,\dot{e}=e,0)/de = (d/de)\int_0^{2\pi}(1+e\cos\varphi)^{-2}\,d\varphi = -2\int_0^{2\pi}\cos\varphi(1+e\cos\varphi)^{-3}\,d\varphi =$$

$$= -2e^{-1}\int_0^{2\pi}(1+e\cos\varphi-1)(1+e\cos\varphi)^{-3}\,d\varphi = -2e^{-1}\int_0^{2\pi}(1+e\cos\varphi)^{-2}\,d\varphi +$$

$$+2e^{-1}\int_0^{2\pi}(1+e\cos\varphi)^{-3}\,d\varphi = -2e^{-1}\mathbf{I_0}(e,\dot{e}=e,0) + 2e^{-1}\mathbf{I_{0\text{-}}}(e,\dot{e}=e,0)\,.$$

Combining (20) and (21) gives:

(22) $-2e^{-1}\mathbf{I_0}(e,\dot{e}=e,0) + 2e^{-1}\mathbf{I_{0\text{-}}}(e,\dot{e}=e,0) = 3(1-e^2)^{-1}\mathbf{I_0}(e,\dot{e}=e,0)$.

Multiplication by $e(1-e^2)$ leads to:

(23) $[-2(1-e^2)-3e^2]\mathbf{I_0}(e,\dot{e}=e,0) = -2(1-e^2)\mathbf{I_{0\text{-}}}(e,\dot{e}=e,0)$.

Another multiplication of the above equality by $e(u) \neq 0$ gives that:

(24) $(2+e^2)e\mathbf{I_0}(e,\dot{e}=e,0) - 2e(1-e^2)\mathbf{I_0}(e,\dot{e}=e,0) = 0$,

that means (taking into account the results(17) and (18)) that (16) is true, and, consequently, the relation (10) is again proved.

4.7. Case $n = 0$, $e(u) \neq 0$, $e(u) - \dot{e}(u) \neq 0$.

The linear relation (10) takes the form:

(25) $e^2(e-\dot{e})\mathbf{I_2}(e,\dot{e},0) = -2(1-e^2)\dot{e}\mathbf{I_{0\text{-}}}(e,\dot{e},0) + [2e^2(e-\dot{e})+2\dot{e}]\mathbf{I_0}(e,\dot{e},0) + (e^2-2e\dot{e})\mathbf{I_1}(e,\dot{e},0)$.

We again shall use the explicit analytical expressions for $n = 0$, derived for the integrals $\mathbf{I_{0\text{-}}}(e,\dot{e},0)$, $\mathbf{I_0}(e,\dot{e},0)$, $\mathbf{I_1}(e,\dot{e},0)$ and $\mathbf{I_2}(e,\dot{e},0)$ ([17], formulas 3a), 3b), 3c) and 3h); see also formulas (48), (49) and (50)in the paper [15]). According to the formula 3h) in [17], we are able to write for $\mathbf{I_{0\text{-}}}(e,\dot{e},0)$ the following expression:

(26) $\mathbf{I_{0\text{-}}}(e,\dot{e},0) = \boldsymbol{A}(e,\dot{e})[2(1-e^2)\dot{e}]^{-1}\{(2e^3-4e^5+2e^7-6e^2\dot{e}+10e^4\dot{e}-4e^6\dot{e}+6e\dot{e}^2 - 5e^3\dot{e}^2+2e^5\dot{e}^2)[1-(e-\dot{e})^2]^{1/2} - 2(e-\dot{e})^3(1-e^2)^{5/2}\}$,

where we have used the notation (47) from paper [15] for $\boldsymbol{A}(e,\dot{e})$:

(27) $\boldsymbol{A}(e,\dot{e}) = 2\pi\dot{e}^{-2}(1-e^2)^{-3/2}[1-(e-\dot{e})^2]^{-1/2}$.

Taking into account the expression [50] from the same paper [15]:

(28) $\mathbf{I_2}(e,\dot{e},0) = \boldsymbol{A}(e,\dot{e})\{(-1+e^2+e\dot{e})[1-(e-\dot{e})^2]^{1/2} + (1-e^2)^{3/2}\}$,

we compute the left-hand-side of the relation (25), which we intent to prove:

(29) $e^2(e-\dot{e})\mathbf{I_2}(e,\dot{e},0) = e^2(e-\dot{e})\boldsymbol{A}(e,\dot{e})\{(-1+e^2+e\dot{e})[1-(e-\dot{e})^2]^{1/2}+(1-e^2)^{3/2}\} =$
$= \boldsymbol{A}(e,\dot{e})\{(-e^3+e^5+e^2\dot{e}-e^3\dot{e}^2)[1+(e-\dot{e})^2]^{1/2}]+e^2(e-\dot{e})(1-e^2)^{3/2}\}$.

Now we begin to evaluate the right-hand-side of (25). That is:

(30) $-2(1-e^2)\dot{e}\mathbf{I_{0\text{-}}}(e,\dot{e},0)+[2e^2(e-\dot{e})+2\dot{e}]\mathbf{I_0}(e,\dot{e},0)+(e^2-2e\dot{e})\mathbf{I_1}(e,\dot{e},0) =$
$= \boldsymbol{A}(e,\dot{e})\{(2e^3-4e^5+2e^7-6e^2\dot{e}+10e^4\dot{e}-4e^6\dot{e}+6e\dot{e}^2-5e^3\dot{e}^2+2e^5\dot{e}^2)\times$
$\times[1-(e-\dot{e})^2]^{1/2}-2(e-\dot{e})^3(1-e^2)^{5/2}\}+$
$+\boldsymbol{A}(e,\dot{e})[2e^2(e-\dot{e})+2\dot{e}]\{e\dot{e}[1-(e-\dot{e})^2]^{1/2}-e(e-\dot{e})(1-e^2)[1-(e-\dot{e})^2]^{1/2}+$
$+(e-\dot{e})^2(1-e^2)^{3/2}\}+$
$+\boldsymbol{A}(e,\dot{e})(e^2-2e\dot{e})\{(e-\dot{e}-e^3)[1-(e-\dot{e})^2]^{1/2}-(e-\dot{e})(1-e^2)^{3/2}\} =$
$= \boldsymbol{A}(e,\dot{e})\{(-2e^3+4e^5-2e^7+6e^2\dot{e}-10e^4\dot{e}+4e^6\dot{e}-6e\dot{e}^2+5e^3\dot{e}^2-2e^5\dot{e}^2)\times$
$\times[1-(e-\dot{e})^2]^{1/2}+(2e^3-2e^5-6e^2\dot{e}+6e^4\dot{e}+6e\dot{e}^2-6e^3\dot{e}^2-2\dot{e}^3+2e^2\dot{e}^3)(1-e^2)^{3/2}+$
$+(2e^4\dot{e}+2e\dot{e}^2-2e^3\dot{e}^2)[1-(e-\dot{e})^2]^{1/2}+$
$+(2e^7-2e^5-2e^2\dot{e}+6e^4\dot{e}-4e^6\dot{e}+2e\dot{e}^2-4e^3\dot{e}^2+2e^5\dot{e}^2)[1-(e-\dot{e})^2]^{1/2}+$
$+(2e^5+2e^2\dot{e}-6e^4\dot{e}-4e\dot{e}^2+6e^3\dot{e}^2+2\dot{e}^3-2e^2\dot{e}^3)(1-e^2)^{3/2}+$
$+(e^3-e^5-3e^2\dot{e}+2e\dot{e}^2+2e^4\dot{e})[1-(e-\dot{e})^2]^{1/2}+(-e^3+3e^2\dot{e}-2e\dot{e}^2)(1-e^2)^{3/2}\} =$
$= \boldsymbol{A}(e,\dot{e})\{(-e^3+e^5-e^3\dot{e}^2+e^2\dot{e})[1-(e-\dot{e})^2]^{1/2}+(e^3-e^2\dot{e})(1-e^2)^{3/2}\}$.

This coincides with the right-hand-side of (29). Consequently, the relations (25) and (10) are proved also for the case $n = 0$, $e(u) \neq 0$ and $e(u) - \dot{e}(u) \neq 0$.

To summarize the situation, we note that we have two linear relations between the integrals $\mathbf{I_2}(e,\dot{e},n)$, $\mathbf{I_1}(e,\dot{e},n)$, $\mathbf{I_0}(e,\dot{e},n)$ and $\mathbf{I_{0\text{-}}}(e,\dot{e},n)$, namely, the equalities (4) and (10). They are both valid for arbitrary integer/noninteger powers n (we consider physically reasonable values of n between – 1 and + 3), arbitrary values of the eccentricity $e(u)$ (provided that $|e(u)| < 1$) and its derivative $\dot{e}(u)$ (provided that $|e(u) - \dot{e}(u)| < 1$). To proceed further, we shall multiply (4) by $e^2(e - \dot{e})$ and also multiply (10) by $[e + (n - 1)\dot{e}]$. The result is the following:

(31) $e^2(e-\dot{e})[e+(n-1)\dot{e}]\mathbf{I_2}(e,\dot{e},n) \equiv e^2(e-\dot{e})(-e+n\dot{e})\mathbf{I_0}(e,\dot{e},n) -$
$-e^2(e-\dot{e})[1+e(e-\dot{e})]\mathbf{I_1}(e,\dot{e},n) = [e+(n-1)\dot{e}](n-2)(1-e^2)\dot{e}\mathbf{I_{0\text{-}}}(e,\dot{e},n)+$
$+[e+(n-1)\dot{e}][2e^2(e-\dot{e})-(n-2)\dot{e}]\mathbf{I_0}(e,\dot{e},n)+$
$+[e+(n-1)\dot{e}][e^2+(n-2)e\dot{e}]\mathbf{I_1}(e,\dot{e},n)$.

The second equality in the above expression gives, in one's turn, a new linear relation between the integrals $\mathbf{I_1}(e,\dot{e},n)$, $\mathbf{I_0}(e,\dot{e},n)$ and $\mathbf{I_{0\text{-}}}(e,\dot{e},n)$:

(32) $e[-2e^2-e^4+(-2n+4)e\dot{e}+2e^3\dot{e}-(n-1)(n-2)\dot{e}^2-e^2\dot{e}^2]\mathbf{I_1}(e,\dot{e},n) =$
$= (n-2)[e\dot{e}-e^3\dot{e}+(n-1)\dot{e}^2-(n-1)e^2\dot{e}^2]\mathbf{I_{0\text{-}}}(e,\dot{e},n)+$
$+[3e^4-(n-2)e\dot{e}+(n-5)e^3\dot{e}-(n-1)(n-2)\dot{e}^2-(n-2)e^2\dot{e}^2]\mathbf{I_0}(e,\dot{e},n)$.

The derivation of the above result (32) supposes that the both multipliers $e(e - \dot{e}) \neq 0$ and $[e + (n - 1)\dot{e}] \neq 0$. As already investigated above, the vanishing of these two multipliers does not invalidate the relations (4)

and (10). Their left-hand-sides will be simply equal to zero. We stress, that our purpose in this section **4** is to eliminate the integral $\mathbf{I_0}(e,\dot{e},n)$ by means of the establishment of linear relations between $\mathbf{I_0}(e,\dot{e},n)$, $\mathbf{I_{0-}}(e,\dot{e},n)$ and $\mathbf{I_{0+}}(e,\dot{e},n)$. Just from this point of view, we shall consider the particular cases $e(e - \dot{e}) = 0$ and $[e + (n - 1)\dot{e}] = 0$. Firstly, we shall resolve the problem namely for these two particular cases and, after that, we shall return to the equality (32).

4.8.1. Case $e(u)[e(u) - \dot{e}(u)] = 0$, $[e(u) + (n - 1)\dot{e}(u)] = 0$, $\dot{e}(u) \neq 0$.

The relations (4) and (10) take the forms (having in mind that $- e = (n - 1)\dot{e}$):

(33) $0 = (2n - 1)\dot{e}\mathbf{I_0}(e,\dot{e},n) - \mathbf{I_1}(e,\dot{e},n)$, or $\mathbf{I_1}(e,\dot{e},n) = (2n - 1)\dot{e}\mathbf{I_0}(e,\dot{e},n)$,

(34) $0 = (n - 2)(1 - e^2)\dot{e}\mathbf{I_{0-}}(e,\dot{e},n) - (n - 2)\dot{e}\mathbf{I_0}(e,\dot{e},n) + [e^2 + (n - 2)e\dot{e}]\mathbf{I_1}(e,\dot{e},n)$.

Combining the above results (33) and (34), we obtain:

(35) $[(n - 2)\dot{e} - (2n - 1)e^2\dot{e} - (2n - 1)(n - 2)e\dot{e}^2]\mathbf{I_0}(e,\dot{e},n) = (n - 2)(1 - e^2)\dot{e}\mathbf{I_{0-}}(e,\dot{e},n)$.

As already mentioned above, $- e = (n - 1)\dot{e}$ for this particular case, and we are able to write (35) as:

(36) $\dot{e}[(n - 2) - (n - 1)(2n - 1)\dot{e}^2]\mathbf{I_0}(e,\dot{e},n) = (n - 2)(1 - e^2)\dot{e}\mathbf{I_{0-}}(e,\dot{e},n)$.

If $\dot{e}(u) = 0$, this equality cannot be used for determination of $\mathbf{I_0}(e,\dot{e},n)$. But if $\dot{e}(u) \neq 0$, (which is the situation *in our case* !), we have:

(37) $[(n - 2) - (n - 1)(2n - 1)\dot{e}^2]\mathbf{I_0}(e,\dot{e},n) = (n - 2)(1 - e^2)\mathbf{I_{0-}}(e,\dot{e},n)$.

Of course, (37) may be useful only if the multiplier $[(n - 2) - (n - 1)(2n - 1)\dot{e}^2] \neq 0$.

It is worthy to note, that in the present case, which we consider, the left-hand-side of the relation (79) from paper [15]:

(38) $2e(e - \dot{e})\mathbf{I_1}(e,\dot{e},n) = (n - 2)(e - \dot{e})(1 - e^2)\mathbf{I_{0-}}(e,\dot{e},n) +$
$+ (n + 1)e[(e - \dot{e})^2 - 1]\mathbf{I_{0+}}(e,\dot{e},n) + [3e + (n - 2)\dot{e}]\mathbf{I_0}(e,\dot{e},n)$

is also equal to zero. This provides another possibility to exclude the integral $\mathbf{I_0}(e,\dot{e},n)$. That is:

(39) $[3e + (n - 2)\dot{e}]\mathbf{I_0}(e,\dot{e},n) = - (n - 2)(e - \dot{e})(1 - e^2)\mathbf{I_{0-}}(e,\dot{e},n) +$
$+ (n + 1)e[1 - (e - \dot{e})^2]\ \mathbf{I_{0+}}(e,\dot{e},n)$.

Taking about the present particular case, the equality $e(u) = - (n - 1)\dot{e}(u)$ we express (39) in the form:

(40) $(2n - 1)\dot{e}\mathbf{I_0}(e,\dot{e},n) = - (n - 2)n\dot{e}(1 - e^2)\mathbf{I_{0-}}(e,\dot{e},n) +$
$+ (n - 1)(n + 1)\dot{e}(1 - n^2\dot{e}^2)\ \mathbf{I_{0+}}(e,\dot{e},n)$.

If $\dot{e}(u) = 0$, this equality cannot be useful for the determination of $\mathbf{I_0}(e,\dot{e},n)$. But if $\dot{e}(u) \neq 0$, we have:

(41) $(2n-1)\mathbf{I_0}(e,\dot{e},n) = -n(n-2)(1-e^2)\mathbf{I_{0-}}(e,\dot{e},n) + (n-1)(n+1)(1-n^2\dot{e}^2)\,\mathbf{I_{0+}}(e,\dot{e},n)$.

If $(2n-1) = 0$ (i.e., $n = 1/2$), we are not able to eliminate the integral $\mathbf{I_0}(e,\dot{e},n)$, using the above relation (41). But nevertheless, $[(n-2) - (n-1)(2n-1)\dot{e}^2] = n - 2 = -3/2 \neq 0$, and we may then use (37) to eliminate $\mathbf{I_0}(e,\dot{e},n)$. Consequently, if $\dot{e}(u) \neq 0$, the linear relations (37) and (41) ensure the elimination of $\mathbf{I_0}(e,\dot{e},n)$ in the case **4.8.1**, which particular case implies $e(u) = -(n-1)\dot{e}(u)$. *We strongly emphasize, that the later equality* ***must not*** *be considered, in general, as a first order ordinary differential equation for the eccentricity* $\boldsymbol{e(u)}$*, whose solution is* $\boldsymbol{e(u) = constant \times exp[-(n-1)^{-1}u]}$. Though such a situation may be (in principle) a subject of special investigation. In the present paper, we limit our computations only to concrete values $u \equiv ln\ p$ of the focal parameter p, which are able to cause troubles (e.g. singularities) in the derived by us linear relations between the seven integrals $\mathbf{I_{0-}}(e,\dot{e},n)$, $\mathbf{I_{0+}}(e,\dot{e},n)$ and $\mathbf{I_j}(e,\dot{e},n)$, ($\mathbf{j} = 0,1,\ldots,4$) (see formulas (1), (2) and (3)). We do not expect that these divergences do scope the whole range of the accretion disc. Such a pathological situation would imply that the accretion disc model itself is very wrong. So, we consider the possible "singular values" of the independent variable $u \equiv ln\ p$ as isolated points or "small" (in some sense) intervals, which do not determine the global structure of the accretion flow.

4.8.2. Case $e(u)[e(u) - \dot{e}(u)] = 0$, $[e(u) + (n-1)\dot{e}(u)] = 0$, $\dot{e}(u) = 0$.

To conclude the considerations in the paragraph **4.8**, ***we return*** to the case $\dot{e}(u) = 0$. By the hypotheses of this paragraph, $\dot{e}(u) = 0$ implies also that $e(u) = 0$. But the situation is now very trivial, simply because

$\mathbf{I_0}(0,0,n) = \int_0^{2\pi} d\varphi = 2\pi$. There does not arise the necessity to represent the integral $\mathbf{I_0}(0,0,n)$ by means of a linear combination of the integrals $\mathbf{I_{0-}}(0,0,n)$ and $\mathbf{I_{0+}}(0,0,n)$. It is worthy to note, that in the limits $e(u) \to 0$ and $\dot{e}(u) \to 0$, the relations (37) and (41) (depending on the condition $n = 1/2$ or $n \neq 1/2$, correspondingly) give the same value 2π for $\mathbf{I_0}(0,0,n)$, although they are derived under the assumption $\dot{e}(u) \neq 0$. That is an indication for a continuous transition of the values of $\mathbf{I_0}(0,0,n)$ through the "singular" value $\dot{e}(u) = 0$. We shall not handle here in a strict mathematical manner this circumstance. Out goal is to prove that the integral $\mathbf{I_0}(0,0,n)$ is possible to be

removed from the dynamical equation ([15], equation (4)) of the accretion disc.

4.8. Case $e(u)[e(u) - \dot{e}(u)] = 0, [e(u) + (n - 1)\dot{e}(u)] \neq 0.$

We shall prove now that under the above conditions $\dot{e}(u)$ ***cannot*** be equal to zero. If we suppose the opposite, namely that $\dot{e}(u) = 0$, then from the second condition $[e(u) + (n - 1)\dot{e}(u)] \neq 0$ it follows that $e(u) \neq 0$. But from the first equality we have $e(u) - \dot{e}(u) = 0$, or $e(u) = 0$. We obtain a ***contradiction***. Hence, $\dot{e}(u) \neq 0$. We shall consider the following subclasses:

4.9.1. Case $e(u)[e(u) - \dot{e}(u)] = 0, [e(u) + (n - 1)\dot{e}(u)] \neq 0, e(u) = 0.$

From $[e(u) + (n - 1)\dot{e}(u)] \neq 0$ it follows $(n - 1)\dot{e}(u) \neq 0$. We already just proved that $\dot{e}(u) \neq 0$. Consequently, $n \neq 1$, otherwise this case **4.9.1** will be ***inconsistent***. The relation (10) reduces to (11) and if $n \neq 2$, after dividing the both sides of (11) by $(n - 2)\dot{e}$, the result is $\mathbf{I_0}(0,\dot{e},n \neq 1, 2) = \mathbf{I_{0\text{-}}}(0,\dot{e},n \neq 1, 2)$. In the case $n = 2$, we have

$\mathbf{I_0}(0,\dot{e},2) = \int_0^{2\pi}(1 - \dot{e}cos\varphi)^{-3}\, d\varphi = \mathbf{I_{0\text{-}}}(0,\dot{e},2)$. Therefore, for all admissible n (i.e., $n \neq 1$) we again obtain the equality

(42) $\quad \mathbf{I_0}(0,\dot{e},n \neq 1) = \mathbf{I_{0\text{-}}}(0,\dot{e},n \neq 1)$.

4.9.2. Case $e(u)[e(u) - \dot{e}(u)] = 0, [e(u) + (n - 1)\dot{e}(u)] \neq 0, e(u) \neq 0.$

If $e(u) \neq 0$, it follows that $e(u) = \dot{e}(u) \neq 0$. The second condition implies that $n\dot{e}(u) \neq 0$. Consequently, we must reject the value $n = 0$, otherwise the case **4.9.2** will be ***inconsistent***. Direct computation shows that:

(43) $\quad \mathbf{I_0}(e,\dot{e} = e,n \neq 0) = \int_0^{2\pi}(1 + ecos\varphi)^{n-2}\, d\varphi = \mathbf{I_{0+}}(e,\dot{e} = e,n \neq 0)$,

according to the definitions (2) and (3). Of course, the equality $\mathbf{I_0}(e,\dot{e} = e,n) = \mathbf{I_{0+}}(e,\dot{e} = e,n)$ is also true for $n = 0$, out of the considered present case.

The above analysis again demonstrates the possibility to express the integral $\mathbf{I_0}(e,\dot{e},n)$ in terms of the integrals $\mathbf{I_{0\text{-}}}(e,\dot{e},n)$ and $\mathbf{I_{0+}}(e,\dot{e},n)$, without to put in use the earlier eliminated integrals $\mathbf{I_j}(e,\dot{e},n)$ ($\mathbf{j} = 1,2,3,4$).

4.9. Case $e(u)[e(u) - \dot{e}(u)] \neq 0$, $[e(u) + (n-1)\dot{e}(u)] = 0$.

The above conditions impose two restrictions over the power n in the viscosity law $\eta = \beta\,\Sigma^n$. From the second equality $[e(u) + (n-1)\dot{e}(u)] = 0$ it follows that:

(44) $e(u) - \dot{e}(u) = -n\dot{e}(u)$ **.**

If $n = 0$, this will vanish the difference $e(u) - \dot{e}(u)$, in contradiction to the hypothesis that $e(u)[e(u) - \dot{e}(u)] \neq 0$. Moreover, if $n = 1$, the second equality also will imply that $e(u) = 0$, again in contradiction to the first requirement. The equality (44) also imposes the requirement $\dot{e}(u) \neq 0$, to avoid vanishing of the difference $e(u) - \dot{e}(u)$. Therefore, *to avoid* ***inconsistency*** of the case **4.10**, we must preliminary exclude the possibilities that some of the equalities $n = 0$, $n = 1$ and $\dot{e}(u) = 0$ (or combinations of them) are appearing. In our consideration, two different subclasses must be investigated separately.

4.10.1. Case $e(u)[e(u) - \dot{e}(u)] \neq 0$, $[e(u) + (n-1)\dot{e}(u)] = 0$, $n \neq 1/2$.

To establish a linear dependence between integrals $\mathbf{I_{0\text{-}}}(e,\dot{e},n)$, $\mathbf{I_{0+}}(e,\dot{e},n)$ and $\mathbf{I_0}(e,\dot{e},n)$, we shall use the relations (4) and (38), which we have already proved to be valid for arbitrary integer/noninteger powers n ($-1 \leq n \leq +3$), $e(u)$ ($|e(u)| < 1$) and $\dot{e}(u)$ ($|e(u) - \dot{e}(u)| < 1$). In the present case $[e(u) + (n-1)\dot{e}(u)] = 0$, i.e., the left-hand-side of (4) is equal to zero. Multiplying (4) by $2e(u)[e(u) - \dot{e}(u)] \neq 0$, we obtain:

(45) $2e(e-\dot{e})(-e+n\dot{e})\mathbf{I_0}(e,\dot{e},n) - [1+e(e-\dot{e})]2e(e-\dot{e})\mathbf{I_1}(e,\dot{e},n) =$
$= 2e(e-\dot{e})(-e+n\dot{e})\mathbf{I_0}(e,\dot{e},n) - [1+e(e-\dot{e})](n-2)(e-\dot{e})(1-e^2)\mathbf{I_{0\text{-}}}(e,\dot{e},n) -$
$- [1+e(e-\dot{e})](n+1)e[(e-\dot{e})^2 - 1]\mathbf{I_{0+}}(e,\dot{e},n) -$
$- [1+e(e-\dot{e})][3e+(n-2)\dot{e}]\mathbf{I_0}(e,\dot{e},n) = 0,$

where we have applied the equality (38). After some algebra, the second equality may be transformed to the next form, representing the linear dependence between $\mathbf{I_0}(e,\dot{e},n)$, $\mathbf{I_{0\text{-}}}(e,\dot{e},n)$ and $\mathbf{I_{0+}}(e,\dot{e},n)$, which we are searching for:

(46) $[3e + 5e^3 + (n-2)\dot{e} - (n+7)e^2\dot{e} + (n+2)e\dot{e}^2]\mathbf{I_0}(e,\dot{e},n) =$
$= (n-2)(-e + e^5 + \dot{e} + e^2\dot{e} - 2e^4\dot{e} - e\dot{e}^2 + e^3\dot{e}^2)\mathbf{I_{0\text{-}}}(e,\dot{e},n) +$
$+ (n+1)(e - e^5 + e^2\dot{e} + 3e^4\dot{e} - e\dot{e}^2 - 3e^3\dot{e}^2 + e^2\dot{e}^3)\mathbf{I_{0+}}(e,\dot{e},n)$.

It is useful to rewrite the relation (45) into another (***equivalent*** to (46))form, which will allow us to reveal more clearly the conditions, making possible to eliminate $\mathbf{I_0}(e,\dot{e},n)$ using (45). For this purpose, we employ that under the hypothesis, valid for the considered case **4.10**, $[e(u) + (n-1)\dot{e}(u)] = 0$. Consequently, we can rewrite (44) in an equivalent form:

(47) $e(u) = -(n-1)\dot{e}(u)$, or $-e(u) = (n-1)\dot{e}(u)$,

in order to eliminate $e(u)$ from (45), obtaining for analysis a more simple expression. The result is:

(48) $(2n-1)\dot{e}[1 + 3n(n-1)\dot{e}^2]\mathbf{I_0}(e,\dot{e},n) =$
$-n(n-2)\dot{e}[1 + n(n-1)\dot{e}^2][1 - (n-1)^2\dot{e}^2]\mathbf{I_{0-}}(e,\dot{e},n) -$
$-(n^2-1)\dot{e}[1 + n(n-1)\dot{e}^2](n^2\dot{e}^2 - 1)\mathbf{I_{0+}}(e,\dot{e},n)$.

Clearly, $\dot{e}(u)$ cancels out (we have already shown at the beginning of the case **4.10**, that $\dot{e}(u)$ cannot be equal to zero). It is also evident that (48) cannot be useful if $n = 1/2$, because its left-hand-side is then equal to zero; for this reason, we supposed in the hypotheses of the subclass **4.10.1** that $n \neq 1/2$. It remains to check is it possible that the third multiplier in the left-hand-side of (48) may happen to be zero for the given value of the argument $u \equiv ln\ p$, i.e.:

(49) $1 + 3n(n-1)\dot{e}^2(u) = 0 \quad <=> \quad 3\dot{e}^2(u)n^2 - 3\dot{e}^2(u)n + 1 = 0$.

From the above equality it is obvious that $n(n-1) < 0$, i.e., the multipliers n and $(n-1)$ have opposite signs (remember that $n \neq 0, 1/2$ and $+1$). Let us consider the two alternative possibilities:

(i) $n < 0$ & $(n-1) > 0$. => We obtain a ***contradiction***, because it is impossible to have simultaneously $n < 0$ and $n > +1$.

(ii) $n > 0$ & $(n-1) < 0$ => $0 < n < +1$. Taking into account that $n \neq 1/2$, we conclude that n belongs to the union of the open intervals $(0,1/2)\mathbf{U}(1/2,1)$ if the equality (49) holds. We shall use the positivity of n in the deriving of the next inequalities.

Having in mind the equalities (44) and (47) (valid for the case **4.10**), and the restrictions $|e(u)| < 1$ and $|e(u) - \dot{e}(u)| < 1$ (valid for any value of u), we are able to rewrite them in the following way:

(50) $n\,|\dot{e}(u)| < 1$,

(51) $(1-n)\,|\dot{e}(u)| < 1$.

Summation of these two inequalities immediately gives:

(52) $|\dot{e}(u)| < 2$.

Let us consider (49) as a quadratic equation for the power n. The discriminant of this equation is:

(53) $\mathbf{D} = 9\dot{e}^4(u) - 12\dot{e}^2(u) \equiv \dot{e}^4(u)[9 - 12\dot{e}^{-2}(u)]$.

We are seeking only for real solutions of the equation (49), which means that $\mathbf{D} \geq 0$. Therefore:

(54) $\dot{e}^2(u) \geq 4/3 \quad => \quad |\dot{e}(u)| \geq 2/\sqrt{3} > 1$.

Combining (52) and (54) leads to the limitations:

(55) $1 < 2/\sqrt{3} \leq |\dot{e}(u)| < 2$.

The two solutions of the equation (49) are:

(56) $n_{1,2} = (6\dot{e}^2)^{-1}(6\dot{e}^2 \pm \sqrt{\mathbf{D}})$.

Strictly speaking, the value $n = +1$ is already excluded and the discriminant **D** must be positive. Correspondingly, (55) has to be corrected as $2/\sqrt{3} < |\dot{e}(u)|$. Moreover, n belongs to the union $(0,1/2)\mathbf{U}(1/2,1)$, and if we choose in (56) the sign "+", the solution for n will be greater than +1. To avoid the contradiction, we must select the sign "–". Then the solution of the equation (49) is:

(57) $n = 1 - [1/4 - (3\dot{e}^2)^{-1}]^{1/2}$.

Obviously, $n < +1$, and the condition $n > 0$ is also satisfied, because $\dot{e}^{-2} > -9/4$. It seems that there is a *possibility* to exist a relation between $\dot{e}(u)$ and the power n belonging to $(0,1/2)\mathbf{U}(1/2,1)$, namely, the equality (57), which implies nullification of the multiplier $[1 + 3n(n-1)\dot{e}^2(u)]$ in the linear relation (48). We shall now show that such a *possibility*, in fact, ***cannot be realized***. Let us accept that the equality (49) is realized. We compute the common factor $[1 + n(n-1)\dot{e}^2(u)]$, which presents in the both terms in the right-hand-side of (48).

(58) $[1 + n(n-1)\dot{e}^2(u)] \equiv [1 + 3n(n-1)\dot{e}^2(u)] - 2n(n-1)\dot{e}^2 = -2n(n-1)\dot{e}^2 \neq 0$,

because $n \neq 0, +1$ and $\dot{e}(u) \neq 0$. Then the equality (48) takes the form (under the condition (49)):

(59) $0 = -n(n-2)[1 - (n-1)^2\dot{e}^2]\mathbf{I_{0-}}(e,\dot{e},n) - (n^2-1)(n^2\dot{e}^2 - 1)\mathbf{I_{0+}}(e,\dot{e},n)$,

where a cancellation by $\dot{e}[1 + n(n-1)\dot{e}^2] \neq 0$ is performed. Expressing $\dot{e}^2(u)$ through the power n, using again (49), we have:

(60) $\dot{e}^2(u) = -[3n(n-1)]^{-1}$, (remember that in the considered case n belongs to the union of the open intervals $(0,1/2)\mathbf{U}(1/2,1)$).

Substitution of this equality into (59) leads to:

(61) $(n-2)(4n-1)\mathbf{I_{0-}}(e,\dot{e},n) + (n+1)(-4n+3)\,\mathbf{I_{0+}}(e,\dot{e},n) = 0$.

At the present stage we shall use a result, which will be proved in a forthcoming paper; namely, the integrals $\mathbf{I_{0-}}(e,\dot{e},n)$ and $\mathbf{I_{0+}}(e,\dot{e},n)$ are *linearly independent* functions on $e(u)$ and $\dot{e}(u)$. This is the reason, for which we prefer to eliminate the integrals $\mathbf{I_4}(e,\dot{e},n)$, $\mathbf{I_2}(e,\dot{e},n)$, $\mathbf{I_1}(e,\dot{e},n)$ and $\mathbf{I_0}(e,\dot{e},n)$ from the dynamical equation ([15], equation (4)) of the accretion disc, and to remain the integrals $\mathbf{I_{0-}}(e,\dot{e},n)$ and $\mathbf{I_{0+}}(e,\dot{e},n)$. The situation with the integral $\mathbf{I_3}(e,\dot{e},n)$ is at present unclear. The linear independence between $\mathbf{I_{0-}}(e,\dot{e},n)$ and $\mathbf{I_{0+}}(e,\dot{e},n)$ implies that the coefficients before these two integrals in the linear relation (61) are identically equal to zero:

(62) $(n-2)(4n-1) = 0 \quad \Rightarrow \quad n = 1/4$, because $n \neq 2$,

(63) $(n+1)(-4n+3) = 0$.

From the later equation (63) we have two different possible solutions: (i) $n = -1$, or (ii) $n = 3/4$. Both they contradict to the solution $n = 1/4$, implied from (62). But (62) and (63) must hold simultaneously. Consequently, this controversial situation means that $[1 + 3n(n-1)\dot{e}^2]$ ***cannot be equal to zero*** for any value $u \equiv ln\ p$ and the *possible* relation between n and $\dot{e}(u)$, admitted by the equality (57) also ***cannot be realized***. As a final result, we conclude that $(2n-1)\dot{e}[1 + 3n(n-1)\dot{e}^2] \neq 0$. This closes the consideration of the case **4.10.1**.

4.10.2. Case $e(u)[e(u) - \dot{e}(u)] \neq 0$, $[e(u) + (n-1)\dot{e}(u)] = 0$, $n = 1/2$.

We directly compute from the definition (3) (for $\mathbf{j} = 0$) that:

$$\text{(64)} \qquad \mathbf{I_0}[e = (1/2)\dot{e},\dot{e},1/2] = \int_0^{2\pi}[1 - (n-1)\dot{e}cos\varphi]^{-3/2}(1 - n\dot{e}cos\varphi)^{-3/2}\, d\varphi =$$

$$= \int_0^{2\pi}[1 - (\dot{e}^2/4)cos^2\varphi]^{-3/2}\, d\varphi > 0\ ,$$

where we have used that for $n = 1/2$ we have $e(u) = (1/2)\dot{e}(u)$ (formula (47)) and $(e - \dot{e}) = -n\dot{e}(u) = -(1/2)\dot{e}(u)$ (formula (44)).

It is known from the analysis, that the definition of the complete elliptic integral of the second kind $\mathbf{E}(k^2)$ is given by ([16], formula **771.**):

$$\text{(65)} \qquad \mathbf{E}(\pi/2,\, k) \equiv \mathbf{E}(k^2) = \int_0^{2\pi}(1 - k^2 sin^2\varphi)^{1/2}\, d\varphi\ .$$

The condition (55) ensures that $(\dot{e}^2/4) < 1$ and then:

$$\text{(66)} \qquad \mathbf{I_0}[e = (1/2)\dot{e},\dot{e},1/2] = 2[1 - (\dot{e}^2/4)]^{-1}\mathbf{E}(\dot{e}^2/4) +$$
$$+ 2[1 - (\dot{e}^2/4)]^{-1/2}\mathbf{E}\{-(\dot{e}^2/4)[1 - (\dot{e}^2/4)]^{-1}\} \underset{\dot{e}\,(u) \to 0}{\longrightarrow} 2\pi\ .$$

We only among the other things show the above formula, in order to manifest the existence of an explicit analytical expression for the integral $\mathbf{I_0}[e = (1/2)\dot{e},\dot{e},1/2]$. We shall not perform here the derivation of the relation (66). It turns out, that just this result goes to be of use. In fact, the hypotheses, made in the subclass **4.10.2.**, leads to the conclusion that the investigated relation (48) simply transforms to the identity $0 = 0$ if $n = 1/2$ and $e(u) = (1/2)\dot{e}(u)$. Thus, in the considered case, it becomes useless for the determination of $\mathbf{I_0}[e = (1/2)\dot{e},\dot{e},1/2]$. To see this, let us evaluate the linear relation (48) (preliminary canceling out the multiplier $\dot{e}(u) \neq 0$) in the case when $n = 1/2$:

(67) $0 = (3/4)[1 - (1/4)\dot{e}^2]^2\,\mathbf{I_{0\text{-}}}(e,\dot{e},1/2) - (3/4)\,)[1 - (1/4)\dot{e}^2]^2\,\mathbf{I_{0+}}(e,\dot{e},1/2)$.

