ESSENTIAL SPACEFLIGHT DYNAMICS AND MAGNETOSPHERICS

Boris V. Rauschenbakh
Michael Yu. Ovchinnikov
Susan McKenna-Lawlor

Space Technology Library

ESSENTIAL SPACEFLIGHT DYNAMICS AND MAGNETOSPHERICS

THE SPACE TECHNOLOGY LIBRARY

Published jointly by Microcosm Press and Kluwer Academic Publishers

An Introduction to Mission Design for Geostationary Satellites, J.J. Pocha
Space Mission Analysis and Design, 1st edition, James R. Wertz and Wiley J. Larson
Space Mission Analysis and Design, 2nd edition, Wiley J. Larson and James R. Wertz
Space Mission Analysis and Design, 3rd edition, James R. Wertz and Wiley J. Larson
Space Mission Analysis and Design Workbook, Wiley J. Larson and James R. Wertz
Handbook of Geostationary Orbits, E.M. Soop
Spacecraft Structures and Mechanisms, From Concept to Launch, Thomas P. Sarafin
Spaceflight Life Support and Biospherics, Peter Eckart
Reducing Space Mission Cost, James R. Wertz and Wiley J. Larson
The Logic of Microspace, Rick Fleeter
Space Marketing: A European Perspective, Walter A.R. Peeters
Fundamentals of Astrodynamics and Applications, David A. Vallado
Influence of Phychological Factors on Product Development, Eginaldo S. Kamata

---

The Space Technology Library Editorial Board

Essential Spaceflight Dynamics and Magnetospherics

by

Boris V. Rauschenbakh

Department of Theoretical Mechanics,
Moscow Institute of Physics and Technology, Moscow, Russia

Michael Yu. Ovchinnikov

Keldysh Institute of Applied Mathematics,
Russian Academy of Sciences, Moscow, Russia,
and Department of Theoretical Mechanics,
Moscow Institute of Physics and Technology, Moscow, Russia

Susan McKenna-Lawlor

Space Technology Ireland Ltd,
National University of Ireland,
Maynooth, Co. Kildare, Ireland

KLUWER ACADEMIC PUBLISHERS
NEW YORK, BOSTON, DORDRECHT, LONDON, MOSCOW

eBook ISBN: 0-306-48027-1
Print ISBN: 1-4020-1063-X

Visit Kluwer Online at: http://kluweronline.com
and Kluwer's eBookstore at: http://ebooks.kluweronline.com

This book is dedicated to our esteemed parents

Victor Yakovlevich Rauschenbakh and Leontina Fedorovna Rauschenbakh
Yuriy Fedorovich Ovchinnikov and Antonina Spiridonovna Ovchinnikova
James Christopher McKenna and Margaret McKenna, née Keane

with deep appreciation of their endless love, guidance and inspiration.

Preface

The present book **Essential Spaceflight Dynamics and Magnetospherics** describes, in the first instance, some of the key aspects of celestial mechanics and spaceflight dynamics. We begin with classical two and three body problems illustrative of the aesthetic aspects of applying analytical methods of investigation to celestial mechanics. Then, osculating orbital elements are introduced as well as analysis techniques sufficient to evaluate the influence of various disturbing forces on spacecraft. Next a theory of manoeuvres is outlined and the methodology of making interplanetary trajectory corrections. Ideas involving various approaches to orbital element determinations using measured data are also considered.

The forces applied to a spacecraft can result in the development of torques that influence attitude motion and the effects of the most important of these are described in terms of equilibrium positions, periodic motions, steady-state and transient motions. Also considered is the problem of attitude control of a spacecraft using active and/or passive methods of orientation and stabilization. In addition, a more advanced treatment of the development of attitude control systems is provided.

A description of the Earth's magnetic and gravitational fields allows us to clarify the relationship between natural features of the Earth's environment and the requirements of mission design, orbit construction and approaches to attitude control. A detailed Addendum provides an overview of circumstances on the Sun that render interplanetary space a very hazardous environment for spacecraft and for 'man in space'. The influence of this environment on spacecraft performance and survival is then presented, together with an outline of some of the mitigating strategies that can be invoked. A feature of the Addendum is the indication it provides of the challenges that the

next generation of space experiments will pose to mission designers. It is accompanied by a separate set of references since if refers to ongoing work in space physics rather than to classical material.

A set of appendices contain mathematical tools useful in making derivations and in obtaining formulae, but without the inclusion of detailed numerical data. There are many suitable reference books that provide material in this regard, some of which are listed in the references to individual Chapters. Other references relevant to the Chapters are contained in footnotes.

While it is assumed that readers of this book are familiar with mathematical analysis, differential equation theory, analytical mechanics and the theory of stability, the intention of the authors in Chapters 1 – 13 is to describe the main aspects of celestial mechanics and spaceflight dynamics without the use of complex mathematical formulae or the provision of examples of associated problem solving. This approach (breadth as opposed to depth) is for those readers who will deal in a general way only with space related topics. Chapters 14 – 22, which employ a more mathematically advanced treatment, are intended for readers interested at a specialist level in attitude dynamics and in the development of attitude control systems.

In the first fourteen chapters diagrams are frequently used to illustrate features of introductory spaceflight and attitude dynamics in lieu of complicated formula. In this regard, those innovative diagrams that were drawn manually by our co-author Boris Rauschenbakh, who died in 2001, are used in his memory with minimum changes. Only their inscriptions were rendered into English and, in some cases, an arrow over a letter or bold font was introduced to indicate a vector symbol. We hope that these visual presentations, with the historical dimension they provide through their association with Academician Rauschenbakh, will be attractive to those who are beginners in the area of astrodynamics, as well as to those who are already specialists, for whom they can serve to illuminate subtle details and provide new insights. The addition in footnotes of short historical biographies reflects the role of several centuries of international mathematicians in the achievement of space navigation and control.

The content concerned with spaceflight dynamics sprang initially from two courses of lectures for B.Sc., M.Sc. and Ph.D. students presented at the Moscow Institute of Physics and Technology. This

material has been significantly expanded here as well as modulated by the practical experience of the authors in developing international space projects in engineering, control, mechanics, and space physics over a period extending from the early days of this discipline up to the present time.

It is a pleasure to thank our associates, colleagues and friends for their support during the preparation of this book, and for the many important general discussions and personal contacts with them that stimulated new ideas and contributed a depth of experience. Among those we would particularly like to acknowledge in this regard are D.Okhotsimsky, V.Sarychev, N.Trukhan, V.Pen'kov, V.Sazonov, V.Rauschenbakh, I.Ritous and V.Kozminikh, all of whom made greatly valued inputs. Also, we extend our thanks to V.Beletsky, F.Graziani, M.Hapgood and M.Dryer who read our manuscript and made several important remarks.

Michael Yu.Ovchinnikov,
Moscow, Russia
ovchinni@keldysh.ru

Susan McKenna-Lawlor
Maynooth, Ireland
stil@may.ie

About the Authors

Rauschenbakh, Boris Victorovich (1915-2001). Doctor of Technical Science, Member of the Russian Academy of Sciences and the International Academy of Astronautics. One of the founders of practical cosmonautics in Russia. Worked in the 'Energia' Enterprise with Sergey Koroljev and later occupied the Chair of Theoretical Mechanics at the Moscow Institute of Physics and Technology.

Ovchinnikov, Michael Yurjevich, Doctor of Science in Physics and Mathematics at the Keldysh Institute of Applied Mathematics of the Russian Academy of Sciences and also a Professor at the Moscow Institute of Physics and Technology. Expert in mechanics, spaceflight dynamics and attitude control for small satellites.

McKenna-Lawlor, Susan, Emeritus Professor at the National University of Ireland, Maynooth. Guest Professor of the Chinese Academy of Sciences. Doctorate in Space Physics. Director of Space Technology Ireland, Ltd. Pioneering experiments flown on European, American, Russian and Chinese Space Agency Missions. Member of the International Academy of Astronautics. Recipient of the Tsiolkovsky Medal for Cosmonautics.

Table of Contents

Chapter 1

Unperturbed Orbital Motion. Two-Body Problem

As the two-body problem provides the classical basis for the analysis of any perturbed motion, it is suitable to begin an exploration of problems in celestial mechanics by considering the two-body problem in detail. Another name for this problem is the *problem of Kepler*,* who formulated it based on precise astronomical observations by Brahe†.

1.1 Equations of Motion

In accordance with Newton's Law of Gravitation, a force

$$\mathbf{F}_1 = -k^2 m_1 m_2 \frac{1}{r^2} \frac{\mathbf{r}}{|\mathbf{r}|},$$

acts on a particle with mass m_1 situated in the gravitational field of a particle with mass m_2, where $k^2 = 6.67 \cdot 10^{-11}\,\mathrm{m}^3/(\mathrm{kg} \cdot \mathrm{s}^2)$ is the *universal gravitational constant* and $\mathbf{r}$ is the the radius-vector of m_1 with respect to m_2. A field where the force depends on distance in

*Kepler, Johannes (1571–1630), German astronomer who discovered that the Earth and planets travel about the Sun in elliptical orbits. He deduced three fundamental laws that govern planetary motion.

†Brahe, Tycho (1546–1601), Danish nobleman who made numerous accurate astronomical observations used by Johannes Kepler in deducing his three fundamental laws of planetary motion.

this way is called a *Newtonian field.* The equations of particle motion are as follows

$$m_1 \frac{d^2\bar{\mathbf{r}}_1}{dt^2} = \mathbf{F}_1, \qquad m_2 \frac{d^2\bar{\mathbf{r}}_2}{dt^2} = \mathbf{F}_2, \qquad \mathbf{F}_2 = -\mathbf{F}_1, \tag{1.1}$$

where r_1 and r_2 are the radius-vectors of the particles with respect to an origin O_a in inertial space (Fig. 1.1). Let us define the radius-vector

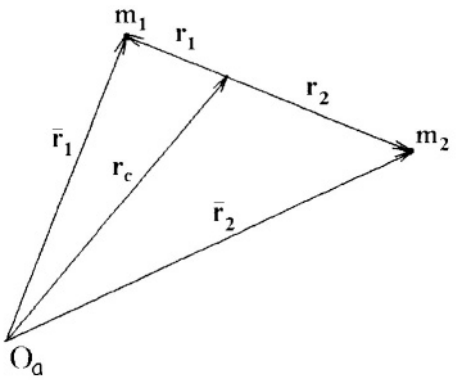

Figure 1.1. Two-particle system

$\mathbf{r}_c$ of the center of mass of this two-particle system using the formula

$$m_1\bar{\mathbf{r}}_1 + m_2\bar{\mathbf{r}}_2 = (m_1 + m_2)\mathbf{r}_c. \tag{1.2}$$

Differentiate this relationship twice with respect to time and substitute the two first expressions from (1.1) on the left side of the relationship obtained. Then, taking into account the last expression from (1.1), we get on dividing by $(m_1 + m_2)$

$$\frac{d^2\mathbf{r}_c}{dt^2} = 0. \tag{1.3}$$

Integrating this expression twice with respect to time we obtain

$$\frac{d\mathbf{r}_c}{dt} = \mathbf{C}_1$$

and

$$\mathbf{r}_c = \mathbf{C}_1 t + \mathbf{C}_2$$

where $\mathbf{C}_1$ and $\mathbf{C}_2$ are two vector constants of the *first integrals* (six scalar first integrals) that describe the linear and uniform motion of the center of mass of the two-particle system. A *first integral* of differential equations (or of a mechanical system the motion of which these equations describe), is a function dependent, generally speaking, on the coordinates, on the velocity of the mechanical system and on time. This function remains constant for any motion of this system.

Obtain now the remaining first integrals of the two-particle system. Locate a coordinate origin at its center of mass. Then, owing to the definition (1.2) after the substitutions

$$\vec{\mathbf{r}}_j = \mathbf{r}_c + \mathbf{r}_j \quad (j = 1, 2) \tag{1.4}$$

the following relation is valid

$$\mathbf{r}_2 = -\frac{m_1}{m_2}\mathbf{r}_1 \tag{1.5}$$

where $\mathbf{r}_1$ and $\mathbf{r}_2$ are the radius-vectors of the particles with respect to their joint center of mass. Consider now the motion of one of the particles, for instance m_1, and rewrite the expression for the force $\mathbf{F}_1$ acting on it

$$\mathbf{F}_1 = -\frac{\mathbf{r}_1}{|\mathbf{r}_1|}k^2 m_1 m_2 \frac{1}{r^2},$$

where $r = |\mathbf{r}_1| + |\mathbf{r}_2|$ is a magnitude of the vector $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$ which, as above, is the radius-vector of the first particle with respect to the second. Transform this formula as follows

$$\mathbf{F}_1 = -\frac{\mathbf{r}_1}{r^3}k^2 m_1 m_2 \left(1 + \frac{|\mathbf{r}_2|}{|\mathbf{r}_1|}\right) = -k^2 m_1 m_2 \frac{1 + \dfrac{m_1}{m_2}}{r^3}\mathbf{r}_1 = -k^2 m_1 M \frac{\mathbf{r}_1}{r^3}$$

where $M = m_1 + m_2$. Substituting this expression into the first equation of (1.1) we obtain

$$\frac{d^2\mathbf{r}_1}{dt^2} = -\mu\frac{\mathbf{r}_1}{r^3} \tag{1.6}$$

where $\mu = k^2 M$ (μ is called the *gravitational parameter*) and (1.3) and (1.4) are taken into account. Using relation (1.5) we get

$$\frac{d^2\mathbf{r}_2}{dt^2} = -\mu\frac{\mathbf{r}_2}{r^3}. \tag{1.7}$$

Consider next the mutual motion of the particles. Subtracting (1.6) from (1.7) we write

$$\frac{d^2\mathbf{r}}{dt^2} = -\mu\frac{\mathbf{r}}{r^3}. \tag{1.8}$$

Equations (1.6), (1.7) and (1.8) show that the motion of two particles around their joint center of mass, is similar to the motion of either particle relative to the other.

Let x, y, z be the projections of the vector $\mathbf{r}$ onto the axes of an *inertial reference system* (IRS) and write equation (1.8) in the coordinate form

$$\frac{d^2x}{dt^2} = -\mu\frac{x}{r^3}, \qquad \frac{d^2y}{dt^2} = -\mu\frac{y}{r^3}, \qquad \frac{d^2z}{dt^2} = -\mu\frac{z}{r^3}. \tag{1.9}$$

Six scalar first integrals thus remain to be obtained.

Note that all these results are valid for rigid bodies as well as for particles if the rigid bodies concerned have a centrally symmetric distribution of mass. In this sense the names 'two-body problem' and 'two-particle problem' are equivalent.

The word 'inertial' used in the name of the inertial reference system just introduced is conditional in nature because the possibility to consider a reference system as inertial depends on the accuracy with which the motions of the bodies (SC and celestial) considered here should be described. Usually for interplanetary missions an inertial reference system with its origin at the Sun's center is used. Missions around the Earth are frequently considered with respect to an inertial reference system with its origin located at the Earth's center.

1.2 Integral of Energy

Multiplying both sides of (1.8) by $2d\mathbf{r}/dt$ we obtain the scalar expression

$$2\frac{d\mathbf{r}}{dt}\frac{d}{dt}\left(\frac{d\mathbf{r}}{dt}\right) = -\frac{\mu}{r^3}2\mathbf{r}\frac{d\mathbf{r}}{dt}.$$

Using the obvious equality

$$\frac{d}{dt}(\mathbf{r}^2) = \frac{d}{dt}(r^2)$$

and integrating the previous equation we can write

$$\left(\frac{d\mathbf{r}}{dt}\right)^2 = \frac{2\mu}{r} + h.$$

Representing $d\mathbf{r}/dt$ by the vector $\mathbf{V}$, which has magnitude V, we obtain

$$V^2 = \frac{2\mu}{r} + h \tag{1.10}$$

which is the *integral of energy.* The motion is here seen to be determined by a *constant of energy* h such that at an infinite distance between the two particles

if $h > 0$	V is a real quantity
if $h = 0$	$V = 0$
if $h < 0$	V should be an imaginary quantity.

It follows from the third case that particles cannot reach an infinite distance from each other (such motion is called 'finite').

1.3 The Area Integral

From (1.8) we write

$$\mathbf{r} \times \frac{d^2\mathbf{r}}{dt^2} = 0.$$

Integrating this equation we obtain a first integral in vector form which is called the *area integral*

$$\mathbf{r} \times \frac{d\mathbf{r}}{dt} = \mathbf{c}. \tag{1.11}$$

This constant vector $\mathbf{c}$ is the *angular momentum* normalized by the particle mass. Let $\mathbf{i}, \mathbf{j}, \mathbf{k}$ be unit vectors of the IRS with origin at the second particle (we already saw that it does not matter with respect to which particle relative motion is considered). Then we write $\mathbf{c} = c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}$ and $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ so that the first integral (1.11) has the form

$$c_1 = y\frac{dz}{dt} - z\frac{dy}{dt},\ c_2 = z\frac{dx}{dt} - x\frac{dz}{dt},\ c_3 = x\frac{dy}{dt} - y\frac{dx}{dt}. \tag{1.12}$$

Multiplying (1.11) by vector **r** we obtain the scalar product

$$\left(\mathbf{r} \times \frac{d\mathbf{r}}{dt}\right)\mathbf{r} = c_1 x + c_2 y + c_3 z = 0. \tag{1.13}$$

This shows that the coordinates x, y, z of the moving particle satisfy the equation of a plane which passes through that particle where the origin of the IRF is located. The trajectory of motion of the particle lies in this plane.

If the unit vector **k** is aligned with vector **c**, then the expressions

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j}, \qquad \frac{d\mathbf{r}}{dt} = \frac{dx}{dt}\mathbf{i} + \frac{dy}{dt}\mathbf{j}, \qquad \mathbf{c} = c\mathbf{k}$$

are valid and $c = x\,dy/dt - y\,dx/dt$. Let us now introduce polar coordinates (r, θ) and, using the formulae $x = r\cos\theta$ and $y = r\sin\theta$, carry out the transformations

$$c = x\frac{dy}{dt} - y\frac{dx}{dt} = r\cos\theta\left(\frac{dr}{dt}\sin\theta + r\frac{d\theta}{dt}\cos\theta\right)$$
$$-r\sin\theta\left(\frac{dr}{dt}\cos\theta - r\frac{d\theta}{dt}\sin\theta\right) = r^2\frac{d\theta}{dt}.$$

The area integral is then derived in the form

$$c = r^2\frac{d\theta}{dt} = \text{const}\,.$$

From this expression we may make two conclusions:

1. Since $d\theta/dt = c/r^2$ then the greater the distance between the particles, the less is the angular velocity of **r**.

2. In a small interval of time Δt, the radius-vector **r** turns through a small angle $\Delta\theta$ covering a sector area ΔS (Fig. 1.2). To within an accuracy of the order of $\mathrm{O}\left((\Delta\theta)^2\right)$, the sector area is determined by the expression $\Delta S = r^2\Delta\theta/2$. From the limit

$$\frac{dS}{dt} = \lim_{\Delta t\to 0}\frac{\Delta S}{\Delta t} = \frac{1}{2}r^2\frac{d\theta}{dt}.$$

we obtain $dS/dt = c/2$. The derivative *dS/dt* which is called the *area velocity* is a constant when the motion takes place in a central

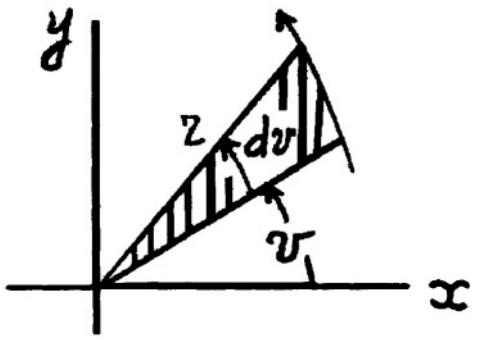

Figure 1.2. Area swept out by the radius-vector $\mathbf{r}$ in a small interval of time

gravitational field. The area S swept out by the radius-vector r during the time interval t_1 to t_2 is determined by the integral

$$S = \int_{t_1}^{t_2} dS = \frac{1}{2} c(t_2 - t_1).$$

1.4 Laplace's Integral

Next define the scalar

$$r' = \mathbf{r}\mathbf{V} = \mathbf{r}\frac{d\mathbf{r}}{dt}.$$

Then the equality

$$\frac{dr'}{dt} = \left(\frac{d\mathbf{r}}{dt}\right)^2 + \mathbf{r}\frac{d^2\mathbf{r}}{dt^2}$$

is valid. Taking into account the equation of motion (1.8) and the energy integral (1.10), rewrite this equation in the following form

$$\frac{dr'}{dt} = \frac{2\mu}{r} + h - \frac{\mu}{r} = \frac{\mu}{r} + h. \tag{1.14}$$

Differentiate this equality with respect to time and take into account

$$r' = \mathbf{r}\frac{d\mathbf{r}}{dt} = r\frac{dr}{dt}.$$

Then we get

$$\frac{d^2r'}{dt^2} = -\frac{\mu}{r^2}\frac{dr}{dt} = -\mu\frac{r'}{r^3}. \tag{1.15}$$

Equation (1.15) is formally similar to the scalar equations (1.9). Substituting r^3 from (1.15) in each of the three equations (1.9) we obtain

$$x\frac{d^2r'}{dt^2} - r'\frac{d^2x}{dt^2} = 0, \quad y\frac{d^2r'}{dt^2} - r'\frac{d^2y}{dt^2} = 0, \quad z\frac{d^2r'}{dt^2} - r'\frac{d^2z}{dt^2} = 0. \quad (1.16)$$

To integrate these equations the following auxiliary expressions can be written

$$\frac{d}{dt}(\mathbf{r}\frac{dr'}{dt}) = \frac{d\mathbf{r}}{dt}\frac{dr'}{dt} + \mathbf{r}\frac{d^2r'}{dt^2}, \quad \frac{d}{dt}(r'\frac{d\mathbf{r}}{dt}) = \frac{dr'}{dt}\frac{d\mathbf{r}}{dt} + r'\frac{d^2\mathbf{r}}{dt^2}.$$

Substitute the components of vectors $\mathbf{r}\,d^2r'/dt^2$ and $r'\,d^2\mathbf{r}/dt^2$ from these relations in (1.16) and integrate the resulting equations to obtain three constant expressions

$$f_1 = x\frac{dr'}{dt} - r'\frac{dx}{dt}, \quad f_2 = y\frac{dr'}{dt} - r'\frac{dy}{dt}, \quad f_3 = z\frac{dr'}{dt} - r'\frac{dz}{dt} \quad (1.17)$$

which may be interpreted as components of the constant vector

$$\mathbf{f} = \mathbf{r}\frac{dr'}{dt} - r'\frac{d\mathbf{r}}{dt}.$$

This is the vector form of *Laplace's integral*[‡] where vector $\mathbf{f}$ is *Laplace's vector.* The expression

$$\mathbf{f} = \frac{d\mathbf{r}}{dt} \times \mathbf{c} - \mu\frac{\mathbf{r}}{r}$$

provides another form of Laplace's integral.

From the theory of differential equations it is known that a set of autonomous differential equations of order n has only $n-1$ first integrals which are all independent of time. We obtained above the scalar first integral of energy (1.10), three scalar first integrals of area

[‡]Laplace, Pierre-Simon (1749–1827), French mathematician, astronomer and physicist, Professor of the École Militare. He studied applications of Newton's theory of gravitation to the Solar System and discovered the invariability of planetary mean motions. Taking into account perturbations, he showed planetary motions to be secular and that the overall Solar System is stable. Laplace introduced the concept of potential — a function whose directional derivative at every point is equal to the component of the field vector of intensity in a given direction. Also, the term *celestial mechanics* was introduced by Laplace.

(1.12) and three scalar first integrals of Laplace. None of these depend on time. Their total number is greater by 2 than is allowed for the set of autonomous equations (1.8) of order 6. This means that the seven first integrals are not all independent and we need to limit their number to 5 independent integrals.

Laplace's vector relates to the vector **c** and to the constant of energy h through the expressions

$$\begin{aligned} \mathbf{cf} &= \left(\mathbf{r} \times \frac{d\mathbf{r}}{dt}\right)\left(\mathbf{r}\frac{dr'}{dt} - r'\frac{d\mathbf{r}}{dt}\right) \\ &= \left(\mathbf{r} \times \frac{d\mathbf{r}}{dt}\right)\mathbf{r}\frac{dr'}{dt} - \left(\mathbf{r} \times \frac{d\mathbf{r}}{dt}\right)r'\frac{d\mathbf{r}}{dt} = 0, \qquad (1.18) \\ f^2 &= \mu^2 + hc^2 \end{aligned}$$

so, the magnitude of vector **f** is dependent on the constants of energy and of angular momentum. Since vector **f** is perpendicular to vector **c** and consequently lies in the plane of motion, and also its magnitude is fixed, only its direction in this plane can be chosen.

Let us next consider a consequence of the existence of one independent scalar Laplace integral

$$\begin{aligned} \mathbf{rf} &= xf_1 + yf_2 + zf_3 \\ &= x\left(x\frac{dr'}{dt} - r'\frac{dx}{dt}\right) + y\left(y\frac{dr'}{dt} - r'\frac{dy}{dt}\right) + z\left(z\frac{dr'}{dt} - r'\frac{dz}{dt}\right) \\ &= r^2\frac{dr'}{dt} - (r')^2. \end{aligned}$$

Using the Laplace identity $(\mathbf{a} \times \mathbf{b})^2 = a^2b^2 - (\mathbf{ab})^2$, and the integral of energy, we derive

$$\begin{aligned} c^2 &= (x^2 + y^2 + z^2)\left[\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2\right] \\ &- \left(x\frac{dx}{dt} + y\frac{dy}{dt} + z\frac{dz}{dt}\right)^2 = r^2\left(\frac{2\mu}{r} + h\right) - (r')^2. \end{aligned}$$

Extract the expression $(r')^2$ from this formula and put it in the previous equality to obtain the scalar product **rf**. Then, taking into account (1.14) we get

$$\mathbf{rf} = xf_1 + yf_2 + zf_3 = r^2\left(\frac{\mu}{r} + h\right) + c^2 - r^2\left(\frac{2\mu}{r} + h\right) = c^2 - \mu r.$$

The trajectory of the particle can now be determined by combining this relation with the equation of the plane (1.13)

$$c^2 - xf_1 - yf_2 - zf_3 = \mu r, \qquad c_1x + c_2y + c_3z = 0.$$

Let us now introduce a new reference system $O\xi\eta\zeta$, where $O\xi$ is directed along the vector **f**, and $O\zeta$ along the vector **c**. This can be done due to the validity of the relationship **cf** = 0. Then, since $c_1 = c_2 = 0$, $c_3 = c$ and $f_1 = f$, $f_2 = f_3 = 0$, the previous system assumes the form

$$\mu r = c^2 - \xi f, \qquad \zeta = 0$$

where $r = \sqrt{\xi^2 + \eta^2}$. Employ next polar coordinates (radius-vector r and *true anomaly* υ). Then, using the standard transformations $\xi = r\cos\upsilon$, $\eta = r\sin\upsilon$, we obtain $\mu r = c^2 - fr\cos\upsilon$ and resolve it with respect to r

$$r = \frac{p}{1 + e\cos\upsilon} \tag{1.19}$$

where the expressions

$$p = \frac{c^2}{\mu}, \qquad e = \frac{f}{\mu}. \tag{1.20}$$

define the *parameter* p and the orbit *eccentricity* e.

Hence, particles in the framework of the two-body problem move along conic, sections (hyperbola, parabola or ellipse) created by the intersection of a plane with the surface of a cone. At $\upsilon = 0$ the magnitude of r is a minimum. The corresponding point of the trajectory is called *pericenter* while the furthermost point of the trajectory at $\upsilon = \pi$ is called *apocenter*. The apocenter is pertinent only for elliptical trajectories. Since $\upsilon = 0$, the vector **r** is directed along the vector **f**. This is why one can say that Laplace's vector is directed towards the pericenter. The straight line coincident with the semi-major axis of the orbit and directed along Laplace's vector is called the *line of apsides*.

The fact that any eccentricity e is determined by the magnitude of Laplace's vector follows from (1.20). As $0 \leq e < 1$; $e = 1$ or $e > 1$ the trajectory is an ellipse, a parabola and a hyperbolic curve respectively. A closed elliptical trajectory is frequently called an *orbit*, although real

orbits are never closed. In order to express e through other parameters one needs to use formulae (1.18) and (1.20) to obtain

$$e = \frac{\sqrt{\mu^2 + hc^2}}{\mu} = \sqrt{1 + \frac{c^2}{\mu^2}h} \tag{1.21}$$

and

$$c = \sqrt{p\mu}, \qquad h = \frac{\mu^2}{c^2}(e^2 - 1) = \frac{\mu}{p}(e^2 - 1). \tag{1.22}$$

The relationship between the coordinates and explicit time is determined by the sixth first integral and we obtain it using the polar reference system introduced above. As was already shown, the area integral in this reference system has the form

$$r^2 \frac{dv}{dt} = c. \tag{1.23}$$

Substituting here the expression for r from (1.19) we obtain after integration

$$t - \tau = \frac{p^2}{c} \int_0^v \frac{dv}{(1 + e\cos v)^2}, \tag{1.24}$$

where τ is a constant called the *epoch* (that is the moment of time when a moving particle first passes the pericenter of its trajectory). Note, that although the quadrature (1.24) provides a formal solution of the problem, it is not convenient to use this expression in an analytical treatment of motion.

Six first integrals and six constants corresponding to them have now been obtained. Owing to our initial choice of a suitable reference system, three of these constants are equal to zero, because the plane (x, y) is the plane of motion and the axis Ox is directed along vector **f** to the pericenter. These latter constants specify the position of the orbital plane in space and also the orientation of the orbit in this plane. The solution has thus only three non-zero constants and these are p, e, τ.

1.5 Kepler's Laws

It was shown above that motion of a particle in the two-body problem is executed along a conic section with the central attracting body

placed at its focus. Such motion is called *Keplerian* and the path of the particle is called a *Keplerian orbit* or a *Keplerian trajectory.* Also it was shown that a radius-vector sweeps out equal areas during equal intervals of time. Further to these results we may now consider the period of revolution (T_{rev}) of a body moving along an elliptical orbit. It is known that the area of an ellipse S is equal to πab, where a and b are the semi-major axis and the semi-minor axis respectively. Then, from the formula for a 'swept-out' area

$$S = \frac{1}{2}c(t_2 - t_1)$$

we get

$$\pi ab = \frac{1}{2}cT_{rev}.$$

Again, from the definition of p an expression $c = \sqrt{p\mu}$ follows and so

$$2\pi ab = \sqrt{p\mu}T_{rev}.$$

From analytical geometry $b = a\sqrt{1-e^2}$ and $p = a(1-e^2)$. Thus, the relation

$$T_{rev} = 2\pi\frac{a^{3/2}}{\sqrt{\mu}} \tag{1.25}$$

is valid. Hence, if two bodies move around any much larger central body (in the absence of mutual attraction between the two orbiting bodies) so that the value of μ is substantially determined by the mass of the central body rather than by the masses of the orbiting bodies (note that the expression for μ introduced in (1.6) is dependent on both masses), then the orbital periods T_1 and T_2 are in accordance with the equality

$$\left(\frac{T_1}{T_2}\right)^2 = \left(\frac{a_1}{a_2}\right)^3.$$

It is worth mentioning that this result and that concerning particle motion along a conic section are valid only for a Newtonian field. The result that the velocity with which a rotating vector sweeps out a particular area is constant is valid for any central field.

These theoretical results can be compared with results published by Kepler in the form of three laws

1. Each planet moves around the Sun in an ellipse with the Sun at one focus (a trajectory of motion lies in a plane).

2. The line from the Sun to a planet (the radius-vector **r**), sweeps out equal areas in equal time intervals.

3. The ratio of the square of the period to the cube of the semi-major axis, is the same for all planets in our solar system.

derived on the basis of his analysis of accurate observations of planetary motions made by Tycho Brahe. The first and second laws were published in 1609 and the third in 1619.

1.6 Kepler's Equation

Equation (1.24) provides a relationship between coordinates and time, but it has a highly complicated form. This is a reason to use instead *Kepler's equation.* Fig.1.3 shows an ellipse with semi-axes a and b

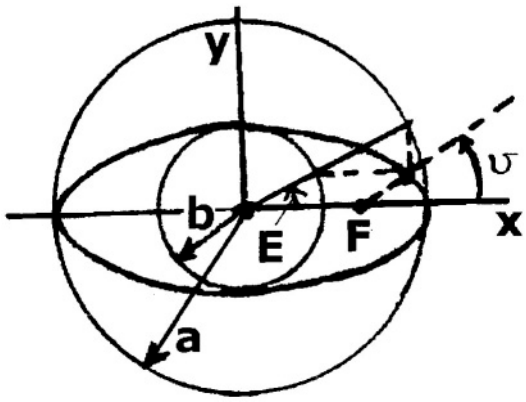

Figure 1.3. Ellipse construction using two concentric circles

constructed using concentric circles with radii a and b. Introduce the *eccentric anomaly* E to determine the coordinates of points which are situated on the ellipse using

$$x = a\cos E, \qquad y = b\sin E \tag{1.26}$$

Then the trigonometric expressions

$$r = a(1 - e\cos E), \qquad \tan\frac{v}{2} = \sqrt{\frac{1+e}{1-e}}\tan\frac{E}{2} \tag{1.27}$$

are valid, and the formulae

$$r\cos v = a(\cos E - e), \qquad r\sin v = b\sin E$$

are true in terms of Fig. 1.3. Squaring both parts of these equations we obtain an equation in r^2. From this, an expression for r follows

$$r = [a^2(\cos E - e)^2 + b^2\sin^2 E]^{1/2}.$$

Now, going to half-angles and taking into account this expression for r we obtain

$$2r\cos^2\frac{v}{2} = a\cos E - ae + a - ae\cos E = 2a(1-e)\cos^2\frac{E}{2},$$
$$2r\sin\frac{v}{2}\cos\frac{v}{2} = 2b\sin\frac{E}{2}\cos\frac{E}{2} = 2a\sqrt{1-e^2}\sin\frac{E}{2}\cos\frac{E}{2}.$$

Dividing the second equation by the first, we obtain an expression for $\tan(v/2)$. The angle E uniquely determines commonly used coordinates of the ellipse points (r, v).

The integral of area (1.23) and the energy integral (2.1) have, using (1.22), the forms

$$r^2\frac{dv}{dt} = \sqrt{\mu p} = \sqrt{\mu a(1-e^2)}$$

and

$$\left(\frac{d\mathbf{r}}{dt}\right)^2 = \left(\frac{dr}{dt}\right)^2 + r^2\left(\frac{dv}{dt}\right)^2 = \mu\left(\frac{2}{r} - \frac{1}{a}\right)$$

respectively. Excluding dv/dt we obtain

$$\left(\frac{dr}{dt}\right)^2 = \frac{\mu}{r^2 a}[a^2e^2 - (a-r)^2].$$

Now introduce *mean motion* using

$$n_{ev} = \frac{2\pi}{T_{rev}} = \frac{\sqrt{\mu}}{a^{3/2}}$$

where the orbital period is taken from (1.25). The mechanical sense of n_{ev} is *mean angular velocity* of revolution around a central body. Then the formulae

$$\frac{dr}{dt} = n_{ev}\frac{a}{r}\sqrt{a^2e^2 - (a-r)^2}$$

and

$$n_{ev}dt = \frac{rdr}{a\sqrt{a^2e^2 - (a-r)^2}}$$

are valid. Substitute E for r using (1.27), then

$$n_{ev}dt = (1 - e\cos E)dE.$$

Integrating this equality, we obtain

$$n_{ev}(t - t_0) = (E - E_0) - e(\sin E - \sin E_0).$$

Choose now a reference point t_0 corresponding to $E_0 = 0$. Then, as is obvious from the second equation of (1.27), $\upsilon = 0$ and consequently the point $E_0 = 0$ corresponds to pericenter. Hence, t_0 can be used as the epoch τ. In this case the previous equation has the form

$$n_{ev}(t - \tau) = E - e\sin E$$

and

$$E - e\sin E = M \tag{1.28}$$

where $M = n_{ev}(t - \tau)$ is called the *mean anomaly.* This corresponds to the turning angle of a radius-vector of a moving body which rotates around an attracting body with a permanent angular velocity equivalent to its mean motion n_{ev}.

Equation (1.28) is called *Kepler's equation* and it is noted that, despite its simple form, it can not be solved using elementary functions. This equation can, however, be used to obtain answers to the following problems.

- For a given position of a moving body (i.e. when r and υ are given), determine the corresponding time t.
- For a given t, determine r and υ.

The solution of the first problem is elementary since r and v determine E by formula (1.27) and we can immediately obtain t from the mean anomaly M (1.28).

The second problem requires E to be determined for a specified M which is related to the given t through the expression $M = n(t - \tau)$ using (1.28). Solving this is possible employing approximate methods (for instance, by a *successive approximations method*), such as that described on page 260. Determining the relation of the $(n + 1)$-th approximation of E_{n+1} from the previous E_n using Kepler's equation we may write

$$E_{n+1} = e \sin E_n + M.$$

It can be shown that such a procedure converges to a unique solution $\overline{E}$ which is the root of the equation

$$E = e \sin E + M.$$

In practice, when E increases from $-\infty$ to $+\infty$ the function $E + e \sin E$ monotonically increases because its first derivative is always positive. It follows that Kepler's equation has a unique solution for every value of M. If the magnitude of the eccentricity e is sufficiently small, it is convenient to represent the variables r and v by a time series.

Chapter 2

Qualitative Analysis of the Properties of Orbits

Next we employ the integral of energy when considering, and interpreting, the properties of elliptical orbits.

The expression for the energy constant (1.22) and the geometrical relation $p = a(1-e^2)$ allow us to express the energy integral in a form

$$V^2 = \mu\left(\frac{2}{r} - \frac{1}{a}\right) \tag{2.1}$$

also called the *vis-viva equation* [2]. Note that a comparison of the energy integral (1.10) with expression (2.1) gives $h = -\mu/a$, so that the semi-major axis actually provides a measure of the energy of elliptical orbits. Let us consider now the evolution of such orbits based on an analysis of formula (2.1).

2.1 Orbit Evolution due to an Instant Velocity Pulse

If we accept the assumption that imparting an instantaneous velocity increment ΔV (or *instant velocity pulse*) to a *spacecraft* (SC) does not change its coordinates but only changes its velocity then, as shown by (2.1), when V^2 changes instantly (due to an 'instant pulse'), a also changes instantly (i.e. a transfer from orbit *1* to another orbit *2* occurs). It is shown in Fig.2.1 how an initially circular orbit is

modified by the application of an instant pulse ΔV directed 'along the velocity' and 'opposite to the velocity' vector (these are turned *acceleration* and *deceleration* pulses respectively). As illustrated, the greatest geometric difference in the orbits takes place in each case at that location opposite to the point where the pulse was applied.

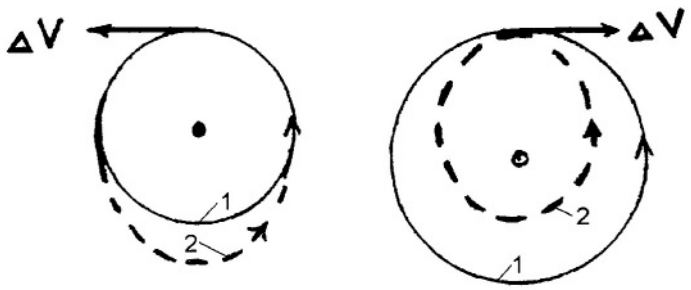

Figure 2.1. Effect of applying an instant pulse along the velocity vector (left) and opposite to the velocity vector (right) of a SC

2.2 Orbital Transfer

Orbital transfer requires two pulses if the initial and final orbits do not initially have any common points (for instance in the case of a transfer from one circular orbit to another). In Fig.2.2, *two-pulse transfer* from circular orbit *1* to circular orbit *3* via a transfer orbit *2* using pulses ΔV_1 and ΔV_2 is illustrated. It is assumed that all orbits lie in the same plane. Transfer ellipse *2* is called *Hohmann's ellipse**, and the

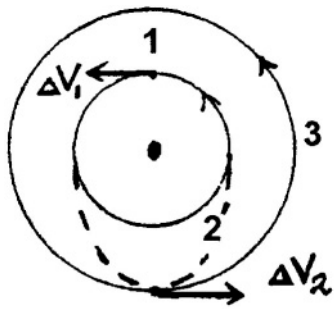

Figure 2.2. Two-pulse transfer via Hohmann's ellipse *2*

associated orbital manoeuvre is called the *Hohmann trasfer.*

2.3 SC Braking in the Atmosphere

Consider a SC in an elliptical orbit about the Earth and suppose that braking (which can be approximated by an instant 'braking pulse'), takes place in the atmosphere at each pericenter passage, thereby producing a progressive decrease in the altitude of the apocenter. In consequence, the original orbit shrinks and becomes less elliptical (Fig.2.3). It is easy to recognize this situation through the ongoing reduction of the revolution period T_{rev} which is proportional (see (1.25)) to the 3/2 power of the semi-major axis a. As the whole orbit lies within the atmosphere, this braking is so intensive that a SC can descend to the Earth's surface from a critical altitude of ~ 100 km.

Figure 2.3. SC braking in the Earth's atmosphere

*Hohmann, Walter (1880–1945), German specialist in spaceflight mechanics. He demonstrated (W.Hohmann, Die Erreichbarkeit der Himmelskorper, Oldenbourg, Munich, 1925) that transfer from one circular orbit to another via an ellipse tangential to both orbits provides, from the point of view of energy consumption, the optimum optionIn the same period Russian researcher and engineer Tsander, Fridrich Arturovich (1887–1933) analyzed interplanetary trajectories and obtained a similar result (F.A.Tsander. Flights to Other Planets (*Theory of Interplanetary Voyages*), 1924–1925. In the Selected Papers of the Pioneers of Rocket Engineering: Kibalchich, Tsiolkovsky, Tsander, Kondratyuk, Moscow, Nauka Publ., 1964, pp.277–369).

2.4 Energy Consumption Required for Interplanetary Flight

The energy integral (2.1) may be rewritten in the form

$$V^2 = \frac{\mu}{r}\left(2 - \frac{r}{a}\right).$$

If r represents the distance from the center of gravity to the launch site (i.e. the Earth's radius) and $2a$ is the distance from the Earth to the orbit of a target planet (Fig.2.4), then it is clear that the velocity V required for flight depends only slightly on a. For the Moon the ratio $r/a \approx 6/400$ is a negligible quantity in comparison with the magnitude (2) of the first term in brackets. That is why according to this formula, obtained within the framework of the two-body problem, the velocities required for speeding-up a flight to the Moon and to Mars have comparable values. However, in practice the values concerned differ owing to the necessity, in each case, to take into consideration the attraction of the Sun.

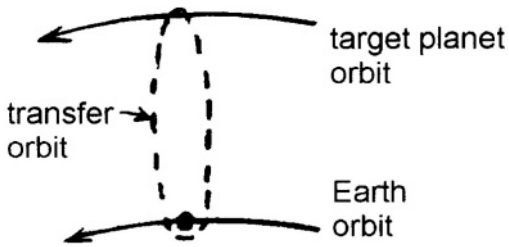

Figure 2.4. Interplanetary flight

2.5 Circular Velocity

Substituting $a = r = \text{const}$ in (2.1) we obtain the velocity of circular motion for orbits of arbitrary radius

$$V_c = \sqrt{\frac{\mu}{r}}. \tag{2.2}$$

Consequently, the velocity of circular motion changes according to the inverse-square-root of the distance r from the central attracting body.

Hence, an infinite number of such velocities exist. One of these, which corresponds to the minimum energy necessary for orbital injection, is called in western technical literature the *circular velocity*,

$$V_{circ} = \sqrt{\frac{\mu}{R_e}} \tag{2.3}$$

where, in the case of motion around the Earth, R_e is the Earth's radius and μ a gravitational parameter of the Earth. When motion around other planets is considered, the appropriate parameters should be substituted in (2.3). Sometimes (mostly in Russian technical literature), this circular velocity is called the *first cosmic (circular) velocity* which is convenient in order to distinguish it from the velocity of a SC V_c in a circular orbit of arbitrary radius.

Compare two idealized schemes for the orbital injection of a SC from an airless, spherical planet (Fig.2.5). In the first scheme the SC is provided with an instant pulse in the horizontal direction, which injects it into a circular orbit of infinitesimal altitude. In the second case the SC is first launched vertically, then given a horizontal pulse sufficient to produce a velocity that is less than the velocity of circular motion attained in the first case and only just enough to maintain an elliptical orbit. The first trajectory intersects points *A* and *C*

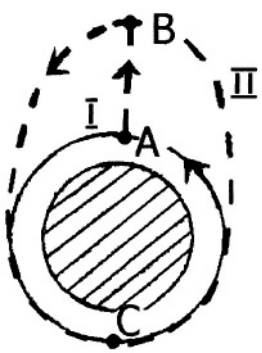

Figure 2.5. Two idealized schemes for SC orbital injection from an airless spherical planet

at an infinitesimal distance above the planet's surface. The second trajectory includes the lift from *A* to *B*. The resulting elliptical orbit passes through the points *B* and *C*.

Table 2.1. Magnitudes of circular velocity V_{circ} for various planets (unit: km/s)

Moon	1.7	Jupiter	42.2
Mercury	3.0	Saturn	25.1
Venus	7.3	Uranus	15.4
Earth	7.9	Neptune	16.6
Mars	3.6	Pluto	3.8

Both orbits touch each other in the point *C* where they have the same magnitude of r. Orbit *II* is characterized by a value of the semi-major axis a, which is larger than that of the semi-major axis of orbit *I*. Hence, the velocity of the body revolving in orbit *II* is greater than the velocity of the body revolving in orbit *I*. Taking into account that the beginning of speed-up (point *A*) is coincident for both orbits, it is necessary to expend more energy to reach orbit *II* than to reach orbit *I*. This means that minimum energy consumption corresponds to circular motion around the planet at the infinitesimal altitude. The circular velocity concerned is calculated using formula (2.3).

In practice, due to the presence of the Earth's atmosphere, the trajectory of injection into an orbit starts vertically from the launch site in order to minimize passage through the atmosphere. Thereafter, to reduce the carrying out of work against gravitational force during the speed-up phase, the trajectory is turned gradually towards the horizontal. The magnitudes of circular velocity V_{circ} for various planets are presented in Table 2.1.

Rewriting (1.25), we obtain a relationship between the SC period of revolution and the orbit size (semi-major axis)

$$a = \left(\frac{\mu T_{rev}^2}{4\pi^2} \right)^{1/3}$$

Substituting for T_{rev} the *sidereal day* corresponding to the revolution of a particular planet, provides us with the semi-major axis of the orbit of a SC having an orbital period equal to the period of the planet's spin. A sidereal day is defined as the period required for one spin revolution of the planet about its spin axis relative to inertial space. One sidereal day on the Earth is 23 h 56 min 4 s, whereas one *solar*

day is 24 h. We thereby obtain the semi-major axis of what is called *geo-synchronous orbit.* Each geo-synchronous orbit is characterized by a particular inclination. Generally, a SC moves relative to the Earth's surface in the same meridian plane (travelling from North to South, then returning to the north across the equatorial plane). Among geo-synchronous orbits there exists one characterized by a fixed position of the SC relative to the Earth. This orbit is circular, lies in the equatorial plane and is called the *geo-stationary orbit.* Its radius is 42 164 km.

2.6 Escape Velocity

If at speed-up the SC velocity reached is sufficiently large, then the resulting trajectory will not be finite. The minimum velocity in this case corresponds to that which provides motion along a parabola. Considering the parabola as an extreme case of an ellipse with semi-major axis $a \to \infty$, then, in terms of formulae (2.1) and (2.3), we obtain

$$V = \sqrt{\frac{2\mu}{R_e}}. \tag{2.4}$$

Comparing this velocity with circular velocity we can define the *escape velocity* as

$$V_{esc} = \sqrt{2} V_{circ}. \tag{2.5}$$

Thus, the escape velocity (sometimes called the *parabolic velocity*) or the *second cosmic velocity* is that minimum velocity sufficient to move a SC (in the framework of the two-body problem) to an infinite distance from its attracting body. The magnitudes of the escape velocity V_{esc} for various planets are presented in Table 2.2. In actuality, speed-up to V_{esc} does not mean the removal of a SC to infinity as, besides the Earth, the Sun exerts an influence and forces the SC to move around it in an ellipse.

2.7 Hyperbolic Escape Velocity

The minimum velocity required to allow a SC to exit the Solar System can be calculated using integrals of energy for the motion of the

Table 2.2. Magnitudes of escape velocity V_{esc} for various planets (unit: km/s)

Moon	2.4	Jupiter	59.6
Mercury	4.3	Saturn	35.5
Venus	10.3	Uranus	21.8
Earth	11.2	Neptune	23.4
Mars	5.0	Pluto	5.4

SC in the gravitational fields of the Sun and the Earth. To leave the Solar System the SC has to be provided with the escape velocity $V_{esc\,S}$. Using (2.1) we obtain the magnitude of this velocity for the Earth's orbit $V_{esc\,S} = (2\mu_S/r_{SE})^{1/2} \approx 42.1$ km/s, where μ_S is the gravitational parameter of the Sun and r_{SE} is the radius of the Earth's orbit around the Sun (a circular approximation to the Earth's orbit is considered). The velocity of the Earth's orbital motion $V_{circ\,E} = (\mu_S/r_{SE})^{1/2} \approx 29.8$ km/s. The value of the additional velocity V_{add} required to achieve escape from the Solar System depends on its direction with respect to the direction of $V_{circ\,E}$. The minimum value $V_{add} = V_{esc\,S} - V_{circ\,E} \approx 12.3$ km/s can be enough if V_{add} is aligned with the velocity of the Earth's orbital motion. The SC should already have this 'extra' velocity when leaving the gravitational field of the Earth so that the integral of energy for the SC in the Earth's field has the form

$$V^2 = \frac{2\mu}{r} + V_{add}^2.$$

To determine the total velocity required when starting from the Earth (at $r = R_e$) and acquiring V_{add}, we derive from this formula a set of expressions for the minimum magnitude of this velocity

$$V_{h\,esc} = \left(V_{add}^2 + 2\mu/R_e\right)^{1/2} = \left[(V_{esc\,S} - V_{circ\,E})^2 + V_{esc}^2\right]^{1/2} \approx 16.6\,\text{km/s}$$

which is called the *hyperbolic escape velocity* (sometimes the *third cosmic velocity*). The maximum magnitude of the velocity required to leave the Solar System is obtained when the additional velocity V_{add} is oppositely directed to the orbital velocity of the Earth. In this case, about 72.8 km/s is required overall to achieve escape from the Solar System.

Chapter 3

Perturbed Motion

In celestial mechanics when two massive particles move in orbits under their mutual gravitation (two-body problem), the motion is considered to be unperturbed and is called Keplerian. All cases of motion which deviate from Keplerian are associated with *perturbed motion.* The n-body problem provides a classical example of perturbed motion.

3.1 The n-Body Problem

Let n particles interact according to Newton's Law of Universal Gravitation and obtain the first integral of their motion. Denote the particle masses by $m_0, m_1, \ldots, m_{n-1}$ and their radius-vectors with respect to an origin O_a in inertial space by $\mathbf{r}_0, \ldots, \mathbf{r}_{n-1}$. Then the force $\mathbf{F}_i$ exerted on the i-th particle by the j-th one is determined by the expression

$$\mathbf{F}_i = k^2 m_i m_j \frac{\mathbf{r}_j - \mathbf{r}_i}{r_{ij}^3}.$$

Therefore, we may write

$$m_i \frac{d^2\mathbf{r}_i}{dt^2} = k^2 \sum_0^{n-1(i)} m_i m_j \frac{\mathbf{r}_j - \mathbf{r}_i}{r_{ij}^3}, \quad (i = 1, \ldots, n-1). \tag{3.1}$$

Here $r_{ij} = |\mathbf{r}_j - \mathbf{r}_i|$; $\sum_0^{n-1(i)}$ is the sum from $j = 0$ to $j = n-1$, except for the term with $j = i$. We are able to rewrite equation (3.1) in

scalar form through making projections onto the axes of the inertial reference system $O_a\xi\eta\zeta$

$$\begin{aligned} m_i\frac{d^2\xi_i}{dt^2} &= k^2\sum_0^{n-1(i)} m_i m_j\frac{\xi_j-\xi_i}{r_{ij}^3}, \\ m_i\frac{d^2\eta_i}{dt^2} &= k^2\sum_0^{n-1(i)} m_i m_j\frac{\eta_j-\eta_i}{r_{ij}^3}, \\ m_i\frac{d^2\zeta_i}{dt^2} &= k^2\sum_0^{n-1(i)} m_i m_j\frac{\zeta_j-\zeta_i}{r_{ij}^3}. \end{aligned} \tag{3.2}$$

Introduce now a function

$$V = k^2 S\frac{m_i m_j}{r_{ij}}$$

where S represents a sum of terms with indices i and j but with $i=j$ absent and $i>j$. Then, (3.2) may be rewritten as

$$m_i\frac{d^2\xi_i}{dt^2} = \frac{\partial V}{\partial\xi_i}, \quad m_i\frac{d^2\eta_i}{dt^2} = \frac{\partial V}{\partial\eta_i}, \quad m_i\frac{d^2\zeta_i}{dt^2} = \frac{\partial V}{\partial\zeta_i}. \tag{3.3}$$

Multiplying equations (3.3) consequtively by $\frac{d\xi_i}{dt}$, $\frac{d\eta_i}{dt}$, $\frac{d\zeta_i}{dt}$ and then combining them we get

$$\sum_0^{n-1} m_i\left(\frac{d\xi_i}{dt}\frac{d^2\xi_i}{dt^2}+\frac{d\eta_i}{dt}\frac{d^2\eta_i}{dt^2}+\frac{d\zeta_i}{dt}\frac{d^2\zeta_i}{dt^2}\right) = \frac{dU}{dt}$$

or

$$T - V = h \tag{3.4}$$

where

$$T = \frac{1}{2}\sum_0^{n-1} m_i\left[\left(\frac{d\xi_i}{dt}\right)^2+\left(\frac{d\eta_i}{dt}\right)^2+\left(\frac{d\zeta_i}{dt}\right)^2\right]$$

is the kinetic energy of the system; V is a *force function* which coincides with *potential energy* in value but differs from it in sign, while h is a constant of the energy integral represented by (3.4). A further nine first integrals can be obtained using the following arguments. The system of n particles is acted on by internal forces only and, therefore, the resultant of these forces is equal to zero. A similar comment

can be made about the resultant torque of the internal forces. The relations

$$\sum_0^{n-1} m_i \frac{d^2\mathbf{r}_i}{dt^2} = 0, \qquad \sum_0^{n-1} \mathbf{r}_i \times m_i \frac{d^2\mathbf{r}_i}{dt^2} = 0$$

follow from (3.1). Integrating these equations we get

$$\sum_0^{n-1} m_i \frac{d\mathbf{r}_i}{dt} = \mathbf{a}, \quad \sum_0^{n-1} m_i \mathbf{r}_i = \mathbf{a}t + \mathbf{b}, \quad \sum_0^{n-1} m_i \left(\mathbf{r}_i \times \frac{d\mathbf{r}_i}{dt}\right) = \mathbf{c}. \quad (3.5)$$

These expressions each contain three vector constants so that nine scalar first integrals are obtained. It is easy to interpret them in terms of mechanics. The first two relationships imply that the center of mass of the system moves rectilinearly and uniformly. The third relationship indicates that the angular momentum of the system is conserved.

Further the third integral (3.5) allows us to specify a convenient reference system $O\xi'\eta'\zeta'$ with its origin situated at the center of mass of the system O with its axis $O\eta'$ directed along vector **c.** A plane $O\xi'\zeta'$ is thereby fixed in space with respect to the inertial reference system that moves with the system's center of mass. This plane is usually called *Laplace's plane.* In the two-body problem it coincides with that plane containing the trajectory of particle motion.

Thus, ten first integrals, including the integral of energy, are obtained. Generally, these are referred to as *classical.* Numerous attempts to find other first integrals pertaining to the n-body problem have been made but without success. In this connection, in the XIX-th century Bruns* (1887), Poincaré† (1889), and Painlevé‡ (1898) proved that the equations of motion (3.1) have neither other independent algebraic first integrals, or first integrals expressed in terms of unique transcendental functions if $n > 2$.

*Bruns, Henry Ernst (1848–1919), German astronomer, geodesist and mathematician, Member of the Berlin Academy of Science.

†Poincaré, Jules Henri (1854–1912), French mathematician, mechanician, theoretical astronomer, and philosopher of science. He introduced the concept of automorphic functions, made a contribution to the theory of orbits, developed new mathematical models in astronomy and methods for the investigation of periodic motions.

‡Painlevé, Paul (1863–1933), French mathematician. His main work was devoted to differential equations and to the theory of functions of complex variables.

First integrals allow us to reduce the order of a system of equations. The classical first integrals allow us to reduce the order of system (3.1) by 10. Moreover, it is possible to reduce this order by a further 2, using the fact that the internal forces acting do not depend on the coordinates of the particles but only on their mutual separations. In order to achieve this, we can eliminate time t from the system by introducing a new independent variable φ and adding a quadrature

$$\varphi = \int \frac{d\varphi}{dt} dt.$$

One of the n–body problem variables can be used as variable φ. We may then divide the equations in (3.2) by, for instance, the last equation from (3.2), thereby reducing by 12 the order of the equations describing the n-body problem. If we consider the Solar System, which is composed of 10 large bodies, then the order of the equations is equal to 60 (because the number of two-order vector equations of motion is equal to 10 and each radius-vector has dimension 3). The order can thereby be reduced to 48. However, in practice, to study a system of 60 simple-form equations is easier than to deal with 48 non-symmetrical equations.

3.2 Planetocentric Form of the Relative Motion Equations

Let the relative motion of a system of particles be studied in the framework of the n-body problem and let the mass of one particle be much greater than that of the others. We have here an analog of the motion of the planets around the Sun.

Let the particle with mass m_0 be the largest one. Impose on this particle the origin of a reference system O_0xyz with axes which are parallel to the corresponding axes of the reference system $O_a\xi\eta\zeta$ and obtain the formulae

$$\xi_i = \xi_0 + x_i, \quad \eta_i = \eta_0 + x_i, \quad \zeta_i = \zeta_0 + z_i \tag{3.6}$$

of coordinate transformation from ξ_i, η_i, ζ_i. Equations (3.2) after di-

viding by m_i have the form

$$\frac{d^2\xi_i}{dt^2} = k^2 \sum_0^{n-1^{(i)}} m_j \frac{x_j - x_i}{r_{ij}^3} = -k^2 m_0 \frac{x_i}{r_{0i}^3} + k^2 \sum_1^{n-1^{(i)}} m_j \frac{x_j - x_i}{r_{ij}^3},$$
$$\frac{d^2\xi_0}{dt^2} = k^2 \sum_0^{n-1^{(0)}} m_j \frac{x_j - 0}{r_{0j}^3} = k^2 m_i \frac{x_i}{r_{0i}^3} + k^2 \sum_1^{n-1^{(i)}} m_j \frac{x_j}{r_{0j}^3}$$

where these representative equations are written for coordinate x only. In the second equation the i-th term $(i \neq 0)$ is highlighted by extracting it from the total sum and writing it individually. This term is indicated by superscript index 'i' in the sum notation $\sum_1^{n-1^{(i)}}$. For brevity we assume that $r_{0i} = r_i$. Subtracting now the second equation from the first one we get

$$\frac{d^2x_i}{dt^2} = -k^2(m_0 + m_i)\frac{x_i}{r_i^3} + k^2 \sum_1^{n-1^{(i)}} m_j \left(\frac{x_j - x_i}{r_{ij}^3} - \frac{x_j}{r_j^3} \right).$$

Let us now introduce the *perturbation function* $R = \sum R_i$, where

$$R_i = k^2 \sum_1^{n-1^{(i)}} m_j \left(\frac{1}{r_{ij}} - \frac{x_i x_j + y_i y_j + z_i z_j}{r_j^3} \right).$$

In this case, the equations of motion can be written in what is called *planetocentric form*

$$\begin{aligned} \frac{d^2x_i}{dt^2} + k^2(m_0 + m_i)\frac{x_i}{r_i^3} &= \frac{\partial R_i}{\partial x_i}, \\ \frac{d^2y_i}{dt^2} + k^2(m_0 + m_i)\frac{y_i}{r_i^3} &= \frac{\partial R_i}{\partial y_i}, \\ \frac{d^2z_i}{dt^2} + k^2(m_0 + m_i)\frac{z_i}{r_i^3} &= \frac{\partial R_i}{\partial z_i}, \ (i = \overline{1, n-1}). \end{aligned} \tag{3.7}$$

If we assume $R_i = 0$, then system (3.7) decomposes into $3(n-1)$ independent equations, where each equation has a form similar to the equation of motion in the two-body problem. The mechanical interpretation of this result is as follows. Equality $R_i = 0$ means that there is no gravitational attraction acting between the planets orbiting the Sun. Thus, the n-body problem decomposes into $(n-1)$

separated two-body problems. The inequality $R_i \neq 0$ is valid but, for the Solar System, the derivatives of R_i with respect to the coordinates are small. Therefore, the motion of the planets around the Sun can, on the whole, be correctly described in the framework of the two-body problem, and the effect of mutual planetary attraction can be considered to comprise a weak perturbation. The weakness of R_i is a result of the large inequality $m_0 \gg m_i$.

3.3 The Three-Body Problem

As the n-body problem is not in general solved, it is natural to investigate the first case for which $n > 2$ (namely $n = 3$). In 1913[§] Sundman[¶] obtained a complete solution of the three-body problem in the form of a convergent series. However, its rate of convergence is so slow that it is not possible to use this series in practical calculations. Already in 1933 estimations showed that to carry out calculations with a precision acceptable for astronomical year-books, about $10^{8\,000\,000}$ terms would have to be employed. Thus, Sundman's solution is of theoretical rather than of practical interest. More easily visualized results can be obtained through imposing additional restrictions on the initial motion and on the relationships between particle masses. The restrictions on the initial motion lead us to *Lagrange's case* and *Euler's case.*

3.3.1 The Lagrange and Euler Cases

Euler and Lagrange addressed the problem as to whether initial conditions of motion exist in respect of three mutually attracting bodies with arbitrary masses which move along Keplerian orbits.

Let the three bodies have finite masses m_0, m_1, m_2. Radius-vectors $\mathbf{r}_0, \mathbf{r}_1, \mathbf{r}_2$ join these particles (Fig.3.1). The relationship

$$\mathbf{r}_0 + \mathbf{r}_1 + \mathbf{r}_2 = 0 \tag{3.8}$$

[§]K.F.Sundman. Mémoire sur le problèm des trois corps. *Acta Math.,* 1913, Vol.36, pp.105–179.

[¶]Sundman, Karl Fritiof (1873–1949), astronomer and mathematician of Swedish origin. Longtime Professor of Astronomy at the University of Helsinki. Studied collisions between particles.

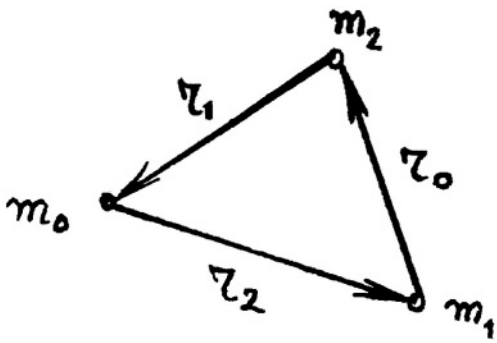

Figure 3.1. Three-Body Problem. Arrangement of particles and radius-vectors

is evident.