The second inequality (55) $|\dot{e}(u)| < 2$ allows to cancel out the multiplier $(3/4)[1 - (1/4)\dot{e}^2]^2 \neq 0$. Taking into account (47) for $n = 1/2$ (i.e., $e(u) = (1/2)\dot{e}(u)$), we rewrite (67) as:

(68) $$\mathbf{I_{0\text{-}}}[e = (1/2)\dot{e},\, \dot{e},\, 1/2] = \int_0^{2\pi}[1 - (\dot{e}^2/4)cos^2\varphi]^{-3/2}[1 + (1/2)\dot{e}cos\varphi]^{-1}d\varphi =$$

$$= \int_0^{2\pi}[1 - (\dot{e}^2/4)cos^2\varphi]^{-3/2}[1 - (1/2)\dot{e}cos\varphi]^{-1}d\varphi \equiv \mathbf{I_{0+}}[e = (1/2)\dot{e},\dot{e},1/2].$$

The above result can be also checked if we compute the difference:

(69) $$\mathbf{I_{0\text{-}}}[e = (1/2)\dot{e},\dot{e},1/2] - \mathbf{I_{0+}}[e = (1/2)\dot{e},\dot{e},1/2] =$$

$$= -\dot{e}\int_0^{2\pi} cos\varphi[1 - (\dot{e}^2/4)cos^2\varphi]^{-5/3}d\varphi = 0.$$

The later equality is evident from the fact that $cos(\pi + \varphi) = -cos\varphi$.

We conclude with the investigation of that particular case, corresponding to the nullification of the factorized multiplier $e^2(e - \dot{e})[e + (n - 1)\dot{e}]$. Now we return to the linear relation (32), which is useful under the condition $e^2(e - \dot{e})[e + (n - 1)\dot{e}] \neq 0$. In order to obtain the factor $2e(e - \dot{e})\mathbf{I_1}(e,\dot{e},n)$, we multiply (32) by $2(e - \dot{e})$. In this way, we shall get as a factor the left-hand-side of the relation (38), i.e. we use (38) to eliminate the integral $\mathbf{I_1}(e,\, \dot{e},\, n)$. The final result is a linear relation only between the integrals $\mathbf{I_0}(e,\dot{e},n)$, $\mathbf{I_{0\text{-}}}(e,\dot{e},n)$ and $\mathbf{I_{0+}}(e,\dot{e},n)$:

(70)
$$\begin{aligned}
&[-2e^2 - e^4 + (-2n + 4)e\dot{e} + 2e^3\dot{e} - (n - 1)(n - 2)\dot{e}^2 - e^2\dot{e}^2] \times \\
&\times (n - 2)(e - \dot{e})(1 - e^2)\mathbf{I_{0\text{-}}}(e,\dot{e},n) + \\
&+ [-2e^2 - e^4 + (-2n + 4)e\dot{e} + 2e^3\dot{e} - (n - 1)(n - 2)\dot{e}^2 - e^2\dot{e}^2] \times \\
&\times e(n + 1)[(e - \dot{e})^2 - 1]\mathbf{I_{0+}}(e,\dot{e},n) + \\
&+ [-2e^2 - e^4 + (-2n + 4)e\dot{e} + 2e^3\dot{e} - (n - 1)(n - 2)\dot{e}^2 - e^2\dot{e}^2] \times \\
&[3e + (n - 2)\dot{e}]\mathbf{I_0}(e,\dot{e},n) = \\
&= 2(n - 2)(e - \dot{e})[e\dot{e} - e^3\dot{e} + (n - 1)\dot{e}^2 - (n - 1)e^2\dot{e}^2]\mathbf{I_{0\text{-}}}(e,\dot{e},n) + \\
&+ 2(e - \dot{e})[3e^4 - (n - 2)e\dot{e} + (n - 5)e^3\dot{e} - (n - 1)(n - 2)\dot{e}^2 - \\
&- (n - 2)e^2\dot{e}^2]\mathbf{I_0}(e,\dot{e},n) .
\end{aligned}$$

After some algebra, the above equality may be put into a form, which represents $\mathbf{I_0}(e,\dot{e},n)$ as a linear combination of the integrals $\mathbf{I_{0\text{-}}}(e,\dot{e},n)$ and $\mathbf{I_{0+}}(e,\dot{e},n)$:

(71) $[6e^3 + 9e^5 + 6(n-2)e^2\dot{e} + 3(n-8)e^4\dot{e} + 3(n-1)(n-2)e\dot{e}^2 - 3(2n-7)e^3\dot{e}^2 +$
$+ n(n-1)(n-2)\dot{e}^3 + 3(n-2)e^2\dot{e}^3]\mathbf{I_0}(e,\dot{e},n) = (n-2)[-2e^3 + e^5 + e^7 - 2(n-2)e^2\dot{e} +$
$+ (2n-1)e^4\dot{e} - 3e^6\dot{e} + (-n^2 + 3n - 2)e\dot{e}^2 + (n^2 - 3n - 1)e^3\dot{e}^2 + 3e^5\dot{e}^2 + n(n-1)\dot{e}^3 +$
$+ (-n^2 + n + 1)e^2\dot{e}^3 - e^4\dot{e}^3]\mathbf{I_{0-}}(e,\dot{e},n) + (n+1)[2e^3 - e^5 - e^7 + 2(n-2)e^2\dot{e} -$
$- 2(n-3)e^4\dot{e} + 4e^6\dot{e} + (n-1)(n-2)e\dot{e}^2 + (-n^2 + 7n - 11)e^3\dot{e}^2 - 6e^5\dot{e}^2 +$
$+ 2(n-2)^2e^2\dot{e}^3 + 4e^4\dot{e}^3 - (n-1)(n-2)e\dot{e}^4 - e^3\dot{e}^4]\mathbf{I_{0+}}(e,\dot{e},n) \equiv$
$\equiv (n-2)(1-e^2)(e-\dot{e})[-2e^2 - e^4 - 2(n-1)e\dot{e} + 2e^3\dot{e} - n(n-1)\dot{e}^2 - e^2\dot{e}^2]\mathbf{I_{0-}}(e,\dot{e},n) +$
$+ (n+1)[1-(e-\dot{e})^2]e[2e^2 + e^4 + 2(n-2)e\dot{e} - 2e^3\dot{e} + (n-1)(n-2)\dot{e}^2 + e^2\dot{e}^2] \times$
$\times \mathbf{I_{0+}}(e,\dot{e},n).$

This is the linear relation for which we are seeking. Clearly, it may be relevant to the problem of the elimination of $\mathbf{I_0}(e,\dot{e},n)$, only if the multiplier before this integral is different from zero. The investigation of the case when this does not happen is much more complicated than the situation with the other integrals. We shall not investigate in the present paper the possible ineligibility to apply formula (71). We only illustrate graphically (Fig. 1) for two concrete numerical values of the power n (n = +2.4 and n = – 0.4) that such a trouble really exists.

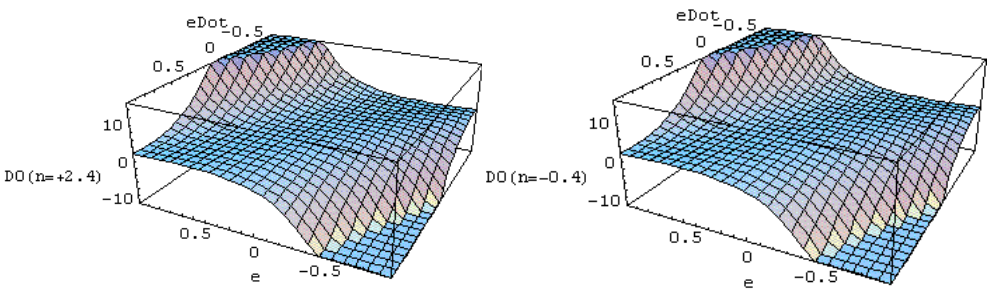


Fig. 1. Two graphics of the coefficient $\mathbf{D_0}(e,\dot{e},n) \equiv 6e^3 + 9e^5 + 6(n-2)e^2\dot{e} + 3(n-8)e^4\dot{e} + +$ $3(n-1)(n-2)e\dot{e}^2 - 3(2n-7)e^3\dot{e}^2 + n(n-1)(n-2)\dot{e}^3 + 3(n-2)e^2\dot{e}^3$ *for two different (arbitrary chosen) values of the power n*; ***top***: n = + 2.4 *and* ***down***: n = – 0.4. *Both* $e(u)$ *and* $\dot{e}(u)$ *take values from* – 0.99 *to* + 0.99 .

5. Conclusions and comments

The last paragraph **4** of the present investigation, in combination with the results in the earlier paper [15], clearly demonstrate that between the seven integrals $\mathbf{I_k}(e,\dot{e},n)$, ($\mathbf{k}$ = 0-, 0+, 0, 1, 2, 3, 4) exist linear relations, which ensure the opportunity to eliminate four of them in the dynamical equation ([15], equation (4)), governing the *space* structure of the *stationary* elliptical accretion discs, according to the model of Lyubarskij et al. [7].

More concretely, we are able to express the integrals $\mathbf{I_4}(e,\dot{e},n)$, $\mathbf{I_2}(e,\dot{e},n)$, $\mathbf{I_1}(e,\dot{e},n)$ and $\mathbf{I_0}(e,\dot{e},n)$ only by means of the integrals $\mathbf{I_{0-}}(e,\dot{e},n)$ and $\mathbf{I_{0+}}(e,\dot{e},n)$ (see definitions (1), (2) and (3)). It turn outs, that the later two integrals $\mathbf{I_{0-}}(e,\dot{e},n)$ and $\mathbf{I_{0+}}(e,\dot{e},n)$ are linearly independent functions of the eccentricity $e(u)$ and its derivative $\dot{e}(u) \equiv de(u)/du$ for each orbit of the disc particles. This statement is quoted in advance and its proof will be given in a forthcoming paper. In such a way, the dynamical equation may be set free from some of the above cited integrals, appearing as a consequence of the azimuthal-angle averaging under the derivation of the dynamical equation in the research of Lyubarskij et al. [7]. This gives some simplification of this equation and may be eventually useful for a finding of its solution by means of analytical methods. Concerning the integral $\mathbf{I_3}(e,\dot{e},n)$, until now, we are not in a condition to eliminate it through the other integrals, *using only linear relations*. The availability of a linear dependence or independence between $\mathbf{I_{0-}}(e,\dot{e},n)$, $\mathbf{I_{0+}}(e,\dot{e},n)$ and $\mathbf{I_3}(e,\dot{e},n)$ will be a subject of a forthcoming analysis. It also remains to establish the utmost limits, under which we shall be able to attain, in our attempts to solve the dynamical equation of the elliptical disc by purely analytical methods. It is possible that this approach may turn out to be only partially successful. We hope that even in this less optimistic situation, the obtained analytical results will be useful for the more clear interpretation both of the *intermediate* calculations and the *further* necessary numerical simulations, leading to the finding of the final solution itself. Most probably, (and unfortunately), it seems that our simplifications of the dynamical equation, governing the *space* structure of the *stationary* elliptical discs ([15], equation (4) and references therein), will be relevant essentially only to the model of Lyubarskij et al. [7]. Other, much more complicated, and more realistic model of elliptical accretion discs, is developed by Ogilvie [18]. There are considered complex-valued eccentricities of the particle orbits. This mathematical approach allows to overcome the restriction of orbits sharing only a common longitude of periastron (i.e. all apse lines of the orbits are in line with each other), which limitation is an essential feature of the examined by us model of Lyubarskij et al. [7]. Unlike the later 2-dimensional analytical simulation, in the full model of Ogilvie ([18], section 3), the basic equations, governing the fluid disc, are written in 3 dimensions. In the case of elliptical discs of Lyubarskij et al. [7], the structure of the accretion flow is prescribed by ***one*** ordinary differential equation, while in the paper of Ogilvie ([18], section 4.4) ***a system of four*** ordinary differential equations must be solved. We have to

take into account, that the theory, presented in [18], goes considerably beyond the previous analytical treatments of the eccentric discs. Consequently, we have to expect that the mathematical treatment of the problem will require (in principle) more complicated ways, in order to solve the structure and the dynamics of the elliptical discs. Of course, resolving the more simple case, described in [7], we hope that at least some of the established properties will be presented also (in some sense) in the real discs, observed in the nature. It will be also useful even for the process of formulation of the more realistic accretion discs models, approaching the characteristics of these objects, investigated by the methods of the observational astronomy. It is clear, that including the considerations of new details or more precisely described processes (for example, including the vertical motions in the disc), will give a better agreement between the theories and observations. But working out of models, like that of Lyubarskij et al. [7], which are not very much appropriate to approximate the really existing accretion flows, indicated by the astronomical observations, is nevertheless an unavoidable step in the direction of their more complete and perfect understanding.

References

1. Z d z i a r s k i, A. A., R. K a w a b a t a, S. M i n e s h i g e. Viscous propagation of mass flow variability in accretion discs., Monthly Not. Royal Astron. Soc., 399, 2009, № 3, pp.1633–1640.
2. F o r g a n, D., K. R i c e. Stellar encounters: a stimulus for disc fragmentation ?, Monthly Not. Royal Astron. Soc., 400, 2009, № 4, pp. 2022–2031.
3. R o y, N., A. K. R a y. Fractal features in accretion discs., Monthly Not. Royal Astron. Soc., 397, 2009, № 3, pp. 1374–1385.
4. M a r z a r i, F., H. S c h o l l, P. T h é b a u t, C. B a r u t e a u. On the eccentricity of self-gravitating circumstellar disks in eccentric binary systems., Astron. & Astrophys., 508, 2009, № 3, pp. 1493–1502.
5. M a r t i n, R. G., J. E. P r i n g l e, C. A. T o u t. The shape of an accretion disc in a misaligned black hole binary., Monthly Not. Royal Astron. Soc., 400, 2009, № 1, pp. 383–391.
6. F r a g n e r, M. M., P. P. N e l s o n. Evolution of warped and twisted accretion discs in close binary systems., Astron. & Astrophys., 511, 2010, February–March, article № A77.
7. L y u b a r s k i j, Y u. 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, pp. 583–596.
8. 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, pp. 337–355.

9. M u r p h y, G. C., J. F e r r e i r a, C. Z a n n i. Large scale magnetic fields in viscous resistive accretion disks. I. Ejection from weakly magnetized disks., Astron. & Astrophys., 511, 2010, February–March, article № A82.
10. H u b b a r d, A., E. G. B l a c k m a n. New constraints on turbulent transport in accretion discs., Monthly Not. Royal Astron. Soc., 398, 2009, № 2, pp. 931–942.
11. L e s u r, G., G. I. Ogilvie. On the angular momentum transport due to vertical convection in accretion discs., Monthly Not. Royal Astron. Soc. Letters, 404, 2010, № 1, pp. L64–L68.
12. K h e s a l i, A., K. F a g h e i. Time dependence of advection-dominated accretion flow with a toroidal magnetic field., Monthly Not. Royal Astron. Soc., 398, 2009, № 3, pp. 1361–1367.
13. Z a l a m e a, I., A. M. B e l o b o r o d o v. Mini-discs around spinning black holes., Monthly Not. Royal Astron. Soc., 398, 2009, № 4, pp. 2005–2011.
14. Z h a n g, H., Y. W a n g, F. Y u a n, F. D i n g, X. L u o, Q. H. P e n g. Is the energy generation rate of nuclear reactions in hot accretion flows important ?, Astron. & Astrophys., 502, 2009, № 2, pp. 419–422.
15. D i m i t r o v, D. V. Thin viscous elliptical accretion discs with orbits sharing a common longitude of periastron.V. Linear relations between azimuthal-angle averaged factors in the dynamical equation., Aerospace Research in Bulgaria, 24, 2010, (in print).
16. D w i g h t, H. B. Tables of integrals and other mathematical data., Fourth edition, New York, MacMillan company, 1961.
17. D i m i t r o v, D. V. 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, 2006, pp.16–28.
18. O g i l v i e, G. I. Non-linear fluid dynamics of eccentric discs., Monthly Notices Royal Astron. Soc., 325, 2001, № 1, pp. 231–248.

ТЪНКИ ВИСКОЗНИ ЕЛИПТИЧНИ АКРЕЦИОННИ ДИСКОВЕ С ОРБИТИ, ИМАЩИ ОБЩА ДЪЛЖИНА НА ПЕРИАСТРОНА. VI. ОПРОСТЯВАНЕ НА ДИНАМИЧНОТО УРАВНЕНИЕ

Д. Димитров

Резюме

Ние продължаваме серията от статии, посветени на изследването и опростяването на динамичното уравнение, определящо структурата на *стационарните* елиптични акреционни дискове. Тези проучвания са в рамките, определени от модела на Любарски и др. [7]. В добавка към предишните проучвания, ние намираме още една зависимост между коефициентите на това обикновено диференциално уравнение от втори

ред, което ни дава възможност да елиминираме ефективно най-малко четири от тях. Това е в курса на нашия подход да намалим броя на тези функции, зависещи от ексцентрицитета, неговата производна и степенния показател $\boldsymbol{n}$ в закона за вискозита $\boldsymbol{\eta = \beta \Sigma^{n}}$. Те се появяват в уравнението в течение на процеса на усредняване (т.е., при интегрирането) по азимуталния ъгъл на елиптичните орбити. На сегашния стадии на изследванията, остават все още три интеграла от указания тип. С изключение на случая на целочислени стойности на $\boldsymbol{n}$, техните аналитични решения не са известни. Във връзка с линейната зависимот или независимост на тези функции (това е предмет на нашите бъдещи проучвания), динамичното уравнение на елиптичните акреционни дискове може да бъде разцепено на една система от съответстващ брой по-прости уравнения за неизвестните ексцентрицитети на орбитите на частиците. Такъв един подход е в съответствие с нашата основна линия, прекарвана през споменатата серия от статии, да се постигне колкото се може по-голям прогрес в решаването на задачата с помощта на *чисто аналитични* методи. И само когато по-нататъшният напредък по този способ (ако крайното решение не е вече достигнато) става толкова сложен, че се оказва в безизходно положение, чак тогава да се използват числените моделирания.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 24, 2012, Sofia

CONSEQUENCES FROM CONSERVATION OF THE TOTAL DENSITY OF THE UNIVERSE DURING THE EXPANSION

Dimitar Valev

Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: valev@gbg.bg

Abstract

The recent Cosmic Microwave Background (CMB) experiments have shown that the average density of the universe is close to the critical one and the universe is asymptotically flat (Euclidean). Taking into account that the universe remains flat and the total density of the universe Ω_0 is conserved equal to a unit during the cosmological expansion, the Schwarzschild radius of the observable universe has been determined equal to the Hubble distance $R_s = \frac{2GM}{c^2} = R \sim cH^{-1}$, where M is the mass of the observable universe, R is the Hubble distance and H is the Hubble constant. Besides, it has been shown that the speed of the light c appears the parabolic velocity for the observable universe $c = \sqrt{\frac{2GM}{R}} = v_p$ and the recessional velocity $v_r = Hr$ of an arbitrary galaxy at a distance r > 100 Mps from the observer, is equal to the parabolic velocity for the sphere, having radius r and a centre, coinciding with the observer. The requirement for conservation of $\Omega_0 = 1$ during the expansion enables to derive the Hoyle-Carvalho formula for the mass of the observable universe $M = \frac{c^3}{2GH} \sim 10^{53} kg$ by a new approach.

1. Introduction

The problem for the average density of the universe $\overline{\rho}$ acquired significance when it was shown that the General Relativity allowes to reveal the large-scale structure and the evolution of the universe by simple cosmological models [1-3]. Crucial for the geometry of the universe appears

the dimensionless total density of the universe $\Omega_0 = \frac{\overline{\rho}}{\rho_c}$, where $\overline{\rho}$ is the average density of the universe and ρ_c is the critical density of the universe. In the case of $\Omega_0 < 1$ (open universe) the global spatial curvature is negative and the geometry of the universe is hyperbolic and in the case of $\Omega_0 > 1$ (closed universe) the curvature is positive and the geometry is spherical. In the special case of $\Omega_0 = 1$ (flat universe) the curvature is zero and the geometry is Euclidean. Until recently scarce information has been available about the density and geometry of the universe. The most reliable determination of the total density Ω_0 is by measurements of the dependence of the anisotropy of the Cosmic Microwave Background (*CMB*) upon the angular scale. The recent results have shown that $\Omega_0 \approx 1 \pm \Delta\Omega_0$, where the error $\Delta\Omega_0$ decreases from 0.10 [4, 5] to 0.02 [6]. The fact that Ω_0 is so close to a unit is not accidental since only at $\Omega_0 = 1$ the geometry of the universe is flat (Euclidean) and the flat universe was predicted by the inflationary theory [7]. The total density Ω_0 includes densities of baryon matter $\Omega_b \approx 0.05$, cold dark matter $\Omega_c \approx 0.22$ [8] and dark energy $\Omega_\Lambda \approx 0.73$, producing an accelerating expansion of the universe [9, 10]. The found negligible *CMB* anisotropy $\frac{\delta T}{T} \sim 10^{-5}$ indicates that the early universe has been very homogeneous and isotropic [11]. Three-dimensional maps of the distribution of galaxies corroborate homogeneous and isotropic universe on large scales greater than 100 *Mps* [12, 13].

2. Consequences from conservation of the total density of the universe during the expansion

The flat geometry of the universe allows to solve some cosmological problems in the Euclidean space. The finite time of the cosmological expansion H^{-1} (age of the universe) and the finite speed of the light *c* set a finite particle horizon $R \sim cH^{-1}$ beyond which no material signals reach the observer. Therefore, for an observer in an arbitrary location, the universe appears a three-dimensional, homogeneous and isotropic sphere having finite "radius" (particle horizon) equal to the Hubble distance $R \sim cH^{-1}$,

where $H \approx 70\ km\ s^{-1}\ Mps^{-1}$ [14] is the Hubble constant and $H^{-1} \approx 1.37 \times 10^{10}$ *years* is the Hubble time (age of the universe).

The fact that the total density of the universe Ω_0 is close to a unit is fundamental since only $\Omega_0 = \frac{\overline{\rho}}{\rho_c} = 1$ supplies flat geometry of the universe. There are no arguments to assume the recent epoch privileged in relation to the other epochs; therefore, the universe always remains flat, and *the total density of the universe Ω_0 is conserved equal to a unit during the cosmological expansion:*

(1) $$\Omega_0 = \frac{\overline{\rho}}{\rho_c} = 1$$

The critical density of the universe [15] is determined from equation (2):

(2) $$\rho_c = \frac{3H^2}{8\pi G} \approx 9.5 \times 10^{-27}\ kg\ m^{-3},$$

where G is the universal gravitational constant.

Considering $\overline{\rho} = \frac{3M}{4\pi R^3}$, where M and R are the mass and the Hubble distance ("radius") of the observable universe, and replacing ρ_c with expression (2) in (1) we obtain:

(3) $$\frac{2MG}{R^3 H^2} = 1$$

Replacing $H \sim cR^{-1}$ in (3) we obtain:

(4) $$R = \frac{2GM}{c^2}$$

Obviously, (4) appears the formula for the Schwarzschild radius of the mass of the observable universe M [16]. Therefore, *the Schwarzschild*

radius of the observable universe R_s *is equal to the Hubble distance* $R_s = R \sim cH^{-1} \sim 1.37 \times 10^{10}$ *light years*.

From (4) we find:

(5) $$c = \sqrt{\frac{2GM}{R}}$$

Evidently, (5) is the formula of the parabolic velocity for the Hubble sphere, i.e. the sphere having mass *M* and a radius, equal to the Hubble distance $R \sim cH^{-1}$. Therefore, *the speed of the light c appears the parabolic velocity* v_p *for the observable universe.*

Below, we find that the recessional velocity $v_r = Hr$ of an arbitrary galaxy at a distance *r* > 100 *Mps* from the observer is equal to the parabolic velocity of a sphere, having radius *r* and a centre, coinciding with the observer. As mentioned at the end of the Introduction, the universe is homogeneous and isotropic on large scales greater than 100 *Mps*. Therefore, the average density ρ_r of a sphere having radius *r* > 100 *Mps* is equal to the average density of the universe $\overline{\rho}$:

(6) $$\rho_r = \frac{3m}{4\pi r^3} = \overline{\rho} \approx \rho_c = \frac{3H^2}{8\pi G},$$

where *m* is the mass of the total matter in the sphere.

We find from equation (6):

(7) $$H = \sqrt{\frac{2Gm}{r^3}}$$

Replacing *H* in the Hubble law $v_r = Hr$ we obtain the recessional velocity of a galaxy:

(8) $$v_r = Hr = \sqrt{\frac{2Gm}{r}}$$

Equation (8) coincides with the formula for the parabolic velocity of a sphere, having radius *r* and a centre, coinciding with the observer.

Finally, the requirement for conservation of the total density of the universe equal to a unit during the expansion allows to estimate the total mass of the observable universe *M*. Actually, replacing $R \sim cH^{-1}$ in (3) we find:

$$M = \frac{c^3}{2GH} \approx 8.8 \times 10^{52}\ kg \tag{9}$$

Obviously, this mass is close to the mass of the Hubble sphere M_H:

$$M_H = \frac{4}{3}\pi R^3 \overline{\rho} \sim \frac{4\pi c^3 \rho_c}{3H^3} = \frac{c^3}{2GH} \tag{10}$$

Formula (9) has been derived independently by dimensional analysis without consideration of the average density of the universe in [17, 18] and practically coincides with the Hoyle-Carvalho formula for the mass of the universe [19, 20], obtained by a totally different approach.

3. Conclusions

The recent *CMB* experiments have shown that the average density of the universe is close to the critical one and the universe is asymptotically flat. The flat geometry of the universe allows to solve some cosmological problems in the Euclidean space. Taking into account that the universe remains flat and the total density of the universe Ω_0 is conserved equal to a unit during the expansion, the Schwarzschild radius of the observable universe has been determined equal to the Hubble distance $R_s = \frac{2GM}{c^2} = R \sim cH^{-1}$, and the speed of the light *c* appears the parabolic velocity for the observable universe $c = \sqrt{\frac{2GM}{R}} = v_p$. Besides, the recessional velocity $v_r = Hr$ of an arbitrary galaxy at a distance *r* > 100 *Mps* from the observer, is equal to the parabolic velocity of a sphere, having radius *r* and a centre, coinciding with the observer.

The requirement for conservation of $\Omega_0 = 1$ during the cosmological expansion enables to derive the Hoyle-Carvalho formula for the mass of the observable universe $M = \frac{c^3}{2GH}$ by a new approach.

References

1. F r i e d m a n, A., Uber die Krummung des Raumes, Z. Physik, 10, 1922, 377-386.
2. L e m a i t r e, G., Un univers Homogene de Masse Constante et de Rayon Croissant, Ann. Soc. Sci. Brux., 47A, 1927, 49-59.
3. E i n s t e i n, A., W. de S i t t e r, On the Relation between the Expansion and the Mean Density of the Universe, Proc. Nat. Acad. Sci. USA, 18, 1932, 213-214.
4. de B e r n a r d i s, P. et al., A flat Universe from High-Resolution Maps of the Cosmic Microwave Background Radiation, Nature, 404, 2000, 955-959.
5. B a l b i, A. et al., Constraints on Cosmological Parameters from MAXIMA-1, Astrophys. J., 545, 2000, L1-L4.
6. S p e r g e l, D. N. et al., First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters, Astrophys. J. Suppl. Series, 148, 2003, 175-194.
7. G u t h, A. H., The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems, Phys. Rev. D, 23, 1981, 347-356.
8. P e a c o c k, J. A. et al., A Measurement of the Cosmological Mass Density from Clustering in the 2dF Galaxy Redshift Survey, Nature, 410, 2001, 169-173.
9. R i e s s, A. G. et al., Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant, Astron. J., 116, 1998, 1009-1038.
10. P e r l m u t t e r, S. et al., Measurements of Omega and Lambda from 42 High-Redshift Supernovae, Astrophys. J., 517, 1999, 565-586.
11. B e n n e t t, C. L. et al., Four-Year COBE DMR Cosmic Microwave Background Observations: Maps and Basic Results, Astrophys. J., 464, 1996, L1-L4.
12. S h e c t m a n, S. A. et al., The Las Campanas Redshift Survey, Astrophys. J., 470, 1996, 172-188.
13. S t o u g h t o n, C. et al., Sloan Digital Sky Survey: Early Data Release, Astron. J., 123, 2002, 485-548.
14. M o u l d, J. R. et al., The Hubble Space Telescope Key Project on the Extragalactic Distance Scale. XXVIII. Combining the Constraints on the Hubble Constant, Astrophys. J., 529, 2000, 786-794.
15. P e e b l e s, P. J. E., Physical Cosmology, Princeton University Press, Princeton, New-Jersey, 1971.
16. S c h w a r z s c h i l d, K., On the Gravitational Field of a Mass Point According to Einstein's Theory, Sitzungsberichte der Koniglich Preussischen Akademie der Wissenschaften zu Berlin, 1916, 189-196.

17. V a l e v, D., Determination of total mechanical energy of the universe within the framework of Newtonian mechanics, Suppl. Compt. rend. Acad. bulg. Sci.: 'Fundamental Space Research', 2010, 241-243; http://arxiv.org/abs/0909.2726
18. V a l e v, D., Estimations of total mass and energy of the universe, 2010, http://arxiv.org/abs/1004.1035
19. H o y l e, F., The Structure and Evolution of the Universe, Proceedings of 11th Solvay Conference in Physics, Edited by R. Stoops, 1958, Brussels.
20. C a r v a l h o, J. C., Derivation of the Mass of the Observable Universe, Int. J. Theor. Phys., 34, 1995, 2507-2509.

ПОСЛЕДИЦИ ОТ ЗАПАЗВАНЕТО НА ТОТАЛНАТА ПЛЪТНОСТ НА ВСЕЛЕНАТА ПРИ РАЗШИРЯВАНЕТО

Д. Вълев

Резюме

Съвременните изследвания на космическия микровълнов фон показват, че средната плътност на Вселената е близка до критичната, а Вселената е асимптотически плоска (Евклидова). Вземайки под внимание това, че Вселената остава плоска, а тоталната плътност на Вселената Ω_0 се запазва равна на единица при космологичното разширяване, е установено че Шварцшилдовият радиус на наблюдаемата Вселената е равен на разстоянието на Хъбъл $R_s = \frac{2GM}{c^2} = R \sim cH^{-1}$, където M е масата на наблюдаемата Вселена, R е разстоянието на Хъбъл, а H е константата на Хъбъл.

Освен това е показано, че скоростта на светлината c се явява параболична скорост за наблюдаемата Вселена $c = \sqrt{\frac{2GM}{R}} = v_p$, а скоростта на отдалечаване $v_r = Hr$ на произволна галактика на разстояние $r > 100\ Mnc$ от наблюдателя, е равна на параболичната скорост за сферата, имаща радиус r и център съвпадащ с наблюдателя. Изискването за запазване на $\Omega_0 = 1$ при космологичното разширяване дава възможност да се изведе формулата на Хойл-Карвальо за масата на наблюдаемата Вселена $M = \frac{c^3}{2GH} \sim 10^{53} kg$ по нов начин.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 24, 2012, Sofia

ESTIMATIONS OF TOTAL MASS AND DENSITY OF THE OBSERVABLE UNIVERSE BY DIMENSIONAL ANALYSIS

Dimitar Valev

Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: valev@gbg.bg

Abstarct

The dimensional analysis has been carried out by means of three fundamental constants - the speed of the light in vacuum (c), the universal gravitational constant (G) and the Hubble constant (H). The mass dimension quantity $m \sim c^3/(GH) \sim 10^{53}$ *kg, derived by this approach, practically coincides with Hoyle-Carvalho formula for the mass of the observable universe, obtained by a totally different approach. It has been shown that this value is close to the mass of the Hubble sphere. Besides, by dimensional analysis it has been found that the total density of the universe, including the dark matter and the dark energy, is of the order of a magnitude of the critical density of the universe* $\rho \sim H^2/G \sim \rho_c \approx 10^{-26}$ *kg* m^{-3}.

Key words: dimensional analysis; fundamental constants; mass of the universe; total density of the universe

1. Introduction

The observable universe consists of the galaxies and other matter that we can in principle observe from Earth in the present day, because light (or other signals) from those objects has had time to reach us since the beginning of the cosmological expansion. Because the universe is homogeneous and isotropic in large scale, the distance to the edge of the observable universe is roughly the same in every direction. Therefore, the observable universe is a three dimensional sphere centered on the observer.

Until recently scarce information has been available concerning the mass, density and geometry of the universe. Recent observations have shown that the average density of the bright matter (stars, galaxies, quasars, etc.) is less than 1 % of the critical density of the universe ρ_c [1].

$$(1) \qquad \rho_c = \frac{3H^2}{8\pi G} \approx 10^{-26}\ kg\ m^{-3}$$

On other side, the rotational curves of spiral galaxies [2] and the stability of rich clusters of galaxies [3] infer total density of the matter $\Omega_M = \rho_M / \rho_c \sim 0.25$, in units of the critical density ρ_c. According to the Big Bang Nucleosynthesis, the density of the ordinary (baryonic) matter is $\Omega_B < 0.05$ [4], whereas the density of the cold dark matter $\Omega_C > 0.20$ [5].

The most trustworthy total matter density $\Omega_0 = \overline{\rho} / \rho_c$ has been determined by measurements of the dependence of the anisotropy of the Cosmic Microwave Background (*CMB*) upon the angular scale. The recent observations show that $\Omega_0 \approx 1 \pm \Delta\Omega$, where the error $\Delta\Omega = 0.02$ [6, 7], i.e. the density of the universe is close to the critical one and the universe is asymptotically flat (Euclidean). Recent distant SNeIa observations have shown accelerating expansion of the universe produced by the dark energy, possessing density $\Omega_\Lambda > 0.70$ [8 – 11]. Therefore, the bulk of density of the universe consists from dark energy and cold dark matter $\Omega_0 = \Omega_B + \Omega_C + \Omega_\Lambda \approx 1$.

In this paper we have shown that the mass and density of the observable universe can be estimated by dimensional analysis using the fundamental constants – the speed of the light in vacuum (*c*), the gravitational constant (*G*) and the Hubble constant (*H*), without any information for the total density of the universe.

The dimensional analysis is a conceptual tool often applied in physics to understand physical situations involving certain physical quantities [12 – 14]. It is routinely used to check the plausibility of the derived equations and computations. When it is known, the certain quantity with which other determinative quantities would be connected, but the form of this connection is unknown, a dimensional equation is composed for its finding. In the left side of the equation, the unit of this quantity q_0 with its dimensional exponent has been placed. In the right side of the equation, the

product of units of the determinative quantities q_i rise to the unknown exponents n_i, has been placed $[q_0] \sim \prod_{i=1}^{n} [q_i]^{n_i}$, where *n* is positive integer and the exponents n_i are rational numbers. Most often, the dimensional analysis is applied in mechanics and other fields of modern physics, where there are many problems with a few determinative quantities. Many interesting and important problems related to the fundamental constants have been considered in [15 – 17].

The discovery of the linear relationship between recessional velocity of distant galaxies, and distance $v = Hr$ [18] introduces new fundamental constant in physics and cosmology – the famous Hubble constant (*H*). Hubble constant determines the age of the universe H^{-1}, the Hubble distance cH^{-1}, the critical density of the universe ρ_c, and other large-scale properties of the universe. Because of the importance of the Hubble constant, in the present paper we include *H* in the dimensional analysis. Thus, the Hubble constant will represent the cosmological phenomena in new derived fundamental mass. According to the recent cosmology, the Hubble 'constant' slowly decreases with the age of the universe, but there are indications that other constants, especially gravitational and fine structure constants also vary with time [19, 20]. That is why, the Hubble constant could deserve being treated on an equal level with the other constants.

2. Estimation of total mass of the observable universe by dimensional analysis

The Plank mass $m_P \sim \sqrt{\frac{\hbar c}{G}}$ has been derived from Planck [21] by dimensional analysis using the fundamental constants – *c*, *G* and the reduced Plank constant ($\hbar$). Since the constants *c*, *G* and $\hbar$ represent three very basic aspects of the universe (i.e. the relativistic, gravitational and quantum phenomena), the Plank mass appears to a certain degree a unification of these phenomena. The Plank mass have many important aspects in modern physics. One of them is that the energy equivalent of

Planck mass $E_P = m_P c^2 \sim \sqrt{\frac{\hbar c^5}{G}} \sim 10^{19}\ GeV$ appears as unification energy of the fundamental interactions [22].