The equations of relative motion in planetocentric form for particles with masses m_1 and m_2 with respect to m_0 and of m_2 with respect to m_1 are, respectively

$$\begin{aligned}
\frac{d^2\mathbf{r}_2}{dt^2} &= -k^2(m_0+m_1)\frac{\mathbf{r}_2}{r_2^3} + k^2 m_2\left(\frac{-\mathbf{r}_1-\mathbf{r}_2}{r_0^3} - \frac{-\mathbf{r}_1}{r_1^3}\right) = \\
&= -k^2(m_0+m_1)\frac{\mathbf{r}_2}{r_2^3} + k^2 m_2\left(\frac{\mathbf{r}_0}{r_0^3} + \frac{\mathbf{r}_1}{r_1^3}\right), \\
\frac{d^2\mathbf{r}_1}{dt^2} &= -k^2(m_0+m_2)\frac{\mathbf{r}_1}{r_1^3} + k^2 m_1\left(\frac{-\mathbf{r}_2-\mathbf{r}_1}{r_0^3} + \frac{\mathbf{r}_2}{r_2^3}\right) = \\
&= -k^2(m_0+m_2)\frac{\mathbf{r}_1}{r_1^3} + k^2 m_2\left(\frac{\mathbf{r}_0}{r_0^3} + \frac{\mathbf{r}_2}{r_2^3}\right), \qquad (3.9) \\
\frac{d^2\mathbf{r}_0}{dt^2} &= -k^2(m_1+m_2)\frac{\mathbf{r}_0}{r_0^3} + k^2 m_0\left(\frac{-\mathbf{r}_2-\mathbf{r}_0}{r_1^3} - \frac{-\mathbf{r}_2}{r_2^3}\right) = \\
&= -k^2(m_1+m_2)\frac{\mathbf{r}_0}{r_0^3} + k^2 m_0\left(\frac{\mathbf{r}_1}{r_1^3} + \frac{\mathbf{r}_2}{r_2^3}\right).
\end{aligned}$$

In order that the motion be Keplerian, it has to satisfy the following equations

$$\frac{d^2\mathbf{r}_i}{dt^2} = -k^2 h_i \frac{\mathbf{r}_i}{r_i^3} \qquad (i=0,1,2) \qquad (3.10)$$

where h_i are unknown constants. In other words, the three identities

$$\begin{aligned}
(m_0+m_1-h_2)\frac{\mathbf{r}_2}{r_2^3}-m_2\left(\frac{\mathbf{r}_0}{r_0^3}+\frac{\mathbf{r}_1}{r_1^3}\right)&=0,\\
(m_0+m_1-h_1)\frac{\mathbf{r}_1}{r_1^3}-m_1\left(\frac{\mathbf{r}_0}{r_0^3}+\frac{\mathbf{r}_2}{r_2^3}\right)&=0,\\
(m_1+m_2-h_0)\frac{\mathbf{r}_0}{r_0^3}-m_0\left(\frac{\mathbf{r}_1}{r_1^3}+\frac{\mathbf{r}_2}{r_2^3}\right)&=0
\end{aligned}$$

must be valid. With regard to $\mathbf{r}_0+\mathbf{r}_1+\mathbf{r}_2=0$ we get the equalities

$$\begin{aligned}
\mathbf{r}_0\left(\frac{m_2}{r_0^3}+\frac{m_0+m_1-h_2}{r_2^3}\right)+\mathbf{r}_1\left(\frac{m_2}{r_1^3}+\frac{m_0+m_1-h_2}{r_2^3}\right)&=0,\\
\mathbf{r}_1\left(\frac{m_0}{r_1^3}+\frac{m_1+m_2-h_0}{r_0^3}\right)+\mathbf{r}_2\left(\frac{m_0}{r_2^3}+\frac{m_1+m_2-h_0}{r_0^3}\right)&=0,\\
\mathbf{r}_2\left(\frac{m_1}{r_2^3}+\frac{m_2+m_0-h_1}{r_1^3}\right)+\mathbf{r}_0\left(\frac{m_1}{r_0^3}+\frac{m_2+m_0-h_1}{r_1^3}\right)&=0
\end{aligned}\qquad (3.11)$$

which are relevant to the following cases.

- If the expressions contained in parenthesis are equal to zero.
- If the magnitudes contained in parenthesis are not equal to zero but the vectors $\mathbf{r}_i$ are collinear.

Lagrange's Case

In the first case, the condition provides three pairs of equations

$$\begin{aligned}
\frac{m_0+m_1-h_2}{r_2^3}&=-\frac{m_2}{r_0^3}=-\frac{m_2}{r_1^3},\\
\frac{m_1+m_2-h_0}{r_0^3}&=-\frac{m_0}{r_1^3}=-\frac{m_0}{r_2^3},\\
\frac{m_2+m_0-h_1}{r_2^3}&=-\frac{m_1}{r_2^3}=-\frac{m_1}{r_0^3}.
\end{aligned}$$

From these, we obtain the equalities $r_0=r_1=r_2$ and, consequently, $h_0=h_1=h_2=m_0+m_1+m_2$. Let us now differentiate the equality (3.8) with respect to time

$$\frac{d\mathbf{r}_0}{dt}+\frac{d\mathbf{r}_1}{dt}+\frac{d\mathbf{r}_2}{dt}=0.$$

It is clear that if, at an arbitrary time, the vectors of particle velocities compose an equilateral triangle (Fig.3.1) then, at the next moment of time, the equalities $r_1 = r_2 = r_0$ are satisfied again (i.e. at any time this equality is true). Thus, (see equations (3.9), (3.8)) the condition that the motion of each of the three bodies be Keplerian is always fulfilled.

Since three particles can be situated at the vertices of an equilateral triangle in two ways (m_0, m_1, m_2, m_0) and (m_0, m_2, m_1, m_0), two solutions of the formulated problem exist. Both of these solutions and their corresponding mass configurations were obtained by Lagrange in 1772[‖] while studying Eulerian solutions and they are called *Lagrangian.*

Euler's Case

In the second case, three particles m_0, m_1, m_2 must always be situated in a straight line. This is possible if their mutual separations and initial velocities are appropriately chosen. There are three particle sets with the following sequences

$$m_0\, m_1\, m_2; \qquad m_0\, m_2\, m_1; \qquad m_1\, m_0\, m_2.$$

These particle configurations and corresponding solutions were obtained by Euler in 1767[**] and now are called *Eulerian.*

In 1951 K.Stumpff[††] showed that, for three-body motion, there are no initial conditions which produce Keplerian trajectories other then those described above.

3.4 Restricted Three-Body Problem

The next simplification is to assume that the mass of one body is negligible with respect to the others. In this case two 'large' bodies

[‖] J.L.Lagrange. *Oeuvres.* Paris, 1873, Vol.6, pp.272–292.

[**] L.Euler. De moto rectilineo trium corporumse mutio attrahentium. *Novo Comm. Acad. Sci. Imp. Petrop.,* 1767, Vol.11, pp.144–151.

[††] Stumpff, Karl (1895–1970), German by birth, he was appointed a Professor of Astronomy at Graz and later at G'ottingen. His first book 'Foundations of Periodic Research' (1937) anticipates by nearly thirty years part of Colley and Tukey's work on 'Fast Fourier Transforms'. Also, he analyzed the orbits of SC in the gravitational fields of the Earth, Moon and Sun.

are not affected by a 'small' one. Therefore, the relative motion of large bodies is completely described within the framework of the two-body problem. The third small body moves along a non-Keplerian orbit and follows a complex trajectory under the effect of the two attracting bodies. This specific case of the three-body problem is named the *restricted three-body problem.*

This method was initially used by astronomers to study the motion of the Sun–Earth–Moon system. More recently it has been applied in the development and study of interplanetary trajectories because SC motion in the vicinity of a planet may be considered to fit the framework of a three-body problem involving the Sun, a planet and a SC.

3.4.1 Hill's Surface

Let bodies *S*, *J*, *P* have the respective masses $m_1 \geq m_2 \gg m$ and suppose the two large bodies *S* and *J* to move along circular orbits. This specific case is called the *circular restricted three-body problem.* It is convenient to assume that *m* is negligible i.e. tends to zero. Locate the origin of a reference system $Oxyz$ at the center of mass of the two bodies *S* and *J*. Situate the plane Oxy in the plane of motion of the two bodies. The axis Ox is aligned with a vector from *S* to *J* (Fig.3.2). Let d_1 and d_2 be the distances of the bodies *S* and *J*

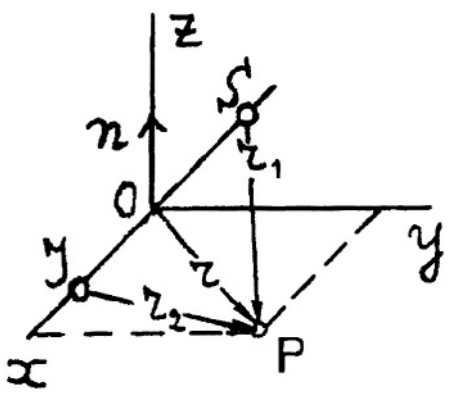

Figure 3.2. Circular restricted three-body problem. Reference system and notations

from the origin of the reference system so that their coordinates in this

system are $(-d_1, 0, 0)$, $(d_2, 0, 0)$ and $R = d_1 + d_2$. Let n be the angular velocity of the axis Ox with respect to inertial space, and align axis Oz along the vector of angular velocity so that n is a positive number. An axis Oy with the two axes defined above completes the Cartesian reference system. According to Kepler's law (1.25), and taking into consideration the relation $n = 2\pi/T_{rev}$, we obtain the expressions

$$n^2(d_1 + d_2)^3 = \mu, \qquad \mu = k^2(m_1 + m_2). \tag{3.12}$$

Consider the motion of a particle P using (x, y, z) coordinates. The absolute angular velocity components of this particle are determined by the formulae $(\frac{dx}{dt} - ny, \frac{dy}{dt} + nx, \frac{dz}{dt})$. The kinetic energy of the particle is

$$T = \frac{m}{2}\left[\left(\frac{dx}{dt} - ny\right)^2 + \left(\frac{dy}{dt} + nx\right)^2 + \left(\frac{dz}{dt}\right)^2\right].$$

Write Lagrange's equation of motion

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}} - \frac{\partial T}{\partial q} = \frac{\partial V}{\partial q},$$

denoting differentiation with respect to time t by a dot $(\dot{})$ and a generalized coordinate by q. The force function V now has the form

$$V = k^2\left(\frac{mm_1}{r_1} + \frac{mm_2}{r_2} + \frac{m_1m_2}{R}\right).$$

The last term, which is a constant, may be dropped. Let us introduce *potential*

$$U = k^2\left(\frac{m_1}{r_1} + \frac{m_2}{r_2}\right)$$

and write Lagrange's equations in the form

$$\begin{aligned} \ddot{x} - 2n\dot{y} - n^2x &= U_x, \\ \ddot{y} + 2n\dot{x} - n^2y &= U_y, \\ \ddot{z} &= U_z. \end{aligned}$$

Introduce now a function

$$\Omega = \frac{1}{2}n^2(x^2 + y^2) + k^2\left(\frac{m_1}{r_1} + \frac{m_2}{r_2}\right). \tag{3.13}$$

Then the equations of motion of the particle P are given by

$$\ddot{x} - 2n\dot{y} = \Omega_x,$$
$$\ddot{y} + 2n\dot{x} = \Omega_y, \quad \ddot{z} = \Omega_z.$$

Multiplying these equations by $2\dot{x}$, $2\dot{y}$, $2\dot{z}$ respectively, combining them and integrating, we get *Jacobi's integral* *

$$(\dot{x})^2 + (\dot{y})^2 + (\dot{z})^2 = 2\Omega - C', \tag{3.14}$$

where C' is a constant. This integral is reminiscent of an energy integral but it should not be forgotten that on the left-hand side appears the relative velocity of particle P rather than its absolute value. The first integral of motion allows us to carry out a qualitative analysis of the motion of particle P.

Surface $(\dot{x})^2 + (\dot{y})^2 + (\dot{z})^2 = 0$ separates areas in which the relative velocity is, respectively, real and imaginary. From (3.14) it follows that the equation of this surface is

$$2\Omega = C'. \tag{3.15}$$

To simplify the analysis we now select an appropriate system of units. Let this be chosen so that $k = 1$ and $d_1 + d_2 = 1$. Then, from equality (3.12), the expression

$$n^2 = m_1 + m_2$$

follows. Equation (3.15), having regard to (3.13), can now be written in the form

$$(m_1 + m_2)(x^2 + y^2) + \frac{2m_1}{r_1} + \frac{2m_2}{r_2} = C'. \tag{3.16}$$

Equation (3.16) defines a *zero-velocity surface* located in the reference system $Oxyz$ rotating with bodies S and J. Sometimes a zero-velocity surface is referred to as a zero-relative velocity surface in order to emphasize that a relative velocity is considered.

* Jacobi, Karl Gustav Jacob (1804-1851), German mathematician. Professor at Königsberg who, with Niels Henrik Abel, established the theory of elliptic functions. He invented the functional determinant that bears his name. Also, he studied partial differential equations of the first order and their applications in dynamics and in celestial mechanics.

Comparison of equations (3.14) and (3.13) shows that a decrease in the constant C' at a given point (in coordinates (x, y, z)) corresponds to an increase in the relative velocity magnitude. We may use this fact to analyze the variation of the zero-velocity surface configuration in accordance with the variation of the relative velocity magnitude of particle *P*. We begin this analysis for the case of slow relative velocities (which corresponds to large C').

If C' is large, then one can see from equation (3.16) that this is true in three areas of space (x, y, z) where, respectively, large values of $x^2 + y^2$ (large r_1 and r_2 simultaneously); small r_1 or small r_2 occur. It is convenient for these three cases to rewrite equation (3.16) in three different forms

$$\begin{aligned} (m_1 + m_2)(x^2 + y^2) &= C' - \varepsilon_1(x, y, z); \\ 2\frac{m_1}{r_1} &= C' - \varepsilon_2(x, y, z); \\ 2\frac{m_2}{r_2} &= C' - \varepsilon_3(x, y, z) \end{aligned} \tag{3.17}$$

where ε_1, ε_2, ε_3 are functions of (x,y,z) with relatively small magnitudes.

Putting functions ε_i equal to zero, the first equation from (3.17) is the equation of a circular cylinder with axis coincident with the axis Oz. The second and third equations in (3.17) are the equations of spheres with their respective centers situated at *S* and *J*. Thus, the zero-velocity surface (3.16) disintegrates into three separate surfaces. Since the terms ε_i are small they slightly deform these surfaces which, in consequence, are respectively described as constituting a quasicylinder and quasispheres. The quasicylinder has its least diameter in the plane (x, y) and the quasispheres have an egg-shaped form, with the sharp end directed towards the origin of the reference system.

In response to an increase in relative velocity (that is of a reduction in C'), the quasicylinder diameter decreases and the sizes of the quasispheres grow. Thus, originally separated surfaces begin to approach each other and to merge. This process is shown in Fig.3.3 where the evolution of the sections of zero relative velocity surface (3.16) which occur in planes Oxy and Oxz, are represented. Drawings from left to right correspond to successive reductions in C'. Areas, where motion is impossible are hatched.

To explain the mechanical sense of what is represented in Fig.3.3, it is recalled that, in accordance with a reduction in C', there is an

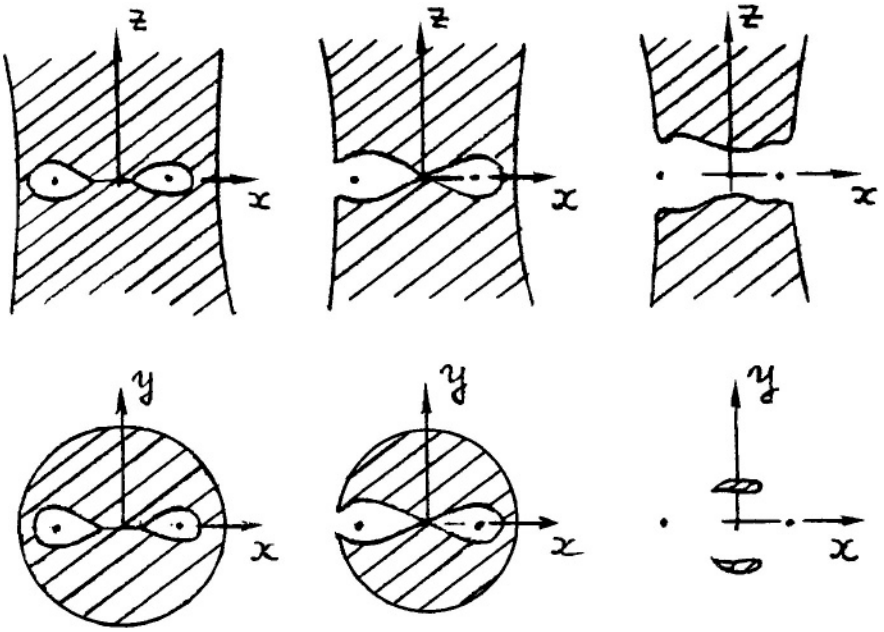

Figure 3.3. Evolution of the zero relative velocity surface (3.16) cross sections in planes Oxy and Oxz

increase in the relative velocity of particle P and the area of its accessibility increases. At very small relative velocities (left diagram), the particle P remains close to the powerful attractive centers S and J. An increase in the relative velocity of P is accompanied by an increase in its area of accessibility so that, for example, in the right-hand diagram the particle 'starting' from both S and J can go to infinity.

This method was originally used by Hill† in 1878 to analyze the motion of the Moon with respect to the Sun and the Earth, assuming the mass of the Moon to be negligible relative to the other bodies. Hill showed that the two left diagrams in Fig.3.3 correspond to the Moon moving deeply inside the quasisphere centered at the Earth. It can then be said, following Hill, that the motion of the Moon relative to the Earth is stable.

Jacobi's integral has sometimes been used to identify periodic co-

†Hill, George William (1838–1914), American astronomer. In 1877–1878 he published two papers on lunar motion in which the theory of linear differential equations with periodic coefficients was established.

mets. Assuming that the Solar System is composed of the Sun, Jupiter and a new comet, the constant C' is obtained by observation. At the next appearance of this comet a necessary condition of its identity with the earlier apparition, is the closeness of the magnitude of C' to the previous value.

3.4.2 Singular Points of a Zero-Velocity Surface

As is well-known, the singular point of a surface $F(x, y, z)$ should satisfy the conditions $F_x = 0$, $F_y = 0$, $F_z = 0$. Consider now the properties of the singular points of surface (3.15). On this surface the equality $(\dot{x})^2 + (\dot{y})^2 + (\dot{z})^2 = 0$ is valid by definition and, consequently, $\dot{x} = \dot{y} = \dot{z} = 0$. At the singular points, the equalities $\Omega_x = 0$, $\Omega_y = 0$, $\Omega_z = 0$ are additionally satisfied, and from (3.14) we obtain equalities

$$\ddot{x} - 2n\dot{y} = \ddot{y} + 2n\dot{x} = \ddot{z} = 0.$$

Having regard to $\dot{y} = \dot{x} = 0$,

$$\ddot{x} = \ddot{y} = \ddot{z} = 0.$$

Hence, the particle P if it is situated at the singular point of surface (3.15) and characterized by a constant C' of appropriate magnitude, is not only motionless in the reference system $Oxyz$ but, also, its acceleration in this system is equal to zero. Thus, they are points where particle P is, relatively, in an equilibrium position. Under the conditions formulated above, particle P can remain forever at such points.

A search for equilibrium points can be carried out using standard methods of analysis associated with lengthy calculations. However, we choose here another way to obtain them based on already derived results. Since in the singular points of surface (3.15) particle P is motionless with respect to the bodies S and J, and since also the distance between bodies S and J is constant, all three rotate around their common center of mass in a plane Oxy. For bodies S and J this follows from the definition of plane Oxy. For particle P it follows from the condition $\ddot{z} = 0$ that otherwise non-compensated acceleration appears along the axis Oz. Also, as the orbits of S and J are circular

and the orbit of particle P is circular, all three orbits are Keplerian. It was already shown in Section 3.3.1 that when three bodies with arbitrary masses move along Keplerian orbits, only two cases corresponding to mutually motionless configurations exist (Lagrange and Euler's cases). Thus the singular points of surface (3.15) belong to those configurations described by Lagrange[‡] and Euler[§].

If particle P is in the vicinity of a singular point of the zero-velocity surface and if it has a small relative velocity, then its later behaviour depends on the character of the acceleration field $\ddot{x}, \ddot{y}, \ddot{z}$. It may be such that acceleration will 'withdraw' particle P from the singular point or, conversely, it will bring P towards the singular point. Thus, in the first case the equilibrium position is unstable and in the second case it is a steady-state position. In the case of stability, a deviation of P from an equilibrium position (due to the presence of small relative velocity and/or small coordinate deviation) results, owing to the absence of mechanical energy dissipation, in non-damped librations of P in the vicinity of the singular point. These singular points are usually called *points of libration.*

Denote the coordinates of a libration point in a rotating reference system $Oxyz$ by x_0, y_0, z_0 and the small deviations of P from this libration point by ξ, η, ζ. Then, on applying usual linearization techniques, it is possible to formulate and to solve the following problem concerning the stability of the motion of P in the vicinity of the libration point [16]. Motion of particle P in the plane Oxy is here our

[‡] Lagrange, Joseph-Louis (1736–1813). Born in Italy, worked at the Berlin Academy but eventually settled in Paris where he was a Professor at the École Polytechnique. He introduced key concepts and developed innovative methods in mathematical analysis, differential equations, variational computation and number theory. He mathematically formalized mechanics, reducing statics to the *virtual displacement principle* and reducing dynamics to this principle combined with the *principle of D'Alembert.* Also, he proved the theorem of equilibrium stability now called the *Lagrange–Dirichlet's theorem* and derived equations of motion using generalized variables (*Lagrange's equations of the second kind*).

[§] Euler, Leonard (1707–1783). Eminent mathematician, mechanician, physicist and astronomer. Of Swiss origin, he worked primarily at the St.Petersburg Academy of Sciences and at the Berlin Academy. He made outstanding contributions inter alia to geometry, calculus, analysis, celestial mechanics, mathematical physics, optics, ship-building and the theory of music. He was author of more than 850 publications, and his influence on mathematics and theoretical physics has been immense.

main interest. The corresponding linear equations of motion have the form

$$\ddot{\xi} + \ldots = 0, \quad \ddot{\eta} + \ldots = 0$$

and, hence, the characteristic equation of the system is of the fourth order

$$\lambda^4 + a\lambda^2 + b = 0$$

where factors a and b depend on the ratio of the masses m_1 and m_2. Analysis of this equation shows that collinear points of libration L_1, L_2, L_3 (Euler's case) are always unstable but triagonal points of libration L_4, L_5 (Lagrange's case) are stable when

$$\frac{m_2}{m_1 + m_2} < 0.038, \qquad \text{or} \qquad \frac{m_2}{m_1 + m_2} > 1 - 0.038. \tag{3.18}$$

The stable points of libration are remarkable in that, on being placed in their vicinity, a particle will remain always situated there if the corresponding constant C' is of an appropriate magnitude. No energy is required to maintain this position. For example, these properties can be used in designing a scheme to transmit signals from the invisible side of the Moon (A) to the Earth (C), (see Fig.3.4). Here

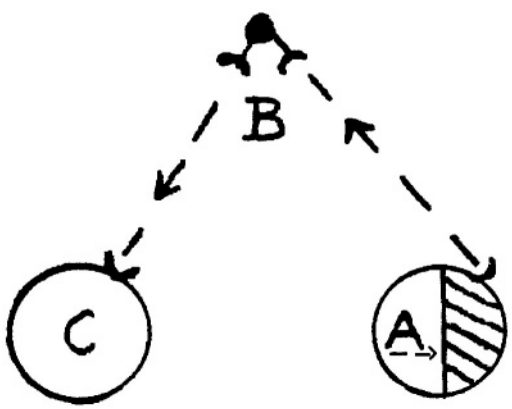

Figure 3.4. Configuration to transmit signals from the invisible side of the Moon (A) to the Earth (C)

since the re-transmission point (B) is fixed relative to bodies *A* and *C*, the antennae at *A* and *B* can maintain permanent directions, thus obviating the need for tracking. The stability of configuration (*A*, *B*, *C*) follows from (3.18) since the magnitude $m_2/(m_1 + m_2)$ is ≈ 0.01.

Stable points of libration also act as ‘debris baskets’ in space where dust and other fine colliding particles with a suitable value of the constant C' collect.

Chapter 4

Gravispheres

In a number of applications the n-body problem can be broken down into $(n-1)$ two-body problems if one of the bodies is very large and the others are very small. We will illustrate this situation using the example of an interplanetary mission. In this connection, let the trajectory of a SC starting from the Earth pass planet Mars. Assume that the SC does not affect the motion of the planets around the Sun and that motion is determined at the beginning of the flight by the system pair (SC–Earth); thereafter by the pair (SC–Sun) and near the planet by the pair (SC–Mars). If the mission proceeds further we have again the pair (SC–Sun) etc.

The actual trajectory may be represented by an approximate one composed of pieces of conic sections, constructed by deriving solutions of individual two-body problems in the sequence indicated above. The procedure for joining the segments is as follows. The parameters of a terminal point of the trajectory created in the system (SC–Earth) should coincide with the parameters of the initial point of a segment created in the system (SC–Sun) and so on. The parameters concerned are the positions of the SC and the vectors of its absolute velocity. This approach makes it necessary to determine the joining points of the segments introduced above in order to compose the required trajectory. This approximate method of modeling interplanetary trajectories by joining up pieces of conic sections is referred to either as the *method of conjugated conic sections* or the *patched conic method.*

Suppose that each planet is surrounded by a hypothetical, non-geometric, sphere separating area of space, in which the gravitational

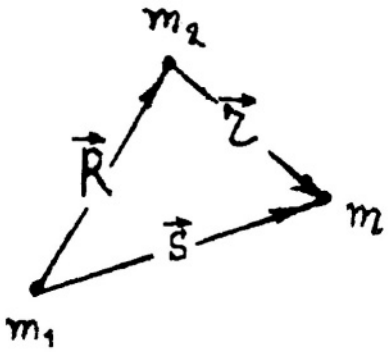

Figure 4.1. Mutual arrangement of the Sun (m_1), planet (m_2), and SC (m)

fields of the planet and the Sun are, respectively, dominant. There are several different approaches to constructing these gravitational spheres which are called *gravispheres.* We present below three of them namely, the *sphere of attraction,* the *sphere of influence* and *Kislik's sphere.*

For all three, we consider that the Sun's mass is m_1, that of the planet is m_2 and the SC's mass is *m*, subject to the condition $m_1 \gg m_2$ while mass *m* is negligible with respect to both m_1 and m_2.

4.1 Sphere of Attraction

The most natural definition of a gravisphere is that surface where the values of the attracting forces on a SC exerted by the Sun and by a particular planet are equal each other. This criterion is used in defining the *sphere of attraction.* Determine now the radius of this sphere.

Let **R** and **s** be the respective heliocentric radius-vectors of m_2 and *m* and let **r** be the geocentric radius-vector of *m* (Fig.4.1), By virtue of the condition $m_1 \gg m_2$ the border of the sphere of attraction of a planet lies at a small distance from it and, hence, $R \approx s$. Let the values of the attracting forces of the Sun and of the planet acting on the SC *m* be equal to each other,

$$\frac{k^2 m m_1}{s^2} = \frac{k^2 m m_2}{r^2}.$$

Then, after reduction, we obtain

$$m_1 r^2 = m_2 s^2 \approx m_2 R^2$$

and

$$r = R\sqrt{\frac{m_2}{m_1}}. \tag{4.1}$$

Expression (4.1) for the radius of the sphere of attraction shows that the inequality $r \ll R$ is valid (i.e. the border of the sphere of attraction lies near the planet).

Although the above considerations appear to be very natural they contain a gross blunder. This proceeds from the inherent assumption that m_1 and m_2 are mutually motionless. Actually, not only m is under the influence of gravitational fields but so also are m_1 and m_2 which are mutually attracted. Therefore, the problem of gravispheres should be handled using a dynamic rather than a static approach.

4.2 Sphere of Influence (Laplace's Gravisphere)

Laplace considered the problem of the motion of a small body m in the gravitational fields of bodies m_1 and m_2 and solved it through numerical integration of the equations of motion. This raised the question of choosing a reference system in which such integration requires the least effort. Suppose that in any point of space a body m is under the gravitational attraction of two bodies m_1 and m_2. We can then choose either a *heliocentric* or a *planetocentric* reference system. We designate here values referred to in the heliocentric system by index *1* and in the planetocentric system by index *2*. The total acceleration of m can be represented by either $\mathbf{B}_1 + \mathbf{b}_1$ or $\mathbf{B}_2 + \mathbf{b}_2$. Here, $\mathbf{B}_j$ denotes the 'main' acceleration developed by a body which is at the origin of the j-th reference system (j= 1,2). Similarly, $\mathbf{b}_j$ denotes acceleration developed by another body which is considered to be the perturbing one. It is obvious that the smaller the perturbing acceleration and, hence, the closer the problem considered is to the two-body problem, the more convenient it is to carry out the calculations. Hence, between two possible reference systems, it is necessary to choose that in which the perturbing acceleration is the lesser.

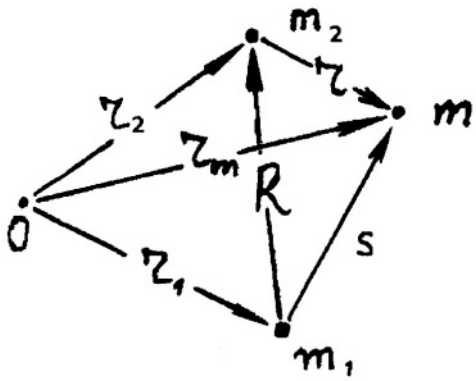

Figure 4.2. Mutual arrangement of the Sun (m_1), planet (m_2) and a body (m) with respect to a point O

In practice, as the body m moves, it can go from an area where use of the heliocentric reference system is preferred to an area where it is better to employ the planetocentric system (or visa versa). Laplace suggested that the *relative perturbing acceleration* (i.e. the ratio of the perturbing to the main acceleration) be employed as a 'key' parameter and, thence, defined a border between the two areas mentioned above using the equality

$$\frac{|\mathbf{b}_1|}{|\mathbf{B}_1|} = \frac{|\mathbf{b}_2|}{|\mathbf{B}_2|}. \tag{4.2}$$

That surface surrounding a planet, on which condition (4.2) is satisfied is called the *sphere of influence.*

We now write down the equations of motion of the Sun (m_1), the planet (m_2) and a body (m) in a certain reference system with origin O (Fig.4.2)), taking into account that the mass of m is vanishingly small.

$$\begin{aligned} \frac{d^2\mathbf{r}_1}{dt^2} &= -k^2 m_2 \frac{\mathbf{r}_1 - \mathbf{r}_2}{R^3}, \\ \frac{d^2\mathbf{r}_2}{dt^2} &= -k^2 m_1 \frac{\mathbf{R}}{R^3}, \\ \frac{d^2\mathbf{r}_m}{dt^2} &= -k^2 \left[m_1 \frac{\mathbf{r}_m - \mathbf{r}_1}{s^3} + m_2 \frac{\mathbf{r}_m - \mathbf{r}_2}{r^3} \right]. \end{aligned} \tag{4.3}$$

Writing next the obvious geometrical equality

$$\mathbf{r}_2 = \mathbf{r}_1 + \mathbf{R}, \qquad \mathbf{r}_m = \mathbf{r}_1 + \mathbf{s} \qquad \mathbf{R} + \mathbf{r} = \mathbf{s}. \tag{4.4}$$

we can use these formulae to re-write (4.3) as follows

$$\frac{d^2\mathbf{r}_1}{dt^2} = -k^2 m_2 \frac{-\mathbf{R}}{R^3},$$

$$\frac{d^2\mathbf{r}_2}{dt^2} = -k^2 m_1 \frac{\mathbf{R}}{R^3},$$

$$\frac{d^2\mathbf{r}_m}{dt^2} = -k^2 \left[m_1 \frac{\mathbf{s}}{s^3} + m_2 \frac{\mathbf{r}}{r^3} \right].$$

Subtracting the first from the second and the third equations and using (4.4), we obtain

$$\frac{d^2\mathbf{R}}{dt^2} = -k^2 (m_1 + m_2) \frac{\mathbf{R}}{R^3},$$

$$\frac{d^2\mathbf{s}}{dt^2} = -k^2 \left[m_1 \frac{\mathbf{s}}{s^3} + m_2 \left(\frac{\mathbf{r}}{r^3} + \frac{\mathbf{R}}{R^3} \right) \right]. \tag{4.5}$$

The first equation is simple and implies that a planet (m_2) moves around the Sun (m_1) in a Keplerian orbit. The second equation describes motion of the body (m) with respect to the Sun under the influence of a perturbation. It thus can be said that both orbit in a heliocentric reference system. The first term contained in square brackets describes the 'main' part of the acceleration generated by the Sun. The second term represents the perturbation caused by the gravitational field of the planet.

To write down an expression for the motion of the body (m) in the planetocentric reference system, it is necessary to subtract the first from the second equation in (4.5) and this results in the expression

$$\frac{d^2\mathbf{r}}{dt^2} = -k^2 \left[m_2 \frac{\mathbf{r}}{r^3} + m_1 \left(\frac{\mathbf{s}}{s^3} - \frac{\mathbf{R}}{R^3} \right) \right]. \tag{4.6}$$

The first term contained in square brackets describes the acceleration generated by the planet and the second the perturbing acceleration generated by the Sun.

The second equation of (4.5) and equation (4.6) describe the same motion and are completely equivalent to each other. The choice as to which equation is better to use in a particular situation is consequent on the simplicity of the associated computations. To determine the radius of the sphere of influence, substitute the corresponding terms

from (4.5) and (4.6) in equation (4.2) to obtain

$$\frac{\left|m_2\left(\frac{\mathbf{r}}{r^3}+\frac{\mathbf{R}}{R^3}\right)\right|}{\left|m_1\frac{\mathbf{s}}{s^3}\right|}=\frac{\left|m_1\left(\frac{\mathbf{s}}{s^3}-\frac{\mathbf{R}}{R^3}\right)\right|}{\left|m_2\frac{\mathbf{r}}{r^3}\right|}. \tag{4.7}$$

To simplify the terms contained in this expression we introduce the following approximations, under the assumption that r/R is small

$$\left|\frac{\mathbf{s}}{s^3}\right|\approx\frac{1}{R^2}; \qquad \left|\frac{\mathbf{r}}{r^3}\right|\approx\frac{1}{r^2}; \tag{4.8}$$

$$\left|\frac{\mathbf{r}}{r^3}+\frac{\mathbf{R}}{R^3}\right|\approx\frac{1}{r^2}\left|\frac{\mathbf{r}}{r}+\left(\frac{\mathbf{r}}{R}\right)^2\frac{\mathbf{R}}{R}\right|\approx\frac{1}{r^2}. \tag{4.9}$$

Also,

$$\left|\frac{\mathbf{s}}{s^3}-\frac{\mathbf{R}}{R^3}\right|=\left|\frac{\mathbf{R}+\mathbf{r}}{|\mathbf{R}+\mathbf{r}|^3}-\frac{\mathbf{R}}{R^3}\right|. \tag{4.10}$$

Since

$$|\mathbf{R}+\mathbf{r}|^{-3}=\left[(\mathbf{R}+\mathbf{r})^2\right]^{-3/2}\approx\frac{1}{R^3}\left(1-3\frac{\mathbf{R}\mathbf{r}}{R^2}\right)$$

then, retaining only magnitudes with order up to r/R, we obtain an approximation for the term (4.10) contained in (4.7)

$$\left|\frac{\mathbf{s}}{s^3}-\frac{\mathbf{R}}{R^3}\right|\approx\left|\frac{\mathbf{R}}{R^3}\left(1-3\frac{\mathbf{R}\mathbf{r}}{R^2}\right)+\frac{\mathbf{r}}{R^3}-\frac{\mathbf{R}}{R^3}\right|=\frac{1}{R^3}\left|\mathbf{r}-3\frac{\mathbf{R}\mathbf{r}}{R^2}\mathbf{R}\right|. \tag{4.11}$$

Hence, we can now write a simplified form of equation (4.7) having regard to (4.8) and (4.11) and omitting the symbol of approximation

$$\left(\frac{m_2}{m_1}\right)^2\left(\frac{R}{r}\right)^5=\left|\frac{\mathbf{r}}{r}-3\frac{\mathbf{R}\mathbf{r}}{Rr}\frac{\mathbf{R}}{R}\right|.$$

If we designate the angle between vectors $\mathbf{R}$ and $\mathbf{r}$ by ψ then, from the last equality, we obtain an expression for the radius of the sphere of influence

$$r=R\left(\frac{m_2}{m_1}\right)^{2/5}\frac{1}{F(\psi)^{1/5}} \tag{4.12}$$

where

$$F(\psi)=\left|\frac{\mathbf{r}}{r}-3\frac{\mathbf{R}}{R}\cos\psi\right|. \tag{4.13}$$

Table 4.1. Radii of the planetary spheres of influence (unit: 10^6 km)

Mercury	0.11	Saturn	54.6
Venus	0.62	Uranus	51.8
Earth	0.93	Neptune	86.8
Mars	0.58	Pluto	15.1
Jupiter	48.8		

Expression (4.13) is composed of two unit vectors $\mathbf{r}/r$ and $\mathbf{R}/R$ and, due to the theorem of cosines, the equality $F^2(\psi) = 1 + 3\cos^2\psi$ is valid. Then we can write $1 \leq F^2(\psi) \leq 4$. Thus, function $F^{1/5}(\psi)$ is close to unity. Hence, the surface called the sphere of influence is indeed close to being a geometric sphere (radius variation $\approx 15\,\%$) and an approximate expression for its radius is

$$r \approx R\left(\frac{m_2}{m_1}\right)^{2/5}.$$

Sometimes we use the term 'gravisphere' instead of 'sphere' in order to emphasize that the accelerations generated by gravitational attractions are the main reasons to introduce this sphere.

It is seen that the size of the sphere of influence depends on two factors, namely the distance of a particular planet from the Sun and the relevant mass ratio. We present in Table 4.1 the radii of the planetary spheres of influence of a number of planets. The radius of the sphere of influence surrounding the Moon within the pair Earth-Moon is 66 200 km.

We now estimate the magnitude of the expression $|\mathbf{b}|/|\mathbf{B}|$ at the border of a particular sphere of influence. Using expressions already obtained above we can write

$$\frac{|\mathbf{b}_2|}{|\mathbf{B}_2|} = \frac{m_1}{m_2}r^2\left|\frac{\mathbf{s}}{s^3} - \frac{\mathbf{R}}{R^3}\right| \approx \frac{m_1 r^3}{m_2 R^3}F(\psi). \tag{4.14}$$

Using relation (4.12) to substitute for r/R we obtain

$$\frac{|\mathbf{b}_2|}{|\mathbf{B}_2|} = \left(\frac{m_1}{m_2}\right)^{1/5}F^{2/5}(\psi). \tag{4.15}$$

Table 4.2. The magnitudes of the ratio $|\mathbf{b}|/|\mathbf{B}|$ at the borders of particular spheres of influence

Earth	0.08
Mars	0.05
Jupiter	0.25

The magnitudes of this ratio for several planets, taking $F = 1$, are given in Table 4.2. Thus, for the terrestrial planets, the value of the relative perturbing acceleration caused by gravitational attraction is estimated to be only a few percent at the borders of their relevant spheres of influence. For large planets like Jupiter, these values can reach 25%.

The value of the relative perturbing acceleration of the Sun is estimated to be of the order of 8% at a distance of 900 000 km from the Earth. The Moon is located at a distance of about 380 000 km from the Earth. The relative perturbing acceleration varies in proportion to the cube of the distance from a planet (see (4.14)). Thus, the perturbing acceleration of the Sun near the Moon is of the order of less than 1% while, at the trajectory of a SC orbiting the Earth, it vanishes.

4.3 Kislik Gravisphere

The sphere of influence considered above does not provide a perfect solution to the problem of finding the best surface of separation between regions dominated, respectively, by 'solar' and 'planetary' influences. Laplace searched for the best surface, having regard to the convenience of performing numerical integration of the motion equations. The sphere of influence method fulfills this requirement. However, this method proceeds from a local approach where a point on a trajectory that satisfies (4.2) is sought. An integral approach which takes into account the properties of the whole trajectory rather then those of individual points, is preferable. On adopting an integral approach, the real trajectory is approximated by joining together a sequence of segments of conic sections. The problem of defining the 'goodness of

fit' between the real and the approximated trajectories then needs to be addressed.

At the border of separation of regions dominated, respectively, by 'solar' and 'planetary' influences, it is necessary to take into account the effect on the motion of a SC of the Sun and the planets. Thus, a sequence of restricted three-body problems appears, made-up from individual trios composed of the elements SC-Sun-planet.

Taking into account the real properties of planetary orbits in the Solar System, it is possible to use the solution of the circular restricted three-body problem (see Section 3.4.1). The first integral of this system is Jacobi's integral. Consequently, to use this integral for finding a new gravisphere* is reasonable because the first integral constant describes the trajectory as a whole rather than its local properties at particular points.

Consider a model trajectory of SC motion which is at once simple, admits analytical treatment and, at the same time, closely approximates the basic properties peculiar to real trajectories. For this purpose we introduce the simple trajectory of a SC that lies in the plane of the Sun and the Earth. The approximate trajectory is vertical at its beginning (rocket launch) and directed through the centre of the Earth. On leaving the Earth's gravisphere, the trajectory is an ellipse with a focus situated in the Sun. Let us find on this model trajectory a point A which marks the transformation from rectilinear lifting to elliptical motion around the Sun. This composite trajectory should be as close as possible to the real one which is executed at all times under the influence of both the Earth and the Sun.

Consider the trajectory of vertical lifting within the framework of a two-body problem (Earth–SC). The integral of energy (1.10) then has a form

$$h = V_{abs}^2 - \frac{2k^2 m_2}{r_2}$$

where V_{abs} is the absolute velocity of the SC, r_2 is the distance from the Earth to the SC and we neglect the mass of the SC in comparison with the mass of the Earth m_2.

Within the framework of a circular restricted three-body problem (Sun–Earth–SC), we introduce a new system of units

*M.D.Kislik. Spheres of Influence of Large Planets and the Moon, *Cosmic Research.* 1964, Vol.2, Issue 6, pp.853–858.

$$\begin{array}{lll} \text{unit of} & \text{mass} & — \quad m_1 + m_2, \\ \text{unit of} & \text{length} & — \quad d_1 + d_2, \\ \text{unit of} & \text{velocity} & — \quad V_{hel} \end{array}$$

appropriate for our consideration, where m_1 and m_2 are the masses of the Sun and the Earth respectively; d_1 and d_2 are the corresponding distances of the Sun and the Earth to their center of mass and V_{hel} is the heliocentric velocity of the Earth. In our existing notation $R = d_1 + d_2$ (see Section 3.4.1).

In the new scale of units the angular velocity of the system Sun–Earth $n = 1$ since $n = V_{hel}/R$. Taking into account that the Earth moves around the Sun in a circular orbit, its velocity is defined by (2.3). Recalling further that $\mu = k^2(m_1 + m_2)$ and having regard to the newly introduced units we obtain $k = 1$. This allows us to transform the integral of energy for the point of start (supposed to be instantaneous) to the following form

$$h = V_*^2 - \frac{2\kappa}{r_{20}}, \tag{4.16}$$

where

$$\kappa = \frac{m_2}{m_1 + m_2}, \tag{4.17}$$

r_{20} is the distance of the SC from the center of the Earth at the initial moment of starting and V_* is the corresponding value of the SC's absolute velocity in the framework of a two-body problem.

The constant C' of Jacobi's integral due to the equalities (3.13) and (3.14) has the form

$$C' = n^2(x^2 + y^2) + 2k^2\left(\frac{m_1}{r_1} + \frac{m_2}{r_2}\right) - V_{rel}^2,$$

V_{rel} is the relative velocity of the SC and, in the new system of units,

$$C' = r^2 + \frac{2\kappa}{r_2} + 2\frac{1-\kappa}{r_1} - V^2. \tag{4.18}$$

Here V is the unitless relative velocity of the SC and the transformation from $x^2 + y^2$ to r_2 is made owing to the fact that the trajectory of the SC lies in the plane Oxy.

Let us now obtain a relationship between V_* and the relative velocity V at the moment of starting. At this initial moment $r_2 = r_{20}$

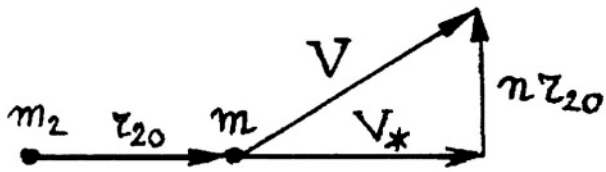

Figure 4.3. A relationship between V_* and the relative velocity V at the moment of starting

and from the triangle of velocities illustrated in Fig.4.3 it is seen that, as $n = 1$

$$V^2 = V_*^2 + r_{20}^2.$$

Hence, for the initial moment of time denoted by index "zero" , the constant of Jacobi's integral (4.18) has the form

$$C' = r_0^2 + \frac{2\kappa}{r_{20}} + 2\frac{1-\kappa}{r_{10}} - V_*^2 - r_{20}^2.$$

We make the assumption that the start is executed from a particle m_2 and not from the surface of a planet with non-zero radius. Using the integral of energy (4.16) and letting $r_{20} \to 0$, we get

$$C' = \kappa^2 - 4\kappa + 3 - h, \tag{4.19}$$

and also $r_{10} \to 1$ and $r_0 \to 1 - \kappa$. The above magnitude of C' (4.19) corresponds to the moment of starting under idealized conditions. Along the real trajectory at each point the equality

$$C' = \text{const} \tag{4.20}$$

pertains. If we consider the real trajectory and the model one (composed from a rectilinear line and the arc of an ellipse) again and calculate the magnitude of C' using (4.18) for points along the model trajectory, then we at once meet an invalidity of (4.20). This is quite natural because the composed trajectory does not coincide with the trajectory which an exact solution of the circular restricted three-body problem would provide.

Construct a function

$$\overline{C} = r^2 + \frac{2\kappa}{r_2} + 2\frac{1-\kappa}{r_1} - V^2 \tag{4.21}$$

based on the expression (4.18) for points along an elliptical part of the model trajectory, which is an approximate trajectory of SC motion in the region dominated by the Sun. Let us now find the deviation of this function from the exact value of Jacobi's integral constant

$$\Delta C = \overline{C} - (\kappa^2 - 4\kappa + 3 - h). \tag{4.22}$$

The value of ΔC depends on parameters that include r_{2A} and α. The parameter mentioned determines a position of the point A at which the rectilinear and elliptical parts of the trajectory are joined. The angle α determines the direction of start from the Earth (Fig.4.4).

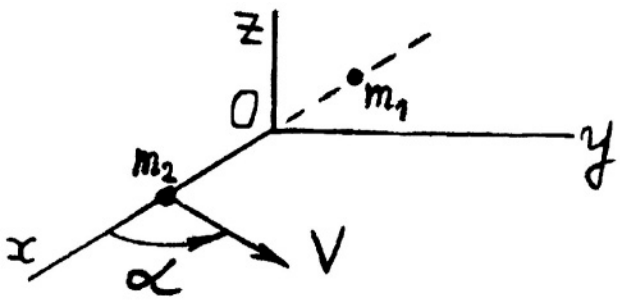

Figure 4.4. Angle α definition

To minimize the contribution of the relatively minor parameter α, we introduce a function

$$J = \frac{1}{2\pi} \int_0^{2\pi} (\Delta C)^2 d\alpha \tag{4.23}$$

which provides a value of $(\Delta C)^2$ averaged by α. We seek now a point A where the function J has a minimum

$$\frac{\partial J}{\partial r_{2A}} = 0. \tag{4.24}$$

The magnitude of r_{2A} obtained from equation (4.24) corresponds to a location of A where the initial part of the elliptical trajectory has that value of Jacobi's integral constant which is closest to the value of C' at the moment of start. We can expect that, using such an initial value, the elliptical model trajectory will be closest to the real one.

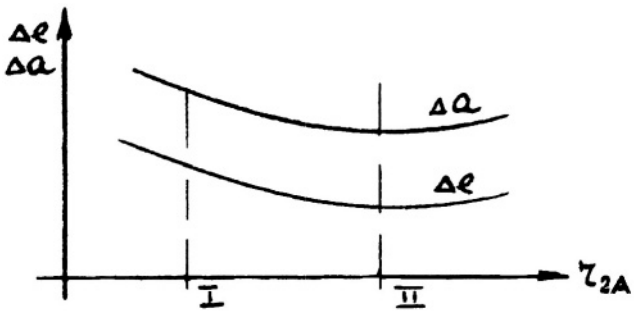

Figure 4.5. Errors in the determination of the principal semi-axis (Δa) and of the eccentricity (Δe) vs r_{2A}

Omitted here are detailed calculations which give the following solution for (4.24)

$$r_{2A} \approx 1.12(\kappa)^{1/3}.$$

It was supposed above that start from the Earth is performed in the plane Oxy and, indeed, practically all interplanetary missions are executed in the *ecliptic* plane. If this is not the case the final conclusion will not be substantially changed since, averaging by angle α, and additionally by the angle between vector $\mathbf{V}$ and the plane Oxy, we get

$$r_{2A} \approx 1.16(\kappa)^{1/3}.$$

Approximating the magnitudes of the factor, neglecting m_2 in comparison with m_1 in the parameter κ and returning to the dimensional variable, we finally obtain

$$r = 1.15R\left(\frac{m_2}{m_1}\right)^{1/3} \tag{4.25}$$

where r is the radius of a sphere called the *Kislik gravisphere.*

The assumption concerning minimum J insures that the approximate trajectory is the best among those obtained above. Comparative calculations can demonstrate that this is the case. In Fig.4.5 errors in the determination of the semi-major axis (Δa) and in the magnitude of the eccentricity (Δe) obtained as a difference between the approximated and the real values, are shown. In the neighbourhood of *II* an obvious minimum is observed where the numerical magnitudes of r_{2A} are close to the magnitudes determined using formula (4.25).

Table 4.3. Radii of Kislik gravispheres (unit: 10^6 km)

Mercury	0.37	Saturn	108
Venus	1.68	Uranus	116
Earth	2.48	Neptune	193
Mars	1.79	Pluto	47.1
Jupiter	88.1		

The neighbourhood of I corresponds to the radius of Laplace's gravisphere (4.12). It is seem from the diagram that joining-up trajectory segments using Laplace's local approach gives a worse approximation than is obtained using an integral approach. Values of the radii of the Kislik gravispheres of the planets are presented in Table 4.3. These radii are a factor of 2–3 times larger than the radii of the corresponding spheres of influence.

Since Jacobi's integral constant depends on the orbital parameters of a SC in the restricted three-body problem, the use of Kislik's gravisphere can increase the accuracy of trajectory determination and also result in a more accurate computation of various parameters of SC motion.

Chapter 5

Equations of Motion in Terms of Osculating Elements

5.1 Reference System and Orbital Elements

Although Cartesian reference systems have, hitherto, been employed in this text, in celestial mechanics it is useful to utilize a spherical reference system based on a representation of the celestial sphere.

Take first a *rectangular Cartesian reference system* $Oxyz$ with origin O at the center of the Earth (Fig.5.1). (All Cartesian reference systems which will be introduced later are rectangular also). Let the plane Oxy coincide with the plane of the Earth's equator and let the axis Ox be directed to the Vernal equinox. The straight line of this direction is created by the intersection of the ecliptic with the Earth's equatorial plane. If a celestial body M considered as a particle, moves around the Earth then, within the framework of the two-body problem, its trajectory lies in a plane called an *orbital plane* that passes through the center of the Earth. The point of crossing of the plane Oxy by the body as it moves along its orbit from the southern to the northern hemisphere is termed an *ascending node.* Similarly, the point of transition from the northern to the southern hemisphere is termed a *descending node.* A straight line through the ascending and

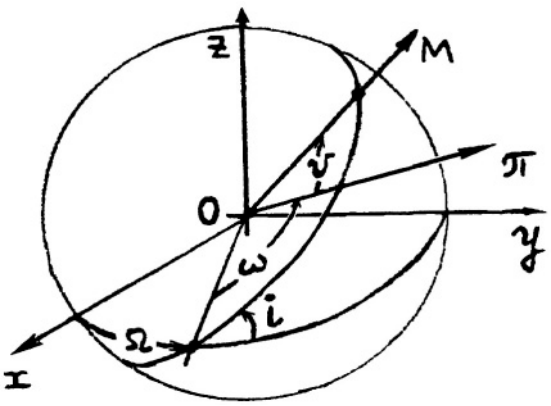

Figure 5.1. Definition of angles Ω, ω, i, v

descending nodes is called the *line of nodes.* The angle between the axis Ox and that segment of the line of nodes to which the ascending node belongs is called the *longitude of the ascending node,* or the *right ascension of the ascending node* and it is designated by Ω. The angle between the equatorial plane and the orbital plane is called the *inclination of the orbital plane* and it is designated by i. Straight lines from the origin O to the pericenter π and to the instantaneous position of the body in its orbit, provide the following angles in the orbital plane: the angular distance of pericenter from the ascending node ω, which is called the *argument of pericenter* (or the *argument of perigee* if we consider the orbit around the Earth), and the angular distance of the body from pericenter, which is called the *true anomaly* v. The sum of these two angles

$$u = \omega + v. \tag{5.1}$$

is termed the *argument of latitude.*

The angles Ω and i completely define the position of the orbital plane with respect to the reference system $Oxyz$. Angle u provides the direction to the instantaneous position of the body. The magnitude r of the radius-vector $\mathbf{r}$ of the body with respect to the Earth provides the distance of the body from the center of the Earth. The two parameters u and r completely define the position of the body.

Next consider a sphere of unit radius which surrounds the origin O. Let the arcs of two great circles be inscribed on this sphere through crossing it by the equatorial plane and by the orbital plane (Fig.5.2).

The intersection point of these two arcs in the segment of the ascending

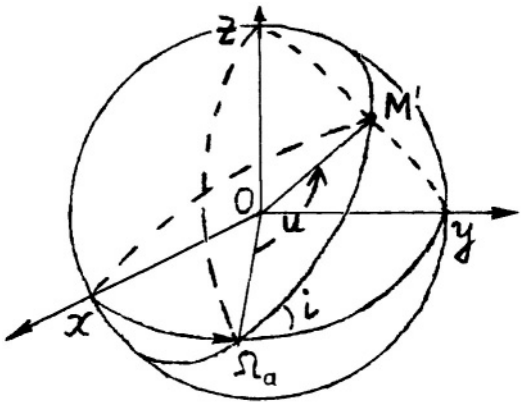

Figure 5.2. Translation from Cartesian to spherical co-ordinates

node we designate by Ω_α.

Let us now establish a relationship between Cartesian coordinates and the spherical coordinates introduced above. The radius-vector **r** of the body (line *OM* in Fig.5.1) intersects the arc of unit radius in a point M' (Fig.5.2). It is obvious that

$$x = r\cos(M'Ox), \qquad y = r\cos(M'Oy), \qquad z = r\cos(M'Oz), \quad (5.2)$$

where the arguments of cosines are the angles with apices at the origin *O*.

Apply now the cosine formula of spherical trigonometry

$$\cos a = \cos b \cos c + \sin b \sin c \cos A, \quad (5.3)$$

where the lower case letters represent the sides of spherical triangles (i.e. the arcs of great circles defined by angles with apices at the point *O*), and capital letters represent dihedral angles between planes of great circles. Oppositely situated sides and angles that correspond to each other are designated by identical letters. Applying the cosine law to spherical triangles $xM'\Omega_\alpha$, $\Omega_\alpha M'y$, $M'z\Omega_\alpha$ and using the equalities (5.2) we obtain the relations

$$\begin{aligned} x &= r(\cos\Omega\cos u - \sin\Omega\sin u\cos i), \\ y &= r(\sin\Omega\cos u + \cos\Omega\sin u\cos i), \\ z &= r\sin u\sin i. \end{aligned} \quad (5.4)$$

On combining (5.4) with

$$r = \frac{p}{1 + e\cos v}, \quad u = \omega + v, \quad t - \tau = \frac{p^2}{c}\int_0^v \frac{dv}{(1 + e\cos v)^2},$$

which we obtained before, (see (1.19), (5.1), and (1.24)), the Cartesian coordinates of a celestial body can be expressed as explicit functions of time

$$x = x(t), \qquad y = y(t), \qquad z = z(t). \tag{5.5}$$

The values of Ω, i, ω, p, e, τ are constants of the first integrals of motion. Expressions (5.5) constitute a general solution of (1.9).

5.2 Equations of Perturbed Motion

The general solution of (1.9) corresponds to the classical case involving two particles moving under the influence of their mutually attractive forces. It was shown above that the introduction of a third particle or of $n > 3$ particles results, generally speaking, in considerably more complex motions. The problem becomes yet further complicated if we take into account that planets may not always be represented by particles (quite often it is necessary to consider them as rigid bodies) and if we include not only the influence of gravitational force but also of solar radiation pressure, the resistive effect of the upper atmospheres of planets, the action of SC thrusters etc. Assume that a SC moves about a body not only under the action of that body's attractive force, but also under the action of a variety of disturbing forces. The equations of motion (1.9) may then be presented in the following form

$$\begin{aligned} &\frac{d\dot{x}}{dt} = X_0 + X, \quad \frac{d\dot{y}}{dt} = Y_0 + Y, \quad \frac{d\dot{x}}{dt} = Z_0 + Z, \\ &\frac{dx}{dt} = \dot{x}, \qquad \frac{dy}{dt} = \dot{y}, \qquad \frac{dz}{dt} = \dot{z}, \end{aligned} \tag{5.6}$$

where

$$X_0 = -\mu\frac{x}{r^3}, \quad Y_0 = -\mu\frac{y}{r^3}, \quad Z_0 = -\mu\frac{z}{r^3}$$

are accelerations produced by the central body considered as a particle, and X, Y, Z are accelerations exerted on the SC as a result of all the external forces acting on it (excluding the force of central body attraction).

5.2.1 Introduction of Osculating Elements

The expressions (5.5) are not a solution of (5.6). However, we may use them to convert the six original variables x, y, z, $\dot{x}$, $\dot{y}$, $\dot{z}$ to six new variables Ω, i, ω, p, e, τ. On performing these transformations we obtain

$$\begin{aligned} \frac{d\Omega}{dt} &= f_1(\Omega, i, \omega, p, e, \tau), \\ \frac{di}{dt} &= f_2(\Omega, i, \omega, p, e, \tau), \\ &\ldots, \\ \frac{d\tau}{dt} &= f_6(\Omega, i, \omega, p, e, \tau). \end{aligned} \tag{5.7}$$

The original variables x, y, z, $\dot{x}$, $\dot{y}$, $\dot{z}$ determine at each moment t the position of the SC and its velocity vector (i.e. each combination defines a point along the trajectory and also the tangent to that point). The overall trajectory is the envelope of a set of tangents.

The new variables Ω, i, ω, p, e, τ give at each moment t an instantaneous conic section; the position of the SC and its velocity vector. The trajectory of the SC is thus the envelope of a set of conic sections. This representation of the trajectory is meaningful since the perturbing accelerations X, Y, Z are much less than the accelerations X_0, Y_0, Z_0. Thus the use of Ω, i, ω, p, e, τ allows us to obtain an approximation to the SC trajectory over a large interval of time. The original six variables x, y, z, $\dot{x}$, $\dot{y}$, $\dot{z}$ for any t yield only instantaneous coordinates and the corresponding velocity vector and do not provide any indication of prior or subsequent SC motion. The transition from rather simple equations (5.6) to the more bulky system (5.7) is useful only in the case of small perturbing accelerations. Since the actual trajectory is obtained as an envelope of instantaneous conic sections, each such conic section is called an *osculating* (or *tangential*) section; the six elements Ω, i, ω, p, e, τ are correspondingly termed *osculating elements* and the time t is termed the *moment* or *epoch of osculation.*

Further, using the formulae for Keplerian motion, the variables Ω, i, ω, p, e, τ provide the position and velocity of a SC at time t. Use of equalities (5.5) to transfer from system (5.6) to (5.7) results in unreasonably bulky calculations. To overcome this, equations of

motion (5.7) can be derived in an explicit form using the first integrals of motion.

5.2.2 Definition of the 'Main Operation'

Let

$$F(x, y, z, \dot{x}, \dot{y}, \dot{z}) = C \tag{5.8}$$

be one of the first integrals of Keplerian motion. In this case the equations

$$\ddot{x} = X_0, \quad \ddot{y} = Y_0, \quad \ddot{z} = Z_0,$$

$$\frac{\partial F}{\partial t} + \frac{\partial F}{\partial x}\dot{x} + \frac{\partial F}{\partial y}\dot{y} + \frac{\partial F}{\partial z}\dot{z} + \frac{\partial F}{\partial \dot{x}}X_0 + \frac{\partial F}{\partial \dot{y}}Y_0 + \frac{\partial F}{\partial \dot{z}}Z_0 = 0.$$

are valid.

If the motion is perturbed, equality (5.8) remains valid since we consider the value of C not to be a constant but, rather, a function of time. Thus, taking into account equation (5.6), we obtain the equality

$$\frac{\partial F}{\partial t} + \frac{\partial F}{\partial x}\dot{x} + \ldots + \frac{\partial F}{\partial \dot{z}}(Z_0 + Z) = \frac{dC}{dt}.$$

Since at any moment of time the coordinates and velocities of true and osculating motions coincide, by subtracting, termwise, the last equality from the previous one we obtain

$$\frac{dC}{dt} = \frac{\partial F}{\partial \dot{x}}X + \frac{\partial F}{\partial \dot{y}}Y + \frac{\partial F}{\partial \dot{z}}Z. \tag{5.9}$$

This formula is sometimes called the *main operation.* It is obtained in practice by initially differentiating the first integrals (5.8) with respect to t (assuming that the coordinates and time contained in function F are constants) and, thereafter, substituting the perturbing accelerations for the derivatives of velocities.

5.2.3 Equations for Osculating Elements

Apply now the main operation to the scalar area integrals for Keplerian motion (1.12)

$$\frac{dc_1}{dt} = yF_z - zF_y = M_x,$$

$$\frac{dc_2}{dt} = zF_x - xF_z = M_y, \qquad (5.10)$$
$$\frac{dc_3}{dt} = xF_y - yF_x = M_z.$$

Here F_x, F_y, F_z are the components of the perturbing accelerations and M_x, M_y, M_z are the corresponding components of the torque due to these perturbing accelerations.

Let us now make a transformation from the coordinates $Oxyz$ to those coordinates $O\xi\eta\zeta$ corresponding directly to an instantaneous osculating conic section. We direct axis $O\xi$ along the line of nodes (towards the point Ω_α), axis $O\zeta$ along the vector of momentum **c**, and suitably align axis $O\eta$ so that a right handed reference system is created (Fig.5.3).

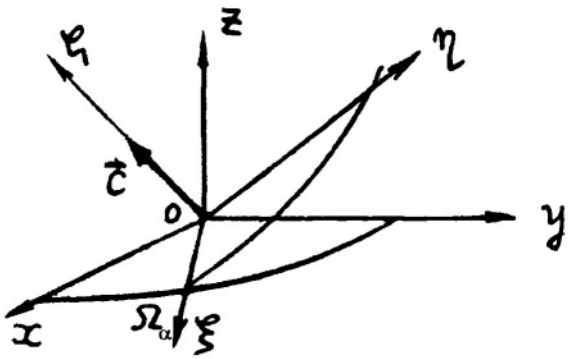

Figure 5.3. Transformation from coordinates $Oxyz$ to coordinates $O\xi\eta\zeta$

In this new reference system equations (5.10) are written as

$$\frac{dc_\xi}{dt} = M_\xi, \qquad \frac{dc_\eta}{dt} = M_\eta, \qquad \frac{dc_\zeta}{dt} = M_\zeta. \qquad (5.11)$$

The projections dc_ξ, dc_η, dc_ζ arise due to variations of the vector **c**, occasioned by its elementary turns around axis $O\eta$ and axis $O\xi$, and to the variation of its module $c = \sqrt{\mu p}$ (see (1.22)).