Analogously, we seek a mass dimension quantity *m* composed from the fundamental constants – the speed of the light (*c*), the gravitational constant (*G*) and the Hubble constant (*H*), using dimensional analysis. Here we replace Planck constant with Hubble constant assuming that the quantum phenomena described by $\hbar$ are negligible in relation to the cosmological phenomena described by *H*. Consequentially, so we can write equation (2):

$$m = kc^{n_1} G^{n_2} H^{n_3}, \tag{2}$$

where n_1, n_2 and n_3 are unknown exponents to be determined by matching the dimensions of both sides of the equation and *k* is dimensionless parameter of the order of magnitude of unit. Using the symbol *L* for length, *T* for time, *M* for mass, and writing "[*x*]" for the dimensions of some physical quantity *x*, we have the following:

$$\begin{aligned} [c] &= LT^{-1} \\ [G] &= M^{-1}L^3T^{-2} \\ [H] &= T^{-1} \end{aligned} \tag{3}$$

The dimensions of the left and right sides of the equation (2) must be equal. Therefore:

$$[m] = [c]^{n_1}[G]^{n_2}[H]^{n_3} \tag{4}$$

Taking into account the dimensions of quantities in formula (4) we obtain:

$$L^0T^0M^1 = (LT^{-1})^{n_1}(L^3T^{-2}M^{-1})^{n_2}(T^{-1})^{n_3} = L^{n_1+3n_2}T^{-n_1-2n_2-n_3}M^{-n_2} \tag{5}$$

We find the system of linear equations from (5):

$$\begin{aligned} n_1 + 3n_2 &= 0 \\ -n_1 - 2n_2 - n_3 &= 0 \\ -n_2 &= 1 \end{aligned} \tag{6}$$

The determinant Δ of the system is:

$$\Delta = \begin{vmatrix} 1 & 3 & 0 \\ -1 & -2 & -1 \\ 0 & -1 & 0 \end{vmatrix} = -1 \quad (7)$$

The determinant $\Delta \neq 0$. Therefore, the system has an unique solution. We find this solution by Kramer's formulae (8):

$$n_1 = \frac{\Delta_1}{\Delta}$$
$$n_2 = \frac{\Delta_2}{\Delta}, \quad (8)$$
$$n_3 = \frac{\Delta_3}{\Delta}$$

where $\Delta_1 = -3$, $\Delta_2 = 1$ and $\Delta_3 = 1$.

Therefore, the exponents $n_1 = 3, n_2 = -1, n_3 = -1$. Replacing obtained values of exponents in equation (2) we find formula (9) for the mass m_1:

$$m \sim \frac{c^3}{GH} \quad (9)$$

First of all, the formula (9) has been derived by dimensional analysis in [23]. This formula is close to the Hoyle formula [24] for the mass of the universe $M = c^3/(2GH)$ and perfectly coincides with Carvalho formula [25] for the mass of the universe $M \sim c^3/(GH)$, obtained by totally different approaches, Steady State Theory [26] and Large Number Hypothesis, respectively [19].

The strict linearity of the Hubble law has been confirmed in [27] by SNeIa observations. The Hubble sphere is the sphere where the recessional velocity of the galaxies is equal to the speed of the light in vacuum *c*, and according to the Hubble law $v = c$ when $r = cH^{-1}$. Thus, the Hubble sphere appears a three-dimensional sphere, centered on the observer, having radius $r = cH^{-1}$ and density $\overline{\rho} \approx \rho_c$. Therefore, the mass of the Hubble sphere is:

(10) $$M_H \approx \frac{4}{3}\pi r^3 \rho_c = \frac{4}{3}\pi \frac{c^3}{H^3}\frac{3H^2}{8\pi G} = \frac{c^3}{2GH}$$

Obviously, the mass of the Hubble sphere (10) coincides with the Hoyle formula for the mass of the universe and differs from formula (9) with dimensionless parameter $k = ½ \sim 1$.

The recent experimental values of c, G and H are used – c = 299 792 458 $m\ s^{-1}$, $G = 6.673\times10^{-11}\ m^3\ kg^{-1}\ s^{-2}$ [28] and $H \approx 70\ km\ s^{-1}\ Mps^{-1}$ [29]. Replacing this values in (9) we obtain $m \sim 1.76\times10^{53}\ kg$. Therefore, the enormous mass $m \sim c^3/(GH)$ would be identified with mass of the observable universe. Strictly speaking, the size of the observable universe determines from the cosmological horizon and depends from cosmological model, yet the former roughly coincides with Hubble distance.

3. Estimation of total density of the observable universe by dimensional analysis

Analogously, a quantity ρ having dimension of density could be composed by means of the fundamental constants c, G and H:

(11) $$\rho = kc^{n_1}G^{n_2}H^{n_3},$$

where k is a dimensionless parameter of the order of magnitude of unit.

By dimensional analysis, we find the system of linear equations:

(12) $$\begin{aligned} n_1 + 3n_2 &= -3 \\ -n_1 - 2n_2 - n_3 &= 0 \\ -n_2 &= 1 \end{aligned}$$

The determinant of the system is $\Delta = -1 \neq 0$. Therefore, the system has a unique solution, which we obtain by the Kramer's formulae again:

(13) $$\begin{aligned} n_1 &= \frac{\Delta_1}{\Delta} = 0 \\ n_2 &= \frac{\Delta_2}{\Delta} = -1 \\ n_3 &= \frac{\Delta_3}{\Delta} = 2 \end{aligned}$$

Replacing obtained values of the exponents in equation (11) we find formula (14) for the density ρ:

$$(14) \qquad \rho \sim \frac{H^2}{G} \approx 7.93 \times 10^{-26}\ kg\ m^{-3}$$

As mentioned in Section 1, the recent Cosmic Microwave Background (*CMB*) observations have shown that the total density of the universe $\overline{\rho}$ is close to the critical density ρ_c:

$$(15) \qquad \overline{\rho} = \Omega\rho_c \approx \rho_c = \frac{3H^2}{8\pi G} \approx 10^{-26}\ kg\ m^{-3}$$

It deserves to note that the formula (14) correctly determines the dependence of the total density of the universe on the Hubble and gravitational constants. Besides, the density ρ derived by means of the fundamental constants *c*, *G* and *H* coincides with formula (15) for the total density of the universe with accuracy to a dimensionless parameter $k = 3/(8\pi)$ of the order of a magnitude of a unit.

The formula (14) could be derived by means of other set of fundamental constants, namely ($\hbar$, *G*, *H*). Actually, a quantity ρ having dimension of density could be composed by means of the fundamental constants $\hbar$, *G* and *H*:

$$(16) \qquad \rho = k\hbar^{n_1} G^{n_2} H^{n_3},$$

where *k* is a dimensionless parameter of the order of magnitude of unit.

By dimensional analysis, we obtain the respective system of linear equations:

$$(17) \qquad \begin{aligned} 2n_1 + 3n_2 &= -3 \\ -n_1 - 2n_2 - n_3 &= 0 \\ n_1 - n_2 &= 1 \end{aligned}$$

The determinant of the system $\Delta = -5 \neq 0$ and the system have a solution coinciding with the solution of the system (12):

$$n_1 = \frac{\Delta_1}{\Delta} = 0$$

$$n_2 = \frac{\Delta_2}{\Delta} = -1 \qquad (18)$$

$$n_3 = \frac{\Delta_3}{\Delta} = 2$$

Replacing the exponents in (16) we again obtain equation (14) for the total density of the universe. Therefore, the mass dimension quantity composed by means of the set of fundamental constants ($\hbar$, *G*, *H*) coincides with equation (14) found by means of the set of constants (c, *G*, *H*). Thus, the dimensional analysis automatically rejects inappropriate determinative quantities from the equations, such as *c* and $\hbar$ in the above examined cases.

4. Conclusions

The dimensional analysis has been carried out by means of three fundamental constants - the speed of the light in vacuum (*c*), the universal gravitational constant (*G*) and the Hubble constant (*H*). The mass dimension quantity $m \sim c^3/(GH) \sim 10^{53}$ *kg*, derived by this approach, practically coincides with Hoyle-Carvalho formula for the mass of the observable universe, obtained by a totally different approach. It has been shown that this value is close to the mass of the Hubble sphere. Besides, by dimensional analysis it has been found that the total density of the universe, including the dark matter and the dark energy, is of the order of a magnitude of the critical density of the universe $\rho \sim H^2/G \sim \rho_c \approx 10^{-26}$ *kg* m^{-3}. It deserves to note that these formulae have been derived without consideration of any cosmological model and the formula for the total mass of the observable universe (9) has been obtained by means of the fundamental parameters *c*, *G* and *H* only, with no information for the total density of the universe.

References

1. Peebles, P., Physical Cosmology, Princeton Univ. Press, Princeton, 1971.
2. Persic, M., P. Salucci, F. Stel, The universal rotation curve of spiral galaxies - I. The dark matter connection MNRAS, 281, 1996, 27-47.
3. Evrard, A., The intracluster gas fraction in X-ray clusters - Constraints on the clustered mass density, MNRAS, 292, 1997289-297.

4. O l i v e, K., Big-bang nucleosynthesis, Eur. Phys. J. C 15, 2000, 133-135.
5. P e a c o c k, J. et al., A measurement of the cosmological mass density from clustering in the 2dF Galaxy Redshift Survey, Nature, 410, 2001, 169-173.
6. S p e r g e l, D. et al., First-year Wilkinson microwave anisotropy probe (WMAP) Observations: Determination of cosmological parameters, ApJS, 148, 2003, 175-194.
7. K o m a t s u, E. et al., Seven-year Wilkinson microwave anisotropy probe (WMAP) observations: cosmological interpretation, ApJS, 192, 2011, article id. 18.
8. R i e s s, A. et al., Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant, AJ, 116, 1998, 1009-1038.
9. P e r l m u t t e r, S. et al., Measurements of Omega and Lambda from 42 High-Redshift Supernovae, ApJ, 517, 1999, 565-586.
10. K u z n e t s o v a, N. et al., A New Determination of the High Redshift Type Ia Supernova Rates with the Hubble Space Telescope Advanced Camera for Surveys, ApJ, 673, 2008, 981-998.
11. M o r o k u m a, T. et al., Subaru FOCAS Spectroscopic Observations of High-Redshift Supernovae, Publ. Astron. Soc. Japan, 62, 2010, 19-37.
12. B r i d g m a n, P., Dimensional Analysis, Yale Univ. Press, Yale, 1922.
13. K u r t h, R., Dimensional Analysis and Group Theory in Astrophysics, Pergamon Press, Oxford, 1972.
14. P e t t y, G., Automated computation and consistency checking of physical dimensions and units in scientific programs, Software – Practice and Experience, 31, 2001, 1067-1076.
15. L e v y-L e b l o n d, J., On the conceptual nature of the physical constants, Riv. Nuovo Cim., 7, 1977, 187-214.
16. D u f f, M., L. O k u n, G. V e n e z i a n o, Trialogue on the number of fundamental constants, J. High En. Phys., Issue 03, 2002, id. 023.
17. B a r r o w, J., The Constants of Nature: From Alpha to Omega, Jonathan Cape, London, 2002.
18. H u b b l e, E., A Relation between distance and radial velocity among extra-galactic nebulae, Proc. Nat. Acad. Sci., 15, 1929, 168-173.
19. D i r a c, P., The Cosmological Constants, Nature, 139, 1937, 323.
20. W e b b, J. et al., Further Evidence for Cosmological Evolution of the Fine Structure Constant, Phys. Rev. Lett., id. 87, 2001, 091301.
21. P l a n c k, M., The Theory of Radiation, Dover Publications, NY, 1906 (translated in 1959).
22. G e o r g i, H., H. Q u i n n, S. W e i n b e r g, Hierarchy of interactions in unified gauge theories, Phys. Rev. Lett., 33, 1974, 451-454.
23. V a l e v, D., Determination of total mechanical energy of the universe within the framework of Newtonian mechanics, Compt. Rend. Acad. Bulg. Sci., Spec. Issue, 2009, 233-235 [arxiv:0909.2726].
24. H o y l e, F., The Structure and Evolution of the Universe, Proc. 11th Solvay Conference in Physics, Brussels, edited by R. Stoops, 1958.
25. C a r v a l h o, J., Derivation of the mass of the observable universe, Int. J. Theor. Phys., 34, 1995, 2507-2509.

26. B o n d i, H., T. G o l d, The steady-state theory of the expanding universe, MNRAS, 108, 1948, 252-270.
27. R i e s s, A., W. P r e s s, R. K i r s h n e r, A Precise Distance Indicator: Type IA Supernova Multicolor Light-Curve Shapes, ApJ, 473, 1996, 88-109.
28. M o h r, P., B. T a y l o r, CODATA recommended values of the fundamental physical constants 1998, J. Phys. Chem. Ref. Data, 28, 1999, 1713-1852.
29. M o u l d, J. et al., The Hubble space telescope key project on the extragalactic distance scale. XXVIII. Combining the constraints on the Hubble constant, ApJ., 529, 2000, 786-794.

ОЦЕНКИ НА ТОТАЛНАТА МАСА И ПЛЪТНОСТ НА НАБЛЮДАЕМАТА ВСЕЛЕНА ПОСРЕДСТВОМ АНАЛИЗ НА РАЗМЕРНОСТИТЕ

Д. Вълев

Резюме

Проведен бе анализ на размерностите с помощта на три фундаментални константи, а именно – скоростта на светлината във вакуум (c), универсалната гравитационна константа (G) и константата на Хъбъл (H). Величината с размерност на маса $m \sim c^3/(GH) \sim 10^{53}$ *kg*, изведена посредством този метод практически съвпада с формулата на Хойл-Карвальо за масата на наблюдаемата Вселена, получена посредством съвършенно различен метод. Показано бе че, тази стойност е близка до масата на сферата на Хъбъл. Освен това, посредством анализ на размерностите бе установено, че тоталната плътност на Вселената, включително тъмната материя и тъмната енергия, е от порядъка на критичната плътност на Вселената $\rho \sim H^2/G \sim \rho_c \approx 10^{-26}$ *kg m*$^{-3}$.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 24, 2012, Sofia

MESOPAUSE TEMPERATURE RETRIEVAL FROM SATI MEASURED SPECTRA WITH A BACKGROUND WITH LINER INTENSITY COMPONENT. RESULTS FROM THE NUMERICAL SIMULATION

Atanas Atanassov

Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: At_M_Atanassov@yahoo.com

Abstract

The Spectral Airglow Temperature Imager is a ground-based spectral instrument for spatial registration of airglow emissions. The basic aim of the instrument development is the investigation of gravity waves based on the spatial characteristics of the temperature field at the altitude of mesopause and its evolution in the time. The temperature retrieval is based on matching measured and preliminary calculated synthetic spectra.

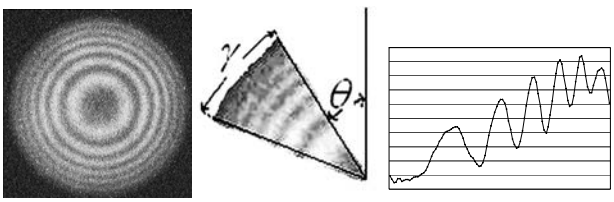
Possibilities are presented for generalization of the basic regression equation which connects the measured and the synthetic spectra. A linear change of the background (slant background) for the entire filter transmittance interval was presumed. Numerical experiments by Monte-Karlo simulation were conducted.

The presented results show a bigger stability of the proposed approach in comparison with the traditional one, without considering the slant background.

Introduction

The Spectral Airglow Temperature Imager (SATI) is a ground-based Fabry-Perot spectrometer for spatial registration of airglow emissions [1, 2]. The SATI instrument was originally intended for investigation of internal gravity waves, propagating at the altitude of the mesopause [3]. This is possible by processing the registered images (Fig. 1a). Twelve sector spectra are determined from each image (fig. 1b, c). These spectra are compared with synthetic spectra, calculated in advance for different temperatures for determination of the mesopause temperature where the

respective emissions are radiated. The radiation maximum of these airglow emissions is disposed at a mean altitude of 97 km for the transitions $O_2(b^1\Sigma_g^+ - X^3\sum_g^-)$. The processing of images from one night séance yield the rotational temperature and the emission rate sector time series [1, 2] determined for different points from the annular sky segment at the altitude of the mesopause. The propagation and the cinematic parameters of the internal gravity waves are determined on the basis of the analysis of these time series [3].



a) *b)* *c)*

Fig. 1. a) Original image; b) image sector; and c) sector spectrum

A version of SATI- 3SZ was manufactured at the Department of the Solar-Terrestrial Influences Laboratory in Stara Zagora [4] in collaboration with the Space Instrumentation Laboratory at the Centre for Research in Earth and Space Science (CRESS), York University, Toronto, Canada. Additionally, algorithms for retrieval of the rotational temperature from the registered O_2 spectra were investigated and developed [5-6].

Temperature retrieval

The determination of the mesopause temperature by SATI-registered images is based on the comparison of the measured spectra with a set of patterns – calculated in advance for different values of the temperature of the radiated oxygen molecules spectra which correspond to transitions $O_2(b^1\Sigma_g^+ - X^3\sum_g^-)$. These transitions correspond to wave longitude in the interval 864÷868 nm. The maximum radiation of the SATI-registered signal is at the height of the mesopause. The comparison of the measured spectrum M with the synthetic spectra $S_{T,}$ calculated in advance for different temperatures for long-time night interval allows to determine the equilibrium temperature of the oxygen molecules, radiating the signal,

incoming toward the Earth. This is done on the basis of the following linear regression equation:

(1) $$M = E.\tilde{S}_{T_{rot}} + B$$

In the last equation, the physical meaning of the multiplicative coefficient E is an integral emission intensity for all filter transmittance intervals. The meaning of B is a mean intensity of the background. The solution of regression equation (1) yields the values of E and B. The rotational temperature T_{rot} is determined on the basis of minimizing the functional

(2) $$\delta_{L,T_{rot}} = \sqrt{\frac{1}{p_2 - p_1}\sum_{p=p_1}^{p_2}(M_p - E.\tilde{S}_{T_{rot},p} - B)^2.w_p}$$

The additive constant B in regression equation (1) allows the admission of even background intensities for the entire spectral interval. The weight coefficients w_p satisfy condition $\sum_p w_p = 1$.

Generalization of the basic regression equation for a case of slant background

The conducted measurements are accepted to be correct and usable when a low background signal is retrieved. Besides, the background signal is accepted even for the entire filter transmittance spectral interval. In all cases when spectral pollution from a secondary source is available, the respective measurements are eventually removed.

We will accept a linear change of the background by frequency and generalization of the basic regression equation (1)

(3) $$M = E.\tilde{S}_{T_{rot}} + B^0 + B^{\angle}$$

The background B was presented by two components $\overline{B}$ and $B^{\angle}$ in equation (3). Component $\overline{B}$ is even by intensity for the entire spectral interval of the background which was passed from the filter; component $B^{\angle}$ is presumed dependant linearly in the frame of the same spectral interval. Both components can be changed for the time of every next measurement as a result of the change of the direction of the incoming signal in relation to the stars. Component $B^{\angle}$ is unknown in the space and time and can be determined by minimization of the generalized version of (2)

(4) $$\delta_{L;T_{rot}} = \sqrt{\frac{1}{p_2 - p_1}\sum_{p=p_1}^{p_2}(M_p - E.\tilde{S}_{T_{rot},p} - B^0 - B^{\angle})^2.w_p}$$

The solution of regression task (3) by the same tools as (1) is possible after transformation to

(5) $$M - B^{\angle} = E.\tilde{S}_{T_{rot}} + B^0$$

or

(5') $$M^{\angle} = E.\tilde{S}_{T_{rot}} + B^0,$$

In (5') $M^{\angle}$ is a modified measured spectrum. The minimization is based already on

(6) $$\delta_{L;T_{rot},B^{\angle}} = \sqrt{\frac{1}{p_2 - p_1}\sum_{p=p_1}^{p_2}(M_p^{\angle} - E.\tilde{S}_{T_{rot},p} - B^0)^2.w_p}$$

In the new regression task, the rotational temperature is found when (6) has a minimum along with the additional condition for slant background.

The reasons for admission of uneven background are purely physical and are connected with the natural conditions of the measurement. Except the natural airglow emissions, other sources of a natural background are also available.

Monte-Karlo simulation

A series of simulations was performed as for each temperature in the interval (110÷300°K) one "measured" spectrum was generated. This was made as the relative synthetic spectrum $S_{T_{rot}}$ was transformed towards the image space by choosing typical values of the filter parameters (refractive index and central wavelength) in order to conduct the entire experiment.

(7) $$\tilde{S}_{T_{rot},p} = \sum_{\lambda_p}^{\lambda_p+1} S_{T_{rot},\lambda}$$

In the last formula λ_p is the wavelength, corresponding to the p^{-th} pixel in the one-dimensional space of the measured spectra [5]. A multiplicative gaining coefficient E is applied; with this coefficient the values of the spectrum for every pixel are fitted to the real ones:

(8) $$M_{T_{rot,p}} = E.\tilde{S}_{T_{rot,p}}$$

A random noise was added to the spectrum with a generator of random numbers $\phi(p) = \Phi.(RANDOM - .5)$ with amplitude Φ . Two background components - even and uneven, are added to the spectrum, thus obtained, and we finally have

(9) $$M_{T_{rot,p}} = E.\tilde{S}_{T_{rot,p}} + \overline{B} + B^{\angle} + \phi_p$$

The aim of the investigation is to establish the behavior of the original approach for temperature retrieval without taking into account component $B^{\angle}$ as well as the behavior of the modified one, taking into account this component.

For each temperature spectrum, generated in the observed interval (110-300K), "polluted" with uneven background, 20 tests were performed for temperature retrieval by the two approaches.

Uneven background with linear change and different intensity at both ends b2= b1 + b(rand-.5) was generated by every test. Value b1 is the background intensity at the first end of the spectrum, however, at the other end we have b2= b1 + b.(rand-.5); where b is the doubled maximum of the possible difference of the background amplitude at this end and ξ- is the random number within the interval (-.5,5). A series of experiments was conducted with different values of even and uneven backgrounds $(\overline{B}, B^{\angle})$=(20,20; 20,15; 20,5) (Table 1). Experiments with different values of the random noise amplitude Φ =(1;3;5) and different gain coefficients E=(25,50,100) were conducted.

Table 1

E = 25 Noise = 5 B = (20;20)	E = 50 Noise = 5 B = (20;20)	E = 100 Noise = 5 B = (20;20)
E = 25 Noise = 5 B = (20;15)	E = 50 Noise = 5 B = (20;15)	E = 100 Noise = 5 B = (20;15)
E = 25 Noise = 5 B = (20;5)	E = 50 Noise = 5 B = (20;5)	E = 100 Noise = 5 B = (20;5)

Results from the Monte-Karlo simulation experiments

Figures 2 and 3 contain histograms which present the temperature retrieval process by applying the two models M and B. The initial temperatures, at which "the measured spectra" were generated, were 170°K for fig.2 and 210°K for fig.3 by intensities 25, 50 and 100 respectively. During these experiments, 1000 tests were performed for temperature retrieval with addition of random noise and slant background. The difference between the two approaches is obvious. The retrieval temperatures were distributed in the ranges 67° and 95°, respectively, for initial temperatures of 170° and 210° and with model intensity 25. At higher intensity of the modeled signal of 50 units, the intervals are narrowed to 33°K and 46°K, respectively, and to 16°K and 22°K at intensity of 100, which correspond to reduction of the significance of the slant background. When applying the temperature retrieval model B, the intervals in which the temperatures are distributed at intensity 25, are 28°K and 57°K, 14°K and 19°K at intensity 50 and 7°K and 10°K at intensity 100, respectively.

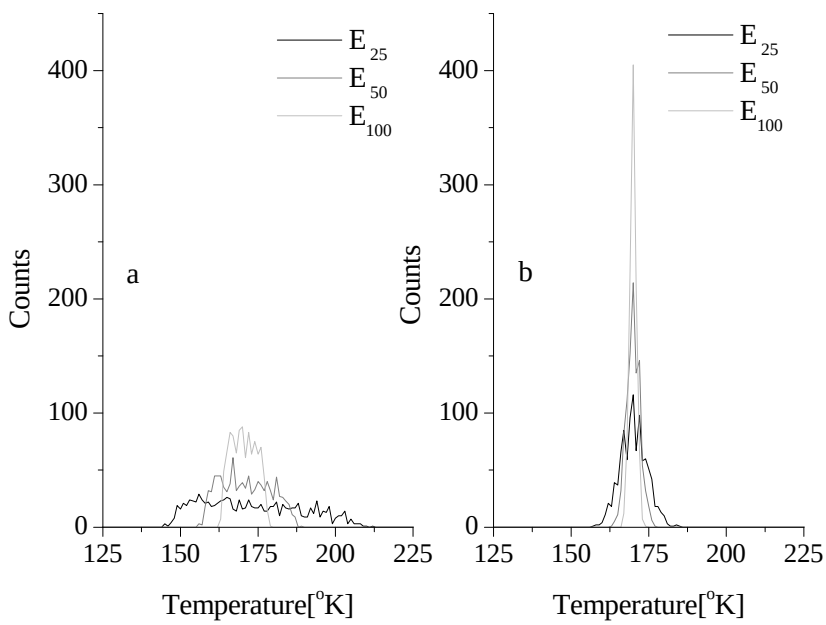


Fig. 2. Histograms of retrieved temperatures by original temperature 170°K- (a) by original method and (b) by uneven background reduction

Figures 4 and 5 show the mean square deviation and the absolute errors for the retrieved values of the temperature in series from 20 tests for each initial temperature in the interval (110-300°K). Figures a-c refer to intensity E of 25 and to values of the slant background component of 20, 15 and 5, respectively. The intensity is 50 units in the figures from the second column (d-f) and in the third one (g-i) - 100 units, with analogous background. It is obvious that the retrieved values fall within a very large interval, without considering the slant background. The values which were retrieved by model B fall within a rather narrow interval.

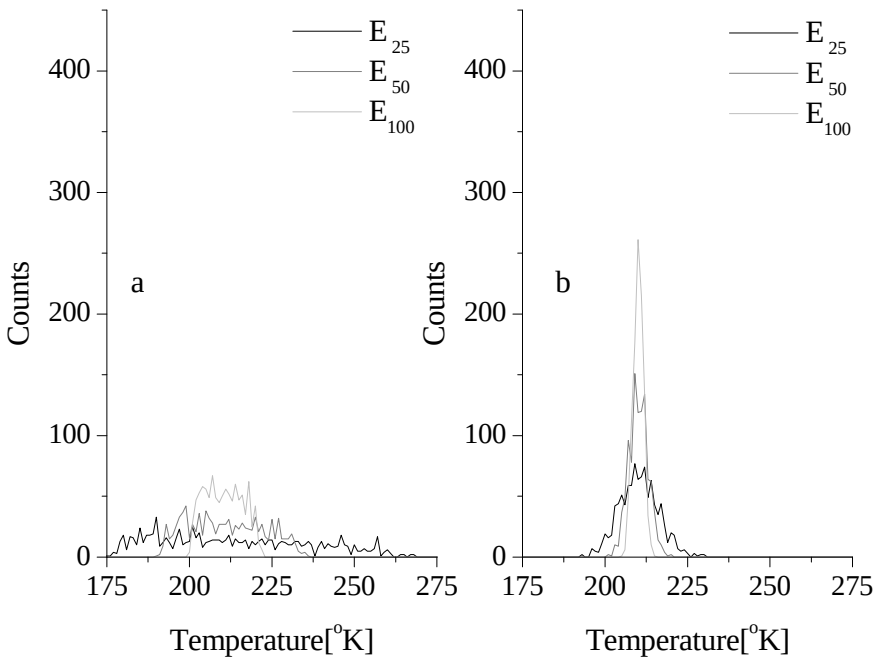


Fig. 3. Histograms of retrieved temperatures by original temperature 210°K- (a) by even component only and (b) by slant background

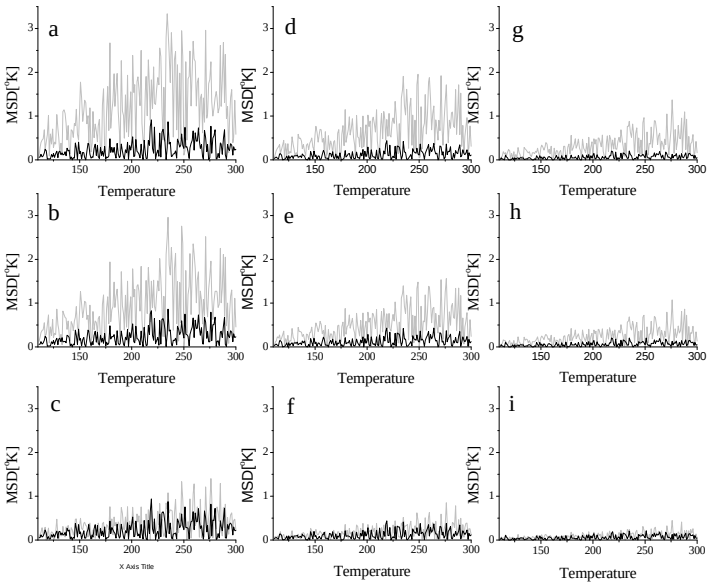


Fig. 4. Mean Standard deviation for E=25 and B< (20,15,5) respectively (a-c); analogously E= 50 and B< (20,15,5) respectively (d-f) and E= 100 and B< (20, 15, 5) (g-i)

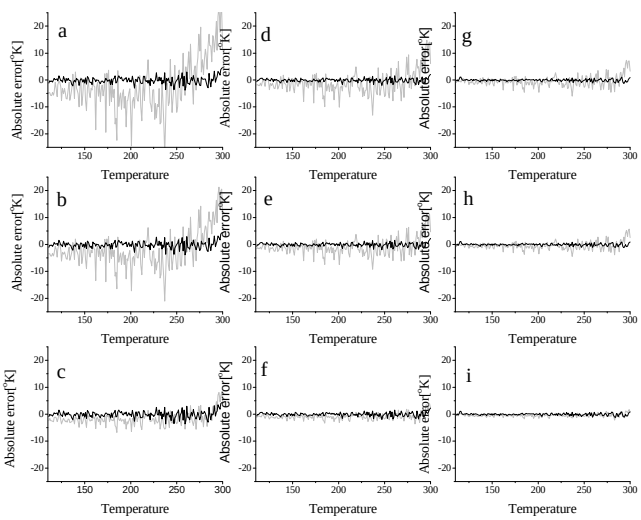


Fig. 5. Absolute error for E=25 and B< (20,15,5) respectively (a-c); analogously E= 50 and B< (20,15,5) respectively (d-f) and E= 100 and B< (20, 15, 5) (g-i)

The absolute error grows by applying of model M with increasing of the retrieved temperature. Respective increasing by applying of model B is minimal.

Some asymmetry in absolute errors for model M is obviously, which speak probably for potential systematic error by retrieving of temperatures

Conclusions

If a slant background component is available in the measured spectra, the proposed approach would yield better temperature results. In all cases, the mean results from the temperature retrieval, taking into consideration the uneven background, will be in the range of one degree for almost the entire temperature interval of the simulation (110-300K).

Except the development of the basic calculation model with addition of uneven background, which is random in the time by its character, very interesting appears a similar development with a significant growth of the background, connected with additional plank sources. The investigation of the spectral pollution with a solar spectrum during sunrise or sunset, or the Moon reflected solar light when the Moon is over the horizon, are current tasks, focusing the attention.

The application of the present approach for minimization of errors by real data is a very attractive idea, aimed at decreasing the noise in the nocturnal course of the sector temperatures, determined by SATI-registered images. In real conditions, along with the possible slant background, most probably other parameters of the calculation models will also have a random character. Thus, the errors of the filter parameters determination will lead to errors of the synthetic spectra transformation [5]. These errors represent specific discrepancies between the transformed synthetic and measured spectra. Therefore, the errors, connected with the slant background, will compete with the noise in the measured spectrum and the unsystematic random noise of the filter parameters determination. The analysis of the influence of these errors is a possible next stage of the investigation.

References

1. Wiens, R., A. Moise, S. Brown, S. Sargoychev, R. Peterson, G. Shepherd, M. Lopez-Gonzalez, J. Lopez-Moreno, R. Rodrigo, SATI: a spectral airglow temperature imager. Adv. Space Res. 19, 677-680, 1997.

2. Sargoytchev, S., S. Brown, B. Solheim, Y.-M. Cho, G. Shepherd, M. López-Gonzalez, Spectral airglow temperature imager (SATI) – a ground based instrument for temperature monitoring of the mesosphere region. Appl. Opt. 43, 5712-5721, 2004.
3. Zhang, S., R. Peterson, R. Wiens, G. Shepherd, Gravity waves from O_2 nightglow during the AIDA '89 campaign I: emission rate/temperature observations. J. Atmos. Terr. Phys., 55 (3), 355-375, 1993.
4. Petkov, N., A. Atanassov, B. Benev, K. Kanev, G. Hristov, L. Bankov, S. Sargoytchev, M. Shepherd, SATI-3SZ - a Spectral airglow temperature imager in Stara Zagora station: possibilities and first results, Proceedings of conference "Fundamental Space Research", Sunny Beach, Bulgaria, 21-28 Sep 2008, 347-350, 2008. http://www.stil.bas.bg/FSR/PDF/TOP4Petkov_Nencho2261626.pdf
5. Atanassov, At., Determination of Sati Instrument Filter Parameters by Processing Interference Images, Compt. rend. Acad. bulg. Sci., 62, No 8, 2009, 993-1000
6. Atanassov, At., Dark Image Correction of Spectrograms Produced by Sati Instrument, Compt. rend. Acad. bulg. Sci., 63, No 4, 2010, 583-592.

ОПРЕДЕЛЯНЕ НА ТЕМПЕРАТУРАТА НА МЕЗОПАУЗАТА ОТ СПЕКТРИ, ИЗМЕРЕНИ С ИНСТРУМЕНТА SATI ПРИ НАКЛОНЕНА КОМПОНЕНТА НА ФОНА. РЕЗУЛТАТИ ОТ ЧИСЛЕНИ ЕКСПЕРИМЕНТИ

Ат. Атанасов

Резюме

Spectral Airglow Temperature Imager е наземен спектрален инструмент за пространствена регистрация на емисии на нощно светене. Основната цел на развитието на инструмента е изследването на гравитационни вълни, основано на пространствените характеристики на температурното поле на височината на мезопаузата и тяхната еволюция във времето. Определянето на температурата се основава на сравняване на измерените спектри с предварително изчислени синтетични спектри.

Представени са възможности за обобщаване на основното регресионно уравнение, което свързва измерените и синтетичните спектри. Предполага се линейно изменение на фона (наклонен фон) за целия интервал на пропускане на филтъра. Проведена е симулация по метода Монте-Карло.

Представените резултатите показват по-голяма стабилност на предлагания подход в сравнение с традиционния, без разглеждане на наклонен фон.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 24, 2012, Sofia

OVERVIEW OF THE ATMOSPHERIC IONIZING RADIATION ENVIRONMENT MONITORING BY BULGARIAN BUILD INSTRUMENTS

Tsvetan Dachev[1], Borislav Tomov[1], Yury Matviichuk[1], Plamen Dimitrov[1], Frantisek Spurny[2], Ondrej Ploc[2], Yukio Uchihori[3], Erwin Flueckiger[4], Karel Kudela[5], Eric Benton[6]

[1]Space Research and Technology Institute – Bulgarian Academy of Sciences, tdachev@bas.bg, btomov@bas.bg, yumat@bas.bg, pdimitrov1957@abv.bg
[2]Nuclear Physics Institute – Czech Acad. Sci., ploc@ujf.cas.cz
[3]National Institute of Radiological Sciences, Chiba, Japan, uchihori@nirs.go.jp
[4]Physikalisches Institut, University of Bern, Switzerland, flueckiger@space.unibe.ch
[5]Ústav experimentálnej fyziky SAV, Oddelenie kozmickej fyziky, Košice, Slovakia, kkudela@kosice.upsj.sk;
[6]Oklahoma State University, Stillwater, OK, USA, eric.benton@okstate.edu

Abstract

Humans are exposed to ionizing radiation all the time, and it is known that it can induce a variety of harmful biological effects. Consequently, it is necessary to quantitatively assess the level of exposure to this radiation as the basis for estimating risks for their health. Spacecraft and aircraft crews are exposed to elevated levels of cosmic radiation of galactic and solar origin and to secondary radiation produced in the atmosphere, the vehicle structure and its contents. The aircraft crew monitoring is required by the following recommendations of the International Commission on Radiological Protection (ICRP) (ICRP 1990), the European Union (EU) introduced a revised Basic Safety Standards Directive (EC 1997) which, inter alia, included the exposure to cosmic radiation. This approach has been also adopted in other official documents (NCRP 2002). In this overview we present the results of ground based, mountain peaks, aircraft, balloon and rocket radiation environment monitoring by means of a Si-diode energy deposition spectrometer Liulin type developed first in Bulgarian Academy of Sciences (BAS) for the purposes of the space radiation monitoring at MIR and International Space Station (ISS). These spectrometers-dosemeters are further developed, calibrated and used by scientific groups in different countries. Calibration procedures of them are performed at different

accelerators including runs in the CERN high-energy reference field, simulating the radiation field at 10 km altitude in the atmosphere and with heavy ions in Chiba, Japan HIMAC accelerator were performed also. The long term aircraft data base were accumulated using specially developed battery operated instrument in 2001-2009 years onboard of A310-300 aircrafts of Czech Air Lines, during 24 about 2 months runs with more than 2000 flights and 13500 flight hours on routes over the Atlantic Ocean mainly. The obtained experimental data are compared with computational models like CARI and EPCARD. The mountain peak measurements are performed with Liulin-6S, Liulin-6MB and Liulin-6M internet based instruments. They use internet module to generate WEB page, which is posted online. The obtained deposited energy spectra, dose and flux data are transmitted via LAN interface by HTTP and FTP protocols. They work online for different periods between 2005 and 2011 at Jungfrau (3453 meters Above Mean Sea Level (AMSL) http://130.92.231.184/); at Lomnicky stit (2633 meters AMSL http://147.213.218.13/) and Moussala (2925 meters AMSL http://beo-db.inrne.bas.bg/moussala/) peaks in Switzerland, Slovakia and Bulgaria. 4 small size battery operated instruments were flown on balloon over south France in June 2000 and NASA balloon over New Mexico, USA on 11th of June 2005. 1 instrument was used in rocket experiment January 2008.