Consider the increment dc_ξ. An elementary turn of vector **c** around the axis $O\eta$ through an angle $d\alpha$ gives $dc_\xi = cd\alpha$ (Fig.5.4). In association with the turning of vector **c**, the plane $O\xi\eta$ turns to a new position $O\xi_1\eta$. Thus, in the fixed plane Oxy the ascending node Ω_α is

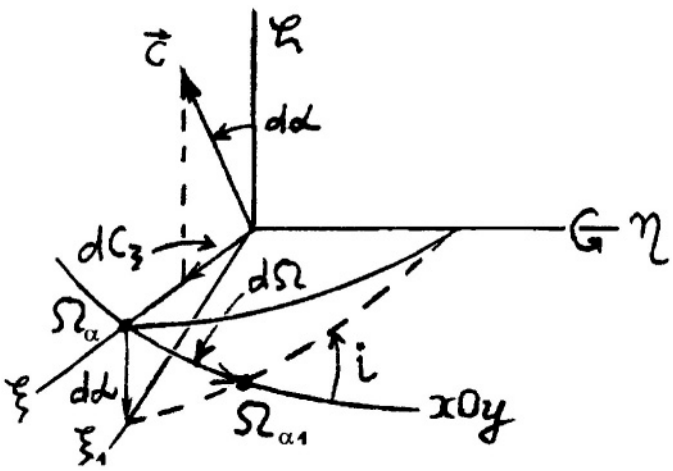

Figure 5.4. Relationship between elementary turns dc_ξ and $d\Omega$

displaced (with a displacement magnitude $d\Omega$) to a new position $\Omega_{\alpha 1}$. This gives the relation $d\alpha = \sin i\, d\Omega$. Therefore, the formula

$$dc_\xi = \sqrt{\mu p} \sin i\, d\Omega \tag{5.12}$$

is valid. Similarly, the component dc_η is associated with an elementary turn of vector **c** around the axis $O\xi$ through an angle di (Fig.5.5) and it is evident that the following relationship

$$dc_\eta = -\sqrt{\mu p}\, di \tag{5.13}$$

is true. The component dc_ζ concerns the variation of the magnitude of vector **c** and is determined by differentiating the expression for this magnitude

$$dc_\zeta = dc = \frac{1}{2}\sqrt{\frac{\mu}{p}} dp. \tag{5.14}$$

Let us obtain now the vector components of the perturbing accelerations torque (Fig.5.6). Let the SC be situated at a point M in the plane $O\xi\eta$. It is recalled that this plane coincides with the instantaneous plane of the osculating orbit. Thus, the position of M is determined by the angle u and the radius-vector **r**. Let the perturbing acceleration be decomposed into three orthogonal components S, T, W directed respectively: along **r**; perpendicular to **r** within the

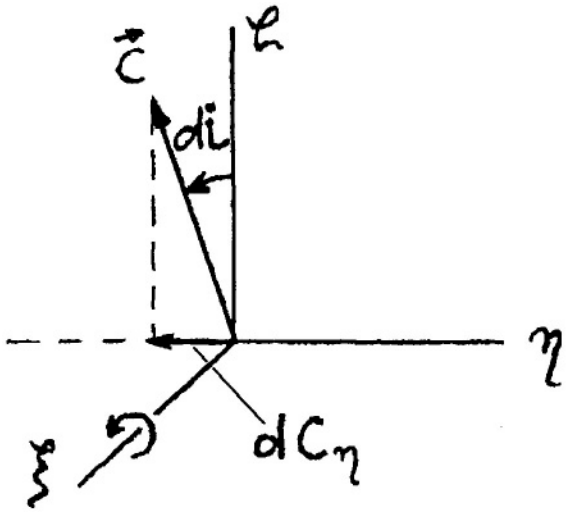

Figure 5.5. Relationship between elementary turns dc_η and di

plane $O\xi\eta$ in the direction of body motion and perpendicular to the latter plane; parallel to the axis $O\zeta$. Positive directions of S, T and W are assigned along the radius-vector $\mathbf{r}$, the vector of motion of M and along the vector of momentum respectively. It is now obvious that

$$M_\xi = Wr\sin u, \qquad M_\eta = -Wr\cos u, \qquad M_\zeta = Tr. \tag{5.15}$$

Substituting the values obtained from (5.12), (5.13), (5.14) and (5.15) in (5.11) we get

$$\frac{d\Omega}{dt} = \frac{r}{\sqrt{\mu p}}\frac{\sin u}{\sin i}W, \quad \frac{di}{dt} = \frac{r}{\sqrt{\mu p}}\cos u\, W, \quad \frac{dp}{dt} = 2\sqrt{\frac{p}{\mu}}rT. \tag{5.16}$$

These equations are three of the six required equations of motion written in terms of osculating elements.

Let us next derive equations for e and ω. Variations in ω originate, generally speaking, in two sources: displacement of the direction to the pericenter π and displacement of the ascending node Ω_α from which the angle ω is measured. Consider initially the consequence of the first mentioned displacement. The angular distance from the ascending node Ω_α to the direction of a celestial body M is determined by the expression

$$u = \omega + v.$$

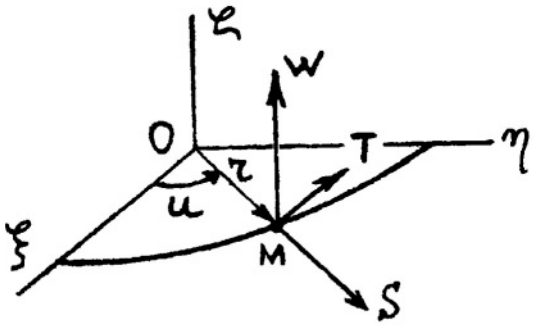

Figure 5.6. Definition of the vector components of the perturbing accelerations torque

In the case considered u = is constant since the ascending node Ω_α is assumed not to be displaced while M is taken to be fixed within the framework of the main operation (see (5.9)). Thus we get

$$\frac{d\omega}{dt} + \frac{dv}{dt} = 0. \tag{5.17}$$

It is necessary to stress that the value dv/dt does not describe motion of the SC itself but rather represents a variation owing to the motion of that direction with respect to which the angle v is measured (the direction to pericenter).

The second reason for a variation in ω is displacement of the ascending node Ω_α giving (see Fig.5.7) the relationship

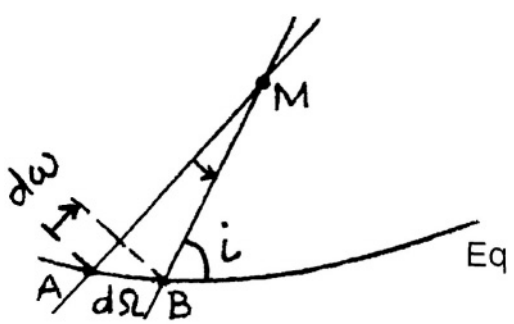

Figure 5.7. Variation of ω

$$d\omega = -\cos i d\Omega. \tag{5.18}$$

The fixing of M (by virtue of the main operation) makes the displacement of ω consequent only on turning of the orbital plane. Let the orbital plane originally be intersected by the plane of the equator Eq at a point A then, owing to an elementary turn of the orbital plane, the intersection occurs at point B. The resulting change in the angular distance of M from the node corresponds to the difference between AM and BM. The combination of both variations in ω (5.17) and (5.18) results finally in the expression

$$d\omega = -dv - \cos i d\Omega. \tag{5.19}$$

Let us continue now our steps to obtain equations for e and ω by deriving the two components of the SC velocity vector $V_r = dr/dt$, $V_n = r(dv/dt)$ in the orbital plane. In the chosen reference system $O\xi\eta\zeta$ by virtue of $c = \sqrt{\mu p}$, the area integral gives

$$r^2 \frac{dv}{dt} = \sqrt{\mu p}. \tag{5.20}$$

Then, using equalities (1.19) and (5.20) we may write

$$V_r = \frac{dr}{dv}\frac{dv}{dt} = \sqrt{\frac{\mu}{p}} e \sin v \tag{5.21}$$

and

$$V_n = r\frac{dv}{dt} = \sqrt{\frac{\mu}{p}}(1 + e\cos v). \tag{5.22}$$

The resulting expressions (5.21) and (5.22) are first integrals of system (1.9) as they link the constants e and p of these integrals with the variables $V_r(t)$, $V_n(t)$ and $v(t)$.

Let us now apply the main operation (see (5.9)) to the first integrals of Keplerian motion (5.21) and (5.22), having written, instead of the derivatives dV_r/dt and dV_n/dt, the corresponding perturbing accelerations S and T so that

$$\frac{de}{dt}\sin v + e\frac{dv}{dt}\cos v = \sqrt{\frac{p}{\mu}} S + \frac{1}{2}\frac{1}{p}\frac{dp}{dt} e \sin v,$$

$$\frac{de}{dt}\cos v - e\frac{dv}{dt}\sin v = \sqrt{\frac{p}{\mu}} T + \frac{1}{2}\frac{dp}{dt}(1 + e\cos v).$$

Next, having substituted in these equations dp/dt from (5.16) and derived, thereby, the derivatives de/dt and dv/dt, we get

$$e\frac{dv}{dt} = -\sqrt{\frac{p}{\mu}}S\cos v + \sqrt{\frac{p}{\mu}}T\left(1+\frac{r}{p}\right)\sin v \tag{5.23}$$

and

$$\frac{de}{dt} = \sqrt{\frac{p}{\mu}}S\sin v + \sqrt{\frac{p}{\mu}}T\left(1+\frac{r}{p}\right)\cos v + e\frac{r}{\sqrt{\mu p}}T. \tag{5.24}$$

Using (5.19) and (5.23), we then obtain

$$\frac{d\omega}{dt} = \frac{1}{e}\left[\sqrt{\frac{p}{\mu}}S\cos v + \sqrt{\frac{p}{\mu}}T\sin v - e\frac{r}{\sqrt{\mu p}}W\sin u\cot i\right]. \tag{5.25}$$

Equalities (5.24) and (5.25) together with (5.16) make up five equations of motion written in terms of osculating elements. To complete the system it is still necessary to derive an equation for $d\tau/dt$. This is not done here because this equation plays no role in what follows and the expression for $d\tau/dt$ is only quoted. Now we write the complete set of equations of motion through osculating elements as follows

$$\begin{aligned}
\frac{d\Omega}{dt} &= \frac{r}{\sqrt{\mu p}}\frac{\sin u}{\sin i}W,\\
\frac{di}{dt} &= \frac{r}{\sqrt{\mu p}}\cos u\, W,\\
\frac{dp}{dt} &= 2\sqrt{\frac{p}{\mu}}rT,\\
\frac{de}{dt} &= \sqrt{\frac{p}{\mu}}S\sin v + \sqrt{\frac{p}{\mu}}T\left(1+\frac{r}{p}\right)\cos v + e\frac{r}{\sqrt{\mu p}}T,\\
\frac{d\omega}{dt} &= \frac{1}{e}\left[\sqrt{\frac{p}{\mu}}S\cos v + \sqrt{\frac{p}{\mu}}T\sin v - e\frac{r}{\sqrt{\mu p}}W\sin\ u\cot i\right],\\
\frac{d\tau}{dt} &= \frac{r^2}{e\mu}\left[-(\cos v - eN\sin v)S + \frac{p}{r}NT\right],
\end{aligned} \tag{5.26}$$

where

$$N(v) = 2\frac{p^2}{r^2}\int_0^v \frac{\cos v dv}{(1+e\cos v)^3}.$$

Usually, the dependence of the osculating elements on t is of minor interest because, as a rule, the perturbing accelerations *S, T, W* are not explicit functions of time. More often, instead of t, a new independent variable corresponding to the angular coordinate of orbital motion (in many cases u) is introduced.

5.2.4 Transformation to the Argument of Latitude

Consider now the area integral. To avoid already used notations we employ the symbol σ to denote the length of an arc on the unit sphere (Fig.5.8). Then the area integral takes the form

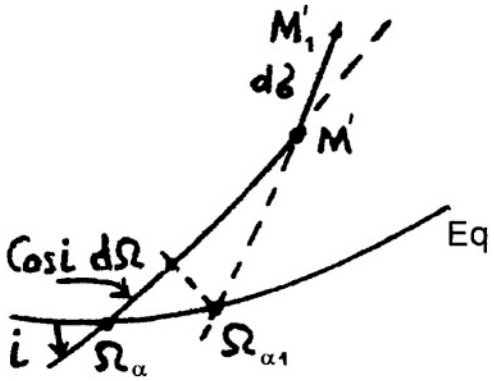

Figure 5.8. Variation of the argument of latitude u

$$r^2 \frac{d\sigma}{dt} = \sqrt{\mu p}. \tag{5.27}$$

Starting at a time t, a point M' on the unit sphere representing the instantaneous position of a celestial body moves through an angular distance $d\sigma$ to point M_1' during interval dt. Consider the variation in the argument of latitude accompanying this displacement. Firstly, it increases along $d\sigma$ and, secondly, the reference point varies as it passes from Ω_α to $\Omega_{\alpha 1}$. The result is that we get

$$du = d\sigma - \cos i d\Omega.$$

Then the first integral (5.27) assumes the form

$$r^2 \left(\frac{du}{dt} + \cos i \frac{d\Omega}{dt} \right) = \sqrt{\mu p}.$$

Taking the value $\frac{d\Omega}{dt}$ from (5.26), we obtain the relationship

$$\frac{du}{dt} = \frac{\sqrt{\mu p}}{r^2} \left(1 - \frac{r^3}{\mu p} W \cot i \sin v \right). \tag{5.28}$$

Using expression (5.28) we now change the independent variable t in (5.26) to the independent variable u. In this process the right-hand terms that are linear functions of S, T, W become fractional, thereby causing considerable difficulties in integrating the resulting equations. We may neglect several terms by taking into account that W is small. Thus, instead of the exact equality (5.28) we can use the approximate equation

$$\frac{du}{dt} = \frac{\sqrt{\mu p}}{r^2}. \tag{5.29}$$

Next, instead of the exact system (5.26) we can write approximate equations of motion in terms of osculating elements

$$\begin{aligned}
\frac{d\Omega}{du} &= \frac{r^3}{\mu p}\frac{\sin u}{\sin i}W, \\
\frac{di}{du} &= \frac{r^3}{\mu p}\cos u\, W, \\
\frac{dp}{du} &= \frac{2r^3}{\mu}T, \\
\frac{de}{du} &= \frac{r^2}{\mu}\left[S\sin v + \left(1+\frac{r}{p}\right)T\cos v + \frac{er}{p}T\right], \\
\frac{d\omega}{du} &= \frac{r^2}{e\mu}\left[-S\cos v + \left(1+\frac{r}{p}\right)T\sin v\frac{er}{p}W\sin u\cot i\right], \\
\frac{d\tau}{du} &= \frac{r^4}{e\sqrt{\mu^3 p}}\left[-(\cos v - eN\sin v)S + \frac{p}{r}NT\right],
\end{aligned} \tag{5.30}$$

These equations are valid when $e \neq 0$.

Chapter 6

Braking of a SC in the Earth's Atmosphere

In the upper atmosphere at an altitude of several hundred kilometers above the surface of the Earth, slow braking of a SC occurs. Consider the change of the SC's orbit during such braking.

6.1 Qualitative Analysis of Osculating Element Evolution

Assume that the Earth is spherical and that the atmospheric density decreases with increasing altitude in an identical way at all geographic locations (i.e. that the configuration concerned has the property of central symmetry). Suppose also that motion of the SC in the atmosphere is affected only by the force of resistance, so that there is no lifting aerodynamic force. Then, the force of aerodynamic drag is subject to the law

$$R_x = c_x \frac{\rho V^2}{2} F \tag{6.1}$$

where c_x is a drag coefficient, ρ is the relevant atmospheric density, V is the velocity of the SC with respect to the atmosphere and F the square of the cross-sectional area of the SC. From Fig.6.1 we get expressions for components S and T of the acceleration

$$S = -\frac{R_x}{m}\frac{V_r}{V} = -c_x \frac{F}{m}\frac{\rho V}{2} V_r; \tag{6.2}$$

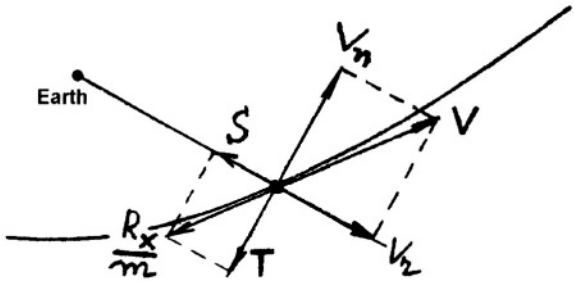

Figure 6.1. Acceleration components due to atmospheric drag

$$T = -\frac{R_x}{m}\frac{V_n}{V} = -c_x \frac{F}{m}\frac{\rho V}{2} V_n. \tag{6.3}$$

As we suppose an aerodynamic lifting force to be absent

$$W = 0. \tag{6.4}$$

Components S and T of the total acceleration of braking can be easily calculated using the relations (6.2), (6.3), (5.21), and (5.22). To integrate (5.30) it is necessary to define a relation $\rho = \rho(r)$. Different models of this relation exist. However, to use them requires extensive calculation. We therefore confine ourselves to a qualitative analysis of the solution of system (5.30).

Equality (6.4) results from the first two equations of (5.30) in the form

$$\frac{d\Omega}{du} = 0, \qquad \frac{di}{du} = 0,$$

whence it at once follows that

$$\Omega = \text{const}, \qquad i = \text{const}\,. \tag{6.5}$$

This means that the orbital plane does not change its position in space. Therefore, aerodynamic braking can be deduced to involve the problem of planar, rather than of spatial, motion.

We now introduce

$$C = c_x \frac{F}{m}, \tag{6.6}$$

where C is called the *ballistic parameter.* Then the third equation of (5.30) gives the following change in the orbital parameter p for one

revolution of the SC around the Earth

$$\Delta p = \int_0^{2\pi} \frac{2r^3}{\mu} T du = -\frac{C}{\mu} \int_0^{2\pi} r^3 \rho V V_n du. \tag{6.7}$$

All the variables in this integral are positive (for V_n this follows from (5.22) because the inequality $e < 1$ is valid for elliptical trajectories around the Earth and the assertion is obvious for the other variables). Therefore, during aerodynamic braking, there is a monotonic decrease of the *parameter*,

$$\Delta p < 0 \tag{6.8}$$

This corresponds to a decrease in the size of the orbit of the SC.

Let us analyze the variation in the eccentricity of the orbit for one revolution of the SC around the Earth. The fourth equation of (5.30) after suitable transformations gives

$$\begin{aligned} \Delta e = -\frac{C}{2\sqrt{\mu}} \int_0^{2\pi} \frac{r^2 \rho V}{\sqrt{p}} \Big[e \sin^2 v + \cos v (1 + e \cos v) \\ \times \Big(1 + \frac{1}{1 + e \cos v} \Big) + e(1 + e \cos v) \frac{1}{1 + e \cos v} \Big] du \\ = -\frac{C}{2\sqrt{\mu}} \int_0^{2\pi} \frac{r^2 \rho V}{\sqrt{p}} (e + \cos v) du. \end{aligned} \tag{6.9}$$

The factor before the round brackets inside the integral is certainly positive, while the first term within the round brackets is also positive and the second is alternating. Let us consider in more detail the expression

$$\int_0^{2\pi} \frac{r^2 \rho V}{\sqrt{p}} \cos v du. \tag{6.10}$$

Suppose the perturbing accelerations S and T are sufficiently small that, in some particular time interval, the alternating values p, e and ω can be neglected. Let this time interval be of the order of the orbital revolution time. In this case, during one revolution $p \approx \text{const}$, $e \approx \text{const}$, $\omega \approx \text{const}$. Consequently, omitting the approximation sign we write the equality

$$du = dv + d\omega = dv. \tag{6.11}$$

Then, substituting (6.11) in (6.10) we obtain

$$\int_0^{2\pi} \frac{r^2\rho V}{\sqrt{p}} \cos \upsilon du = \int_0^{2\pi} \frac{r^2\rho V}{\sqrt{p}} \cos \upsilon d\upsilon. \tag{6.12}$$

In the latter integral, the parameter p is assumed to be constant. For typical orbits the ellipticity is usually small and, consequently, we consider the values r and V to be functions that weakly depend on υ. On the other hand, the atmospheric density ρ varies strongly with r (as an exponential function). The value υ varies for one turn around the planet by 2π. Therefore, the value of the integral (6.12) depends primarily on the term

$$\int_0^{2\pi} \rho \cos \upsilon d\upsilon. \tag{6.13}$$

The profiles of the functions $\cos \upsilon$ and $\rho = \rho(\upsilon)$ are shown in Fig.6.2. When constructing the plot of $\rho(\upsilon)$ it was taken into account that

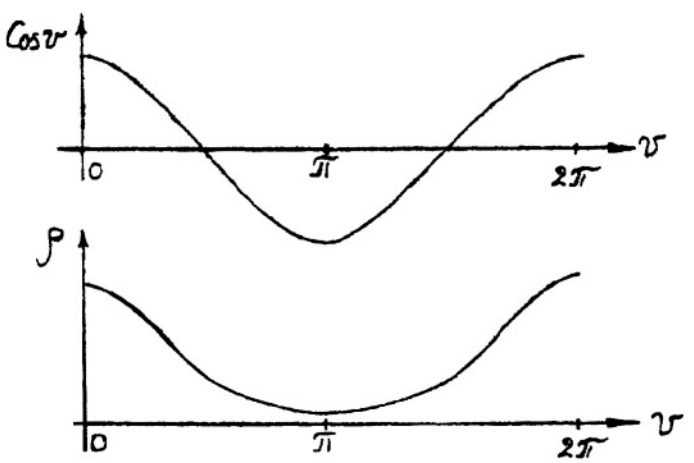

Figure 6.2. Profiles of the functions $\cos \upsilon$ and $\rho = \rho(\upsilon)$

at perigee ρ reaches its maximum value, while at apogee ρ is at its minimum. Clearly, the sign of (6.13) depends on those υ where ρ has its greatest value and thus $\rho(\upsilon)$ may be considered to act as a weight function. Therefore, it is possible to write

$$\int_0^{2\pi} \rho \cos \upsilon d\upsilon > 0,$$

and, consequently,

$$\int_0^{2\pi} \frac{r^2 \rho V}{\sqrt{p}} \cos v du > 0.$$

Overall, the right-hand side of equality (6.9) is negative and, therefore,

$$\Delta e < 0. \tag{6.14}$$

This corresponds to a monotonic reduction of the eccentricity of the SC's orbit (i.e. the shape of the orbit tends to become circular).

The inequalities (6.5), (6.8) and (6.14) show that SC braking by virtue of atmospheric drag results in the orbit decreasing in size and approaching a circular shape, without changing its orientation in space. It is evident that, if the original orbit is circular ($e = 0$), this circularity will be maintained during orbit evolution due to braking. It is pertinent to recall here that the conclusions stated above are based on the use of an approximate relationship. Thus, the SC's orbit in actuality reduces along a spiral which, in each orbit, is close to a circle.

These results allow us to make a qualitative analysis of the lifetime of a SC. Rough estimates show that a body moving in a circular orbit at an altitude near 100 km is effected by such strong atmospheric braking that the SC reaches the surface of the Earth in less than one revolution. Therefore, it is possible to say within the framework of osculating element techniques that the SC ceases to exist after the altitude of its trajectory becomes less than a critical altitude $H_{crit} \sim 100$ km.

From (1.19) it follows that for perigee $(v = 0)$

$$r = \frac{p}{1+e}.$$

Therefore, for orbits close to a circle $(e \approx 0)$ we can take $\Delta p \approx \Delta r$. Simultaneously, the altitude change ΔH for one orbit is determined by

$$\Delta H \approx \Delta p. \tag{6.15}$$

The sum Δp for N orbits characterizes the altitude change in perigee for these N revolutions. Also, Δp is proportional to the ballistic coefficient C and to the density ρ, where ρ is strongly dependent on the altitude of the orbit (see (6.7)). The relationship between Δp and N

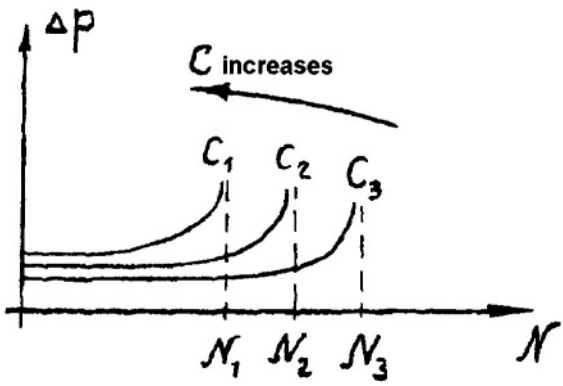

Figure 6.3. Change of the orbital parameter p as a function of the number of orbits N executed at different values of C (the arrow shows the direction of increasing C)

for a range of values of C at some initial altitude of perigee H_π is illustrated in Fig.6.3. This diagram shows that, at first, Δp changes rather slowly and, consequently, the altitude at perigee does not substantially decrease in the process of orbiting around the Earth. Later, Δp begins to sharply increase. The perigee is then quickly lowered and, after a number of revolutions, the SC reaches a critical value H_{crit}, beyond which the model of motion based on the use of osculating elements no longer applies. The number of revolutions N_1, N_2, N_3 executed by the SC before H_{crit} is reached (i.e. the lifetime of the SC) decreases as C increases (Fig.6.3).

6.2 The Descending SC 'Paradox'

It appears evident that the velocity of a SC should decrease due to the effect of atmospheric drag. Consider the case of a SC moving in an orbit that is close to circular where, due to the braking effect of the rarefied atmosphere, the SC is slowly descending to the Earth along a spiral trajectory. Assume that the angle α between the vector $\mathbf{V}$ of SC velocity and the transverse orbit direction does not vary (i.e. that $d\alpha/d\varphi = 0$, which is valid in a rarefied atmosphere), where φ is the polar angle of the radius-vector $\mathbf{r}$ of the SC center of mass (Fig.6.4). Introduce the gravitational force on the SC

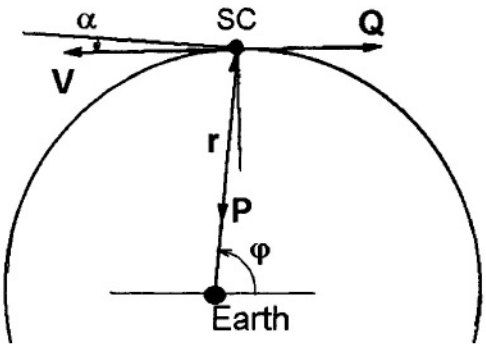

Figure 6.4. SC descending along a spiral trajectory (the magnitudes of the forces are not scaled)

$$\mathbf{P} = -\frac{\mu_g m}{r^3}\mathbf{r}$$

where m is the SC mass. Let the braking force, which is in opposition to **V**, be designated by **Q**. Project the vector equation of motion of the SC

$$m\dot{\mathbf{V}} = \mathbf{P} + \mathbf{Q} \tag{6.16}$$

onto the direction of the vector **V**

$$m\dot{V} = P\sin\alpha - Q \tag{6.17}$$

and onto the perpendicular to this direction

$$mV\dot{\varphi} = P\cos\alpha. \tag{6.18}$$

The acceleration vector $\dot{\mathbf{V}}$ is represented as the resultant of a vector with magnitude $\dot{V}$ along vector **V** in (6.17) and a vector with magnitude $V\dot{\varphi}$ in the perpendicular direction in (6.18).

On substituting for $\dot{\varphi}$ in(6.18) using the equality $V\cos\alpha = r\dot{\varphi}$, and the expression for **P**, we obtain a relationship between the variables V and r

$$V^2 = \frac{\mu_g}{r}. \tag{6.19}$$

Taking into account the geometric relationship $\dot{r} = -V\sin\alpha$ and the expression for **P**, we may now write

$$m\dot{V} = m\frac{dV}{dr}\frac{dr}{dt} = m\frac{dV}{dr}(-V\sin\alpha) =$$

$$-\frac{m\sin\alpha}{2}\frac{dV^2}{dr} = \frac{\mu^g m\sin\alpha}{2r^2} = P\frac{\sin\alpha}{2}.$$

Combining this final expression for $m\dot{V}$ with (6.17), we get $Q = P \sin \alpha/2$ and, consequently,

$$m\dot{V} = Q. \tag{6.20}$$

This indicates that the velocity V increases with the braking force Q, thus giving rise to what is called the *descending SC 'paradox'* (namely that although the SC is affected by a braking force, its velocity increases). Actually, there is no paradox here because more detailed consideration shows that a component of the gravitational force acts along the velocity vector (see (6.17)) and, since the magnitude of the braking force is twice less than this component, SC acceleration occurs.

Let us now consider SC motion in the deep atmosphere along the final segment of its trajectory (at altitudes less than approximately 100 km), where the assumption that $d\alpha/d\varphi = 0$ is no longer valid since α increases faster than does φ. It can be qualitatively shown [1] that, in these circumstances

$$\frac{1}{\cos \alpha} \frac{d\alpha}{dt} = \frac{g}{V} \tag{6.21}$$

where g is the gravitational acceleration, which is taken to be constant near the Earth. Substituting for V in (6.21) its value averaged along the final trajectory segment and integrating this equation, we can obtain a relationship between angle α and time t which indicates that the final descent is almost along a plumb-line. This is because the averaged value of V is typically several tens of times less than the orbital velocity of the SC.

Chapter 7

Terrestrial Nonsphericity and SC Motion

To analyze the effects of the nonsphericity of the Earth's gravitational field on SC motion requires initial consideration of the theory of gravitational potential.

7.1 Introduction to Models of the Earth's Gravitational Field

The Earth's gravitational field lacks exact spherical symmetry because this body is not itself spherical, while also the mass in its interior is not symmetrically distributed. The largest contributor to the Earth's asymmetry is its equatorial bulge, which is such that the polar radius differs from the equatorial radii by about 21 km. The equatorial radii differ from one another by a few hundred meters.

We now introduce the potential U of the Earth's gravitational field using a *spherical harmonic* expansion and spherical coordinates (r, λ, φ),where r is the geocentric distance, λ is the geographic longitude measured from the Greenwich meridian and φ is the geocentric latitude

$$U = \frac{\mu}{r}\left[1 - \sum_{n=2}^{\infty} J_n \left(\frac{R_{ev}}{r}\right)^n P_n(\sin\varphi) \right. \tag{7.1}$$
$$\left. + \sum_{n=2}^{\infty}\sum_{m=1}^{n} J_{nm}\left(\frac{R_{ev}}{r}\right)^n P_n^m(\sin\varphi)\cos m(\lambda - \lambda_{nm})\right].$$

Here μ is a gravitational parameter of the Earth and R_{ev} the mean equatorial radius; P_n^m are associated Legendre's functions of degree n and order m; P_n are Legendre polynomials of order n; J_n, J_{nm} are harmonic coefficients, λ_{nm} are phase angles corresponding to J_{nm}. The relationships between J_{nm}, λ_{nm} and C_{nm}, S_{nm} are

$$J_{nm}^2 = C_{nm}^2 + S_{nm}^2, \quad \lambda_{nm} = \frac{1}{m} \arctan \frac{S_{nm}}{C_{nm}} \quad (n \geq 2,\ 1 \leq m \leq n).$$

The coefficients J_n are called *zonal harmonic* coefficients. If $n \neq m$, C_{nm} and S_{nm} are called *tesseral harmonic* coefficients. If $n = m$ they are called *sectoral harmonic* coefficients.

We need to mention here that a force $\mathbf{F}$ exerted on a particle of mass m at an external point of space is represented by

$$\mathbf{F} = m\nabla U$$

where the potential is given by (7.1).

Zonal Harmonics

Zonal harmonics are defined by the P_n polynomials, where the dependence of the potential on longitude vanishes and the field is symmetrical about the polar axis. For any $P_n(\sin\varphi)$ there are n circles of latitude along which P_n is zero and hence $(n+1)$ zones in which the function is alternately increasing and decreasing. Term J_2 accounts for most of the Earth's gravitational departure from a perfect sphere and thus reflects the Earth's oblateness. Representative low-order zonal harmonic coefficients are $J_2 = 1.082 \cdot 10^{-3}$, $J_3 = -2.53 \cdot 10^{-6}$, $J_4 = -1.61 \cdot 10^{-6}$. It can be seen that J_2 is about 400 times larger than the next largest value J_3.

Sectoral Harmonics

Sectoral harmonics occur when $n = m$. Whereas the polynomials $P_{nn}(\sin\varphi)$ are constant, the terms $\cos n\lambda$ and $\sin n\lambda$ are zero for $2n$ different values of λ. Hence the lines along which the functions $P_{nn}(\sin\varphi)\cos n\lambda$ and $P_{nn}(\sin\varphi)\sin n\lambda$ vanish, are meridians of longitude which divide the sphere into $2n$ sectors, alternately positive and negative.

Tesseral Harmonics

The functions $P_{nm}(\sin\varphi)\cos m\lambda$ and $P_{nm}(\sin\varphi)\sin m\lambda$ are referred to as tesseral harmonics when $n \neq m$ since the Earth is associatively divided into a mosaic of domains which are alternately positive and negative. There are $(n-m)$ circles of latitude along which $P_{nm}\sin\varphi$ is zero, whereas the terms $\cos m\lambda$ and $\sin m\lambda$ vanish along $2m$ meridians of longitude. The tesseral and sectoral harmonics take into account anomalous mass distributions in longitudinal regions. Although their effect is much smaller than that associated with zonal harmonics, this contribution becomes important when it is not averaged out over a long period of SC motion. This occurs in the case for a SC in geosynchronous orbit. For SC in other orbits, the effect vanishes due to averaging over revolutions around the Earth and only zonal harmonics then need to be taken into account. The cross section of the Earth in its equatorial plane is not a circle but an ellipse. The primary tesseral harmonic is designated by J_{22} and the corresponding longitude of symmetry, which is denoted by λ_{22} has a value of –14.7° relative to the Greenwich meridian.

7.2 Oblateness of the Gravitational Field

Now we may start to analyze the effects of nonsphericity of the Earth's gravitational field. Introduce a reference system $Oxyz$. Its origin O is placed at the Earth's center of mass (Fig.7.1).

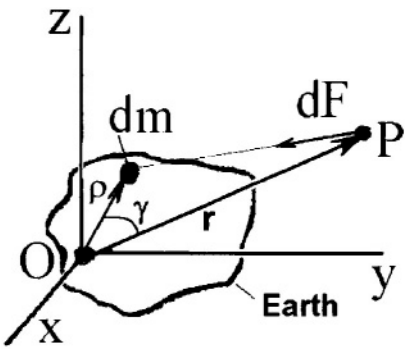

Figure 7.1. The aspherical Earth and an attracted particle P

The partial derivatives with respect to the coordinates determine the gravitational acceleration of particle P with mass m_0

$$\frac{F_x}{m_0} = \frac{\partial U}{\partial x}, \qquad \frac{F_y}{m_0} = \frac{\partial U}{\partial y}, \qquad \frac{F_z}{m_0} = \frac{\partial U}{\partial z}.$$

A common representation of the potential of the Earth's gravitational field is given by (7.1). Taking onto account the notations adopted in (5.6) we may write

$$\frac{\partial U}{\partial x} = X_0 + X, \qquad \frac{\partial U}{\partial y} = Y_0 + Y, \qquad \frac{\partial U}{\partial y} = Z_0 + Z. \tag{7.2}$$

Following the geometric method of MacCullagh* which gives an intuitive feeling for gravity and inertia† and using Fig.7.1, we can represent the cosine of the angle γ constructed by vectors $\boldsymbol{\rho}$ and $\mathbf{r}$ from the Earth's origin to an element dm and to an attracted particle P, using the scalar product of the vectors (i.e. $\cos\gamma = \mathbf{r}\boldsymbol{\rho}/(r\rho)$). The distance l between P and dm is calculated through the expression

$$l = r\left[1 - 2\frac{\rho}{r}\cos\gamma + \left(\frac{\rho}{r}\right)^2\right]^{1/2}.$$

Consider the general formula for the potential of an attracting body

$$U = k^2 \int \frac{dm}{r}. \tag{7.3}$$

Using the expression

$$\left[1 - 2\frac{\rho}{r}\cos\gamma + \left(\frac{\rho}{r}\right)^2\right]^{-1/2} = \sum_{n=0}^{\infty}\left(\frac{\rho}{r}\right)^n P_n(\cos\gamma)$$

which is valid for Legendre's polynomials $P_n(\cos\gamma)$ we can rewrite (7.3) in the following form

$$U = \frac{k^2}{r}\int \sum_{n=0}^{\infty}\left(\frac{\rho}{r}\right)^n P_n(\cos\gamma)dm. \tag{7.4}$$

*McCullagh, James (1809-1847). Irish mathematician, Professor at Trinity College Dublin. Fellow of the Royal Society. An outstanding geometer, he was especially interested in the ellipsoid and other surfaces of the second order. Also, he studied the physical problem of light propagation.

†J.MacCullagh. On the Attraction of Ellipsoids with a New Demonstration of Clairut's Theorem. *Transactions of the Royal Irish Academy.* Dublin, 1855, Vol.22, Part 1, pp.379–395.

The first three terms of the potential expansion (7.4) have the specific form

$$U = \frac{k^2}{r}\int dm + \frac{k^2}{r^2}\int \rho\cos\gamma dm + \frac{k^2}{2r^3}\int \rho^2(3\cos^2\gamma - 1)dm.$$

Integrating over the Earth's body allows us to write the expression for the potential as

$$U = \frac{\mu}{r} + \frac{k^2}{2r^3}(A_E + B_E + C_E - 3J_P) \tag{7.5}$$

where A_E, B_E, C_E are the principal central moments of inertia of the Earth and J_P its moment of inertia about the axis *OP*. The form of the potential (7.5) is called *MacCullagh's approximation.*

The first term in (7.5) represents the potential of the Earth considered as a particle and determines X_0, Y_0, Z_0. These components obey an inverse square relationship with respect to r. The second term in (7.5) determines the perturbing accelerations *X*, *Y*, *Z*.

Consider within the framework of (7.5) the conditions under which the Earth has a gravitational field equivalent to that of a particle. Such a situation is realized in the following two cases.

- If $r \to \infty$, since at a large distance, the shape of the Earth may be neglected, it can be treated as a particle.
- If the Earth is spherical, $A = B = C = J_P$ for any direction *OP*.

Only nonsphericity of the Earth can result in the appearance of non-zero perturbing accelerations. Nonsphericity is made evident through two effects.

- The force of attraction on a particle *P* by the Earth is not governed by a relation proportional to $1/r^2$ but includes a term of order $1/r^4$.
- The equipotential surfaces cease to be spheres with centers at a particular origin as the potential *U* can, generally speaking, depend on a direction *OP* on which J_P also depends. This expresses the noncentrality of the gravitational field (the force of the attraction on *P* by the Earth does not pass through the origin because the derivatives of *U* along directions normal to *OP* cease to be zero).

These conclusions are based on an analysis of the first two terms in the series expansion. However, they continue to be valid when the other terms in the series are taken into account.

7.3 Calculation of Perturbing Accelerations

The next approximation after a sphere to the shape of the Earth, is an ellipsoid of rotation. In this case, the potential can be expressed by

$$U = \frac{\mu}{r} - \frac{\varepsilon}{3r^3}(3\sin^2\psi - 1) \tag{7.6}$$

where ψ is the geographic latitude and $\varepsilon = 3\mu R_e^2 J_2/2$ is a constant which is proportional to the oblateness of the ellipsoid.

The first term in (7.6) gives the potential of the Earth considered as a particle or as a uniform sphere. The second term determines the contribution of the component defined by the difference between the Earth's true shape and that of a sphere.

Let us obtain, using equality (7.6), the components *S*, *T*, *W* of a perturbing acceleration. Differentiating *U* along the radius r and perpendicular to it in a meridional plane along a length s, we determine the radial and meridional components of the perturbing acceleration

$$g_r = \frac{\partial U}{\partial r} = -\frac{\mu}{r^2} + \frac{\varepsilon}{r^4}(3\sin^2\psi - 1), \tag{7.7}$$

$$g_m = \frac{\partial U}{\partial s} = \frac{\partial U}{\partial \psi}\frac{d\psi}{ds} = -\frac{\varepsilon}{r^4}\sin 2\psi. \tag{7.8}$$

The latitude component is equal to zero by virtue of symmetry. The second term of equality (7.7) is a perturbing acceleration directed along r. This is the component *S* of the acceleration, so that

$$S = \frac{\varepsilon}{r^4}(3\sin^2\psi - 1). \tag{7.9}$$

To obtain the equality (7.8) the obvious relationship $ds = rd\psi$ was used.

To derive *W* and *T* it is necessary to calculate the projections of the meridional acceleration (7.8). Let an orbital plane intersect the plane of the equator at an angle i (Fig.7.2). Consider a spherical triangle

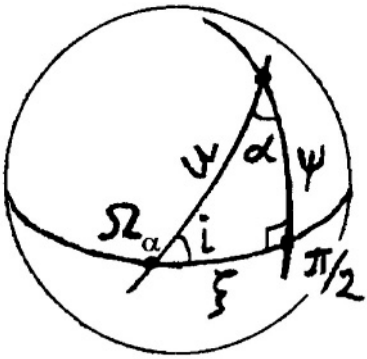

Figure 7.2. Spherical triangle formed by three arcs of great circles

formed by three arcs of great circles corresponding to the planes of the equator, the orbit, and the meridian. The dihedral angles of this spherical triangle are i, $\pi/2$, and angle α. The 'sides' of this triangle are the arcs u, ψ, and some arc ξ.

Let us denote by lower case letters the arcs forming the sides of the spherical triangle and by capital letters the dihedral angles between the planes in which these arcs lie. Then, using the theorem for a spherical triangle

$$\frac{\sin a}{\sin A} = \frac{\sin b}{\sin B} = \frac{\sin c}{\sin C} \tag{7.10}$$

we obtain for the case considered

$$\frac{\sin u}{\sin \pi/2} = \frac{\sin \psi}{\sin i}$$

whence we get

$$\sin \psi = \sin u \cdot \sin i. \tag{7.11}$$

Then, in view of (7.9) the formula for the acceleration component S is

$$S = \frac{\varepsilon}{r^4}(3 \sin^2 u \cdot \sin^2 i - 1). \tag{7.12}$$

Using the cosine theorems for a spherical triangle, we obtain the following formulae for cosines of a spherical right triangle

$$\cos a = \cos b \cdot \cos c, \qquad \cos B = \cos b \cdot \sin C.$$

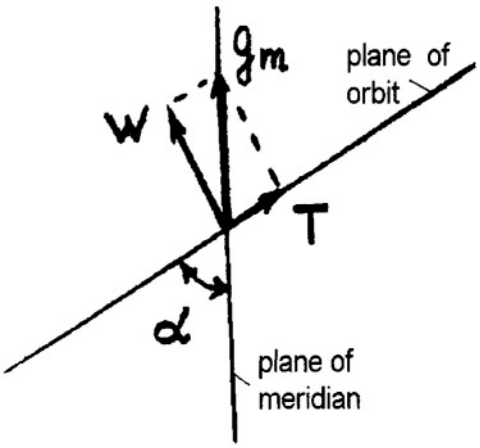

Figure 7.3. Components of the acceleration g_m in the orbital and meridional planes

These expressions give three required relationships

$$\begin{aligned} \cos u &= \cos\psi \cdot \cos\xi, \\ \cos\alpha &= \cos\xi \cdot \sin i, \\ \cos i &= \cos\psi \cdot \sin\alpha. \end{aligned} \tag{7.13}$$

For a point on a unit sphere corresponding to the intersection of orbital and meridional planes, the components of the acceleration g_m are shown in Fig.7.3, namely

$$T = g_m \cos\alpha, \qquad W = g_m \sin\alpha.$$

Using now the expressions (7.8), (7.11), and (7.13), we obtain

$$T = -\frac{\varepsilon}{r^4} \sin 2u \cdot \sin^2 i, \tag{7.14}$$

$$W = -\frac{\varepsilon}{r^4} \sin 2i \cdot \sin u. \tag{7.15}$$

7.4 Evolution of the Equatorial Orbit

To demonstrate the effect of central, but not Newtonian, gravitational attraction we will first analyze the evolution of that orbit in which a perturbing acceleration is directed towards the Earth's center of mass.

This kind of acceleration takes place in equatorial orbits and illustrates a difference between the attraction of a particle by a rigid body and the mutual attraction of two particles. This difference is due to the appearance in the former case of a term proportional to $1/r^4$ in the expression for the acceleration.

If the orbit is equatorial $(i = 0)$ we get from (7.12), (7.14), and (7.15)

$$S = -\frac{\varepsilon}{r^4}, \qquad T = 0, \qquad W = 0. \tag{7.16}$$

Again assuming, as we did earlier, the presence of only weak disturbances, we now further suppose that, during one orbit, the values Ω, i, p, e, ω are essentially constant and that their actual variations are described by discontinuous changes $\Delta\Omega$, Δi, Δp, Δe, $\Delta\omega$ at the end of each revolution of the Earth. This corresponds to the replacement of a smooth curve by a, closely related, step-function.

Equations (5.30) and (7.16) at once give expressions for the derivatives of the first three variables

$$\frac{d\Omega}{du} = 0, \quad \frac{di}{du} = 0, \quad \frac{dp}{du} = 0, \tag{7.17}$$

reflecting the constancy of these values. Consequently,

$$\Delta\Omega = 0, \quad \Delta i = 0, \quad \Delta p = 0. \tag{7.18}$$

From (5.30) and (7.16) using (1.19) and (6.11) we can write

$$\Delta e = -\int_0^{2\pi} \frac{\varepsilon}{\mu r^2} \sin v\, du = -\frac{\varepsilon}{\mu p^2} \int_0^{2\pi} (1 + e\cos v)^2 \sin v\, dv = 0. \tag{7.19}$$

Therefore, a particle on being perturbed by a central force does not change its plane of motion $(\Delta i = 0, \Delta\Omega = 0)$ and there is no change in the shape and size of its orbit $(\Delta p = 0, \Delta e = 0)$. The orbital plane does not turn because the perturbing acceleration W normal to this plane is equal to *zero*. The orbit does not change its size and shape because the total energy of the SC moving along a closed curve in a conservative field can not change and also the central force can not change the momentum of the SC (p and e are the constants, see (1.20) and (1.21)).

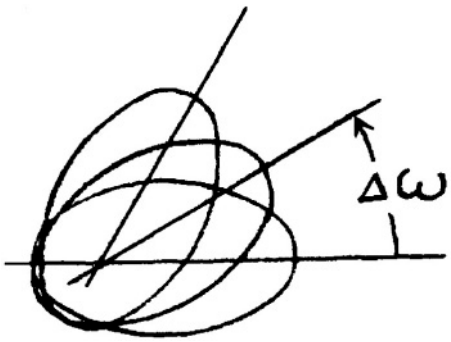

Figure 7.4. Precession of the trajectory of a SC following a step-function approximation

Let us write a formula for the variation of ω using (5.30) and (7.16)

$$\Delta\omega = \int_0^{2\pi} \frac{\varepsilon}{e\mu r^2} \cos v du = \frac{\varepsilon}{e\mu p^2} \int_0^{2\pi} (1 + e\cos v)^2 \cos v dv = \frac{2\pi\varepsilon}{\mu p^2}. \quad (7.20)$$

The angle ω, which defines the angular position of perigee with respect to the ascending node, is the single perturbed osculating element which varies. Its variation $\Delta\omega$ is constant in each orbit and related to the rotation of Laplace's vector, which is directed along the line of apsides. The trajectory of SC motion under our assumption that the changing of osculating element values follows a step-function has the form shown in Fig.7.4.

7.5 Precession of the Inclined Orbit

If the perturbing acceleration contains components caused by noncentrality of the gravitational field, motion occurs which is called *precession of the orbit.* Let us consider a case arising due to the oblateness of the Earth, which provides an example of such precession.

Let the inclination of the orbital plane satisfy the inequality $i \neq 0$. Then, substituting the magnitudes of S, T, W taken from (7.12), (7.14), and (7.15) into system (5.30), and integrating them using appropriate assumptions (in particular, keeping the osculating elements as constants during integration over one revolution), gives us the ex-

pressions

$$\Delta\Omega = -\frac{\varepsilon \cos i}{\mu p^2} 2\pi, \quad \Delta i = 0, \quad \Delta p = 0,$$
$$\Delta e = 0, \quad \Delta\omega = \frac{\varepsilon}{\mu p^2}(4 - 5\sin^2 i)\pi. \tag{7.21}$$

Reasons for the invariance of the shape and size of the ellipse ($\Delta p = 0$, $\Delta e = 0$) are the same as those already discussed in Section 7.4, where the variation of the direction to pericenter ($\Delta\omega \neq 0$) was also described. In contrast to the previously discussed cases (braking of a SC in the atmosphere and evolution of the equatorial SC orbit), the orbital plane now rotates ($\Delta\Omega \neq 0$) while preserving its initial orbital inclination ($\Delta i = 0$). Such a rotation of the orbital plane is called *precession.* The use of this term suggests a relationship between the phenomenon considered and the theory of the gyroscope. We will next demonstrate this relationship using qualitative reasoning.

Suppose that a SC moves around the Earth in a circular orbit. In a gedanken (thought) experiment, spread the mass of that SC along its orbit to produce a rotating ring having angular momentum Γ (Fig.7.5). This rotating ring can act like a gyroscope. Let the orbital plane of the ring *AB* make an angle i with the plane of the equator *CD.* Thus, the angular momentum vector Γ is inclined at an angle i with respect to the axis of symmetry of the Earth *NS.*

As was already stated, the spherical Earth can be deemed equivalent to a particle. Therefore, new effects can be expected if the Earth is considered to be extended by an axisymmetric body that forms a belt around the sphere which 'fills it out' to the shape of an ellipsoid of revolution. In Fig.7.5 the resulting extension in cross-section is indicated by shading.

Consider now the gravitational interaction between the 'belt' and the gyroscopic ring. In this connection, (see the diagram) the point *A* of the ring is mainly attracted by area *C* of the belt (force F_A). A similar relation exists with respect to point *B* and area *D* (force F_B). These forces produce a torque *M,* which causes the ring to precess. The vector of angular momentum Γ resultingly moves on a cone preserving a constant angle i. The rotation of plane ΓON caused by the precession, is characterized by $\Delta\Omega$ not equal to 0 while $\Delta i = 0$. Both of these results follow from (7.21).

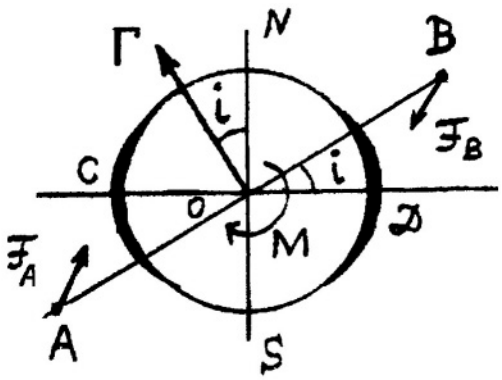

Figure 7.5. Model for interpreting the effect of the Earth's oblateness on the inclined orbit of a SC

Let us now consider the dependence between orbital precession and the angle i. The first expression in (7.21) shows that at $i = \pi/2$ there is no precession of the orbit $(\Delta\Omega = 0)$. This is because, for polar orbits which are located in a meridional plane, due to the axisymmetric model of the Earth considered, the resultant of forces similar to F_A and F_B (Fig.7.5) pass through the origin and no torque is developed.

If now we put $i = 0$ (i.e. motion is executed in the plane of the equator), the first expression in (7.21) gives a maximum value of the displacement $\Delta\Omega = -2\pi\varepsilon/(\mu p^2)$ and this corresponds to the fastest realizable precession. However, this formal result is contradicted by the circumstance that, in this case, the resultant of forces like F_A and F_B pass through the origin so that there is actually no precession. Let us consider this inconsistency in more detail.

7.6 Clarification of 'Inconsistency'

A monotonic increase in the velocity of orbital precession as the orbit of a SC approaches ever closer to the equatorial plane, results from the fact that this precession is measured by $\Delta\Omega$. At $i = 0$ the node Ω disappears as, in this case, the orbital plane does not intersect the plane of the equator but coincides with it. To avoid the complications attending the use of an approach based on spherical geometry we will consider this situation qualitatively (Fig.7.6). Let $\Delta\beta$ represents the dihedral angle between two consecutive positions of the orbital

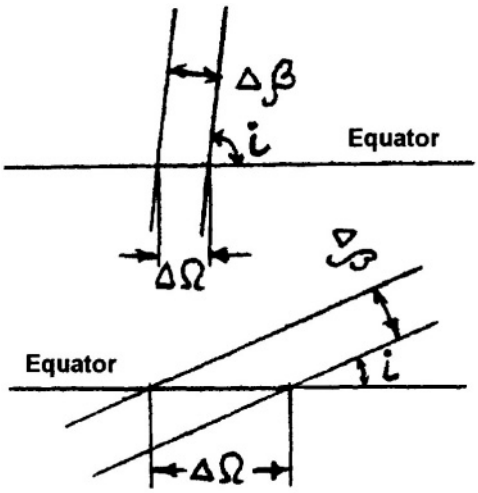

Figure 7.6. Illustration of the orbital plane displacement $\Delta\beta$ as $i \to 0$

plane. With the reduction of angle i, $\Delta\Omega$ measured in the plane of the equator increases while $\Delta\beta$ remains constant. In the planar case at $\Delta\beta = \text{const}$ $\Delta\Omega \to \infty$ as $i \to 0$. On a sphere the limiting value $\Delta\Omega$ is finite.

If we measure the precession by $\Delta\beta$ then, on combining the relation $\Delta\beta = \Delta\Omega \sin i$ with the first formula in (7.21), we obtain

$$\Delta\beta = -\frac{\pi\varepsilon}{\mu p^2} \sin 2i. \tag{7.22}$$

In conformity with common sense, $\Delta\beta = 0$ for polar and equatorial orbits while $\Delta\beta$ reaches a maximum value for $i = \pi/4$.

At $i = 0$

$$\Delta\omega = \frac{4\pi\varepsilon}{\mu p^2} \tag{7.23}$$

which obviously contradicts formula (7.20) where the same element has half this magnitude. However, the mismatch obtained is only apparent. The angle ω is measured from the node Ω. In the case of an equatorial orbit, the angles describing the longitude of an ascending node and the direction to pericenter lie in the same plane and consequently can be added algebraically. If we chose in this plane any direction fixed in inertial space, then the change in the direction to pericenter relative to the fixed direction is determined by the sum

$\Delta\Omega+\Delta\omega$. The change in the direction to pericenter in inertial space is equal, in the already considered case (7.21) to $\varepsilon[-2\pi/(\mu p^2)+4\pi/(\mu p^2)]$ while in the case discussed in (7.18) and (7.20) it is $\varepsilon[0+2\pi/(\mu p^2)]$. Therefore, both solutions give similar motion in inertial space. A difference between the two values $\Delta\omega$ relates to the fact that in the first case the value $\Delta\omega$ was obtained in a reference system fixed in inertial space $(\Delta\Omega=0)$, and in the second case it was obtained with respect to a rotating system $(\Delta\Omega=-2\pi\varepsilon/(\mu p^2))$. (In the first case we put $i=0$ and in the second case we assumed that $i\to 0$.)

It is now a proper time to return to (7.17). It is instructive to pay attention to the singularity of equation for $d\Omega/du$ from (5.30) at $i=0$, which was not considered before. If the limit $i\to 0$ is approached in this equation for $d\Omega/du$ and also in the expression (7.15) for W, we again reach the result obtained above for $i=0$ (i.e. through using formulae for the arbitrary values of i and proceeding to the limit $i\to 0$). This implies that (7.18) corresponds to the case of a reference system fixed in inertial space.

Overall, it can be said that the Earth's oblateness results in a precession of the orbital plane of a SC around the axis of symmetry of the Earth, as well as rotation of the orbit in this plane. Depending on the inclination of the orbit different cases can arise.

7.7 Orbits with Specific Inclinations

The precession of the orbital plane is an essential feature of spaceflight dynamics. At the orbital inclination typical of the majority of Russian SC, the value of precession is of the order of 4° per day (equivalent to approximately 400 km at the equator). This effect has both negative and positive consequences. We will consider here a positive consequence for star map creation. Let a SC be oriented so that the optical axis of an on-board telescope is always directed to the zenith. Consequently, a photograph of the sky around the zenith can be made. Over one orbit the SC has the capability to photograph the stars located in a zone forming a narrow ring which is symmetrical relative to the orbital plane. If precession of the orbit did not exist, the same stars would be photographed during every orbit. However, at the above mentioned rate of precession, all parts of the sky are photographed

within 45 days without the expenditure of thruster fuel to turn the orbital plane. Later it will be shown that such a manoeuvre would be 'expensive' in terms of the expenditure of energy.

The last equality in (7.21) indicates how the direction to pericenter varies. This change depends on the inclination of the orbital plane i. For a value of i satisfying the condition $4 - 5\sin^2 i = 0$ the magnitude $\Delta\omega = 0$. At this value of i, the orbital plane precesses but the line of apsides does not turn in this plane. The value of i satisfying this condition is 63.4°.

Orbits with this inclination are widely used for communications SC. In particular, the series of Russian communications SC *Molnia* uses $i = 63.4°$ with typical apogee and perigee altitudes of about 39 400 km and 1000 km respectively; eccentricity 0.72 and argument of perigee 270°. Such a highly elliptical orbit has a period of revolution of 12 hours. Due to this choice of parameters the SC remains above the northern hemisphere for approximately 11 hours. Once a day, in every two revolutions the SC occupies a position in the orbit with the same sub-satellite point. The duration of an uninterrupted communications session is about 8 hours. Three satellites launched into orbits with Ω differing by 120° ensures a continuous communications service over the territory of Russia. About one hundred and forty *Molnia* type SC were launched [30] in this series.

Elliptical orbits with the same inclination (63.4°) but twice the period of revolution (i.e. 24 hours); apogee and perigee altitude about 46 300 km and 25 300 km, respectively, eccentricity 0.25, argument of perigee 270° are used by the communications SC series *Tundra.*

Another orbit widely used by designers of space systems is the *sun-synchronous (SSO)* or *sun-stationary* orbit. Again, the oblateness of the Earth is a key factor in providing the useful evolution of this orbit. The inclination is such that the rate of evolution of the ascending node is equal and opposite to the angular rate of the Earth's radius-vector with respect to the Sun in its yearly revolution. The angular position of the Sun relative to the orbital plane thus remains the same for a long time. This type of orbit is suitable for meteorological applications and for Earth observations since the same conditions of illumination are maintained from one revolution to another. Meteorological SC in the *Meteor* series (since *Meteor-28* which was orbited in 1977) have employed sun-synchronous orbits. SC in the *LANDSAT* series (the first

launched in 1972) and in the *SPOT* series (the first launched in 1986), also have used sun-synchronous orbits for long term remote sensing. From the first formula in (7.21) it can be seen that the inclination required depends on the altitude of the orbit. For orbits with altitude about 650 km the inclination should be near 98°. For higher orbits with altitudes about 1 500 km, the required inclination increases up to approximately 102°.

In Section 2.5 it is mentioned that a specific circular orbit (called a *geo-stationary* orbit) exists for which the period of SC revolution is equal to the Earth's spin rotation period and the sub-SC point is consequently 'fixed' on the Earth's surface. Further to our analysis of the effect of the non sphericity of the Earth on the dynamics of a SC (illustrated in Fig.7.5), we will consider next in a similar manner the effect of terrestrial nonsphericity on a SC in geo-stationary orbit. The oblateness of the Earth, which is represented by the second zonal harmonic J_2 in the expansion of the potential of the gravity field, does not result in asymmetry in the Earth's equatorial plane. The second tesseral harmonic J_{22}, however, confers on the equatorial cross section a slight ellipticity. In consequence of these two harmonics, the Earth is approximated by a ellipsoid with the following principal semiaxes: two equatorial (6 378 266.30 m and 6 378 053.70 m), and polar (6 356 774.72 m) [19]. The difference between the equatorial semiaxes is less by, almost, an order of two than the difference between the equatorial and polar semiaxes. Of course, there are a number of other representations of the figure of the Earth and we quoted here values for the semiaxes adopted in but one of these models.

The ellipticity of the Earth's equatorial cross section can be indicated diagrammatically (as in the case of polar oblateness in Fig.7.5) by shading. The nonuniformity of the Earth associated with the second tesseral harmonic J_{22} can be represented by a belt around the axis lying in the equatorial plane, thereby creating the maximum axis of the Earth's ellipsoid. In Fig.7.7 and in Fig.7.8 this belt is indicated by shading at *A* and *B*. We consider *A* and *B* to be, in effect, attracting particles. Due to the symmetry of the gravitational field in the equatorial plane, four equilibrium positions in geo-stationary orbit exist designated in Fig.7.7 and Fig.7.8 by the numbers *1*, *2*, *3*, *4*. Consider first the stability of point *1* (Fig.7.7). In a reference system spinning with the Earth, equilibrium is determined by equating

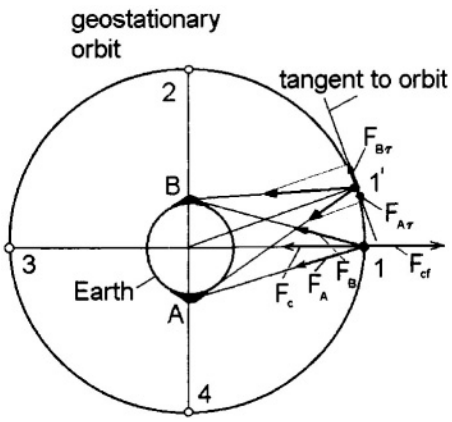

Figure 7.7. Illustration of the instability of equilibrium position 1 in geo-stationary orbit

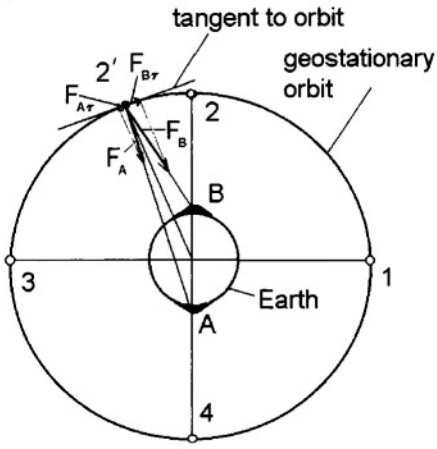

Figure 7.8. Illustration of the stability of equilibrium position 2 in geo-stationary orbit

attracting and inertial forces applied to a SC. In the case illustrated, the null-resultant of the centrifugal force $\mathbf{F}_{cf}$ directed outside the center of spinning and three representative attracting forces $\mathbf{F}_c$, $\mathbf{F}_A$, and $\mathbf{F}_B$ provide SC equilibrium. Now suppose the SC to be displaced along its orbit to the position *1'*. The projection of the resultant on the line joining the center of mass of the SC and the center of the Earth is equal to zero but the projection of the resultant on the tangent to the orbit is not equal to zero. This is because point *1'* is closer to *B* than to *A*. Therefore, the magnitude of $\mathbf{F}_b$ is greater than the magnitude of $\mathbf{F}_A$ and, also, the angle between the direction to the Earth's center and to *B* is greater than the angle between the direction to the Earth's center and to *A* (Fig.7.7). Thus, the projection $F_{B\tau}$ is greater then the projection $F_{A\tau}$. Their resultant causes the SC situated at *1'* to move away from equilibrium position 1. Consequently, position *1* is unstable (and similarly position *3*).

Consider next the situation in the vicinity of position 2 (see Fig.7.8). For simplicity, the attracting forces $\mathbf{F}_A$ and $\mathbf{F}_B$ are only shown at point *2'*, which is slightly displaced along the orbit from point *2*. Stability is provided by the projection of the resultant of these forces on a tangent directed to equilibrium position *2* which causes the SC at *2'* to move towards it. A similar situation pertains at *4*. Hence, equilibrium po-

sitions *2* and *4* located at the intersection of the geo-stationary orbit and the extension of the minimum axis of the Earth's ellipsoid are stable. These two stable points, which are situated at 75° E and 255° E geographic longitude, are called *gravitational valleys* and they are used for providing stable locations for geostationary SC. For instance, the Russian Geo-stationary Operational Meteorological Satellite *GOMS* which was launched in 1994 utilizes the 76°50′ E stationary point at an altitude of about 36 000 km. Positioning of SC at other points along the geostationary orbit requires station-keeping manoeuvres to maintain a given location.

Chapter 8

SC Motion in the Field of Two Attracting Centers

We again consider here the planetary form of equations in the framework of the three-body problem. In contrast to previous considerations, however, we show how quantitative estimates can be obtained through analysis of the equations written using osculating elements.

To establish the problem, assume that two celestial bodies (say the Sun and the Earth) in the gravitational field of which the SC moves are particles (spherically uniform bodies), and that there is no atmosphere around them. Further suppose that the SC moves inside the Earth's *sphere of influence* and that, consequently, the Sun develops a weak acceleration with components S, T, W.

The equation (4.6) of SC motion in a geocentric reference system contains the expression

$$\mathbf{b} = -k^2 m_1 \left(\frac{\mathbf{s}}{s^3} - \frac{\mathbf{R}}{R^3} \right)$$

representing perturbation acceleration. By virtue of equality (4.11) this can be rewritten as follows

$$\mathbf{b} = -\frac{k^2 m_1}{R^2} \cdot \frac{r}{R} \left(\frac{\mathbf{r}}{r} - 3 \frac{\mathbf{Rr}}{Rr} \cdot \frac{\mathbf{R}}{R} \right). \tag{8.1}$$

Let us now obtain expressions for S, T and W. We assume that the motion is planar (i.e. that the trajectory of the SC lies in the plane of the Earth's orbit). Let the SC which has a negligibly small mass m

move along an elliptical orbit around the Earth, which has mass m_2 (Fig.8.1). The Sun with mass m_1 is so far away that the disturbances developed by it at m can be considered to be small. As is well known,

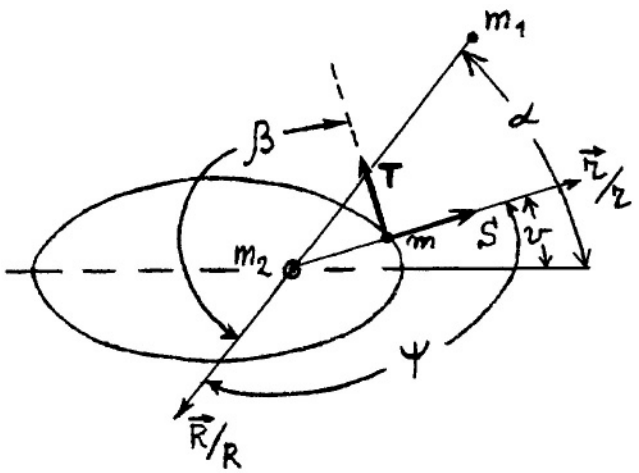

Figure 8.1. Motion of a SC in the field of two attracting centers

the angle α which determines the direction to the Sun changes by approximately one angular degree in one day. We can regard such a change in α as unimportant when time intervals comparable with the duration of SC revolution around the Earth are considered. In what follows, we will assume that α is constant. Further, for simplicity but without loosing generality, we will confine our analysis to the case where the Sun, the Earth and the SC are co-planar.