1. Introduction

Ionization in the lower atmosphere is dominated by radionuclides in the Earth's crust. Over deep water, there are few dissolved radionuclides so that the ionization is dominated by radiation incident on the top of the atmosphere. The ionization over the landmass is complicated and depends on many physical and chemical factors. V. F. Hess studies found the ionization rates to decrease with altitude up to 500 meters followed by a steady increase at higher altitudes to where the ground level rate is matched at 1500 meters. For this discovery, Hess would receive a Noble prize in physics (1936) [1].

The space radiation sources as Galactic cosmic rays (GCR) and Solar cosmic rays (SCR) penetrate deep in the atmosphere where the primary protons generate a cascade of particles (protons, neutrons, pions, muons, electrons and gamma quants) [2]. The first reactions of the cosmic rays with the atmosphere occur at altitudes above 20 km. Down at altitudes 19-20 km is recognized the so-called Photzer maximum [3]. This is the main maximum of the ionization and of dose rates profile in the Earth atmosphere and reach about 3 $\mu Gy\ h^{-1}$. At aircraft altitudes (10-12 km) the neutron flux dominate and generate about 10% of the absorbed dose but 59% of the ambient equivalent dose. Earth magnetic field shields atmosphere from the primary and secondary cosmic rays that is why the maximum of the

latitudinal profile is at high magnetic latitudes. Close to the magnetic equator is formed the absolute minimum of the ionizing radiation at aircraft altitudes [4,5].

The solar activity also modulates the atmospheric ionizing radiation through the modulation of the GCR flux. Interplanetary magnetic field, which is embedded in the solar wind shield the heliosphere from GCR and this, is the reason for observation of maximum of the GCR flux and respectively the dose rates in atmosphere close to the minimum of the solar cycle. In reverse close to the maximum of the solar cycle a minimum of the GCR flux is observed [5].

1. Description of the Liulin type instruments used for atmospheric radiation monitoring

The main purpose of Liulin type Spectrometry-Dosimetry Instruments (LSDIs) is cosmic radiation monitoring in the atmosphere at the workplaces. LSDI measures the amplitudes of the pulses generated by the incoming particle and rays radiation in the silicon detector, which is proportional to the deposited energy and respectively to the absorbed dose in Gray. These amplitudes are organized in 256 channels spectrum of the deposited energy in the silicon detector, which is further used for precise calculation of the absorbed and equivalent doses and for characterization of the type and energy of the incoming radiation. Up to now more than 20 LSDIs were developed, build and used on the ground, in aircraft altitudes, in Low Earth Orbits (LEO) and inside and outside of the Earth magnetosphere and on the Moon orbit [6].

First use of Liulin type LSDI were in the Mobile Radiation Exposure Control System - Liulin-E094, which contains 4 active individual dosemeters and worked successfully between May and August 2001 on board of US Laboratory module of the International Space Station (ISS) as a part of the ESA Dosimetric mapping experiment leaded by Dr. Günter Reitz, DLR, Germany [7, 8].

The LSDI functionally is low mass, low power consumption or battery operated dosimeter. The smallest one built till now is the RADOM instrument (98 grams) used for measurement of the near Moon radiation environment on the first Indian Moon satellite – Chandrayaan-1 in 2008-2009 [9, 10]. The largest modifications (450 grams) are these with 2 lithium-ion batteries, Global Positioning System (GPS) receiver and Secure

Digital (SD) card for about a month of independent recording of the radiation environment and the UTC time and geographic coordinates at aircraft altitudes [11]. (See Figure 1.)

Fig. 1. LSDI with GPS receiver and SD card for continues dose monitoring at aircrafts

Except the already mentioned devices, since 2001 following examples of LSDIs have been used on ground, on board aircrafts, balloons and rockets:

Mobile Dosimetry Units MDU-5 and 6 was used for more than 13500 hours between 2001 and 2009 on Czech Airlines (CSA) aircraft at different routes as comparison measurement with aircraft crew individual dosimetry. The experiments and data analysis were managed by Prof. F. Spurny [12];

The Liulin-MDU-2 instrument work successfully during the flight of French balloon up to 32 km altitude in the region of the Gap town in Southern France on 14th of June 2000. This experiment was performed by the Nuclear Physics Institute, Czech Academy of Sciences [13].

One battery-powered LSDIs of Liulin-4J type perform dosimetric measurements of the ionizing radiation environment at ~20 km altitude aboard NASA's Lockheed ER-2 high altitude research aircraft in October-November 2000 from Edwards Air Force Base (AFB) in Southern California and flew over the border region dividing Central California from Central Nevada [14].

Three battery-powered LSDIs were operated during the 8 June 2005 certification flight of the NASA Deep Space Test Bed (DSTB) balloon at Ft. Sumner, New Mexico, USA. The duration of the flight was about 10 hours [15];

Liulin-R was successfully launched on HotPay2 rocket from Andoya Rocket Range (ARR), Norway, on 31st of January, 2008 at 19:14:00 and rising up to 380 km altitude, as a part of an EU financed scientific program called eARI (ALOMAR eARI project) [16];

Liulin-6S, Lilun-M, Liulin-6MB and Liulin-6R are internet based instruments. They use internet module to generate web page. The obtained deposited energy spectra data are transmitted via LAN interface by HTTP and FTP protocols. They worked for different periods since 2005 at Jungfrau (Switzerland) 3453 meters Above Mean Sea Level (AMSL) http://130.92.231.184/), Moussala (Bulgaria) 2925 meters AMSL http://beo-db.inrne.bas.bg/moussala/ and Lomnicky Stit (Slovakia) 2633 meters AMSL http://147.213.218.13/ peaks and at ALOMAR observatory in Norway (http://128.39.135.6/) [17]. The three peak instruments are working well till now (March 2011) and their data can be obtained online on the mentioned above addresses;

Very similar instruments to the Mobile Dosimetry Units MDU-5 and 6 are used by scientific groups in Spain [18] and Germany [19] for radiation measurements at aircrafts.

2.1. Block diagram explanation

LSDI usually contains: one semiconductor detector, one charge-sensitive preamplifier, a fast 12 channel analog-to-digital converter (ADC), discriminator, real time clock, 2 or more microcontrollers and a flash memory. Different modifications of LSDI use additional modules such as: UV sensitive photo diodes, temperature sensor, Global Positioning System (GPS) with antenna and receiver, display (see Figure 2.), multimedia card (MMC) or SD cards. Figure 3 presents a generalized block schema of Liulin type spectrometers.

Fig. 2. LSDI with LCD display

The unit is managed by the microcontrollers through specially developed firmware. Plug-in links provide the transmission of the stored on the memory data toward the standard Personal Computer (PC) or toward the telemetry system of the carrier. A computer program in PC is used for the full management of the LSDI through standard serial/parallel or USB communication port. The same program stores the full data sets on the PC and visualizes the data for preliminary analysis.

Different power supplies were used in the different instruments. They are presented on the upper part of Figure 1 and include 3.6 V or 7.2 V rechargeable or primary batteries, 28 V or 43 V DC aircraft and satellite power and 110 V, 400 Hz AC aircraft power line.

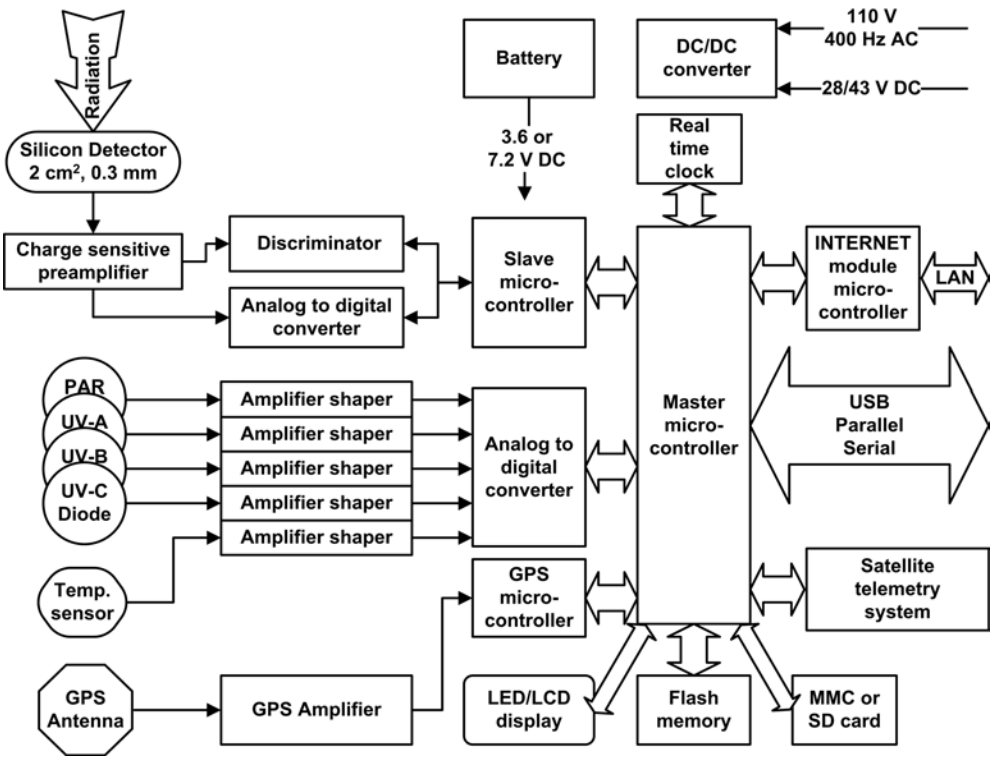


Fig. 3. Generalized block-diagram of Liulin type instruments

1.2. Dose interpretation procedure

The spectrometers measure the pre-amplified amplitude of pulses generated by particles or quantum hitting the detector. The amplitude is proportional by a factor of 240 $mV.MeV^{-1}$ to the energy loss in the detector and respectively to the dose and Linear Energy Transfer (LET). By the 12 bit ADC these amplitudes are digitized and organized in a 256-channel spectrum using only the first 8 bits of the ADC. The dose D [Gy] by definition is one Joule deposited in 1kg. We calculate the absorbed dose by dividing the summarized energy deposition in the spectrum in Joules to the mass of the detector in kilograms.

$$D = K\sum_{i=1}^{256} ik_i A_i m_D^{-1} \quad (1)$$

where m_D is the mass of the detector in kg, k_i is the number of pulses in channel "i", A_i is the amplitude in volts of pulses in channel "i", $K.i.k_i.A_i$ is the deposited energy (energy loss) in Joules in channel "i". K is a coefficient. All 256 deposited dose values, depending on the deposited energy for one exposure time, form the deposited energy spectrum.

In 2001 F. Spurny developed a procedure, which allows the calculation of the ambient dose equivalent from the deposited energy spectrum [20]. The procedure was further developed by O. Ploc [21].

2. Calibration results

LSDIs were calibrated in wide range of radiation fields. First it was irradiated in gamma and neutron (^{137}Cs, ^{60}Co, AmBe, and ^{252}Cf) radiation fields [22]. The calibrations revealed that except for charged energetic particles, the detector has high effectiveness toward gamma rays. Detector's neutron effectiveness depends on their energy [22].

LSDIs have been calibrated in the CERN-EU energy reference field behind the concrete shield [23]. The fluence energy spectra of neutrons registered there are very similar to the spectra on the aircraft and/or balloon [20].

Eight batteries operated LSDIs were tested in CERN-EU high-energy reference field in July 2003. All 8 instruments were irradiated at the same time by exposing their Si-diode surfaces parallel to the concrete wall at the distance of 15 cm between the diodes and the side wall of the concrete

shielding. The relative LSDIs dose rates depend by correction factors from their positions. The correction factors assessment is based on the results obtained by CERN collaborators Mitaroff & Silari [18], taking into account that the most part of doses is due to low LET component (muon background) of the field. The time structure of the beam was within a pulse cycle lasting 16.8 s and particles were impinging on the target for 5.1 s. The level of beam intensity is monitored using a precise ionization chamber (PIC). It is expressed in terms of number of PIC impulses per 1 spill of the accelerator.

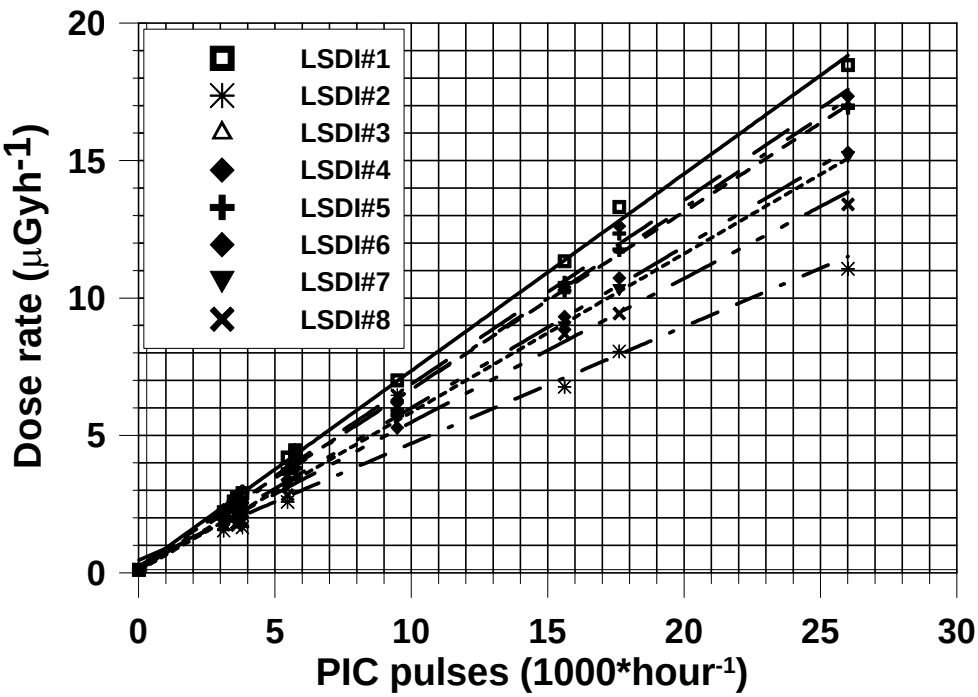


Fig. 4. Linear regressions of the total absorbed dose rates in silicon as a function of the beam intensity for 8 different LSDIs.

Figure 4 represents the obtained by the eight different LSDIs linear regressions of the total dose in Silicon detectors rates as a function of the beam intensity represented with the number of PIC impulses. Nice linear increase of the dose rates in each LSDI is observed when the beam intensity increase. The differences in the slope of the linear curves are found to

depend mainly from the position in the field and therefore their reference values [23].

LSDIs were calibrated, including the dependence of energy deposition in the detector on the direction of the incident radiation at the cyclotron facilities of the Indiana University [24], University of Louvain, Belgium [7] and of the National Institute of Radiological Sciences-STA, Chiba, Japan [25] with protons of energies up to 190 MeV. In all protons calibrations good agreement was obtained between the experimental and simulated with GEANT4 code spectra.

Finally LSDIs was calibrated at heavy ions fluxes up to 500 MeV/u iron ions at HIMAC, Japan irradiation facility [7, 25]. Heavy ion experiments also confirmed that the dosimeter can measure the fragmentation of heavy ions and it was established that instrument's energy resolution is sufficient to distinguish the charge peaks of the individual fragment ions [25].

In-flight intercalibration between LSDI and commercial available TEPC were obtained by Canadian group during a flight from Singapore to London on 2 December 2003 and show nice agreement [26].

4. Experimental Results

4.1. General presentation of the deposited dose spectra shapes and slopes

Figure 5 presents examples of the averaged spectra shapes and slopes from ground, mountain peak Jungfrau, aircraft and spacecraft (Please look the top part of the figure). The individual spectra seen on this figure are obtained after averaging of various numbers of primary spectra and are plotted in coordinates Deposited energy/Deposited dose rate. The main idea of the figure is that the spectra shapes and slopes characterize the predominant type of radiation where the data are taken from [6]. Spectra are grouped by the predominant type of radiation: Lowest blue shadowed is from Galactic Cosmic Rays (GCR), while middle (yellow shadowed) is from protons from South Atlantic Anomaly (SAA). The top magenta shadowed group of curves is from Outer Radiation Belt (ORB) electrons.

From bottom to top the spectra are arranged depending on the value of the deposited dose rates seen in the middle part of the figure. Lowest is from ground natural radiation of 0.12 μGy h^{-1}, while the highest spectra of 9000 μGy h^{-1} is from relativistic electrons measurements at ISS [27].

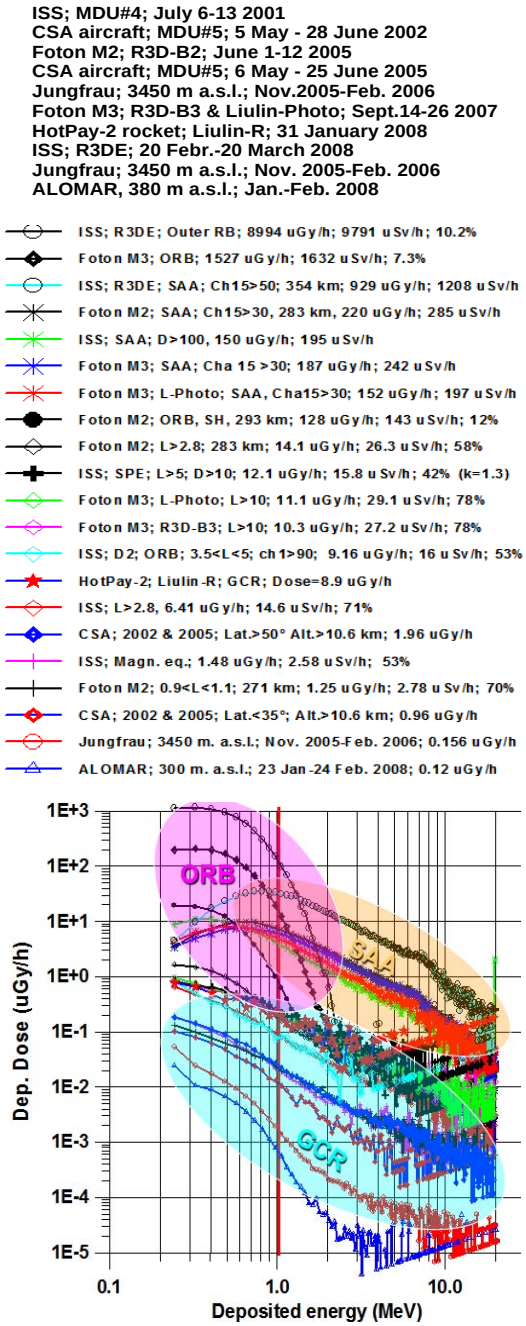


Fig. 5. General presentation of the deposited dose spectra obtained during different experiments in the atmosphere and space.

4.2. Ground based and mountain peak results

The ground based measurements in usual low solar activity conditions obtain the activity from the radioactive isotopes in the ground below the instrument. This type and the Radon gas radiation predominates there and gives about 60% of the total dose, which usual values vary between 0.04 and 0.15 μGy/h. The GCR primary and secondary particles give only 13% of these doses [28]. These type of data were obtained by us by the Liulin-6R instrument located at the ALOMAR observatory at about 380 meters above the see level during the flight of HotPay-2 rocket experiment up to 380 km from Andoya Rocket Range, Norway on January 31st 2008 [16]. As expected on Figure 5 this is the lowest curve.

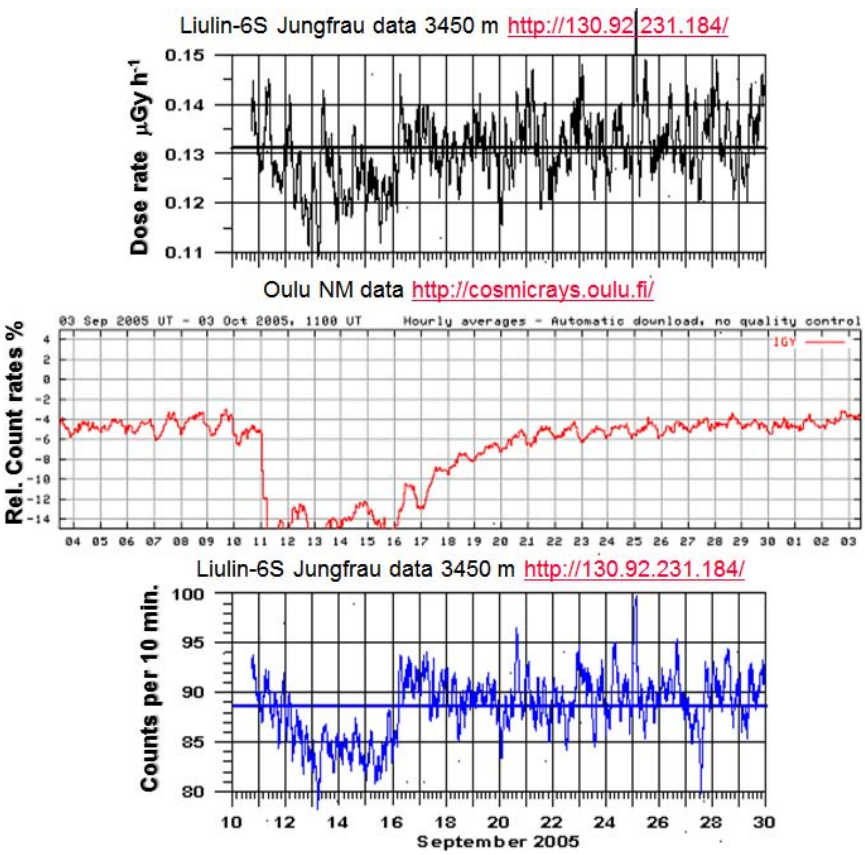


Fig. 6. Comparison of the variations of the Jungfrau dose and count rate with the Oulu NM relative count rate data around the Forbush decrease in September 2005.

The absorbed dose spectra measured at the 3 mountain peaks Jungfrau (3453 m), Mousala (2925 m) and Lomnicky stit (2633 m) are more similar to the aircraft spectra than to the ground based measurements (Please see Figure 5).

The comparison of the Jungfrau count and dose rate data (3450 m) with the Oulu NM http://cosmicrays.oulu.fi/ relative count rate data (Please see Figure 6) obtained around the Forbush decrease in September 2005 show that Liulin type spectrometers can be used effectively to monitor the amount of primary and secondary GCR particles at mountain peaks. The ALOMAR station data don't show dependence by Forbush decrease.

4.3. Aircraft results

The aircraft spectra on Figure 5 are in the middle of the blue shadowed area and are very similar but with lower doses to the spacecraft GCR spectra.

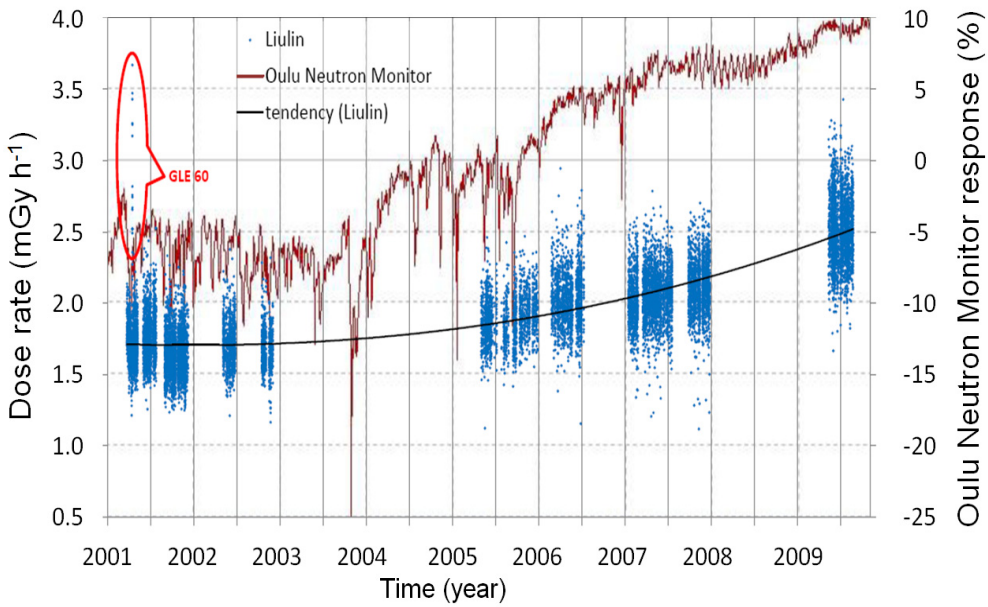


Fig. 7. Variations of the average deposited dose rate for transatlantic flights at altitude 10.6 km

Figure 7 summarizes all data obtained by 2 Liulin type instruments on CSA aircraft between 2001 and 2009. All data in the period 2001-2007 was collected by the MDU#5 instrument. The data in 2009 was measured with a new build instrument, which have almost same characteristics as MDU#5. More than 64000 measurements with 10 minutes resolution are presented on the figure. Each patches of data were obtained in about 1-3 months of continues measurements campaign. Mostly aircraft flights on the destinations Prague - New York and Prague - Toronto at fixed altitude of 35000 feet (10.6km) are used. The cut-off rigidity varies between 0.16 and 2.0 GV when the latitude changes between 50 and 65°.

On the X axis is plotted the date between January 2001 and October 2009. On the left hand Y axis the measured absorbed dose rate in the silicon of the detector is plotted. The right hand Y axis is for the Oulu Neutron Monitor response in percent. The Oulu data http://cosmicrays.oulu.fi/ are seen on the figure as continues heavy black line, which varies in average between -7% in the maximum of the solar activity (2001-2004) and +9% in the minimum of solar activity in 2009.

The Liulin data rises in average from about 1.75 to 2.5 $\mu Gy\ h^{-1}$. This tendency is presented on the plot by polynomial fit of data shown as black line through them. The dose rates obtained during the solar proton event and Ground Level Enhancement on 15th of April 2001 (GLE 60) (Spurny and Dachev, 2001) form the absolute maximum in the data and are specially mentioned in the left hand side of the picture. The increase of the GCR data in 2009 shows single points, which are comparable with those obtained during GLE 60. The calculated apparent dose equivalent dose rates shows very similar to the presented at Figure 1 variations but in an average range from 4-6 $\mu Sv\ h^{-1}$. Some extreme high measurements in 2009 reach values of 11 $\mu Sv\ h^{-1}$.

Figure 8 was specially designed to present how closer the measured GCR dose rates and fluxes on aircraft and spacecraft are. There are 2 panels on the figure. The X axes is for the geographic latitude in the range from 0 to 70° in the Northern hemisphere. The data in the figure are selected from relatively narrow longitudinal range – ±40° from the Greenwich meridian. Two facts allow us to conclude that only GCR data are separated: 1) This latitudinal and longitudinal range is away from the region of the South Atlantic Anomaly (SAA); 2) There are no Solar Proton Events in the mentioned above time intervals.

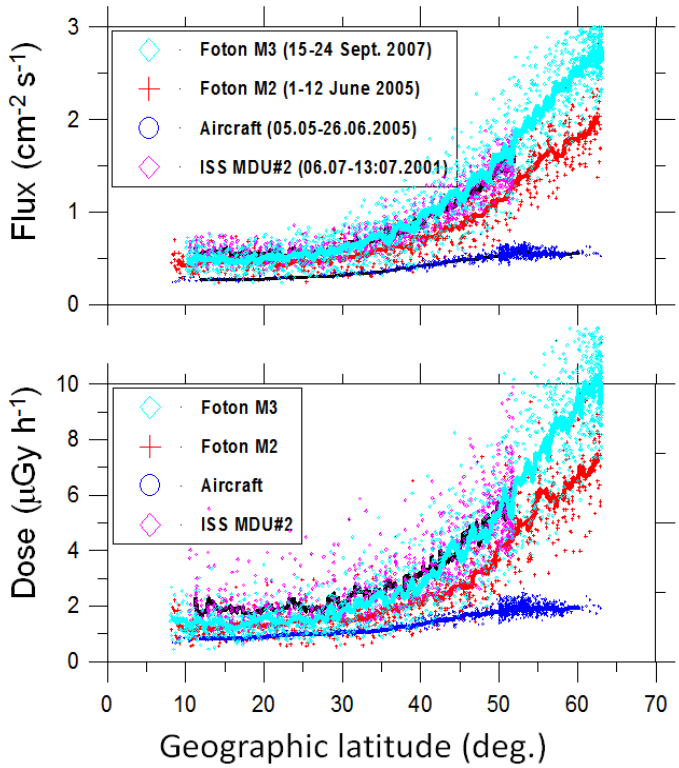


Fig. 8. Latitudinal profiles of the dose rate and fluxes at aircraft, Foton M2/M3 spacecraft and ISS.

In the panels are presented the measured absorbed dose rates (bottom panel) and fluxes (top panel) at 4 vehicles, which data are taken for the periods and altitudes as follows: Aircraft -05.05-26.06 2005 at 10.6 km; Foton M2 1-12 June 2005 at 260 km; Foton M3 15-24 September 2007 at 267 km; ISS (MDU#2) 6-13 July 2001 at 393 km.

The main results from the analysis of Figure 7 are: 1) All latitudinal profiles shows similar shape with minimum at low latitudes and rising values toward high latitudes; 2) In the range 10-30° the values are practically independent from the latitude. The averaged dose rates in this range are 0.66 μGy h^{-1} at aircraft, 1.34 μGy h^{-1} at Foton M3/M3 satellite and 1.93 μGy h^{-1} at ISS. Simple calculations reveal that the ratios of the dose rates in this range at altitudes 10.6, 260 and 393 km are as 1:2:3 i.e. the GCR component of the Earth radiation environment is attenuated only 3

times from the Earth magnetic field and atmosphere on it path from space to the ground; 3) The aircraft dose rates and fluxes shows almost fixed values in the range 50-60°.

4.5. Balloon results

One MDU of the Liulin-4C system has been exposed on the balloon launched the 14 June 2000 at the Gap (France). The altitude, the effective dose profile calculated by means of the CARI-6 code and the dose in Si, D(Si), profiles directly measured with Liulin-4C are presented in the Figure 9 [13]. All dosimetric data well present the Photzer maximums [3] at about 1 hour and 5.2 hours after the launch.

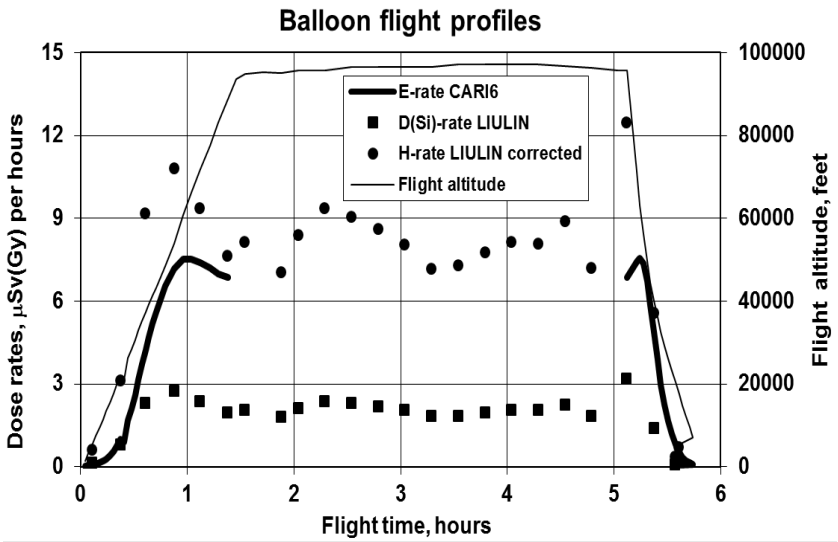


Fig. 9. Profiles obtained by one of the MDUs of Liulin-4C system during balloon flight over the French Gap town on 14th of June 2000.

We transform the dose rates in Si determined on the balloon board to the dose equivalent rates (H-rate) based on the CERN reference field. The results obtained are graphically presented in the Figure 9, together with the values calculated by means of the CARI-6 code [29]. One can see from Figure 9 that the values of dose equivalent deduced from the Liulin-4C data on the base of CERN calibration are much closer to the effective dose

values calculated by means of the CARI6 code. They are generally higher, about 20-30 %, due to the difference in the relative contributions of low and high LET components in on balloon board and CERN reference field.

Three battery-powered LSDIs were operated during the 8 June 2005 certification flight of the NASA Deep Space Test Bed (DSTB) balloon at Ft. Sumner, New Mexico, USA. The duration of the flight was about 10 hours [15].

The DSTB was launched from Ft. Sumner, NM at 09:45 Mountain Daylight Time (MDT). The three Liulin-4 MDUs measured particle flux and dose rate as functions of time at one minute intervals during. Figure 10 shows flux as a function of time as measured by the three MDUs, as well as altitude as provided by the DSTB GPS receiver, together barometric altitude, as functions of time.

All three MDUs measured similar flux and dose rate profiles and these profiles correlate well with the altitude profiles. Following launch at 09:45 MDT, there is a rapid increase in flux and dose rate as the balloon gains altitude. Both flux and dose rate reach a maximum at ~70,000 ft. altitude (21.3 km) and then fall off as altitude continues to increase. The altitude of maximum flux and dose rate is the Photzer Maximum, the altitude at which the showers or cascades of secondary particles produced by primary cosmic rays interacting with the constituent nuclei of the atmosphere are most intense. Shortly before 12:00 MDT, the DSTB attained its maximum cruising altitude of ~120,000 ft. (36.5 km) and both flux and dose rate levelled off. Flux and dose rate remained fairly constant for the remainder of the flight and only began to change at 18:45 when the DSTB gondola was released from the balloon and began its rapid descent toward the ground. The flux and dose rate measured by MDU #5 during the high altitude cruise phase shows considerably more variation than do the measurements made by MDU #1 and MDU #2. This is because MDU#5 was exposed beneath the shielding carousel at the centre of the DSTB platform and the carousel was repeatedly rotated during the flight in order to test its operation. As a result, the shielding environment immediately above the MDU #5 detector repeatedly changed over the course of the flight.

The most interesting observation from these results is that higher values of flux and dose are for MDU #2 under 5 g cm^{-2} Al shielding and for MDU #5 under the carousel, and not for MDU #1 which was relatively unshielded. This result runs contrary to expectations that the larger amounts of shielding would attenuate the flux and thereby reduce the dose rate.

4.3. Rocket results

Liulin-R was successfully launched on HotPay2 rocket from Andoya Rocket Range (ARR), Norway, on 31st of January, 2008 at 19:14:00 and rising up to 380 km altitude, as a part of an EU financed scientific program called eARI (ALOMAR eARI project) [16].

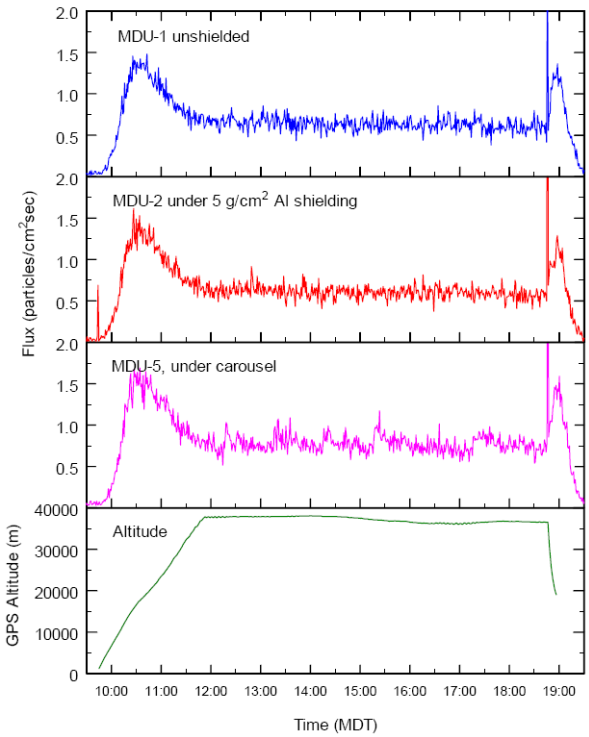


Fig. 10. Flux profiles measured by the three Liulin-4 MDUs exposed during the 8 June 2005 DSTB certification flight. Also shown is the GPS altitude profile in meters.

Figure 10 represent the obtained dose rate (red line) and flux (blue line) data in dependence by the altitude of the rocket. We believe the ascending data up to 150 km of altitude are corrupted by the rocket vibrations, which were infused by the burning in this phase engine. Next between altitude of 200 km up to the apogee of the rocket flight (380 km) the dose rate and flux remain almost fixed in both ascending and descending parts of the flight. The dose rate values of about 10 μGy h^{-1} are close to the

observed at similar altitudes GCR values on spacecraft [6]. Also the obtained deposited energy average spectrum, shown on Figure 5 with blue line and red stars is exactly inside of the bunch of curves from satellites.