By virtue of our assumption of the planar character of the motion we can write

$$W = 0. \tag{8.2}$$

The perturbing acceleration S is the projection of vector $\mathbf{b}$ on the direction $\mathbf{r}$

$$S = \mathbf{b} \cdot \frac{\mathbf{r}}{r} = -\frac{k^2 m_1}{R^2} \cdot \frac{r}{R}(1 - 3\cos^2\psi)$$

which can be rewritten using the expression $\psi - v + \alpha = \pi$ obtained from Fig.8.1

$$S = -\frac{k^2 m_1}{R^2} \cdot \frac{r}{R}[1 - 3\cos^2(\alpha - v)]. \tag{8.3}$$

As the perturbation acceleration T is the projection of $\mathbf{b}$ on the direction of a unit vector $\mathbf{e}_T$ normal to $\mathbf{r}$

$$T = \mathbf{b}\mathbf{e}_T = -\frac{k^2 m_1}{R^2} \cdot \frac{r}{R}(0 - 3\cos\psi\cos\beta).$$

After simple trigonometric transformations using the relation $\beta = \alpha - \upsilon + \pi/2$ from Fig.8.1, we get

$$T = \frac{3k^2 m_1}{R^2} \cdot \frac{r}{R}\cos(\alpha - \upsilon)\sin(\alpha - \upsilon). \tag{8.4}$$

The values obtained W (8.2), S (8.3) and T (8.4) allow us to find the variations of all osculating elements after one revolution of the SC around the Earth using the techniques discussed in Section 6.1. The equality (8.2) at once gives

$$\Delta\Omega = 0, \qquad \Delta i = 0. \tag{8.5}$$

If now, instead of parameter p, we use the related semi-major axis of the ellipse $a = p/(1 - e^2)$, it is easy to show that

$$\Delta a = 0. \tag{8.6}$$

The latter equality can be explained as follows. Since the Sun is assumed to be fixed $(\alpha = \text{const})$, the joint gravitational field of the two attracting centers considered is conservative. In this case, motion along a closed elliptical trajectory cannot change the energy of the SC and, consequently, semi-major axis a is constant since it is an indicator of energy (see the comments following formula (2.1)).

Bulky calculations allow us to obtain

$$\Delta e = -15\pi\frac{m_1}{m_2}\left(\frac{a}{R}\right)^3 e\sqrt{1 - e^2}\cos\alpha\sin\alpha. \tag{8.7}$$

Thus, as a result of the Sun's perturbation, the orbit does not change either its plane of motion or its size (semi-major axis) but it changes its shape (eccentricity). This effect does not appear in the case of a circular orbit (e = 0), or for strongly elongated orbits that approach being parabolic (e = 1). The greatest effect, as follows from relation (8.7), is particular to orbits with $e = 1/\sqrt{2} \approx 0.7$, that is orbits with significant ellipticity.

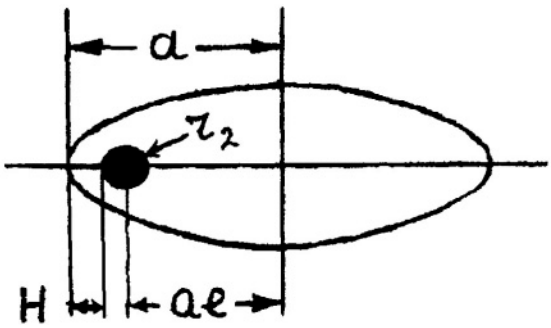

Figure 8.2. Plot to calculate using formula (8.8) a critical altitude H below which a SC falls to Earth

A change in the ellipticity of its orbit can result in a SC colliding with the Earth. Let us consider this process. As illustrated in Fig.8.2 the altitude H of the SC's orbit over the surface of the spherical Earth with radius r_2 is equal to

$$H = a(1 - e) - r_2. \tag{8.8}$$

If the increment of altitude

$$\Delta H = -a\Delta e \tag{8.9}$$

has the proper sign and satisfies the inequality

$$|\Delta H| > H,$$

then the SC impacts the Earth (we exclude here from consideration the braking effect of the Earth's atmosphere).

It should be noted that the character of this process is that altitude increases or decreases depending on the sign of Δe (see (8.7)) which, in turn, depends on the angle α which describes the angular position of the Earth with respect to the Sun.

If we do not assume that the motion is planar, the component W of the perturbing acceleration is not equal to zero. This makes the investigation more complicated but the general conclusion does not vary.

Detailed analysis of the perturbing effect of a second attracting body was performed by Lidov* who showed, as a dramatic example, that if the Moon were placed in an orbit at a 90° inclination to the ecliptic plane, it would fall to Earth within four and a half years†. This example provides insight as to why the planets in the Solar System occupy a plane. (The highly inclined orbits of the satellites of Uranus are a consequence of the oblateness of the planet).

A practical example of the effect is provided by the unexpected reentry of the Soviet SC *Luna-3* in 1959 (which took the first picture of the reverse side of the Moon). Its geocentric orbit was inclined at 80° relative to the ecliptic plane and it initially had altitude of apogee 480 000 km and altitude of perigee 47 000 km. After 11 orbits around the Earth this SC re-entered the Earth's atmosphere due to the perturbing effects exerted by the Moon and the Sun [19].

Another example is provided by the reduced life time of the American SC *Explorer-6* from an initially expected 20 to 2 years. This SC executed a highly elliptical orbit.However, due to the perturbation of the Moon's attraction, its perigee varied between 250 and 160 km every three months (angle α in (8.7)), thereby causing the orbit to rapidly degrade.

*Lidov, Mikhail L'vovich (1926–1993), Principal Researcher at the Keldysh Institute of Applied Mathematics of the Russian Academy of Sciences and a professor at the Lomonosov Moscow State University. He made estimates of the Earth's atmospheric density through processing the tracking data of *Sputnik-1*; made an input to the three-body problem and performed multiple applied investigations in the field of spaceflight dynamics under the aegis of the Keldysh-Okhotsimsky Scientific School of *Applied Celestial Mechanics.*

†M.L.Lidov. The Evolution of Orbits of Artificial Satellites of Planets under the Action of Gravitational Perturbations of External Bodies. *Planet. Space Sci.,* 1962, Vol.9, pp.719–759.

Chapter 9

Elements of SC Manoeuvring Theory

9.1 Statement of the Problem

Next we consider SC manoeuvres, that is the making of deliberate changes in a SC trajectory. Such manoeuvres are necessary in cases such as the following:

- to correct a SC's interplanetary trajectory on its approach to a target plane;
- to change the parameters of the orbit of a SC around a planet;
- to achieve rendezvous with another SC;
- to use a landing pulse to transfer a SC to a landing trajectory.

Manoeuvres are achieved by activating thrusters to vary the velocity vector of motion. The 'cost' of a pulse that enables particular orbital elements to change, is evaluated in terms of propellant consumption. Here not absolute, but relative, propellant consumption with respect, for example, to the total mass of the SC is of interest. Usually the relative propellant consumption is estimated in terms of *characteristic velocity* or ΔV*-budget* (that is the velocity imparted by the thrusters

to the SC in field-free space). In this case the *formula of Tsiolkovsky** is valid

$$V_x = w \ln \frac{m_0}{m} \tag{9.1}$$

where V_x is the current value of the characteristic velocity; w is the exhaust velocity; m_0 is the initial mass of the wet SC and m is the current value of SC mass.

It is obvious that the relative mass m_T of the fuel spent at a particular moment of time T is determined by the expression

$$\frac{m_T}{m_0} = \frac{m_0 - m}{m_0}. \tag{9.2}$$

From (9.1) and (9.2) we obtain

$$V_x = w \ln \left(1 - \frac{m_T}{m_0}\right)^{-1}$$

This relationship allows us, instead of speaking about fuel expenditure m_T, to speak about the quantity of characteristic velocity expended. Also, instead of the content of the fuel tanks we can refer to 'a reserve of characteristic velocity' etc. This way of estimating the propellant reserve is appropriate as the manoeuvring of a SC involves changing its velocity. Consequently it is more informative to say, for example, that a SC still has a reserve 100 m/s of characteristic velocity than to say that it still has 200 kg of fuel available.

Let us assume that a change in the velocity vector caused by the thruster takes place during a short time interval t_n. This interval t_n is understood to be short in comparison with some typical time scale of the motion, for example, the period of SC revolution. The shortness of t_n allows us to consider the manoeuvre to be produced by the application of an instantaneous pulse. This assumption is realized in practice since the duration of thruster activity is usually measured in terms of tens of seconds to minutes, while the SC revolution period is not less than one and half hours. (The relatively long boosts required to achieve interplanetary velocity or to land successfully on the surface

*Tsiolkovsky, Konstantin Eduardovitch (1857–1935), Russian scientist and inventor. He provided the first proof that it is possible to use a rocket for interplanetary missions and is regarded as the founder of present day cosmonautics.

of an atmosphereless planet are excluded here from consideration since these cases are more pertinent to studies in rocket dynamics.)

The instantaneous character of the SC velocity change has the consequence that, in the equations of motion, the change of velocity is not related to a change of coordinates, of the true anomaly υ or of the argument of latitude u.

Consider now the equations of SC motion in osculating elements, choosing time t as the independent variable (5.26). Integrate this system assuming t_n to be short and the characteristic velocity change imparted to the SC to be small. These assumptions allow us to suppose that, during the interval of thruster firing t_n, the osculating elements do not change continuously but, rather, jump at the moment of thruster switch off. Thus, for example, the change of element Ω is determined by the expression

$$\Delta\Omega = \frac{r}{\sqrt{\mu p}} \frac{\sin u}{\sin i} \int_0^{t_n} W\,dt = \frac{r}{\sqrt{\mu p}} \frac{\sin u}{\sin i} \Delta V_W \tag{9.3}$$

as per the first equation of (5.26), where the symbol Δ denotes a linear increment. Since now the perturbing acceleration W (more exactly a component of the acceleration associated with thruster firing which perturbs the initial trajectory) is small and t_n is short, the change in the SC velocity vector components ΔV_W is small and can be represented by

$$\Delta V_W = \int_0^{t_n} W\,dt. \tag{9.4}$$

A small ΔV_W results in a small change in $\Delta\Omega$. As already mentioned above, we consider a control pulse ΔV_W imparted to the SC to be an instantaneous one. The reaction of the SC to this pulse is an instantaneous change in the angle Ω by an amount $\Delta\Omega$.

Completely similar reasoning can be used in considering changes in the other osculating elements i, p, e, ω, τ. Thus, as in the case of ΔV_W, the non-zero components ΔV_S and ΔV_T of instantaneous change in the velocity vector associated with the acceleration components S and T can be represented in a form similar to that used in (9.4). The full vector of instantaneous change of SC velocity $\Delta\mathbf{V}$ with components $(\Delta V_S, \Delta V_T, \Delta V_W)$ produces an instantaneous change in

the individual osculating elements given by the increments $\Delta\Omega$, Δi, Δp, Δe, $\Delta\omega$, $\Delta\tau$. We will confine further consideration of SC manoeuvres to those geometrical parameters of SC motion representing the position of the orbital plane (Ω, i), and the size and shape of the orbit (p, e).

The smallness of both the magnitude of $\Delta\mathbf{V}$ and the increments of osculating elements $\Delta\Omega$, Δi, Δp, Δe, $\Delta\omega$, $\Delta\tau$, allows us to use the *superposition principle* and write

$$\begin{aligned}
\Delta\Omega &= A_{11}\Delta V_S + A_{12}\Delta V_T + A_{13}\Delta V_W,\\
\Delta i &= A_{21}\Delta V_S + A_{22}\Delta V_T + A_{23}\Delta V_W,\\
\Delta p &= A_{31}\Delta V_S + A_{32}\Delta V_T + A_{33}\Delta V_W,\\
\Delta e &= A_{41}\Delta V_S + A_{42}\Delta V_T + A_{43}\Delta V_W.
\end{aligned} \tag{9.5}$$

The coefficients of this linear transformation $A_{11}, \ldots, A_{43}$ have the following notations

$$\begin{aligned}
A_{11} &= \frac{\partial\Omega}{\partial V_S}, & A_{12} &= \frac{\partial\Omega}{\partial V_T}, & A_{13} &= \frac{\partial\Omega}{\partial V_W},\\
A_{21} &= \frac{\partial i}{\partial V_S}, & A_{22} &= \frac{\partial i}{\partial V_T}, & A_{23} &= \frac{\partial i}{\partial V_W},\\
A_{31} &= \frac{\partial p}{\partial V_S}, & A_{32} &= \frac{\partial p}{\partial V_T}, & A_{33} &= \frac{\partial p}{\partial V_W},\\
A_{41} &= \frac{\partial e}{\partial V_S}, & A_{42} &= \frac{\partial e}{\partial V_T}, & A_{43} &= \frac{\partial e}{\partial V_W}.
\end{aligned} \tag{9.6}$$

To obtain, for example, the value $\dfrac{\partial\Omega}{\partial V_S}$ let $\Delta V_S = \Delta V_T = 0$ then we can write

$$\begin{aligned}
\frac{\partial\Omega}{\partial V_W} &= \frac{\Delta\Omega}{\Delta V_W}, & \frac{\partial i}{\partial V_W} &= \frac{\Delta i}{\Delta V_W},\\
\frac{\partial p}{\partial V_W} &= \frac{\Delta p}{\Delta V_W}, & \frac{\partial e}{\partial V_W} &= \frac{\Delta e}{\Delta V_W}.
\end{aligned} \tag{9.7}$$

The relation between an increment of an osculating element and a component of the vector $\Delta\mathbf{V}$ is given by formulae such as (9.3) from which it, in particular, follows that

$$\frac{\Delta\Omega}{\Delta V_W} = \frac{r}{\sqrt{\mu p}}\frac{\sin u}{\sin i}. \tag{9.8}$$

It also follows from the first equality in (9.7) and from (9.8) that A_{13} coincides with the coefficient of the perturbing acceleration W in the

first equation of (5.26). Other values of the coefficients in (9.6) can be similarly obtained.

A matrix of these coefficients

$$A = \begin{pmatrix} 0 & 0 & \frac{\partial \Omega}{\partial V_W} \\ 0 & 0 & \frac{\partial i}{\partial V_W} \\ 0 & \frac{\partial p}{\partial V_T} & 0 \\ \frac{\partial e}{\partial V_S} & \frac{\partial e}{\partial V_T} & 0 \end{pmatrix} \tag{9.9}$$

giving a complete representation of the properties intrinsic to the manoeuvre considered can be written in the explicit form

$$A = \begin{pmatrix} 0 & 0 & \frac{r}{\sqrt{\mu p}} \frac{\sin u}{\sin i} \\ 0 & 0 & \frac{r}{\sqrt{\mu p}} \cos u \\ 0 & 2r\sqrt{\frac{p}{\mu}} & 0 \\ \sqrt{\frac{p}{\mu}} \sin v & \sqrt{\frac{p}{\mu}} \left(1 + \frac{r}{p}\right) \cos v + \frac{er}{\sqrt{\mu p}} & 0 \end{pmatrix} \tag{9.10}$$

which is consequently named the *matrix of manoeuvre.* In particular, for a circular orbit $(e = 0, p = r)$, the matrix of manoeuvre has the form

$$A = \begin{pmatrix} 0 & 0 & \frac{r}{\sqrt{\mu p}} \frac{\sin u}{\sin i} \\ 0 & 0 & \frac{r}{\sqrt{\mu p}} \cos u \\ 0 & 2r\sqrt{\frac{p}{\mu}} & 0 \\ \sqrt{\frac{p}{\mu}} \sin v & 2\sqrt{\frac{p}{\mu}} \cos v & 0 \end{pmatrix}. \tag{9.11}$$

An analysis of certain manoeuvres (below) made initially for the case of a circular orbit, allows us to derive a number of analytical expressions that provide estimations which are also valid for elliptical orbit evolution.

9.2 Manoeuvre of Changing the Orbital Plane Orientation

This manoeuvre involves changing Ω and i. Therefore, we may consider now the two first rows of matrix A. Using the expressions (9.5) and (9.11) we get

$$\Delta\Omega = \frac{r}{\sqrt{\mu p}} \frac{\sin u}{\sin i} \Delta V_W, \tag{9.12}$$

$$\Delta i = \frac{r}{\sqrt{\mu p}} \cos u \Delta V_W. \tag{9.13}$$

Recall that for a circular orbit of arbitrary radius the equalities

$$r = p = a, \quad e = 0, \quad V_c = \sqrt{\frac{\mu}{r}}, \quad T_{rev} = 2\pi \frac{r^{3/2}}{\sqrt{\mu}} \tag{9.14}$$

are valid (see Section 1.5 and Section 2.5, formulae (1.25) and (2.2) respectively). Using equalities (9.14), we now rewrite the relations (9.12) and (9.13) as follows

$$\Delta\Omega = \frac{1}{V_c} \frac{\sin u}{\sin i} \Delta V_W, \tag{9.15}$$

$$\Delta i = \frac{1}{V_c} \cos u \Delta V_W. \tag{9.16}$$

These formulae allow us to specify two important aspects of manoeuvres (namely their efficiency and their 'cost'). The efficiency of a manoeuvre is strongly associated with the coordinate u, which determines that point along a SC trajectory at which the manoeuvre 'instant pulse' is injected. Let us consider first this factor.

At $u = 0$ the longitude of the ascending node Ω_α can not be changed (see (9.15)), while the inclination i undergoes its greatest change (see (9.16)). At $u = \pi/2$ the inverse of this situation occurs. These effects can be explained as follows.

Consider first the case $u = 0$. The appropriate scheme of manoeuvre is illustrated in Fig.9.1. The initial position of the orbital plane is represented by the straight line AA which passes through the ascending node Ω_α. Let the SC have a velocity $\mathbf{V}$ when crossing the equatorial plane and, at this moment, allow an additional velocity $\Delta\mathbf{V}_W$ to be injected perpendicular to the initial orbital plane. The

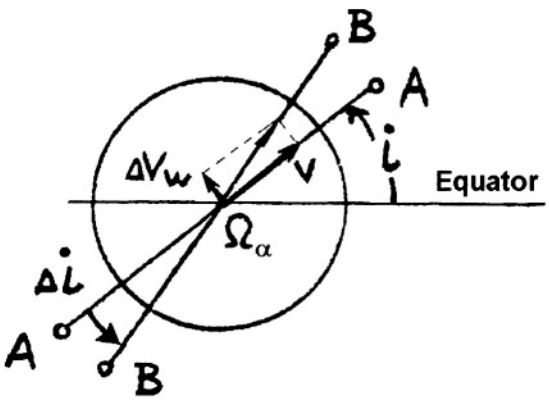

Figure 9.1. Manoeuvre performed at $u = 0$ to change the inclination of the orbital plane of a SC

resultant vector of SC velocity is the sum of $\mathbf{V}$ and $\Delta\mathbf{V}_W$ and it defines the position of the resulting new orbital plane, which is represented by the line *BB*. The ascending node Ω_α is not shifted (i.e. the equality $\Delta\Omega = 0$ is satisfied. The orbit inclination i is changed so that $\Delta i \neq 0$.

Consider next the case $u = \pi/2$ and that the pulse of velocity $\Delta\mathbf{V}_W$ is imparted to the SC at this point in the orbit. For simplicity, we assume that $i = \pi/2$ (i.e. we consider the initial orbit of the SC to be in the plane of the meridian). When $u = \pi/2$, the SC is situated over the North Pole (designated by *N* in Fig.9.2). The additional velocity pulse $\Delta\mathbf{V}_W$ is injected perpendicular to the SC's orbital plane *AA*. The resultant vector of SC velocity defines the position of the new orbital plane which is represented by the line *BB*. The intersection of the two orbital planes occurs in a straight line connecting the poles of the Earth (the line of nodes). The new orbit again lies in the plane of the meridian, so that again the angle $i = \pi/2$ (i.e. $\Delta i = 0$). However, the ascending node Ω_α moves along the equator to a new position corresponding to the new meridian plane in which the SC orbits. In Fig.9.2 this angular displacement is designated by $\Delta\Omega \neq 0$.

Consider now a second important factor to be taken into account in choosing a manoeuvre strategy, namely the associated 'cost'. In the beginning of this Chapter it was shown that characteristic velocity provides an appropriate measure of expenditure.

If a SC is launched into an equatorial orbit from either Russia or

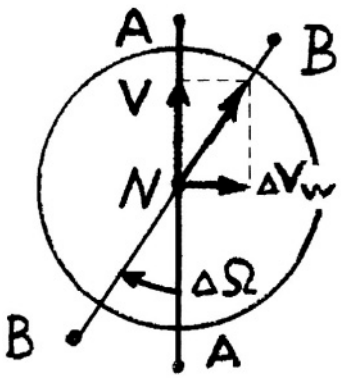

Figure 9.2. Manoeuvre to change the argument of the ascending node of a SC's orbital plane with $i = \pi/2$ performed at $u = \pi/2$

Kazakhstan, manoeuvres must be performed to change the inclination of its orbital plane because the angle of inclination cannot be less than the geographic latitude at the point of launch and the inclination of SC orbits launched from either of these locations is not less than, roughly, one radian (the latitude of the more southerly cosmodrome (Baikonur) is 46° N). Formula (9.16) shows that $\Delta V_W = V_c$ is required to change the orbital plane inclination by one radian, even if the manoeuvre is executed at the optimal point with $u = 0$. The manoeuvre required in this case can not be regarded as minor. Due to inherent nonlinearity, formula (9.16) does not provide quantitatively correct results for Δi but demonstrates only a trend. However, it can be deduced that to transfer a SC from LEO (low Earth orbit) to an equatorial orbit, the characteristic velocity required is of the same order of magnitude as that expended to launch the SC into its initial orbit. Thus, to achieve such an orbit transfer using on-board thrusters is impractical since it would require a booster with equivalent capability to that of the launcher. In this respect, Russian astronautics operates under more severe constraints than American because the most southerly site available to launch Russian rockets is situated at *Baikonur* in Kazakhstan (46° N, 63° E) while the American Kennedy Space Flight Center is located on *Cape Canaveral* at 28° N, 80° W. A new Russian launch facility established at *Svobodny* after the dissolution of the former Soviet Union (51° N, 128° E), although more favorably sited than *Plesetsk* (63° N, 41° E), can only partially solve

this problem.

If we confine the discussion to small changes in inclination then, to turn the orbital plane by one degree at the equator, the magnitude of ΔV_W should be 140 m/s (formula (9.16)). Formula (9.15), which is similar to (9.16), shows that a high 'cost' (measured again in terms of characteristic velocity), is also involved in changing the argument of the ascending node Ω_α. The large increase in $\Delta\Omega$ as $i \to 0$ is a mathematical artifact and does not concern the reality of the process. This effect we already met when considering orbital precession due to the oblateness of the Earth (see Section 7.6). When planning space missions it is important, because of the cost factor, to avoid situations involving the manoeuvre of changing the orbital plane orientation.

An example of avoiding the re-orientation orbit manoeuvre is provided by history's first international spaceflight, when an American SC *Apollo* with a three man crew docked with the Soviet two man SC *Soyuz* in July 1975. The launches of these two SC were not carried out simultaneously in both countries since, in this case, the planes of their orbits would not have coincided. Rather, the SC *Soyuz* was orbited from Baikonur first. Then, when the rotating Earth had carried the SC *Apollo* located at Cape Canaveral to the orbital plane of the SC *Soyuz* (which to a first approximation was fixed in inertial space), *Apollo* was launched into the orbital plane of *Soyuz*. Due to this technique the precise rendezvous process could, thereafter, be realized without any requirement to change either the orbital plane inclination or the argument of the ascending node.

We will next consider manoeuvring in an orbital plane.

9.3 In-plane Manoeuvres

Consider now those parameters of a SC trajectory describing its size (parameter p) and shape (eccentricity e). The third row of the matrix of manoeuvre A in (9.11) shows that a change in p is consequent only on the component ΔV_T, noting that $\partial p/\partial V_T = 2r(p/\mu)^{1/2}$. If we assume that the orbit is circular it can be shown using (9.14) that

$$\frac{\partial p}{\partial V_T} = \frac{T_{rev}}{\pi}. \tag{9.17}$$

A typical SC revolution period $T_{rev} \approx 90$ min so that

$$\frac{\partial p}{\partial V_T} \approx \frac{90 \cdot 60}{\pi} \approx 1.7 \frac{\text{km}}{\text{m/s}}. \tag{9.18}$$

Thus, an instant, velocity pulse of 1 m/s increases the orbital parameter p by 1.7 km. In fact, a pulse of approximately 140 m/s, which would suffice only for a hardly noticeable turn in the orbital plane of 1°, enables the altitude of the orbit to be changed by approximately 240 km. This is a significant amount if we take into account that the altitude of a SC in LEO is typically only several hundred kilometers above the surface of the Earth.

Consider next the change in shape of the orbit produced by varying the parameter e starting from the case of a circular orbit. As follows from the lowest line of the matrix of manoeuvre A (9.11) corresponding to transformation (9.5), a change in e is related to both a pulse ΔV_T and a pulse ΔV_S. The total effect is their superposition. We will next study the effects of these pulses individually.

9.3.1 Effect of a Tangent Pulse ΔV_T

The change Δe_T in eccentricity produced by a pulse ΔV_T is described by the corresponding element of the matrix of manoeuvre and can be calculated using

$$\Delta e_T = 2\sqrt{\frac{p}{\mu}} \cos v \, \Delta V_T. \tag{9.19}$$

A difficulty that arises in attempting to utilize this pulse relates to the associated difficulty of recognizing the true anomaly v. This is measured from the direction to pericenter but, for a SC moving in a circular orbit, this direction is not defined. A way to overcome the problem is to change the shape of the orbit using a two-step process. In the first step, an infinitesimal pulse immediately precedes the main pulse. As it is infinitesimal, the first pulse deforms the circular orbit into an, infinitely close, elliptical orbit. This elliptical orbit, irrespective of its value of e, has already a direction to pericenter. In Fig.9.3 the initial circular orbit *I* and infinitesimally close elliptical orbit *II* are shown, where orbit *II* results from the application of the pulse ΔV_T ($\Delta V_T > 0$ and $\Delta V_T \to 0$) at an arbitrary point of the orbit. It is seen in the figure that the point where the infinitesimal pulse is

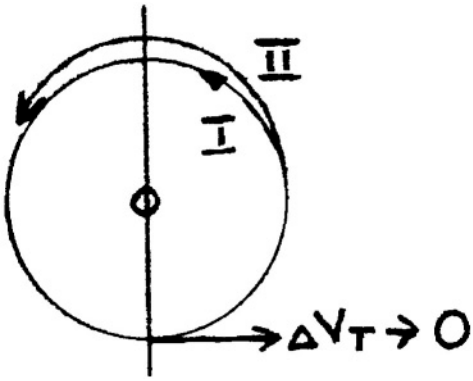

Figure 9.3. Initial circular orbit I and the close elliptical orbit II (produced by applying an infinitesimally small pulse ΔV_T)

applied becomes a pericenter and, therefore, $\upsilon = 0$. Then, via (9.19) and (9.14) we get an expression for the finite increment Δe_T attained under the application of the main, finite, pulse ΔV_T

$$\Delta e_T = \frac{2}{V_{circ}} |\Delta V_T|. \tag{9.20}$$

The absolute value of ΔV_T is used to make this formula valid when both pulses ΔV_T (the infinitesimal and the main one) are either positive or negative (i.e. are parallel or antiparallel respectively to the vector of SC velocity). In the second case when such pulses are applied in the negative direction, the kinetic energy of the SC decreases and the total energy decreases. The semi-major axis of the orbit then also decreases, since it provides a measure of the total energy (see Section 2). In this situation the point of pulse application becomes an apocenter and $\upsilon = \pi$ should be substituted in formula (9.19). It is a point worth mentioning that, whatever the direction of ΔV_T, a positive increment Δe_T is produced because we started from a circular orbit (as was stated in the beginning of this Section). To analyze the effect of applying ΔV_T to transform an elliptical orbit, formula (9.19) can be used. However, in this case the sign of Δe_T depends at what point of the elliptical orbit, and in what direction the pulse is applied.

9.3.2 Effect of a Normal Pulse ΔV_S

If a pulse ΔV_S is imparted to a SC moving in a circular orbit, then the corresponding element of the manoeuvre matrix A (9.11) is equal to $(p/\mu)^{1/2} \sin v$ and the uncertainty regarding recognizing the angle v in circular orbits is overcome using a similar technique to that adopted above (Section 9.3.1). Let an infinitesimal pulse ΔV_S directed towards an attracting center (Fig.9.4) be imparted to a SC at the point B. Then the vector of velocity of the SC turns through an infinitesimal angle ε without changing its magnitude, if we neglect the terms of higher order in the expansion with respect to ε. Therefore, the semi-major axis of the orbit, which provides a measure of the total energy, is not changed. Thus, in this approximation, a new trajectory infinitesimally close to a circle II is obtained from the initial circle I by turning about the point B through the angle ε. Consequently, point D becomes the nearest point to the attracting center and can be considered to represent pericenter. This means that a direction to pericenter is available (from the attracting center to point D) and the position of point B on the initial orbit is characterized by $v = 3\pi/2$. Thus we can write

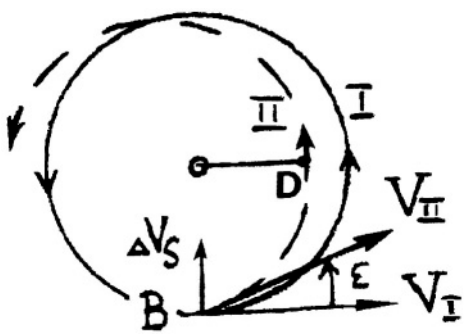

Figure 9.4. Obtaining the direction to pericenter (point D) using an infinitesimal pulse ΔV_S

$$\Delta e_S = \frac{1}{V_c} |\Delta V_S|. \tag{9.21}$$

The absolute value is used here because $\sin v < 0$ and $\Delta V_S < 0$ (the latter since the imparted pulse of velocity was directed towards the attracting center). If the pulse ΔV_S were oppositely directed ($\Delta V_S > 0$), formula (9.21) would remain valid as in this case the circular orbit

would turn clockwise and the point B is characterized by $\upsilon = \pi/2$ so that $\Delta e_S > 0$.

Expressions (9.20) and (9.21) do not provide any insight into the characteristic velocity change required to carry out realistic SC manoeuvres. To gain this insight, we will consider now the important manoeuvre of landing a SC on the Earth.

9.3.3 Manoeuvre of Landing a SC

To land a SC on the surface of an atmosphereless planet, a pulse should be applied to lower the pericenter of the orbit to the surface of the planet. In the case of a planet with an atmosphere (the Earth), to land a SC on its surface it is sufficient to push it out of its orbit into an elliptical trajectory passing through the top of the dense atmosphere. Braking then results from air resistance. An altitude in the dense atmosphere can be defined which corresponds to that altitude resulting in landing the SC after less than one revolution around the Earth. In practice the altitude of this upper boundary of the dense atmosphere is of the order of 100 km. In what follows, we will assume that the lowering of any point of the orbit to an altitude less than 100 km means that the SC has landed. Entrance into the dense atmosphere should be performed using a sloped trajectory (such that the entry angle between the local horizontal and the tangent to the trajectory is less than a few degrees). Otherwise the aerodynamic braking will be too vigorous and the g-load fatal to any on-board astronauts. For unmanned SC the value of the g-load is not such a critical factor but too steep an entry angle may mean that too large a landing pulse was imparted to the SC and, in consequence, the expenditure of characteristic velocity was too high.

Let a SC move around the Earth along a circular orbit. By imparting a suitable landing pulse $\Delta\mathbf{V}$, it can be transferred to an elliptical orbit with an appropriate conditional perigee, where *conditional perigee* is the perigee of a SC orbit corresponding to the case of an atmosphereless Earth. Clearly, it should be low enough to ensure that effective braking of the SC in the atmosphere occurs. Its value is naturally associated with the upper boundary of the dense atmosphere. It is evident that, after imparting the pulse ΔV to the SC in any direction, its new trajectory will have a segment that is situated below the

initial circular orbit. Therefore, the SC can be lowered by the pulse imparted in any arbitrary direction. The single exception is when a SC is given a pulse in the direction of its velocity. Then all points of the new orbit lie above the initial circular orbit except for that point at which the SC received this pulse. This latter point is a common one for both orbits. Therefore, in this case the perigee of the new orbit has the same value as the initial radius and landing is not achieved. Usually landing pulses apply 'braking'. Such pulses, for example, are used in landing the Russian *Soyuz* series of manned SC.

Let us consider a pulse directed against the SC velocity vector $\Delta \mathbf{V}_T$, that is, $\Delta \mathbf{V}_T < 0$, which is called a *pure braking pulse.* In this case the SC is given a pulse ΔV_T, which produces a change in both p and e as the matrix of manoeuvre A shows (see (9.9)). The corresponding elements of the matrix A are obtained in an explicit form from (9.11). An altitude reduction in the vicinity of the conditional perigee is now sought. At this point $\upsilon = 0$ and equation (1.19) gives

$$r = \frac{p}{1+e}.$$

Let the lowering of the altitude have a value ΔH sufficient to land the SC. Then, the change of r at the perigee $\Delta r = \Delta H$ or

$$\Delta H = \Delta p - p\Delta e. \tag{9.22}$$

This expression was obtained using a linear approximation, taking into account that $e = 0$ for the initial orbit. The use of a linear approximation is justified by the fact that, usually the SC to be landed moves above the Earth's surface at an altitude of a few hundred kilometers whereas the radius of the Earth is almost 6 400 km.

Let us compare both terms on the right side of (9.22). The first of them is equal to $T_{rev}\Delta V_T/\pi$ (see (9.17)) and the second is equal to $2r|\Delta V_T|/V_c$ (see 9.20). Since $T_c = 2\pi r/V_c$ it is obvious that their absolute values coincide. This means that the altitude changes in the vicinity of perigee owing to changes in e and p are of similar order for nearly circular orbits and exactly coincide for circular orbits. We can conclude from the above that the cost expressed in terms of characteristic velocity of changing H through Δe is equivalent to changing H through Δp and, thus, manoeuvring through changing e is reasonable.

To arrive at a numerical estimate of the cost in terms of characteristic velocity of carrying out the manoeuvre of perigee lowering by braking, estimation (9.18) can be combined with the conclusion obtained immediately above to show that, at $\Delta V_T < 0$, the following expression is valid

$$\Delta H \approx -3.4|\Delta V_T| \text{ km} \tag{9.23}$$

where ΔV_T is measured in units of m/s.

Expression (9.23) allows us to estimate the value of the characteristic velocity required to achieve landing. Let, for example, the initial altitude of the orbit be 400 km. To reach conditional perigee (which is lower than 100 km and therefore lies inside the dense atmosphere), it is necessary to achieve $\Delta H \geq 300$ km. The formula (9.23) for *pure braking* gives a numerical value $|\Delta V_T| \geq 90$ m/s. Practical experience confirms that the required landing pulse is indeed of the order of 100 m/s. This value is low compared with the pulse required to execute a significant orbital plane turn. Therefore, in mission planning, a strategy of manoeuvring should be developed to avoid or to decrease as much as possible, the necessity to change the angular position of the orbital plane.

An example of the possibility to substitute an orbital plane turn manoeuvre by an in-plane manoeuvre is provided by the case where it is necessary to make a given location on the Earth's surface a sub-SC point, when that point passes the orbital plane concerned either too late or too early when the SC is at the required latitude. Here, we use the fact that the orbital plane of the SC is fixed in inertial space (excluding a precession effect, which can be easily taken into account), while the Earth performs diurnal rotation beneath. Coincidence can be achieved by suitably adjusting the period of rotation of the SC. The same adjustment can be achieved in a short time by applying one 'large' pulse or, over a longer time, by accumulating, during several revolutions, the displacements produced by several relatively small pulses (see below).

9.3.4 Change of the Period of SC Revolution

The expression for T_{rev} (1.25) shows that, in order to change T_{rev}, it is necessary to change the semi-major axis a of the orbit. Taking into account that a depends only on SC energy, it follows that the SC

velocity **V** can be most effectively changed by applying a pulse ΔV_T. Using again a linear approximation we may write

$$\frac{\partial T_{rev}}{\partial V_T} = \frac{3\pi}{\sqrt{\mu}}\sqrt{a}\frac{\partial a}{\partial V_T}. \tag{9.24}$$

From geometry $p = a(1 - e^2)$ and, therefore,

$$\frac{\partial p}{\partial V_T} = (1 - e^2)\frac{\partial a}{\partial V_T} - 2ea\frac{\partial e}{\partial V_T}.$$

For a circular orbit $(e = 0)$ we get

$$\frac{\partial p}{\partial V_T} = \frac{\partial a}{\partial V_T} \tag{9.25}$$

and, using equalities (9.14) and (9.17)

$$\frac{\partial T_{rev}}{\partial V_T} = 3\sqrt{\frac{a}{\mu}}T_{rev} = \frac{3T_{rev}}{V_c}. \tag{9.26}$$

Assume, as an example, that we need a SC to transit a particular geographic point situated at the equator but that this point is located at 160 km to the East at the moment when the SC crosses the equator (i.e the SC crosses the equator too early). Suppose that $V_c = V_{circ} =$ 7 900 m/s, $T_{rev} = 90$ min and $i = 90°$. Apply a pulse $\Delta V_T = 10$ m/s to increase the velocity of the SC. This produces an increase in T_{rev} of 20 s which corresponds to a delay in the time of crossing the equator by the SC in each revolution. The lateral displacement of the geographic point with respect to the orbital plane of the SC at the moment when it crosses the equator is 10 km in these 20 s. It follows that, by accumulating the displacements of this point with respect to the SC orbit over one day $(\approx 160$ km), the point is successfully traversed after some 16 orbits.

The technique described above can be used to rendezvous two SC since it takes into account the time when the SC launched first passes the SC launched second and then provides for the synchronization of their orbital planes.

Chapter 10

Interplanetary Trajectory Corrections

The correction of a SC trajectory is a particular case of manoeuvring. In a broad sense a correction to the trajectory is made with the aim of achieving long-term targeting whereas a manoeuvre, generally speaking, results in a significant trajectory change. Among the cases previously considered, correction of the period of revolution of a SC may be considered to be of the first type while the manoeuvre of landing can be associated with the second type.

In the present Chapter the correction of an interplanetary SC trajectory through reducing inherent start errors will be considered with the purpose of establishing a trajectory configuration in the vicinity of a target planet, suitable to either land the SC on its surface or perform a flyby at a given distance. The accuracy of targeting required in a planetary neighbourhood is rather high. Thus, when placing a SC in orbit around Mars using the Martian atmosphere for braking, adherence to the nominal (precalculated) trajectory near the planet should be maintained to an accuracy of the order of *tens* of kilometers. On the other hand, the velocity error at the moment of shutting down a booster engine near the Earth, is about 1 m/s and this results in a large error near Mars of *several tens of thousands of kilometers.* This example shows the practical requirement for interplanetary trajectory correction.

Since the on-board propulsion reserve for thruster correction is always rather limited, the pulses available for correction are not large

and, consequently, they are applied using a linear approximation technique. The gross error of tens of thousands of kilometers mentioned above which should be corrected for can, however, be considered small in terms of the overall length (*hundreds of millions* of kilometers) of the trajectory.

The distances which are dealt with in correcting interplanetary trajectories close to a planet are small in comparison with typical heliocentric distances, but they cease to be small when considered in a reference system centered on a target planet. In this framework, the change of the gravitational field of the Sun in the neighbourhood of a target planet is small whereas the gravitational field of the planet itself changes very significantly at the same location. The task of correcting an interplanetary trajectory can thus be greatly simplified if it is possible to neglect the role of the gravitational field of the planet. To implement this approach, a special technique is used which is described below.

10.1 SC Motion Close to a Target Planet

Consider the motion of a SC in the immediate proximity of a target planet, for example inside its *gravisphere* (in Fig. 10.1 the boundary of the gravisphere of the target planet is indicated by a dashed semicircle). Let the trajectories which would be followed by a SC in the

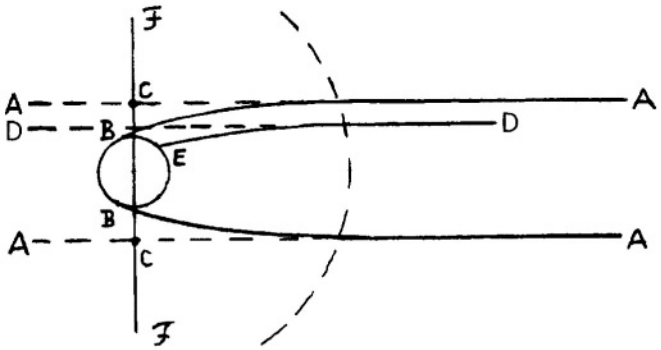

Figure 10.1. Motion of a SC inside the gravisphere of a target planet

absence of the planet be denoted by straight lines *AA*. A narrow, short, bundle of trajectories is considered which, by virtue of the local

conditions far from the Sun, can be deemed to comprise mutually parallel segments of straight lines. If now we take into account the gravitational field of the planet, the influence of this field on the trajectory is exerted after the SC transits the boundary of the gravisphere. In Fig.10.1 the resulting SC trajectories are designated by the lines *AB*. All trajectories, which lie inside that area bounded by the lines *AB* impact the surface of the planet. The trajectories illustrated allow us to differentiate between a family of *impacting* trajectories and a family of *flyby* ones.

Each trajectory in the vicinity of the planet corresponds, on a one-to-one basis, to that unperturbed trajectory followed in the absence of the planet's gravitational field (for example, impacting trajectory *DE* corresponds with trajectory *DD*). In this circumstance, we may introduce a *target plane FF* passing through the center of the planet and normal to the bundle of trajectories that would be present in the absence of the planet's gravitational field. In Fig.10.1 this family of 'fictitious' unperturbed trajectories (unperturbed means that they are built without taking into account the planet's attraction), which corresponds with the family of 'real' impacting trajectories, intersects the target plane *FF* in the points *C,* thereby defining the boundary of an area on the target plane. This is the area through which fictitious nonperturbed trajectories must pass in order that 'real' trajectories impact the surface of the planet. It is noted that other geometrical features of the real trajectories which are of interest, for example, impact at a particular planetary location; flyby at a given distance; the position of pericenter etc., can be mapped onto this target plane.

The technique described above allows us to consider, instead of real trajectories, fictitious trajectories constructed without taking into account the gravitational field of the planet. Thus, the problem of determining the position of a fictitious trajectory crossing a point within the target plane is now required to be solved instead of the problem of the location of a real trajectory relative to a real planet. This approach is designed to allow us to select the terminal parameters of the problem so that it can be solved using a linear approximation. Local non-linear effects, such as attraction by the planet, can be considered separately, for example using non-linear Keplerian relationships which give a good approximation to the actual motion of a SC inside the grav-

isphere*. The initial conditions defining such planet-centered motion coincide with the corresponding parameters of heliocentric motion in a regime unperturbed by the planet at the boundary of the gravisphere. Thus, the dashed semi-circle in Fig. 10.1 defines a location at which heliocentric and planet-centered motions can be merged.

10.2 Segment of a Nominal Heliocentric SC Trajectory

Consider a heliocentric segment of a nominal SC trajectory passing through the center *O* of a planet which is considered to be a massless point (Fig. 10.2). Let another point *B* be placed on the trajectory at

Figure 10.2. Heliocentric segment of a nominal SC trajectory passing through the center O of a planet

some distance s from the point of mission start (to correspond with the end of the active segment of the start). The flight time t corresponds to the distance s from the end of the active start. Introduce a Cartesian reference system $Bxyz$ with its origin situated at point *B*. The axes of this system are appropriately oriented relative to the stars. Such a system is convenient from the practical point of view since the orientation of the SC before its correcting thrusters are fired is defined using certain stars (including the Sun) as reference points.

Let a pulse of velocity $\Delta\mathbf{V}$ be imparted to the SC at point *B*. Without loss of generality, let such pulses satisfy the equality $|\Delta\mathbf{V}| = 1$. This defines a *sphere of unit pulses* with its center situated at *B* and denoted in Fig. 10.2 by a dashed sphere. Each point on the unit

*An example of such motion inside a gravisphere is given in Chapter 12, where a gravity-assist manoeuvre is described.

sphere corresponds to an individual pulse, and is associated with a particular, subsequent, SC trajectory. Let the SC takes a time t_k to reach the point O from B along its nominal trajectory. Consider those points along a set of trajectories originating at point B which individually correspond to a distance reached in a time t_k. Introduce a Cartesian reference system $O\xi\eta\zeta$ at point O (Fig. 10.2). Axis $O\zeta$ is directed along the velocity vector of a SC moving along its nominal trajectory. Axis $O\xi$ is directed to the center of the Sun and axis $O\eta$ completes the right-hand system. Thus, the target plane coincides with the plane $O\xi\eta$.

If the SC moves along its nominal trajectory without a pulse $\Delta\mathbf{V}$ being applied, the coordinates of the SC are $\xi = \eta = \zeta = 0$ at the moment t_k. If a pulse $\Delta\mathbf{V}$ with components V_x, V_y, V_z is imparted, then the coordinates of the SC have other values at the moment t_k. Under the assumptions of a small velocity pulse and the admissibility of adopting a linear approximation, we may write expressions for coordinates ξ, η, ζ at the moment t_k in the following way

$$\begin{aligned}
\xi &= \frac{\partial \xi}{\partial V_x}\Delta V_x + \frac{\partial \xi}{\partial V_y}\Delta V_y + \frac{\partial \xi}{\partial V_z}\Delta V_z,\\
\eta &= \frac{\partial \eta}{\partial V_x}\Delta V_x + \frac{\partial \eta}{\partial V_y}\Delta V_y + \frac{\partial \eta}{\partial V_z}\Delta V_z,\\
\zeta &= \frac{\partial \zeta}{\partial V_x}\Delta V_x + \frac{\partial \zeta}{\partial V_y}\Delta V_y + \frac{\partial \zeta}{\partial V_z}\Delta V_z.
\end{aligned} \tag{10.1}$$

We do not assume in this procedure that $|\Delta\mathbf{V}| = 1$. Before going further, it is necessary to pay attention to the following basic circumstance. A SC trajectory is described by six parameters and it is possible to choose them for $t = t_k$ in different ways (for example using six osculating elements or using six values ξ, η, ζ, $\dot{\xi}$, $\dot{\eta}$, $\dot{\zeta}$). Correction of the trajectory is made by imparting a pulse $\Delta\mathbf{V}$. Only three independent components for correcting a trajectory are, thereby, available. It is, thus, possible to define properly only three of the six parameters describing the new trajectory obtained after the pulse $\Delta\mathbf{V}$ is applied. Formulae (10.1) show that if the correction is performed for coordinates ξ, η, ζ, the components of velocity $\dot{\xi}$, $\dot{\eta}$, $\dot{\zeta}$ are not controlled. By properly selecting a pulse $\Delta\mathbf{V}$, it is possible to reach the target plane's coordinates $\xi = \xi_k$, $\eta = \eta_k$, $\zeta = \zeta_k$ at the moment $t = t_k$ but it is not possible, generally speaking, to simultaneously provide the prescribed values of SC velocity at the moment $t = t_k$.

Imparting a unit pulse of velocity $|\Delta\mathbf{V}| = 1$ to a SC (10.1) results, generally speaking, in an *affine transformation* of a sphere $(\Delta V_x)^2 + (\Delta V_y)^2 + (\Delta V_z)^2 = 1$ into an ellipsoid. Therefore, points on the SC trajectories at the moment $t = t_k$ are situated on the surface of an ellipsoid with origin at the point *O*. This ellipsoid was obtained by transformation of the sphere of unit pulses and it is named an *ellipsoid of influence.* In Fig.10.2 the outline of this ellipsoid is designated by dashes. The matrix of affine transformation (10.1)

$$A = \begin{pmatrix} \frac{\partial\xi}{\partial V_x} & \frac{\partial\xi}{\partial V_y} & \frac{\partial\xi}{\partial V_z} \\ \frac{\partial\eta}{\partial V_x} & \frac{\partial\eta}{\partial V_y} & \frac{\partial\eta}{\partial V_z} \\ \frac{\partial\zeta}{\partial V_x} & \frac{\partial\zeta}{\partial V_y} & \frac{\partial\zeta}{\partial V_z} \end{pmatrix}. \tag{10.2}$$

is called the *matrix of correction.*

Matrix (10.2) is a function of the location of the center *B* of the sphere of unit pulses along the trajectory. It can be denoted by $A(s)$ or by $A(t)$, where t is the flight time. The relation of the matrix of correction *A* to the location where a correction is performed, raises the issue of selecting an optimum location for the correction point *B,* having regard to minimizing propellant consumption. To solve this problem we will consider the properties of the required correction in more detail.

10.3 Properties of the Correction

If the determinant of the matrix *A* is not equal to zero then, in response to an affine transformation, a straight line is transformed into a straight line and a plane into a plane. Consider now some special directions (straight lines) and special planes related respectively to the sphere of unit pulses and to the ellipsoid of influence.

Among the directions of unit pulses $\Delta\mathbf{V}$, there is a direction which provides values $\xi = 0, \eta = 0, \zeta \neq 0$ by virtue of the nonsingularity of (10.1). Strictly speaking, there are two such directions lying on one straight line that correspond to $\zeta > 0$ and $\zeta < 0$ respectively. We do not distinguish between them but, rather, speak about one direction, taking into account that both directions belong to the same

straight line. This direction for unit pulses is named the *null-direction,* meaning that such pulses do not change the position of the point where the SC trajectory crosses the target plane. In particular, the nominal trajectory, which intersects the target plane at $\xi = 0, \eta = 0$ remains nominal after a pulse is imparted to the SC along the null-direction.

Let us now direct one axis of a new Cartesian reference system with origin at the center of the sphere of unit pulses along the null-direction, and designate this axis by Bz'. Any other direction of the vector $\Delta\mathbf{V}$ with unit magnitude has a projection onto the axis Bz' that is less than unity. This means that the corresponding points on the ellipsoid of influence have coordinates $|\zeta| < \zeta_1$, where ζ_1 is the absolute value of this coordinate for a pulse applied in the direction Bz'. Therefore, one of the principal central axes of the ellipsoid of influence coincides with the axis $O\zeta$ and the two other principal axes lie in a plane perpendicular to $O\zeta$ that passes through the center of the ellipsoid (that is they lie in a plane $O\xi\eta$). Let us designate the directions coincident with these latter principal axes of the ellipsoid by $O\xi'$ and $O\eta'$. Then, the axes of the planet-centered reference system $O\xi'\eta'\zeta'$ are directed along the principal central axes of the ellipsoid of influence.

The system $O\xi'\eta'\zeta'$, which was obtained by an affine transformation from the unit sphere, corresponds to a Cartesian system of axes installed in the sphere. Let us designate this reference system by $Bx'y'z'$. Then the plane $Bx'y'$ perpendicular to the null-direction Bz' corresponds to the target plane $O\xi\eta$ (or equivalently to the plane $O\xi'\eta'$). Usually, the plane $Bx'y'$ is named the *plane of optimal correction.* This name is associated with the fact that, since the vector of the pulse $\Delta\mathbf{V}$ is located precisely in this plane, the corresponding points of the ellipsoid of influence are located precisely in the target plane. Thus, the ΔV-budget or *characteristic velocity* of correction is expended only in producing a lateral displacement of the trajectory. If the purpose of correction is only to displace the trajectory in a lateral direction (for example, to cause a trajectory to pass through a particular point in the target plane), this is the most budget-effective way to achieve that aim. This is because any additional component of $\Delta\mathbf{V}$ along the null-direction requires budget expenditure, but does not produce lateral displacement in the target plane.

The above discussion suggests that the application of correcting

pulses along the null-direction is always undesirable. However, this is not correct. A pulse directed along the axis Bz' changes the coordinate ζ of the SC, that is the distance from the SC up to the target plane $t = t_k$. Therefore, a pulse along the null-direction corrects the time of arrival of the SC at a prescribed point in the vicinity of the planet. Correction of the arrival time of a SC is sometimes extremely important. For example, suppose that it is required to land a SC on the surface of Venus. Owing to the high temperature of the atmosphere, the life time of the scientific instruments on-board is limited by the time they take to heat up to a destructive value. This time is of the order of one hour. Therefore, the process of landing in the Venusian atmosphere should be monitored using direct transmissions to the Earth. Such monitoring is only possible if Venus is above the horizon of the Earth with respect to the location of the ground based mission control center. If the trajectory measurements predict that the SC will reach the planet when it will be under the horizon during the data transmission session, correction of the arrival time is mandatory.

By virtue of what has been said above, two kinds of basic manoeuvre can be distinguished. In the first, the position of crossing a target plane by the SC trajectory is corrected. This is called *two-parameter correction.* In the second case, the time of intercepting the target plane is corrected also. This is called *three-parameter correction.* Let us consider further two-parameter correction.

10.4 Two-parameter Correction

The crossing of the sphere of unit pulses by the plane of optimal correction and the crossing of the ellipsoid of influence by the target plane create, respectively, a circle of unit pulses in the plane of optimal correction and an ellipse of influence in the target plane that are related to each other. Both of these curves are represented in Fig.10.3. Elongation of the ellipse of influence informs us about the nonequivalence (in the sense of the expenditure of characteristic velocity), of directions in the target plane.

Let the actual trajectory of the SC pass through some point E in the target plane instead of through the origin. The vector of displacement OE intersects the ellipse of influence in a point F. A point P

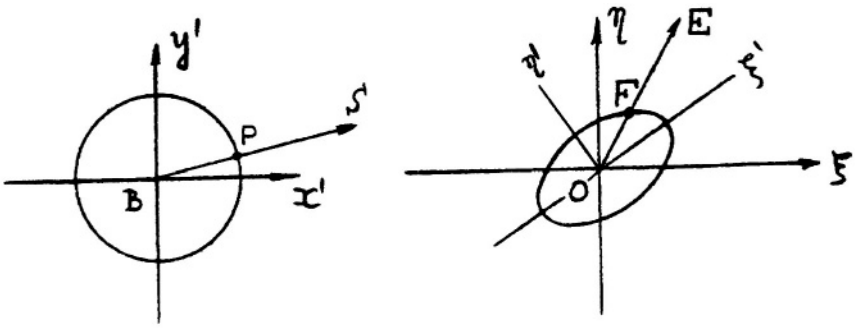

Figure 10.3. Circle of unit pulses in the plane of optimal correction and the ellipse of influence in the target plane

corresponds to the point F on the circle of unit pulses. Therefore, to obtain a vector of correction equal to OE, it is necessary to impart a pulse to the SC in the direction BP. A correcting pulse $\Delta \mathbf{V}_k$ does not, generally speaking, have a magnitude equal to unity. Owing to the linearity of the transformation this is expressed by

$$|\Delta \mathbf{V}_k| = |BS| = \frac{OE}{OF} BP = \frac{OE}{OF} 1. \tag{10.3}$$

The inverse proportionality of $|\Delta \mathbf{V}_k|$ to the value of OF, characterizes the nonequivalence of directions in the target plane. Equalities (10.3) and Fig. 10.3 show that less power expenditure is required to make corrections in the general direction of the axis $O\xi'$ than in the general direction of the axis $O\eta'$. These considerations allow us to develop an optimum strategy for interplanetary trajectory correction.

Up until now, the position of the point of correction on the interplanetary trajectory was always considered to be given. We will next study how to choose the optimum location of this point.

10.5 Optimum Location of the Correction Point on Interplanetary Trajectories

It is possible to argue that the longer a SC trajectory is tracked, the greater is the accuracy of determining the relevant trajectory parameters and that corrections to the trajectory should, therefore, be carried

out as late as is possible. Consequently, the distance s from the start point should be as long as possible. On the other hand, the longer is s, the closer the SC is to the target planet and the larger is the characteristic velocity required to realize a particular correction manoeuvre. To choose the correction point appropriately, it is necessary to find an optimum value of s that combines the required accuracy of both trajectory measurements and manoeuvre execution. In what follows, we will assume that these requirements are already fulfilled, thereby allowing us to regard the problem of locating the point of correction as a task in celestial mechanics.

Let us consider ellipses of influence for different positions of a SC on a nominal trajectory to Mars, as the duration of the flight increases (Fig.10.4). These ellipses are presented for the first day of the flight and for intervals thereafter separated by one month based on a trajectory planned in advance. As the actual trajectory was close to nominal, the ellipses could be used for flight control. It is seen that

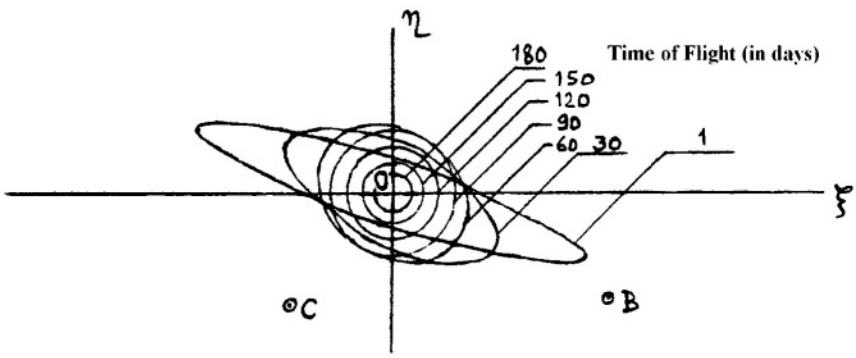

Figure 10.4. Ellipses of influence for different positions of a SC on a nominal trajectory to Mars as the flight duration increases

if, for example, the actual trajectory is characterized by a point B in the target plane (gross error OB), correction should be carried out as early as possible. If it is characterized by point C, correction should be carried out during the second or third month of flight. The figure demonstrates that postponement of correction until the sixth month, would increase the required propellant budget by approximately a factor of three. This example illustrates an important aspect of operational decision making with respect to choosing the time of in-flight correction manoeuvres.

The analysis presented above was performed on the assumption that the rank of the matrix of correction A (10.2) is equal to three. We will next consider a situation where this matrix is of rank two.

10.6 Singularity of the Correction Matrix

Consider the circumstance of executing a trajectory correction when the affine transformation (10.1) is singular and the rank of matrix (9.26) is equal to two. In this affine transformation, three-dimensional space is mapped onto a plane. Thus, the sphere of unit pulses is transformed, not into an ellipsoid, but into an ellipse (formally the magnitude of one of the principal axes of the ellipsoid reduces to zero). In consequence, the cross-section of this 'ellipsoid' (now it is degenerated to an ellipse) obtained through crossing it by the target plane provides, generally speaking[†], not an ellipse of influence but a straight line 'segment of influence'. Formally, one semi-axis of the ellipse of influence reduces to zero and, in consequence, trajectory correction in a direction perpendicular to this straight line is impossible (since the value representing distance OD in (10.3) is equal to zero). Thus, on reducing the rank of the matrix of correction to two, situations arise in which the implementation of an arbitrary correction is impossible[‡]. We will illustrate this by an example.

Let a SC moving along a circular, heliocentric, orbit be situated at point B and it is required to correct the trajectory in the target plane passing through point O (Fig.10.5). The angular distance between points B and O is equal to π. Construct at the point B a sphere of unit pulses, choosing three mutually orthogonal directions defined by the vectors ΔV_S, ΔV_T, ΔV_W (see also Section 9.1). It follows from matrix (9.10) when $\Delta V_W = 0$, that a pulse applied in the plane defined by ΔV_S and ΔV_T, can change the trajectory of the SC without changing the angular position of this plane. In the target plane, the

[†]We exclude here the case where this ellipse coincides with the target plane.

[‡]If the rank of the matrix of correction is equal to unity, the sphere of unit pulses is mapped into a segment of a straight line (now two semi-axes of the ellipsoid reduce to zero). Crossing of this segment by the target plane produces a point 'of influence' and trajectory correction is impossible.

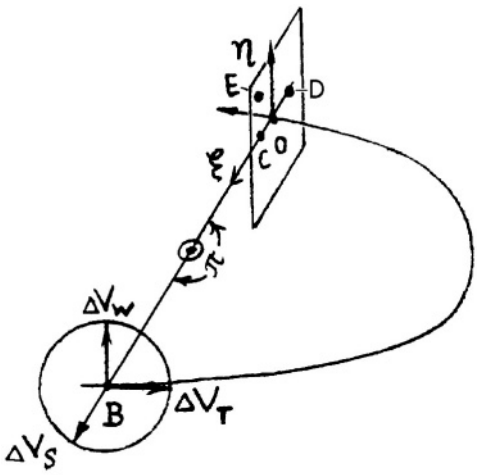

Figure 10.5. Case corresponding to the singularity of the matrix of correction of rank two

pulse provides a displacement of the trajectory along the axis $O\xi$. Then, any $\Delta\mathbf{V}$ satisfying the equality $|\Delta\mathbf{V}| = 1$ on being applied in the orbital plane $(\Delta V_W = 0)$, provides the displacement of a bundle of trajectories in the target plane along the straight line segment *CD*.

Assume that it is required to make the SC trajectory pass through the point *E*. It was shown immediately above that this cannot be achieved by any pulse applied in the orbital plane. Let us now use the last remaining possibility and impart the pulse ΔV_W to the SC as it moves along its orbit *I* (Fig.10.6). In this case, the orbital plane turns by virtue of the second row of the manoeuvre matrix (9.10) and a new orbit *II* is followed. Considering the correspondence between this manoeuvre and our previous analysis of (9.12) and (9.13), the point of imparting the pulse ΔV_W can be deemed to lie on a line of nodes *BO,* and the angle u (argument of latitude in the formulae mentioned above) is measured from this line. Thus, this pulse is able to turn the orbital plane through an angle $\Delta\alpha$ (the analog of Δi in (9.13)) around the line *BO,* but it can not produce an angular displacement of the line *BO* (the analog of $\Delta\Omega$ in (9.12)). Consequently, there is no displacement of the crossing point of orbit *II* in the target plane along the axis $O\eta$. Indeed, no pulse imparted to the SC at point *B* can produce a displacement of the trajectory along axis $O\eta$, and it is thus, in particular, impossible to pass a trajectory through point *E*.

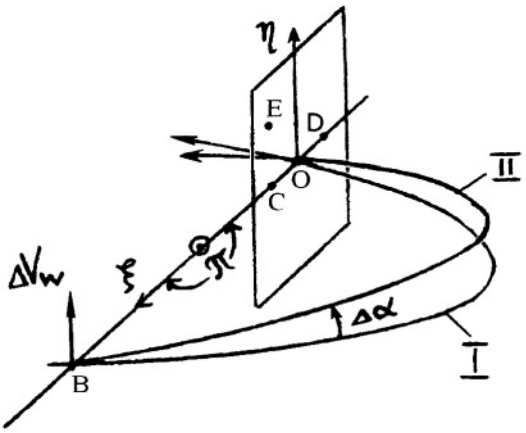

Figure 10.6. Result of applying a pulse ΔV_W at point B

This is a consequence of the degeneration of the ellipse of influence into the line segment *CD* under the constraint of the singularity of the matrix (10.2). Singularity of (10.2) derives from the fact that all the elements situated in its second row are equal to zero.

We will next outline what it is necessary to do to obtain a trajectory that passes through point *E*.

10.7 Correction Using the Singular Matrix

Consider first a strategy to traverse a point on the axis $O\eta$ in the target plane under the constraint of singularity of the correction matrix. The strategy may be designed to be optimum in terms of minimum expenditure of characteristic velocity. Traversal of $O\eta$ can be achieved by transferring the correction point B to a new position B^*. It is geometrically obvious, that the greatest displacement in the direction η is realized if the angular distance between the point of correction and the target plane is $\pi/2$ (Fig.10.6). Here the new and initial planes have the greatest linear displacement as measured along 'a vertical' (in the Figure) between corresponding points on the two considered trajectories.

Next consider the problem of traversing the point E (Fig.10.6) under the constraint of singularity of the correction matrix. It should

be noted that transfer of the correction point to B^* is not optimal with respect to achieving displacement of the trajectory along the axis $O\xi$. In relation to landing pulses, it was shown that the most effective pulse for displacing a SC trajectory in its orbital plane is the ΔV_T pulse, directed along the SC velocity vector, and that the greatest displacement takes place at an angular distance $\sim \pi$ from where the pulse is imparted.