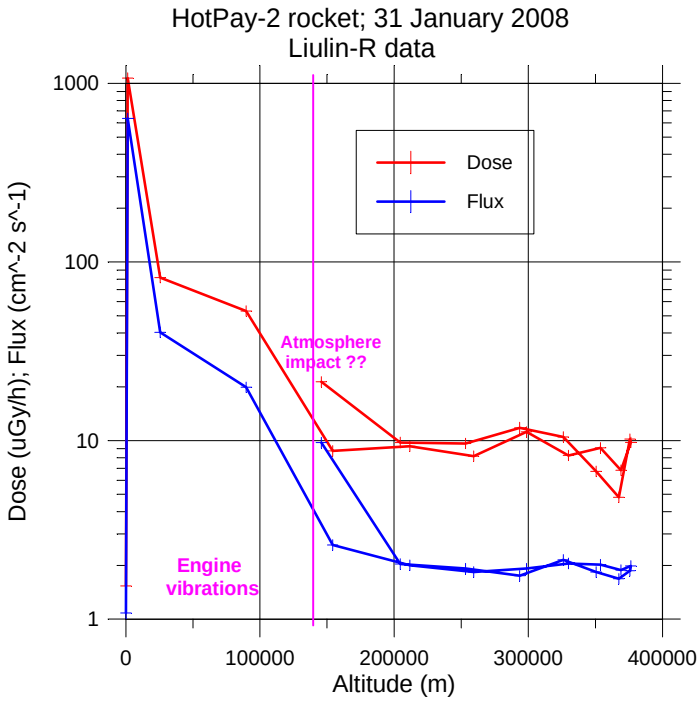


Fig. 11. Liulin-R data during the flight of HotPay-2 rocket experiment up to 380 km on January 31st 2008, Andoya Rocket Range, Norway.

Further when on the descending part of the trajectory the rocket reached the higher atmosphere density regions the data were interrupted probably because of the generated impact in the atmosphere.

Altitudinal, latitudinal or longitudinal dependence of the flux and dose are not observed along the HotPay-2 trajectory above 200 km altitude.

Conclusions

The presented LSDI data at various carriers prove very well the ability of these instruments to be used for monitoring of the atmospheric ionizing radiation environment.

Main advantage of the Liulin type spectrometers are their low weight (~100 g), low power consumption (~100 mW), low cost (~ 10000 Euro). The high scientific and application value of the obtained data is coming mainly from the extensive calibrations at different accelerators and from well-developed data analysis procedures.

Acknowledgments

This work is partially supported by the Bulgarian Academy of Sciences and contract DID 02/8 with the Bulgarian Science Fund.

References

1. W i l s o n, J. W., D. M a i d e n, P. G o l d h a g e n, H. T a i, J. S h i n n, Chapter 2: Overview of Atmospheric Ionizing Radiation. (AIR). http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20030063010_2003072785.pdf

2. De A n g e l i s et al., Radiat. Res., 156, 689-694, 2001.

3. K r a i n e v, M., W. W e b b e r, The medium energy galactic cosmic rays according to the spacecraft and stratospheric data, Preprint № 11, Lebedev Physical Institute, Moscow, 2005.

4. S h e a, M., D. S m a r t, Vertical Cutoff Rigidities for Cosmic Ray Stations Since 1955, 27th International Cosmic Ray Conference, Contributed Papers, 10, 4063-4066, 2001.

5. Group of Experts established under Article 31 of the Euratom treaty, EUROPEAN COMMISSION, Radiation Protection 140, Cosmic Radiation Exposure of Aircraft Crew, Compilation of Measured and Calculated Data, Final Report of EURADOS WG 5 to the Directorate-General for Energy and Transport, Directorate H — Nuclear Energy, Unit H.4 — Radiation Protection, 2004.

6. D a c h e v, Ts. Characterization of near earth radiation environment by Liulin type instruments, Adv. Space Res. 44, 1441-1449, (2009). doi:10.1016/j.asr.2009.08.007

7. D a c h e v, Ts. et al. Calibration results obtained with Liulin-4 type dosimeters. Adv. Space Res. 30, 917-925, (2002). doi:10.1016/S0273-1177(02)00411-8

8. R e i t z, G. et al. Space radiation measurements on-board ISS-the DOSMAP experiment. Rad. Prot. Dos. 116, 374-379, (2005). http://rpd.oxfordjournals.org/cgi/content/abstract/116/1-4/374

9. D a c h e v, Ts. et al. Monitoring Lunar radiation environment: RADOM instrument on Chandrayaan-1, Current Science, 96, 544-546, (2009). http://www.ias.ac.in/currsci/feb252009/544.pdf

10. D a c h e v, Ts. et al. Monitoring of the Earth and Moon radiation environment by the RADOM instrument on Indian Chandrayyan-1 satellite. Preliminary results. 40th

Lunar and Planetary Science Conference, The Woodlands, Texas, USA, March 2-27, 2009. http://www.lpi.usra.edu/meetings/lpsc2009/pdf/1274.pdf

11. D i m i t r o v, Pl., et al. Liulin type spectrometers - last developments. Proceedings of Fundamental Space Research Conference. Sunny Beach, Bulgaria, September 2008, ISSBN 978-954-322-316-9, 334-337, (2008). http://www.stil.bas.bg/FSR/PDF/TOP4Dimitrov_Plamen237123.pdf

12. S p u r n y, F. and D a c h e v T. New results on radiation effects on human health, Acta geophysica, 57, 125-140, 2009. http://www.springerlink.com/content/t2364384842lk5v8/

13. D a c h e v, Ts., F. S p u r n ý, B. T o m o v, P. D i m i t r o v, Y. M a t v i i c h u k, Analysis of the Long Term Dosimetry Results Obtained During Commercial Aircraft Flights, Proceedings of 8th STIL-BAS conference, 131-134, Sofia, December, 2001.

14. U c h i h o r i, Y., et al, Radiation Measurements aboard NASA ER-2 High Altitude Aircraft with the Liulin-4J Portable Spectrometer", Advances Space Research, 32, 41-46, 2003.

15. B e n t o n, E., Deep Space ICCHIBAN: An International Comparison of Space Radiation Dosimeters aboard the NASA Deep Space Test Bed, 10th Workshop for Radiation Monitoring on ISS, Chiba, Japan, 7-9 September 2005. http://www.oma.be/WRMISS/workshops/tenth/pdf/08_benton.pdf

16. T o m o v, B. et al. Galactic and solar cosmic rays study by ground and rocketborne space radiation spectrometers-dosimeters-Liulin-6R and Liulin-R, Proceedings of Fundamental Space Research Conference, ISSBN 978-954-322-316-9, 252-257, 2008. http://www.stil.bas.bg/FSR/PDF/TOP5Tomov_Borislav2242058.pdf

17. D a c h e v, Ts. et al. New Bulgarian build spectrometry-dosimetry instruments–short description, Proceedings of 11-th International Science Conference on Solar-Terrestrial Influences, 195-198, Sofia, 2005. http://www.stil.bas.bg/11conf/Proc/195-198.pdf

18. S á e z V e r g a r a, J. and R. D o m i n g u e z – M o m p e l l R o m á n, The Implementation of Cosmic Radiation Monitoring in Routine Flight Operation of IBERIA Airline of Spain: 1 Y Of Experience of in-Flight Permanent Monitoring, Radiation Protection Dosimetry, 136(4):291-296 2009.

19. W i s s m a n n, F., S. B u r m e i s t e r, E. D ö n s d o r f, B. H e b e r, M. H u b i a k, T. K l a g e s, F. L a n g n e r, T. M ö l l e r, M. M e i e r, Field calibration of dosemeters used for routine measurements at flight altitudes, Radiation Protection Dosimetry, 140, 4, 18 May 2010, Pages 319-325, 2010.

20. S p u r n y, F. and D a c h e v T. On board aircrew dosimetry with a semiconductor spectrometer, Radiation Protection Dosimetry, 100, 525-528, 2002. http://rpd.oxfordjournals.org/cgi/content/abstract/100/1-4/525

21. P l o c, O., F. S p u r n y, Ts. D a c h e v, Use of Energy Depositing Spectrometer for Individual Monitoring of Aircrew, Radiat Prot Dosimetry, 144 (1-4), 611-614, 2011. doi: 10.1093/rpd/ncq505

22. S p u r n y, F., P l o c O. and J a d r n í k o v á I. Spectrometry of Linear Energy Transfer and dosimetry measurements onboard spacecrafts and aircrafts, ISSN 1547-4771, Physics of Particles and Nuclei Letters, 2009, 6, 70–77. © Pleiades Publishing, Ltd., 2009.

23. M i t a r o f f, A. and S i l a r i M. The CERN-EU High-energy reference field (CERF) facility for dosimetry at commercial flight altitudes and in space. Rad. Prot. Dos. 102 7-22 (2002).

24. D a c h e v, Ts. et al. Analysis of the cyclotron facility calibration and aircraft results obtained by LIULIN-3M instrument, Adv. Space Res., 32, 67-71, (2003). doi:10.1016/S0273-1177(03)90372-3

25. U c h i h o r i, Y. et al. Analysis of the calibration results obtained with Liulin-4J spectrometer-dosimeter on protons and heavy ions, Radiation Measurements, 35, 127-134, (2002). doi:10.1016/S1350-4487(01)00286-4

26. G r e e n et al., An empirical approach to the measurement of the cosmic radiation field at jet aircraft altitudes, Adv. Space Res. 36, 9, 1618-1626, 2005.

27. D a c h e v, Ts., B. T o m o v, Yu. M a t v i i c h u k, P. D i m i t r o v, N. B a n k o v, Relativistic Electrons High Doses at International Space Station and Foton M2/M3 Satellites, Adv. Space Res., 1433-1440, 2009. doi:10.1016/j.asr.2009.09.023

28. Ionizing Radiation in our Environment, http://www.who.int/ionizing_radiation/env/en/, March 2011.

29. http://www.faa.gov/data_research/research/med_humanfacs/aeromedical/radiobiology/cari6/

ПРЕГЛЕД НА РЕЗУЛТАТИТЕ ОТ ИЗМЕРВАНЕ НА ЙОНИЗИРАЩИ ЛЪЧЕНИЯ В АТМОСФЕРАТА ПО ДАННИ ОТ ПРИБОРИ, РАЗРАБОТЕНИ В БЪЛГАРИЯ

Ц. Дачев, Б. Томов, Ю. Матвийчук, П. Димитров, Ф. Спурни, О. Плоц, Ю. Учихори, Е. Флукигер, К. Кудела, Е. Бентон

Резюме

Хората са изложени на йонизиращо лъчение през цялото време, и се знае, че то може да предизвика различни вредни биологични ефекти. Следователно е необходимо да се оцени количествено нивото на експозиция от това излъчване, което да е основа за оценка на рисковете за тяхното здраве. Екипажите на космически кораби и самолети са изложени на високи нива космическа радиация от галактически и слънчев произход, както и на вторична радиация, създадена в атмосферата и структурите на превозното средство.

Мониторингът на екипажите на самолети се изисква съгласно препоръките на Международната комисия за радиационна защита (ICRP) (ICRP 1990) и на Директивата на Европейския съюз (ЕС), която въвежда основните ревизирани стандарти за безопасност (ЕС 1997), които включват и експозицията от космическа радиация. Този подход е приет и в други официални документи (NCRP 2002). В този преглед ние представяме резултатите за радиационната обстановка на планински върхове, самолети, балони и ракети, получени с помощта на спектрометри от типа Люлин, които измерват депозираната енергия със силициев диод. Те са разработени в Българска академия на науките (БАН) за целите на радиационния мониторинг на станцията МИР и на Международната космическа станция (МКС). Тези спектрометри-дозиметри са доразвити, калибрирани и използвани от научни колективи от различни страни. Тяхното калибриране е проведено на различни ускорители, включително в CERN във високо-енергийно радиационно поле, което симулира условията на 10 км надморска височина в атмосферата, както и с тежки йони в ускорителя - HIMAC в Чиба, Япония. Дългосрочна база данни е създадена чрез използването на специално разработен батериен прибор на борда на самолети от типа A310-300 на Чешките авиолинии за периода 2001-2009 г. Данните са от 24 сесии, всяка от които по около 2 месеца. Те съдържат повече от 2000 полета (13 500 летателни часа) по маршрути главно над Атлантическия океан. Получените експериментални данни са сравнени с изчислителни модели като CARI и EPCARD. Измерванията на планински върхове са направени с приборите "Люлин-6S", "Люлин-6MB" и "Люлин-6M", които използват интернет модул за генериране на WEB страница, в която се публикуват онлайн получените енергийни спектри, дозата и потока чрез LAN интерфейс с протоколите HTTP и FTP. Данните са за различни периоди между 2005 и 2011 г. на върховете Юнгфрау (3453 метра над морското равнище) (http://130.92.231.184/); на Ломнички щит (2633 метра над морското равнище) (http://147.213.218.13/) и Мусала (2925 метра над морското равнище) (http://beo-db.inrne.bas.bg/moussala/) в Швейцария, Словакия и България съответно. 4 малки по размер батерийни прибора са използвани при полет на балон над южна Франция през юни 2000 г. и в балон на НАСА над Ню Мексико, САЩ на 11 юни 2005 г. 1 е използван в ракетен експеримент през януари 2008 г.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 24, 2012, Sofia

RESEARCH OF THE RELATIONSHIPS BETWEEN LIGHT DISPERSION AND CONTRAST OF THE REGISTERED IMAGE AT DIFFERENT BACKGROUND BRIGHTNESS

Stiliyan Stoyanov, Garo Mardirossian

Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: mardirossian@space.bas.bg

Abstract

The light diffraction is for telescope apparatuses an especially important characteristic which has an influence on the record image contrast from the eye observer.

The task of the investigation is to determine to what degree the coefficient of light diffraction influences the record image brightness.

The object of the theoretical research are experimental results provided from a telescope system experiment in the process of observation of remote objects with different brightness of the background in the fixed light diffraction coefficients and permanent contrast of the background in respect to the object.

The received values and the ratio of the image contrast to the light diffraction coefficient is shown in a graphic view. It's settled that with increasing of the value of background brightness in permanent background contrast in respect to the object, the image contrast sharply decrease. The relationship between the increase of the light diffraction coefficient and the decrease of the brightness of the project image from telescope apparatuses can be observed.

Light dispersion is optical device which characteristics affect the contrast of the image recorded by the observer's eye [1, 5, 6, 8, 10, 11].

Many objects disperse light falling onto them, so the brightness's values along the various directions appear to be strong. According to Lambert's law [2], the brightness of a light-dispersing surface is equal in all directions. This assertion may be assumed only as an approximation.

Let σ be a small area with brightness β equal in all directions.

The light flow ψ emitted from area σ along the normal constituent of angle φ is calculated. Isolating the bodily angle dφ located between two ring cones, generated by the rotation about normal N of two lines forming angles φ and φ + dφ, produces apparently:

(1) $d_{\varphi} = 2\pi.\sin\varphi\, d\varphi$.

The light intensity within this spatial angle is constant. Therefore, the light flow within the bodily angle dφ will be:

(2) $d\phi = I_{\varphi} d\varphi = 2\pi B\sigma \sin\varphi\cos\varphi\, d\varphi$.

To determine the light flow ψ emitted by the area within the whole hemisphere, the above expression must be integrated within the limits from 0 to π/2. Then: ψ = πBσ.

(3) $M = \frac{\phi}{\sigma} = \pi B$.

The above shows that to brightness $B = 1\,\text{cd/m}^2$ corresponds lightness: $M = 3{,}14\,\text{lm/m}^2$.

The surface properties of each diffusely dissipating body differ greatly from those of the ideal light dissipater, i.e., the brightness in the various directions is different. To provide numerical characteristics of surface brightness change in various directions, the light dissipation factor for a given surface is used, i.e., the ratio of the brightness of the surface along an arbitrary direction and the brightness of an ideal dissipater, placed under the same illumination conditions. The light dissipation factor is usually denoted by β [9].

The task is to investigate whether the dissipation factor β affects the brightness of the recorded image.

The subject of theoretical research are the results obtained by an experiment with observation telescopic system [3] represented in Table 1 during the observation of remote objects with various background brightnesses ranging between 10^{-2} and 10^{-3} cd/m^2 with given light dissipation factors: $\beta_1 = 0.1$; $\beta_2 = 0.2$; $\beta_3 = 0.3$ and constant contrast of the object's background K = 0.3.

Table 1

Light dissipation factor							X
	β_ϕ = 0,01	β_ϕ = 0,1	β_ϕ = 1	β_ϕ = 10	β_ϕ = 100	β_ϕ = 10^3	
β_1 = 0,1	0,2999	0,2981	0,2431	0,0901	0,0125	0,0013	$\overline{X_1}$ = 0,1566
β_2 = 0,2	0,2986	0,2868	0,2054	0,0517	0,0064	0,0006	$\overline{X_2}$ = 0,1415
β_3 = 0,3	0,2979	0,2806	0,1124	0,0379	0,0043	0,0004	$\overline{X_3}$ = 0,1239
							$\overline{X}$ = 0,1407

In the last column of Table 1, the obtained data is presented, considered as values of the brightness x for the group of factors β_1, β_2, β_3, i.e., z = 3, where the mean group values are denoted by $\overline{X_1}$, $\overline{X_2}$, $\overline{X_3}$ and the overall mean value $\overline{X}$ for the considered brightnesses n = 6 are calculated using formulae [7]:

(4) $$\overline{X} = \frac{1}{2}\sum_{i=1}^{z} xij \qquad i = 1,2...z$$

(5) $$\overline{X} = \frac{1}{zn}\sum_{i=1}^{z}\sum_{i=1}^{n} = \frac{1}{z}\sum_{i=1}^{z}\overline{xi} \ .$$

The hypothesis H which must be verified suggests that the light dissipation factor β does not affect brightness, while the alternative hypothesis suggests the opposite. To check up the zero hypotheses H, the averaged data from the 18 performed studies must be processed. The data processing includes calculation of the square sums $\varsigma, \varsigma_A, \varsigma_R$ using formulae:

(6) $$\varsigma = \sum_{i=1}^{Z}\sum_{i=1}^{n}\left[x_{ij} - \overline{x}\right]^2$$

(7) $$\varsigma_A = \sum_{i=1}^{z}\sum_{i=1}^{n}\left[\overline{x} - \overline{x}_i\right]^2 = n\sum_{i=1}^{z}(\overline{x} - \overline{x}_i)^2$$

(8) $$\varsigma_R = \sum_{i=1}^{z}\sum_{i=1}^{n}(x_{ij} - \overline{x}_i)^2$$

while the dispersions S^2, S_A^2 and S_R^2 are evaluated using formulae [4]:

(9) $$S^2 = \frac{2}{\nu} = \frac{2}{k\pi - 1}$$

(10) $$S_A^2 = \frac{\varsigma_A}{\nu_A} = \frac{\varsigma_A}{k-1}$$

(11) $$S_R^2 = \frac{\varsigma_R}{\nu_R} = \frac{\varsigma_R}{k(\pi - 1)} \ .$$

The obtained values are shown in Table 2.

Table 2

Types of square sums	Square sum	Degree of freedom	Dispersion evaluation
Total	ς = 0,331486	ν 6 = 17	S^2 = 0,019499
By factors	ς_A = 0,000537	ν_A = 2	S_A^2 = 0,000268
Residual	ς_R = 0,022063	ν_R = 15	S_R^2 = 0,022063

The calculation of the disperse ratio F is performed using formula:

(12) $$F = \frac{S_A^2}{S_R^2} = 0,\ 0121831.$$

The obtained disperse ratio (12) is compared with the table value F_T at significance level $\alpha = 0.05$ [2] and it is observed that $F > F_T$, which evidences that light dissipation affects image brightness.

Accounting to the fact that the contrast K depends on the object's brightness B_{ob} and the background B_b, K may be determined from:

(13) $$K = \frac{B_{об} - B_{ф}}{B_{ф}}$$

and, accounting to the additional brightness ΔB, due to light dissipation, which may be written as:

(14) $$\Delta B = (B_{об} + B_{ф})$$

the contrast of the image K′ recorded by a visual optic system during observation of a remote object will be equal to:

$$K` = \frac{B_{об} - B_{ф}}{B_{ф} + \beta(B_{об} + B_{ф})} = \frac{K}{1 + \beta(B_{об} + B_{ф}} = \frac{K}{1 + \Delta B}. \quad (15)$$

From expression (15) it follows that, with definite object contrast with respect to the surrounding background, the image contrast K will be reduced, while the light dissipation factor increases.

In Fig. 1, the curves for the appropriate dissipation factors are shown. Apart from the image contrast's reduction with the light dissipation factor β's increase, the curves presented in Fig. 2 also reveal that the contrast K′ of the recorded image drops abruptly when the background's brightness exceeds (24…30) cd/m^2, i.e., the specified background contrast with respect to the object, which is 0.3, does not provide proper image of the observed remote objects. Therefore, at some given contrast of the object with respect to the surrounding background, the contrast of the recorded image K′ is reduced while the light dissipation factor increases. At background brightness within the range from 10^{-2} to 10^3 cd/m^2 it may be shown that, when the value of background brightness increases, while the background contrast with respect to the object K = 0.3, the image's contrast drops abruptly.

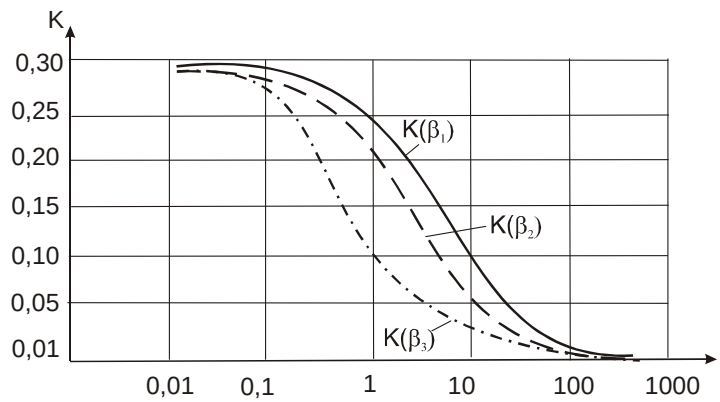


Fig. 1. Dependence of the image contrast on the light dissipation factor

The graphic relationship displays reduction of the image contrast with increase of the light dissipation factor β. Moreover, when the background's brightness exceeds 24…30 cd/m^2, the contrast of the recorded image drops abruptly.

References

1. Г е ц о в, П. Космос, екология, сигурност. НБУ, София, 2002, 211 с.
2. Д н е п р о в с к и й, С., Контрастность приборов, Оптико-механическое производство, под ред. Фрайберг М. Оборонгиз, 1998.
3. Ж е к о в, Ж. Оптични методи и средства за откриване на отдалечени обекти от борда на космически летателни апарати. Универс. изд. "Еп. К. Преславски", Шумен, 2006, 308 с.
4. К е н д а л л, М., А. С т ю а р т. Статистические выводи и связи, Наука, М., 1973.
5. М а н е в, А., К. П а л а з о в, С. Р а й к о в, В. И в а н о в. Комбиниран спътников мониторинг на температурната аномалия през август 1998 г., "Сборник доклади на IX-та национална конференция с международно участие. Основни проблеми на слънчево–земните въздействия, 21–22 ноември 2002 г., София, стр. 153 – 156.
6. М а р д и р о с я н, Г. Аерокосмически методи в екологията и изучаването на околната среда. Акад. издат. „Проф. Марин Дринов", 2003, 208 с.
7. С т о я н о в, С. Оптоелектронен спектрофотометър за изследване на атмосферния озон. Докторска дисертация. София, СНС по ВИН при ВАК, 2010, 168 с.
8. С т о я н о в, С. Оптични методи за изследване на атмосферния озон, Изд. "Фабер", Велико Търново, 2009, 231 с.
9. С т о я н о в, С. Приложна оптика, Изд. „Фабер", Велико Търново,2009,234 с.
10. С т о я н о в, С., И. Х р и с т о в, Г. М а р д и р о с я н, Ж. Ж е к о в. Оптико-електронен тракт в апаратура за регистриране на оптични атмосферни емисии. Сб. научни трудове „Научно технологичен трансфер", ИКИ – БАН, 2000, с. 42 – 45.

ИЗСЛЕДВАНЕ НА ЗАВИСИМОСТИТЕ МЕЖДУ СВЕТОРАЗСЕЙВАНЕТО И КОНТРАСТА НА РЕГИСТРИРАНОТО ИЗОБРАЖЕНИЕ ПРИ РАЗЛИЧНА ЯРКОСТ НА ФОНА

С. Стоянов, Г. Мардиросян

Резюме

Светоразсейването е важна харктеристика на оптичните уреди, която влияе върху контраста на регистрирания образ. Повърхността на всяко дифузно разсейващо тяло в значителна степен се различава по свойства от идеалния светоразсейвател, т. е. яркостите в различни посоки се оказват различни. За да се характеризира числено изменението на яркостта на повърхнина в различни направления, се използва коефициент на светоразсейване за дадена повърхност, като се разбира яркостта на тази повърхност, в произволна посока, към

яркостта на идеален разсейвател, намиращ се в същите условия на осветеност.

Обект на разработката са резултати от изследване на отдалечени обекти при различна яркост на фона с определени коефициенти на светоразсейване.

Получените стойности и отношението на контраста на образа към светоразсейването е представено в графичен вид. Наблюдава се връзка между нарастване на коефициента на светоразсейване и намаляване на яркостта на образа. От представената графика се вижда, че освен намаляване на контраста на образа с нарастване на коефициента на светоразсейване β, контрастът $K^{|}$ на регистрирания образ рязко спада над яркост на фона (24... 30) cd/m^2 т. е. зададеният контраст на фона спрямо обекта 0,3 не осигурява качествен образ на наблюдавания отдалечен обект. Следователно при определен контраст на обект спрямо заобикалящия го фон, контрастът на регистрирания образ $K^{|}$ се намалява с нарастване на коефициента на светоразсейване. При яркост на фона в диапазона от 10^{-2} до 10^3 cd/m^2 се установява, че при нарастване на стойността на яркостта на фона, при постоянен контраст на фона и обекта К = 0,3, контрастът на образа стремително спада.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 24, 2012, Sofia

STUDYING THE RELATIONSHIP BETWEEN SOME ATTRIBUTES OF CONIFEROUS FORESTS AND SPECTRAL DATA FROM THE ASTER SATELLITE SENSOR

Petar Dimitrov, Eugenia Roumenina

Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: petarkirilov@mail.bg

Abstract

The paper presents the results from a study aiming to assess the relationship between spectral data from the satellite sensor Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) and five forest attributes measured in 29 plots in coniferous stands in the Rila Mountain. Biomass, volume and canopy cover showed moderate correlation with the radiance in the near-infrared band. Stand density was weakly correlated with spectral data, and basal area did not show statistically significant correlation. The relationship between ground-based and satellite data was modelled by ordinary least squares (OLS) and reduced major axis (RMA) regression. The models for predicting the different forest attributes had relative standard error of estimate between 14.4% (for canopy cover) and 51.8% (for volume).

1. Introduction

The information demand for different forest attribute data used to characterize productivity, structure and environmental functions of forests have increased in the recent years, along with the need for sustainable forest management. Remote sensing offers increasing variety of data types and the potential of these data for assessing forest characteristics is now extensively studied (Lefsky et al., 2001). Optical remote sensors are the most commonly used for forest research. Sensors like NOAA AVHRR are very useful to assess forest biomass at global scale (Dong et al., 2003). For studies at larger scale, the use of high-resolution data, such as Landsat, ASTER or

SPOT is needed. These sensors are suitable to map the highly fragmented forests in Bulgaria.

Forest attributes cannot be measured directly by remote sensing methods. However, they are related with stands' physiognomy and architecture, which in turn influence the reflectance characteristics captured by optical sensors (Franklin 2001). The reflectance in different spectral bands is correlated with ground-measured forest attributes, allowing their assessment and mapping by multispectral images. This empirical approach is widely used by the remote sensing community (Muukkonen and Heiskanen, 2005; Zheng et al., 2004; Hall et al., 2006).

To determine correlations, ground data at both plot level (Anaya et al., 2009; Healey et al., 2006) and stand level are used. In this work, data from a set of ground plots are used, because this allows gathering data for more forest attributes than these available in the stand datasets provided by forest authorities.

Coniferous forests in Bulgaria cover over 1.2 million hectares, which represents 31 % of the forest area (NSI, 2008). These forests are very important not only as a considerable timber resource, but also for their soil protecting, water regulating and other ecological functions. To characterize and manage them, information for many parameters is required. Dendrometry parameters describing the overstory layer, for example canopy cover, density, basal area, volume and biomass, are most commonly measured and used in practice. These parameters are related with many processes and forest functions; for example, canopy cover affects surface runoff and natural regeneration of stands (Raev, 1980; Stoyanova, 2006). Recently, increasing attention has been paid to the assessment of forest biomass for ecosystem productivity and carbon cycle studies (Houghton, 2005).

Information about forest attributes is obtained from the forest inventory, which provides accurate and relatively exhaustive data on a regular basis. In Bulgaria, forest inventories are carried out usually at ten-year intervals and data are gathered on stand level. Because of its flexibility (possibility for more frequent updates over arbitrary territory), remote sensing will be used in future more often as an additional source of forest information.

The objective of the study is to assess the relationship between some commonly used forest attributes and the spectral data from the Advanced

Spaceborne Thermal Emission and Reflection Radiometer (ASTER) instrument onboard the Terra satellite for coniferous forests.

The study area comprises part of the north-west section of the Rila Mountain and falls within the boundaries of the *Govedartsi* forestry section withn the *Borovets* State Forestry. The region features several types of coniferous forests (Petkov et al., 1966), comprising mainly Scots Pine (*Pinus sylvestris* L.) and Norway Spruce (*Picea abies* (L.) Karst.) and more rarely, Silver Fir (*Abies alba* Mill.) and Macedonian Pine (*Pinus peuce* Griseb.).

2. Materials and methods

2.1. Data

This study uses two types of data – ground-based and satellite. Ground-based data comprise measurements carried out in 29 test plots. The measurements were carried out in 2009 and 2010, whereas the test plots were distributed throughout the territory, so that stands of various age and biomass were comprised. Accounting for the inevitable errors during satellite images' geometric correction, the aspiration in specifying the test plots was that they fell among homogeneous stands, far from roads, cuttings or other features. Depending on the trees' age and density, the plots were sized between 5x5 m and 30x30 m and were commensurate with the ASTER pixel size. In each plot, the species and the diameter at breast height (dbh) of all trees higher than 2 meters were marked. The trees' height was measured by a heightmeter, whereas only in some plots, the heights were estimated by a stand height curve. The collected data were used to calculate the basal area (m^2/ha), the density (pcs/ha), the stem volume (m^3/ha) and the biomass (t/ha). The volume and the biomass were calculated by methods described by Beruchashvili and Zhuchkova (1997). In it, the stem volume is found as a function of the height, the breast diameter, and the stem form quotient. The biomass is calculated by multiplying the stem's volume by the density of the relevant tree species, adding to it the biomass of the branches, leaves and roots, determined as a percent of the stam's mass for the relevant tree species. Moreover, for each plot, the canopy cover was determined based on photos taken by a digital camera assembled on a tripod and levelled so that the optical axis was vertical (Cohen et al., 2003). The photos were taken by a standard zoom lens, adjusted so as to ensure a view angle of 30° along the frame's short side. On the smaller test plots, one photo was taken in the centre of the plot, and on the larger test plots – four non-

overlapping photos were taken. The canopy cover percent was determined by a round pallette (of diameter equal to the length of the frame's short side) imposed on the frame's centre. The statistical description of the data collected from the test plots is presented on *Table 1*.

The ASTER image used for the analysis dates from 1 October 2008 and features processing level 1A. Regretfully, in the images obtained after April 2007, the bands from the short-wave infrared range are unusable on account of a fault in the SWIR sensor. Therefore, the analysis uses only the three bands from the visible and near-infrared range (VNIR) (the thermal bands are not subject to this study). The three VNIR bands (band1-green (0.52-0.60 μm), band2-red (0.63-0.69 μm) and band3- near-infrared (0.76-0.86 μm)) were orthorectified in the ENVI software product as one HDF file. For the purpose, the SRTM digital elevation model (DEM) was used, with cell size of 90 m (USGS, 2006) and 9 GPS ground control points, identified on the red band. The orthorectification error was 2.2 pixels. ENVI applied automatically the calibration coefficients provided in the HDF file and converted the data into 32-bit radiance values ($W/m^2/sr/\mu m$). The image was resampled after the nearest neighbour method and had pixel size of 15 m.

Table 1. Statistical description of the data collected from the test plots

	Min	Max	Mean	St.dev.
Volume ($m^3\,ha^{-1}$)	81	983	480,0	287,1
Biomass ($t\,ha^{-1}$)	41	518	252,3	150,9
Density ($\#\,ha^{-1}$)	175	19200	3078	5030,9
Basal area ($m^2\,ha^{-1}$)	23,2	84,4	49,4	16,3
Canopy cover (%)	51,1	96,4	76,9	13,9

2.2. Data processing and analysis

To assess the relationships between ground-measured parameters and satellite data, the radiance values from the image's three spectral bands for the relevant pixel in which each test plot falls were derived. Apart from the three spectral bands, to assess the forest attributes, two spectral indices were used as well: Normalized Difference Vegetation Index (NDVI = band3-band2/ band3+band2) and Simple Ratio (SR = band3/band2) (Tucker 1979).

The relationship between ground-based data and satellite variables was assessed through the correlation coefficient (*r*) (*Table 2*) and through

graphs (*Figure 1*), after which for part of the forest attributes, regression models were developed. Models were developed only for the canopy cover, the biomass and the volume, since it was established that the relation of the basal area and stand density and the ASTER data is insignificant (*Figure 1*, *Table 2*).

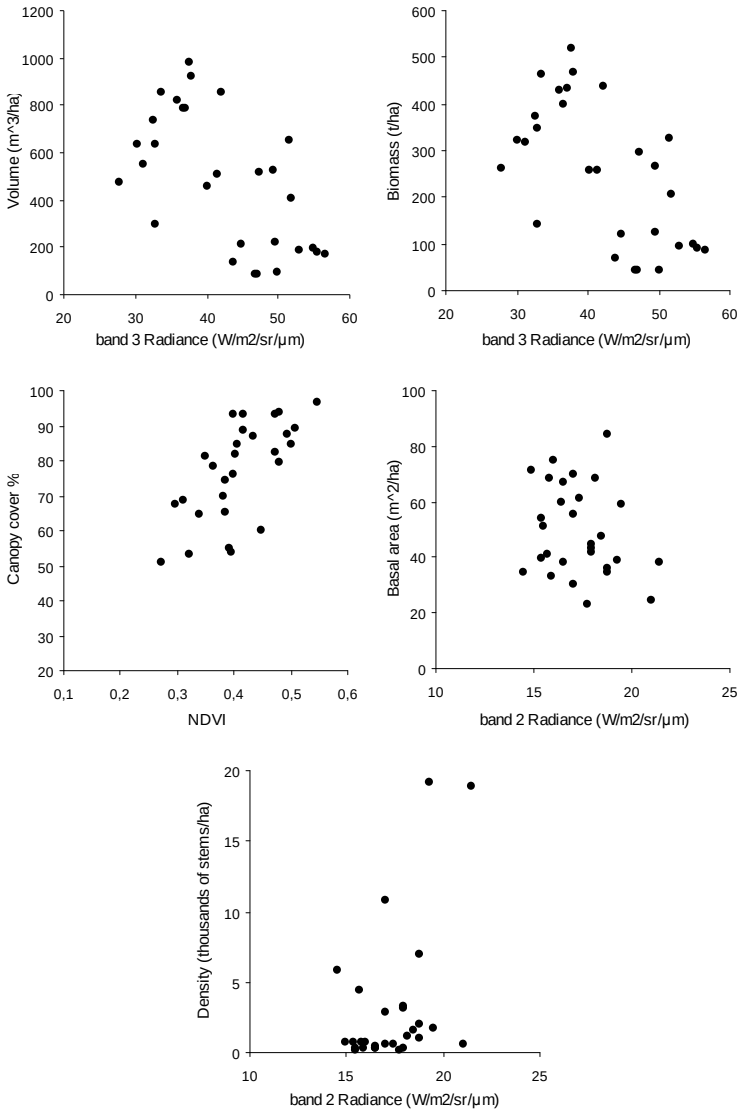


Fig. 1. Scatterplots of ground-measured forest attributes against the corresponding satellite variables, most strongly correlated with them.

To model statistically the relationships between forest attributes and spectral data, linear regression analysis was used since the established relationships were linear or close to the linear. Two types of regression were tested – the traditional ordinary least squares (OLS) method and the reduced major axis (RMA) method. The reduced major axis regression is used, when the measurement errors in the independent variable should be accounted for. Regression analysis by this method was made using a free-of-charge web-based application (Bohonak and Kim van der Linde, 2004).

To predict each forest attribute, the three spectral bands and spectral indices were tested consecutively as independent variables. In the final version, the band or index for which the coefficient of determination (r^2) was highest was chosen. With all forest attributes, the best predictor was the spectral band or index which is most strongly correlated with them. Addition of a second independent variable did not improve the models.

As predictors in the models, the third band (x_3) and NDVI (x_{NDVI}) were used. The volume and the biomass models have the following form:

(1) $$y = a + b * x_3,$$

and this for the canopy cover is:

(2) $$y = a + b * x_{NDVI},$$

where *a* and *b* are regression parameters.

2.3. Validation

To assess the accuracy of regression models, cross validation procedure was used. In it, the value of each test plot is predicted based on the observations from the other plots. This is necessitated by the small volume of the sample which does not allow part of it to be allocated for validation. The accuracy of the models was assessed by the standard error of estimate (*s*):

(4) $$s = \sqrt{\frac{\sum (y - y')^2}{(n-2)}},$$

where:

y is the actual value, *y′* is the predicted value, and *n* is the number of observation pairs. To allow for comparing the errors between the forest

parameters, the standard error of estimate was also presented as a percentage of the parameter's mean actual value (s_r).

3. Results

The radiance in the near infrared band (band3) is most strongly correlated with the volume and the biomass measured on the plots, the correlation coefficients being accordingly -0.60 and -0.62 (*Table 2*). The strongest correlation with canopy cover is demonstrated by NDVI (r=0.67). The density and especially, the basal area, demonstrate very poor correlation with all satellite variables. In all cases, the correlation coefficients are not high which is also prompted by the scattered distribution of the graph points in *Figure 1*. A relatively clearer linear relationship is observed only between canopy cover and NDVI.

Table 2. Correlation coefficients of forest attributes and spectral data

	band1	band2	band3	NDVI	SR
Volume (m^3ha^{-1})	-0.467*	-0.451*	**-0.603*****	-0.453*	-0.480**
Biomass (t ha^{-1})	-0.497**	-0.490**	**-0.618*****	-0.449*	-0.475**
Basal area (m^2ha^{-1})	-0.136	**-0.234**	-0.211	-0.072	-0.123
Density (#/ha)	0.389*	**0.432***	0.373*	0.245	0.224
Canopy cover (%)	0.247	0.109	0.553**	**0.667*****	0.655***

*p < .05 **p < .01 ***p < .001

Table 3 presents the parameters and the accuracy assessments of traditional regression models (ordinary least squares method) for the three forest attributes most strongly correlated with satellite data. Satellite spectral data explain only between 36% and 45% of the variations of the three forest attributes, the canopy cover model featuring the highest r^2, while timber volume coefficient of determination is lowest. The relative standard error of estimate varies between 13.7% and 48.6% for the different parameters. The cross validation results provide somewhat higher values for the relative standard error of estimate.

Table 4 presents the regression results after the RMA method. Again, the canopy cover model features the smallest error, and the stem volume model features the greatest error. During the validation, as well as during the modelling itself, s_r are higher with this type of regression, than with the ordinary least squares regression.

Table 3. Results from the ordinary least squares (OLS) regression

	Model				Validation	
	Parameters*	r^2	s	s_r	s	s_r
Volume (m^3ha^{-1})	a =1342.1 b =-20.3	0.36	233.3	48.6	248.4	51.8
Biomass (t ha^{-1})	a =717.4 b =-10.9	0.38	120.8	47.9	128.6	51.0
Canopy cover (%)	a =21.5 (p=0.09) b =135.2	0.45	10.5	13.7	11.1	14.4

* The p values lower than 0,001 are not shown; s-standard error of estimate; s_r-relative standard error of estimate.

Table 4. Results from the reduced major axis (RMA) regression.

	Model				Validation	
	Parameters	r^2	s	s_r	s	s_r
Volume (m^3ha^{-1})	a =1911.0 b =-33.7	0.36	260.7	54.3	276.9	57.7
Biomass (t ha^{-1})	a =1004.0 b =-17.7	0.38	134.3	53.2	142.5	56.5
Canopy cover (%)	a =-6.1 b =202.5	0.45	11.5	15.0	12.1	15.8

4. Discussion and conclusion

4.1. Correlations

All considered characteristics of the stand except for the basal area are statistically significantly correlated with at least one of the three ASTER bands. Nevertheless, the values of r are low, since no strict linear relationship is observed (*Figure 1*). The near-infrared band appears to be the best volume and biomass predictor. On the overall, the spectral indices NDVI and SR manifest significantly lower correlations with forest attributes compared to the spectral bands themselves. Only canopy cover is strongly correlated with NDVI.

The volume and the biomass of the studied coniferous stands are negatively related to spectral bands and indices. With growth of the trees and increase of the forest's age, the reflectance in the visible spectrum range decreases. The same holds for the reflectance in the near-infrared range, although green vegetation is a good reflector in this part of the spectrum. The reason for this tendency is that the reflectance in the near-infrared range

is strongly affected by the shadows between the individual tree crowns, which take shape and become thicker with increase of the forest's age. Canopy cover is positively correlated with spectral bands. This contradicts previous studies (Hall et al., 2006), where negative correlation has been established. The discrepancy may be explained by the different structure and peculiarities of the forests in the examined regions. In this study, forests with great (70-80%) canopy cover prevail, while the stands examined by Hall et al. (2006) feature mostly canopy cover between 30 and 60%. Actually, the correlation established in this study represents only the interval for canopy cover between 50 and 100%. The physical explanation of the observed positive correlation is related with the fact that the greatest canopy cover is typical for young planted stands which look bright on the image. The forests with smaller canopy cover look darker, since these are mostly mature stands where the shadowing of the adjacent crowns and soil is strong.

4.2. Assessment of the regression models

To model the relationships between forest attributes and spectral characteristics, liner regression was used. The obtained standard errors of estimate are high, but close to those from previous studies (Hyyppä et al., 2000). Thus, for instance, the non-linear model used by Muukkonen and Heiskanen (2005) to predict volume features relative RMSE error of 44.8% (calculated using independent data), while in this study, the relative standard error of estimate is 51.8 % (*Table 3*). In the first case, however, the authors use data at stand level.

The cross validation of the biomass assessment model showed that the relative error is 51%. For comparison, the error with a similar study using data from Landsat TM is 47% (Lefsky et al., 2001). Canopy cover assessed using ASTER data has relative error of 14.4%. For comparison, Hall et al. (2006) model the canopy cover with relative RMSE error of 12%, using bands 3 (0.63-0.69 μm), 4 (0.75-0.90 μm) and 7 (2.09-2.35 μm) of Landsat ETM+.

4.3. Comparison of both regression methods

The accuracy of the models, discussed in the previous section, refers to the results from the ordinary least squares regression. The use of the reduced major axis (RMA) method results in greater standard errors which

might be expected. Nevertheless, RMA regression is often used with satellite data to determine different forest characteristics (Healey et al., 2006). The advantage of this method is that, with it, the measured quantity variance is preserved during its modelling (Cohen et al., 2003). As may be seen from *Figure 2,* with the ordinary least squares method, the modelled biomass values feature smaller variance compared to the real ones. On the

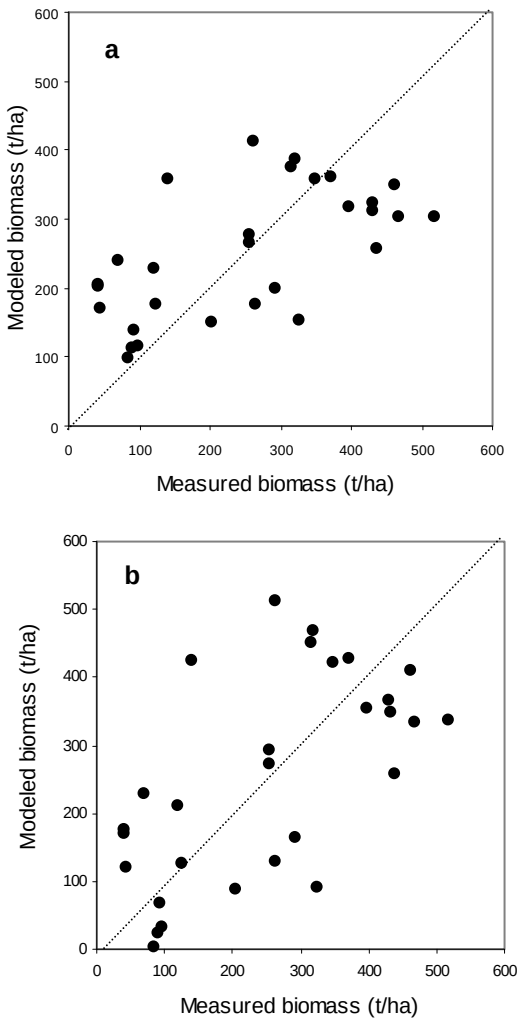


Fig. 2. Measured versus modelled biomass; (a) after the ordinary least squares method, and (b) after the reduced major axis method

other hand, with the reduced major axis method, the variation of the real values is preserved during modelling. Cohen et al. (2003) also point out that when there are errors in the measurement of both the independent and the dependent variable, it is not reasonable to use the ordinary least squares method. This is exactly the case when applying remote sensing to determine forest vegetation characteristics, since errors are observed both in ground-based data, as well as in the reflectance measured by the satellite sensors. Both methods have their advantages and disadvantages; therefore, they are used simultaneously in remote sensing.

***

The established correlations between the attributes of coniferous forests in the Rila Mountain and spectral data from the ASTER satellite sensor comply with the results under similar conditions in the boreal forests of Europe and North America. Most forest attributes are most strongly correlated with the radiance values in the near-infrared band. The absolute value of the maximal correlation coefficients for the various parameters varies between 0.23 (for the basal area) and 0.67 (for the canopy cover), while the relative standard errors of estimate lie in the interval between 14.4% (for the canopy cover) and 51.8% (for the volume). The results show that ASTER data may be used with the greatest fidelity to assess the canopy cover of coniferous forests in the region. The achievement of more accurate assessments, with admissible error levels for various applications, requires improvement of data geometric precision, and the use of greater number of spectral bands. Studies show that predictions accuracy increases when, besides spectral data, additional information is used, such as stand age map (Zheng et al., 2004) or canopy cover map. These possibilities will be examined in a future publication.

References

1. A n a y a, J., E. C h u v i e c o, A. P a l a c i o s-O r u e t a. Aboveground biomass assessment in Colombia: A remote sensing approach. *Forest Ecology and Management*, 257, 2009, 1237–1246.
2. B e r u c h a s h v i l i, N., V. Z h u c h k o v a. Methods for complex physical-geographical investigations: The manual. Moscow University Press, Moscow, 1997, 320 p. (In Russian).
3. B o h o n a k, A. J. & K i m v a n d e r L i n d e. 2004. RMA: Software for Reduced Major Axis regression, Java version. Website: http://www.kimvdlinde.com/professional/rma.html

4. C o h e n, W. B., T. K. M a i e r s p e r g e r, S. T. G o w e r, D. P. T u r n e r. An improved strategy for regression of biophysical variables and Landsat ETM+ data. *Remote Sensing of Environment*, 84, 2003, 561–571.
5. D o n g, J., R. K a u f m a n n, R. M y n e n i, C. T u c k e r, P. K a u p p i, J. L i s k i, W. B u e r m a n n, V. Alexeyev, M. Hughes. Remote sensing estimates of boreal and temperate forest woody biomass: Carbon pools, sources, and sinks. *Remote Sensing of Environment*, 84, 2003, 393– 410
6. F r a n k l i n, S. E. Remote sensing for sustainable forest management. Lewis Publishers. 2001, 407 p.
7. H a l l, R. J., R. S. S k a k u n, E. J. A r s e n a u l t, B. S. C a s e. Modeling forest stand structure attributes using Landsat ETM+ data: Application to mapping of aboveground biomass and stand volume. *Forest Ecology and Management*, 225, 2006, 378–390.
8. H e a l e y, S. P., Z. Y a n g, W. B. C o h e n, D. J. P i e r c e. Application of two regression-based methods to estimate the effects of partial harvest on forest structure using Landsat data. *Remote Sensing of Environment*, 101, 2006, 115–126
9. H o u g h t o n, R. A. Aboveground Forest Biomass and the Global Carbon Balance. *Global Change Biology*, 11, 2005, 945–958, doi: 10.1111/j.1365-2486.2005.00955.x.
10. H y y p p ä, J., H. H y y p p ä, M. I n k i n e n, M. E n g d a h l, S. L i n k o, Y i-H o n g Z h u. Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes. *Forest Ecology and Management*, 128, 2000, 109-120
11. L e f s k y, M. A., W. B. C o h e n, T. A. S p i e s. An evaluation of alternate remote sensing products for forest inventory, monitoring, and mapping of Douglas-fir forests in western Oregon. *Can. J. For. Res.*, 31, 2001, 78-87
12. M u u k k o n e n, P., J. H e i s k a n e n. Estimating biomass for boreal forests using ASTER satellite data combined with standwise forest inventory data. *Remote Sensing of Environment*, 99, 2005, 434 – 447.
13. NSI. 2008. Statistical yearbook 2008. National Statistical Institute, Republic of Bulgaria, Sofia.
14. P e t k o v, P, N. P e n e v, M. M a r i n o v, S. N e d j a l k o v, Z. N a o u m o v, D. G a r e l k o v a, G. A n t o n o v. Forest types and organization of forestry in technical section Govedartsi of the Samokov Forestry. *Gorskostopanska nauka (Forest Science)*, Vol. III, № 2, 1966, 87-109 (in Bulgarian)
15. R a e v, Iv. Surface runoff in spruce stands undergoing active economic intervention. *Newsletter of the Bulgarian Geographical Society*, Vol. XVIII (XXVIII), 1980. (In Bulgarian).
16. S t o y a n o v a, N. Natural regeneration of the forest ecosystems of *Picea abies* (L.) Karsten in the Rila Mountain. In: Ivan Raev (Ed.) Environment and structure of the Norway spruce forests in the Rila Mountain, Pensoft. Sofia-Moscow, 2006, 131-145 (in Bulgarian).
17. T u c k e r, C. J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of the Environment*, 8, 1979, 127-150.

18. USGS. 2006. Shuttle Radar Topography Mission, 3 Arc Second scene SRTM_ff03_p183r031, Filled Finished-A 2.0, Global Land Cover Facility, University of Maryland, College Park, Maryland, February 2000.
19. Z h e n g, D., J. R a d e m a c h e r, J. C h e n, T. C r o w, M. B r e s e e, J. L e M o i n e, S o u n g-R y o u l R y u. Estimating aboveground biomass using Landsat 7 ETM+ data across a managed landscape in northern Wisconsin, USA. *Remote Sensing of Environment*, 93, 2004, 402– 411.

ИЗСЛЕДВАНЕ НА ВРЪЗКАТА МЕЖДУ НЯКОИ ТАКСАЦИОННИ ПОКАЗАТЕЛИ НА ИГЛОЛИСТНИ НАСАЖДЕНИЯ И СПЕКТРАЛНИТЕ ДАННИ ОТ САТЕЛИТНИЯ СЕНЗОР ASTER

П. Димитров, Е. Руменина

Резюме

В статията са представени резултатите от проведено изследване за оценка на връзката между спектралните данни от сателитния сензор Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) и пет таксационни показателя, измерени в 29 пробни площадки в иглолистни насаждения в Рила планина. Биомасата, обемът и склопеността показват умерена корелация с яркостта в близкия инфрачервен канал. Гъстотата на дървостоя е слабо корелирана със спектралните данни, а кръговата площ не показва статистически значима корелация. Връзката между наземните и сателитните данни е моделирана с помощта на регресионен анализ по метода на най-малките квадрати и на редуцираната главна ос. Съставените регресионни модели за оценка на отделните таксационни параметри имат относителни стандартни грешки на оценките в интервала от 14,4 % (за склопеността) до 51,8 % (за обема).

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 24, 2012, Sofia

DOMESTIC PARKING ESTIMATION USING REMOTELY SENSED DATA

Ramzi Ahmed

e-mail: Ramzi.Ahmed@hotmail.com

Keywords: *Parking supply and demand – Remote sensing - Road Extraction*

Abstract

Parking is an integral part of the traffic system everywhere. Provision of parking facilities to meet peak of demands parking in cities of millions is always a real challenge for traffic and transport experts. Parking demand is a function of population and car ownership which is obtained from traffic statistics. Parking supply in an area is the number of legal parking stalls available in that area. The traditional treatment of the parking studies utilizes data collected either directly from on street counting and inquiries or indirectly from local and national traffic censuses. Both methods consume time, efforts, and funds. Alternatively, it is reasonable to make use of the eventually available data based on remotely sensed data which might be flown for other purposes. The objective of this work is to develop a new approach based on utilization of integration of remotely sensed data, field measurements, censuses and traffic records of the studied area for studying domestic parking problems in residential areas especially in informal areas. Expected outcomes from the research project establish a methodology to manage the issue and to find the reasons caused the shortage in domestics and the solutions to overcome this problems.

Problem

Parking arrangements in urban areas are usually problematic since parking demand is always growing in accordance with social and economic development while parking supply is limited to the available spaces on the streets which are originally provided for relatively low traffic volumes of the early development stages. For all types of land-use this problem is markedly remarkable and especially in residential land-use where the streets

are practically planned and designed with relatively narrow carriage ways. Consequently, parking demand and parking supply has to be seriously investigated for both existing and planned residential areas specially in developing countries like Egypt. Egypt's population still grows - the annual population growth rate was 1.70 per cent by 2010. Also, the increase in car ownership levels creates parking problems and affects liveability in residential areas. All of these leads to issues and shortage of on-street parking makes problems for residents. This problem will be very complicated in the future if a policy and solution will not be established. In this study we provide insight in both parking supply and demand in residential areas and explore possible solutions.

Overview

Photogrammetry has been in studying of moving, automobiles speed in a traffic control system. An automobile on the highway is photographed from a police automobile behind. Hallert concluded that photogrammetry can doubtless be of great value for the practical application to traffic control (Hallert, 1971). Also, photogrammetry has been applied in the reconstruction of traffic accidents. Wolf and Janseen concluded that photogrammetry assumes an extremely important role in accident reconstruction since the above described accident related information quickly changes or disappears altogether (Wolf and Janseen 1980). Photogrammetry and Parking Studies, Aerial photography has been used to collect the necessary data for domestic parking (El-Nokrashym and et al ., 1992; Ramzi, 1995). The application of photogrammetry in demographic studies deals with the residents as individuals of the population. However, in parking studies residents are further viewed as road users and the surrounding spaces will be viewed as spaces also available for traffic purposes. The traditional treatment of the parking studies utilizes data collected either directly from on street counting and inquiries or indirectly from local and national traffic censuses. Both methods consume time, efforts, and funds. Since the successful launch of very high resolution sensors, especially IKONOS-II with 1 m Ground Sample Distance (GSD) and QuickBird with 0.61 GSD, many researchers have considered them as possible substitutes of the classical aerial photos used for cartographic purposes at large scales (Fraser, 2002; Kay et al., 2003; Chmiel et al., 2004; Pecci et al., 2004). Satellite remote sensing has displayed a large potential to obtain information on urban housing development state. Without this

information on urban housing development state, an effective urban planning is hardly possible. Satellite data less than 1 m spatial resolution are available now. Satellite data have been used to detect the changes of large land-use areas (Jensen et al., 1993; Macleod and Congalton, 1998; Ridd and Liu, 1998; Prakash and Gupta, 1998). QuickBird imagery are accurate enough for mapping purpose up to scale 1:2500 (Ramzi and et al ., 2009), The urban planners require up-to-date information to make and implement the city plans. Therefore, the maps should be produced and regularly updated with the changes (Ashraf, 2004; Mahmoud, 2004). 2D second order rectification models are more appropriate for cases where the perspective and elevation effects are smaller like in our case to minimize the effect of tilt and relief displacements (Ashraf et al, 2007; Mohamed, 2006; Elghazali, 2005). Wang and Yun Zhang, (2003) extracted roads from Quickbird images using classification techniques. Qualitative analysis of visual interpretation of single QuickBird imagery explained that netrwork roads and built-up areas can be easy identify and extracted (Ramzi et al., 2008).

Study area:

In this study, Qabaa city Egypt, which has a flat topography is chosen as a test area The test area is covered by Geoeye1- , 0.5m resolution, panchromatic standard, date 06/09/2011. The Area of Interest includes different types of man made features and the characteristics of narrow roads low level of economic and high prices of land, rapid growing of population and lack of up to date large scale maps.

Methodology

The general methodology is to compare parking demand with supply to identify the parking deficit. Parking demand is based on the type and amount of various land uses in the study area. Parking supply is based on available on-street and off-street parking inventory in the study area. The general approach to achieve the objectives of the project can be described in the following points:

1. Data collection

These will include the following items:

- Remotely sensed data

  VHRS images, GCPs, CPs and maps covering the study area.

- Census data or Demographic data of population.

Population can be estimated from mono rectified image using many methods. In this study population can estimated from mono rectified image Geoeye-1 by producing large scale map after finding the number of residential building, area of each building and the average number of stories. Based on the average size of each apartment, the average number of stories and the average household size population can be estimated using equations (1) and (2). After collection the following items from statistics Year book and field:

Total number of apartments/building = (Area of building average size of each apartment) * average number of stories (1)

Population = Total number of apartments * Average household size (2)

- Traffic data:

a. Parking stall

a. Parking space dimensions or Parking stall:

The area necessary for parking supply is dependent upon the space needed by the parked vehicles and the space needed for parking maneuver. The total space is dependent upon the geometry of parking stalls, i.e. parallel, angled or right angled parking. For on street parking in GCR, the 1977 Cairo University study gives average values based on field observations in CBD of Cairo. It could be concluded as a result of this study that for parallel parking, the necessary space corresponds to curb length of 4.5m per car, while for perpendicular parking the needed curb length is reduced to 2.5m.

b. Car ownership.

Car ownership can be determine from statistics year book year 2006 last census and forecasting it to year 2011.

c. Field Observation and measurements

Average size of each apartment
Average number of stories
Off street parking
legal parking supply
illegal parking or domestic parking at night.

Parking Estimates:

Parking demand can be defined as the number of vehicles expected to be looking for parking opportunity within a certain area for a specific time period. For such residential area parking demand is usually peaks at night. It is a function in car ownership and can be estimated from the following equation:

Parking demand = Population * Car ownership ... equation (3)

Parking supply in an area is the number of legal parking stalls available in that area. This means that parking supply is the summation of on street and off street parking. To estimate parking supply for Qabaa city from mono rectified satellite image by extracting the following data: measuring all road widths and all curb faces lengths, and converting this lengths to parking spaces or cars based on roads widths and the possible type of parking

Results

Five ground control points (GCPs) and nine check points (CPs) has been used for geometric correction of .Geoeye-1 image using 2D second order polynomial function. The result shows that the total RMS error on GCPs and CPs meets the specifications of large scale maps. Image enhancement techniques improve the quality of an image as perceived by a human. In our case edge enhancement has been used. In this study, visual classification of buildings, parcels and roads in the study using Geoeye-1 rectified images and ARCGIS software has been done. In summary, the results show that:

- All buildings can be identified by the pattern that they make in conjunction with the roads. Individual houses and other buildings can also be identified as dark and light tones.
- All roads with their categories secondary, local access and Alley are visible due to their shape (straight in many cases) and their generally bright tone contrasting against the other darker features. Also, from recertified images all road widths can be measured and classified. After that roads function and parking type has been classified according to their widths.

Table 1 shows road Categories, widths, function and Parking type. Based on measuring from rectified satellite image, statistical and using equation (1),

(2) and (3) both parking supply and demand has been estimated. Table 2 shows parking supply and demand

Table 1. Road Categories, widths, function and Parking type

Road categoriey	Width (m)	Function	Parking Type
Secondry	More than 8m	Traffic and servicing	Parallel or angled parking
Local access	5-8 m	Traffic and servicing	Parallel parking
Alley	Less than 5m	servicing	Parallel parking
Alley (closed)	Less than 5m	servicing	Parallel parking

Field Observation:

- Average size of each apartment=107 m2
- Average number of stories = 8 story
- legal parking supply = 3916 car
- Off street parkingt = 134 car
- Un occupied parking stalls (due to observations) =71 car
- illegal parking or domestic parking at night=258 car
- Parking types are on and off street parking and garage parking. On-street parking is limited with streets widths and lengths it contributes with small ratio in parking supply. Streets are not well designed to meet the requirements of parking supply in informal areas. Streets should also allow for moving vehicles and pedestrians.
- There are many illegal parking causes problems of moving traffic inside the study area especially at night.
- There are much other activity in residential areas such commercial, industrial and others activity leads to there is many obstruction on the platform prevent pedestrians from using the curb from parking.
- Absent of town planning this leads to narrow streets.
- Absent of garages in most buildings.

- High prices of land leads limited vacant space sand the height of buildings more than the allowable height.
- Many residents put obstructions in front of their building prevent people from parking especially in front of shops.

It has been found as presented in table 2 that: Percentage of error in calculation parking supply fro satellite image and field observations is -2%. Percentage of error in calculation shortage in parking stall from satellite image and field observations is 11.8%. Taken into consideration off street parking and un-occupied parking stalls rectified images lead to more accurate results in calculation parking demand from satellite image.

Table 2. Parking supply and demand

Region	Method	Parking supply (car)	Parking demand (cars	Shortage (car)	% error
Qabba	Rectified image	3835	4194	359	11.8
	Field	3916	4237	321	

Suggested soluations:

- Based on the rectified geoeye-1 satellite image and the produced large scale map, it has been suggest two vacant spaces, one in middle and the other in the east of the study area. This vacant spaces can be used as open parking garage to solve the shortage in parking spaces.
- Law of buildings should be applied to control the problem of domestic parking.
- Absent of town planning leads to creation informal residential regions.

Conclusion:

From the research the following conclusions can be drawn:

- Parking arrangements in urban areas are usually problematic since parking demand is always growing in accordance with social and economic development while parking supply is limited.

- Very high resolution satellite image can be used for estimation both parking demand and parking supply.

- It has been proved that Geo eye-1 satellite imagen can play an important role in calculating parking demand and supply with sufficient accuracy.

- The value of the method of collecting data from remotely sensed data lies in its ability to produce representative results for large study areas in a very short time.

- Field observations are very essential to calculate parking demand.

Recommendations

- Re-calculating the results using stereo satellite images.

- The proposed methodology with appropriate modifications should be encouraged to be applied for parking studies in urban land uses other then the residential like commercial land use, the central business districts in particular, industrial land use, cultural and recreational land uses, etc.

References

1. A s h r a f, M., B. A t y e a, I. A h m e d, S. A m i r a, Quantitative and qualitative assessment of planmetric information extraction from quick bird images, SENS 2007, Varna, Bulgaria, 2007.
2. A s h r a f, M. A., Assessment of producing large scale maps from the optical space russian imager, A Dissertation Faculty of Engineering, Ain Shams University, 2004.
3. Central agency for public mobilisation and statistics, "Statistical year book, Year 2006.
4. C h m i e l, J., S. K a y and P. S p r u y t, Orthorectification and geometric quality assessment of very high spatial resolution satellite imagery for common agricultural policy purpose. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 35(Part B4): 5 p. Istambul, 2004 (On DVD).
5. E l-g h a z z a l i, S. S., Assessment of Indian remote sensing satellite (IRS) imaging for production and updating of 1:25000 planimetric city maps, M. Sc. Thesis, Faculty of engineering, Cairo University, 2005.

6. E l-N o k r a s h y, M. A., M. Y. V i r a i m, M. A. E i l a w a, A. I. R a m z i, Evaluation of Using Photogrammetry in Parking Studies, CERM Magazine, El-azher university, 1992, vol.(14) no(11), Pp 403-415.
7. H a l l e r t, B., Photogrammetry for traffic control, close \range photogrammetry, January 26-29, 1971.
8. Interm report for the improvement of parking in CBD, Cairo University, 1977.
9. Greater Cairo region Tranportation Master plan studying the Arab republic of Egypt, Interm report for Urban study and traffic improvement, October, 1988.
10. J e n s e n, J. R., D. J. C o w e n, J. D. A l t h a u s e n, and S. N a r u m a l a n i, An Evaluation of the CoastWatch Change Detection Protocol in South Carolina, Photogrammetric Engineering & Remote Sensing, 1993, 59(6), pp. 1039-1046.
11. K a y, S.; P. S p r u y t, K. A l e x a n d r o u, Geometric quality assessment of orthorectified VHR space image data. Photogrammetric Engineering & Remote Sensing, 2003, 69 (5):484-491.
12. M a c l e o d, R. D., and R. G. C o n g a l t o n, A Quantitative Comparison of Change-Detection Algorithms for Monitoring Eelgrass from Remotely Sensed Data, Photogrammetric Engineering & Remote Sensing, 1998, 64(3), pp. 207-216.
13. M a h m o u d, S. M., Updating maps using high Resolution satellite imagery as an alternative to traditional techniques, Faculty of Engineering, Zagazig University, 2004.
14. M o h a m e d, L. K., Geometric quality assessment of various sensor modeling techniques, M. Sc. Thesis, Faculty of Engineering, Cairo University, 2006.
15. P e c c i, J.; F. C a n o and G. M a z a, Generación de una ortoimagen QuickBird del año 2003 de la Comunidad Autónoma de la Región de Murcia: metodología y resultados. XI Congreso Métodos Cuantitativos, Sistemas de Información Geográfica y Teledetección 2004. Murcia, Spain, 2004, Vol. I: 301-312.
16. P r a k a s h, A. and R. P. G u p t a, Land-use mapping and change detection in a coal mining area - a case study in the Jharia coalfield, India, International Journal of Remote Sensing, 1998, 19(3), pp.391- 410.
17. R a m z i, A. I, N. G e o r g i e v and R. N e d k o v, Mapping using ortho ready panchromatic QuickBird images, CERM Magazine, El-azher University, June 2009, vol.(31) no(2), Pp 403-415.
18. R a m z i, A. I. and R. N e d k o v, Qualitative analysis of visual interpretation of single QuickBird imagery, Fourth scientific conferences, SENS 2008, Bulgarian Academy of Sciences, 2008, pp 102- 105.
19. R a m z i, A. I., M. Sc., in Civil Engineering, Civil Engineering Department, Faculty of Engineering, El-Minia, Egypt, titled *Evaluation of Using Photogrammetry in Parking Studies*, 1995.
20. R i d d, M. K. and J. L i u, A Comarison of Four Algorithms for Change Detection in an Urban Environment, Remote Sensing of Environment, 1998, 63,pp. 95-100.
21. W a n g, Y u n Z h a n g, Extraction of urban road network using Quickbird Pansharped Multispectral and Panchromatic Imagery by Performing Edge-Aided Post-Classification, 2003.
http://www.geoict.net/Resources/Publications/RW_UNB_GRS S_2003.pdf

22. W o l f, R. and J. J a n s s e e n, , "Photogrammetry in traffic accident reconstruction", fall technical meeting "ACSM " American congress of surveying & Mapping OFOct. 7-10, 1980.

ОЦЕНКА НА ПАРКИРАНЕТО В СТРАНАТА С ПОМОЩТА НА ДИСТАНЦИОННИ ДАННИ

Рамзи А. И.

Ключови думи: Възможности и търсене в паркирането – Дистанционни изследвания – Извличане на пътища

Резюме

Паркирането е неразделна част от системата на уличното движение в цял свят. Осигуряването на паркинги за задоволяване на потребностите от паркиране в милионните градове винаги е било предизвикателство за експертите по улично движение и транспорт. Нуждата от паркиране е функция на населението и броя на притежаваните коли, който може да бъде установен от статистиката на уличното движение. Паркинговата доставка в дадена област се изразява с броя на наличните законни паркоместа в тази област. Традиционният подход в паркинговите проучвания използва данните, получени пряко, чрез преброяване на улицата и запитвания, или непряко, от местните и националните статистически данни за уличното движение. И двата метода изискват време, сили и средства. Другата разумна възможност е да се използват евентуалните налични данни, получени дистанционно от спътници, изстреляни за други цели. Целта на настоящата работа е да се развие нов подход на базата на съвместното използване на дистанционни данни, полеви измервания, преброявания и записи за уличното движение в изследваната област за изследване на проблемите, свързани с паркирането в жилищните и особено, в неформалните области на страната. Очакваните резултати от изследователския проект са да се въведе методология за управление на проблема и да се установят причините, предизвикващи недостиг в паркирането в страната и решенията за преодоляване на този проблем.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 24, 2012, Sofia

GEOMAGNETIC SURVEYS ON THE TERRITORY OF SOFIA AIRPORT

Bozhidar Srebrov, Iliya Cholakov

National Institute for Geophysics, Geodesy and Geography –
Bulgarian Academy of Sciences
e-mail:srebrov@geophys.bas.bg

Abstract

Geomagnetic surveys on the territory of Sofia Airport for measuring of geomagnetic declination and to obtain the suitability of the sit for compass calibration pad are described in the paper.

Considered is the applicability of the world geomagnetic field models, as example IGRF, when geomagnetic data with high accuracy are needed.

Introduction

There are tow primary objectives in performing airport geomagnetic surveys. The first is to determine geomagnetic declination at suitable sit and second to obtain the suitability of the sit for compass calibration pad. Geomagnetic declination is determined using procedures developed by International Association for Geomagnetism and Aeronomy (IAGA) [1]. The suitability of the sit for compass calibration pad is assessed using the standards outlined in Federal Aviation Administration (FAA) of the United States. In Bulgaria Geomagnetic Service at National Institute for Geophysics, Geodesy and Geography provide airport geomagnetic surveys for many years.

Because of the increased requirements for flight safety in Bulgaria last years teams from the National Institute for Geophysics, Geodesy and Geography performed measurements of the geomagnetic declination at many airports, including Sofia Airport [2]. Also in 2008 were carried out

surveys at the Sofia Airport to obtain the suitability of the sit for compass calibration pad.

Measurement of geomagnetic declination on Sofia Airport

Geomagnetic declination on the Sofia airport was first measured at the threshold of the runways (this is the point for take off and landing of the airplanes). Large anomalies were detected at the runway thresholds. For example, the difference between the measured declination on the eastern and on the western part of Sofia airport was $4^0 25'$. Our opinion is that the large difference is due to the concrete with substantial steel reinforcement of the runway. Knowing that the airplanes use the geomagnetic declination in the air far of the runway we decided to make the measurements away from the runway. At the Sofia Airport the measurements were made at the nearest repeat station of the Bulgarian geomagnetic repeat stations network. In this point the value in 2001 was $3^0 10'$. This value is reduced in epoch 2010.5 by using the data for geomagnetic declination changes from geomagnetic observatory Panagyurishte (Bulgaria) and the value is $3^0 53'$.

Obtain the suitability of the sit for calibration pad on Sofia Airport

FAA requirements for design, location and construction of a compass calibration pad are detailed in FAA AC (advisory circular) 150/5300-13 Appendix 4. The advisory circular may be obtained at the internet site [3].The criteria of a site location suitability in accordance with the AC are:

-The difference between geomagnetic and geography north must be uniform across the site.

- The range of declination must be less than one half degree (from 0,3 to 3 meters above the base and 75 meters off the center).

Section 5 of the FAA Advisory Circular provides guidelines for locating a suitable site for calibration pad. The general requirements for location are:

- Locate a calibration pad 90 meters from power and communication cables and other aircraft.
- Locate a calibration pad at last 180 meters from large magnetic objects, such as buildings, railroad track, high voltage transmission lines, or cables with direct current.

- Locate calibration pad off the site of a taxiway or runway to satisfy local clearances.

In the late spring of 2008 were carried out preliminary site checks on the territory of Sofia airport and as a result two sites were proposed for calibration of the compasses of airplanes types Airbus 320 and Boeing 737. The first one (Site 1) is situated at the end of taxiway K and N on the airport and the second one (Site 2) is the western ground for anti-icing system (Fig.1).

Fig. 1. Site 1 and Site 2 on the Sofia Airport map

On these two sites and then a magnetic total field survey was carried out. A geomagnetic total field survey of an existing new calibration pad is essential to determine the suitability of the site because preliminary checks may not show buried ferrous metals which can impact the magnetic field. The equipment used for survey was proton magnetometer PMP 5A. Procedure for total field survey is:

- Mark the center point with wooden stake.
- Make total field readings (mean value of 5 readings) at the center point and approximately every 5 meters along N-S and E-W lines out to 35 meters.

If the total field has a range of 75 nT over a 75 meter area, it will meet the FAA requirements. Total field survey results are used to rank potential compass calibration pad sites. Following the total field survey, a single site must be chosen for construction of the calibration pad after which a detailed geomagnetic declination survey must be made to confirm the site meet the FAA requirements.

The results from the total field geomagnetic surveys on the territory of Sofia airport are presented further down.