Apply first a pulse ΔV_1 parallel to the SC velocity vector at B, an angular distance $\sim \pi$ from the target plane, to obtain a required coordinate ξ. A second pulse ΔV_2 in the direction ΔV_W is then imparted at the point B^*, an angular distance of $\sim \pi/2$ from the target plane, to obtain a displacement along the axis $O\eta$ to point E (see Fig.10.7 where I is the initial orbit, II is the orbit obtained after the first

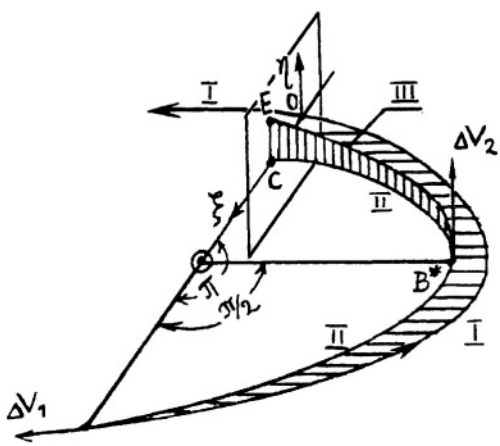

Figure 10.7. Optimum strategy for two-point correction in the case of degeneration of the manoeuvre matrix

correction, and III is the orbit obtained after the second correction).

The example considered testifies that, in some cases, multiple corrections provide the optimum effect (neglecting inherent errors in orbit prediction and in the execution of individual correction manoeuvres).

Chapter 11

Rendezvous Manoeuvring

Let a rendezvous between two SC occur in the gravitational field of a planet (for example the Earth). Assume that one SC is 'passive'(i.e. it moves in a geocentric orbit without manoeuvring), and that another 'active' SC performs the rendezvous manoeuvre. This latter situation is typical when two SC rendezvous. We will discuss first control-free relative motion of the two SC.

11.1 Control-Free Relative Motion

Consider the motion of the active SC relative to the passive one. We define their positions in space through radius-vectors $\mathbf{r}_a$ and $\mathbf{r}_p$ respectively. The motion of both SC is governed by the central Newtonian gravitational field of the Earth. Also a control force is applied to the active SC to provide the control acceleration $\mathbf{a}_c$. The equations of SC motion thus have the form

$$\frac{d^2\mathbf{r}_a}{dt^2} = -\frac{\mu \mathbf{r}_a}{r_a^3} + \mathbf{a}_c, \quad \frac{d^2\mathbf{r}_p}{dt^2} = -\frac{\mu \mathbf{r}_p}{r_p^3}. \tag{11.1}$$

Introduce a vector $\mathbf{s}$ to describe the position of the active SC with respect to the passive one so that $\mathbf{r}_a = \mathbf{r}_p + \mathbf{s}$ and assume that the magnitudes of vectors $\mathbf{r}_p$ and $\mathbf{s}$ satisfy the inequality $|\mathbf{s}| \ll |\mathbf{r}_p|$, since close relative motion of the SC pair is considered. The inequality

allows us to write a set of formulae

$$r_a^3 = ((\mathbf{r}_p + \mathbf{s})^2)^{3/2} = r_p^3 \left(1 + 2\frac{\mathbf{r}_p\mathbf{s}}{r_p^2} + \frac{s^2}{r_p^2}\right)^{3/2} \approx r_p^3 \left(1 + 3\frac{\mathbf{r}_p\mathbf{s}}{r_p^2}\right),$$
$$\frac{1}{r_a^3} \approx \frac{1}{r_p^3}\left(1 - 3\frac{\mathbf{r}_p\mathbf{s}}{r_p^2}\right).$$

Terms of order s^2/r_p^2 and higher are neglected. Then the first equation in (11.1) due to the second equation can be written, using $\mathbf{s}$ as

$$\frac{d^2\mathbf{s}}{dt^2} = -\frac{\mu\mathbf{s}}{r_p^3} + 3\frac{\mu\mathbf{r}_p}{r_p^3}\cdot\frac{\mathbf{r}_p\mathbf{s}}{r_p^2} + \mathbf{a}_c \tag{11.2}$$

where, on the right side of the equation, the terms of order $(1/r_p^2)(s/r_p)$ and the control acceleration are kept and the symbol of equality is used instead of the symbol of approximation. This equation describes the motion of the active SC in the neighbourhood of the passive SC in inertial space. If we consider their mutual motion in a reference system that rotates with respect to inertial space with angular velocity $\boldsymbol{\omega}$, then equation (11.2) attains another form

$$\frac{d^2\mathbf{s}}{dt^2} + \frac{d\boldsymbol{\omega}}{dt}\times\mathbf{s} + 2\boldsymbol{\omega}\times\frac{d\mathbf{s}}{dt} + \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{s}) = -\frac{\mu\mathbf{s}}{r_p^3} + 3\frac{\mu\mathbf{r}_p}{r_p^3}\cdot\frac{\mathbf{r}_p\mathbf{s}}{r_p^2} + \mathbf{a}_c \tag{11.3}$$

where the derivative of $\mathbf{s}$ means its 'local' variation relative to the rotating reference system. Sometimes this is called the *local derivative.* Define the vector $\mathbf{s}$ through its projections (x, y, z) on the axes of a Cartesian reference system with its origin at the center of mass of the passive SC. The third axis aligns with $\mathbf{r}_p$, the second aligns with Laplace's vector $\mathbf{c}$ and the first axis completes the reference system*.

In circular orbit this reference system rotates with constant angular velocity (denote this by ω_{orb}) and the vector equation (11.3) can then be written in the form of scalar equations as follows

$$\frac{d^2x}{dt^2} + 2\omega_{orb}\frac{dz}{dt} = a_x, \quad \frac{d^2y}{dt^2} + \omega_{orb}^2 y = a_y,$$
$$\frac{d^2z}{dt^2} - 2\omega_{orb}\frac{dx}{dt} - 3\omega_{orb}^2 z = a_z. \tag{11.4}$$

*We will introduce this reference system in more detail in Section 15.2 and call it the *orbital reference system*

These are called either *Hill's* equations or the *Clohessy-Wiltshire*[†] *equations.*

In the absence of control acceleration when the components $a_x = a_y = a_z = 0$ in the above reference system, the third equation in (11.4) transforms to the equation of a linear oscillator

$$\frac{d^2z}{dt^2} - 2\omega_{orb}\frac{dx}{dt} - 3\omega_{orb}^2 z = 0.$$

Substituting

$$\frac{dx}{dt} = \frac{1}{2\omega_{orb}}\left(\frac{d^2z}{dt^2} - 3\omega_{orb}^2 z\right)$$

expressed from this equation into the first equation in (11.4) we obtain

$$\frac{d^3z}{dt^3} + \omega_{orb}^2\frac{dz}{dt} = 0$$

which can also be transformed to a linear oscillator equation with respect to dz/dt. This means that, in a circular orbit containing a pair of active and passive SC, one can oscillate (given appropriate initial conditions) relative to the other. There is no inconsistency here with regard to one SC following a Keplerian orbit while the other does not since what is considered is *relative* motion.

11.2 Approaches to Rendezvous Manoeuvring

It is helpful now to use the approach already adopted when considering interplanetary trajectory correction (Chapter 10) since, in both cases, one body (a planet or a passive SC) moves in an unknown Keplerian orbit and the active SC moves along a trajectory designed to intercept it. Also, both bodies should reach the interception point at the same moment. These requirements can result either in the arrival of a SC at a planet or in the docking of two SC. In each of these cases, the trajectory of the manoeuvring SC is calculated to ensure interception

[†]W.H.Clohessy and R.S.Wiltshire. Terminal Guidance System for Satellite Rendezvous. *Journal of Aerospace Science,* 1960, Vol.27, Sept., pp.653–658.

and, in both cases, manoeuvring is reduced to correcting those errors inevitably introduced when injecting a SC along a pre-calculated trajectory.

Consider next that phase of the rendezvous of two SC when they are only a few kilometers apart so that on-board means of guidance can be employed. Let measurements made aboard the active SC be sufficient to provide adequate information concerning the relative motion of the SC, and also concerning the motion of the active SC with respect to the Earth. It is then possible to design a target plane at the center of mass of the passive SC, and to calculate the matrix manoeuvre of the active SC so as to perform those corrections necessary to realize a rendezvous. This task is easier than the problem of interplanetary manoeuvring already considered (using simplifying assumptions) in Chapter 10 because the passive SC does not have a significant gravitational field producing non-linear effects.

The task of rendezvous manoeuvring, however, differs from those manoeuvres considered earlier in that, for interplanetary transfers, either two or three-parameter correction of the trajectory is sufficient whereas, when two SC rendezvous, a *six-parameter* correction is required. For example, to dock two SC it is necessary not only to bring both SC to one point, but also to balance their velocities. Therefore, it is necessary to correct three linear coordinates (or two coordinates in the target plane and the time of plane interception), and also three components of the active SC velocity.

It was already shown in Section 10.2 that, since the velocity vector pulse $\Delta\mathbf{V}$ has three components, it is possible to correct only three parameters using this pulse. Therefore, to achieve six-parameter correction, at least double correction of the trajectory is required. The easiest way to represent the double correction theoretically is as follows. First, a pulse $\Delta\mathbf{V}_1$ is applied so that, at some subsequent moment t the active SC reaches the passive SC. At this moment t an instantaneous pulse $\Delta\mathbf{V}_2$ is imparted to the active SC. This pulse cannot result in an instantaneous change in the coordinates of the active SC but, rather, provides an instantaneous change in its velocity. The pulse $\Delta\mathbf{V}_2$ is chosen so as to balance the velocities of both SC. The task of making an active SC rendezvous with a passive SC is thus reduced to imparting a 'pursuit' $\Delta\mathbf{V}_1$ pulse to the active SC and then, in the immediate proximity of the passive SC, imparting a 'braking'

pulse $\Delta\mathbf{V}_2$ to the active SC to neutralize the relative velocity. It can be shown using control theory that such a rendezvous technique is optimum in that it minimizes the expenditure of characteristic velocity and it is named the *method of free trajectories.*

A disadvantage of the above method is that, in addition to requiring information on the relative motion of the two SC, it is necessary to also have information on the motion of the active SC with respect to the Earth. This means that we need to know the orbit of the active SC prior to the beginning of the rendezvous manoeuvre and also its orientation relative to the inertial reference system, in order to be able to impart to the SC a $\Delta\mathbf{V}_1$ with the necessary magnitude and direction. Obtaining such information about the SC motion to the required level of accuracy is a complex technical and mathematical problem and, therefore, a more simple technique of manoeuvring (presented below), based only on knowledge of the relative motion of the two SC has been developed.

Let B be the passive SC and C the active one (Fig. 11.1). Choose B as the origin of the *relative reference system* in which the motion of C with respect to B is described. Then, the position C of the ac-

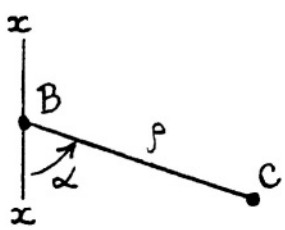

Figure 11.1. Relative positions of the passive SC B and the active SC C during the rendezvous process

tive SC is defined by the distance $BC = \rho$ and by the angle α, which is formed by the straight lines BC and xx. The latter line, which contains B, moves in inertial space parallel to itself and comprises one axis of the relative reference system (here only the planar case is considered). Radar or gyroscopic devices mounted on the active SC measure ρ, $\dot{\rho}$, and $\dot{\alpha}$. If now we force the active SC to manoeuvre so that $\dot{\alpha} = 0$ is always satisfied, then the SC moves along a straight line BC with respect to the relative reference system. To implement a rendezvous, it is necessary to impart a 'pulling' pulse to the active

SC so that $\dot{\rho} < 0$, and to impart a braking pulse which gives $\dot{\rho} = 0$ immediately before the two SC can touch. This basically simple rendezvous procedure is complicated by the need to satisfy the condition $\dot{\alpha} = 0$. 'Natural' motion in the period between the pulling and braking pulses is along a curved trajectory because both SC move at this time in Keplerian orbits around the Earth. To straighten the natural curvature of the trajectory of relative motion, it is required to maintain the condition $\dot{\alpha} = 0$ by imparting to the active SC, using a thruster, a force perpendicular to *BC*. This force does not change $\dot{\rho}$ and it is, thus, a 'spurious' one, requiring extra expenditure of fuel that does not contribute to the rendezvous manoeuvre purpose — which is to reduce ρ. The corresponding consumption of characteristic velocity is the price paid for simplifying the guidance control system of the active SC. Also, knowledge of the geocentric orbit and SC orientation is not necessary and the control algorithm is very simple (in the early days of space flight, no on-board computer was thus required). This manoeuvring method (where the vector of relative velocity of the active SC $\dot{\rho}$ remains parallel to itself all the time), is called the *method of parallel sighting* and it was used when docking the Soviet experimental SC *Cosmos-186* with *Cosmos-188* (30th of October, 1967) and also when docking manned missions of the SC *Soyuz* with the orbital station *Salyut* [11].

The consumption of characteristic velocity using the method of parallel sighting approaches that consumption inherent in the method of free trajectories ever more closely, the 'straighter' is the trajectory of uncontrolled SC motion between the pulses of pulling and braking. Obviously, the degree of 'straightness' increases as the initial distance *BC* is reduced. Therefore, at rather small initial distances (a few kilometers), use of a rather simple method of parallel sighting is preferable while, at initial distances of tens of kilometers where the method of free trajectories is employed, more sophisticated instrumentation is required. The procedure of rendezvous maneuvering described above, in which at the moment of termination the coordinates of both SC centers of mass coincide, is certainly idealized. This process begins when the distance between both SC is of the order of tens of meters. After that, the phase of *docking* starts which is realized by thrusters that provide translational and angular displacement of the active SC required for accurate soft contact and docking with the passive SC.

Chapter 12

Gravity-Assist Manoeuvre

All the manoeuvres considered previously are based on changing a SC trajectory through thruster ignition. However, there is also a possibility to change the orbit of a SC by utilizing the gravitational field of a celestial body, as was suggested first by Kondratyuk* in 1918†.

If a SC moves along some heliocentric elliptical orbit then, within the framework of the two-body problem, this elliptical orbit never changes. If, however, the SC flies by a planet, the gravitational field of that planet can significantly alter the heliocentric orbit followed by the SC. That change in the trajectory and velocity of a SC produced during a planetary flyby due to the ambient gravitational field, constitutes what is called a *gravity-assist manoeuvre.*

12.1 Description of the Manoeuvre

Let us suppose as was already done in Chapter 4 that, inside the gravisphere of a planet, SC motion is determined by the planet's gravitational field while outside the boundary its motion is determined by the gravitational field of the Sun. Consider the motion of the SC inside the gravisphere of a planet in a planet-centered reference system. As will be shown below, introducing the gravisphere allows us to

*Kondratyuk, Yury Vasiljevich (1897–1941). Russian inventor who carried out research in astronautics and rocket engineering.

†Yu.V.Kondratyuk. For Those Who Are Studying to Build (1918-1919). In the Selected Papers of the Pioneers of Rocket Engineering: Kibalchich, Tsiolkovsky, Tsander and Kondratyuk, Moscow, Nauka Publ., 1964, pp.501-536.

treat the motion of the SC in the neighbourhood of the planet in the framework of a two-body problem and to associate the gravispheric boundary with infinity.

We start from a determination of SC velocity at the gravispheric boundary. When the SC enters the gravisphere, its planetocentric velocity at the boundary is greater than zero. If the velocity had been zero at the boundary, the SC would have 'dropped' to the Earth while gaining the planetary escape velocity — zero velocity at infinity corresponds to the escape velocity at the surface of a planet.

Since the velocity of the SC in a planet-centered reference system at the gravispheric boundary is greater than zero, it follows (from the two-body problem), that the SC moves relative to the planet with hyperbolic velocity. Thus, when considering SC motion in the vicinity of a planet, we may use relationships that are valid for hyperbolic trajectories. For hyperbolic motion the constant of energy h in the integral of energy (1.10) is positive and, at infinity, where $r = \infty$ and $V = V_\infty$, this formula reduces to

$$h = V_\infty^2. \tag{12.1}$$

By definition the gravitational field of the planet vanishes outside its gravisphere. Thus, acceleration of the moving SC by the planet's gravitational field begins only when it intercepts the gravispheric boundary. The velocity at the boundary of the gravisphere can be denoted by V_∞ and determined using formula (12.1). On the other hand, for any orbit defined by a conic section, equality (1.22), that is,

$$h = \frac{\mu}{p}(e^2 - 1)$$

is valid. Substituting in this formula an expression for the eccentricity e of a hyperbola with semi-major axis a and semi-minor axis b, and for the parameter p (which is valid for any conic section)

$$e = \frac{\sqrt{a^2 + b^2}}{a}, \quad p = \frac{b^2}{a}$$

we obtain

$$h = \frac{\mu}{a}.$$

Combining this with formula (12.1) gives

$$a = \frac{\mu}{V_\infty^2}. \tag{12.2}$$

The trajectory of SC motion is illustrated in Fig. 12.1. The planet is

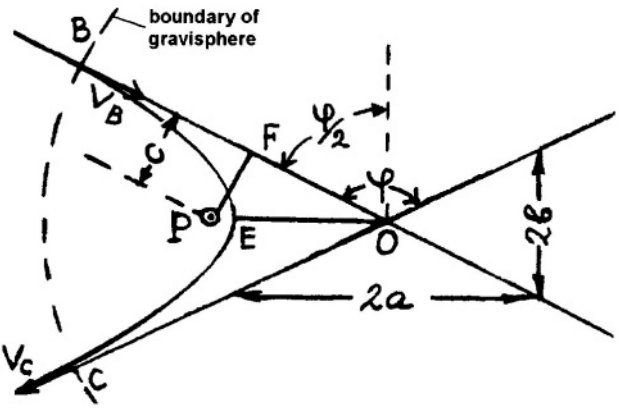

Figure 12.1. Trajectory of a SC during a gravity-assist manoeuvre in a planetocentric reference system

deemed to be located at the point *P*, and the boundary of its gravisphere is represented by the dashed line *BC*. The SC enters the gravisphere parallel to the asymptote (because this is its theoretical direction as $r \to \infty$) with a velocity V_B which is equal to V_∞. After flyby the SC. following an hyperbolic trajectory, exits the gravisphere along another asymptote at point *C*, with a velocity V_C, which is also equal to V_∞ by virtue of the law of energy conservation. During this overall process, there has been a turn of the velocity vector through an angle defined by the asymptotes of the hyperbola.

Let us denote this angle by φ. From geometry, the relationship *OE* = *FO* is known, since point *F* is obtained through the orthogonal projection of *P* onto *BO*. Also, from the diagram

$$\tan\frac{\varphi}{2} = \frac{FO}{FP} = \frac{a}{c}. \tag{12.3}$$

The value c is named the *aiming distance*[‡].

[‡] This name is associated with the fact that c is the distance between the direc-

Combining equalities (12.2) and (12.3) we get the expression

$$\tan\frac{\varphi}{2} = \frac{\mu}{cV_\infty^2}, \tag{12.4}$$

which relates the turn of the vector of velocity relative to the planetocentric reference system (angle φ) with the aiming distance (c); the velocity at the boundary of the gravisphere (V_∞) and the gravitational parameter (μ) of the planet.

Expression (12.4) shows that one can effectively vary angle φ by changing the aiming distance c, which is determined by the point F at which the nonperturbed trajectory of the SC crosses the target plane. To change V_∞ is usually inconvenient. This latter velocity derives from previous requirements in composing the segments of the heliocentric trajectory. Also, to change V_∞ by rocket-engine ignition requires the expenditure of characteristic velocity, whereas the main purpose of a gravity-assist manoeuvre is to change a trajectory without employing rocket engine thrust.

The maximum value of the angle φ is π, which corresponds to $c = 0$. Therefore, considering the planet as a particle, it is theoretically possible for a SC to be 'turned back' in a planetocentric reference system. The maximum value of the turn angle φ is obtained at $c = R$, where R is the radius of the planet. Taking now into account the expression for circular velocity (2.3), we obtain, using (12.4), the expression

$$\tan\varphi_{max} = \frac{V_{circ}^2}{V_\infty^2} \tag{12.5}$$

where V_{circ} is the circular velocity associated with the gravitational field of the planet (Section 2.5) and φ_{max} is the maximum accessible value of the turn angle of the SC velocity vector. This value is large because the planetocentric velocity of the SC 'at infinity' V_∞ is significantly less than its circular velocity V_{circ} at the planet.

---

tions of the velocity vector at infinity and the center of the planet. The asymptotes of the hyperbola in this case each looks like the sighting line of a gun which, as is well known, does not coincide with the flight trajectory of a bullet.

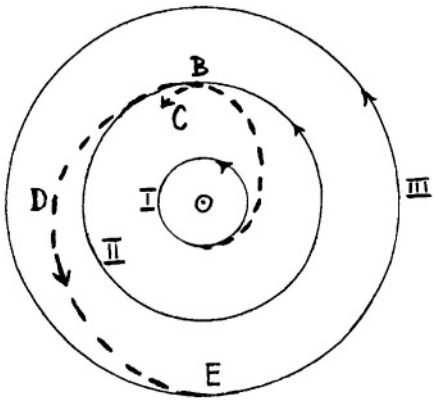

Figure 12.2. Scheme for interplanetary transfer

12.2 Application of the Gravity-Assist Manoeuvre for Interplanetary Missions

Consider first the application of the gravity-assist manoeuvre for interplanetary transfer. Fig.12.2 shows a part of the Solar System containing the trajectories of three planets. Circle *I* corresponds to the orbit of the Earth. A SC is required to reach the orbit of a planet represented by circle *III*. This task, however, requires the SC to be boosted at its start from the Earth to such a high velocity that this cannot be technically achieved. Suppose a launcher to be only capable of speeding up the SC to such a velocity that its trajectory becomes an ellipse (denoted in Fig.12.2 by *C*) tangent to the point *B* in the orbit *II* of an intermediate planet, which is itself also called *II*. If the time of liftoff is chosen so that both the SC and planet *II* appear in the vicinity of point *B* simultaneously, the planet will overtake the SC. This follows from the fact that the total energy of a body in Keplerian motion is proportional to the semi-major axis of its orbit and, here, the semi-major axis of the planetary orbit *II* is greater than the semi-major axis of the SC's elliptical orbit *C*.

For simplicity, suppose that planet *II* is a particle, and that the aiming distance chosen for rendezvous of the SC with the planet is zero. At the beginning of the rendezvous (i.e. when the planet is 'overtaking' the SC), the heliocentric velocity of the planet is repre-

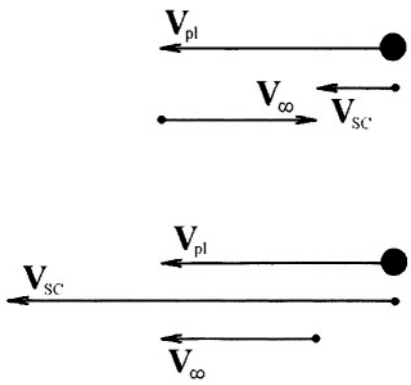

Figure 12.3. Diagrams of the velocities before (upper) and after (lower) a gravity-assist manoeuvre

sented by $\mathbf{V}_{pl}$ and the heliocentric velocity of the SC by $\mathbf{V}_{SC}$. The vector of the *relative* velocity of the SC,

$$\mathbf{V}_{\infty} = \mathbf{V}_{SC} - \mathbf{V}_{pl},$$

is then directed to the right side, as per Fig.12.2. The vector $\mathbf{V}_{pl}$ is, meanwhile, directed to the left side (see the upper diagram in Fig12.3). During the manoeuvre the vector of planetocentric velocity of the SC $\mathbf{V}_{\infty}$ is reversed (i.e. it turns through an angle $\varphi = \pi$). Thus, after the SC has passed the planet the vector $\mathbf{V}_{\infty}$ is directed to the left and the relative velocity of the SC is given by

$$-\mathbf{V}_{\infty} = -\mathbf{V}_{SC} + \mathbf{V}_{pl}.$$

This means that the SC speeded up in the vicinity of the planet by $2\mathbf{V}_{\infty}$, such that its heliocentric velocity

$$\mathbf{V}_{pl} + (-\mathbf{V}_{SC} + \mathbf{V}_{pl})$$

attained a magnitude greater than V_{pl}, i.e. it became greater than the heliocentric velocity of planet *II* (see the lower diagram in Fig12.3). As a result, the SC transfers to elliptical orbit *D* which intersects the orbit *III* of the third planet in point *E*. The transfer problem to reach orbit *III* has thus been solved through using the gravitational field of an intermediate planet to speed up the SC. This example provides only

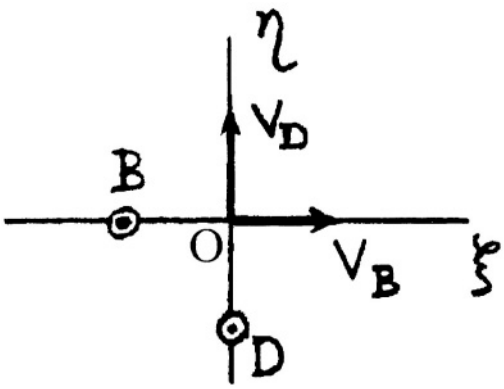

Figure 12.4. Target plane of a planet used for a gravity-assist manoeuvre

the concept of a manoeuvre with $\varphi = \pi$ because an aiming distance of zero cannot be realized in practice.

Another application of the gravity-assist manoeuvre involves changing the direction of SC motion. Formulae (12.4) and (12.5) allow us to choose the aiming distance c which provides a required value of the trajectory turn angle in a planetocentric reference system (with the limitation $\varphi \leq \varphi_{max}$). It is especially noted, that the parameter providing the trajectory turn angle (i.e. aiming distance c) should be accurately realized in order to obtain a particular desired value of φ. This implies the necessity to make precise three-parameter corrections to the trajectory along the segment between planets *I* and *II*.

12.3 Inclination Change of the Heliocentric Orbit of a SC

The gravity-assist manoeuvre is also capable of producing a change in the inclination of the heliocentric orbit of a SC. Consider an example. Let plane $O\xi\eta$ be the target plane of a planet which is used for a gravity-assist manoeuvre (Fig. 12.4), and let axis $O\xi$ lie in the plane of the SC orbit. The aiming distance c is such that the turn angle φ of the SC trajectory is $\pi/2$. Suppose that the SC crosses the target plane at a point B lying on the axis $O\xi$. After turning around the planet, the SC displays a planetocentric velocity $\mathbf{V}_B$ directed parallel to the axis $O\xi$ while the original plane of the SC orbit is maintained.

If, instead, the SC crosses the target plane at point D (after a preliminary manoeuvre), then the SC velocity on its exit from the planetary gravisphere is $\mathbf{V}_D$ and directed parallel to the axis $O\eta$. In this case, the inclination of the plane of the SC orbit is changed by $\pi/2$ with respect to the initial one.

Thus, depending on where we choose the crossing point of the target plane to be, it is possible to turn the planetocentric velocity vector of the SC in any direction and, consequently, turn the heliocentric velocity vector and the plane of heliocentric motion.

Gravity-assist manoeuvres are widely, and effectively, used for interplanetary missions. The first time such a manoeuvre was realized was during the flight of the Soviet SC *Luna-3* in 1959, when a pioneering picture of the reverse side of the Moon was taken. The simple transfer trajectory shown in Fig. 12.5 presents an obvious candidate orbit. In this scenario the SC would move along an elliptical trans-

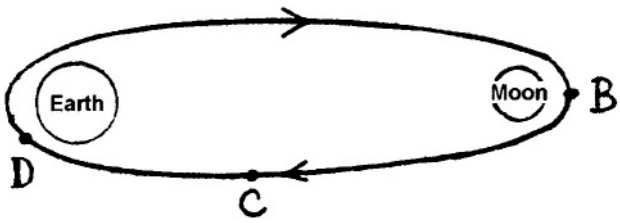

Figure 12.5. Simple transfer trajectory to the Moon

fer orbit and at point B a picture of the Moon could be taken. In 1959 a disadvantage of this trajectory was that, on returning to the Earth, the SC would approach from the South and could not, along the orbit segment CD closest to the Earth, be contacted from the territory of the former Soviet Union. To overcome this a trajectory was chosen involving a gravity-assist manoeuvre. Detailed analysis of such trajectories was performed by Egorov[§]. First, the semi-major

[§]Egorov, Vsevolod Alexandrovich (1930–2001), Principal Researcher at the Keldysh Institute of Applied Mathematics of the Russian Academy of Sciences and a professor at the Lomonosov Moscow State University. In the framework of the restricted three-body problem he classified trajectories of direct approach to the Moon and researched fundamental properties of these trajectories under the aegis of the Keldysh-Okhotsimsky Scientific School of *Applied Celestial Mechanics (V.A.Egorov. Specific Problems of a Flight to the Moon. Physics–Uspekhi, 1957, Vol.63, N 1a, pp.73–117).*

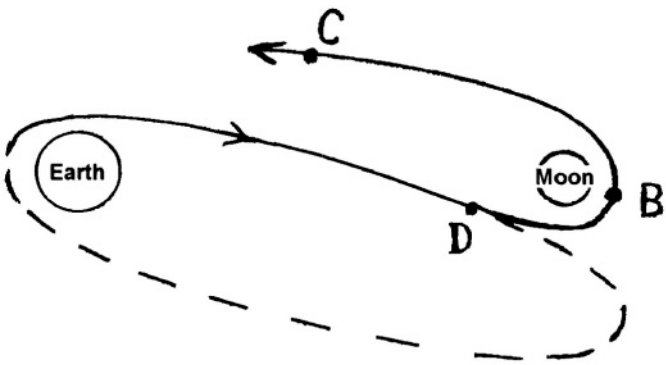

Figure 12.6. Trajectory followed by the SC *Luna-3* using a gravity-assist manoeuvre

axis of the elliptical transfer orbit of the SC was lowered 'under' the Moon's orbit around the Earth (see Fig. 12.6 where the transfer orbit is shown in a plane perpendicular to the orbit of the Moon). Without the gravitational field of the Moon the SC would have proceeded, after point *D,* along an orbit represented by the dashed line. However, under the effect of the Moon's gravitational field, the trajectory could follow the segment *DC* (with picture taking at the point *B*). As a result of adopting this scheme, the SC approached the Earth from the North and, during practically all the time of its return flight, data were transmitted to the Soviet mission control stations.

Without the use of gravity-assist manoeuvres, in particular for speeding up SC, it would be impossible to realize the interplanetary missions of today. The first use of this manoeuvre in the vicinity of Venus was made by the U.S. SC *Mariner-10* while *en route* from the Earth to Mercury in 1974. Also, U.S. SC *Pioneer-11* was the first to execute a gravity-assist manoeuvre at Jupiter in 1974 while on the way to Saturn. The European SC *Giotto* performed the first gravity-assist manoeuvre at the Earth of an observing SC coming from deep space in 1990.

Chapter 13

About Orbit Determination Using Measured Data

To control a SC during flight, it is necessary to know its orbit accurately. An orbit is determined using computer processing of trajectory measurements. The measurements concerned usually comprise the distance to the SC and its radial velocity at appropriate moments of time $t_1, t_2, \ldots, t_n$. Such measurements may be made using radio or laser instruments installed at different locations on the Earth. Their coordination and processing would be an elementary task if they were absolutely accurate and independent and their number were equal to six, the number of osculating elements determining a Keplerian orbit which can be used to approximate an actual orbit for a certain time interval. Since, however, SC motion is always perturbed, such measurements should be made frequently in order to monitor orbit evolution.

Overall, the problem of orbit determination is complicated by the fact that, with the purpose of increasing accuracy, the number of measurements made considerably exceeds the necessary minimum of six, so that a problem of processing large quantities of data arises. If there are no constraints on computer memory and on the duration of the calculations, a classical *least square method* of data processing developed by Gauss* can be used for orbit determination. Consider an

*Gauss, Carl Frederich (1777–1855), outstanding German mathematician and astronomer who made major inputs to mathematics, physics and astronomy. He invented the method of least squares. A renowned memoir on potential theory was

outline of this method.

13.1 Least Square Method

Let m parameters of SC motion x_i, $(i = \overline{1,m})$ be measured. They can be presented as known unique functions of orbital elements $a_1, \ldots, a_6$ and time t

$$x_i = x_i(a_1, a_2, \ldots, a_6, t), \quad (i = \overline{1,m}) \tag{13.1}$$

where a bar over the numbers indicates that index i varies from 1 to m, running through all intermediate integer values.

If $m = 6$ and the measurements have been made at known moments of time, we can resolve this system with respect to unknown orbital elements $a_1, a_2, \ldots, a_6$ and obtain the Keplerian orbit of the SC (although of course with some inaccuracy).

If $m > 6$, then using equalities (13.1) we can write equations

$$\begin{aligned} \Phi_1(a_1, a_2, \ldots, a_6) &= b_1, \\ &\ldots, \\ \Phi_m(a_1, a_2, \ldots, a_6) &= b_m, \end{aligned} \tag{13.2}$$

where Φ_i are known functions, $(i = \overline{1,m})$, and b_i are measured values containing measurement errors. Different kinds of measurements (for example, distances, angles, velocities) at times $t_1, t_2, \ldots$ can be taken into account. The problem to be solved is as follows. We need to find values of the variables a_k, $(k = \overline{1,6})$, that satisfy equations (13.2), assuming an initial approximation to the values a_k to be available,

$$a_1 = a_1^{(0)}, \ldots, a_6 = a_6^{(0)}. \tag{13.3}$$

Denote the differences between the approximated $(a_k^{(0)})$ and true (a_k) values of the orbital elements by the symbol Δ_i

$$a_k - a_k^{(0)} = \Delta_k, \qquad (k = \overline{1,6}). \tag{13.4}$$

If the functions Φ_i are differentiated, we can write the approximate equality

$$\Phi_i(a_1, \ldots, a_6) = \sum_{k=1}^{6} A_{ik}\Delta_k + \Phi_i(a_1^{(0)}, \ldots, a_6^{(0)}), \tag{13.5}$$

---

just one of his many significant contributions to the field of applied mathematics.

where the partial derivatives $A_{ik} = \dfrac{\partial \Phi_i}{\partial a_k}$, $(i, k = \overline{1,6})$ are calculated at the point $(a_1^{(0)}, \ldots, a_6^{(0)})$. Then, substituting (13.5) in (13.2), we obtain

$$\begin{aligned} &\sum_{k=1}^{6} A_{1k}\Delta_k = B_1, \\ &\ldots, \\ &\sum_{k=1}^{6} A_{mk}\Delta_k = B_m, \end{aligned} \tag{13.6}$$

where

$$B_i = b_i - \beta_i, \qquad (i = \overline{1,m})$$

and $\beta_i = \Phi_i(a_1^{(0)}, \ldots, a_6^{(0)})$.

The number of equations in (13.6) is more than the number of variables Δ_k contained there. Thus, we cannot solve (13.6) with respect to Δ_k using regular mathematical methods. It is necessary to introduce another formulation of the problem which will allow us to obtain a solution. Since we can select initial values $a_i^{(0)}$ of the variables a_i we may choose them so that the *measured* magnitudes are close to the *calculated* values. Calculated means that we substitute the initial values of the variables a_i in the mathematical model of motion of the SC and compute quantities corresponding to the measured ones (b_i). In our case the mathematical model comprises a set of functions $\Phi_i(a_1, a_2, \ldots, a_6)$, $(i = \overline{1,m})$ but also, it may, for example, be a set of differential equations).

Introduce the *loss function*

$$T = \left[\sum_{k=1}^{6} A_{1k}\Delta_k - B_1\right]^2 + \ldots + \left[\sum_{k=1}^{6} A_{mk}\Delta_k - B_m\right]^2, \tag{13.7}$$

which is the sum of the squares of the differences between the *measured* and *calculated* quantities. The other condition for solving (13.6) is that the loss function should be a minimum. This corresponds to obtaining variables Δ_k, which provide minimum error in the determination of the variables a_k.

Necessary conditions for minimum T are

$$\frac{\partial T}{\partial \Delta_k} = 0, \ (k = \overline{1,6}). \tag{13.8}$$

In contrast to (13.6), the number of equations (13.8) is equal to the number of variables $\Delta_1, \ldots, \Delta_6$ which refine the initial approximation $a_k^{(0)}$ by

$$a_k^{(1)} = a_k^{(0)} + \Delta_k, \ (k = \overline{1,6}). \tag{13.9}$$

This approximation could, in turn, be improved by repeating the procedure described above using $a_k^{(1)}$ instead of $a_k^{(0)}$. The process of iterating refinements ceases when the difference between two consecutive values of the calculated elements a_i becomes less than some assigned value. A proof of convergence of the method described and presentation of different available versions of the method are omitted here.

The least square method is widely used in science and engineering because of its simplicity and universality. However, it has inherent disadvantages. In the discipline of astronautics, the main disadvantage is its 'static' character. It is well adapted to processing a fixed array of measurements. Let us suppose that this array is continuously being filled up with new measurements but we are not allowed to wait for the end of the measurements. Then the task is as follows. Let, at the moment t, m measurements be available. We determine the orbital elements provided by these measurements. At the moment $t_2 > t_1$ we get the next measurement and the total quantity of measurements is now $m + 1$. Generally speaking, determination of the orbital elements can then be performed with higher accuracy because the quantity of measurements has increased. Therefore, after each new measurement session, the orbital elements should again be determined. Using the least square (or some similar) method, after each new measurement session it is necessary to perform the awkward calculation procedure described above. It is thus desirable, having obtained the orbital elements by processing m previous measurements, to update them, taking into account the next measurement only, and not process all the previous measurements again. This motivation gave rise to the concept of *recurrent methods.*

13.2 Concept of Recurrent Methods

Let some vector of state **a** have components $(a_1, \ldots, a_6)$ which are obtained by processing an array of m measurements. We indicate

this by the index m in the notation of vector **a**. Let the $(m+1)$-th measurement be accomplished and denoted by $\mathbf{b}_{m+1}$. Then, to implement the idea mentioned above, it is necessary to construct an algorithm Ψ such that the equality

$$\mathbf{a}_{m+1} = \Psi(\mathbf{a}_m, \mathbf{b}_{m+1}) \tag{13.10}$$

is valid. Such methods have been developed. They are usually called *recurrent filtering algorithms.* We will consider next an elementary task that illustrates the essence of the related procedures.

Suppose that it is necessary to find the most probable value of a length a of which multiple measured values are available. The algorithm to calculate the most probable value is first reduced to the computation of the mean value a_m of this length over a number of measurements b_i,

$$a_m = \frac{1}{m}\sum_{i=1}^{m} b_i. \tag{13.11}$$

Here the index m of the length a indicates that its magnitude is obtained by processing m measurements. Let the $(m+1)$-th measurement be made. Then the improved value of a is as follows

$$a_{m+1} = \frac{1}{m+1}\sum_{i=1}^{m+1} b_i. \tag{13.12}$$

Taking the difference $a_{m+1} - a_m$ and performing simple transformations, we can express the next improved approximation of the length a, using the previous approximation a_m and the next measurement b_{m+1}, by

$$a_{m+1} = \frac{ma_m + b_{m+1}}{m+1}. \tag{13.13}$$

This latter expression is similar to (13.10). The efficiency of the recurrent algorithm (13.13) in determining the length a is demonstrated as follows. Let m be equal to 1000. Then to obtain a_{m+1} using (13.12) requires the fulfillment of about 1000 elementary operations, while the use of (13.13) needs only four operations. In astronautics and spaceflight dynamics where the processing of large scale arrays of continuously accumulating measurements is necessary, recurrent algorithms are widely used. This allows, for example, at any moment of time,

the most probable value of a SC vector of state **a** to be continuously updated, practically at the rate of measurement.

A version of the recurrent method called Kalman filtering is very common for on-board determination of the attitude and orbital parameters of SC. To use this method a model of measurements and a model of SC motion are each required [32].

Chapter 14

Introduction to Attitude Control

In previous chapters, spaceflight dynamics were considered using the assumption that a SC is a particle. Nevertheless, the importance of controlling the angular position (attitude) of the SC was implied. In considering manoeuvring problems, it was always assumed that a pulse of velocity could be imparted to a SC in a necessary direction. Since such a pulse is provided by a thruster, capability is implied to turn the SC to a required attitude.

A change in the angular position of a SC with respect to an external reference system is called *attitude motion.* The introduction of a deliberate change in the position of the body-fixed axes of a SC relative to an external reference system, is called *attitude control.* Maintenance of a given angular position to a required accuracy is called *stabilization* of the SC. The system providing either a change in, or maintenance of, SC attitude is called either the *attitude control system* or the *stabilization system.*

14.1 Active, Passive and Combined Attitude Control Systems

The mode of SC attitude motion is determined by tasks to be solved by various payload and auxiliary systems during the mission. For instance, during a radio-transmission session, an antenna is required to

point towards a control station; to another SC or to a mobile, ground based, facility. For remote sensing of the Earth's surface, pointing a camera towards a chosen surface area is required. The observation of stars or other remote celestial objects requires pointing towards these objects and stabilization of the SC relative to inertial space. Again, to increase the power output of solar arrays, or to maintain a temperature regime, it may be necessary to orient the SC relative to the Sun. Thus, the problem of SC orientation and stabilization belongs among the basic tasks to be solved during SC development, design and operation.

Orientation of a SC can be achieved using methods of active or passive control, or through their combination. This depends on the mission concerned. Attitude control systems may thus be divided into three main classes, namely active, passive and combined systems.

Active attitude control systems require an on-board power supply; a computer; attitude sensors; attitude determination and control software and actuators. Magnetometers, sun-sensors, star-sensors, measuring gyroscopes, infrared-sensors etc, are used to make measurements providing information about the actual attitude motion of a SC. Control gyroscopes, thrusters, flywheels, magnetic coils, magnetorquers, gimbaled control booms etc, are used as actuators to provide a control torque.

Passive attitude control systems do not require on-board sources of energy, sensors, software or actuators. Their operation is based on the development of restoring and damping torques, achieved through designing an appropriate shape for the SC body, a suitable tensor of inertia and through installing special elements on the SC (for instance, a permanent magnet).

Attitude control systems variously utilize the interactions of the SC with the Earth's gravitational and magnetic fields; solar radiation pressure; the aerodynamic drag created through movement of the SC in the ambient atmosphere and the property of a rigid body spinning about its axis of maximum moment of inertia to maintain the orientation of this axis relative to inertial space.

A system using the interaction between a SC and the gravitational field of the Earth is called a *gravity-gradient attitude control system.* Due to the difference in the gravitational force exerted on elements of the SC situated at different distances from the center of the Earth,

and to the presence of a centrifugal force produced due to the rotation of the SC relative to inertial space, a *gravity-gradient torque* is developed. This torque depends on the tensor of inertia of the SC, and also on its orientation relative to the local vertical and to the orbit normal. If the torque returns the SC to a required stable position, it is called a *restoring torque.* If the torque causes the SC to assume an undesired stable position it is called a *tumbling torque.* A damping device should be installed on-board a SC to provide asymptotic stability for its required stable position.

A system using the interaction between the magnetic dipole moment of a SC and the Earth's magnetic field to develop a restoring torque, is called a *magnetic attitude control system.* This torque causes the SC to follow the direction of the local vector of intensity of the geomagnetic field (as the arrow in a compass follows the ambient magnetic field). Damping elements are also required on-board to provide asymptotic stability.

The pressure exerted on different parts of a SC surface due to incident solar photons, or to the molecules of the atmosphere, results in each case, in the development of a restoring torque when the axis of orientation of the SC does not coincide with the direction of the photon, or of the molecular, flow. The behaviour of a SC equipped with attitude control systems that interact either with solar radiation or with the ambient atmosphere, is similar to that of a vane in windy weather. Such installations are called *solar radiation attitude control systems* and *aerodynamic attitude control systems* respectively. The associated damping effect is very weak in both cases and, to achieve asymptotic stability, other physical principles are applied.

The capability of a spinning SC to keep the orientation of its axis of spin fixed in inertial space is based on the 'internal' properties of a spinning rigid body. Any interactions with external fields or with the environment, produce perturbing torques that adversely affect this orientation. A damping device has to damp out nutation motion that displaces the spin-axis from its position while, at the same time, not affecting the rate of SC spin. The required damping device is called a *nutation damper.* It exploits that friction between moving parts of a SC body that pertains when the SC rotates about an axis that is not coincident with its axis of spin.

Sensors are not needed to provide attitude control. However, they

are often installed to make attitude determinations, in order to support the interpretation of certain payload measurements. Another reason to make attitude determinations is to identify proper times at which to change the configuration of the attitude control system.

It is next useful to compare the advantages and disadvantages of passive and active attitude control systems.

— An active system can provide both a high precision of orientation and a fast time-response. A passive system cannot do this. To seek from a passive control system both high precision and a fast time-response is to impose contradictory requirements. However, due to the long life time of a SC, the inherently slow time-response of passive systems is not a significant defect.

— An active system can develop a large control torque and perform complex, pre-programmed, rotation and re-orientation procedures. A passive attitude control system develops a weak restoring torque. This results in their being rather strong constraints on the initial conditions of SC motion under which a required orientation can be achieved. In consequence, accurate pre-calculation and a judicious choice of relevant structural parameters is required.

— The reliability of active attitude control systems containing numerous active elements is usually provided through redundancy. This results in an increase in the total mass, as well as in the cost of, the SC. Passive systems are light, have a simple design and, consequently, are characterized by low cost, high reliability, practically unlimited life time and parameter stability.

— Passive systems are ecologically clean in comparison with the active systems provided by gas thrusters (the discharge of spent gases can result in the degradation of optical surfaces and lenses, thus making it impossible to execute some experiments). Also, spinning elements of active systems can give rise to vibrations that are deterimental to some experiments.

— Passive attitude control systems are preferred for certain applications. These include small, low-cost, SC with long life times, dedicated to space exploration. The exploration missions concerned do not require the performance of complex, pre-programmed, angular manoeuvres or the attainment of precise orientation. For such SC, strong constrains on the total mass, dimensions, power resources and life time usually make passive attitude control systems the single acceptable

choice.

An orientation precision of several degrees is sufficient for, non-complex, space based, communication systems. This precision is also acceptable for environmental monitoring and for making observations of the Earth's surface from space if a suitable technology for the manufacture of the sensing instruments is adopted. For example, tunable acousto-optical filters with microchannel plates* as the recording devices require, due to their high sensitivity, only short exposure times. They can thus be substituted for CCD-cameras, with a consequent reduction in the required accuracy and stability of SC orientation.

Attempts of SC designers to avoid the individual disadvantages of active and passive attitude control systems, have resulted in the development of combined systems. Examples of such systems include:
— a gravity-gradient attitude control system with a gimbal controlled boom;
— a fixed (post deployment) boom with active magnetorquers, to provide damping and uniqueness in the orientation of a SC;
— a system with a *pitch flywheel* spinning at a constant angular rate so as to maintain the SC pitch axis perpendicular to the orbital plane, and active magnetorquers to turn the SC about the flywheel spin axis.

Combined attitude control systems can also be used for multi-body SC to appropriately orient their individual units with the accuracy each requires. These systems foster an increase in the working life of a SC because some elements remain in operation all the time (for example passive or active elements working in 'save mode'). Other elements that provide higher accuracy but require more energy can be stimulated from time to time for short intervals.

Such control systems foster an increase in the working life of a SC because elements of the attitude control system can operate at reduced to minimum strength, generating only a weak control torque and operating during short time intervals. To increase the duration of active operation of a particular space system, we may use a *low cost SC* which can be readily replaced in case of failure. The emergence on the market of low cost SC, has radically altered the approach of users to mission planning and to SC development.

With respect to low cost SC, which sometimes also are called *small*

*I.D.Rodionov and M.Yu.Ovchinnikov. Optical Tomograph of the Universe. *Acta Astronautica,* April 1999, Vol.44, Issue 2-3, pp.211-213.

SC, it is proper to clarify what is implied. There are presently two trends in the development of small SC.

The first trend is characterized by the low cost of such SC. This is achieved through the use of: off-the-shelf technologies; a short time required for development and fabrication and the use of a piggy-back launch. This approach results in cost reductions that make small SC especially attractive for companies embarking on space activity and for universities.

The second trend is to use payload and on-board systems developed with the aid of modern high (micro and nano) technology. Cost in this case is not a critical criterion for designers and 'smallness' is only an external attribute.

For small SC, passive and combined attitude control systems are widely used because of their inherent simplicity, reliability, short duration of development and fabrication and low cost.

Let us next review some particular aspects of attitude operations and, thereafter, consider the general background of SC attitude dynamics required for the development of passive systems.

14.2 Scheme for SC Active Attitude Control

Let $Ox_1x_2x_3$ be a body-fixed reference system such that its axes are the principal central axes of the tensor of inertia of a SC. Then, the attitude motion of the SC assuming that it does not contain any internal moving elements (i.e. the SC is a rigid body), can be described using *Euler's dynamical equations*

$$\begin{aligned} A\dot{\omega}_1 + (C-B)\omega_2\omega_3 &= M_1, \\ B\dot{\omega}_2 + (A-C)\omega_3\omega_1 &= M_2, \\ C\dot{\omega}_3 + (B-A)\omega_1\omega_2 &= M_3. \end{aligned} \tag{14.1}$$

Here, ω_1, ω_2, ω_3 represent projections of the SC absolute angular velocity vector $\boldsymbol{\omega}$ onto the respective axes Ox_1, Ox_2, Ox_3. Also M_1, M_2, M_3 are projections of the torque M onto the same axes; *A, B, C* are moments of inertia relative to the introduced axes. In practice, the angular velocity of the SC and, consequently of its projections, are

small (the reason for this will be made clear later). Thus, we can neglect terms containing products of the angular velocity projections. Euler's dynamical equations then have the simple form

$$\begin{aligned} A\dot{\omega}_1 &= M_1, \\ B\dot{\omega}_2 &= M_2, \\ C\dot{\omega}_3 &= M_3. \end{aligned} \tag{14.2}$$

The projections of the torque at the right side of equations (14.2) are of two kinds, namely, the active control torque $\mathbf{M}_a$ generated by the attitude control system and the external perturbing torque $\mathbf{M}_b$ developed by the interaction between the SC and its environment. Let the torque $\mathbf{M}_b$ vanish so that $\mathbf{M} = \mathbf{M}_a$. Then, we can choose a variant of $\mathbf{M}_a$ that simplifies the logic of the attitude control process. If it is assumed that the angles φ_1, φ_2, φ_3 of turn about the introduced axes of the body-fixed reference system are small, we may write the kinematic relationships in the following form

$$\omega_1 = \dot{\varphi}_1, \qquad \omega_2 = \dot{\varphi}_2, \qquad \omega_3 = \dot{\varphi}_3. \tag{14.3}$$

Next, the logic of the attitude control algorithm is selected using the formulae

$$\begin{aligned} M_1 &= M_1(\varphi_1, \dot{\varphi}_1), \\ M_2 &= M_2(\varphi_2, \dot{\varphi}_2), \\ M_3 &= M_3(\varphi_3, \dot{\varphi}_3), \end{aligned} \tag{14.4}$$

assuming that on-board sensors measure the angles φ_1, φ_2, φ_3 and also the components ω_1, ω_2, ω_3 of angular velocity. In this way, Euler's equations (14.2) become mutually independent. In the mechanical sense, this means that the spatial attitude motion of the SC can be described it terms of a combination of three planar motions. As all three of these planar motions are completely similar, we may study any one of them

$$J\ddot{\varphi} = M_a(\varphi, \dot{\varphi}), \tag{14.5}$$

omitting indexing for simplification. Here J denotes the moment of inertia of the SC around its axis of rotation. The dependence of M_a on φ and $\dot{\varphi}$ is inherently non-linear and, because of this, analytical solving

of (14.5) is difficult. The non-linearity of the expression $M_a(\varphi, \dot{\varphi})$ is related, in particular, to the fact that active attitude control is often exerted using a thruster. Such an engine can be in one of two states, either off or at full thrust. The torque $\mathbf{M}_a$ is, thus, a discontinuous function.

Let us assume that the torque $M_a(\varphi, \dot{\varphi})$ causes the values of φ and $\dot{\varphi}$ to be close to zero, and that these values endure for a long time. In other words, the attitude control procedure, starting at a moment $t = 0$ with arbitrary, but finite, values of φ and $\dot{\varphi}$, is described at the moment $t = t_1$ by the relationships $|\varphi| \leq \varphi_1$ and $|\dot{\varphi}| \leq \dot{\varphi}_1$, where φ_1 and $\dot{\varphi}_1$ are small positive values. For any $t > t_1$ these constraints on the magnitudes of φ and $\dot{\varphi}$ are valid. Thus, at $t \geq t_1$ the kinematic equations (14.3) and, following from them, the dynamical equations (14.5), are also valid.

14.2.1 Estimation of Propellant Consumption

We estimate next the consumption of propellant necessary for the maintenance of SC attitude motion with small values of φ and $\dot{\varphi}$. First, we pay attention to an important feature of attitude control in space, namely that the absence of a dense ambient atmosphere results in practically no dissipation of the energy of SC attitude motion during the orientation process. Consequently, even having directed the axis of the SC in the necessary direction with $\varphi = 0$, or more realistically with $|\varphi| \leq \varphi_1$, and having also reduced the angular velocity of rotation to a minimum value satisfying the inequality $|\dot{\varphi}| \leq \dot{\varphi}_1$ (generally speaking, $\dot{\varphi}$ is not equal to zero), we still cannot expect the condition $|\varphi| \leq \varphi_1$ to be maintained. Rather, the residual angular velocity $\dot{\varphi}$ such that $0 < |\dot{\varphi}| \leq \dot{\varphi}_1$, results in a slow drift of the SC axis of orientation from its position with $\varphi = 0$, leading to a violation of the condition $|\varphi| \leq \varphi_1$.

The presence of residual angular velocity is inevitable because neither exact measuring of $\dot{\varphi}$, nor provision of an exact value of the pulse of the thruster reducing $\dot{\varphi}$ to zero, is possible. Thus, permanent angular positioning in a given direction with $\varphi = 0$ is, in practice, replaced by oscillations about this position with amplitude φ_1. This is because, when the magnitude of the angle φ reaches its extreme allowed value φ_1, a thruster 'pushes' the SC in a direction to reduce

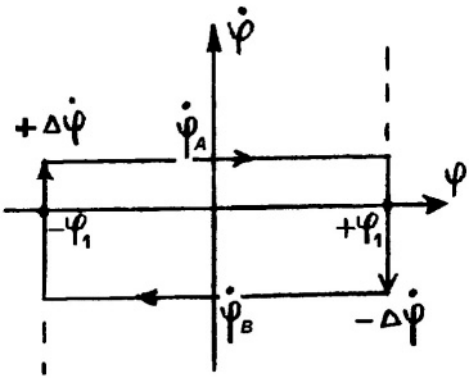

Figure 14.1. Oscillations in the phase plane due to the application of instant pulses

φ to zero. These thruster pulses should, of course, be minimized to avoid excessive consumption of propellant.

Next, we analyze the above mentioned oscillations of the SC about the position $\varphi = 0$ using the *phase-plane method.* Here, φ and $\dot{\varphi}$ are plotted along the abscissa and ordinate respectively. Let an 'instant' pulse $\dot{\varphi}$ developed by a thruster have its minimum possible magnitude $\Delta\dot{\varphi}$. Then, assuming that the SC is maintained in the neighbourhood of $\varphi = 0$ by such pulses, oscillations occur as illustrated in Fig. 14.1.

The logic of control illustrated by Fig. 14.1 is very simple. The thrusters are activated to impart instant pulses $\pm\Delta\dot{\varphi}$ to the SC when the phase-plane trajectory reaches *change-over lines* described by the equations $\varphi = +\varphi_1$ and $\varphi = -\varphi_1$. On the left, vertical, change-over line, a pulse $-\Delta\varphi$ is applied and, on the right vertical line, a pulse $+\Delta\varphi$ is imparted. The change-over lines are represented by dashed lines. Motion along the horizontal straight lines is free. Along the vertical lines it is forced by thruster action and, therefore, propellant consumption is required. We will now estimate this consumption.

Let a thrust P be developed by a thruster with arm L (L is the distance between the center of mass of the SC and the line of thrust), so that the torque developed is PL. Because the thrust $P = w\dot{m}$, where w is the exhaust velocity of a propulsive mass expelled from the nozzle of the thruster, and $\dot{m}$ is the *specific consumption* of the

propellant (i.e. the consumption of propellant per unit time), then

$$\Delta\dot{\varphi} = \frac{wL\dot{m}}{J}\Delta t. \tag{14.6}$$

Here Δt is the firing time of the thruster. The assumption that the pulse is instantaneous was already made above. This is a reasonable idealization as Δt is small compared with the period of oscillation T_k. However, now it is necessary to take into account that Δt is finite, because the process of imparting the pulse $\Delta\dot{\varphi}$ to the SC is considered, and Δm, which represents the consumption of propellant during one firing of the thruster, is given by

$$\Delta m = \dot{m}\Delta t. \tag{14.7}$$

Thus, using equality (14.7) we can transform (14.6) as follows

$$\Delta\dot{\varphi} = \frac{wL}{J}\Delta m. \tag{14.8}$$

This expression shows that the value $\Delta\dot{\varphi}$ is proportional to the quantity of consumed propellant Δm but the duration of firing (Δt) is not relevant. The assumption concerning the application of an instant pulse, made when constructing the phase-plane diagram (Fig. 14.1) is thus a reasonable one.

Obviously, during a complete cycle of oscillation with period T_k, two firings of the thruster are made and the total consumption of propellant is

$$2\Delta m = \frac{2J}{wL}\Delta\dot{\varphi}. \tag{14.9}$$

The specific consumption of propellant averaged over the cycle (i.e. the *time-averaged specific consumption*), is determined by

$$\overline{m} = \frac{2\Delta m}{T_k}. \tag{14.10}$$

It is seen from Fig. 14.1 that the value T_k is determined by

$$T_k = \frac{2\varphi_1}{|\dot{\varphi}_A|} + \frac{2\varphi_1}{|\dot{\varphi}_B|} = 2\varphi_1\frac{\Delta\dot{\varphi}}{|\dot{\varphi}_A\dot{\varphi}_B|}, \tag{14.11}$$

when the equality $\Delta\dot{\varphi} = |\dot{\varphi}_A| + |\dot{\varphi}_B|$ is taken into account. Substitution of (14.9) and (14.11) in (14.10) allows us to write

$$\overline{m} = \frac{J}{wL} \frac{|\dot{\varphi}_A \dot{\varphi}_B|}{\varphi_1}. \tag{14.12}$$

The time-averaged specific consumption $\overline{m}$ depends on the magnitudes of the angular velocities $|\dot{\varphi}_A|$ and $|\dot{\varphi}_B|$ over the oscillation period T_k. It is easy to show, taking into account the relationship $|\dot{\varphi}_A| + |\dot{\varphi}_B| = \Delta\dot{\varphi}$, that the maximum value $\overline{m}_{max}$ of $\overline{m}$ is reached when $\Delta\dot{\varphi}/2 = |\dot{\varphi}_A| = |\dot{\varphi}_B|$. Thus, from (14.12) we obtain a formula for $\overline{m}_{max}$

$$\overline{m}_{max} = \frac{J}{wL} \frac{(\Delta\dot{\varphi})^2}{4\varphi_1}. \tag{14.13}$$

In practice, the ratio between $|\dot{\varphi}_A|$ and $|\dot{\varphi}_B|$ is a random value because it is a result of transient motion preceding the oscillation phase. SC designers should provide a reserve of propellant to allow for possible variations in this ratio and, thus, calculation of the required propellant load should be based on (14.13). This relationship indicates the strategy of attitude control in the absence of an external perturbing torque. It is seen that the maximum time-averaged specific consumption of propellant decreases in inverse proportion to the orientation accuracy φ_1 and in proportion to the square of the instant pulse $\Delta\dot{\varphi}$.

Designers of SC should, generally speaking, develop the attitude control system so that (a) it does not provide more precise orientation than is required and (b) the control pulses are as weak as possible.

14.2.2 Effect of a Constant Perturbing Torque

Consider attitude control in the presence of an external perturbing torque $\mathbf{M}_b$. Introduce, to simplify the calculations, two assumptions, namely that (a) the external perturbing torque is constant ($\mathbf{M}_b = \text{const}$) and (b) that its magnitude is 'large'. The reason for the second assumption will become clear further on. At nonzero M_b equation (14.5) attains the following form

$$J\ddot{\varphi} = M_a(\varphi, \dot{\varphi}) + M_b. \tag{14.14}$$

Suppose that the magnitude of the control torque M_a is much greater than the magnitude of the perturbing torque M_b when the thruster is fired. Then it is possible to use our previous assumption that, on firing a thruster, there is an instant change $\Delta\dot{\varphi}$ in the angular velocity of the SC. The attitude motion of the SC is in consequence composed of two types of motion. The first is the instant change of angular velocity produced on firing the thruster (the action of the torque M_b is associatively negligible). The second is motion under the action of the torque M_b when the thruster is switched off.

Let the state $(\varphi, \dot{\varphi})$ correspond to a point A (called a *phase-plane point*), in the phase plane shown in Fig. 14.2. Point A results from imparting an instant pulse $+\Delta\dot{\varphi}$ to the SC which then reaches an angle of orientation $\varphi = -\varphi_1$. After this, the phase-plane point moves

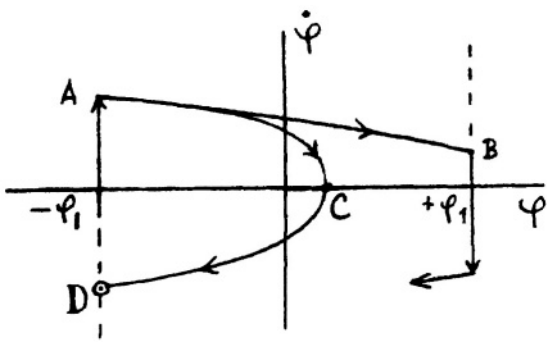

Figure 14.2. Oscillation process under the action of instant pulses in the phase plane in the presence of a perturbing torque

to the right along a parabola with its axis of symmetry coincident with the abscissa[†]. If the magnitude of M_b is 'small', the phase-plane point reaches the change-over line $\varphi = +\varphi_1$ at B, where the SC receives a pulse $-\Delta\dot{\varphi}$. If the magnitude of M_b is 'large', it can return to the change-over line $\varphi = -\varphi_1$ without reaching the change-over line $\varphi = +\varphi_1$ (i.e. it follows the parabola ACD).

Consider next the period of SC oscillation for the case of a large torque M_b. The parabola ACD returns the phase-plane point to the

[†]To show that, under the action of a constant torque M_b the phase-plane point follows this trajectory, one needs to integrate equation (14.14) with $M_a = 0$; then to exclude explicit time from the expressions for $\varphi = \varphi(t)$ and $\dot{\varphi} = \dot{\varphi}(t)$ in order to obtain the relationship between $\dot{\varphi} = \dot{\varphi}(\varphi)$ shown in Fig.14.2 for $M_b < 0$.

change-over line $\varphi = -\varphi_1$. However, point D is located, generally speaking, at a distance from A which is not equal to $\Delta\dot{\varphi}$. Therefore, the application of this pulse at D does not return the SC to A and so the cycle of oscillation is not completed. To complete the cycle, sequential motion along two parabolas is required.

Let the sequence start at point A with SC angular velocity $\dot{\varphi}_A$ (Fig. 14.3). At the point D, by virtue of parabolic symmetry, the

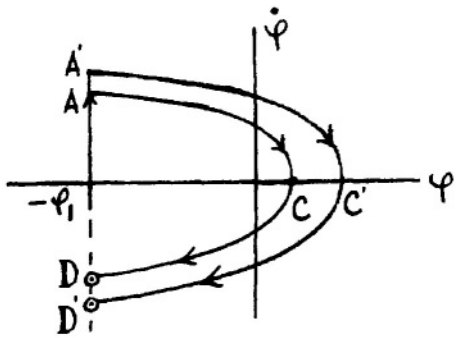

Figure 14.3. Cyclic oscillation process along two sequential parabolas

relation $\dot{\varphi}_D = -\dot{\varphi}_A$ is valid. Pulse $+\Delta\dot{\varphi}$ transfers the phase-plane point along the vertical change-over line to point A' with ordinate $-\dot{\varphi}_A + \Delta\dot{\varphi}$. Moving along the second parabola, the phase-plane point reaches a point D' on the change-over line which is symmetrical with respect to point A', again by virtue of parabolic symmetry. Then, the ordinate of point D' is $+\dot{\varphi}_A - \Delta\dot{\varphi}$. Finally, a pulse $+\Delta\dot{\varphi}$ returns the phase-plane point to A with ordinate $+\dot{\varphi}_A$, and the cycle is completed.

Let us next compute the period of a complete cycle of oscillation of the SC. First, we compute the duration of motion along the first parabola. In this case, equation (14.14) has the form

$$J\ddot{\varphi} = M_b.$$

Integrating, we obtain the solution

$$J\dot{\varphi} = M_b t + c, \tag{14.15}$$

where c is a constant of integration. Taking into account that at $t = 0$ the velocity $\dot{\varphi} = \dot{\varphi}_A$, we obtain $c = J\dot{\varphi}_A$. Next, considering half of

the arc AC of the parabola (Fig.14.3) and denoting the duration of motion along the first parabola by T_1, on substituting $\dot{\varphi}$ from (14.15), we can write

$$-J\dot{\varphi}_A = M_b\frac{T_1}{2}.$$

Again, denoting the duration of motion along the second parabola by T_2, we similarly obtain

$$-J(-\dot{\varphi}_A + \Delta\dot{\varphi}) = M_b\frac{T_2}{2}.$$

We can now write an expression for the total period of oscillation $T_k = T_1 + T_2$, that is,

$$T_k = 2\frac{J\Delta\dot{\varphi}}{|M_b|}. \tag{14.16}$$

Here again the inequality $M_b < 0$ was taken into account.

During a complete cycle of oscillation of the SC, two pulses $\Delta\dot{\varphi}$ were imparted and, therefore, the same quantity of propulsion propellant discussed in Section 14.2.1 was expended, (see (14.9)). The time-averaged specific consumption defined by expression (14.10) on substituting T_k from (14.16), is given by

$$\overline{m} = \frac{|M_b|}{wL}. \tag{14.17}$$

On comparing (14.17) with the relationship (14.12) we derived earlier, basic differences between them immediately emerge. In the absence of an external perturbing torque, the consumption of propellant (14.13) depends on the orientation accuracy φ_1 and on the pulse $\Delta\dot{\varphi}$. In the case of a 'large' perturbing torque, the consumption of propulsion propellant (14.17) is determined only by this torque and does not depend on the characteristics of the attitude control system (the orientation accuracy, the applied pulses etc.).

Experience shows that, for a large SC, the counteraction of external perturbing torques is a major task of the attitude control system requiring the expenditure of most of the propellant. Consequently, the reduction of such propellant expenditure is a key problem for SC designers. Such a reduction can be achieved (see above) by customizing the design of the SC to combat M_b but not by changing the parameters of the attitude control system.

Two specific cases have already been considered, that is, when the external torque was absent and when it was 'large'. Analysis of a more general case, the details of which are omitted here owing to its complexity, shows that the consumption $\overline{m}$ combines the two terms represented by (14.17) and (14.13).

Finally, it is necessary to note that the two terms of the propellant consumption discussed above have each a completely different nature. The consumption required to compensate for external perturbing torques cannot be avoided. On the other hand, the consumption when $\mathbf{M}_b = 0$ is due to the absence of a 'natural' dissipation of mechanical energy in space. In other words, if one could provide asymptotic stability of the equilibrium position, there would be no need to impart a pulse $\Delta\dot{\varphi}$ to ensure an accuracy φ_1. A method to achieve this in practice involves the use of gyroscopes.

14.3 SC Gyros for Attitude Control

Spinning rotors installed on a SC to change or to maintain its orientation, or to determine the attitude and angular velocity of the SC, are called *gyroscopes* or *gyros.* We will consider here only the application of generating a control torque. Gyros have angular momentum. By changing their vector of angular momentum, we can apply a reaction torque to a SC to turn its body, thereby providing attitude control. We will discuss next the general properties of gyros without referring to their specific design.

Let $\mathbf{G}$ be the vector of total angular momentum of a SC and let $\mathbf{G}_0$ be its initial value. Then, the attitude motion of the SC can be described by

$$\mathbf{G} = \mathbf{G}_0 + \int_0^t \mathbf{M}_b dt, \tag{14.18}$$

where $\mathbf{M}_b$ is the total vector of the external perturbing torques. The orientation thrusters are considered to be switched off. In analyzing (14.18), it is necessary to distinguish between two cases:

- the torque $\mathbf{M}_b$ changes with time so that $\mathbf{G}$ is a bounded function $|\mathbf{G}| \leq G_1 > 0$ (for example, $\mathbf{M}_b$ either is a periodic function of time with a zero average value, or $\mathbf{M}_b = 0$);

- the torque $\mathbf{M}_b$ changes so that $|\mathbf{G}| \to \infty$ with time (for example, $\mathbf{M}_b$ either is a periodic function of time with a nonzero average value, or $\mathbf{M}_b = \text{const} > 0$).

The angular momentum $\mathbf{G}$ of the SC consists of two components,

$$\mathbf{G} = \mathbf{K} + \mathbf{H} \tag{14.19}$$

where $\mathbf{K}$ represents the angular momentum of the SC body and $\mathbf{H}$ is the angular momentum of the gyros. The moments of inertia of the SC body necessary for determining $\mathbf{K}$ are calculated assuming that (a) the axes of the gyros are fixed relative to the SC body and (b) the angular momentum $\mathbf{H}$ is calculated with respect to the SC body.

The task of the attitude control system is to implement a required attitude motion of the SC. This motion is completely specified by $\mathbf{K}$. Therefore, suppose that the vectors $\mathbf{K}(t)$ and $\mathbf{G}(t)$ are given. The first vector is determined by the required orientation of the SC, the second by the external torques using (14.18). Thus, the value of $\mathbf{H}$ necessary to provide the required attitude motion of a SC affected by known external torques is determined by

$$\mathbf{H}(t) = \mathbf{G}(t) - \mathbf{K}(t). \tag{14.20}$$

Consider attitude motion of a SC in the vicinity of its steady-state orientation. Then, $\mathbf{K}(t)$ is variable but bounded in magnitude. There are two ways (see above) in which the vector $\mathbf{G}(t)$ changes.

14.3.1 Bounded Angular Momentum of a SC

In the first way, the magnitude of the vector $\mathbf{G}(t)$ is bounded. The task of the SC designer is to choose a combination of gyros with a magnitude of $\mathbf{H}(t)$ that exceeds the necessary magnitude of $\mathbf{H}(t)$ defined by (14.20) so that there is spare capability.

In general, the set of n control gyros has total angular momentum

$$\mathbf{H} = \sum_{i=1}^{n} \mathbf{h}_i = \sum_{i=1}^{n} \boldsymbol{\omega}_i J_i, \tag{14.21}$$

where $\mathbf{h}_i$ represents the angular momentum; $\boldsymbol{\omega}_i$ the angular velocity and J_i the moment of inertia of the i-th gyro. The equality (14.21)

shows that, by varying $\boldsymbol{\omega}_i$, change of the vector **H** and, consequently, attitude control of the SC can be achieved. This is realized by changing the magnitude of $\boldsymbol{\omega}_i$ through either accelerating or braking the spin rotation of the gyro, or by turning the vector $\boldsymbol{\omega}_i$ as a result of turning the spin axis of the gyro.

The first method (i.e. changing the magnitude of $\boldsymbol{\omega}_i$), although it is sometimes used to control a SC, is not economical. This is because changing the spin rate requires an expenditure of energy which increases in proportion to the square of the spin rate, whereas the useful effect (changing the angular momentum), is only proportional to the spin rate itself. The advantage of the method lies in the simplicity of the gyro design, due to the fact that the spin axis is fixed with respect to the SC body and, thus, no gimbal is used. This method because of its simplicity, coupled with the reliability and low cost of the gyros, is frequently utilized aboard small SC.

The second method is to turn the spin axis of the gyro. This requires, for its realization, more complex hardware and software than the first. Its advantage lies in the rather small associated consumption of energy, which is lost only in overcoming friction in the bearings of the spin axis to maintain a constant spin rate, and in compensating for other such losses. Irrespective of the design of the gyros, attitude control of a SC can be achieved if the magnitude of vector $\mathbf{G}(t)$ is bounded and if vector $\mathbf{H}(t)$ satisfies (14.20).

14.3.2 Unbounded Angular Momentum of a SC

If the magnitude of vector $\mathbf{G}(t)$ is unbounded, the necessary condition (14.20) for achieving control is, generally speaking, not satisfied. As the number of gyros, their moments of inertia and maximum spin rates, are limited, the magnitude of the vector **H** is limited also. To maintain the validity of (14.20) during an unbounded increase in the magnitude of vector **G**, the magnitude of vector **K** has to similarly increase. This leads to an unbounded increase in the angular velocity of the SC so that its original orientation is lost. There is then no possibility to control the SC and it is governed by the perturbing torque M_b.

The case where perturbing torques have a constant component is practically always met with and designers should build capability into

the attitude control systems to counteract this effect. A way to do this is to apply an external control torque. The reaction torque developed by a gyro can redistribute the composite angular momentum of the SC between its parts (the gyros and SC body), but cannot change the total value (this follows from the theorem of mechanics concerning the effect of internal forces). It is, therefore, necessary to install a dedicated control loop which applies an external control torque to the SC. Actuators of the control loop could comprise thrusters or magnetorquers that interact with the magnetic field of the Earth.

It might be inferred that the necessity to utilize thrusters as well as gyros makes the use of gyros unattractive. However, that is not true. If gyros are installed on a SC, the large consumption of propellant (14.12) required to provide the SC with precise orientation using thrusters is not called upon. This is because the task can be accomplished by the gyros themselves using not on-board propellant but, rather, the renewable electrical energy resource provided by solar arrays.

Consider next the consumption of propellant determined by (14.18) for overcoming the perturbing torque, taking into account in this formula only the constant component of $\mathbf{M}_b$, because its periodic component is compensated for by the gyros. In practice the constant component is significantly less than the periodic component, and so the use of control gyros considerably reduces the consumption of propellant.

This discussion results in the conclusion that, for SC which are required to have a long life time and accurate orientation, the use of gyros is unavoidable. This is because gyros spend regeneratable energy, in contrast to the propellant consumed by thrusters. Also, the electrical energy consumed by gyros does not, in general, depend on the accuracy of orientation or on the angular velocity of the SC. Further, the electrical energy expended is, in practice, rather modest.

Chapter 15

SC Affected by a Gravity-Gradient Torque

To describe the angular motion of a SC affected by a torque due to the gravitational field of the Earth, we need to introduce several reference systems; to derive an expression for the gravity-gradient torque and to obtain the associated equations of motion. We also should consider how to determine angular equilibrium positions and periodic motions of SC under the influence of such a torque and provide ways to analyze their stability. Before this, however, we will introduce several general assumptions that will also be applied in subsequent Chapters.

15.1 General Assumptions

We assume that a SC is a rigid body. This is the case if the natural frequencies of the SC structure are much higher than the lower frequencies of the external torques affecting the SC (which is true, in particular, for small SC). Also, we assume that the center of mass of the SC moves in a Keplerian orbit (i.e. the orbit is fixed in inertial space), and that the gravitational field of the Earth is central and Newtonian. To show the validity of these two latter assumptions, consider SC motion in the Earth's gravitational field. The associated equations of motion written in terms of osculating elements and averaged over the true anomaly, where the second term of the geopotential is taken

into account, have the following first integrals (see Section 7.5)

$$a = \text{const}, \qquad e = \text{const}, \qquad i = \text{const} \tag{15.1}$$

where a is the semi-major axis, e the eccentricity and i the inclination of the orbital plane. Thus the Earth's oblateness does not affect the form and size of the orbit or its inclination. Orbit evolution consists only of a change of longitude Ω -of the ascending node of the orbit and of its perigee argument ω_π. The equations with right parts averaged over the true anomaly describing this evolution are as follows

$$\frac{d\Omega}{dt} = -\frac{3}{2} J_2 n_{ev} \left(\frac{R_e}{p}\right)^2 \cos i, \tag{15.2}$$

$$\frac{d\omega_\pi}{dt} = -\frac{3}{4} J_2 n_{ev} \left(\frac{R_e}{p}\right)^2 (1 - 5\cos^2 i) \tag{15.3}$$

where $n_{ev} = (\mu_g/a^3)^{1/2}$ is *mean motion* (compare with Section 1.6), μ_g denotes the *gravitational parameter of the Earth,* R_e is the equatorial radius of the Earth and $J_2 = -1\,082.2 \cdot 10^{-6}$ is a zonal harmonic coefficient.

As a result of the Earth's oblateness, the plane of the orbit precesses with angular velocity (15.2), while the orbit itself rotates within this plane with angular velocity (15.3). For a sun-synchronous, low Earth, orbit with altitude less than 1 000 km, the *angular velocity of the line of nodes* (sometimes called the *angular velocity of the ascending node*) is about 0.06 deg/orbit and the angular velocity of the line of apsides (sometimes called the *angular velocity of perigee*), is about 0.2 deg/orbit. These values which are, at least, three orders of magnitude smaller than the angular velocity of SC orbital motion, both contribute to the centrifugal forces that affect the attitude motion of the SC. The torque developed by the centrifugal forces is roughly proportional to the square of the associated angular velocity. Thus, the torque due to the orbital motion of the SC around the Earth, is 10^6 times greater than the torque produced due to the rotation of the orbit of the SC, so that its influence can be neglected. It is thereby demonstrated that the assumptions made above concerning the orbit of the SC and the Earth's gravitational field are reasonable in a general context.

15.2 Reference Systems (IRS / ORS / BRS)

We will now introduce three Cartesian reference systems:
— $O_aX_{a1}X_{a2}X_{a3}$ is an *inertial reference system* (IRS). Its origin O_a is placed at the Earth's center of mass. Axis O_aX_{a3} is aligned along the Earth's axis of rotation. Axes O_aX_{a1}, O_aX_{a2} lie in the Earth's equatorial plane so that axis O_aX_{a1} points to the *Vernal equinox* and axis O_aX_{a2} completes the right-hand orthogonal reference system. The term inertial in the name of this reference system means that it can be considered as absolute, or inertial, with an accuracy acceptable for our present purposes;
— $OX_1X_2X_3$ is an *orbital reference system* (ORS). Its origin O is placed at the SC's center of mass. Axis OX_3 is aligned along the radius-vector from the center of mass of the Earth to the SC center of mass. Axis OX_2 is aligned along the orbit normal. (An observer located on this axis would see a SC orbiting around the Earth in a counter-clockwise direction.) Axis OX_1 completes the right-hand orthogonal system. It is perpendicular to OX_3, lies in the orbital plane and forms an acute angle with the vector of the center of mass velocity of the SC; $\mathbf{E}_1$, $\mathbf{E}_2$, $\mathbf{E}_3$ are unit vectors with respect to the ORS axes;
— $Ox_1x_2x_3$ is a *body-fixed reference system* (BRS). Its axes are directed along the principal axes of the tensor of inertia of the SC (i.e. its origin coincides with the center of mass of the SC) and $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ are unit vectors of the BRS axes.

Let an arbitrary vector $\mathbf{a}$ be represented by its projections a_1, a_2, a_3 on a reference system. We designate the system the vector is projected onto by a sub-index corresponding with the letter used in the notation of the axes of that system. For instance, for projecting onto the axes of a body-fixed reference system $Ox_1x_2x_3$ we write $\mathbf{a} = (a_1, a_2, a_3)_x$.

Consider a transfer from the orbital to the body-fixed reference system. Plane angles α, β, γ (Fig.15.1) can be used for representing this transfer. When these angles are small, they are named *pitch, yaw,* and *roll* respectively. The *transformation matrix* of *directional cosines*

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \tag{15.4}$$

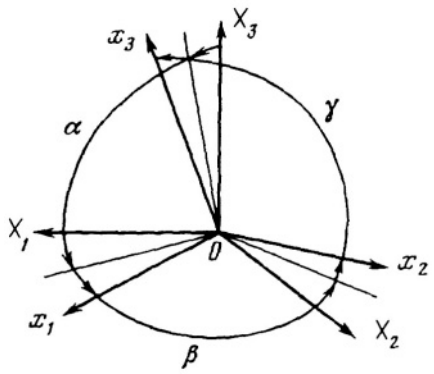

Figure 15.1. Representation of a transfer from the orbital to the body-fixed reference system using plane angles α, β, γ

which transforms the components of a vector from one reference system to another is sometimes called the *directional cosine matrix* and denoted by $\|a_{jk}\|$, because its elements are defined through scalar products of unit vectors $\mathbf{E}_j, \mathbf{e}_k$ $(j,k = 1,2,3)$ by $a_{jk} = \cos(\mathbf{E}_j, \mathbf{e}_k)$. This matrix can be conveniently represented in tabular form (Table 15.1). Such a representation allows us to easily transform a vector

Table 15.1. Tabular form of a matrix of transformation

	x_1	x_2	x_3
X_1	a_{11}	a_{12}	a_{13}
X_2	a_{21}	a_{22}	a_{23}
X_3	a_{31}	a_{32}	a_{33}

specified by its components in one reference system, by summing the paired products of the components of a string-vector in BRS with a string-vector of the matrix to obtain corresponding components of the column-vector in ORS (and vice versa). For example,

$$X_2 = x_1 a_{21} + x_2 a_{22} + x_3 a_{23}.$$

To specify the matrix elements, instead of representing them using the products of three matrices of *elementary rotation* (i.e. rotation

about a particular axis), we write a set of linear transformations that result from consecutive rotations of the reference system (Fig.15.2).

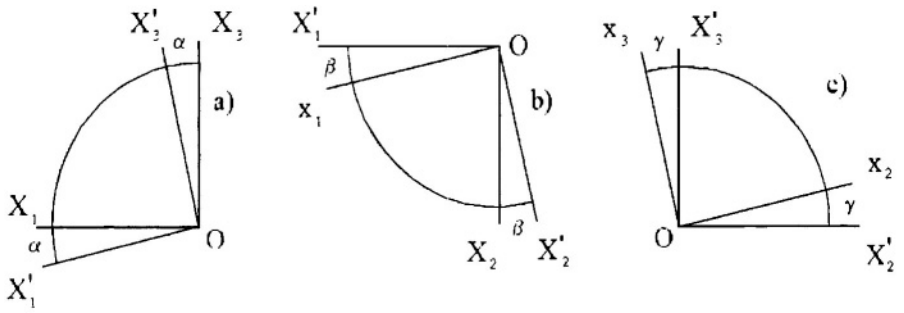

Figure 15.2. Transfer from the orbital to the body-fixed reference system by consecutive rotations through plane angles α, β, γ

Transfer now from ORS to BRS through three elementary rotations. The first rotation is performed around the axis OX_2 through angle α (Fig.15.2(a)), with linear transformation

$$X_1 = X_1' \cos\alpha + X_3' \sin\alpha, \quad X_3 = -X_1' \sin\alpha + X_3' \cos\alpha.$$

The second rotation is executed around the axis OX_3' through angle β (Fig.15.2 (b)) with

$$X_1' = x_1 \cos\beta - X_2' \sin\beta, \quad X_2 = x_1 \sin\beta + X_2' \cos\beta.$$

The last rotation is executed around the axis Ox_1 through angle γ (Fig.15.2 (c)), with

$$X_2' = x_2 \cos\gamma - x_3 \sin\gamma, \quad X_3' = x_2 \sin\gamma + x_3 \cos\gamma.$$

Substituting from one transformation to another we express X_1, X_2, X_3 through linear combinations of x_1, x_2, x_3. The factors of x_1, x_2, x_3 obtained by such substitutions are the elements of the directional cosine matrix A:

$$a_{11} = \cos\alpha \cos\beta$$
$$a_{21} = \sin\beta$$
$$a_{31} = -\sin\alpha \cos\beta$$

$$a_{12} = \sin\alpha \sin\gamma - \cos\alpha \sin\beta \cos\gamma \quad a_{13} = \sin\alpha \cos\gamma + \cos\alpha \sin\beta \sin\gamma$$
$$a_{22} = \cos\beta \cos\gamma \quad a_{23} = -\cos\beta \sin\gamma$$
$$a_{32} = \cos\alpha \sin\gamma + \sin\alpha \sin\beta \cos\gamma \quad a_{33} = \cos\alpha \cos\gamma - \sin\alpha \sin\beta \sin\gamma$$

Since a matrix A performs an orthogonal transformation with a unit determinant, a matrix A^{-1} which performs the reverse transformation can be obtained by transposing matrix A (i.e. $A^{-1} = A^{\top}$).

Similarly we introduce the transformation matrix $||c_{jk}||$ which transforms the components of a vector from ORS to IRS and which can be represented in tabular form (Table 15.2). Elements of the transforma-

Table 15.2. Tabular form of the matrix of transformation from ORS to IRS

	X_1	X_2	X_3
X_{a1}	c_{11}	c_{12}	c_{13}
X_{a2}	c_{21}	c_{22}	c_{23}
X_{a3}	c_{31}	c_{32}	c_{33}

tion matrix can be represented using osculating elements (see Fig. 5.1 and relationship (5.1))

$$
\begin{aligned}
c_{11} &= -\sin u \cos\Omega - \cos u \sin\Omega \cos i \\
c_{21} &= -\sin u \sin\Omega + \cos u \cos\Omega \cos i \\
c_{31} &= \cos u \sin i
\end{aligned}
$$

$$
\begin{aligned}
c_{12} &= \sin\Omega \sin i & c_{13} &= -\cos u \cos\Omega - \sin u \sin\Omega \cos i \\
c_{22} &= -\cos\Omega \sin i & c_{23} &= \cos u \sin\Omega + \sin u \cos\Omega \cos i \\
c_{32} &= \cos i & c_{33} &= \sin u \sin i.
\end{aligned}
$$

15.3 Kinematic Relationships

Obtain *kinematic relationships* between the components $(\omega_1, \omega_2, \omega_3)_x$ of SC absolute angular velocity $\boldsymbol{\omega}$ and the derivatives $\dot{\alpha}, \dot{\beta}, \dot{\gamma}$ of the transfer plane angles. The term *absolute angular velocity* means that this is an angular velocity of the SC with respect to the inertial reference system $O_a X_{a1} X_{a2} X_{a3}$.

The relationship between the absolute angular velocity of the SC $\boldsymbol{\omega}$, its angular velocity with respect to the ORS $\overline{\boldsymbol{\omega}}$ and the *orbital angular velocity* $\omega_{orb}\mathbf{E}_2$ is given by

$$\boldsymbol{\omega} = \omega_{orb}\mathbf{E}_2 + \overline{\boldsymbol{\omega}}. \tag{15.5}$$

The orbital reference system rotates around the orbit normal $\mathbf{E}_2$, which is aligned with the vector of the area integral **c** (see Section 1.3). The magnitude of the angular velocity is determined through the velocity $\dot{v}$ of rotation of the SC radius-vector around the Earth using the formulae (1.23),(1.22) and (1.19) as follows

$$\omega_{orb} = \omega_0(1 + e\cos v)^2 \tag{15.6}$$

where v is the true anomaly; $\omega_0^2 = \mu_g/p^3$ and $p = a(1-e^2)$ is the *parameter*. In a circular orbit where p is equal to the orbit radius, ω_{orb} reduces to ω_0.

Establish the relationships between the components $(\overline{\omega}_1, \overline{\omega}_2, \overline{\omega}_3)_x$ of angular velocity $\overline{\boldsymbol{\omega}}$ of SC rotation relative to the ORS derivatives $\dot{\alpha}, \dot{\beta}, \dot{\gamma}$ through

$$\begin{aligned} \overline{\omega}_1 &= \dot{\alpha}a_{21} + \dot{\gamma}, \\ \overline{\omega}_2 &= \dot{\alpha}a_{22} + \dot{\beta}\sin\gamma, \\ \overline{\omega}_3 &= \dot{\alpha}a_{23} + \dot{\beta}\cos\gamma \end{aligned} \tag{15.7}$$

where the first rotation around axis OX_2 has angular velocity $\dot{\alpha}$; the second rotation around axis OX_3' has angular velocity $\dot{\beta}$ and the third rotation around axis Ox_1 has angular velocity $\dot{\gamma}$.

On substituting the expression for $\overline{\boldsymbol{\omega}}$ in (15.5) we obtain the required kinematic relationships

$$\begin{aligned} \omega_1 &= \dot{\alpha}a_{21} + \dot{\gamma} + \omega_{orb}a_{21}, \\ \omega_2 &= \dot{\alpha}a_{22} + \dot{\beta}\sin\gamma + \omega_{orb}a_{22}, \\ \omega_3 &= \dot{\alpha}a_{23} + \dot{\beta}\cos\gamma + \omega_{orb}a_{23} \end{aligned} \tag{15.8}$$

which will later be used to complete the dynamical equations of SC attitude motion.

15.4 Derivation of the Expression for the Gravity-Gradient Torque

We next derive the expression for the gravity-gradient torque. Let **R** be a radius-vector from the center of mass of the Earth O_a to the SC center of mass O; **r** be a radius-vector from O_a to the small mass

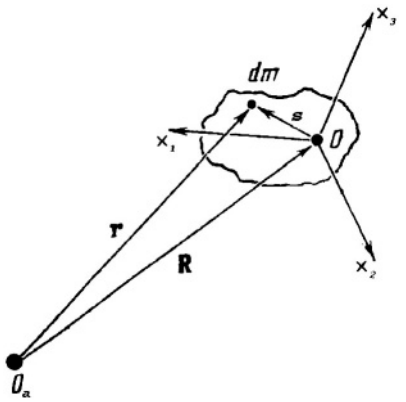

Figure 15.3. Configuration for calculating the gravity-gradient torque

element dm (particle), and $\mathbf{s}$ a radius-vector from this particle to the SC center of mass O (Fig.15.3).

The gravitational force acting on dm is determined by the Law of Universal attraction

$$d\mathbf{F} = -\frac{\mu_g dm}{r^3}\mathbf{r}. \tag{15.9}$$

A general formula for calculating the total torque exerted on a body of volume V by a force composed of elementary forces $d\mathbf{F}$, each applied at a point specified by a radius-vector $\mathbf{s}$ relative to an origin O, has the form

$$\mathbf{M}_0 = \int_V \mathbf{s} \times d\mathbf{F}.$$

Thus, using (15.9) for $d\mathbf{F}$ we obtain an expression for the gravity-gradient torque $\mathbf{M}_O$ with respect to the center of mass of a SC

$$\mathbf{M}_0 = -\mu_g \int_V \frac{\mathbf{s} \times \mathbf{r}}{r^3} dm. \tag{15.10}$$

Calculate the components of vector $\mathbf{M}_0$ in the BRS, using notations $\mathbf{s} = (s_{x1}, s_{x2}, s_{x3})_x$, $\mathbf{R} = (Ra_{31}, Ra_{32}, Ra_{33})_x$, the resultant $\mathbf{r} = \mathbf{R}+\mathbf{s} = (Ra_{31} + s_{x1}, Ra_{32} + s_{x2}, Ra_{33} + s_{x3})_x$ and the vector product

$$\mathbf{s} \times \mathbf{r} \tag{15.11}$$
$$= R[(s_{x2}a_{33} - s_{x3}a_{32}),\ (s_{x3}a_{31} - s_{x1}a_{33}),\ (s_{x1}a_{32} - s_{x2}a_{31})]_x.$$

Write the magnitude of the vector **r** in the form

$$r = R\left((a_{31} + \frac{s_{x1}}{R})^2 + (a_{32} + \frac{s_{x2}}{R})^2 + (a_{33} + \frac{s_{x3}}{R})^2\right)^{1/2}.$$

Assuming that the ratio *s/R* is vanishingly small, we can write the approximation in powers of *s/R* restricted to terms up to $\mathrm{O}(s^2/R^2)$

$$r \approx R\left[1 + \frac{1}{R}\left(a_{31}s_{x1} + a_{32}s_{x2} + a_{33}s_{x3}\right)\right].$$

Also, the series expansion $1/r^3$ may be written

$$\frac{1}{r^3} = \frac{1}{R^3}[1 - \frac{3}{R}(s_{x1}a_{31} + s_{x2}a_{32} + s_{x3}a_{33})] \tag{15.12}$$

where the expression in round brackets represents the projection of vector **s** onto the axis OX_3. Substitute (15.11) and (15.12) in (15.10) and integrate. Take into account the set of equalities

$$\int_V s_{x1}s_{x2}dm = 0, \quad \int_V s_{x1}s_{x3}dm = 0, \quad \int_V s_{x2}s_{x3}dm = 0 \tag{15.13}$$

which are valid because the centrifugal moments of inertia of a SC are equal to zero in the body-fixed reference system considered, due to the fact that its axes are coincident with the principal axes of the tensor of inertia of the SC.

Also, take into account the set of equalities

$$\int_V s_{X1}dm = 0, \quad \int_V s_{X2}dm = 0, \quad \int_V s_{X3}dm = 0 \tag{15.14}$$

which are valid since the origin *O* of the BRS coincides with the SC center of mass. After these substitutions we have

$$\begin{aligned}\mathbf{M}_0 = \frac{3\mu_g}{R^3}\int_V \{&[(s_{x1}^2 + s_{x2}^2) - (s_{x1}^2 + s_{x3}^2)]a_{32}a_{33}, \\ &[(s_{x2}^2 + s_{x3}^2) - (s_{x1}^2 + s_{x2}^2)]a_{31}a_{33}, \\ &[(s_{x1}^2 + s_{x3}^2) - (s_{x2}^2 + s_{x3}^2)]a_{31}a_{32}\}dm.\end{aligned}$$

Zero-value sums $s_{xj}^2 - s_{xj}^2$, $(j = 1, 2, 3)$ are next substituted so that we can use the notations

$$A = \int_V (s_{x2}^2 + s_{x3}^2) dm, \quad B = \int_V (s_{x1}^2 + s_{x3}^2) dm,$$
$$C = \int_V (s_{x1}^2 + s_{x2}^2) dm \tag{15.15}$$

for the principal central moments of inertia of the SC. Then, we can write the required expression for the gravity-gradient torque

$$\mathbf{M}_0 = \frac{3\mu_g}{R^3}[(C - B)a_{32}a_{33},\ (A - C)a_{31}a_{33},\ (B - A)a_{31}a_{32}]_x. \tag{15.16}$$

The elements a_{31}, a_{32}, a_{33} of the transformation matrix A contained in (15.16), are the components of the unit vector $\mathbf{E}_3$ of the local vertical in the body-fixed reference system. The 'physical' sense lies in the coincidence of the local vertical and the gradient of the central Newtonian gravitational field which gives rise to the torque $\mathbf{M}_0$.

For the general case, where the axes of another body-fixed reference system have arbitrary angular positions with respect to the principal axes of the SC tensor of inertia, the torque can be represented by

$$\mathbf{M}_0 = \frac{3\mu_g}{R^3} A^\top \mathbf{E}_3 \times J A^\top \mathbf{E}_3 \tag{15.17}$$

where J is the tensor of inertia of the SC in this new body-fixed reference system. If now $\mathbf{s} = (s_{x1}, s_{x2}, s_{x3})_x$ in this reference system, the tensor of inertia has the form

$$J = \begin{pmatrix} A & -D & -E \\ -D & B & -F \\ -E & -F & C \end{pmatrix}, \tag{15.18}$$

where the centrifugal elements of the tensor are

$$D = \int_V s_{x1}s_{x2}dm,\ E = \int_V s_{x1}s_{x3}dm,\ F = \int_V s_{x2}s_{x3}dm. \tag{15.19}$$

Its diagonal elements *A, B* and *C* have already been introduced in (15.15). Sometimes, instead of *A*, *B*, *C*, *D*, *E*, *F* the corresponding notations J_{11}, J_{22}, J_{33}, J_{12}, J_{13}, J_{23} are used.

15.5 Equations of SC Motion

To describe SC motion we may use Euler's *dynamical equations* which can be represented in vector form by

$$J\dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times J\boldsymbol{\omega} = \frac{3\mu_g}{R^3} A^\top \mathbf{E}_3 \times JA^\top \mathbf{E}_3. \tag{15.20}$$

which is valid for any arbitrary BRS. Consider this equation in its more commonly used scalar form by projecting (15.20) onto the axes of the BRS, which axes coincide with the principal axes of the SC tensor of inertia i.e. $J = \mathrm{diag}(A, B, C)$,

$$\begin{aligned} A\dot{\omega}_1 + (C-B)\omega_2\omega_3 &= \frac{3\mu_g}{R^3}(C-B)a_{32}a_{33}, \\ B\dot{\omega}_2 + (A-C)\omega_1\omega_3 &= \frac{3\mu_g}{R^3}(A-C)a_{31}a_{33}, \\ C\dot{\omega}_3 + (B-A)\omega_1\omega_2 &= \frac{3\mu_g}{R^3}(B-A)a_{31}a_{32}. \end{aligned} \tag{15.21}$$

These three dynamical equations contain six variables (three components of absolute angular velocity and three planar angles). To obtain their solution we should complement (15.21) with *kinematic equations* which express the derivatives of the planar angles in terms of the components of the SC absolute angular velocity. Resolving (15.8) with respect to these derivatives, we obtain Euler's *kinematic equations*

$$\begin{aligned} \dot{\alpha} &= \frac{1}{\cos\beta}(\omega_2\cos\gamma - \omega_3\sin\gamma) - \omega_{orb}, \\ \dot{\beta} &= (\omega_2\sin\gamma + \omega_3\cos\gamma), \\ \dot{\gamma} &= \omega_1 - (\omega_2\cos\gamma - \omega_3\sin\gamma)\tan\beta. \end{aligned} \tag{15.22}$$

Now we make a remark. All three-angle transformation sets are subject to singularity when the angles cannot define a unique configuration. In (15.22) this takes place when $\beta = \pi/2$ and $\beta = 3\pi/2$. In the close vicinity of these configurations, a small turn of the body is followed by a large changing of angles. To avoid this, one can choose another three-angle transformation set since the motion concerned is known in advance or use other variables, for instance quaternions in-

troduced by Hamilton* or elements of a transformation matrix. An increase in the number of new variables up to more than three allows us to avoid the problem of singularity (so called 'gimbal lock'). For example, choosing six elements $a_{11}, \ldots, a_{23}$ of the matrix (15.4) as new variables we can write the consequent kinematic equations, called the *Poisson*† kinematic equations

$$\dot{\mathbf{a}}_j = -\boldsymbol{\omega} \times \mathbf{a}_j, \quad (j = 1, 2) \tag{15.23}$$

where $\mathbf{a}_k = (a_{k1}, a_{k2}, a_{k3})_x$ are the components of unit vectors of axes OX_k of the ORS in projections on the BRS, $(k = 1, 2, 3)$. Another three elements of the matrix not contained in (15.23) can be calculated using the formula $\mathbf{a}_3 = \mathbf{a}_1 \times \mathbf{a}_2$.

We will next consider special motions of a SC affected by a gravity-gradient torque.

*Hamilton, William Rowan (1805-1865) Irish mathematician. Professor at Trinity College Dublin. In his *Theory of Systems of Rays* he predicted the existence of conical refraction and unified the field of optics under the principle of varying action. His reformulation of the equations of motion of Lagrange is a powerful tool in classical and in modern wave mechanics. He invented quaternions.

†Poisson, Siméon-Denis (1781-1840), French mathematician who was a pupil and friend of Laplace and Lagrange. He extended their work in celestial mechanics. Also, he made outstanding contributions in electricity and magnetism.

Chapter 16

SC Motion in a Circular Orbit

We begin with the motion of a SC in circular orbit. On transforming the general equations (15.21), (15.22) of SC attitude motion for the case of a circular orbit ($e = 0$), where $\omega_{orb} = \omega_0$ and $\mu_g/R^3 = \omega_0^2$, we get

$$
\begin{aligned}
&A\dot{\omega}_1 + (C - B)\omega_2\omega_3 = \omega_0^2(C - B)a_{32}a_{33},\\
&B\dot{\omega}_2 + (A - C)\omega_1\omega_3 = \omega_0^2(A - C)a_{31}a_{33},\\
&C\dot{\omega}_3 + (B - A)\omega_1\omega_2 = \omega_0^2(B - A)a_{31}a_{32}.\\
&\dot{\alpha} = \frac{1}{\cos\beta}(\omega_2\cos\gamma - \omega_3\sin\gamma) - \omega_0,\\
&\dot{\beta} = (\omega_2\sin\gamma + \omega_3\cos\gamma),\\
&\dot{\gamma} = \omega_1 - (\omega_2\cos\gamma - \omega_3\sin\gamma)\tan\beta.
\end{aligned}
\tag{16.1}
$$

First, consider the planar motion of a SC in circular orbit, which involves the motion of any two of its principal axes of inertia in the plane of the orbit.

16.1 Planar Motion of a SC

The equations (16.1) have the special solution

$$
\omega_1 = \omega_3 = \beta = \gamma = 0, \qquad \alpha = \alpha(t), \qquad \omega_2 = \omega_2(t), \tag{16.2}
$$

where $\alpha(t)$ and $\omega_2(t)$ satisfy the following equations

$$\dot{\alpha} = \omega_2 - \omega_0, \qquad B\dot{\omega}_2 = -3\omega_0^2(A - C)\sin\alpha\cos\alpha.$$

Eliminating ω_2 we obtain

$$\ddot{\alpha} + 3\omega_0^2\frac{A - C}{2B}\sin 2\alpha = 0, \tag{16.3}$$

which is similar to the equation of motion of a classical pendulum with one degree of freedom. The special solution satisfying the conditions

$$\omega_1^2 + \omega_3^2 + \sin^2\beta + \sin^2\gamma = 0$$

is for motion of a SC in an orbital plane (this is called *planar motion*), and can be contrasted with the solution satisfying the conditions

$$\omega_1^2 + \omega_3^2 + \sin^2\beta + \sin^2\gamma > 0,$$

which is for three dimensional or *spatial motion,* that is, motion where the axes of the SC can assume any arbitrary orientation relative to the orbital plane.

Equation (16.3) describes the natural planar motions of the SC and is integrated using elliptic functions. Multiplying both sides of (16.3) by 2α and integrating over time, we obtain

$$(\dot{\alpha})^2 + 3\omega_0^2\frac{A - C}{B}\sin^2\alpha = \frac{h_0}{B}, \tag{16.4}$$

where h_0 is an *energy integral.* Solving (16.4) with respect to α gives

$$\dot{\alpha} = \pm\sqrt{\frac{h_0}{B} - 3\omega_0^2\frac{A - C}{B}\sin^2\alpha}. \tag{16.5}$$

Before finalizing the integration, we consider SC motion in the phase plane $(\alpha, \dot{\alpha})$. Obtain first the equilibrium positions of the SC and the conditions for their stability. Substituting $\alpha = \text{const}$ into (16.3), we get $\sin\alpha = 0$ which has four, physically distinct, solutions $\alpha = \pi k/2$, k=0,1,2,3.

We will now derive *sufficient conditions for the stability* of these equilibrium positions using an approach based on *Lyapunov's* theorem*

*Lyapunov, Alexander Mikhailovich (1857–1918), Russian mathematician, professor at the Khar'kov University, who made a significant research input to the general theory of stability; deduced the ellipsoidal shapes of a spinning liquid and derived the screw motions of a rigid body in fluid media.

for stability†. We choose the first integral h_0 (16.4) as Lyapunov's function V_L. At the equilibrium position $\alpha = \dot{\alpha} = 0$, this function V_L is equal to zero. Outside of this equilibrium position, under the inequality $A - C > 0$, the function V_L is strictly greater than zero. Thus, function V_L is *positive-definite.* Using (16.3), the derivative of function V_L with respect to time is

$$\dot{V}_L = \frac{\partial V_L}{\partial \dot{\alpha}}\ddot{\alpha} + \frac{\partial V_L}{\partial \alpha}\dot{\alpha} = 2B\dot{\alpha}\left(\ddot{\alpha} + 3\omega_0^2\frac{A-C}{2B}\sin 2\alpha\right) = 0.$$

Necessary conditions for the stability of the trivial solution $\alpha = 0$ are determined from the conditions for the existence of a finite solution of equation

$$\ddot{\alpha} + 3\omega_0^2\frac{A-C}{B}\alpha = 0, \tag{16.8}$$

which is obtained by linearizing equation (16.3) in the vicinity of the equilibrium position $\alpha = 0$. An harmonic solution of this equation exists if the condition $A - C > 0$ is valid. This follows from the condition that imaginary roots of the *characteristic equation*

$$\lambda^2 + 3\omega_0^2\frac{A-C}{B}a = 0$$

†Consider autonomous differential equations

$$\dot{x}_j = X_j(x_1, \ldots, x_n), \qquad (j = \overline{1,n}) \tag{16.6}$$

with the equlibrium solution

$$x_1 = \ldots = x_n = 0, \tag{16.7}$$

which provides $X_j(0, \ldots, 0) = 0$. If there exists a continuous, sign-definite, scalar function, $V_L(x_1, \ldots, x_n)$ (for example, $V_L > 0$ for all $x_1^2 + \ldots + x_n^2 \neq 0$ and $V_L(0, \ldots, 0) = 0$), which has first partial derivatives with respect to all its variables, such that its time derivative $\dot{V}_L$ calculated along the motion described by (16.6) is either of constant sign; opposite to the sign of the function V_L,

$$\dot{V}_L = \sum_{j=1}^{n}\frac{\partial V_L}{\partial x_j}\dot{x}_j \leq 0, \quad \text{for all } x_1^2 + \ldots + x_n^2 \neq 0,$$

or identically equal to zero, then the trivial solution (16.7) of (16.6) is stable. Function V_L is called *Lyapunov's function.* 'Along the motion' means that $\dot{x}_j$ is substituted from (16.6) in $\dot{V}_L$.

corresponding to (16.8), with respect to the *characteristic exponent* λ exist (the roots of characteristic equations are called *characteristic values* or *eigenvalues*).

Thus, a necessary and sufficient condition for stability, $A - C > 0$ has been obtained for the equilibrium position $\alpha = 0$. For the equilibrium position $\alpha = \pi$, we get a similar result.

An investigation of the stability of equilibrium positions with $\alpha = \pi/2$ and $\alpha = 3\pi/2$ can also be performed. A necessary and sufficient condition for the stability of these solutions is $A - C < 0$.

We next consider a diagram of SC motion in the *phase plane* $(\alpha, \dot{\alpha})$ for the case $A - C > 0$ (Fig. 16.1). One of the following three relationships is then valid, depending on the magnitude of the energy integral h_0,

$$\begin{aligned} &\frac{h_0}{B} > 3\omega_0^2 \frac{(A-C)}{B}; && (a) \\ &\frac{h_0}{B} = 3\omega_0^2 \frac{(A-C)}{B}; && (b) \\ &\frac{h_0}{B} < 3\omega_0^2 \frac{(A-C)}{B}. && (c) \end{aligned} \tag{16.9}$$

The points $\alpha = 0$ and $\alpha = \pi$ are stable equilibrium positions while

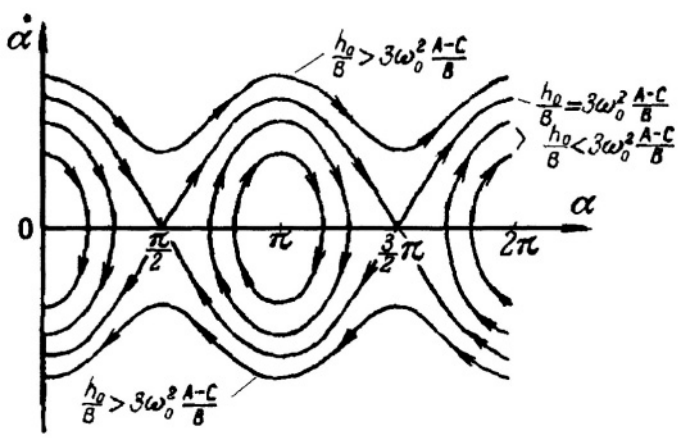

Figure 16.1. Phase-plane diagram of equation (16.3)

points $\alpha = \pi/2$ and $\alpha = 3\pi/2$ are unstable. Among two axes of inertia Ox_1 and Ox_3, the axis Ox_3 corresponding to the minimum moment of

inertia C is directed along the local vertical at the stable position. If h_0 satisfies inequality $(16.9a)$, the SC rotates. If h_0 is sufficiently small that $(16.9c)$ is satisfied, the SC librates relative to the local vertical. Motion along a *separatrix* satisfying equality $(16.9b)$ is asymptotic, that is if $t \to \infty$, then a point in the phase plane tends to approach an unstable equilibrium position but cannot reach that point. More typical motions are rotational and periodic.

If now we consider the case $A - C < 0$, stability is achieved at positions $\alpha = \pi/2$ and $\alpha = 3\pi/2$ when the axis of maximum moment of inertia of the SC coincides with the local vertical.

Next, substituting h_0 from (16.4) in the inequality $(16.9c)$, we obtain the *condition of libration boundedness*

$$|\dot{\alpha}_0| < \omega_0 \sqrt{3\frac{(A-C)}{B}}\,|\cos\alpha_0|. \tag{16.10}$$

Here the index 'zero' in α_0 and $\dot{\alpha}_0$ indicates that the angular position and velocity are pertinent to the moment t_0. Let us estimate allowed values of $\dot{\alpha}_0$ that provide libration motion. From the triangle inequality[†] $B + C > A$ it follows that $A - C < B$ and so $3(A - C)/B < 3$. Recalling that $|\cos\alpha_0| \leq 1$, we obtain from (16.10) the inequality $|\dot{\alpha}_0| \leq \omega_0\sqrt{3}$. For SC in LEO[§] the estimation $\omega_0 \approx 10^{-3}$ rad/s is valid. Thus, SC libration is possible at a very small initial angular velocity.

Let us return now to the integration of equation (16.3). On separating the variables in (16.5) and integrating, we obtain an *elliptic integral*

$$t - t_0 = \int_{\alpha_0}^{\alpha} \frac{d\alpha}{\sqrt{\dfrac{h_0}{B} - 3\omega_0^2\dfrac{(A-C)}{B}\sin^2\alpha}}. \tag{16.11}$$

[†]There are three *triangle inequalities*:

$$A + B \geq C, \quad A + C \geq B, \quad B + C \geq A,$$

which are valid for the principal moments of inertia of any body.

[§]Low Earth Orbits, typically with altitude of the order of 1 000 km.

Introduce the *modulus* k of the elliptic integral through the expression

$$k^2 = \frac{h_0}{3\omega_0^2(A-C)} = \frac{(\dot{\alpha}_0)^2}{3\frac{(A-C)}{B}\omega_0^2} + \sin^2\alpha_0.$$

The case $k^2 < 1$ (see (16.9c)) corresponds to libration. Also introduce a new variable φ through the relationship $\sin\alpha = k\sin\varphi$, from which we obtain $\cos\alpha d\alpha = kcos\varphi d\varphi$. Now, using the expression for the modulus k, we change in (16.11) variable α to variable φ through a set of transformations

$$\omega_0\sqrt{\frac{3(A-C)}{B}}(t-t_0) = \int_{\alpha_0}^{\alpha}\frac{d\alpha}{\sqrt{k^2-\sin^2\alpha}} =$$

$$= \int_{\varphi_0}^{\varphi}\frac{k\cos\varphi d\varphi}{\cos\alpha(\varphi)\sqrt{k^2-k^2\sin^2\varphi}} = \int_{\varphi_0}^{\varphi}\frac{d\varphi}{\sqrt{1-\sin^2\alpha(\varphi)}}.$$

Finally, we obtain

$$\omega_0\sqrt{\frac{3(A-C)}{B}}(t-t_0) = \int_{\varphi_0}^{\varphi}\frac{d\varphi}{\sqrt{1-k^2\sin^2\varphi}}. \qquad (16.12)$$

Next, recall the definition of an *elliptic integral of the first kind* in Legendre[¶] form

$$z = F(k,\varphi) = \int_0^{\varphi}\frac{d\varphi}{\sqrt{1-k^2\sin^2\varphi}}$$

where φ is termed the *amplitude* of z, written $\varphi = \mathrm{am}\, z$, which is induced by inversion of $F(k,\varphi)$. Definition of a sine of amplitude z induces the definition of elliptic function sn, which is called the *elliptic sine*

$$\sin\mathrm{am}\, z = \mathrm{sn}\, z.$$

[¶] Legendre, Andrien-Marie (1752-1833), French mathematician who made important contributions to calculus (particularly to function theory, differential equations and elliptic integrals), number theory and the theory of errors of observation. The Legendre polynomials he introduced are among the most important of the special functions.

Return now to equality (16.12). Transforming its right side and retaining there the elliptic integral, we obtain

$$\omega_0\sqrt{\frac{3(A-C)}{B}}(t-t_0)+\int_0^{\varphi_0}\frac{d\varphi}{\sqrt{1-k^2\sin^2\varphi}}=\int_0^{\varphi}\frac{d\varphi}{\sqrt{1-k^2\sin^2\varphi}}. \quad (16.13)$$

Using the definitions introduced for $F(k,\varphi)$ and sn, and inverting the integral on the right side of (16.13), we get the required solution of equation (16.3)

$$\sin\alpha = k\,\mathrm{sn}\left[\omega_0\sqrt{\frac{3(A-C)}{B}}(t-t_0)+F(k,\varphi_0)\right]. \quad (16.14)$$

It follows from (16.14) that $|\sin\alpha|\le k$ and, hence, $\max|\sin\alpha| = k$. This allows us to calculate the amplitude of SC libration.

The period of libration T is determined from the sequence of expressions

$$T=\frac{4K(k^2)}{\omega_0\sqrt{3\dfrac{A-C}{B}}}=\frac{T_0}{\sqrt{3\dfrac{A-C}{B}}}\cdot\frac{K(k^2)}{\pi/2} \quad (16.15)$$

where K is a *complete elliptic integral of the first kind,*

$$K(k^2)=\int_0^{\pi/2}\frac{d\varphi}{\sqrt{1-k^2\sin^2\varphi}},$$

$T_0 = 2\pi/\omega_0$ is the period of revolution of the SC around the Earth; and $4K$ is the period of $\mathrm{sn}\,z$ (i.e. $\mathrm{sn}(z+4K)=\mathrm{sn}(z)$).

If $k \ll 1$ then, for K, we write a Taylor–series expansion

$$K(k^2)=\frac{\pi}{2}\left[1+\frac{1}{4}k^2+\mathrm{O}(k^4)\right].$$

Substituting this expansion in (16.15) we get an approximate expression for the libration period

$$T=\frac{T_0}{\sqrt{3\dfrac{A-C}{B}}}\left[1+\frac{1}{4}k^2+\mathrm{O}(k^4)\right]. \quad (16.16)$$

If the terms $O(k^2)$ are negligible, the SC libration is similar to the motion of an harmonic oscillator with $T = T_0/[3(A-C)/B]^{1/2}$. The period of non-harmonic libration is always greater than this because $K(k^2) > \pi/2$ (see the last ratio on the right side in (16.15). The value $T_0/\sqrt{3}$ is the minimum period of libration since, due to the triangle inequality $B + C > A$, the maximum value of the ratio $3(A - C)/B$ is 3. If $A \to C$ then the period tends to infinity.

16.2 SC Equilibrium Positions

We now consider the existence of stationary solutions of (16.1) i.e. the equilibrium positions of a SC in circular orbit with respect to the ORS. Substituting $\overline{\omega}_1 = \overline{\omega}_2 = \overline{\omega}_3 = 0$, $\alpha = \alpha_0$, $\beta = \beta_0$, $\gamma = \gamma_0$ in (16.1) we obtain, if the conditions of pairwise inequality $A \neq B \neq C \neq A$ are valid for the SC moments of inertia, the following relationships

$$\begin{aligned} a_{22}a_{23} - 3a_{32}a_{33} &= 0, \\ a_{23}a_{21} - 3a_{33}a_{31} &= 0, \\ a_{21}a_{22} - 3a_{31}a_{32} &= 0. \end{aligned} \tag{16.17}$$

These comprise a closed system of trigonometric equations with respect to the angles α_0, β_0, γ_0 which has, in particular, the *trivial solution*

$$\alpha_0 = \beta_0 = \gamma_0 = 0. \tag{16.18}$$

It can be shown that 24 equilibrium positions of a three-axis SC exist. These correspond to all the possible configurations in which the axes of the BRS can coincide with the axes of the ORS.

The question here arises at to whether solutions $\alpha = \alpha_0$, $\beta = \beta_0$, $\gamma = \gamma_0$, which correspond to tilted positions of the axes of the BRS with respect to the ORS, exist. Introduce the expressions

$$\begin{aligned} \Delta_1 &= 3(a_{21}a_{32} - a_{22}a_{31}), \\ \Delta_2 &= 3(a_{23}a_{31} - a_{21}a_{33}), \\ \Delta_3 &= 3(a_{22}a_{33} - a_{23}a_{32}). \end{aligned} \tag{16.19}$$

The definition of the unit vector $\mathbf{E}_3$ of the ORS, given by $\mathbf{E}_3 = \mathbf{E}_1 \times \mathbf{E}_2$,

provides the relationships

$$(a_{21}a_{32} - a_{22}a_{31}) = a_{13},$$
$$(a_{23}a_{31} - a_{21}a_{33}) = a_{12},$$
$$(a_{22}a_{33} - a_{23}a_{32}) = a_{11}.$$

for the components of this vector in the BRS. Due to the trivial integral $a_{11}^2 + a_{12}^2 + a_{13}^2 = 1$ which is valid for a transformation matrix A, parameters Δ_1, Δ_2 and Δ_3 cannot all be equal to zero at the same time.

Let $\Delta_1 \neq 0$ then, in this case, the first and second equations of (16.17), which are uniform and linear with respect to a_{23}, a_{33} with non-zero determinant equal to Δ_1, have the unique trivial solution $a_{23} = a_{33} = 0$. On substituting $a_{23} = a_{33} = 0$ in (16.19), the equalities $\Delta_2 = \Delta_3 = 0$ are valid. The third equation of (16.17) has also to be satisfied, as well as the trivial integrals of motion $a_{21}^2 + a_{22}^2 + a_{23}^2 = 1$, $a_{31}^2 + a_{32}^2 + a_{33}^2 = 1$ and $a_{21}a_{31} + a_{22}a_{32} + a_{23}a_{33} = 0$ i.e. taking the solution $a_{23} = a_{33} = 0$ into account, the following four relationships

$$a_{21}a_{22} - 3a_{31}a_{32} = 0, \quad a_{21}a_{31} + a_{22}a_{32} = 0,$$
$$a_{21}^2 + a_{22}^2 = 1, \; a_{31}^2 + a_{32}^2 = 1 \tag{16.20}$$

should hold.

If we now introduce new variables φ_1 and φ_2 through $a_{21} = \sin\varphi_1$, $a_{22} = \cos\varphi_1$, $a_{31} = \sin\varphi_2$, $a_{32} = \cos\varphi_2$, then (16.20) leads to two trigonometric equations

$$\sin 2\varphi_1 - 3\sin 2\varphi_2 = 0, \; \cos(\varphi_1 - \varphi_2) = 0 \tag{16.21}$$

because the two last equations of (16.20) have become identities. Thus, the introduction of φ_1, φ_2, facilitated by the two trivial integrals $a_{21}^2 + a_{22}^2 = 1$, $a_{31}^2 + a_{32}^2 = 1$, results in reducing the number of equations to be solved.

Express φ_2 through φ_1 using the second equation of (16.21), and substitute it into the first equation. We then obtain $\sin 2\varphi_1 = 0$, which has solutions $\varphi_1 = \pi k/2$, $(k = 0, 1, \ldots)$. As a result, the solutions of (16.20) expressed through directional cosines have the form

$$a_{21} = a_{22}^2 - 1 = a_{31}^2 - 1 = a_{32} = 0$$

and

$$a_{21}^2 - 1 = a_{22} = a_{31} = a_{32}^2 - 1 = 0$$

which, together with the relationships $a_{23} = a_{33} = 0$, determine eight equilibrium positions of the BRS with respect to the ORS.

Similarly, other cases can be considered. If $\Delta_1 = 0$, then either Δ_2 or Δ_3 is not zero. Let $\Delta_1 = 0$, $\Delta_2 \neq 0$, then from the first and third equations of (16.17), we obtain $a_{22} = a_{32} = 0$ and, consequently, $\Delta_3 = 0$. As in the case where $\Delta_1 \neq 0$, the solutions

$$a_{21} = a_{22} = a_{23}^2 - 1 = a_{31}^2 - 1 = a_{32} = a_{33} = 0$$

and

$$a_{21}^2 - 1 = a_{22} = a_{23} = a_{31} = a_{32} = a_{33}^2 - 1 = 0$$

determine eight other equilibrium positions.

Finally, let $\Delta_1 = 0$ and $\Delta_3 \neq 0$. Then $a_{21} = a_{31} = 0$ and, consequently, $\Delta_2 = 0$ and a further eight solutions exist

$$a_{21} = a_{22} = a_{23}^2 - 1 = a_{31} = a_{32}^2 - 1 = a_{33} = 0$$

and

$$a_{21} = a_{22}^2 - 1 = a_{23} = a_{31} = a_{32} = a_{33}^2 - 1 = 0.$$

It is convenient to present all the solutions of (16.17) in tabular form (Table 16.1). Dashes in this Table mean that configurations

Table 16.1. Equilibrium positions of a SC

—	$a_{21} = a_{22} = 0$	$a_{21} = a_{23} = 0$	$a_{22} = a_{23} = 0$
$a_{31} = a_{32} = 0$	—	$a_{22}^2 = a_{33}^2 = 1$	$a_{21}^2 = a_{33}^2 = 1$
$a_{31} = a_{33} = 0$	$a_{23}^2 = a_{32}^2 = 1$	—	$a_{21}^2 = a_{32}^2 = 1$
$a_{32} = a_{33} = 0$	$a_{23}^2 = a_{31}^2 = 1$	$a_{22}^2 = a_{31}^2 = 1$	—

corresponding to values of the relevant directional cosines in the left column and upper line do not exist. The directional cosines a_{1k} are expressed using the directional cosines a_{2k}, a_{3k} $(k = 1, 2, 3)$. It is seen in this Table, that there are only 24 equilibrium positions of the BRS

with respect to the ORS[‖]. There are no other equilibrium positions that would correspond to tilted positions of the axes of the BRS with respect to the ORS.

16.3 Sufficient Conditions for Equilibrium Stability

Let us turn next to an investigation of the stability of the equilibrium positions obtained above. Write a Jacobi integral of energy [5]

$$\frac{1}{2}(Aw_1^2 + Bw_2^2 + Cw_3^2) + \frac{3}{2}\omega_0^2(Aa_{31}^2 + Ba_{32}^2 + Ca_{33}^2) \\ -\omega_0(Aw_1a_{21} + Bw_2a_{22} + Cw_3a_{23}) = h_0 \qquad (16.22)$$

for equations (16.1). Express this integral using the components of relative angular velocity $(\overline{\omega}_1, \overline{\omega}_2, \overline{\omega}_3)_x$ with respect to the ORS

$$\frac{1}{2}(A\overline{\omega}_1^2 + B\overline{\omega}_2^2 + C\overline{\omega}_3^2) + \frac{3}{2}\omega_0^2(Aa_{31}^2 + Ba_{32}^2 + Ca_{33}^2) \\ -\frac{1}{2}\omega_0^2(Aa_{21}^2 + Ba_{22}^2 + Ca_{23}^2) = h_0. \qquad (16.23)$$

Here on the left side, the first term represents the kinetic energy of relative angular motion; the second term the potential energy of the gravitational field and the third term the potential energy of centrifugal inertial forces.

Use integral (16.23) to determine the stability conditions of equilibrium positions of the SC. For specificity, consider the representative equilibrium positions

$$\overline{\omega}_1 = \overline{\omega}_2 = \overline{\omega}_3 = a_{21} = a_{23} = a_{31} = a_{32} = 0. \qquad (16.24)$$

Table (16.1) shows that there are *four* equilibrium positions described by these relationships since $a_{22}^2 = a_{33}^2 = 1$. Exclude a_{33}^2 and a_{22}^2 from

[‖]This approach to proving the existence of 24 equilibrium SC positions was used by V.A.Sarychev and V.V.Sazonov in *Evaluation of the Boundaries of the SC Librational Motion.* Preprint of the Institute of Applied Mathematics USSR Academy of Sciences, 1974, N 130. Another approach by P.W.Likins and R.E.Roberson is contained in *Uniqueness of Equilibrium Attitudes for Earth-Pointing Satellites.* Journal of the Astronautical Sciences, 1966, Vol.13, Issue 2, pp.87-88.

(16.23) by substituting the trivial integrals

$$a_{33}^2 = 1 - a_{31}^2 - a_{32}^2, \quad a_{22}^2 = 1 - a_{21}^2 - a_{23}^2.$$

Then integral (16.23) can be written as

$$\frac{1}{2}(A\bar{\omega}_1^2 + B\bar{\omega}_2^2 + C\bar{\omega}_3^2) + \frac{3}{2}\omega_0^2[(A-C)a_{31}^2 + (B-C)a_{32}^2] + \frac{1}{2}\omega_0^2[(B-A)a_{21}^2 + (B-C)a_{23}^2] = \bar{h}_0 \qquad (16.25)$$

where $\bar{h}_0 = h_0 - \frac{1}{2}\omega_0^2(3C - B)$ is a new constant.

Denote first integral (16.25) by V_L and confirm with reference to Section 16.1 that (16.25) can be used as a Lyapunov function . The function V_L is equal to zero only at positions (16.24). It is noted that this is a consequence of the substitutions for a_{22} and a_{33} made above which have non-zero values at equilibrium positions. If the inequalities

$$B > A > C \qquad (16.26)$$

are valid, then the function V_L is strictly positive outside the equilibrium positions (16.24). Thus, function V_L is *positive-definite* by virtue of (16.26). The derivative $dV_L/dt = 0$ is calculated along the motion described by (16.1) due to the relationship $V_L = \bar{h}_0$. Hence, by Lyapunov's theorem for stability, the inequalities (16.26) provide *sufficient* conditions for stability at the relative equilibrium positions (16.18).

Thus, it is sufficient for stability at the relative equilibrium positions of a SC in circular orbit that (a) the axis of SC minimum moment of inertia be along the radius-vector connecting the Earth's center with the SC center of mass; (b) the axis of SC maximum moment of inertia be along the orbit normal; (c) the axis of middle moment of inertia be along the orbit tangent. There are only *four* equilibrium positions corresponding to these configurations. They are referred to, speaking figuratively, as motions with 'face forward' ('head up' or 'head down') and 'face back' ('head up' or 'head down'). Although, there are 24 equilibrium positions listed in Table (16.1), the above consideration restricts the number of stable configurations to this four (corresponding to minima of the total energy of motion of the SC with respect to the orbital reference system).

16.4 Necessary Conditions for Equilibrium Stability

Consider spatial librations of a SC in a circular orbit close to the equilibrium position (16.18). Assume that the librations are sufficiently small that $\sin x \approx x$, $\cos x \approx 1$, where x corresponds to the angles α, β, γ. Substitute these approximations in the elements of the transformation matrix A. Then the expressions thus obtained for the directional cosines are substituted into (16.1). This results in the following equations, linearized in the vicinity of solution (16.18)

$$\begin{gathered} A\ddot{\gamma} + (A - B + C)\omega_0\dot{\beta} + (B - C)\omega_0^2\gamma = -3(B - C)\omega_0^2\gamma, \\ B\ddot{\alpha} = -3(A - C)\omega_0^2\alpha, \\ C\ddot{\beta} - (A - B + C)\omega_0\dot{\gamma} + (B - A)\omega_0^2\beta = 0. \end{gathered} \tag{16.27}$$

Here all terms are especially retained on the same side of the equations where corresponding nonlinear terms were placed before linearization. This allows us to deduce the source of different components of the restoring torque dependent on angular displacement.

In the first equation, the restoring torque associated with roll motion is developed by centrifugal inertial forces (third term on the left side) and gravity-gradient forces (term on the right side). In the second equation, which describes pitch motion, the restoring torque is developed by gravity-gradient forces only (term on the right side). In the third equation, the restoring torque associated with yaw motion, is developed by centrifugal inertial forces only (third term on the left side) and there is no term corresponding to the gravity-gradient torque. This means that the yaw component of the restoring torque is developed due only to the rotation of the ORS. Because the first and third equations are mutually coupled, we can speak about the association between components of the restoring torque and individual motions.

On grouping now the similar terms in (16.27), we obtain

$$\begin{gathered} B\ddot{\alpha} + 3(A - C)\omega_0^2\alpha = 0, \\ C\ddot{\beta} - (A - B + C)\omega_0\dot{\gamma} + (B - A)\omega_0^2\beta = 0, \\ A\ddot{\gamma} + (A - B + C)\omega_0\dot{\beta} + 4(B - C)\omega_0^2\gamma = 0. \end{gathered} \tag{16.28}$$

Here we placed first an equation describing pitch motion in the orbital plane which is independent of roll (γ) and yaw (β). If $A - B + C = 0$ (this case corresponds to an infinitesimally thin plate contained in the Ox_1x_3 plane of the BRS), then there is no relation between roll and yaw librations, and all three motions are decoupled.

Next introduce the *unitless inertial parameters* $\theta_A = A/B$ and $\theta_C = C/B$, use the argument of latitude u as a unitless independent variable instead of time t and transform (16.28) to

$$\begin{aligned} \alpha'' + 3(\theta_A - \theta_C)\alpha = 0, \\ \theta_C\beta'' - (\theta_A + \theta_C - 1)\gamma' + (1 - \theta_A)\beta = 0, \\ \theta_A\gamma'' + (\theta_A + \theta_C - 1)\beta' + 4(1 - \theta_C)\gamma = 0 \end{aligned} \tag{16.29}$$

where prime $(')$ denotes a derivative with respect to u.

The *secular equation* corresponding to (16.29) has the form

$$\begin{aligned} \{\rho^2 - 3(\theta_A - \theta_C)\} \\ \times \{\theta_A\theta_C\rho^4 - [\theta_A\theta_C + 3\theta_C(1 - \theta_C) + (1 - \theta_A)(1 - \theta_C)]\rho^2 \\ +4(1 - \theta_A)(1 - \theta_C)\} = 0, \end{aligned} \tag{16.30}$$

ρ is an *eigenvalue frequency* when we seek a solution of (16.29) in the form of harmonic oscillations $\cos(\rho u + \varphi_0)$, where φ_0 is the initial phase of motion.