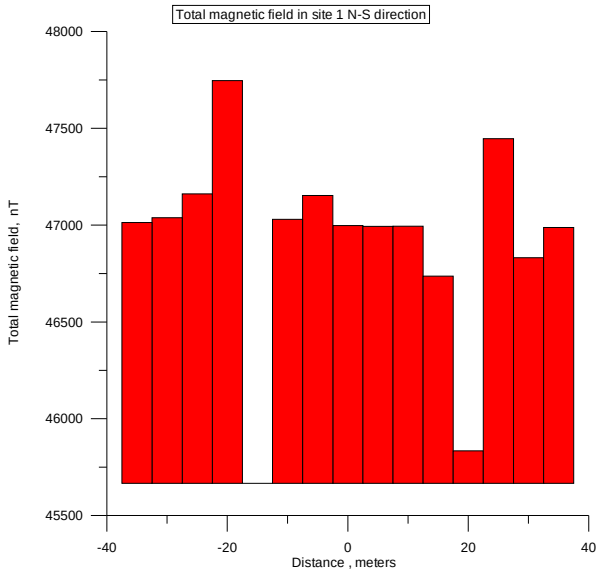


Fig. 2. Total magnetic filed (nT) in site1 N-S direction

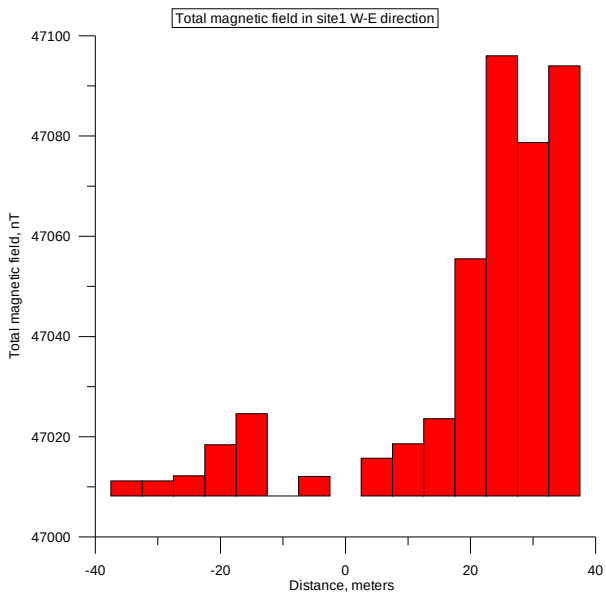


Fig. 3. Total magnetic filed (nT) in site1 W-E direction

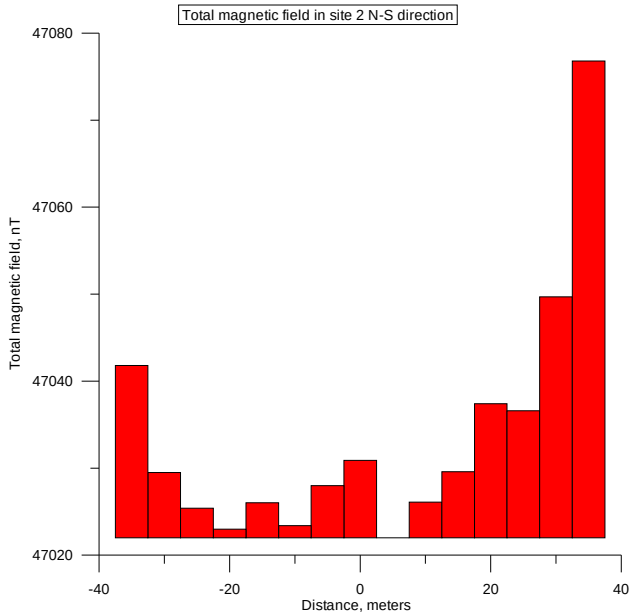


Fig. 4. Total magnetic filed (nT) in site2 N-S direction

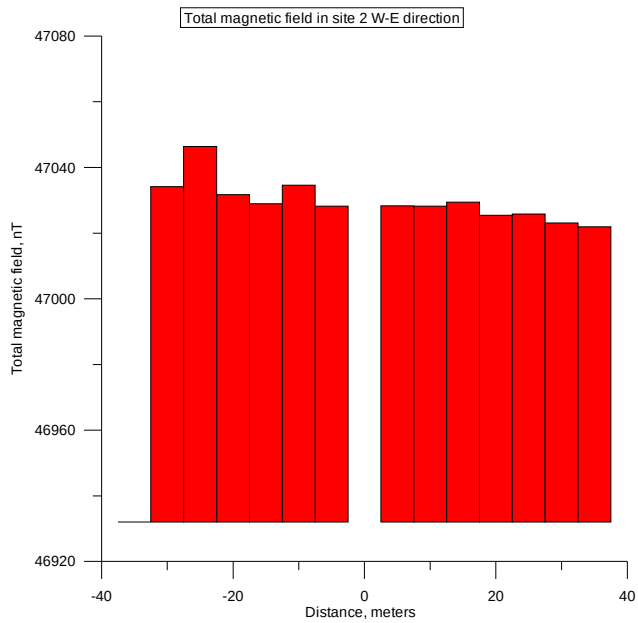


Fig. 5. Total magnetic filed (nT) in site 2 W-E direction

On the Fig. 2 and Fig. 3 are sown the values of the total geomagnetic field measured along N-S and E-W lines for Site 1. As it is clearly visible this site do not meet the FAA requirements because the range of magnetic field changes in N-S direction is more than 2000 nT and in W-E direction is approximately 90 nT.

On the Fig. 4 and Fig. 5 are sown the values of the total geomagnetic field measured along N-S and E-W lines for Site 2. The range of magnetic field changes in N-S direction is approximately 55 nT and in E-W (for 14 measurement points) is less than 30 nT and only in one measurement point is less than 120 nT. Probably this anomalous value has accumulated big measurement error caused by small local ferrous disturber and can be ignored. We can see that Site 2 meet the FAA requirements. The mean value of total magnetic field in Site 2 is 46969,4 nT and the field gradient is lese than of 1 nT/meter in the all area with dimension 75x75 meters .

As a conclusion is established that Site 2 is a suitable sit for building of new compass calibration pad on Sofia Airport.

About using of the world models for obtaining of geomagnetic field component values

A frequently asked question is: Is it possible to use geomagnetic field data, calculated from a world model, for example the International Geomagnetic Reference Field (IGRF) [4], in airport practice? As we know this is a global model of the geomagnetic field. It allows spot values of the geomagnetic field vector to be calculated anywhere on the Earth's core out into space. But for obtaining of the geomagnetic vector component values in a concrete point the model is not very accurate. We will demonstrate this further down. In epoch 2000.5 the model (IGRF-11/2010) for the spot of Geomagnetic Observatory PAG gives for geomagnetic declination a value $3^0 21{,}36'$ and for epoch 2005.5 - $3^0 42{,}24'$, but the measured thru values are $3^0 00{,}5'$ (2000.5) and $3^0 22{,}6'$ (2005.5). The absolute error is more than $0^0{,}21'$ in both cases. Similar are the results and for the total field, in epoch 2000.5 the model (IGRF-11/2010) for the spot of Geomagnetic Observatory PAG gives value 47011 nT and for epoch 2005.5 - 47166 nT. Thus we see that the model IGRF is not applicable when high accuracy is needed. It became clear that this is very important especially for obtaining of geomagnetic declination by the model anywhere on the Earth's core.

We will introduce here an alternative way for obtaining of the geomagnetic field elements values at Sofia Airport in the next ages after the epoch of measurements. For example for the element geomagnetic declination we obtained in epoch 2001.5 the measured declination value at the airport is $3^0 10'$ and in Geomagnetic Observatory PAG is $3^0 04'$. The value difference between these tow space points is $0^0 06'$.This difference is stable in relatively long time interval and thus we can obtain the value at the Sofia Airport in some next epoch by reduction of the observatory value for this same epoch by mentioned above $0^0 06'$. For example in 2007.5 observatory declination value is $3^0 32{,}4'$ and at the Sofia Airport reduced value therefore is $3^0 38{,}4'$. This method is applicable with small error for periods close to 10 years after the epoch of measurements, but for obtaining of values for longer time interval repeat measurements at the territory of airport are necessary.

References

1. N e w i t t, L. R., C. E. B a r t o n and J. B i t t e r l y, Guide for Magnetic Repeat Station Surveys, IAGA, 1996.
2. C h o l a k o v, Il. in J. Rasson and T. Delipetrov (edt.), Geomagnetics for Aeronautical Safety, Springer, Printed in the Nitherlands, 2006.
3. http// www.//faa.gov/ arp / pdf 5300 – 13p2.pdf.
4. P a r k i n s o n, W. D., Introduction to Geomagnetism, Scottish Academic Press, Edinburg and London, 1983.

ГЕОМАГНИТНИ ПРОУЧВАНИЯ НА ТЕРИТОРИЯТА НА ЛЕТИЩЕ СОФИЯ

Б. Сребров, Ил. Чолаков

Резюме

В работата са описани геомагнитни проучвания на територията на летище София за определяне на геомагнитната деклинация в различни епохи и за намиране на подходяща площадка за калибриране компасите на самолети от типове Airbus 320 и Boeing 737. Разгледана е и приложимостта на световните геомагнитни модели, като например IGRF, в случаи на необходимост от геомагнитни данни с висока точност.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 24, 2012, Sofia

ASSESSMENT OF INFORMATION EFFICIENCY AND DATA QUALITY FROM MICROSATELLITE FOR THE NEED OF ECOLOGICAL MONITORING

Roumen Nedkov

Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: rnedkov@space.bas.bg

Abstract

In this paper a methodology for defining and assessment of information efficiency, data quality and their georeferencing is presented. The criteria, which the presented methodology is based on, are related to specific tasks in the implementation of ecological monitoring.

Keywords: *microsatellite, data quality, information efficiency, ecological monitoring*

Introduction

Microsatellites are high technology devices both for space investigation and Earth observation. By using low-cost nano-, micro-, and small satellites, it is possible to solve many theoretical and applied scientific problems. These tasks have been solved so far by using the "everything in one" technology based on big space satellites.

Microsatellites can be defined as "Flying Space Intelligent Multisensors". They are data sources for space information of different efficiency and quality. The usage of microsatellites for solving numerous theoretical and applied scientific problems is directly related to the information efficiency and the quality of data received on the basis of measurements.

The problem for defining information efficiency and quality of the data obtained from the measurements is significant because it is directly correlated to the technological characteristics of the onboard sensors and their functional applications [1].

Methodology for defining and assessment of information efficiency, data quality and their georeferencing

Definition of the criteria for assessment

On the basis of the observed phenomena and processes such as fires and floods and their impact on environment the following criteria for assessment of information efficiency and space data quality are defined:

- The received data should be with a relevant space resolution ensuring the detection and recognition of the phenomena;
- Since fires are connected with temperature variations, the space data must have the relevant spectral and radiometrical resolution;
- In case of flooding, atmospheric pollution, and oil spill, the data should have a high resolution time;
- For a faster localization of these phenomena, their georeferencing should be made in a near real-time mode.

Description of the methodology

The parameters that define the information efficiency and data quality are as follows:

- Space resolution
- Radiometrical resolution
- Spectral resolution
- Temporal resolution

The space dynamics of a specific phenomenon can be approximately presented by the following functions:

(1) $\Delta(X,t) = X_i(t) - Xs$

(2) $\Delta(Y,t) = Y_i(t) - Ys$

where Xs = image resolution on the axis X

Ys = image resolution on the axis Y

$X_i(t)$ and $Y_i(t)$ = temporal variations of the dimensions of the observed phenomenon.

If the values of the functions $\Delta(X, t)$ and $\Delta(Y, t)$ are positive, the phenomenon can be detected and identified on the image.

Spectral and radiometrical resolutions are predefined and correspond to previously known spectral characteristics of the observed phenomenon.The assessment is usually made by comparing the data with other data received from other sensors [1, 3].

The assessment of the temporal resolution is made on the basis of the period of data updating compared with the duration of the time interval in which the phenomenon is observed. If the time period of receiving of new data is less longer than the time duration of the observed phenomenon, the data are defined as efficient [4].

In order to achieve a near real-time mode, the following scheme can be applied in the data georeferencing:

1. Simulation rasters with the same space resolution as the real images are generated on the basis of ground control points with fixed coordinates.

2. The coordinates of the images are georeferenced to the relevant nudes of the coordinate grid on the earth's surface on the basis of the coordinates of the pattern rasters.

3. If there is a displacement of a shot image from the pattern image due to the orbital deviation, the georeferencing is implemented on the basis of the relevant identical points with the same coordinates. In this case, it is not necessary to generate a new coordinate grid but only to add or to remove nudes from the base coordinate grid.

The image displacement ΔX and ΔY to the base image and the relevant identical points **IP1** and **IP2** are shown on Fig. 1.

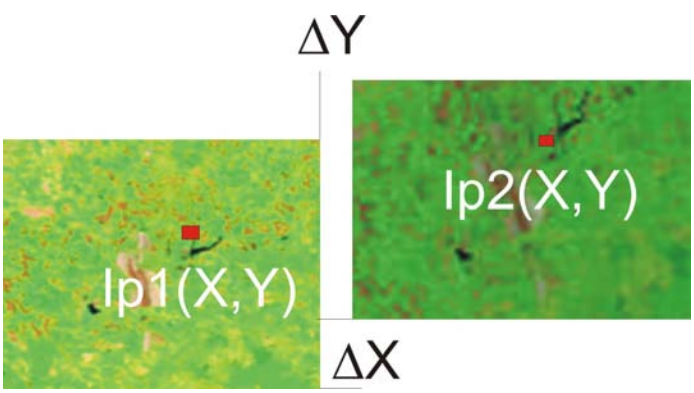


Fig. 1. Image displacement to the base image

The validation of the methodology for defining and assessment of information efficiency and data quality on the basis of the space, spectral, and radiometrical resolution is made by means of a specially-designed device for testing the quality of a multispectral sensor for microsatellite. The results are presented on Fig. 2 [2].

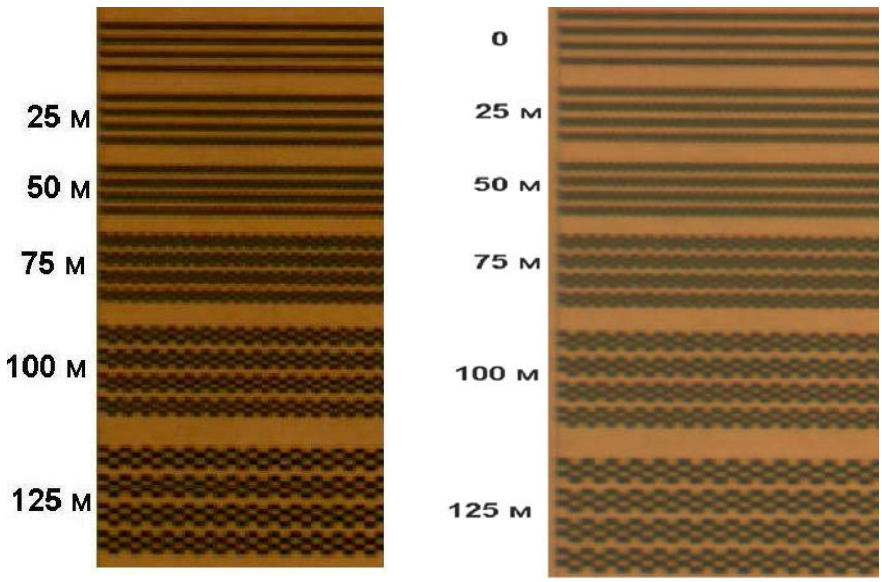


Fig. 2. The left part of the test image indicates the resolution quality of the limb of the image, whereas the right part – the resolution quality of its centre.

The methodology for achieving a near real-time mode in data georeferencing has been used for the Web-based ecological monitoring of atmospheric pollution in different parts of the territory of Bulgaria [4].

Conclusion

1. Criteria for defining and assessment of information efficiency and data quality on the basis of measurements as well as data georeferencing by using microsatellite platforms for ecological monitoring have been formulated.

2. A methodology for defining and assessment of information efficiency, data quality and data georeferencing has been proposed.

3. The results of the implementation of the proposed methodology are positive, which proves its practical applicability.

References

1. V l a d o v, M., D. D o b r o v, R. N e d k o v, M. K o l p a k o v i c h, D. U k r a i n t s e v. The scanner of remote sensing of the Earth. F i f t h S c i e n t i f i c C o n f e r e n c e w i t h I n t e r n a t i o n a l P a r t i c i p a t i o n S E N S 2009, 2 - 4 November 2009, Sofia.
2. V l a d o v, M., D. D o b r o v, R. N e d k o v, M. K o l p a k o v i c h, D. U k r a i n t s e v. The device for testing of a scanner of remote sensing of the Earth. F i f t h S c i e n t i f i c C o n f e r e n c e w i t h I n t e r n a t i o n a l P a r t i c i p a t i o n S E N S 2009, 2 - 4 November 2009, Sofia.
3. Z h u k o v, B., E. L o r e n z, D. O e r t e l. Experience of detection and quantitative characterization of fires during the experimental small satellite mission BIRD. Forschungsbericht, 2005-04. DLR. Köln.
4. Н е д к о в, Р., Е. Р у м е н и н а, Л. Ф и л и п о в, П. Х р и с т о в, М. Д и м и т р о в а, М. З а х а р и н о в а, В. Н а й д е н о в а, Г. Ж е л е в. Web-базиран мониторинг на атмосферните замърсявания в района на община Стара Загора на базата на спътникови данни. SENS 2007/2008, 264-273.

ОЦЕНКА НА ИНФОРМАЦИОННАТА ЕФЕКТИВНОСТ И КАЧЕСТВОТО НА ДАННИТЕ ОТ МИКРОСПЪТНИЦИ ЗА НУЖДИТЕ НА ЕКОЛОГИЧНИЯ МОНИТОРИНГ

Р. Недков

Резюме

В настоящата работа е предложена методика за определяне и оценка на информационната ефективност, качеството на данните от измерванията и тяхното георефериране. Критериите, на които се базира предложената методика, са свързани с определени задачи при извършване на екологичен мониторинг.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 24, 2012, Sofia

SYNTHESIS OF OPTIMAL FILTER OF AUTOMATED CONTROL SYSTEM OF SELF-AIMING UNMANNED AIR VEHICLE WITH FIXED COORDINATOR

Valentina Tsekova

Rakovski Defence and Staff College, Defence Advanced Research Institute
e-mail: valsof20@hotmail.com

Abstract

A structured schema is synthesized and a preliminary function of the the optimal filter of the control system of the self-aiming unmanned air vehicle (UAV) with fixed coordinator using some statisitical methods was obtained .

The conducted investigations showed that if the UAV itself is included in the control system the analysis of the aimed circle will be significantly relieved, as the control system of the UAV with fixed coordinator can be regarded as a following system.

Introduction

In a number of leading in martial relation countries a deep transformation of the armed forces is done in order to adapt them to the threats and the challenges in the new information century. It is in close contact to the new military strategy, which infers that the contemporary war should be conducted in short term and with minimal losses via delivering highly precise strikes from air-space or water-space without immediate contact with the opponent. This approach is known as “distant (non-contact) war”. It becomes possible due to the successful development of new armaments specimen and military techniques which have: improved concealment and safety, global reconnaissance and surveillance systems, navigation and target indication, highly precise striking means and control systems, built on network principal integrated in common reconnaissance-striking (information-striking) systems.

The distant war determines principally new requirements towards the providing to the armies and forces with contemporary armaments and techniques, including highly effective control, reconnaissance and communication systems.

In this regard great attention has been brought to the development of the aero-space component of the armed forces as well as the concept of its use. As plans up to 2020 for development of the military forces of the NATO member countries show, one of the main streams of the MAF (military air force) development is elaboration and acceptance of armaments of unmanned aviation complexes and systems with diverse purpose (including reconnaissance-striking unmanned complexes and their control systems) [2, 4].

In the late sixties in Republic of Bulgaria begins the development of unmanned air vehicle (UAV) of type radio-manipulated targets. On a later stage the development of reconnaissance UAV steps in and it continues up to these days. In certain circumstances the reconnaissance UAV can be transformed into reconnaissance – striking ones. Therefore the investigation and the synthesis of their control systems and their elements is a key phase of their development.

As pointed in [7, 8] for hitting enemy targets with reconnaissance – striking UAV it is essential that the information received by the television or infrared camera installed on the UAV's board to be real-time, in order to adjust its position according to the longitudinal axis of the UAV from 0 to 90 degrees. This information is monitored on a screen by an operator in a land control station. Target-lightening systems are used for improving visibility of targets.

When the operator discovers enemy target in the received TV image, they aim the unmanned air vehicle toward it. In this way the control system (CS) of the UAV starts receiving and processing the signal reflected by the target. In case of high enough amplitude of the accepted information signal the CS starts navigating the UAV autonomously until hitting the target.

At the combat UAV, that reaches the target using controlled and uncontrolled rockets or through straight shot in it (kamikaze type) standing coordinator could be used, which longitudinal axes coincides with the longitudinal axes of the unmanned air vehicle. That allows straight-direction methods where the UAV long axes are directed to the target during the whole flight to be used.

In this case UAV control system could be analyzed as system consisting of coordinator, which measures the target angle position according to the UAV long axes and unmanned vehicle, which eliminates that angle of non-coherence, trying to keep it zero [5].

Optimum Control System

CS should be optimal to be able to define the co-ordinates of the target with minimal possible mistake [7]. During the design and synthesis of the CS usually is needed that UAV characteristics, enemies' targets, informational signals and possible interferences are defined. The structure and scales of CS are defined according to those characteristics. The size of the linear diversion of UAV from the target center is defined as well at the moment when the unmanned vehicle hits the area, which is perpendicular of its flight path and consists as a spot the target center.

In practice, however, there is never a whole set of needed information but only some statistic characteristics. That's why the synthesized CS should be able to change its parameters depending on the conditions of using the UAV, so that its accuracy in general to be equal of the maximum possible for the particular situation.

Due to that above during the design of the optimum CS of self-aiming UAV with fixed coordinator, statistic methods are most reasonable to be used – such as the method of the maximum aposterior possibility and the method of maximum function of probability [1].

In the case above the measured target angle depends on many accidental factors such as: direction, maneuvers, target and UAV autowaverings, UAV target accuracy in the moment of attack, turbulence, etc. All those factors are accidental and their influence over the stochastic process leads to its usual distribution. That's why can be accepted that the measured angle of the attacked target $\lambda(t)$ is commonly distributed.

When the method of maximum aposterior possibility is used the measured parameter value (the angle co-ordinate of the attacked target) is worth due to chosen values of the entering realizations of the received information signal. Therefore for optimal valuation of the unknown parameter (the measured angle) λ is taken the value λ_{exit}, where the aposterior possibility $P_{ps}(\lambda)$ has its maximum.

When the method of maximum aposterior possibility is used a block-diagram of the optimal coordinator is synthesized (coordinator measurement) such as the one shown at fig. 1 [7].

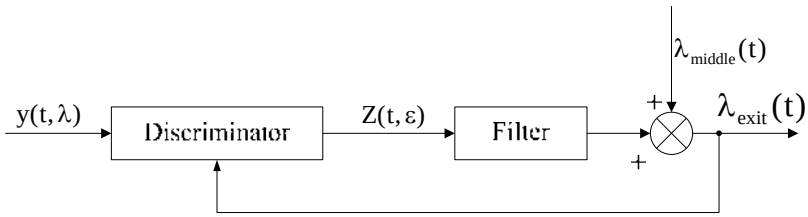


Fig. 1. Block-diagram of optimal coordinator measurer

The received informational signal is processed in the discriminator. Due to that at its exit voltage $Z(t,\varepsilon)$ appears which is proportionally equal to the non-coherence. It comes at the entrance of the filer, which impulse-transitional function is equal to $C^{-1}(t,\tau)$. At the exit filter signal should add middle quantity of the measured angle $\lambda_{middle}(t)$, and in result of it its optimal value is created $\lambda_{exit}(t)$. To have the diagram of the measurement followed, this optimal value needs to be clarified, i.e. $\lambda_{exit}(t)$ to be proceeded by the discriminator.

The aim of the coordinator is to keep the angle of non-coherence about to zero, i.e. to keep tight the axis of the diagram of the orientation of the optical system, which receives the informational signal together with target direction.

During the synthesis of the optimal coordinator according to the set characteristics of the informational signal y(t), the measured angle λ(t) and sound n(t) need to be found optimal structural diagrams of the: discriminator, measurement of the optimal steepness of the pelengation characteristics and filter.

Optimal discriminator structure is defined in [5] as the method of the function probability maximum is used, where the probability equation is worked out. [1]:

$$(1) \qquad \frac{\partial \ln P[y(t),\varepsilon]}{\partial \lambda(t)} = 0,$$

where $P[y(t),\varepsilon]$ – probability function; $\varepsilon = \lambda - \lambda_{exit}$ – target angle diversion of the flat signal discriminatory zone (angle of non-coherence); $\lambda_{exit}(t)$ – angle measured value (value of the measured angle).

The probability function is defined by the functional of the density distribution of the probability of the registered signal P[y(t)], which is regarded as function of the parameter of the non-coherence ε.

The structural schema of the measurer of the optimal steep of the pelengation characteristic is defined as the second derivative of the probability function [1, 6]:

$$(2) \qquad A_{ij} = -\frac{\partial^2 \ln P[y(t), \varepsilon]}{\partial \varepsilon_i \partial \varepsilon_j},$$

where $\mathbf{A_{i,j}}$ – is matrix, characterizing the optimal steep of the pelengation characteristic; i,j = 1,2, ... , n; $\mathbf{y}(t) = y_1(t), y_2(t), ...y_n(t)$ – row vector of the received realization; $\boldsymbol{\varepsilon} = \varepsilon_1, \varepsilon_2, ... ,\varepsilon_п$ – column vector of the measurement non-coherency.

Synthesis of the Optimal Filter

A few methods can be used for defining the structure of the optimal filter, on whose output should be received the optimal value of the measured angle: the method of the minimum square quadratic deviation; the method of non-linear filtration; the method of the maximum of the aposterior probability, etc. In the latter, an integral equation is drawn, which defines the optimum impulse transition function of the filter:

$$(3) \qquad C^{-1}(t,\tau) + \int_{t_0}^{t} C^{-1}(s,\tau) A(s) R(s,\tau) ds = R(t,\tau) \ ,$$

where $A(s)$ is the optimal steep of the pelengation characteristic of the discriminator; $R(t,\tau)$ - correlation function of the information signal**.**

In its general form the equation is hard to solve without specifying the correlation function of the information function and the optimal steep of the pelengation characteristic of the discriminator. Therefore it is necessary to describe them with statements corresponding to their physical nature.

From experience we know that the measured angle λ(t) is static accidental process, whose correlation function depends on the difference of its arguments, i.e. has the form R(t-τ). In this case the correlation coefficient r(t-τ) is experimentally determined and from the received graphic is chosen the empiric formula which can describe it.

For most of the coordinate's measurers the empiric formula of the correlation coefficient can be presented as [3]:

$$r(t-\tau) = e^{-36(t-\tau)} \cos 43(t-\tau) . \quad (4)$$

If taken into consideration that in small values of the time interval $(t-\tau)$ the function $\cos 43(t-\tau) \approx 1$, then the correlation coefficient will be equal to:

$$r(t-\tau) = e^{-36(t-\tau)} . \quad (5)$$

Then the correlation function which corresponds to such correlation coefficient will be equal to:

$$R(t-\tau) = Be^{-\beta(t-\tau)}, \quad (6)$$

where $\beta = 36$ and the value of B is defined by the type of the input information signal.

The function A(t) is a variable stochastic function, characterizing the steep of the pelengation characteristic of the discriminator. It has positive mathematical expectancy (ratio) and changes significantly slower than the correlation function $R(t-\tau)$.

Therefore the integral equation, defining the impulse transition function of the filter can be described as:

$$C^{-1}(t,\tau) + \int_{t_0}^{t} C^{-1}(t,\tau)A(s)R(t-s)ds = R(t-\tau) , \quad (7)$$

i.e. it is transformed into integral equation with different-sided core. If we introduce the designation $C^{-1}(t,\tau) = g(t,\tau)$ and take into account the slow alteration of A(t), then formula (6) will look like:

$$g(t,\tau) + A(t)\int_{t_0}^{t} g(t,s)R(t-s)ds = R(t-\tau) . \quad (8)$$

The simplest way of solving such integral equations is the method of consecutive approximation, i.e.

(9) $$g(t,\tau) = g_0(t,\tau) - g_1(t,\tau)A(t) + g_2(t,\tau)A^2(t) - \ldots\ldots,$$

where:

(10) $$g_0(t,\tau) = R(t-\tau),$$

(11) $$g_1(t,\tau) = \int_{t_0}^{t} g_0(t,\tau)R(t-\tau)d\tau,$$

(12) $$g_n(t,\tau) = \int_{t_0}^{t} g_{n-1}(t,\tau)R(t-\tau)d\tau.$$

Then the first article of the row will be:

(13) $$g_0(t,\tau) = Be^{-\beta(t-\tau)},$$

the second article:

(14) $$g_1(t,\tau) = \int_{t_0}^{t} Be^{-\beta(t-\tau)}R(t,\tau)d\tau = \frac{B^2}{2\beta}\left[1-e^{-2\beta(t-t_0)}\right],$$

the third article:

(15) $$g_2(t,\tau) = \frac{B^3}{2\beta^2}\left[1-e^{-2\beta(t-t_0)}\right]\left[1-e^{-\beta(t-t_0)}\right]$$

and the fourth article:

(16) $$g_3(t,\tau) = \frac{B^4}{2\beta^3}\left[1-e^{-2\beta(t-t_0)}\right]\left[1-e^{-\beta(t-t_0)}\right]^2.$$

The common article of the row can be written in the following appearance, taking into account formulas from (13) to (16):

(17) $$g_n(t,\tau) = \frac{B^{n+1}}{2\beta^n}\left[1-e^{-2\beta(t-t_0)}\right]\left[1-e^{-\beta(t-t_0)}\right]^{n-1}.$$

Thus the row which is the solution to the integral equation will have the look:

(18)

$$g(t,\tau) = Be^{-\beta(t-\tau)} - A(t)\frac{B^2}{2\beta}\left[1-e^{-2\beta(t-t_0)}\right] + A^2(t)\frac{B^3}{2\beta^2}\left[1-e^{-2\beta(t-t_0)}\right]\left[1-e^{-\beta(t-t_0)}\right] - ...$$

$$.....(-1)^{n-1}A^n(t)\frac{B^{n+1}}{2\beta^n}\left[1-e^{-2\beta(t-t_0)}\right]\left[1-e^{-\beta(t-t_0)}\right]^{n-1}.$$

Equation (18) shows that the impulse transition function of the optimal filter, evening the measured value of the angle of optimal coordinator, is presented in the form of an endless row. From practical consideration this entry can be significantly simplified by taking into account the following facts:

- every following article in the row is β times less then the preceding one;
- $1- e^{-\beta(t-\tau)} < 1$;
- with increasing the number of articles in the row (its promotion to a degree) it becomes much smaller than one.

Therefore, with the needed in practice precision the row may be restricted to just the first two – three articles. Then the impulse transiting function of the filter can be noted as:

(19)

$$g(t,\tau) = Be^{-\beta(t-\tau)} - A(t)\frac{B^2}{2\beta}\left[1-e^{-2\beta(t-t_0)}\right] \quad x$$

$$x\left\{1-\frac{A(t)B}{\beta}\left[1-e^{-\beta(t-t_0)}\right]\right\}.$$

Equation (19) can be simplified by accepting $t_0 = 0$. This acceptance is true for all short-term memory systems (systems, which return to their initial state shortly after their reaction to external interference). It can be extended to the remaining systems, if taking into account that t_0 represents a fixed moment of time, close to the moment of aiming of the UAV to the target, and t is the current time moment, which grows fast. In this case $t - t_0 \rightarrow t$. When the UAV approaches the target (it is needed to have the highest precision for measuring its coordinates at that moment), we can consider that $t - t_0 = t$. Thus the impulse transition function takes its final form:

$$g(t,\tau) = Be^{-\beta(t-\tau)} - A(t)\frac{B^2}{2\beta}(1-e^{-2\beta t}) \quad x$$

(20)

$$x\left[1-\frac{A(t)B}{\beta}(1-e^{-\beta t})\right].$$

If we sign:

$$(21) \qquad \frac{B^2}{2\beta}(1-e^{-2\beta t})\left[1-\frac{A(t)B}{\beta}(1-e^{-\beta t})\right] = Д(t),$$

then instead (20) for the impulse transition function we have:

$$(22) \qquad g(t,\tau) = Be^{-\beta(t-\tau)} - A(t)Д(t) \quad .$$

Using the deducted above formula (22) for the impulse transition function of the filter, can be synthesized the structural schema of the optimal filter of coordinator, whose transmission function is equal to:

$$(23) \qquad G(p) = \int_0^s g(t,\tau)e^{-p\tau}d\tau .$$

After replacing of (22) in (23) we have:

$$(24) \qquad G(p) = \left[\frac{A(t)Д(t)}{p} - \frac{B}{p-\beta}e^{-\beta(t-s)}\right]e^{-ps} .$$

If we sign $T = -1/\beta$, then instead of (24) we have the following transmission function of the optimal filter:

$$(25) \qquad G(p) = \left[\frac{A(t)Д(t)}{p} - \frac{TB}{Tp+1}e^{(t-s)/T}\right]e^{-ps} .$$

Using the deducted transmission function (25) the following structural schema of the optimal filter, shown in fig.2 can be composed:

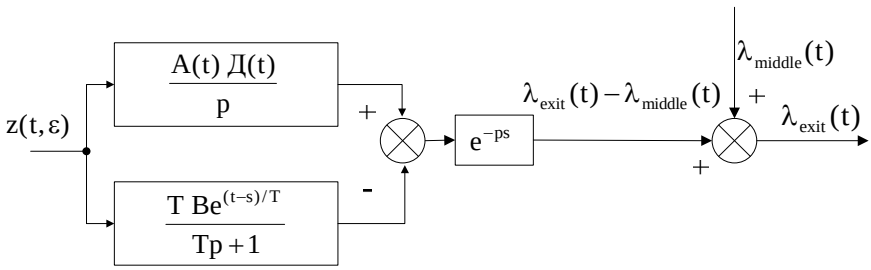


Fig. 2. Optimal filter of coordinator

From it is visible that in order to achieve optimal estimation (value) of the measured angle, the output signal of the optimal discriminator must be let through optimal filter, which consists of the following units: integrating unit with variable magnifying coefficient, proportional to the optimal steep of the pelengation characteristic of the discriminator and dependant on the parameters of the correlation function of the input signal; integrating unit with variable magnifying coefficient, determined by the parameters of the correlation function and unit with constant delay of time s, equal to the time for which the impulse transition function $g(t,\tau) > 0$.

Conclusion

At first glance the above described structural schema of the optimal filter looks hard to put in practice as the ratio of the measured angle - $\lambda_{middle}(t)$ should be added to the output signal of the filter (executive organs of the UAV and the UAV itself). Actually, it only relieves the technical realization as the axis of the UAV with fixed coordinator during the flight is aimed at the target and its orientation in space is the same as the direction of the ratio of the measured angle. Thus the UAV works off only the appearing deviations from this angle.

The deducted structural schema of the optimal filter shows that if the UAV itself is included in the control system, then the analysis of the aiming circle of the UAV can be relieved significantly, as the control system of UAV with fixed coordinator can be regarded as following system.

References

1. Б а к у т, П. и др., Вопросы статистической теории радиолокации, том 2, М.: Соврадио, 1964.
2. З л а т е в, З. Роля, място и задачи на безпилотните авиационни комплекси в системата за национална сигурност, 12.05.2010, http://pan.bg/view_article-6-1602-Rolq-mqsto-i-zadachi-na-bezpilotnite-aviacionni-kompleksi-v-sistemata-za-nacionalna-sigurnost.html
3. П у г а ч е в, В. Теория случайных функций и ее применение к задачам автоматического управления, М.: Физматгиз, 1962.
4. Р о с т о п ч и н, В. Современная классификация беспилотн,х авиационных систем военного назначения, www.uav.ru/stati.php, 18.06.2010.
5. Ц е к о в а, В., Е. Ц е к о в. Оптимален дискриминатор на ъглови координати, в сборник доклади на XII научна конференция с международно участие "Транспорт 2002" на ВТУ "Т. Каблешков"- София, стр. 507 – 511, 2002.
6. Ц е к о в а, В., Е. Ц е к о в. Оптимизиране на стръмността на пеленгационната характеристика на оптимален дискриминатор, в сборник доклади на Юбилейна научна сесия на Факултет „Авиационен" на НВУ „Васил Левски" – Д. Митрополия, том 1, стр.66 – 72, 2003.
7. Ц е к о в а, В., Оптимален измервател на координати за система за управление на безпилотен летателен апарат от ударен тип. Fourth Scientific Conference with International Participation - SENS 2008, Space Research Institute – BAS, Book of papers, pp.201 – 205, 2008.
8. Ц е к о в а, В., Синтезиране на оптимална система за управление на самонасочващ се безпилотен летателен апарат с неподвижен координатор, Годишник на Военна академия, 2009.

СИНТЕЗИРАНЕ НА ОПТИМАЛЕН ФИЛТЪР ОТ СИСТЕМАТА ЗА УПРАВЛЕНИЕ НА САМОНАСОЧВАЩ СЕ БЕЗПИЛОТЕН ЛЕТАТЕЛЕН АПАРАТ С НЕПОДВИЖЕН КООРДИНАТОР

В. Цекова

Резюме

Синтезирана е структурната схема и е получена предавателната функция на оптимален филтър от системата за управление на самонасочващ се безпилотен летателен апарат с неподвижен координатор като са използвани някои статистически методи.