Let ρ_1^2, ρ_2^2 be roots of the biquadratic equation with respect to ρ

$$\begin{aligned} \theta_A\theta_C\rho^4 - [\theta_A\theta_C + 3\theta_C(1 - \theta_C) + (1 - \theta_A)(1 - \theta_C)]\rho^2 \\ +4(1 - \theta_A)(1 - \theta_C) = 0 \end{aligned} \tag{16.31}$$

and ρ_3 be a root of the quadratic equation

$$\rho_3^2 = 3(\theta_A - \theta_C). \tag{16.32}$$

Then the general solution of (16.29) has the form

$$\begin{aligned} \alpha = A_3 \sin(\rho_3 u + \varphi_3), \\ \beta = A_1 \sin(\rho_1 u + \varphi_1) + A_2 \sin(\rho_2 u + \varphi_2) \\ \gamma = k_1 A_1 \cos(\rho_1 u + \varphi_1) + k_2 A_2 \cos(\rho_2 u + \varphi_2) \end{aligned} \tag{16.33}$$

where

$$k_j = \frac{\theta_C\rho_j^2 - (1 - \theta_A)}{\rho_j(\theta_A + \theta_C - 1)} = \frac{\rho_j(\theta_A + \theta_C - 1)}{\theta_A\rho_j^2 - 4(1 - \theta_C)}, \quad (j = 1, 2) \tag{16.34}$$

and A_1, A_2, A_3, φ_1, φ_2, φ_3 are constants, expressed via initial values of the angles α, β, γ and their derivatives.

Roots ρ_1, ρ_2, ρ_3 are real if the following conditions

$$\begin{gathered}\theta_A - \theta_C > 0,\\ (1-\theta_A)(1-\theta_C) > 0,\\ \theta_A\theta_C + 3\theta_C(1-\theta_C) + (1-\theta_A)(1-\theta_C) > 0,\\ [\theta_A\theta_C + 3\theta_C(1-\theta_C) + (1-\theta_A)(1-\theta_C)]^2\\ -16\theta_A\theta_C(1-\theta_A)(1-\theta_C) > 0\end{gathered} \tag{16.35}$$

are met so that:

— $\rho_3^2 > 0$;

— the signs of ρ_1^2 and ρ_2^2 should be the same;

— the sum $\rho_1^2 + \rho_2^2$ should be positive;

— the discriminant of the biquadratic equation should be positive.

If the second inequality in (16.35) is true, then the third inequality is true automatically, since the inertial parameters of the SC satisfy the triangle inequalities and the sufficient conditions for stability (16.26).

The inequalities (16.35) are necessary conditions for the stability of the trivial solution of the linearized equations (16.29). Sufficient conditions (16.26) are expressed via the unitless inertial parameters θ_A and θ_C in the form

$$\theta_A - \theta_C > 0, \qquad 1 - \theta_A > 0. \tag{16.36}$$

Domains in the (θ_A, θ_C) plane (Fig.16.2) determined by the necessary (16.35) and sufficient (16.36) conditions for the stability of the trivial solution (16.18), are indicated through vertical (domains NPL and PQRS) and horizontal (domain NPL) hatching respectively. In domain NPL, the simultaneous fulfilling of necessary and sufficient conditions for stability results in this regime being cross hatched. The infinite semi-strip corresponding to all physically realized rigid bodies is constrained by the triangle inequalities

$$\theta_A + \theta_C - 1 \geq 0, \qquad -\theta_A + \theta_C + 1 \geq 0, \qquad \theta_A - \theta_C + 1 \geq 0.$$

The necessary and sufficient conditions for stability shown in Fig. 16.2 by the domain NPL, are provided by the inequalities

$$B > A > C$$

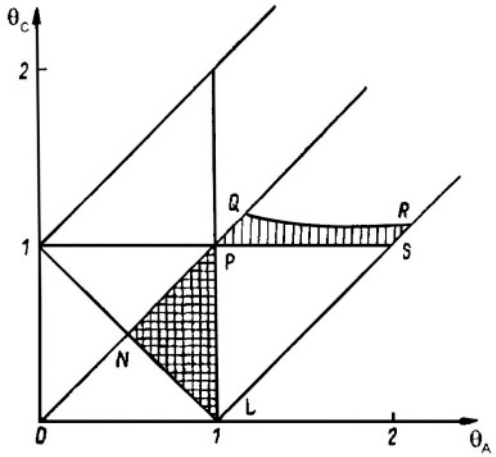

Figure 16.2. Domains of necessary (NPL and PQRS) and sufficient (NPL) conditions for the stability of solution (16.18)

These conditions were derived by Beletsky** in 1959.

** V.V.Beletsky. Libration of a Satellite. Transactions: *Artificial Satellites of the Earth,* Moscow, USSR Academy of Sciences Publ., 1959, Issue 3, pp.13–31.

Chapter 17

SC Motion in an Elliptical Orbit

The equations (15.21) and (15.22) have a special solution relevant to the motion of a SC in an orbital plane (planar motion). Now we will consider those solutions pertaining to motion in elliptical orbits.

17.1 Equation of Planar SC Motion

For an elliptical orbit, planar motion satisfies the following equations

$$B\dot{\omega}_2 = 3\frac{\mu_g}{R^3}(A-C)a_{31}a_{33}, \quad \dot{\alpha} = \omega_2 - \omega_{orb}.$$

Substitute the expressions for the directional cosines a_{31} and a_{33} in these equations and transform them to a single,second order, equation

$$\ddot{\alpha} + \dot{\omega}_{orb} + 3\frac{\mu_g}{R^3}(A-C)\sin\alpha\cos\alpha = 0.$$

Recalling that $\omega_{orb} = \omega_0(1+e\cos\upsilon)^2$ and $\dot{\omega} = \dot{\upsilon}\omega'$, where a prime denotes derivation with respect to the true anomaly υ, then we obtain for υ

$$\dot{\alpha} = \alpha'\dot{\upsilon}, \quad \ddot{\alpha} = [\alpha''\dot{\upsilon} + \alpha'(\dot{\upsilon})']\dot{\upsilon}.$$

Substitute in the above the expression $\dot{\upsilon} = \omega_{orb}$ to obtain the equation* of SC motion

$$\alpha''(1 + e\cos\upsilon) - 2e\alpha'\sin\upsilon + \mu\sin\alpha\cos\alpha = 2e\sin\upsilon. \qquad (17.1)$$

Here $\mu = 3(A - C)/B$ is a *SC inertial parameter* and μ and e satisfy the inequalities $|\mu| \leq 3$ and $0 \leq e < 1$ respectively. Sometimes, it is convenient to use the *mean anomaly M* (see (1.28), Section 1.6) rather than the true anomaly. Using the mean anomaly we can write equation (17.1) in the form

$$\frac{d^2\alpha}{dM^2} + \mu\left(\frac{a}{R}\right)^3\sin\alpha\cos\alpha = 2e\left(\frac{a}{R}\right)^3\sin\upsilon. \qquad (17.2)$$

Introduce a new variable $x = \alpha + \upsilon$. Then from (17.2) we obtain an equation

$$\frac{d^2x}{dM^2} + \frac{\mu}{2}\frac{(1 + e\cos\upsilon)^3}{1 - e^2}\sin 2(x - \upsilon) = 0 \qquad (17.3)$$

which is appropriate, when $|\mu| \ll 1$, for the application of asymptotic methods of solution.

Extract now the particular solution of equation (17.1) that exists for the specific relation between the SC and orbit parameters

$$\mu = 6e.$$

For this, equation (17.1) has the form

$$\alpha''(1 + e\cos\upsilon) - 2e(1 + \alpha')\sin\upsilon + 3e\sin 2\alpha = 0 \qquad (17.4)$$

with the solution

$$\alpha = \upsilon/2$$

which carries the following mechanical meaning. At perigee where $\upsilon = 0$, $\alpha = 0$ and the SC 'faces' towards the Earth. At apogee where $\upsilon = \pi$, $\alpha = \pi/2$, it faces along the tangent to the orbit. On passing the next perigee $(\upsilon = 2\pi)$ $\alpha = \pi$ the SC faces away from the Earth. On the other hand, the SC spins continuously and occupies the same angular positions with respect to the Earth every two revolutions.

*V.V.Beletsky. Motion of an Artificial Satellite About its Center of Mass, Transactions: *Artificial Satellites of the Earth.* Moscow, USSR Academy of Sciences Publ., 1958, Issue 1, pp.25–43.

17.2 Linear Librations

On linearizing equation (17.1) in the vicinity of the solution $\alpha = 0$ which exists for a circular orbit and introducing a new variable z using

$$\alpha = \frac{z}{1 + e\cos v}$$

we obtain a non-uniform Hill type equation, with periodic coefficients and a right side that is periodic with respect to the true anomaly,

$$z'' + \frac{\mu + e\cos v}{1 + e\cos v} z = 2e\sin v. \tag{17.5}$$

This equation cannot be integrated in closed form except for the simple resonant case $\mu = 1$, and we will thus use an approximation method to investigate it.

Let the eccentricity e be much less than unity $(e \ll 1)$ then, to within the approximation $\mathrm{O}(e^2)$, the expression for the coefficient of z in (17.5) can be written in the form

$$\frac{\mu + e\cos v}{1 + e\cos v} = \mu - (\mu - 1)e\cos v$$

and this equation may then be transformed to a non-uniform *Mathieu's equation*

$$z'' + \mu z = e[2\sin v + z(\mu - 1)\cos v]. \tag{17.6}$$

The solution of equation (17.6) for a null-order approximation with respect to e has the form

$$z = a\cos\psi, \tag{17.7}$$

where a is the amplitude and $\psi = \sqrt{\mu}v + \psi_0$ a libration phase with an initial value ψ_0. Let us now obtain a solution of equation (17.6) in the following form

$$\begin{gathered} z = a\cos\psi + eu_1(\psi, v, a), \\ z' = -a\sqrt{\mu}\sin\psi + e\left(\frac{\partial u_1}{\partial\psi}\psi' + \frac{\partial u_1}{\partial v} + \frac{\partial u_1}{\partial a}a'\right), \end{gathered} \tag{17.8}$$

moreover, a and ψ have to satisfy the equations

$$a' = eA_1(a), \qquad \psi' = \sqrt{\mu} + eB_1(a), \tag{17.9}$$

$u_1(\psi, \upsilon, a)$ is a function to be obtained. (We follow here the *van der Pol's method*[†] (see Appendix A)). This form of representation of the sought solution $z(\upsilon)$ is called the 'improved first approximation'.

Substituting (17.8) into (17.6), using the conditions (17.9), and keeping only terms of first order in e, we get after transformation

$$\begin{aligned} &-2e\sqrt{\mu}(A_1 \sin\psi + B_1 a \cos\psi) \\ &+e\left(\frac{\partial^2 u_1}{\partial\psi^2}\mu + 2\frac{\partial^2 u_1}{\partial\psi\partial\upsilon}\sqrt{\mu} + \frac{\partial^2 u_1}{\partial\upsilon^2}\right) + \mu e u_1 = \\ &= e\left\{2\sin\upsilon + \frac{a(\mu-1)}{2}\left[\cos(\upsilon+\psi) + \cos(\upsilon-\psi)\right]\right\}. \end{aligned} \tag{17.10}$$

Cancel e on both sides. On the right side of (17.10) the harmonics $\cos(\upsilon+\psi)$ and $\cos(\upsilon-\psi)$ are present, but the harmonics $\cos\psi$ and $\sin\psi$ are absent. In order that (17.10) be valid, the functions A_1 and B_1 have to satisfy the relations

$$A_1(a) = B_1(a) = 0. \tag{17.11}$$

Now determine u_1 using (17.11) so that the left and right sides in (17.10) are equal to each other. Substitute u_1 in (17.10) in the form

$$u_1 = \gamma_u \sin\upsilon + \alpha_u \cos(\upsilon-\psi) + \beta_u \cos(\upsilon+\psi).$$

The factors of each harmonic are added and the sums obtained equated to zero. Thus, we get expressions for the coefficients

$$\gamma_u = \frac{2}{\mu-1},\ \alpha_u = \frac{a(\mu-1)}{2(2\sqrt{\mu}-1)},\ \beta_u = -\frac{a(\mu-1)}{2(2\sqrt{\mu}+1)}.$$

Finally, a first approximation to the solution of (17.6) has the form

$$\begin{aligned} z = a\cos\psi + \frac{2e}{\mu-1}\sin\upsilon + \\ +e\frac{a(\mu-1)}{2}\left(\frac{1}{2\sqrt{\mu}-1}\cos(\upsilon-\psi) - \frac{1}{2\sqrt{\mu}+1}\cos(\upsilon+\psi)\right). \end{aligned} \tag{17.12}$$

Through (17.11) the equalities $a = \text{const}$ and $\psi_0 = const$ are valid. Apart from a resonance at $\mu = 1$ brought about by the equality between the natural frequency of the SC and the frequency of its orbital

[†]B.van der Pol. A Theory of the Amplitude of Free and Forced Triode Vibration. *Radio Review,* 1920, Vol.1, p.701.

motion (see the forced term $(2\sin v)/(\mu-1)$), a *parametric resonance* at $\mu = 1/4$ caused by the periodic variation of the natural frequency (see (17.5)) appears. The third term which contains a factor *ea* in (17.12), is provided, to a first approximation, by a solution of Mathieu's equation.

17.2.1 Stability of the Solution of Mathieu's Equation

Write the Mathieu equation[‡] in its general form

$$\ddot{y} + (a + b\cos 2t)y = 0. \tag{17.13}$$

This equation describes the behaviour of a mechanical system affected by parametric periodic excitation. The character of its solution is similar to that for libration, depending on the values of the parameters a and b. For one combination of values of these parameters the libration remains bounded. For another they increase infinitely. The phenomenon of librations with infinitely increased amplitude is named *parametric resonance.* In the case of bounded librations, the trivial solution of equation (17.13) is stable. In the unbounded case, there is no stability. The curves along which periodic solutions of (17.13) exist, separate domains of stability and instability in the plane (*a, b*). An example of a partial solution of Mathieu's equation was presented in the previous Section.

Now we will seek a 2π-periodic solution in the following form

$$y = a_1 \sin t + b_1 \cos t + a_3 \sin 3t + b_3 \cos 3t + \ldots. \tag{17.14}$$

Considering only the two first terms, substitute them in (17.13). Transforming the trigonometric function products into sums of these functions we get

$$\left(-a_1 + aa_1 - \frac{ba_1}{2}\right)\sin t + \left(-b_1 + ab_1 + \frac{bb_1}{2}\right)\cos t + \ldots = 0. \tag{17.15}$$

The conditions for equating to zero on the left side of this equation the two retained terms

$$a_1\left(-1 + a - \frac{b}{2}\right) = 0, \qquad b_1\left(-1 + a + \frac{b}{2}\right) = 0$$

[‡]E.Mathieu *Cours de physique mathematique, Paris, 1873.*

allow us to write the conditions for the existence of a non-trivial solution for a_1 and b_1 in the form

$$a \equiv a_l^1 = 1 - \frac{b}{2}, \qquad a \equiv a_r^1 = 1 + \frac{b}{2}.$$

These are the left and right boundaries of the non-stability domain of the trivial solution of Mathieu's equation, both of which start from the point $a = 1$, $b = 0$.

Similarly, the boundaries starting from points $a = 0$, $a = 4$, $a = 9, \ldots$ can be constructed. These have the forms

$$\begin{aligned}
a_r^0 &= -\frac{b^2}{8} + \ldots, \\
a_l^1 &= 1 - \frac{b}{2} + \ldots, & a_r^1 &= 1 + \frac{b}{2} + \ldots, \\
a_l^4 &= 4 - \frac{b^2}{24} + \ldots, & a_r^4 &= 4 + \frac{5b^2}{24} + \ldots, \\
a_l^9 &= 9 + \frac{b^2}{32} - \frac{b^3}{512} + \ldots, & a_r^9 &= 9 + \frac{b^2}{32} + \frac{b^3}{512} + \ldots, \\
&\ldots.
\end{aligned}$$

The domain boundaries are presented in a diagram of stability, where domains of stability are indicated by hatching (Fig.17.1).

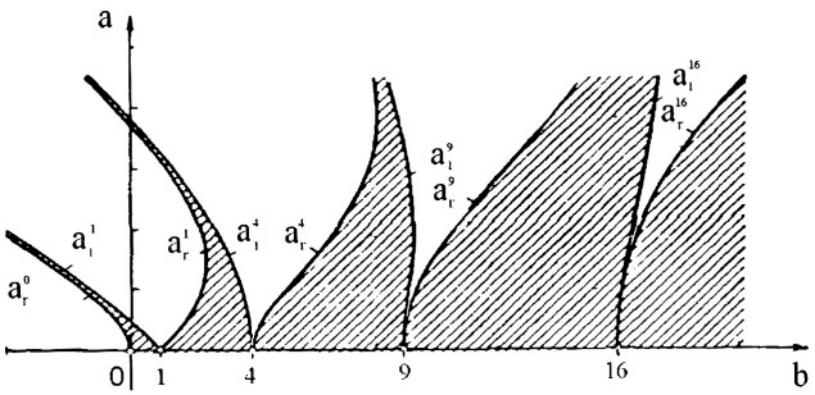

Figure 17.1. Diagram of the stability of the solution of Mathieu's equation (Ince-Strutt's diagram)

If the conditions for parametric resonance are realized, then the amplitude of the librations increases even though damping is present in the system. This behaviour of parametrically excited linear systems, differs from what happens in the case of forced librations. In

the former case, both the energy loss and the input of energy are proportional to the square of the amplitude of libration. In the latter case, the energy input is directly proportional to the amplitude while the energy loss is proportional to the square of the amplitude. This is the reason for the boundedness of forced librations in the presence of damping.

17.3 Non-linear Librations

Let us seek a solution of (17.1) for small, though nonlinear, forced librations, in the form

$$\alpha = a\cos\psi, \qquad \psi = \upsilon + \psi_0, \tag{17.16}$$

where, as in (17.7), ψ_0 and a are constants. First represent $\sin(a\cos\psi)$ in the form of a Fourier-series

$$\sin(a\cos\psi) = 2\sum_{k=0}^{\infty}(-1)^k I_{2k+1}(a)\cos(2k+1)\psi.$$

Here $I_{2k+1}(a)$ is *Bessel's*[§] *function of the first kind* of order $2k+1$. We will consider only the first term in this series

$$\sin(a\cos\psi) \approx 2I_1(a)\cos\psi. \tag{17.17}$$

Substitute (17.16) in (17.1), using (17.17). Summing the factors of the same harmonics of υ and equating the sums for the constant terms, as well as for $\sin\upsilon$ and $\cos\upsilon$, to zero we get three equations

$$\begin{aligned} \tfrac{1}{2}e\cos\psi_0 = 0, \qquad & [2\mu I_1(a) - a]\cos\psi_0 = 0, \\ & [2\mu I_1(a) - a]\sin\psi_0 + 4e = 0. \end{aligned} \tag{17.18}$$

It follows from the first that $\psi_0 = \pm\pi/2$, and from the third, taking into account (17.16),

$$\mu = \frac{a \pm 4e}{2I_1(a)}, \qquad \alpha = \mp a\sin\upsilon. \tag{17.19}$$

[§]Bessel, Friedrich Wilhelm (1784–1846), German astronomer and mathematician, Member of the St.Petersburg and of the Berlin Academy of Sciences. He made a significant contribution to the theory of errors as applied to astronomical observations. Functions he developed to describe specific physical and technical problems, which are now called in his name, are among those most commonly used in physics and engineering.

An example of this solution when $e = 0.01$ is shown in Fig.17.2 [4] where the left (right) branches of the curve correspond, respectively, to minus (plus) in (17.19). For the values of parameter μ lying on the right side relative to the vertical tangent line there are three solutions for each value of μ, in contrast to one solution on the left side relative to this vertical line. For small a, the relationship $I_1(a) \approx a/2$ is

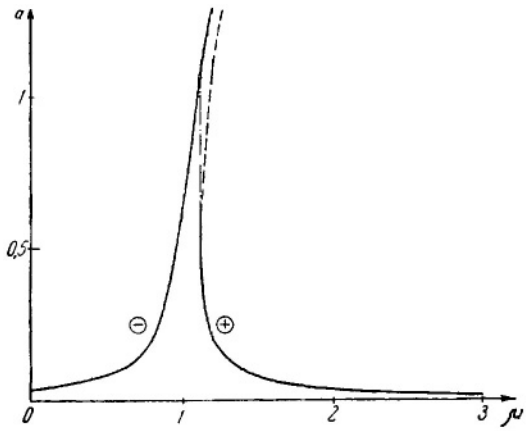

Figure 17.2. Libration amplitude a as a function of μ when $e = 0.01$ (non-resonant case)

valid and the solution obtained coincides with that part of (17.12) corresponding to forced librations.

Consider similarly parametric resonant librations at $\mu = 1/4$ (see (17.12)), making the distinction only that the solution should be in the form

$$\alpha = a\cos(\frac{v}{2} + \psi_0).$$

Then, on executing transformations, we obtain two solutions

$$\sin\psi_0 = 0, \qquad \mu = a\frac{2-3e}{16I_1(a)} \tag{17.20}$$

and

$$\cos\psi_0 = 0, \qquad \mu = a\frac{2+3e}{16I_1(a)}.$$

Hence, it follows that the amplitude is uniquely determined within the domain

$$\frac{1}{4} - \frac{3}{8}e \le \mu \le \frac{1}{4} + \frac{3}{8}e. \tag{17.21}$$

If μ belongs to (17.21), then parametric resonant librations appear. The second solution is not realized. The dependence of libration amplitude a on the inertial parameter μ is shown in Fig.17.3 [4]. Amplitudes

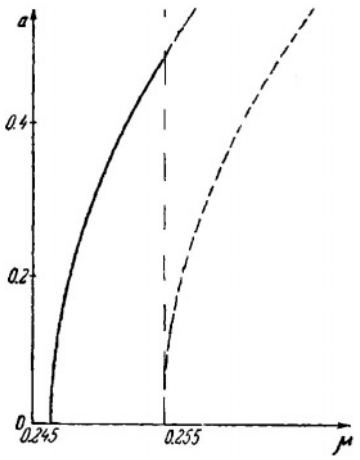

Figure 17.3. The dependence of amplitude a on the inertial parameter μ at $e = 0.01$ (resonant case)

of the solution (17.20) existing in the domain (17.21) are indicated by the solid line. Amplitudes corresponding to non-realized solutions are indicated by dashed lines.

In the next Section we will consider periodic motions in elliptical orbits with arbitrary eccentricity, which support the results obtained above for small eccentricity.

17.4 Periodic Motion of a SC

17.4.1 Solutions for Slightly Elliptical Orbits

Take again equation (17.1) and transform it, for $e = 0$, into the equation of a simple pendulum

$$(2\alpha)'' + \mu \sin 2\alpha = 0, \tag{17.22}$$

which describes motion of a SC in a circular orbit. The integral of energy for equation (17.22) has the form $(\alpha')^2 + \mu \sin^2 \alpha = h_0$. We present here the solution of this equation for two cases, librations and rotations,

libration solution ($k < 1$)**:**

$$\alpha = \arcsin[k \operatorname{sn} \sqrt{\mu}(\upsilon+\upsilon_0)], \quad k = \sqrt{h_0}, \quad T_0 = \frac{4K(k)}{\sqrt{\mu}}; \quad (17.23)$$

rotation solution ($k > 1$)**:**

$$\alpha = \pm \operatorname{am} \frac{\sqrt{\mu}}{k}(\upsilon + \upsilon_0), \quad k = \frac{1}{\sqrt{h_0}}, \quad T_0' = \frac{2kK(k)}{\sqrt{\mu}}, \quad (17.24)$$

where T_0 is the period of SC libration, T_0' is the period of SC rotation and υ_0 is the initial phase. A detailed account of the libration solution (17.23) is contained in Section 16.1.

Since the parameters e and/or μ are small, it is possible to apply asymptotic methods of solution. Adopting, in particular, *Poincare's method* (see Appendix C), it can be shown that a unique, 2π-periodic, solution $\alpha(\upsilon, e)$ of equation (17.1), which is analytic with respect to e, and satisfies the equality $\alpha(\upsilon, 0) = 0$, exists for $e \ll 1$ and $\mu \neq 1$. Equation (17.22) has a stationary solution

$$\alpha_0 = \frac{\pi}{2}n, \qquad n = 0, 1, 2, 3. \quad (17.25)$$

For the non-resonant case a 2π-periodic solution a of (17.1) generated by the stationary solution $\alpha_0 = 0$ (in this case, $\alpha_0 = 0$ is called a *generating solution*), can be obtained as a series in powers of e, for $e \ll 1$

$$\alpha = e\alpha_1 + e^2\alpha_2 + \ldots, \quad (17.26)$$

where factors $\alpha_1, \alpha_2, , \ldots$ have to be determined. To do this, the solution sought (for example (17.26)), is substituted for α in (17.1), and the factors of terms of the same order in e are equated. The resulting set of equations is then solved, beginning from the equation for α_1 and proceeding consecutively. For terms up to order three in e, the solution (17.26) has the form[¶]

$$\alpha = \left[\frac{2e}{\mu - 1} + \frac{4\mu e^3}{(\mu - 1)^4} + \ldots\right] \sin \upsilon$$

[¶]V.A.Zlatoustov, D.E.Okhotsimsky, V.A.Sarychev and A.P.Torzhevsky. Investigation of Satellite Libration in the Plane of an Elliptical Orbit. *Cosmic Research,* 1964, Vol.2, Issue 5, pp.657-666.

$$+\left[\frac{3e^2}{(\mu-1)(\mu-4)}+\ldots\right]\sin 2v \qquad (17.27)$$
$$+\left[\frac{4(10\mu^2-22\mu+9)3e^3}{3(\mu-1)^3(\mu-4)(\mu-9)}+\ldots\right]\sin 3v+\ldots .$$

A general expression for α can be written as

$$\alpha=\sum_{r=1}^{\infty}\sum_{k=0}^{\infty}a_{r+2k}^{(r)}e^{r+2k}\sin rv. \qquad (17.28)$$

In the resonant case, a 2π-periodic solution generated by the stationary solution $\alpha_0=0$, can be obtained as a series in powers of $e^{1/3}$

$$\alpha=-2^{2/3}e^{1/3}\sin v-e\left(\frac{2}{9}\sin v+\frac{1}{12}\sin 3v\right)+2^{-1/3}e^{4/3}\sin 2v$$
$$-2^{4/3}e^{5/3}\left(\frac{119}{158760}\sin v-\frac{1}{144}\sin 3v+\frac{1}{720}\sin 5v\right)+\ldots \qquad (17.29)$$

The solution (17.26) is odd with respect to v and 2π periodic. The condition for oddness and periodicity T are represented by

$$\alpha(-v)=-\alpha(v)$$

and

$$\alpha(v+T)=\alpha(v).$$

Substituting $v=0$ in the first equality, we get $\alpha(0)=0$. Substituting $v=-T/2$ in the second equality, and taking into account the condition of oddness, we get $\alpha(T/2)=0$. Assuming $T=2\pi$, we obtain boundary conditions

$$\alpha(0)=\alpha(\pi)=0 \qquad (17.30)$$

which define the solution $\alpha(v,e)$ of equation (17.1). The first terms of the solution expansion of $\alpha_0(v,e)$ in the series in e is determined by (17.27).

The generating, nonstationary, periodic solutions contained among the librational and rotational solutions (17.23), (17.24) of the pendulum remain periodic at non-zero values of the small parameter e. From the 2π-periodicity of the terms of equation (17.1), it follows that the generating solutions should obey the condition

$$T_0=\frac{4K(k)}{\sqrt{\mu}}=\frac{2\pi m}{n} \qquad (17.31)$$

where m and n are relatively prime natural numbers. This condition determines solutions which describe the motion of a SC when it performs n librations with respect to the local vertical during m orbits. Equation (17.31) has real roots only if $\mu \geq n^2/m^2$. For $e \ll 1$, at least $2m$ periodic solutions of equation (17.1) exist which correspond to the following values of the phase υ_0,

$$\upsilon_0^{(r)} = \frac{\pi r}{n}, \qquad (r = 0, 1, 2, \ldots, 2m-1). \tag{17.32}$$

Denote by $\alpha_{n/m}^{(r)}(\upsilon, e)$ these periodic solutions. They can then be determined in the form of a series in e or, numerically, by solving the one-parameter boundary value problem for (17.1). Sometimes the combination of an equation and its boundary conditions is called a *boundary value problem.*

17.4.2 Numerical Investigation

The results obtained in the previous Section 17.4.1 concern $e \ll 1$. We will now construct periodic solutions for arbitrary values of the eccentricity and, simultaneously, analyze their stability.

The construction of 2π-periodic solutions of (17.1) is performed by numerically solving the boundary value problem.

The stability of 2π-periodic solutions of (17.1) is analyzed through using the, corresponding, linerar equation. Let $\alpha = \alpha^* + y$ be a small variation of the 2π-periodic solution α^* the stability of which is being investigated, where y is a small deviation. Then, the linear equation

$$y''(1 + e\cos\upsilon) - 2ey'\sin\upsilon + \mu y\cos 2\alpha^* = 0, \tag{17.33}$$

which is sometimes called a *linearized equation* for (17.1) in the vicinity of the solution α^* or an *equation in deviations,* describes such perturbed motion.

A *characteristic equation*

$$\xi^2 - 2A\xi + 1 = 0 \tag{17.34}$$

corresponds to (17.33). The coefficients of this equation can be calculated by means of the coefficients of a fundamental matrix (see Appendix B)

$$X(2\pi) = \begin{pmatrix} y_1(2\pi) & y_2(2\pi) \\ y_1'(2\pi) & y_2'(2\pi) \end{pmatrix}$$

with initial conditions

$$y_1(0) = y_2'(0) = 1, \qquad y_1'(0) = y_2(0) = 0.$$

Thus,

$$A = \frac{1}{2}[y_1(2\pi) + y_2'(2\pi)], \tag{17.35}$$

and the last coefficient of the equation is equal to unity due to (B.5).

If $|A| < 1$, then a trivial solution $y = 0$ of (17.33) is stable and the necessary condition for the stability of the solution α^* is met. If $|A| > 1$, then the trivial solution is unstable and, consequently, the solution α^* is unstable also. The boundary of the stability domain is determined by $|A| = 1$.

To extend the generating solutions obtained at $e = 0$ to the domain of arbitrary values of eccentricity, we may use a variant of the *method of parameter extension*[‖].

[‖]Consider an equation $g(x) = 0$, where $g \in R^n$, $x \in R^{n+1}$; R^n is a space of real numbers of n-dimensions. Let x_0 be a point such that the equation $g(x_0) = 0$ is valid. We need to construct a curve L in the space R^{n+1} which is determined by the equation $g(x) = 0$ and passes through the point x_0.

The unit vector $\tau_0 \in R^{n+1}$ directed along the tangent to L at the point x_0 is obtained from

$$g_x \equiv \left.\frac{\partial g(x)}{\partial x}\right|_{x=x_0} = 0$$

Assume that the condition rank $g_x = n$ is met, then the line of the vector τ_0 is determined but not its direction. Let us assign this direction. Consider the point $x_1' = x_0 + \tau_0 h_1$, where $h_1 > 0$. In general, $x_1' \bar{\in} L$, but as $h_1 \to 0$, the distance between x_1' and L has order $O(h_1^2)$. Point $x_1 \in L$ nearest to x_1' we compute using

$$x_1 = x_1' + g_x^T(x_1')\upsilon$$

where $\upsilon \in R^n$ is an unknown vector. Substitute this expression for x in $g(x) = 0$ and expand it as a series in υ. Keeping only terms of zero and first order, we get an equation to determine υ

$$[g_x(x_1')g_x^T(x_1')]\upsilon + g_x(x_1') = 0.$$

By virtue of the assumption rank $g_x = n$, the determinant of this equation is not zero. The final expression for x_1 has the form

$$x_1 = x_1' - g_x^T(x_1')[g_x(x_1')g_x^T(x_1')]^{-1}g(x_1').$$

To apply the method of parameter extension to obtain periodic solutions of (17.1), at least two points on the curve to be constructed should be known. One point can be determined from the generating solution. Another can be obtained through the expansion (17.26), or through similar series calculated for other either stationary (17.25) or non-stationary (17.23), solutions for the case of librational, periodic motions. Let us consider now results of the numerical investigation of 2π-periodic solutions of equation (17.1).

The one-parameter boundary value problem (17.1), (17.30) was solved by the method of parameter extension and the initial values of SC angular velocity $\alpha'(0)$ obtained. With the initial condition $\alpha(0) = 0$, and relevant values of e and μ, this angular velocity totally determines the periodic solution of (17.1). In the plane $(\alpha'(0), e)$, the curves of initial velocity for values of e and μ within a domain where $0 \leq e < 1$ and $|\mu| \leq 3$, are shown in Fig.17.4 [26]. It can be seen in this figure that, for each value of e, there are either one or three initial values of $\alpha'(0)$ in this domain. Denote by $\alpha^{(0)}_{1/1}$ and $\alpha^{(1)}_{1/1}$ solutions which are extended from non-stationary solutions (17.23). The subscript index in these notations associates with the ratio m/n in (17.31). Positive values of $\alpha'(0)$ situated above the dashed line in Fig.17.4 correspond to solution $\alpha^{(0)}_{1/1}$. Negative values of $\alpha'(0)$, when

---

If $\|x_1 - x'_1\| \geq \varepsilon$, where a number $\varepsilon > 0$ characterizes the accuracy of computation then, by taking x_1 as x'_1 we refine x_1 and continue to iterate until the inequality $\|x_1 - x'_1\| < \varepsilon$ is valid.

Unit vector τ_1 directed along the tangent to L at the point x_1, is obtained from the conditions

$$g_x(x_1)\tau_1 = 0, \qquad \tau_1^\top \tau_1 = 1, \qquad \tau_0^T \tau_1 = 1.$$

If h_1 is small enough then, by virtue of $\tau_0^T \tau_1 = 1$, the vector τ_1 coincides with the vector τ_0. When a few such points along the curve L become known, then prediction of the next point can be made more accurately using linear, quadratic etc methods of prediction.

Functions $g(x)$ and $g_x(x)$ for each x are calculated by integrating the differential equations whose periodic solution is being determined, and also their corresponding variational equations. In order that the algorithm works successfully, the equality rank $g_x = n$ should be valid at almost all points of L. This equality can be violated at a point where the corresponding variational equations have a non-trivial periodic solution, characterized by the same period and symmetry as the sought periodic solution. (V.Sarychev, V.Sazonov and N.Melnik. Spatial Periodic Librations of a Satellite Relative to its Center of Mass. *Cosmic Research,* 1980, Vol.18, Issue 5, pp.659-677.)

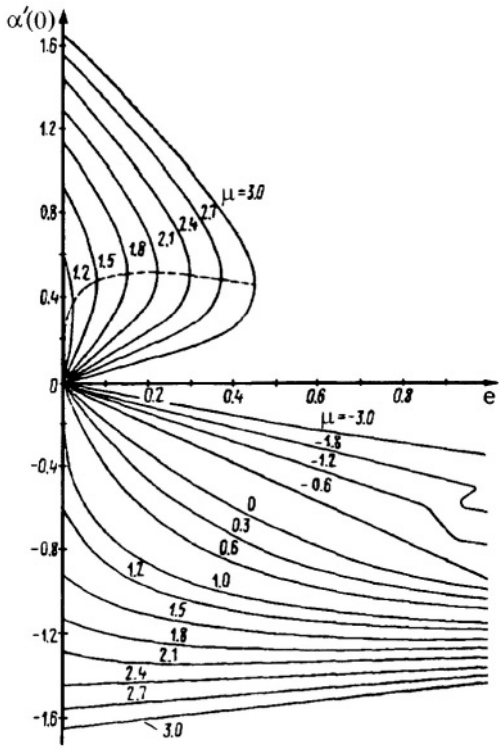

Figure 17.4. Curves of initial velocity for 2π-periodic, odd with respect to v, solutions of equation (17.1)

$\mu < 1$, correspond to the solution $\alpha_{1/1}^{(1)}$. For both generating solutions the point $(\mu = 1,\ e = 0)$ is a *bifurcation point* (i.e. a point where, at least, two solutions appear).

Denote the 2π-periodic solution generated from the stationary solution $\alpha_0 = 0$ by α_0^+ in the range $1 < \mu \leq 3$ and by α_0^- in the range $-3 < \mu \leq 1$. Solution α_0^+ is determined by the initial values $\alpha'(0) > 0$ situated below the dashed line. Initial values of $\alpha'(0) < 0$, when $\mu > 1$, correspond to the solution α_0^-.

The dashed line along which the solutions α_0^+ and $\alpha_{1/1}^{(0)}$ are combined is called a *bifurcation curve.* Along this curve, a single value of $\alpha'(0)$ corresponds to both the α_0^+ and $\alpha_{1/1}^{(0)}$ solutions as determined by the condition of verticality of the tangent to the isoline $\mu = \text{const}$.

Obtain now an approximate expression for the bifurcation curve which divides parameter space into domains within which the number of solutions is different. On substituting the expansion of Bessel's

function of the first kind

$$I_1(a) = \frac{a}{2} - \frac{a^3}{16}$$

in the first equation in (17.19), and solving it with respect to a, we obtain

$$a^3 - \frac{8(\mu-1)}{\mu}a \pm \frac{32e}{\mu} = 0. \tag{17.36}$$

The sign of its discriminant

$$D = -\left[\frac{8(\mu-1)}{3\mu}\right]^3 + \left[\frac{16e}{\mu}\right]^2.$$

specifies the number of real solutions of (17.36). If $D > 0$ this equation has a single real solution and there are three real solutions if $D < 0$. The curve separating domains which correspond to these cases, is determined by the equation $D = 0$ with the solution

$$e_{bif} = \left(\frac{2}{3}\right)^{3/2} \frac{(\mu-1)^{3/2}}{2\sqrt{\mu}}. \tag{17.37}$$

If $e > e_{bif}$ equation (17.1) has a single 2π-periodic solution. If $e < e_{bif}$ it has three solutions for each value of e. The bifurcation curve $e_{bif}(\mu)$ starts at the point ($e = 0$, $\mu = 1$) and terminates at the point ($e = 4/9 \approx 0.444$, $\mu = 3$). For $e > 4/9$, at all allowed values of the inertial parameter μ, a single periodic solution exists.

The bifurcation curve can also be calculated numerically using the condition

$$\frac{\partial \alpha'(0)}{\partial e} = 0$$

which corresponds to the verticality of a tangent to the curve of initial angular velocity $\alpha'(0)$. Another method of calculating the curve uses the conditions for the existence of non-trivial 2π-periodic solutions of the variational equation (17.33) satisfying the boundary conditions $y(0) = y(\pi) = 0$. In the parameter plane (μ, e) (Fig.17.5), the bifurcation curve obtained using this method is represented by $L_1\overline{L}_1$. It starts at the point ($e = 0$, $\mu = 1$) but, in contrast to the curve approximated by (17.37), it terminates at a slightly different point ($e \approx 0.442$, $\mu = 3$). For large e close to unity, there is another, small,

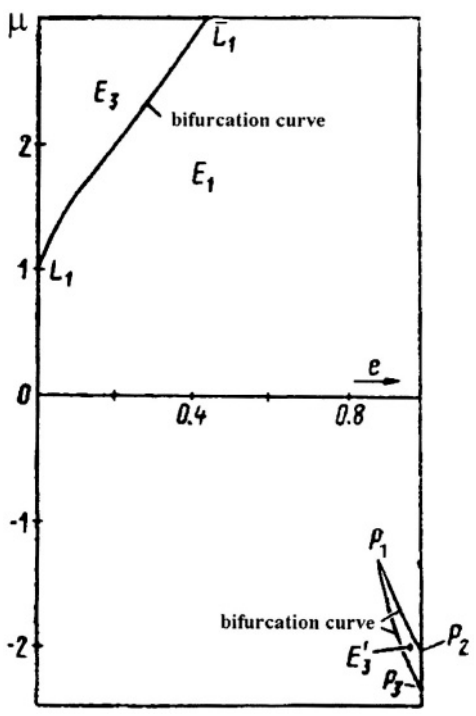

Figure 17.5. Domains where single (E_1) and three (E_3 and E_3') 2π-periodic solutions for each value of e exist

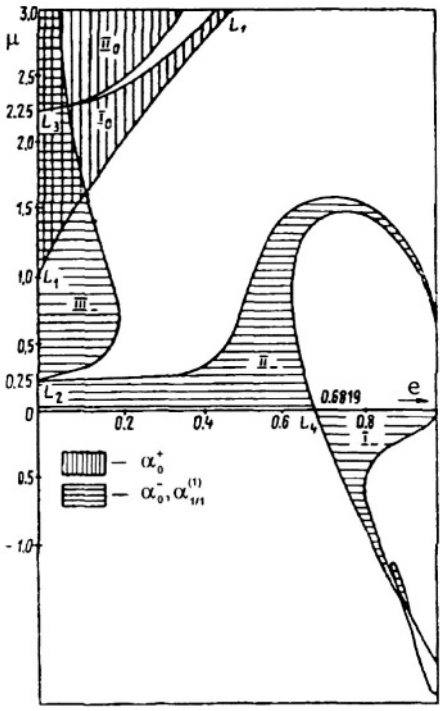

Figure 17.6. Domains of stability of 2π-periodic solutions

domain E_3' where, for each value of e, three 2π-periodic solutions also exist. The domains of stability are shown in Fig.17.6.

Spatial periodic motions of SC in elliptical orbits are calculated and analyzed with respect to bifurcation effects using a similar approach in the reference cited in the footnote on page 214.

Chapter 18

A Spinning Axisymmetric SC in Circular Orbit

A three-axis SC with mutually different moments of inertia can assume, in circular orbit, various equilibrium positions. A dynamically axisymmetric SC with two equal moments of inertia that uniformly spins around its axis of symmetry with its spin axis fixed with respect to the orbital reference system displays, so called, *stationary rotations**. The fixed positions are determined by the spin velocity and by the ratio of the SC moments of inertia. The physical basis for this behavior is a combination of torques produced by the gyroscopic, gravity-gradient and centrifugal forces.

18.1 Equations of Motion

To describe the orientation of a SC with respect to the orbital reference system, we will use classical *Euler's angles* (Fig.18.1). The corresponding matrix of directional cosines $\|a_{jk}\|$ with elements $a_{jk} = \cos(\mathbf{E}_j, \mathbf{e}_k)$,

*Stationary rotation is a specific case of *stationary motion* where the position coordinates (corresponding to a fixed position of the axis of symmetry of the SC with respect to the ORS) remain constant and the cyclic coordinates (angle of spin rotation about the axis of symmetry) vary linearly with time.

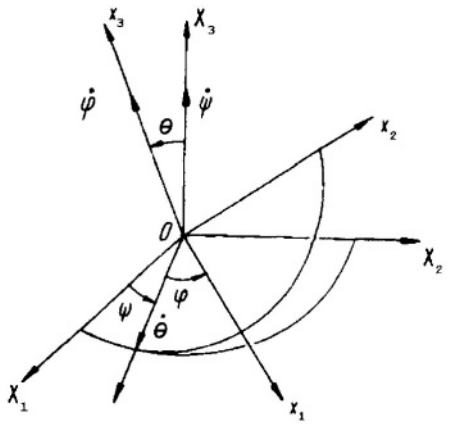

Figure 18.1. SC orientation with respect to the ORS using *Euler's angles* ψ, θ, φ

$(j, k = 1, 2, 3)$ can be defined in tabular form

	x_1	x_2	x_3
X_1	a_{11}	a_{12}	a_{13}
X_2	a_{21}	a_{22}	a_{23}
X_3	a_{31}	a_{32}	a_{33}

Here, we retain our previous notation for elements of the matrix of directional cosines as well as other notations introduced in Chapter 15. However, these cosines and kinematic relations are now expressed using Euler's angles by means of the following formulae

$$\begin{aligned} a_{11} &= \cos\psi\cos\varphi - \sin\psi\cos\theta\sin\varphi, \\ a_{21} &= \sin\psi\cos\varphi + \cos\psi\cos\theta\sin\varphi, \\ a_{31} &= \sin\theta\sin\varphi, \end{aligned}$$

$$\begin{aligned} a_{12} &= -\cos\psi\sin\varphi - \sin\psi\cos\theta\cos\varphi, & a_{13} &= \sin\psi\sin\theta, \\ a_{22} &= -\sin\psi\sin\varphi + \cos\psi\cos\theta\cos\varphi, & a_{23} &= -\cos\psi\sin\theta, \\ a_{32} &= \sin\theta\cos\varphi, & a_{33} &= \cos\theta. \end{aligned}$$

$$\begin{aligned} w_1 &= \dot{\psi}a_{31} + \dot{\theta}\cos\varphi + \omega_0 a_{21}, \\ w_2 &= \dot{\psi}a_{32} - \dot{\theta}\sin\varphi + \omega_0 a_{22}, \\ w_3 &= \dot{\psi}a_{33} + \dot{\varphi} + \omega_0 a_{23}. \end{aligned} \tag{18.1}$$

Euler's equations for a dynamically axisymmetric SC (where $A = B$), have the form

$$\begin{aligned} A\dot{w}_1 + (C-A)w_2w_3 &= 3\omega_0^2(C-A)a_{32}a_{33}, \\ A\dot{w}_2 + (A-C)w_1w_3 &= 3\omega_0^2(A-C)a_{31}a_{33}, \\ C\dot{w}_3 &= 0. \end{aligned} \tag{18.2}$$

18.2 Stationary Motions of a SC

The existence of the first integral

$$\dot{\psi}a_{33} + \dot{\varphi} + \omega_0 a_{23} = w_{30} = \text{const} \tag{18.3}$$

for (18.2) follows from the last equation of (18.2). In addition, there is the general integral of energy (*Jacobi's integral*)

$$\frac{1}{2}[A(\overline{w}_1^2 + \overline{w}_2^2) + C\overline{w}_3^2] + \frac{3}{2}\omega_0^2(C-A)a_{31}^2 + \frac{1}{2}\omega_0^2(A-C)a_{21}^2 = h_0 \tag{18.4}$$

where

$$\overline{w}_1 = w_1 - \omega_0 a_{21}, \quad \overline{w}_2 = w_2 - \omega_0 a_{22}, \quad \overline{w}_3 = w_3 - \omega_0 a_{23} \tag{18.5}$$

and h_0 is constant. Stationary rotation of a SC corresponds to a fixed position of its axis of symmetry with respect to the orbital reference system defined by angles $\psi = \psi_0 = \text{const}$; $\theta = \theta_0 = \text{const}$ and to a uniform rotation of the SC about this axis. When these constant values are inserted in the kinematic relations (18.1), their right-hand sides depend on φ and $\dot{\varphi}$ only. Next, substitute the kinematic relations in the first two equations of (18.2). Replace the derivative $\dot{\varphi}$ by its explicit expression from (18.3)

$$\dot{\varphi} = w_{30} - \omega_0 a_{23},$$

then, taking into account the third expression from (18.1), we obtain

$$\begin{aligned} A\omega_0 a_{22}(w_{30} - \omega_0 a_{23}) + \omega_0(C-A)[w_{30}a_{22} - 3\omega_0 a_{32}a_{33}] &= 0, \\ A\omega_0 a_{21}(w_{30} - \omega_0 a_{23}) + \omega_0(C-A)[w_{30}a_{22} - 3\omega_0 a_{31}a_{33}] &= 0. \end{aligned} \tag{18.6}$$

Among the matrix elements $\|a_{jk}\|$ in (18.6), only the elements a_{21}, a_{22} depend on the angle φ, which is characterized by its uniform variation with time. We may exclude angle φ from (18.6) in the

following manner. Multiply both sides of the first equation by $-\sin\varphi$ and of the second equation by $\cos\varphi$ and combine them. Then multiply both sides of these two equations by $\cos\varphi$ and $\sin\varphi$ respectively and again combine them. After reduction we get the following two equations that do not depend on φ.

$$\begin{aligned}[\overline{\alpha}\overline{\beta} + \cos\psi_0 \sin\theta_0]\sin\psi_0 = 0,\\ [(\overline{\alpha}\overline{\beta} + \cos\psi_0 \sin\theta_0)\cos\psi_0 - 3(\overline{\alpha} - 1)\sin\theta_0]\cos\theta_0 = 0,\end{aligned} \quad (18.7)$$

Here the notations used are

$$\overline{\alpha} = \frac{C}{A}, \qquad \overline{\beta} = \frac{w_{30}}{\omega_0}$$

(recall that $0 \leq \overline{\alpha} \leq 2$, due to the triangle inequalities for a rigid body). This procedure of transformation is equivalent to excluding the *cycling variable* φ by introducing a, so called, body *semi-fixed* reference system, where one axis coincides with the axis of symmetry. The two other axes are individually perpendicular to the first axis and to each other but do not follow the SC's spin rotation around the axis of symmetry.

The equalities (18.7) are valid if at least one co-factor in each of them is equal to zero. This yields four corresponding combinations of equalities. Resolving them, we obtain the following three pairs of equations:

$$\begin{aligned}&\cos\theta_0 = 0, && \sin\psi_0 = 0;\\ &\cos\theta_0 = 0, && \overline{\alpha}\overline{\beta} + \cos\psi_0 \sin\theta_0 = 0;\\ &\sin\psi_0 = 0, && \overline{\alpha}\overline{\beta}\cos\psi_0 - (3\overline{\alpha} - 4)\sin\theta_0 = 0.\end{aligned} \quad (18.8)$$

Simultaneously equating to zero both the expressions contained in square brackets in (18.7), provides us with the equation $\sin\theta_0 = 0$. The solution of this does not belong to the domain of definition of Euler's angles (this is a point of singularity for this set of angles), and we do not take it into consideration.

Let us resolve the three pairs of equations (18.8) with respect to the angles ψ_0 and θ_0 and provide a geometrical interpretation of these solutions in terms of the position of the SC axis of symmetry in the orbital reference system.

- The first pair admits the solution

$$\theta_0 = \frac{\pi}{2} + \pi k, \quad \psi_0 = \pi n, \tag{18.9}$$

for all allowed values of the parameters $\overline{\alpha}$, $\overline{\beta}$, where k and n are integers. The SC axis of symmetry is, in this case, perpendicular to the orbital plane.

- The second pair admits the solution

$$\theta_0 = \frac{\pi}{2} + \pi k, \quad \psi_0 = (-1)^{(k+1)} \arcsin \overline{\alpha}\overline{\beta}, \tag{18.10}$$

for any values of the parameters that satisfy the inequality $|\overline{\alpha}\overline{\beta}| \leq 1$. Here the SC axis of symmetry lies in the plane which is perpendicular to the radius-vector of the center of mass.

- The third pair admits the solution

$$\psi_0 = \pi k, \quad \theta_0 = (-1)^k \arcsin \frac{\overline{\alpha}\overline{\beta}}{3\overline{\alpha} - 4}, \tag{18.11}$$

for any values of the parameters that satisfy the inequalities $|\overline{\alpha}\overline{\beta}/(3\overline{\alpha} - 4)| \leq 1$ and $3\overline{\alpha} - 4 \neq 0$. The SC axis of symmetry now lies in the plane which is perpendicular to the vector of the center of mass velocity.

In Fig.18.2 the position of the SC axis of symmetry Ox_3 with respect to the axes of the orbital reference system is indicated by a bold line. The ranges of possible positions of Ox_3 are denoted by arcs.

Sometimes we will refer to (18.8) as a *stationary solution,* meaning that it includes all the stationary motions described by the three pairs of equations (18.8), without distinguishing between these pairs or indicating which concrete values of the integers k and n are chosen.

The angular velocity of SC spin rotation about its axis of symmetry

$$\dot{\varphi} = \omega_0(\overline{\beta} + \cos\psi_0 \sin\theta_0)$$

is obtained from the third equality of (18.5), taking into account values of θ_0, ψ_0 from one of the solutions of (18.8) and from the first integral (18.3).

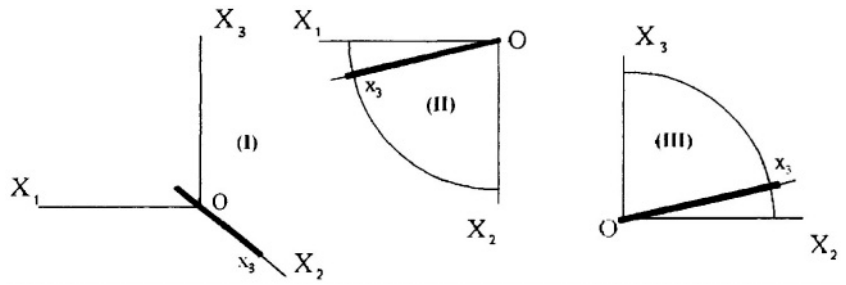

Figure 18.2. Position of the SC axis of symmetry Ox_3 for the three types of stationary rotation

Although the magnitudes of the angles ψ_0 and θ_0 are constant, solutions (18.9), (18.10) and (18.11) are referred to as *stationary motions,* since the SC axis of symmetry performs rotations with respect to inertial space which correspond individually to three types of regular precession with angular velocity ω_0. Recall that ω_0 denotes the angular velocity of rotation of the orbital reference system around the axis OX_2. The angle Θ between the axis of spin rotation Ox_3 and the axis of precession is determined by the relation

$$\cos\Theta = a_{23} = -\cos\psi_0 \sin\theta_0.$$

In correspondence with the figure of the surface that is traced out in inertial space by the SC axis of symmetry, motion (18.9) is named *cylindrical or tube precession.* Similarly, (18.10) is named *hyperboloid precession* while (18.11) is named *conic precession.* Consider next the stability conditions for these motions.

18.3 Conditions of Stability for Stationary Motions

Sufficient conditions for stability are obtained using the first integral (18.4). Necessary conditions are obtained from the existence of imaginary roots of the fundamental equation for (18.1) and (18.2), linearized in the neighborhood of the stationary motion investigated.

- For solution (18.9) at $\theta_0 = \pi$ and $\psi_0 = \pi/2$, sufficient conditions

are

$$\overline{\alpha}\overline{\beta} > 1, \qquad \overline{\alpha}\overline{\beta} + 3\overline{\alpha} - 4 > 0$$

and necessary conditions are

$$(\overline{\alpha}\overline{\beta} - 1)(\overline{\alpha}\overline{\beta} + 3\overline{\alpha} - 4) \geq 0,$$
$$[(\overline{\alpha}\overline{\beta} - 1)^2 + 3\overline{\alpha} - 2]^2 + 4(\overline{\alpha}\overline{\beta} - 1)(\overline{\alpha}\overline{\beta} + 3\overline{\alpha} - 4) \geq 0.$$

- For solution (18.10) the sufficient and necessary conditions coincide and have the form

$$\overline{\alpha} - 1 > 0, \qquad \sin\psi_0 \neq 0.$$

- For solution (18.11) a sufficient condition is

$$\overline{\alpha} - 1 < 0$$

and necessary conditions are either

$$\overline{\alpha} - 1 \leq 0$$

or

$$\overline{\alpha} - \frac{4}{3} \geq 0,$$
$$81\overline{\alpha}^6(\overline{\alpha} - 1)^2\overline{\beta}^4 - 5\overline{\alpha}^2(\overline{\alpha} - 1)(18\overline{\alpha}^2 - 27\overline{\alpha} + 8)(3\overline{\alpha} - 4)^2\overline{\beta}^2 +$$
$$+(3\overline{\alpha} - 4)^4 \geq 0.$$

All of the above conclusions concerning the regular precession of an axisymmetric SC do not involve the angle φ of spin rotation. There is no stability of this latter angle because of the existence of the first integral (18.3). For this reason, there is only stability of the position of the axis of symmetry of the SC in the orbital reference system, and of the three components of SC angular velocity.

Detailed and complete analysis of this problem was carried out by F.L.Chernousko[†]

The stationary motions considered above are more useful for understanding and interpreting the actual motions of SC than for design

[†]F.L.Chernousko. On Satellite Motion around its Center of Mass under the Action of Gravitational Torques. *Applied Mathematics and Mechanics,* 1963, Vol.27, N 3, pp.474–483.

application. A practical example is provided by consideration of a situation when the Russian orbital station *Salyut-7* (which had an almost axisymmetric configuration), tilted in 1985 to the local vertical by tens of degrees. This is interpreted to have been caused by the onset of rotation about its longitudinal axis, induced by the development of a, non-conservative, aerodynamic torque[‡].

An example of the application in motion analysis of (18.11) is its use in estimating the orientation of a, close to, axisymmetric SC in slow spin rotation.

[‡]V.A.Sarychev, M.Yu.Belyaev, S.P.Kuz'min, V.V.Sazonov and T.N.Tyan. Investigation of Attitude Motion of the Salyut-7 Orbital Station for Long Time Intervals. *Acta Astronautica,* 1987, Vol.16, Issue pp.165–192.

Chapter 19

Equilibrium of a Gyrostat

19.1 Equations of Motion

A SC provided with a statically and dynamically balanced axisymmetric rotor (or with a set of such rotors) spinning inside it with constant angular velocity is called a *gyrostat.* Such rotors cause the SC to take up more equilibrium positions relative to the orbital reference system than are obtained utilizing the gravity-gradient torque effect (with consequent interesting practical applications). The effect of uncompensated permanent angular momentum on SC motion may, associatively, be investigated.

Assume in the case of a particular gyrostat that a rotor spins about a fixed axis; that friction between the rotor and the SC body is negligible and that the rotor's motion does not change the moments of inertia of the SC. An expression for the angular momentum of the gyrostat is as follows

$$\mathbf{K} = J\boldsymbol{\omega} + \sum_{i=1}^{n} J_i(\boldsymbol{\omega} + \boldsymbol{\omega}_{ri}), \tag{19.1}$$

where J is the tensor of inertia of the SC body, $\boldsymbol{\omega}$ is a vector of SC absolute angular velocity, J_i is the tensor of inertia of the i-th rotor and $\boldsymbol{\omega}_{ri}$ is the relative angular velocity of the i-th rotor with respect to the SC body (Fig.19.1). Denote the gyrostat tensor of inertia by $\overline{J} = J + \sum_{i=1}^{n} J_i$, then its equation of motion in circular orbit in vector

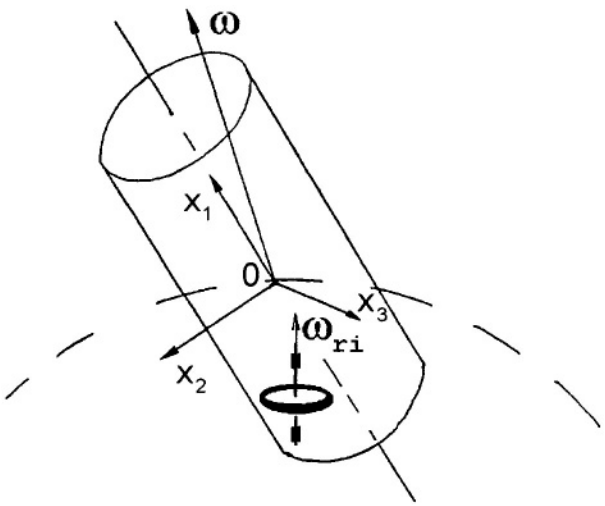

Figure 19.1. A gyrostat (SC with a rotor)

form is as follows

$$\overline{J}\dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times (\overline{J}\boldsymbol{\omega} + \sum_{i=1}^{n} J_i \boldsymbol{\omega}_{ri}) = \omega_0^2 A^\top \mathbf{E}_3 \times J A^\top \mathbf{E}_3.$$

Here A is a transformation matrix (15.4) with elements a_{jk}. Introduce variables h_j, Ω_j using the formulae $\omega_j = \omega_0 \Omega_j$, $\overline{h}_j = \omega_0 h_j$, where $\overline{h}_j$ are projections of the total angular momentum of the rotors on to the j-th axis of a body-fixed reference system with axes coincident with the principal central axes of inertia of the gyrostat. The argument of latitude $u = \omega_0(t - t_0)$ is chosen to be a, unitless, independent variable, where t_0 is the moment of first crossing the ascending node. Then, the dynamical equations of SC motion in scalar form are

$$\begin{aligned} A\Omega_1' + (C-B)(\Omega_2\Omega_3 - 3a_{32}a_{33}) &= h_2\Omega_3 - h_3\Omega_2, \\ B\Omega_2' + (A-C)(\Omega_1\Omega_3 - 3a_{31}a_{33}) &= h_3\Omega_1 - h_1\Omega_3, \\ C\Omega_3' + (B-A)(\Omega_1\Omega_2 - 3a_{31}a_{32}) &= h_1\Omega_2 - h_2\Omega_1. \end{aligned} \tag{19.2}$$

The plane angles α, β and γ introduced in Section 15.2 and corresponding kinematic relationships (15.8) can now be used to define the angular positions of a SC with respect to the ORS.

Express the generalized integral of energy of equations (19.2) using the SC's, unitless, relative velocity components $\overline{\Omega}_1$, $\overline{\Omega}_2$, $\overline{\Omega}_3$

$$\tfrac{1}{2}(A\overline{\Omega}_1^2 + B\overline{\Omega}_2^2 + C\overline{\Omega}_3^2) + \tfrac{3}{2}[(A-C)a_{31}^2 + (B-C)a_{32}^2] \tag{19.3}$$

$$+\frac{1}{2}[(B-A)a_{21}^2+(B-C)a_{23}^2]-(h_1a_{21}+h_2a_{22}+h_3a_{23})=h_0.$$

We next obtain stationary solutions of (19.3), which describe the equilibrium positions of the SC with respect to the ORS. Substituting the zero relative angular velocity components $\overline{\Omega}_1=\overline{\Omega}_2=\overline{\Omega}_3=0$ and the constant angles $\alpha=\alpha_0,\ \beta=\beta_0,\ \gamma=\gamma_0$ in (19.3), we get a set of algebraic equations

$$\begin{aligned}(B-C)(a_{22}a_{23}-3a_{32}a_{33})&=h_2a_{23}-h_3a_{22},\\(A-C)(a_{23}a_{21}-3a_{33}a_{31})&=h_3a_{21}-h_1a_{23},\\(B-A)(a_{21}a_{22}-3a_{31}a_{32})&=h_1a_{22}-h_2a_{21}.\end{aligned}\tag{19.4}$$

These identities when combined with expressions for the directional cosines form a closed set of trigonometric equations from which $\alpha_0,\ \beta_0,\ \gamma_0$ can be determined. To obtain the general solution of (19.4) is a complicated task but it has been successfully accomplished*. In consequence, through a proper choice of gyrostat parameters, any given axis of a SC can be pointed towards the Earth. Also, any required orientation with respect to the ORS can be achieved.

19.2 Particular Cases of Equilibrium Positions

Consider three particular cases of gyrostat orientation.

Case 1 ($h_3=0$). Under this condition, the following particular solution of (19.4) exists

$$\alpha_0=\gamma_0=0,\ (B-A)\sin\beta_0\cos\beta_0+h_2\sin\beta_0-h_1\cos\beta_0=0.\tag{19.5}$$

This corresponds to the attitude of a gyrostat turned around the local vertical through a yaw angle with value (β_0). Thus, if it is necessary to provide a turn with a specific yaw value, we can calculate the required values of h_1 and h_2 using (19.5). Alternatively, if these latter components are known, we can determine the complementary value of β.

*V.A.Sarychev and S.A.Gutnik. Relative Equilibria of a Gyrostat Satellite. *Cosmic Research,* 1984, Vol.22, Issue 3, pp.257–260.