Проведеното изследване показва, че ако към системата за управление се включи и самият БЛА, то анализът на кръга на насочване на БЛА към целта може да се облекчи съществено, тъй като системата за управление на БЛА с неподвижен координатор може да бъде разглеждана като следяща система.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 24, 2012, Sofia

PULSE–ARC PLASMA WELDING AND SURFACING PROCESSES

Rayna Dimitrova, Bojana Tabakova

Technical University of Sofia

Abstract

The paper presents literature survey on pulse-arc plasma welding and surfacing processes. Some issues related with these processes' equipment and materials, variations, application, technologies, and the quality of the welded workpiece are considered.

Introduction

Plasma processes use thermal plasma energy to melt material. The attempts to achieve higher energy density by constricting the electrical arc date a long way back. One of the earliest plasma arc systems was the gas vortex stabilization device introduced by Schonherr in 1909. In this device, gas was fed tangentially into an arc-discharge tube.

In 1922, Gerdien and Lotz designed an arc-stabilizing device by injecting tangentially water to the tube centre. Water whirled along the internal surface and was ejected in the ends. When the arc burned between carbon electrodes and passed through this tube, the water concentrated the arc along the axis, causing high current density. The term *thermal plasma* was first introduced in 1927 by the American physicist Langmuir to denote the fourth aggregate state of matter, i.e. the state of a gas with high dissociation and ionization rate, which apart from neutral atoms and molecules, contains positive and negative charges – ions and electrons. In 1961, the first plasma surface processing equipment was presented, and in 1963, the first plasma welding was introduced.

Peculiarities and characteristics of plasma welding and surfacing

Plasma welding is, in substance, a development of the TIG welding process, however using a different mechanism to transfer thermal energy to the workpiece. Both TIG and plasma welding use a non-melting (tungsten) electrode located in a nozzle, through which the plasma-forming gas is passed. The arc heats this plasma-forming gas, which gets ionized, attains electric conductivity, and is forced out through the nozzle. This ionized gas is defined as plasma. The plasma jet leaves the nozzle at temperature of about 16,700°C (30,000°F) in the form of constricted concentrated jet with precisely controlled direction, which creates a highly favourable seam form coefficient (pool depth-to-width ratio).

Compared to the TIG-arc, plasma features:

- exceptionally high thermal power, energy density, and temperature of the plasma arc;
- cylindrical form of the arc;
- high kinetic energy of the exiting plasma jet;

At low plasma kinetic energy, the melted metal is not blown away, which provides favourable conditions for the welding or surfacing technological processes.

Plasma surfacing is used to lay various metals or alloys on the details' surfaces to improve their operational properties. The laid on metals or alloys feature great hardness, wear-resistance, corrosion resistance and thermal resistance. The depth of the surfaced layer in a passage may reach 4-5 mm. Surfacing in several layers is possible. The process provides to obtain high-quality details with insignificant spending of expensive alloy materials.

Plasma surfacing is applied to lay copper, bronze, or other special alloys on the working surfaces of steel vapour-conducting fixture elements, chromium-nickel alloys covering internal combustion engine valves and more. The process is used successfully during the repair of stamps, press-forms, rolling rolls, and other metal-processing equipment elements.

Plasma welding and surfacing are arc processes, in which the common metal pool is obtained as a result of a forcedly constricted arc between a non-melting electrode and the workpiece (direct arc) or between a non-melting electrode and a concentrating nozzle (indirect arc). No pressure is applied on the welding pool. The process may be implemented with or without additional metal. The arc is concentrated in an ionized plasma column exiting the nozzle's end (Fig. 1).

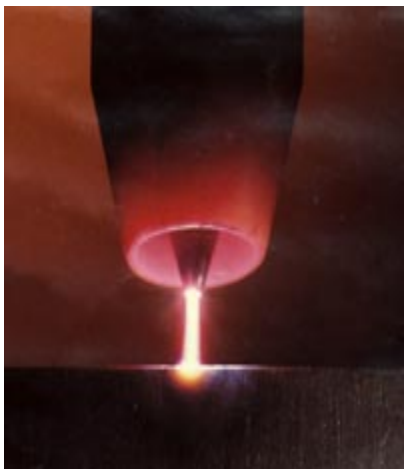


Fig. 1

The plasma-forming gas also provides protection to the molten metal and usually constitutes a part of the basic protection gas, which may be inert gas or inert gas mixture. The gas mainly used as plasma-forming gas is argon. Argon gets easily ionized and therefore provides for the arc's easy ignition. By adding molecular gases (H_2 and N_2), the thermal content of the plasma-forming gas may be increased significantly compared to pure argon (under the same temperature). When the hot gas collides with the relatively colder surface of the workpiece, as a result of atom and ion recombination, the heat from dissociation and ionization is given out accordingly. Thus, the arc's stability is improved and the invested heat amount and the arc's penetration depth into the processed workpiece is increased.

Plasma arc is used to weld non-rusting steels, nickel alloys, titanium alloys, molybdenum, tungsten and more. Compared to TIG welding, the process features a more stable arc and more uniform weld penetration depth. According to its penetration ability, the method occupies intermediate place between electric-arc and electron-beam welding. The arc's column has cylindrical form; therefore, the width of the heated surface depends poorly on the arc's length. The plasma arc provides to obtain heat spot with constant diameter, which results in stabilization of the weld penetration depth. This is particularly important in thin tin welding. The change in the heat spot form is accomplished by using nozzles with different structure (Fig. 2).

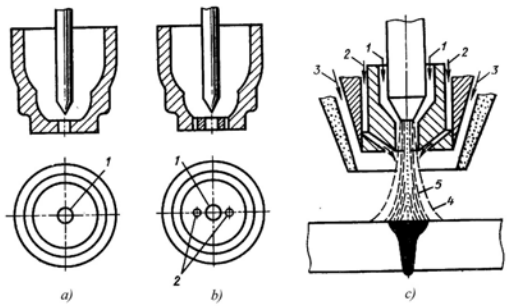


Fig. 2

For instance, when it is necessary to obtain an elongated heat spot in the nozzle (Fig. 2b), two additional openings are made through which cold plasma-forming gas exits, thus reducing the spot's cross size. The use of such a nozzle reduces HAZ's width and increases welding rate by up to 50÷100%. Another variety in the nozzles with additional openings is the formation of a focusing gas flow (Fig. 2c) which constitutes gas mixture consisting of. Ar and He or Ar and H_2, gets ionized with greater difficulty and thus constricts plasma.

Plasma-forming gas is fed tangentially, which provides for proper arc stabilization using a small capacity. Focusing gas is fed through a concentric ring-like channel located between the channels of the plasma-forming gas and the protective gas. Focusing flow is fed at an angle with respect to the arc's longitudinal axis which results in additional constriction of the arc as a result of its cooling. Another possibility for deformation of the heat spot is to use non-homogeneous magnetic field.

Types of plasma arcs and plasmatrons

Depending on the manner of their generation, three main types of plasma arcs and accordingly, three plasma processes may be identified, which are implemented by the respective plasma generation devices, i.e. plasmatrons:

- plasma–arc process with open arc (direct, shifted arc);
- plasma–jet process with closed arc (indirect arc);
- plasma–jet–arc process with open and closed arc – combined process;

Plasma–arc process with open arc (direct, shifted arc)

During open-arc welding 8 (Fig. 3.) the tungsten electrode is cathode 4 and the workpiece is anode 6. The arc passes through water-cooled nozzle 7,

whereat it gets constricted. Argon is used as plasma-forming gas which, depending on the processed metal, may also contain hydrogen (5÷10% for high-alloy Cr-Ni steels) or helium (for Ti and Zr).

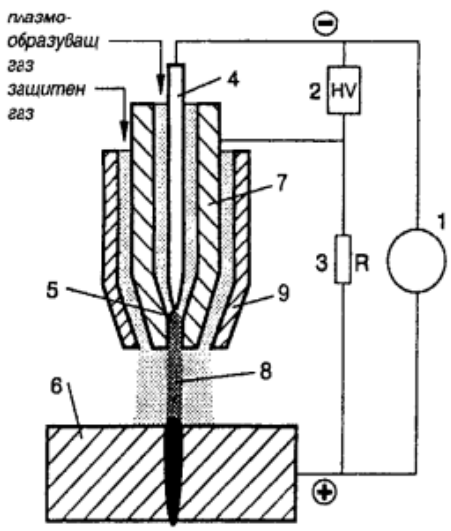


Fig. 3

Plasma-forming gas is fed concentrically around the tungsten electrode, thus also protecting it from oxidation. To provide for the arc's lighting, voltage is applied on the internal nozzle, too. Through high-voltage pulses HV from device 2, the auxiliary (pilot) arc 5 between tungsten electrode 4 and copper nozzle 7 is ignited. To prevent the possible melting of the copper nozzle, resistance R (3) is been included in the current loop of the pilot arc. The pilot arc ionizes the plasma-forming gas, after which main arc 8 jumps onto the workpiece. After the arc has transferred to the base metal, the auxiliary arc is switched-off. Protective gas both protects the welding area and cools the arc column. In some cases, additional concentric jet of cooling gas is let out. The cooling of the arc's column results in its constriction and increase of the thermal source's concentration. As a result, current density around the arc's longitudinal axis increases and the temperatures along the axis of the arc's column reach 3,000K. This provides to weld without flanging widths much greater than those welded after the TIG method. Moreover, linear energy is much smaller and welding rate is up to several times greater. All this has quite a favourable impact on deformations during welding, which

are usually negligibly small. Cooling rates in HAZ are greater and this should be accounted for in welding materials tending to form tempered structures.
To protect the melted area from ambient temperature, additional protective gas is used – 99.95% Ar. It is fed concentrically between the copper nozzle and the gas nozzle. To prevent plasma jet expansion after its exiting the nozzle, additional focusing gas may be used.
During welding of mild steels or low-alloy steels, carbon dioxide may be used as protective gas. On account of the relatively narrow heating area, the deviation of the burner (plasmatron) (Fig. 4) from the seam's line should not exceed 10% of the welded thickness.

Fig. 4

The composition of the protective gas affects the penetration ability of the arc. The addition of hydrogen to argon increases the weld penetration depth. The optimal concentration is 7%. When using helium, the weld penetration depth is smaller compared to argon-hydrogen mixture. In contrast to the TIG process, during welding of stainless steels, the addition of 7.5% hydrogen to the protective gas does not cause formation of pores
During plasma welding, butt seams with sheet material thickness of up to 9.5mm may be implemented without flanging the ends and without using additional metal. With thickness of up to 25mm, V- or U–shaped flanging is required, whereas the flanging angle is smaller compared to the TIG process. The additional metal quantity is reduced up to three times. The process has greatest advantages in welding without flanging. During plasma welding, the additional metal is fed in the back part of the welding pool. The process may be used in all spatial applications and different mechanization levels. Among its various applications, it displays one of its greatest advantages in pipe welding.

Plasma–jet process with closed arc (indirect arc)

The closed arc 5 (Fig. 5) burns inside the plasmatron between tungsten electrode 4 and water-cooled copper nozzle 7. Therefore, only a plasma jet exits the plasmatron. The plasma jet is characterized by a bright-glowing nucleus having a base a little less the size of the nozzle's opening. The nucleus is surrounded by a torch with weaker light emission. The jet length is determined by the arc's power, the nozzle's diameter, and the gas capacity. In case of lamellar exit of the gas, the jet is long and actually, it does not mix with the surrounding environment. In case of turbulent exit of the gas, a short plasma jet is observed. By varying the nozzle's form, the contour of the plasma jet may be preset, thus affecting the distribution of the jet's thermal and power impact on the processed material. The temperature field in the jet is characterized by great radial and axial gradients, reaching temperatures of $32{,}000^0$C near the cathode.

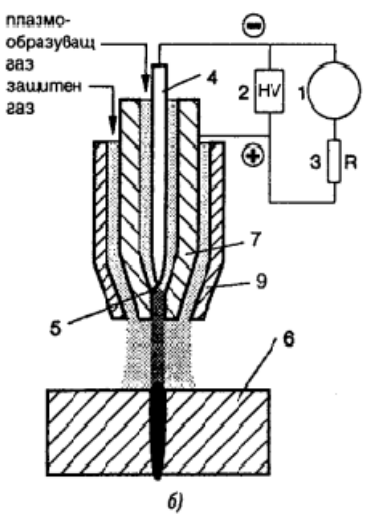


Fig. 5

In engineering practice, plasma jet is characterized by the average mass temperature in the cross section coinciding with the nozzle's butt plane. It may be determined by the relative enthalpy of plasma-forming gas $H = q / G$ (q is the effective arc power in this plane, J/s; G is the mass

capacity of plasma-forming gas, g/s. Table 1 contains the main parameters of the plasma jet for some gases used in practice.

Table 1

Plasma-forming gas	Arc power, KW	Gas capacity, g/s	Internal coefficient of performance (COP) of the plasmatron	Relative gas enthalpy, kcal/m^3	Average mass temperature of the plasma, K
Nitrogen	25	0.5	60	9000	7350
Hydrogen		0.1	80	4350	4075
Air		0.5	50	7760	6925
Argon			40	8450	14100

The main parameters regulating plasma jet thermal characteristics are: current magnitude, plasma-forming gas capacity, and arc length. Their impact on average mass temperature and jet power is shown in Fig. 6

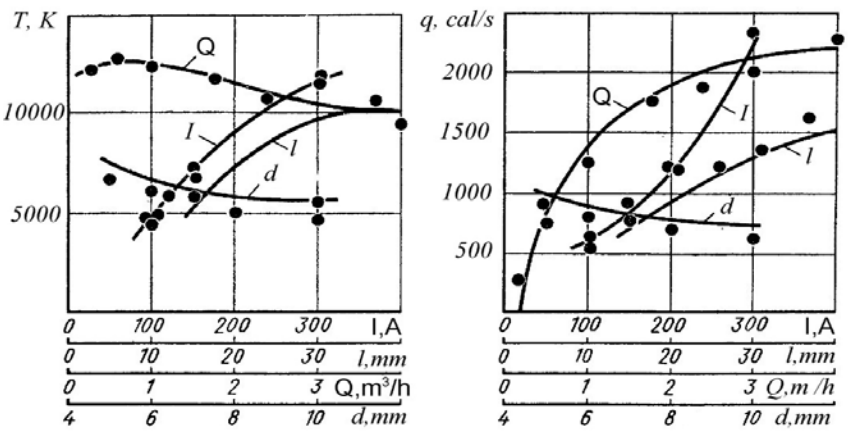


Fig. 6. Impact of the main parameters of the mode: current magnitude I (l=30mm, Q=2.35m^3/h, d=8mm), arc length l (I=200A, Q=2.35m^3/h, d=8mm), argon capacity Q (I=200A, l=30mm, d=8mm), and nozzle diameter d (I=200A, Q=2.35m^3/h, l=30mm) on average mass temperature T and thermal power q

As a result of convective and radiation heat removal, the effective thermal power of the plasma jet q_e is less than q. Increase of current magnitude and decrease of nozzle diameter result in its increase. When plasma-forming gas capacity is small, its increase results in sharp power increase, while with high capacity values its increase actually does not affect power.

Increasing the distance between the nozzle and the heated plane decreases effective power as a result of the increased losses. Convective heat removal increases with turbulent jets. This explains the sharper reduction of effective power with increased capacity. In case of interaction of the plasma jet with a surface perpendicular to its axial axis, thermal flow distribution is close to normal (Fig. 7).

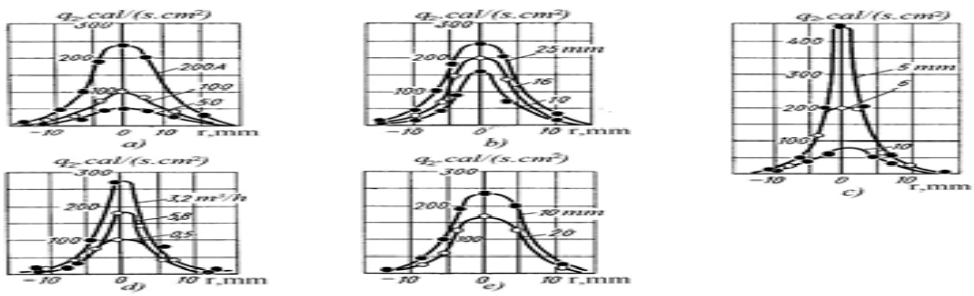


Fig. 7. Impact of current magnitude (a), arc penetration into the nozzle (b), nozzle diameter (c), argon capacity (d), and distance between the nozzle and the workpiece (e) on thermal flow distribution

The maximal density of the thermal flow (q_2) may change from values equivalent to gas flame to values corresponding to a welding arc. It increases with increase of current and arc length and decrease of nozzle diameter. Apart from the thermal impact, plasma jet also has noticeable power impact on the processed material (Fig. 8).

Plasma–jet–arc process with open and closed arc – combined process

The plasma–jet–arc process is implemented by combining open and closed arc in the same plasmatron. Here, two separately regulated direct current sources are used (Fig. 9). One powers closed arc 5 between tungsten electrode 4 and copper nozzle 7, and the other powers open arc 8 between the electrode and the workpiece. In contrast to the open-arc plasmatron, in

this case both arcs burn simultaneously and continuously during the entire welding process. Thus, the closed arc stabilizes further the process and, together with the strongly compressive action of anode nozzle 7, it supports the stable burning of the open arc even at very low current magnitudes. Combined plasmatrons are suitable for both welding and surfacing.

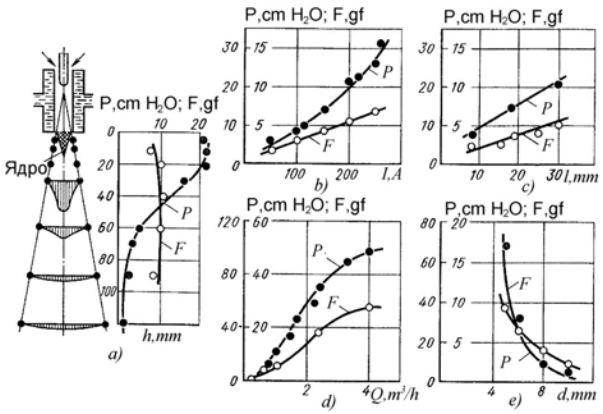


Fig. 8. Impact of the parameters on the pressure along the jet axis and the overall power impact (a – distance to nozzle b – current magnitude, c – arc penetration, d – argon capacity, e – nozzle diameter)

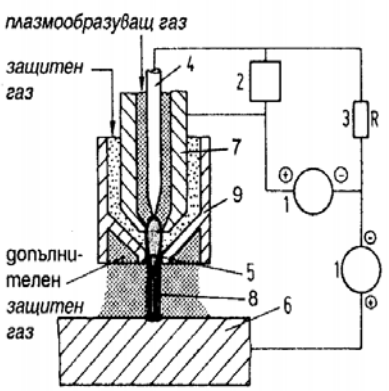


Fig. 9

The analysis of the publications in this field reveals numerous versions of the plasma welding and surfacing technological process:

- using additional material: powder or wire; moreover, wires may be one or two, with solid or tubular cross-section;

- single-arc or double-arc, whereas the arc may burn on the workpiece or on the welding wire, while with the double-arc process one arc burns on the workpiece and the other one burns on the wire or on the canal of the plasma-forming nozzle;
- the polarity of the direct-acting arc may alternate between direct and reverse, depending on the used electrode types;
- usually, argon or argon-hydrogen mixture is used as both plasma-forming and protective gas, but some combinations may be also used, such as: argon as plasma-forming gas, CO_2 – as protective gas;
- depending on the workpiece's complexity and size, surfacing may be implemented manually or by a mechanized process;
- mechanized surfacing may be implemented with no transverse oscillations or with oscillations of various trajectories;
- high-speed plasma surfacing;

Conclusion

The current technology development level, the economic situation, the dynamically changing user needs and requirements impose the need of manufacturing or restoring complex and precise articles within increasingly short time limits and with reduced production costs. An essential trend in the economy of spare parts, raw materials, materials and energy is the introduction of competitive industrial technological processes. Pulse–arc plasma processes have significant potential to produce new high-quality articles, or to repair or restore old ones.

References

1. A n a w a, E. M, A. G. O l a b i. Using Taguchi method to optimize welding pool of dissimilar laser-welded components. Optics & Laser Technology, Volume 40, Issue 2, 2008.
2. A s h b y, M. F., H. S h e r c l i f f, D. C e b o n. Materials: engineering, science, processing and design. Butterworth-Heinemann, 2007.
3. B e r n s H a n s, Microstructural properties of wear-resistant alloys. Wear, Volumes 181-183, Issue 1, 1995.
4. F i s c h e r, A. Well-founded selection of materials for improved wear resistance. Wear, Volume 194, Issues 1-2, 1996.
5. F o u i l l a n d, L., M. E l M a n s o r i, M. G e r l a n d. Role of welding process energy on the microstructural variations in a cobalt base superalloy hardfacing. Surface and Coatings Technology, Volume 201, Issue 14, 2007.

6. G a r c i a – A l l e n d e, P. B., J. M i r a p e i x, O. M. C o n d e, A. C o b o, J. M. L o p e z-H i g u e r a. Spectral processing technique based on feature selection and artificial neural networks for arc-welding quality monitoring. NDT & E International, Volume 42, Issue 1, 2009
7. H o u, R., D. M. E v a n s, J. M c C u r e, A. C. N u n e s, G. G a r c i a. Shielding gas and heat transfer efficiency in plasma arc welding. Welding Journal.Jia Chuan-bao, Wu Chuan-song, 1996.
8. L i u L i m i n g, H a o X i n f e n g. Improvement of laser keyhole formation with the assistance of arc plasma in the hybrid welding process of magnesium alloy. Optics and Lasers in Engineering, Volume 47, Issue 11, November 2009.
9. O b e r h e u s e r, G. Wear protection of surfaces from high pressure grinding rolls: possibilities and limits. International Journal of Mineral Processing, Volumes 44-45, 1996.
10. P a l a n i, P. K., N. M u r u g a n. Selection of parameters of pulsed current gas metal arc welding. Journal of Materials Processing Technology, Volume 172, Issue 1, 20 February 2006.
11. Weldind Handbook Ninth Edition, vol. 2, Welding Processes, Part 1, 2004.
12. Z h a n g Y u-m i n g. Sensing controlled pulse key-holing condition in plasma arc welding. Transactions of Nonferrous Metals Society of China, Volume 19, Issue 2, April 2009.
13. В а с и л е в, В. Технология на възстановяване на детайлите. Русе, 1996.
14. Г у д р е м о н, З. Специальные стали Tl. M. : Металлургия, 1966.
15. Ж е л е в, А. Н. Комплексен термодинамичен подход за оценяване горещата трошливост на металите при заваряване. Дисертация за получаване на научна степен „Доктор на техническите науки", София 1989.
16. Ж е л е в, А. Материалознание, техника и технология, Том II, Технологични процеси и обработваемост, София, 2003.
17. Л о л о в, Н. В. Заваряемост на металите, Част 1: Физични процеси., София 1995.
18. Технология злектрической сварки металлов и сплавов плавлением. Под ред. Б. Е. Патона, Москва, Машиностроение, 1974.
19. Т о н г о в, М. Заваряване, Част Първа, Процеси, София, 2009.

ИМПУЛСНО-ДЪГОВИ ПЛАЗМЕНИ ПРОЦЕСИ ЗА ЗАВАРЯВАНЕ И НАВАРЯВАНЕ

Р. Димитрова, Б. Табакова

Резюме

Направен е кратък обзор по литературни данни на импулсно-дъговите плазмени процеси за заваряване и наваряване. Разгледани са въпроси, свързани с оборудването и материалите, различните вариации на тези процеси, тяхното приложение, технологии за заваряване и наваряване, както и качеството на завареното изделие.

INFORMATION ABOUT THE START OF PROJECT TESTING PROBA-V AND VEGETATION DATA FOR AGRICULTURAL APPLICATIONS IN BULGARIA AND ROMANIA – PROAGROBURO

Eugenia Roumenina, Gheorghe Stancalie, Valentin Kazandjiev

The current VEGETATION 1 & 2 instruments onboard the French SPOT 4 and SPOT 5 satellites will only be available until 2012. For more than 10 years now, these instruments have monitored and mapped the worldwide vegetation every 10 days, thus providing essential information on crop yields, droughts, desertification, changes in the type of vegetation, deforestation, etc. to an ever extending user community. ESA is currently building the Sentinel 3 satellites in view of the European GMES programme. These satellites will contribute to the continuation of the availability of Vegetation type data but will not be operational in due time, thus creating a major time gap in the data continuity. Therefore, Belgium has decided to build a small satellite mission called PROBA-V ("V" standing for Vegetation). In that way, it will be a complement to the Sentinel 3 satellites to be launched after PROBA-V. While being designed as a continuity mission to the SPOT VEGETATION series, PROBA-V will provide some different characteristics, either through the technology used to collect data or through the enhancements in spatial resolution. The Preparatory Programme (http://eo.belspo.be) started in 2010 have two basic objectives:

• to get future users acquainted with these new data sets and their full characteristics and quality,

• to prepare the full exploitation of PROBA-V data sets with respect to the technical enhancements which are planned (spatial resolution in particular).

Spanning a 1-year period, the "PROBA-V Preparatory Programme" is funded by the Belgian federal government and managed by the Belgian Federal Science Policy Office (BELSPO) and the Proba V International Users Committee. In 2010, a call was announced for research proposals in the context of the PROBA-V Preparatory Programme. A number of 12 projects were selected (http://probav-iuc.org), which started their work in 2010. One

of these projects is *Testing PROBA-V and VEGETATION data for agricultural applications in Bulgaria and Romania – PROAGROBURO* with implementation period between 02/12/2010 to 31/12/2011. (http://proagroburo.meteoromania.ro).

The main objective of the project is to assess the quality of the PROBA-V mission as a continuity mission to VEGETATION 1 & 2 by comparison and validation of SPOT-VEGETATION and PROBA-V simulated data for assessing crop condition on chosen test areas for the territory of Bulgaria and Romania. This objective will be achieved by an interdisciplinary team of researchers from the two countries. The team comprises experts in remote sensing of the Earth and GIS technologies, as well as in agrometeorology and agro-forecasts. This team will be supported by the VITO SPS (System Performance Simulator) team which will provide a simulated PROBA-V dataset (http://www.vito.be) prepared based on hyperspectral EO-1/Hyperion data and multispectral SPOT 5 data.

Partners on (of) the *PROAGROBURO* Project are:

The Space and Solar-Terrestrial Research Institute – Bulgarian Academy of Sciences (SSTRI–BAS) is responsible for the overall implementation of the Project. The Principal Investigator of the Project is Assoc. Prof. Dr. Eugenia Roumenina with Promoter from the Romanian National Meteorological Administration (RNMA) – Dr. Gheorghe Stancalie, and Promoter from the National Institute of Meteorology and Hydrology – Bulgarian Academy of Sciences (NIMH–BAS) – Assoc. Prof. Dr. Valentin Kazandjiev.

The test areas in Bulgaria and Romania are chosen in the agricultural environments of Zhiten (Bulgaria), and Fundulea (Romania). Within each of them, a set of 4 test fields for field sampling will be chosen to perform the analysis and comparison of both spectroradiometers: SPOT VEGETATION and PROBA-V. The test area of Zhiten proposed for Bulgarian territory pertains to the Bulgarian Aero-Space Test Sites (BASTS). It is located in Dobrich Region, North-East Bulgaria, The test area on Romanian territory is located in Bargan Plain, South-East Romania.

A methodology to validate simulated PROBA-V and SPOT-VEGETATION data for agricultural applications will be developed. It will encompass three work tasks: building geodatabase; conducting sub-satellite experiments; and combined analysis of satellite and ground-based data. Three sub-satellite experiments for collecting meteorological data for each of the two test areas will evaluate and measure winter crop status during the

growing season. Common methodology will be applied on the two test areas to measure Fraction of Intercepted Photosynthetic Active Radiation (FIPAR), Leaf Area Index (LAI), soil moisture, canopy cover. Field data will be used to assess winter crop status by crop growth models (WOFOST and DSSAT). All field data and PROBA-V simulation images along with SPOT VEGETATION images for the growing season (2010–2011) will be organized in a geodatabase.

As a result of the analysis, the relation between satellite data from the two sensors and the ground-assessed crop status and LAI will be determined and statistically explained. Achieving useful results is based on validation of both sensors with referent ground-based data under the same conditions. The applied methodology has the following expected deliverables: methodological requirements, designing a geodatabase with integrated satellite and in situ biophysical data, establishing correlations between ground-based observations and satellite spectral indices, validating PROBA-V data, spatial statistics and pattern analysis of the simulated PROBA-V and SPOT VEGETATION derived indices and crop growth simulation and yield prediction. The project will also show the potentials of using VEGETATION type satellite data in addition to ground-based measurements and crop growth simulation models. The results are expected to aid the objective comparison of the two sensors, their performance and potential for combined usage in this application field.

The implementation of the project will contribute to the PROBA-V Preparatory Programme by acquiring independent and objective ground-based data that can be used to assess the quality of PROBA-V mission as a continuity mission to VEGETATION 1 & 2. The PROAGROBURO Project is dealing with one of the basic applications of the Vegetation instrument – agriculture, through which the VEGETATION 1 & 2 has gained respect among the user's community. Thus, it is essential to make sure that the PROBA-V mission is challenging the issue with even better quality. The combined use of PROBA-V with SPOT VEGETATION data will help improve agricultural services and products for Romania and Bulgaria. The regular monitoring will also help cope with emergency situations in the field of food production. The project will benefit the PROBA-V mission by providing validating tools and algorithms suitable for PROBA-V data.After the Project's implementation, on the selected test areas, ground-based and satellite data will continue to be collected on a regular basis, which might be included in the *in situ* component of the GMES and GEOSS Programmes.

A NEW BOOK IN THE FIELD OF AEROSPACE RESEARCH AND TECHNOLOGIES

In the end of 2011, the *Prof. Marin Drinov* Publishing House published the book *Semi-Nature Construction of Unmanned Aircraft Control Systems,* which is dedicated to an interesting topic, such as semi-nature construction of unmanned aircraft control systems.

The author, Prof. Petar Getsov, Director of the Space Research and Technology Institute at the Bulgarian Academy of Sciences (SRTI–BAS), is an expert of wide popularity among aviation and space community, who works in the field of control systems and avionics of aerospace aircraft.

The book provides a detailed and rationalized analysis of the modern state of unmanned aircraft, systematizing the spheres of their civil and military application, and proving their indispensability and prospects in complex aerospace experiments. General concept, methodology, and algorithms for semi-nature construction and study of unmanned aircraft are developed. The problems of operator selection and training are considered based on unique formalization of their activities. The possibility for uniting in a common complex the control system and the navigation systems is examined. The book also focuses on the study of man as a control system within the unmanned aircraft control contour, as well as on the on-board microprocessor systems used for the purpose. In the end, construction and experimental solutions investigating unmanned aircraft flight control through analytical, semi-nature, and experimental methods are highlighted.

The book contains 201 pages, 60 figures, and a reference of 237 titles. It is a further development of the habilitation work, defended successfully by the author to obtain the *Professor* scientific title.

The book of Prof. Getsov, *Semi-Nature Construction of Unmanned Aircraft Control Systems,* will be used by civil or military engineers, or researchers in the field of construction and study of unmanned aircraft, as well as by all readers concerned in the problems of unmanned aircraft control systems and satellite navigation systems.

The topic of the book, the discussed up-to-date problems, the high level of studies and quality of obtained results destine it to be a phenomenon on our book market in the field of aerospace research and technologies.

Prof. Garo Mardirossian, DSc

Г е ц о в, П. Полунатурно конструиране на системи за управление на безпилотни летателни апарати. Акад. Издат. “Проф. Марин Дринов”, София, 2011, 201 с.

A NEW BOOK ON NATURAL HAZARDS AND PROTECTION AGAINST THEM

During the second half of 2011, the book *Natural Hazards – Origin, Impact, Protection*, academic issue of AVIT CONSULT, was published. The authors, Prof. Garo Mardirossian, DSc and Assoc. Prof. Dr. Boyko Ranguelov, from the Space Research and Technology Institute, and Prof. Atanas Bliznakov, DSc, from the New Bulgarian University, are widely known experts in the field of study and prevention of natural hazards and ecological catastrophes not only in Bulgaria, but also in abroad.

The book is dedicated to a problem which is very topical both at home and throughout the world – the study of natural hazards, the negative impacts caused by them, and the prevention and protection against them.

The phenomenology, origin, main characteristics, victims, and damages caused by natural ecological hazards are considered, which may hit with varying severity the territory of our country – earthquakes, floods, storms and hurricanes, tsunami, avalanches, thunders, forest fires, landslides, extreme temperatures, hailstorms and more. The available prevention and protection measures for these natural hazards are considered, both engineering and organizational, as well as individual, namely behaviour and action prior to, during, and after the ecological catastrophe. For obvious reasons, engineering measures are described only in brief.

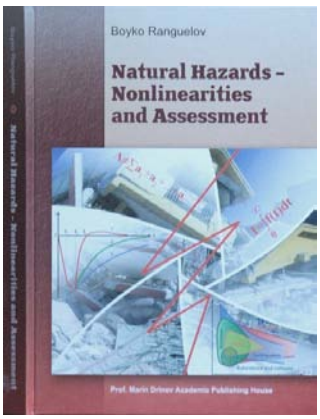
The figures, tables, and the coloured appendix, as well as the alphabetical index boost the better understanding of the topic and perceiving of the book. And the relatively voluminous reference allows the readers with specific interest in the topic to find more detailed information on the discussed phenomena and problems.

The book is intended for a wide readers' audience, but mostly for high school students. It may also be successfully used by primary school teachers to train the students how to protect themselves from natural hazards.

Prof. Radi Radichev

Мардиросян, Г., Б. Рангелов, А. Близнаков. *Природни бедствия – възникване, последици, защита*. Академично издание на „АВИТ КОНСУЛТ", София, 2011, 170 с., ISSN 978-954-92214-2-8.

NATURAL HAZARDS – NONLINEARITIES AND ASSESSMENT. NATURAL HAZARDS COMPLEXITIES, MULTIDISASTERS, METHODOLOGIES, RISK MAPPING, RISK MANAGEMENT, PREVENTION AND PROTECTION



Geoscientists are developing and applying a wide range of methodologies to estimate volcanic and seismic (including tsunami) hazards. Although significant advances in site characterization and model development have been made in the last decade, many challenges remain. Several disastrous earthquakes in the past decade have required a rapid assessment of the underlying causes of the tragic loss of life and property. Earthquake risk reduction and control as a crucial criterion for sustainable development, minimizing social and economic loss and disruption due to earthquakes, requires reliable assessment of seismic hazard, vulnerability of the built environment and risk. All of these provide the critical basis for improved building codes and construction emergency response plans. The European practice needs such approaches due to the increased risk which is definitely dominated by the increased urbanization and the improved quality of life. Dr. Ranguelov is an active participant of several projects developed by the Space Research and Technology Institute of BAS.

His book is devoted to natural hazard studies, where main attention is paid to the expression of the nonlinear properties and influences related to triggering, development and consequences of the natural hazards to environment and society.

Several important issues are presented following the content of this book:

- Fractal properties of the seismotectonic models of the Mediterranean and the Balkan Peninsula and their relationship with the main geological and tectonic structures can help to better understand and develop seismogenic models and achieve better practical results in implementing such approach in seismic hazard mapping.
- It is suggested that primary and secondary damages follow a clear nonlinear pattern. This could be rather helpful when assessing and calculating these damages.
- The deterministic approach using nonlinear functions to investigate human groups behaviour in extreme situations can help a lot in the everyday

management practice of the administrations in case of natural hazards occurrence.

- The nonlinear elastic plate rebound in case of strong earthquakes is a rather new phenomenon established after the strong earthquakes in Sumatra (2004), Chile (2010) and Japan (2011) by the GPS measurements. It can provoke deeper geology and geophysics investigations to discover such effects during the past geological times, when geodynamics was much more active.
- The complex risk assessment in case of multihazard simultaneous action of several hazards (for example earthquakes, tsunamis and landslides) is a big challenge to risk management practice and was developed in some models related to the Black Sea coast.
- The concept of the destructive potential, limited space-time and temporal development of natural hazards could be helpful in damage assessment and the implementation of risk management preventive and protection measures.
- The modern development of early warning systems as a promising tool for human life safety and society sustainability is essential. In case of great natural disasters (such as earthquakes, tsunamis, volcanic eruptions, etc.) these systems can protect and save human lives and preserve the environment.
- The systematization and data base creation about different useful practices and measures against the negative impact of different natural hazards is the genuine way to mitigate their consequences. The cost-benefit analysis in every specific case can help decision makers a lot in the selection and implementation of the most effective measure in each individual case.

The book is a useful tool for urban planners and the Civil Protection authorities and could be of interest not only to scientists, researches, students, but also to the wide public.

Dr. Zdenka Schenkova, Ph.D.
Head of the Department of Geodynamics Institute of Rock Structure and Mechanics - Academy of Sciences of the Czech Republic

Ranguelov, B. Natural Hazards – Nonlinearities and Assessment. Natural hazards complexities, multidisasters, methodologies, risk mapping, risk management, prevention and protection. *Professor Marin Drinov* Academic Publishing House, Bulgaria, 2011. ISBN 978-954-322-419-7