Sufficient conditions for the stability of this equilibrium position, which are obtained using the first integral (19.3) to form Lyapunov's function, are

$$\begin{gathered}(A-C)+(B-A)\sin^2\beta_0>0,\\ (B-A)\cos 2\beta_0+h_1\sin\beta_0+h_2\cos\beta_0>0,\\ [(A-C)+(B-A)\sin^2\beta_0][4(B-C)-(B-C)\sin^2\beta_0\\ +h_2\cos\beta_0]-3(B-C)^2\sin^2\beta_0>0.\end{gathered}\tag{19.6}$$

Case 2 ($h_1 = 0$). This condition results in the following particular solution

$$\alpha_0=\beta_0=0,\ 4(B-C)\sin\gamma_0\cos\gamma_0+h_2\sin\gamma_0+h_3\cos\gamma_0=0.\tag{19.7}$$

This corresponds to the attitude of a gyrostat turned around the local tangent through a roll angle with value (γ_0). Sufficient conditions for the stability of this equilibrium position are

$$\begin{gathered}(A-C)-(B-C)\sin^2\gamma_0>0,\\ 4(B-C)\cos 2\gamma_0+h_2\cos\gamma_0-h_3\sin\gamma_0>0,\\ [(A-C)-(B-C)\sin^2\gamma_0][(B-A)-(B-C)\sin^2\gamma_0\\ +h_2\cos\gamma_0-h_3\sin\gamma_0]-3(B-C)^2\sin^2\gamma_0\cos^2\gamma_0>0.\end{gathered}\tag{19.8}$$

Case 3 ($h_1 = h_3 = 0$). These conditions result in

$$\alpha_0=\beta_0=\gamma_0=0.\tag{19.9}$$

Sufficient conditions for the stability of this, trivial, equilibrium position are

$$(A-C)>0,\quad (B-A)+h_2>0,\quad 4(B-C)+h_2>0.\tag{19.10}$$

If $B = A$ and there are no rotors, then there is no yaw component of the restoring torque. In these circumstances, the axis of dynamical symmetry of the SC orients along the local vertical, and alternative specific orientations of the other SC axes cannot be realized. Rotors can be successfully used to orient a SC in yaw, including the trivial angular position (19.9) (i.e. this latter position can be taken up even

by an axisymmetric SC through using rotors). Sufficient conditions for the stability of (19.9) transformed from (19.10) for $B = A$ are as follows

$$(A - C) > 0, \quad h_2 > 0. \tag{19.11}$$

Consider the solution $\alpha_0 = \beta_0 = \gamma_0 = 0$ when $h_2 \neq 0$ (i.e. when the axis of rotor rotation is directed along the gyrostat axis Ox_2), and the moment of inertia corresponding to this axis is not a maximum. In this case, the stability conditions (19.10) can be satisfied by choosing an appropriate h_2. Recall that, because of the conditions $B > A > C$, the axis Ox_2 of a SC without rotors has to be an axis of maximum moment of inertia. Installation of a rotor with $h_2 > 0$ along this axis weakens this requirement (see the second inequality in (19.10)).

Another feature of a gyrostat is that the rotor provides a unique stable orientation in yaw, because switching the sign of h_2 can cause instability at the position $\alpha_0 = \beta_0 = \gamma_0 = 0$ and stability of the inverse position $\alpha_0 = \beta_0 - \pi = \gamma_0 = 0$. Such uniqueness cannot be attained through utilizing the gravity-gradient torque (Section 16.2).

Consider now the necessary conditions for stability of the trivial equilibrium position. Write equations of motion linearized in the vicinity of (19.9)

$$\begin{gathered} B\alpha'' + 3(A - C)\alpha = 0, \\ C\beta'' + [h_2 - (A - B + C)]\gamma' + [h_2 + (B - A)]\beta = 0, \\ A\gamma'' - [h_2 - (A - B + C)]\beta' + [h_2 + 4(B - C)]\gamma = 0. \end{gathered} \tag{19.12}$$

The characteristic equation of (19.12) has imaginary roots only, that is the necessary conditions for stability are satisfied if

$$\begin{gathered} (A - C) > 0, \\ [h_2 + (B - A)][h_2 + 4(B - C)] > 0, \\ \{A[h_2 + (B - A)] + C[h_2 + 4(B - C)] + [h_2 - (A - B + C)]^2\}^2 \\ -4AC[h_2 + (B - A)][h_2 + 4(B - C)] > 0. \end{gathered} \tag{19.13}$$

When

$$h = \sqrt{h_1^2 + h_2^2 + h_3^2} \to +\infty$$

there are eight equilibrium positions of the gyrostat and, when

$$h \ll \sqrt{A^2 + B^2 + C^2},$$

at least twenty four equilibrium positions of the gyrostat exist. As $h \to 0$ these positions transfer to the corresponding equilibrium positions of a rigid body in the central gravitational field (Section 16.2).

Chapter 20

SC Motion Affected by an Aerodynamic Torque

20.1 General Assumptions

As a SC moves through a rarefied atmosphere, the interaction of the, consequently, counter-flowing molecular stream with the SC body produces different dynamic effects.

- If the center of pressure does not coincide with the center of mass of the SC, then either a restoring or a tumbling torque arises, depending on the relative positions of these centers in the molecular stream. Since the atmosphere is linked to the spinning surface of the Earth, the vector of velocity of the stream does not lie in the orbital plane of the SC.
- SC rotation initiates an effect called 'stream wash', which produces an aerodynamic dissipative torque.
- The atmospheric density is not the same at each point of the surface of the SC but this produces a negligible 'gradient effect' on SC motion.
- The natural, temperature dependent, velocities of the impacting air molecules also only produce a negligible effect on SC motion.

Interaction with a counter-flowing molecular stream can be used to provide passive attitude control of a SC. Early missions during which

this technique was applied include the scientific SC *Cosmos-149* and *Cosmos-320.*

In order to be able to calculate the associated restoring aerodynamic torque, the nature of the interaction between the SC surface and the counter-flowing molecular stream must be studied. Experience shows that the following mode of interaction is the most probable. On impacting with the SC surface, a particle transfers its total energy to the surface and attains temperature equilibrium with the site of impact. After this process the particle exits this site and goes off into space with a temperature velocity equal to the temperature velocity of the surface molecules of the SC. Since this temperature velocity is much lower than the velocity of the counter-flowing molecules, it can be assumed that the impact was absolutely inelastic. In this case the elementary force $d\mathbf{F}$ acting on an element of surface $d\mathbf{S}$ (which is defined by its vector normal) with magnitude equal to the area dS is given by the expression

$$d\mathbf{F} = -\frac{1}{2} c_x(\gamma_V) \rho_a V^2 \frac{\mathbf{V}}{V} dS.$$

Here $\mathbf{V}$ is the vector of the velocity of the element of surface with respect to the flowing molecular stream; $V = |\mathbf{V}|$; ρ_a is the air density at the relevant location and $c_x(\gamma_V)$ is the *drag coefficient* as a function of the *angle of attack* γ_V (i.e. the angle between the vector $\mathbf{V}$ and the vector $d\mathbf{S}$) (Fig. 20.1).

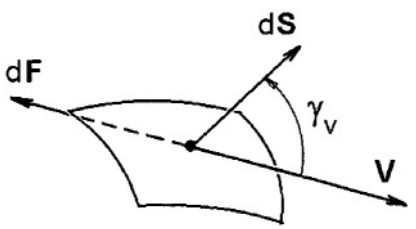

Figure 20.1. Elementary drag force

Consider now the effect of stream wash. As a rotating body moves through the atmosphere, the velocities of its center of mass and the center of pressure differ from each other as a result of the relative velocity caused by rotation of the body. Because of this difference, a nonconservative term proportional to the angular velocity of the

body appears in the expression for aerodynamic torque. This term can provide asymptotic stability of SC motion. The origin of the name 'stream wash' lies in the noncoincidence (relative slant) of the directions of the center of mass velocity and the drag force at the pressure center.

Write the vector equation of SC motion under the effect of an aerodynamic torque $\mathbf{M}_a = \mathbf{r}_a \times \mathbf{F}_a$ caused by the drag force $\mathbf{F}_a = -\rho_a V S c_x \mathbf{V}/2$ concentrated at the center of pressure a

$$J\dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times J\boldsymbol{\omega} = \mathbf{M}_a,$$

where $\mathbf{V} = \mathbf{V}_0 + \boldsymbol{\omega} \times \mathbf{r}_a$ is the velocity of the center of pressure, $\mathbf{V}_0$ is the velocity of the SC center of mass, S is a *cross-sectional area* of the SC, $\boldsymbol{\omega}$ is the angular velocity of the SC with respect to the ORS and $\mathbf{r}_a$ is the radius-vector of the center of pressure with respect to the SC center of mass.

Suppose that the center of pressure is situated on the axis Ox_1 of the BRS. Also assume that both $\mathbf{r}_a$ and c_x do not depend on the orientation of the SC with respect to the vector of its translational velocity so that $\mathbf{r}_a$ and c_x are each constant in the equations of motion. Physically this corresponds to assuming that all interactions of the SC with the atmosphere are reduced to the drag forces which act on a sphere with its center fixed at a given distance from the center of mass of the SC. Since the center of pressure of a sphere coincides with its geometric center, we can replace all the aerodynamic forces by their resultant applied at the center of the sphere. In a complementary application, a sphere that is used to produce an aerodynamic restoring torque on a SC is called an *aerodynamic stabilizer.*

Consider now planar motion of the SC (Fig. 20.2). Let the vector $\mathbf{r}_a$ be antiparallel to the axis Ox_1 of the BRS, which makes an angle α with the axis OX_1 of the ORS. Then, the expression for the magnitude of $\mathbf{V}$ is $V = (V_0^2 + 2V_0 r_a \dot{\alpha} \sin\alpha + r_a^2 \dot{\alpha}^2)^{1/2}$. Symbol $(\dot{})$ means differentiation with respect to time t. The projection of the vector equation of motion on the normal to the orbital plane has the form

$$J\ddot{\alpha} = -\frac{\rho_a S c_x}{2}(r_a V_0 \sin\alpha + r_a^2 \dot{\alpha}^2)(V_0^2 + 2V_0 r_a \dot{\alpha} \sin\alpha + r_a^2 \dot{\alpha}^2)^{1/2},$$

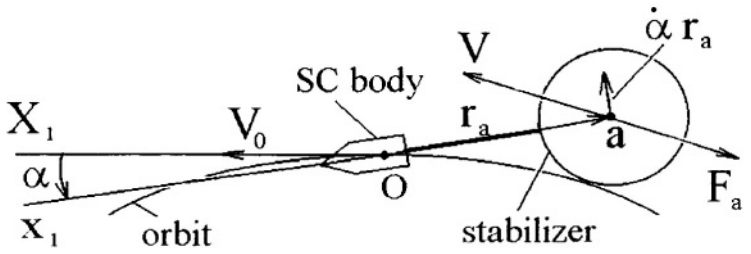

Figure 20.2. SC with aerodynamic stabilizer in the planar case

where J is the moment of inertia of the SC about its principal axis perpendicular to the plane of its orbit. On transforming this we obtain

$$J\ddot{\alpha} = -\frac{\rho_a S c_x r_a V_0^2}{2}\left(\sin\alpha + \frac{r_a\dot{\alpha}^2}{V_0}\right) \times \left(1 + 2\sin\alpha\frac{r_a\dot{\alpha}}{V_0} + (\frac{r_a\dot{\alpha}}{V_0})^2\right)^{1/2}. \qquad (20.1)$$

The energy dissipation effect of stream wash is determined by the relationship $\omega_* L_*/V_0$ which is called *Strouhal's number* Sr, where ω_* is a typical angular velocity of the SC, L_* is a representative SC dimension and V_0 is the velocity of the SC center of mass with respect to the counter-flowing stream. For a representative SC with $\omega_* \simeq 10^{-3}$ rad/sec, $L_* \simeq 1$ m and $V_0 \simeq 8\,000$ m/sec, Strouhal's number is $Sr \simeq 10^{-7}$. We therefore retain in (20.1) only terms up to the first order of the ratio $r_a\dot{\alpha}/V_0$,

$$J\ddot{\alpha} + \frac{\rho_a S c_x r_a V_0^2}{2}(1+\sin^2\alpha)\frac{r_a\dot{\alpha}}{V_0} + \frac{\rho_a S c_x r_a V_0^2}{2}\sin\alpha = 0. \qquad (20.2)$$

The second term here is a, non conservative, dissipating term, and it is assumed that the center of pressure is situated behind the center of mass along the direction of SC motion so that $\mathbf{r}_a = (-r_a,\, 0,\, 0)_x$ and $r_a > 0$.

Equation (20.2) has a solution $\alpha = 0$ and the corresponding variational equation written in terms of the unitless true anomaly is

$$J\alpha'' + \frac{\rho_a S c_x r_a V_0^2}{2}\frac{r_a\omega_0\alpha'}{V_0} + \frac{\rho_a S c_x r_a V_0^2}{2}\alpha = 0. \qquad (20.3)$$

The associated characteristic equation has roots with a negative real part proportional to the factor of α' in (20.3) and, consequently, to Strouhal's number. This means that, in order to decrease the amplitude of SC librations by a factor $e = 2.72\ldots$, some 10^7 sec are required. Such a slow change is negligible when considered relative to the attitude motion of a SC. Thus, we need to retain in the equations of motion only the conservative terms. In the deep atmosphere the effect of stream wash can be significant. It is noted, for example, that the perturbed motion of a bomb equipped with a stabilizing fin is dissipated after its release from an aircraft due to the 'stream wash' effect in the dense surrounding atmosphere.

20.2 Atmospheric Density Approximation

Consider the density of the Earth's upper atmosphere. The density distribution of an homogeneous, ideal, gas with altitude h is determined by the equation of hydrostatic balance

$$dP = -\rho_a g dh \tag{20.4}$$

and by the Clapeyron-Mendeleev formula

$$\rho_a = \frac{PM}{R_0 T}, \tag{20.5}$$

where M is the mean molecular weight, T is the absolute temperature, P is the gas pressure, R_0 is the *universal gas constant* and g is the gravitational acceleration at the altitude h.

Substituting ρ_a from (20.5) in (20.4) and integrating the latter from an initial value h_0 to the altitude h, we obtain formulae for the atmospheric pressure and density

$$P = P_0 \exp\left(-\int_{h_0}^{h} \frac{gM}{R_0 T} dh\right),$$

$$\rho_a = \rho_{a0} \frac{T_0 M}{T M_0} \exp\left(-\int_{h_0}^{h} \frac{gM}{R_0 T} dh\right). \tag{20.6}$$

In general, g, M and T are functions of altitude and time. The main difficulty with regard to atmospheric models is the complexity of the relationships between M, T and time which are of a quasi-cyclic nature (see Section A.1.17). These variations are associated with quasi-periodic changes in the solar energy absorbed by the Earth's atmosphere which are daily, seasonal and half-yearly. There is also a quasi-periodicity associated with the ~ 11 year solar cycle. In low Earth orbits with altitudes in the range ~ 500 – 800 km, variations in atmospheric density by a factor of 10 or more can be consequent on changes in solar activity (see Section A.1.18).

Usually, simplified models of the atmosphere take into account only the altitude profile. The *isothermal model* is among these. This model is obtained from (20.6) at $T = T_0 = \text{const}$, $M = M_0 = \text{const}$

$$\rho_a = \rho_{a0} \exp\left(-\frac{h - h_0}{H_{hmg}}\right), \tag{20.7}$$

where $H_{hmg} = R_0 T_0/(g M_0) = \text{const}$ is the reciprocal of the *scale height* of the atmosphere. Estimations show that, in the altitude range from 300 up to 700 km, the magnitude of H_{hmg} varies almost linearly from 45 km up to 70 km. In an elliptical orbit h varies between altitudes of perigee (h_π) and apogee (h_α). For this reason instead of (20.7) the following expression is usually used

$$\rho_a = \rho_{a\pi} \exp\left(-\frac{h - h_\pi}{H_\pi}\right). \tag{20.8}$$

Here $H_\pi = H(h_\pi)$. Model (20.8) is one of the simplest available and does not even take into account daily variations in density.

For qualitative analysis, analytical expressions for aerodynamic force and torque are usually used. If the SC body has a convex, spheroidal, shape, then axis symmetry allows us to approximate the force and torque using spatial attack angle functions. These could, for instance, be presented in the form of an attack angle cosine series. If the SC center of mass lies on the geometric axis of symmetry then, lifting and drag forces, and also a restoring torque, act on the SC.

To obtain equations of SC motion under the effect of an aerodynamic torque, introduce the following assumptions:

- the Earth and its atmosphere rotate with a common angular velocity;

- the effect of the atmosphere reduces to a drag force applied at the center of pressure and oppositely directed to the velocity of the center of mass of the SC relative to the atmosphere;

- the effect of the atmospheric drag on the translational motion of the SC is negligible, that is we can regard the orbit of the SC as fixed in inertial space.

Taking account of the second assumption, let $\mathbf{r}_p = (r_{p1}, r_{p2}, r_{p3})_x$ be the radius-vector of the pressure center with respect to the origin of the BRS. The projections of the SC center of mass velocity vector $\mathbf{V}_0$ relative to the counter-flowing molecular stream then have the form

$$V_{X1} = \sqrt{\frac{\mu_g}{p}}(1 + e\cos\upsilon) - \Omega_a \rho \cos i, \quad V_{X2} = \Omega_a \rho \sin i \cos\upsilon,$$
$$V_{X3} = \sqrt{\frac{\mu_g}{p}} e \sin\upsilon, \tag{20.9}$$

where Ω_a is the angular velocity of the spin rotation of the Earth and R is the radius-vector of the SC center of mass with respect to the Earth's center (which is determined using $R(1 + e\cos\upsilon) = p$). Expressions (20.9) are obtained taking into account (5.21) and (5.22). Write the aerodynamic torque in vector form

$$\mathbf{M}_a = -\mathbf{r}_a \times \frac{\rho_a V_0 S_m c_x}{2} \mathbf{V}_0. \tag{20.10}$$

Then the dynamic equations of SC motion can be expressed in vector form

$$J\dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times J\boldsymbol{\omega} = \frac{\mu_g}{R^3} A^\top \mathbf{E}_3 \times JA^\top \mathbf{E}_3 - \mathbf{r}_a \times \frac{\rho_a V_0 S_m c_x}{2} \mathbf{V}_0. \tag{20.11}$$

The plane angles α, β, γ are chosen as generalized variables, as was done in Section 15.2 so that the kinematic equations have the form shown in (15.22).

Using now the matrix A of directional cosines (see Section 15.2), we write expressions for the projections of the aerodynamic torque on the axes of the BRS

$$M_1 = -Q\left(r_{p2}\sum_{j=1}^{3}\frac{V_{Xj}}{V}a_{j3} - r_{p3}\sum_{j=1}^{3}\frac{V_{Xj}}{V}a_{j2}\right),$$

$$M_2 = -Q\left(r_{p3}\sum_{j=1}^{3}\frac{V_{Xj}}{V}a_{j1} - r_{p1}\sum_{j=1}^{3}\frac{V_{Xj}}{V}a_{j3}\right), \qquad (20.12)$$

$$M_3 = -Q\left(r_{p1}\sum_{j=1}^{3}\frac{V_{Xj}}{V}a_{j2} - r_{p2}\sum_{j=1}^{3}\frac{V_{Xj}}{V}a_{j1}\right),$$

where $Q = \rho_a V_0^2 S c_x/2$ and $V_0^2 = V_{X1}^2 + V_{X2}^2 + V_{X3}$.

We next consider two aspects of SC dynamics. In the first we investigate the influence on SC motion of the rotation of the atmosphere produced by the spinning Earth. In the second we consider equilibrium positions of a SC in circular orbit, without taking the Earth's spin into account.

20.3 Effect of the Earth's Rotation

Again, let the center of pressure be situated on the axis Ox_1 of the BRS (i.e. $r_{p2} = r_{p2} = 0$) and assume that both r_{p1} and c_x do not depend on the orientation of the SC with respect to the vector of its translational velocity. Then, we can use $r_{p1} = \text{const}$ and $c_x = \text{const}$ in the equations of motion.

For such a practical case, Euler's dynamical equations written using the argument of latitude $u = w_0(t - t_0)$ and the unitless variables Ω_1, Ω_2, Ω_3 ($\omega_j = \omega_0\Omega_j$), have the form

$$A\Omega_1' + (C - B)\Omega_2\Omega_3 = 3(C - B)a_{32}a_{33}, \qquad (20.13)$$

$$B\Omega_2' + (A - C)\Omega_1\Omega_3 = 3(A - C)a_{31}a_{33} - \bar{h}_1\left(\frac{V_{X1}}{V}a_{13} + \frac{V_{X2}}{V}a_{23}\right),$$

$$C\Omega_3' + (B - A)\Omega_1\Omega_2 = 3(B - A)a_{31}a_{32} + \bar{h}_1\left(\frac{V_{X1}}{V}a_{12} + \frac{V_{X2}}{V}a_{22}\right).$$

Here the *aerodynamic parameter* $\bar{h}_1 = -Qr_{p1}/\omega_0^2 = \text{const}$, which has the same dimensions as moments of inertia, is introduced; symbol ($'$) denotes differentiation with respect to the argument of latitude. The negative sign makes $\bar{h}_1$ positive for stable equilibrium positions of the SC. For arbitrary positioning of a center of pressure fixed in the BRS, three aerodynamic parameters $\bar{h}_1$, $\bar{h}_2$ and $\bar{h}_3$ could similarly be introduced.

Using the formulae (20.9). we get the following expressions

$$\frac{V_{X1}}{V} = \frac{1 - \varepsilon\cos i}{\sqrt{(1 - \varepsilon\cos i)^2 + \varepsilon^2\sin^2 i\cos^2 u}},$$

$$\frac{V_{X2}}{V} = \frac{\varepsilon \sin i \cos u}{\sqrt{(1-\varepsilon\cos i)^2 + \varepsilon^2 \sin^2 i \cos^2 u}} \tag{20.14}$$

where $\varepsilon = \Omega_a/\omega_0$, which is approximately equal to 1/16 for low Earth orbits. Although the case for a circular orbit was specifically considered above, these expressions also hold for $e \neq 0$, and the result obtained does not depend on the eccentricity of the orbit.

Now we construct the forced solution of system (20.13) in the form of a series in integer powers of the parameter ε, which is taken to be small

$$\alpha = \alpha_0 + \varepsilon\alpha_1 + \ldots, \quad \beta = \beta_0 + \varepsilon\beta_1 + \ldots, \quad \gamma = \gamma_0 + \varepsilon\gamma_1 + \ldots. \tag{20.15}$$

Here α_j, β_j, γ_j $(j = 0, 1, \ldots)$ are unknown functions of u and of the parameters of the system. Substitute the series (20.15) in (20.13). Sum the terms of the same degree in ε and equate them to zero. Instead now of writing and solving the equations to determine α_0, β_0, γ_0, we adopt the following reasoning process. If there is no aerodynamic effect, equations (20.13) reduce to (16.17) which have a trivial solution (16.18) for an equilibrium position of the SC. The terms of order $O(\varepsilon^0)$ in (20.13) have the forms $\bar{h}_1 a_{12}$ and $\bar{h}_1 a_{13}$ with respect to (20.14) at $\varepsilon = 0$. If $\alpha = \beta = \gamma = 0$, then $a_{12} = a_{13} = 0$. This means that the particular solution $\alpha_0 = \beta_0 = \gamma_0 = 0$ satisfies the equations of zero-order in ε.

Next we write equations for the coefficients α_1, β_1, γ_1 of the first order in the expansion (20.15)

$$\begin{aligned} B\ddot{\alpha}_1 + [3(A-C) + \bar{h}_1]\alpha_1 &= 0, \\ A\ddot{\gamma}_1 + (A + C - B)\dot{\beta}_1 + 4(B-C)\gamma_1 &= 0, \\ C\ddot{\beta}_1 - (A + C - B)\dot{\gamma}_1 + (B - A + \bar{h}_1)\beta_1 &= \bar{h}_1 \sin i \cos u \end{aligned} \tag{20.16}$$

which can be separated into two parts. The first part contains a single equation in α_1 with solution $\alpha_1 = 0$. The second part contains two equations in β_1 and γ_1. The forced solution of the latter pair is obtained in the form $\beta_1 = A_\beta \cos u$ and $\gamma_1 = A_\gamma \sin u$, where expressions for the unknown amplitudes A_β and A_γ are

$$\begin{aligned} A_\beta &= \frac{[4(B-C) - A]\bar{h}_1 \sin i}{(A + C - B)^2 + (\bar{h}_1 - A - C + B)[4(B-C) - A]}, \\ A_\gamma &= -\frac{(A + C - B)\bar{h}_1 \sin i}{(\bar{h}_1 - A - C + B)[4(B-C) - A]}. \end{aligned} \tag{20.17}$$

Now we write the solution (20.15) to within an accuracy of the order of $O(\varepsilon^2)$ in the form

$$\alpha = 0, \tag{20.18}$$

$$\beta = \varepsilon \frac{[4(B-C)-A]\bar{h}_1 \sin i}{(A+C-B)^2 + (\bar{h}_1 - A - C + B)[4(B-C)-A]} \cos u,$$

$$A_\gamma = -\varepsilon \frac{(A+C-B)\bar{h}_1 \sin i}{(\bar{h}_1 - A - C + B)[4(B-C)-A]} \sin u.$$

Analyze next the behaviour of the amplitude of the forced libration at its limits. If $\bar{h}_1 \to 0$, that is if the aerodynamic torque is weak, then A_β, $A_\gamma \to 0$. If $\bar{h}_1 \to \infty$, that is if the aerodynamic torque is dominant, then

$$A_\beta = \varepsilon \sin i, \quad A_\gamma = -\varepsilon \frac{(A+C-B)\sin i}{[4(B-C)-A]}. \tag{20.19}$$

If the gravity-gradient torque is deemed to be negligible so that $A = B = C$, then (20.19) has the following form

$$A_\beta = A_\gamma = \varepsilon \sin i$$

and $A_\beta = A_\gamma \approx 4^\circ$ is valid for a low polar orbit.

20.4 Equilibrium Positions and Conditions for Stability

Now we consider equilibrium positions of a SC in circular orbit without taking the Earth's spin into account. Rewrite the equations of motion (20.13) in the form

$$\begin{aligned} A\Omega_1' + (C-B)\Omega_2\Omega_3 &= 3(C-B)a_{32}a_{33} + \bar{h}_2 a_{13} - \bar{h}_3 a_{12}, \\ B\Omega_2' + (A-C)\Omega_1\Omega_3 &= 3(A-C)a_{31}a_{33} + \bar{h}_3 a_{11} - \bar{h}_1 a_{13}, \\ C\Omega_3' + (B-A)\Omega_1\Omega_2 &= 3(B-A)a_{31}a_{32} + \bar{h}_1 a_{12} - \bar{h}_2 a_{11}. \end{aligned} \tag{20.20}$$

The generalized integral of energy for a SC with an aerodynamic stabilizer, has the form

$$\begin{aligned} &\tfrac{1}{2}(A\bar{\Omega}_1^2 + B\bar{\Omega}_2^2 + C\bar{\Omega}_3^2) + \tfrac{3}{2}(Aa_{31}^2 + Ba_{32}^2 + Ca_{33}^2) \\ &-\tfrac{1}{2}(Aa_{21}^2 + Ba_{22}^2 + Ca_{23}^2) - (\bar{h}_1 a_{11} + \bar{h}_2 a_{12} + \bar{h}_3 a_{13}) = h_0. \end{aligned} \tag{20.21}$$

The approach adopted to find equilibrium positions of a SC under the effect of an aerodynamic torque, is similar to the approach already used to obtain equilibrium positions of a SC under a gravity-gradient torque (Section 16.2). The associated procedure is, however, more complicated, due to the fact that three aerodynamic parameters are additionally contained in (20.20). Integral (20.21) can be used to construct a Lyapunov function to derive the necessary conditions for stability.

Consider as before the special case where the center of pressure is situated on the axis Ox_1, so that $\overline{h}_2 = \overline{h}_3 = 0$ and $\overline{h}_1 = \text{const}$. Then, equations (20.20) and the expression for the integral (20.21) have the respective forms

$$\begin{aligned} A\Omega_1' + (C-B)\Omega_2\Omega_3 &= 3(C-B)a_{32}a_{33}, \\ B\Omega_2' + (A-C)\Omega_1\Omega_3 &= 3(A-C)a_{31}a_{33} - \overline{h}_1 a_{13} \\ C\Omega_3' + (B-A)\Omega_1\Omega_2 &= 3(B-A)a_{31}a_{32} + \overline{h}_1 a_{12} \end{aligned} \tag{20.22}$$

and

$$\begin{aligned} &\tfrac{1}{2}(A\overline{\Omega}_1^2 + B\overline{\Omega}_2^2 + C\overline{\Omega}_3^2) + \tfrac{3}{2}(Aa_{31}^2 + Ba_{32}^2 + Ca_{33}^2) \\ &\quad -\tfrac{1}{2}(Aa_{21}^2 + Ba_{22}^2 + Ca_{23}^2) - \overline{h}_1 a_{11} = h_0. \end{aligned} \tag{20.23}$$

The zero-approximation of α_0, β_0, γ_0 considered in our previous analysis (Section 20.3) corresponds to the absence of the effect of the Earth's spin and, consequently, it provides the presently sought equilibrium position of the SC. Let us now investigate the stability of the trivial equilibrium position described by

$$\alpha = \beta = \gamma = 0 \tag{20.24}$$

using Lyapunov's theorem for stability .

Substitute a_{33}^2 and a_{22}^2 from the trivial integrals $a_{31}^2 + a_{32}^2 + a_{33}^2 = 1$ and $a_{21}^2 + a_{22}^2 + a_{23}^2 = 1$ in (20.23). Multiply both sides of the resulting equality by two. Add the zero-value identity $\overline{h}_1(a_{11}^2 + a_{21}^2 + a_{31}^2 - 1)$ to the left side. Following transformation and on collecting the factors which completely define the perfect square term $(1 - a_{11})^2$, we obtain the first integral of (20.22) in the form

$$\begin{aligned} &A\overline{\Omega}_1^2 + B\overline{\Omega}_2^2 + C\overline{\Omega}_3^2 + [3(A-C) + \overline{h}_1]a_{31}^2 + (B-C)a_{32}^2 \\ &+(B - A + \overline{h}_1)a_{21}^2 + (B-C)a_{23}^2 + \overline{h}_1(1-a_{11})^2 = \overline{h}_0, \end{aligned} \tag{20.25}$$

that can be used as a Lyapunov function for the analysis of the stability of the trivial solution (20.24). Under the following conditions

$$3(A-C)+\bar{h}_1>0, \quad (B-C)>0,$$
$$B-A+\bar{h}_1>0, \quad \bar{h}_1>0 \tag{20.26}$$

the function (20.25) is positive-definite and satisfies the conditions of Lyapunov's theorem of stability. Hence, (20.26) defines sufficient conditions for the stability of the trivial angular position of a SC with a stabilizer. The positive sign of $\bar{h}_1$ required by (20.26), physically corresponds to the location of the center of pressure behind the center of mass with respect to the direction of SC translational motion. To derive the necessary conditions for stability, the method used in Section 16.4 can again be applied.

Let the condition $\bar{h}_2=0$ be satisfied. Then, the equations (20.22) admit the special solution

$$\alpha\neq 0, \quad \beta=\gamma=0,$$

which describes planar motion when α satisfies the equation

$$B\alpha''+3(A-C)\sin\alpha\cos\alpha+\bar{h}_1\sin\alpha-\bar{h}_3\cos\alpha=0. \tag{20.27}$$

The planar solution $\alpha=\text{const}$ can be obtained from the condition

$$3(A-C)\sin\alpha\cos\alpha+\bar{h}_1\sin\alpha-\bar{h}_3\cos\alpha=0$$

which describes planar equilibrium positions of a SC when the axes of the BRS do not coincide with the axes of the ORS (if $\bar{h}_1\neq 0$ and $\bar{h}_3\neq 0$ then there are no solutions $\alpha=\pi/2n$ where $n=0,1,2,\ldots$). We call these solutions the *non-trivial* or *inclined* equilibrium positions. Now we consider another way to produce such inclined positions.

Coming back to the case $\bar{h}_2=\bar{h}_3=0$, we rewrite equation (20.27) in the form

$$B\alpha''+3(A-C)\sin\alpha\cos\alpha+\bar{h}_1\sin\alpha=0. \tag{20.28}$$

The condition for the existence of an equilibrium position

$$3(A-C)\sin\alpha\cos\alpha+\bar{h}_1\sin\alpha=0$$

has two different solutions, $\alpha = 0$ and $\alpha = \pi$, at arbitrary values of the SC parameters, and the solutions

$$\alpha = \pm \arccos\left[-\frac{\bar{h}_1}{3(A-C)}\right], \tag{20.29}$$

if

$$\left|\frac{\bar{h}_1}{3(A-C)}\right| \leq 1. \tag{20.30}$$

Values of SC parameters at which the inequality reduces to an equality, determine the bifurcation point. At this point, a pair of non-trivial solutions (20.29) appears. This is consequent on the attainment of a balance between the gravity-gradient and aerodynamic torques acting on the SC. The first integral

$$B(\alpha')^2 + 3(A-C)(\sin\alpha)^2 - 2\bar{h}_1 \cos\alpha = h_0 \tag{20.31}$$

of equation (20.28) can be used to analyze the stability of these equilibrium positions. This was already done above for the trivial equilibrium position in the spatial case.

Let us here obtain the necessary conditions for stability of the non-trivial solution (20.29). Linearize (20.28) in its vicinity

$$B\Delta\alpha'' + \left[\frac{\bar{h}_1^2 - 9(A-C)^2}{3(A-C)}\right]\Delta\alpha = 0 \tag{20.32}$$

where $\Delta\alpha$ represents a small variation of the variable α. The corresponding characteristic equation has the form

$$B\lambda^2 + \left[\frac{\bar{h}_1^2 - 9(A-C)^2}{3(A-C)}\right] = 0.$$

The condition

$$\frac{\bar{h}_1^2 - 9(A-C)^2}{3(A-C)B} > 0$$

that the roots of the characteristic equation are imaginary, is the necessary condition for stability.

This stability condition is, however, directly opposed to the condition (20.30) arising from the existence of the non-trivial solutions, except for values of the SC parameters satisfying the inequalities

$$-1 < \frac{\bar{h}_1}{3(A-C)} < 0.$$

These inequalities determine the existence of stable non-trivial equilibrium positions of the SC.

Investigation of the effect of the restoring torque over a long time interval using an averaging method showed [4] that the angular momentum vector precesses around the tangent to the orbit at perigee, where the magnitude of the drag force is a maximum. Also, that small nutational librations are superimposed on the precession of the angular momentum vector. Further that, the rotation of the atmosphere caused by the Earth's spin produces a weak variation in the precession velocity.

Chapter 21

SC Motion in the Geomagnetic Field

Let us now consider the simplest practical method of realizing passive SC orientation. This involves providing orientation along the vector **H** of the local *geomagnetic field strength.* In this case a restoring magnetic torque is produced through the interaction between a permanent magnet installed on-board the SC and the local geomagnetic field. The magnetic moment of the magnet is chosen to be sufficiently strong to allow the magnetic torque to govern the motion of the SC with respect to the vector **H**. Hysteresis rods fabricated from soft magnetic material can be used to produce a damping torque.

21.1 The Geomagnetic Field

To a first approximation, the Earth's magnetic field is that of a sphere uniformly magnetized in the direction of its, body centered, dipole axis. This axis cuts the surface of the Earth in the northern and southern hemispheres. The present 'best fit' between this Earth centered dipole and the actual magnetic field is obtained by taking the Southern Pole to be at approximately 78° *S*, 111° *E* (near the *Vostok* polar station, Antarctica) and at 78° *N*, 69° *W* (near Thule, Greenland). These points are called the Geomagnetic Poles. The angle of displacement between the Earth's spin axis and its magnetic axis is of the order of 11°.

The plane through the center of the Earth perpendicular to its

dipole axis is called the *dipole equatorial plane* and the circle in which this latter plane cuts the sphere is called the *dipole equator.* Dipole latitude is reckoned relative to the dipole equator. The semicircles joining the poles are termed *dipole meridians.*

Sources of the geomagnetic field are located in the Earth's deep interior and in localized magnetic anomalies near the surface of the Earth. Transient magnetic disturbances can, in addition, be caused by the motion of charged particles in near Earth space. The geomagnetic field beyond several Earth's radii has a different character from that of the 'near Earth field' and will be separately considered in the Addendum (A.1.15). We discuss here the near Earth field produced by the Earth's internal magnetic sources and by magnetic anomalies.

Usually the potential of the *internal geomagnetic field* U_m which varies inversely with distance from the center of the Earth to a point in space, is represented by an infinite series of spherical harmonics

$$U_m = R_e \sum_{n=1}^{\infty} \sum_{m=0}^{n} (I_n^m \cos m\lambda_R + i_n^m \sin m\lambda_R) \left(\frac{R_e}{R}\right)^{n+1} P_n^m(\cos\theta_R). \tag{21.1}$$

Here R_e is the mean equatorial radius of the Earth; R, λ_R, θ_R are the spherical coordinates of the point in space where the potential is calculated (R is the magnitude of the radius-vector of the point relative to the center of the Earth, λ_R is the geographic longitude of the point and θ_R is the angle between the radius-vector and the spin axis of the Earth which is called the *co-elevation*), $P_n^m(\cos\theta_R)$ is the *associated Legendre function of the first kind of degree n and order m*

$$P_n^m(x) = \frac{(1-x^2)^{m/2}}{2^n \cdot n!} \frac{d^{n+m}}{dx^{n+m}}(x^2-1)^n, \tag{21.2}$$

I_n^m, i_n^m are *Gaussian* coefficients (named after Gauss) that vary slowly with time in correspondence with the secular change in the internal field. The vector of induction of the geomagnetic field with potential (21.1) is determined by the expression

$$\mathbf{B} = -\nabla U_m. \tag{21.3}$$

Since the geomagnetic field varies with time, we may simplify its determination by assuming that the coefficients I_n^m, i_n^m of the series (21.1)

vary step-wise at the moment of passing from one year to another but do not vary during the course of a particular year. Another procedure leading to a more precise analysis, is to interpolate the coefficients using their derivative values. Updated values of the Gaussian coefficients are published every five years by the International Association of Geomagnetism and Aeronomy. This body also sanctions the publication of an International Geomagnetic Reference Field (IGRF), constructed using a combination of several models.

Another way of representing U_m is by the expression

$$U_m = R_e \sum_{n=1}^{\infty} \sum_{m=0}^{n} c_n^m \cos(m(\lambda_R - \lambda_n^m)) \left(\frac{R_e}{R}\right)^{n+1} P_n^m(\cos\theta) \qquad (21.4)$$

with

$$c_n^m = [(I_n^m)^2 + (i_n^m)^2]^{1/2};\ c_n^0 = I_n^0; \qquad (21.5)$$

and

$$\lambda_n^m = \frac{1}{m} \arctan \frac{i_n^m}{I_n^m},\ m \neq 0. \qquad (21.6)$$

Here λ_n^m represents phase angles (with respect to the reference direction to the Greenwich meridian which is denoted by $\lambda_R = 0$ in the equatorial plane) and c_n^m is a measure of the 'total strength' of each term of order (n, m). It is noted that, according to (21.6), the actual quadrant in which λ_n^m lies is determined by the signs of I_n^m and i_n^m. The cosine and sine of the angle λ_n^m are respectively taken to be proportional to I_n^m and i_n^m.

An interpretation of the different terms in the expansion (21.1) has been provided by N.A.Umov*, who showed that each spherical function represents the potential of a *multipole* with axes defined by the order of the function. An individual magnetic moment corresponds to each multipole. Umov, using a specific mathematical procedure, obtained the directions of the axes and the magnitudes of the magnetic moments for functions of the second and third degree employing the coefficients I_n^m and i_n^m (see Fig.21.1). This picture is taken from J.G.Roederer†, who gave a detailed interpretation of these and higher terms of the expansion (21.4) and showed that terms of order 4 and

*N.A.Umov. *Selected Papers.* Leningrad, 1950.

† J.G.Roederer, Geomagnetic Field Distortions and Their Effects on Radiation Belt Particles, *Review Geophys. and Space Phys.,* 1972, Vol.10, pp.599–630.

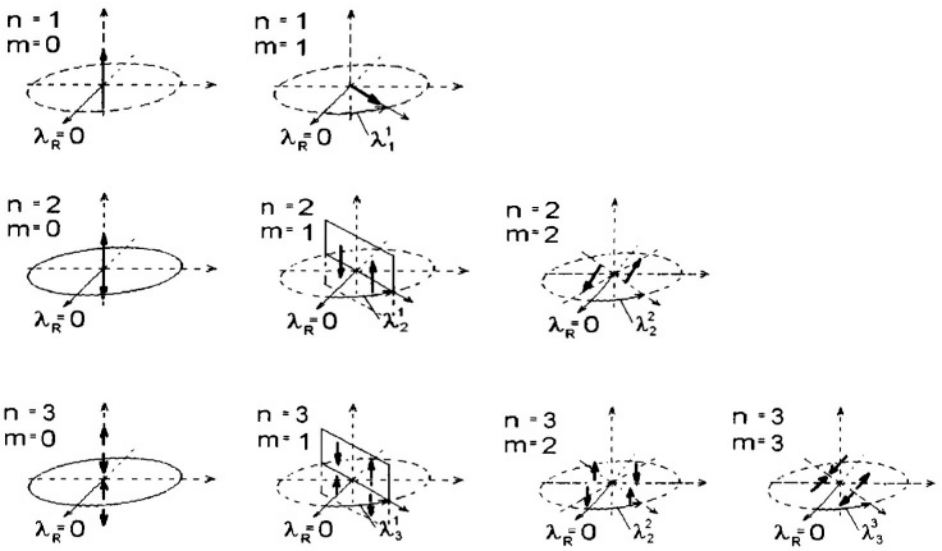

Figure 21.1. Equivalent dipole distributions corresponding to each (n, m) term in the geomagnetic potential expansion (21.4)

higher associate with terrestrial magnetic anomalies. Terms (4, 3) and (5, 4) are, in particular, related to the South African anomaly and (4, 3) and (6, 5) to the South American (Brazilian) anomaly. The multipole moments are proportional to the coefficients c_n^m given by (21.5). and the phase angles are given by (21.6).

21.2 Models of the Geomagnetic Field

Certain terms of the series (21.1) have clear physical interpretations. Let us consider now those terms used in the investigation of SC dynamics under the effect of the geomagnetic field.

21.2.1 Dipole Terms (n=1)

Direct Dipole Model

If we take into consideration the coefficient I_1^0 only $(m = 0)$, the field described by the series (21.1) is the field of a dipole situated at the center of the Earth and oriented antiparallel to its spin axis, i.e. ori-

ented in the direction North to South. This is the, so called, *direct dipole model.* The corresponding equivalent dipole distribution is represented in Fig.21.1 by the case $n = 1$, $m = 0$. Its usefulness lies in the fact that the equations of motion of a SC with regard to the interaction of its magnetic moment with the geomagnetic field are periodic in time. This allows us to investigate the equations of motion using well known techniques based on the theory of differential equations with periodic coefficients. Also, this model is able to take into account the two principle features of the behaviour of the local vector **H** of magnetic field strength during the motion of a SC along its orbit, namely, non-uniform rotation of **H** with respect to the inertial reference system and variations in its magnitude. For general investigations, periodic motion is a suitable approximation to real motion.

Let us represent the vector **H** by means of its projections on the axes of the orbital reference system within the framework of the direct dipole model

$$H_1 = H_0 \sin i \cos u, \quad H_2 = H_0 \cos i, \quad H_3 = -2H_0 \sin i \sin u,$$

where $H_0 = \mu_m/(4\pi R^3)$ is the strength of the geomagnetic field over the equator, μ_m is the Earth's magnetic moment and previously introduced notations for orbital elements are employed. In this model, the vector **H** varies, generally speaking, in both magnitude and direction, circumscribing a closed conical surface for each half-orbit of the SC in the reference system $OX'_{a1}X'_{a2}X'_{a3}$, the axes of which are parallel to the axes of the IRS O_aX_{a1}, O_aX_{a2}, O_aX_{a3} and the origin of which is located at the SC center of mass. For simplicity we consider an orbit with zero-angle of the ascending node.

The projections of **H** on the axes of $OX'_{a1}X'_{a2}X'_{a3}$ have the form

$$\begin{aligned} H_{a1} &= 3H_0 \sin i \cos i \sin^2 u, \\ H_{a2} &= -3H_0 \sin i \sin u \cos u, \\ H_{a3} &= H_0(1 - 3\sin^2 i \sin^2 u). \end{aligned}$$

The conical surface formed by the vector **H** normalized by H_0 during a half-revolution of the SC around the Earth along an orbit with inclination $i = \pi/3$ is shown in Fig.21.2 with respect to this reference system. This cone is symmetrical with respect to the plane $OX'_{a1}X'_{a3}$. The angle which is formed by **H** in this plane is

$$\theta_H = \pi/2 + \arctan \frac{3\sin^2 i - 1}{3 \sin i \cos i}.$$

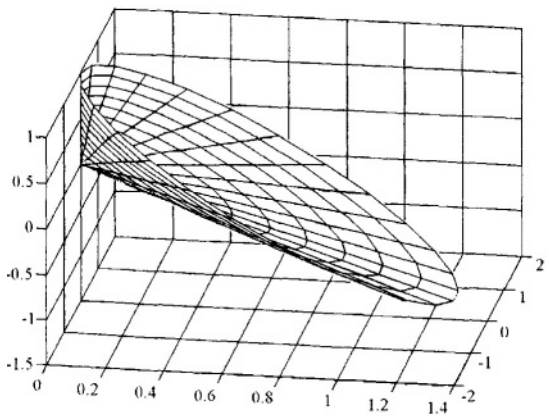

Figure 21.2. The cone formed by vector **H** in $OX'_{a1}X'_{a2}X'_{a3}$ during a half-revolution along an orbit with inclination $i = \pi/3$

It is convenient to represent the motion of vector **H** with respect to the reference system $OX''_{a1}X''_{a2}X''_{a3}$, which is transformed from system $OX'_{a1}X'_{a2}X'_{a3}$, by turning it counterclockwise through an angle Θ_H about a direct line passing through the point O aligned parallel to the line of nodes. The angle Θ_H is determined by the transformed relation

$$\tan\Theta_H = \frac{3\sin 2i}{2(1 - 3\sin^2 i + \sqrt{1 + 3\sin^2 i})}.$$

Calculations show that the cone circumscribed by the vector **H** differs slightly from a circular cone with vertex angle $2\Theta_H$. For an equatorial orbit, the cone circumscribed by vector **H** degenerates into a straight line, which is parallel to the Earth's spin axis. In polar orbit, the cone develops in the orbital plane. For intermediate cases the cone considered lies outside the circular cone but touches it at $u = \pi n/2$, $(n = 1, 2, 3, \ldots)$. The maximum angle between the vector **H** and the nearest generator of the circular cone during half a revolution, is achieved for an orbit with inclination $i \approx 52.1^\circ$ and this angle is $\approx 1.18^\circ$ [6].

Inclined Dipole Model

The term (1, 1) for which $m = 1, n = 1$ (see Fig. 21.1) represents a dipole of moment c_1^1 lying along a direction which makes an angle λ_1^1 in the plane of the geographic equator with the direction to the Greenwich meridian. Coefficients of the potential expansion (21.4) corresponding to the IGRF for the years 2000 – 2005 [31] yield a value for $\lambda_1^1 = 108.4°$ and for $c_1^1 = 5466.3\,nT$. The full $n = 1$ case can thus be interpreted to represent a single, Earth-centered, inclined dipole. The angle between the Earth's spin axis and this dipole is $\approx 169.5°$. The line defined by this Earth-centered dipole vector is the *geomagnetic axis.* Although this is directed southward, it is conventional to use the geographic longitude λ_0 of the northern extension of the geomagnetic axis to define its azimuthal direction. Thus, $\lambda_0 = \lambda_1^1 + 180° = 288.4°\,E$ geographic longitude. The meridian defined by λ_0 is called the *magnetic meridian.*

Thus, taking into account the first three terms of the series (21.1), we get the potential of a dipole inclined to the Earth's spin axis. These three terms make up what is called the *inclined dipole model*, which allows us to take into account the daily rotation of the Earth. Equations of motion obtained within the framework of this model contain terms incorporating the period of the Earth's spin rotation and the period of revolution of the SC around the Earth. It is, however, more complex to perform effective analysis using equations involving quasiperiodic coefficients than periodic ones.

The first there terms in the series (21.1) have the highest magnitudes. The next terms are interpreted to represent potentials produced by different multipoles. This means that the main part of the geomagnetic field has a dipole nature. As the distance from the Earth's surface increases, the magnitudes of the highest harmonics of the magnetic potential, which describe localized anomalies, decrease and the potential becomes closer to that of a dipole. For example, on the Earth's surface, the contribution of the dipole part is $\approx 87\%$ and at an altitude of 10 000 km it is $\approx 96\%$.

21.2.2 Quadrupole Terms $(n = 2)$

Consider now the quadrupole terms $n = 2$. The equivalent pairs of dipoles corresponding to the terms labeled (2, 0), (2, 1) and (2.2) are

shown in Fig.21.1. We can interpret the first-order effect of the (2,0) and (2, 1) terms to constitute a parallel displacement of the position of the Earth's dipole away from the center of the Earth (along and perpendicular to its spin rotation axis).

The secular change of the expansion coefficients causes a slow variation in the position of the magnetic center, which undergoes a westward drift while receding from the Earth's center at a rate of between 2 and 3 km/year. This latter motion means that the eccentricity of the main dipole is increasing with time.

Consider next the (2, 2) case. Since, (see Fig.21.1) the equivalent pair of dipoles lies in a plane perpendicular to the Earth's dipole, the (2, 2) terms cannot be transformed away. The principal first-order effect of the (2, 2) quadrupole is to deform the equatorial component of the magnetic field.

21.2.3 Octupole Terms $(n = 3)$

As shown in Fig.21.1, there are four possible equivalent dipole pair configurations corresponding to $n = 3$. The (3, 1) and (3, 3) terms can be shown to cause a, longitude dependent, inclination of the total field with, however, only a second-order variation in its intensity. Hence, except for a different azimuthal periodicity and phase, their effects are analogous to that of the (2, 2) quadrupole in that they produce warping of the equatorial component of the magnetic field.

The (3, 0) octupole term is responsible, on the other hand, for an azimuthally symmetric compression of the field on and near the equatorial surface and this is associated with a slow secular decrease in magnitude. The (3, 2) octupole causes the field to be compressed or expanded in alternating longitudinal sectors.

The effects of terms of higher order are generally considered collectively and can be similarly interpreted..

Averaged Model

Another model of the geomagnetic field, called the *averaged model*, was used in pioneering work by Zajac[‡]. Accordingly to this model,

[‡]E.E.Zajak. Some Simple Solutions Relating to Magnetic Attitude Control of Satellites. *Proc. of 4th US National Congress of Applied Mechanics, Berkeley,*

vector **H** has a permanent magnitude and uniformly rotates in inertial space with twice the orbital angular velocity ω_0 of a SC (compare with the direct dipole model (see Section 21.2.1)). The magnitude could be defined, for instance, as the mean between maximum and minimum magnitudes of **H** or by the integrated mean magnitude over the orbit. In the framework of this model, equilibrium positions of the SC with respect to **H** exist[§], but these equilibrium positions do not correspond with what is observed in practice.

The general expression (21.1) is used in solving navigational problems by numerical methods when higher order terms, small in magnitude but permanently acting, generate errors in the determination of the orbit and attitude of a SC. In the phase of designing attitude control systems with magnetic elements when a preliminary analysis of SC dynamics is initiated, simple models of the geomagnetic field such as the direct dipole model, the inclined dipole model and even the averaged model are usually used. This opens up opportunities for theoretical analysis which cannot be carried out using the expression (21.1) (i.e. in circumstances when a large number of terms needs to be taken into account). For instance for navigation and attitude determinations as well as for the advanced analysis of SC dynamics, 8×8 or even 10×10 terms are used.

21.3 Equations of Motion

Suppose a SC to be represented by a rigid body containing a permanent magnet. Let the center of mass of the SC move along a Keplerian orbit around the Earth. The Earth's gravitational field is central and Newtonian. The geomagnetic field is approximated by a direct dipole model. To write the equations of SC attitude motion we introduce, in addition to the ORS and BRS, a Cartesian *magnetic reference system* (MRS) $OZ_1Z_2Z_3$ which relates with the vector **H** of the local geomagnetic field at a point *O*. Unit vectors $\mathbf{q}_1$, $\mathbf{q}_2$ and $\mathbf{q}_3$ of the axes of this

CA, *June, 18-21, 1962*, Pergamon Press, 1962, Vol.1, pp.449–456.

[§] V.V.Beletsky and A.B.Novogrebel'sky. Existence of Stable Relative Equilibria of an Artificial Satellite in the Model Magnetic Field. *Astronomical J.*, 1973, Vol.50, Issue 2, pp.327–335.

system have the form

$$\mathbf{q}_1 = \frac{\mathbf{H}}{|\mathbf{H}|},\ \mathbf{q}_3 = \frac{\mathbf{q}_1 \times \mathbf{E}_2}{|\mathbf{q}_1 \times \mathbf{E}_2|},\ \mathbf{q}_2 = \mathbf{q}_3 \times \mathbf{q}_1,$$

where $\mathbf{E}_2$ is the unit vector of the axis OX_2 of the ORS.

The angular position of the BRS $Ox_1x_2x_3$ with respect to the MRS $OZ_1Z_2Z_3$ we denote by angles α, β, γ (Fig.21.3). The corresponding

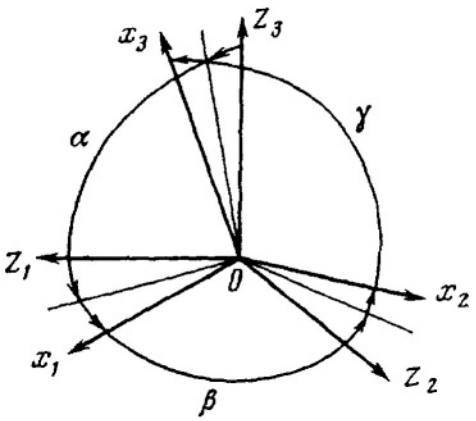

Figure 21.3. Transfer from MRS to BRS using plane angles α, β, γ

transformation matrix $\|b_{jn}\|$ has elements $b_{jn} = \cos(\mathbf{q}_j, \mathbf{e}_n)$, $(j, n = 1, 2, 3)$ determined by

$$\begin{aligned} b_{11} &= \cos\alpha\cos\beta \\ b_{21} &= \sin\beta \\ b_{31} &= -\sin\alpha\cos\beta \end{aligned}$$

$$\begin{aligned} b_{12} &= \sin\alpha\sin\gamma - \cos\alpha\sin\beta\cos\gamma & b_{13} &= \sin\alpha\cos\gamma + \cos\alpha\sin\beta\sin\gamma \\ b_{22} &= \cos\beta\cos\gamma & b_{23} &= -\cos\beta\sin\gamma \\ b_{32} &= \cos\alpha\sin\gamma + \sin\alpha\sin\beta\cos\gamma & b_{33} &= \cos\alpha\cos\gamma - \sin\alpha\sin\beta\sin\gamma \end{aligned}$$

Elements $h_{jn} = \cos(\mathbf{E}_j, \mathbf{q}_n)$, $(j, n = 1, 2, 3)$ of matrix $\|h_{jn}\|$ used to transfer from the system $OZ_1Z_2Z_3$ to the system $OX_1X_2X_3$ have the form

$$h_{11} = \frac{\sin i\cos u}{N},\quad h_{12} = -\frac{\cos i\cos u}{NN_1},\quad h_{13} = \frac{2\sin u}{N_1}$$

$$h_{21} = \frac{\cos i}{N},\quad h_{22} = \frac{N_1\sin i}{N},\quad h_{23} = 0,$$

$$h_{31} = -\frac{2\sin i\sin u}{N},\quad h_{32} = \frac{2\cos i\sin u}{NN_1},\quad h_{33} = \frac{\cos u}{N_1};$$

The angular velocity of the MRS with respect to inertial space is determined by the projections on the axes of the MRS

$$\overline{\Phi}_1 = \frac{3\omega_{orb}(1+\sin^2 u)\cos i}{NN_1^2}, \quad \overline{\Phi}_2 = \frac{3\omega_{orb}(1+\sin^2 u)\sin i}{NN_1},$$

$$\overline{\Phi}_3 = -\frac{3\omega_{orb}\sin 2u \sin 2i}{4N^2 N_1}$$

where $N = \sqrt{1+3\sin^2 i \sin^2 u}$ and $N_1 = \sqrt{1+3\sin^2 u}$. Here, to calculate the projections $\overline{\Phi}_1$, $\overline{\Phi}_2$, $\overline{\Phi}_3$ the following general expressions are used

$$\overline{\Phi}_1 = \omega_{orb}h_{21} + \sum_{k=1}^{3}\dot{h}_{k2}h_{k3}, \quad \overline{\Phi}_2 = \omega_{orb}h_{22} + \sum_{k=1}^{3}\dot{h}_{k3}h_{k1},$$

$$\overline{\Phi}_3 = \omega_{orb}h_{23} + \sum_{k=1}^{3}\dot{h}_{k1}h_{k2}$$

which are presented through derivatives of elements of the matrix $\|h_{jn}\|$.

Now we take into account the external magnetic restoring torque developed by the permanent magnet and the gravity-gradient torque. The equations of SC motion then have the form

$$\begin{aligned}
A\dot{\omega}_1 + (C-B)\omega_2\omega_3 &= \frac{3\mu_g}{R^3}(C-B)a_{32}a_{33} + \mu_0 H(m_{s2}b_{13} - m_{s3}b_{12}) \\
B\dot{\omega}_2 + (A-C)\omega_1\omega_3 &= \frac{3\mu_g}{R^3}(A-C)a_{31}a_{33} + \mu_0 H(m_{s3}b_{11} - m_{s1}b_{13}) \\
C\dot{\omega}_3 + (B-A)\omega_1\omega_2 &= \frac{3\mu_g}{R^3}(B-A)a_{31}a_{32} + \mu_0 H(m_{s1}b_{12} - m_{s2}b_{11}) \\
\dot{\gamma} &= \overline{w}_1 - (\overline{w}_2\cos\gamma - \overline{w}_3\sin\gamma)\tan\beta \\
\dot{\alpha} &= \frac{1}{\cos\beta}(\overline{\omega}_2\cos\gamma - \overline{\omega}_3\sin\gamma) \qquad (21.7) \\
\dot{\beta} &= (\overline{\omega}_2\sin\gamma + \overline{\omega}_3\cos\gamma)
\end{aligned}$$

where $\mathbf{m}_s = (m_{s1}, m_{s2}, m_{s3})_x$ is a magnetic moment of the SC, $H = |\mathbf{H}| = \mu_m N/(4\pi R^3)$; $\mathbf{H} = H(b_{11}, b_{12}, b_{13})_x$; as earlier, $\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_3)_x$ is the vector of absolute angular velocity of the SC; $\overline{\boldsymbol{\omega}} = (\overline{\omega}_1, \overline{\omega}_2, \overline{\omega}_3)_x$ is the vector of SC angular velocity with respect to the MRS and the following relationships for its projections are valid

$$\overline{\omega}_j = \omega_j - \sum_{k=1}^{3}\omega_{orb}\overline{\Phi}_k b_{kj}, \ (j = 1, 2, 3).$$

Using unitless parameters and variables

$$\lambda = \frac{A}{B},\ \mu = \frac{B-C}{A},\ \eta = \frac{m_{s1}\mu_0\mu_m}{4\pi B\mu_g},\ u = u_0 + \omega_0 t,$$
$$\overline{\Omega}_k = \frac{\omega_k}{\omega_0},\ \Phi_k = \frac{\overline{\Phi}_k}{\omega_0},\ (k = 1,2,3), \tag{21.8}$$

equations (21.7) can be written in the form

$$\begin{aligned}
\overline{\Omega}_1{}' &= \mu(\overline{\Omega}_2\overline{\Omega}_3 - 3a_{32}a_{33}),\\
\overline{\Omega}_2{}' &= \frac{1-\lambda}{1+\lambda\mu}(\overline{\Omega}_1\overline{\Omega}_3 - 3a_{31}a_{33}) - \eta N b_{13}, \qquad (21.9)\\
\overline{\Omega}_3{}' &= -(1-\lambda+\lambda\mu)(\overline{\Omega}_1\overline{\Omega}_2 - 3a_{31}a_{32}) + \eta(1+\lambda\mu)Nb_{12},\\
\gamma' &= \overline{\Omega}_1 - (\overline{\Omega}_2\cos\gamma - \overline{\Omega}_3\sin\gamma)\tan\beta - (\Phi_1\cos\alpha - \Phi_3\sin\alpha),\\
\alpha' &= \frac{1}{\cos\beta}(\overline{\Omega}_2\cos\gamma - \overline{\Omega}_3\sin\gamma) + (\Phi_1\cos\alpha - \Phi_3\sin\alpha) - \Phi_2,\\
\beta' &= (\overline{\Omega}_2\sin\gamma + \overline{\Omega}_3\cos\gamma) - (\Phi_1\sin\alpha + \Phi_3\cos\alpha)
\end{aligned}$$

where it is assumed that the magnetic moment of the permanent magnet is directed along the axis Ox_1, so that $m_{s2} = m_{s3} = 0$ and also that the orbit is a circular one.

Consider now the case of an axisymmetric SC with $B = C$. On introducing new variables

$$\Omega_1 = \overline{\Omega}_1,\ \Omega_2 = \overline{\Omega}_2\cos\gamma - \overline{\Omega}_3\sin\gamma,\ \Omega_3 = \overline{\Omega}_2\sin\gamma + \overline{\Omega}_3\cos\gamma,$$

we rewrite (21.9) in the form

$$\Omega_1' = 0,\ \gamma' = \Omega_1 - \Omega_2\tan\beta - \frac{\Pi_2}{\cos\beta}; \tag{21.10}$$

$$\begin{aligned}
\Omega_2' &= -\left(\lambda\Omega_1 - \Omega_2\tan\beta - \frac{\Pi_2}{\cos\beta}\right)\Omega_3 - \eta N\sin\alpha\\
&\quad -3(1-\lambda)(\Lambda_2\cos\beta + h_{32}),\\
\Omega_3' &= \left(\lambda\Omega_1 - \Omega_2\tan\beta - \frac{\Pi_2}{\cos\beta}\right)\Omega_2 - \eta N\cos\alpha\sin\beta \qquad (21.11)\\
&\quad +3(1-\lambda)(\Lambda_2\cos\beta + h_{32}\sin\beta)(-\Lambda_2\sin\beta + h_{32}\cos\beta),\\
\alpha' &= \frac{\Omega_2}{\cos\beta} + \Pi_2\tan\beta - \Phi_2,\ \beta' = \Omega_3 - \Pi_1,
\end{aligned}$$

where

$$\Pi_1 = \Phi_1\sin\alpha + \Phi_3\cos\alpha,\ \Pi_2 = \Phi_1\cos\alpha - \Phi_3\sin\alpha;$$
$$\Lambda_1 = h_{31}\sin\alpha + h_{33}\cos\alpha,\ \Lambda_2 = h_{31}\cos\alpha - h_{33}\sin\alpha.$$

Equations (21.11) describe the motion of the SC axis of symmetry. Equations (21.10) have first integral $\Omega_1 = \text{const}$ and describe SC motion around its axis of symmetry Ox_1. Value Ω_1 is contained in (21.11) as a parameter. The values Ω_2, Ω_3, α, β are assumed to be known from (21.11).

We will next investigate motions that can be adopted as nominal for a SC that is required to be uniaxially oriented along the vector **H**. First, consider the special case of planar motion of the SC axis of symmetry in a *polar* orbit.

21.4 Planar Motion of the SC Axis of Symmetry

For a polar orbit when $\Omega_1 = 0$, the following special solution of equations (21.10) and (21.11) exists

$$\Omega_2 = \Omega_2(u), \quad \alpha = \alpha(u), \quad \Omega_3 = 0, \ \beta = 0, \tag{21.12}$$

where $N = N_1$ and functions $\Omega_2(u)$, $\alpha(u)$ satisfy the following equations with π-periodic coefficients with respect to u

$$\Omega_2' = -\eta N_1 \sin\alpha - 3\frac{1-\lambda}{N_1^2}\left(\frac{1-5\sin^2 u}{2}\sin 2\alpha + \sin 2u \cos 2\alpha\right),$$
$$\alpha' = \Omega_2 - \frac{3(1+\sin^2 u)}{N_1^2}. \tag{21.13}$$

Solution (21.12) corresponds to motion of the SC axis of symmetry in the plane of a polar orbit. We again call this planar motion, in contrast to spatial motion which is characterized by the inequality $\Omega_3^2 + \sin^2\beta > 0$.

21.4.1 Asymptotic Solution

For the case where the magnetic restoring torque governs SC motion ($\eta \gg 1$), we now seek a π-periodic solution of (21.13) in the form of a formal series in negative degree powers of the parameter η

$$\Omega_2(u) = 3\frac{1+\sin^2 u}{N_1^2} + \sum_{k=1}^{\infty}\eta^{-k}\Omega_{2k}(u), \quad \alpha(u) = \sum_{k=1}^{\infty}\eta^{-k}\alpha_k(u). \tag{21.14}$$

The coefficients $\Omega_{2k}(u)$ and $\alpha_k(u)$ contained in (21.14) are determined from an infinite set of recurrent relationships, obtained by substituting the series (21.14) in (21.13) and then equating to zero expressions with the same degree of the parameter η. This set has a unique π-periodic solution satisfying the conditions of symmetry

$$\Omega_2(-u) = \Omega_2(u), \qquad \alpha(-u) = -\alpha(u). \tag{21.15}$$

By means of (21.15) the conditions

$$\Omega_2(\pi) - \Omega_2(0) = 0, \qquad \alpha(\pi) - \alpha(0) = 0$$

of π–periodicity of the solution of equations (21.13) are transformed to boundary conditions

$$\alpha(0) = \alpha(\pi/2) = 0. \tag{21.16}$$

By virtue of (21.13) invariance with respect to the transformation $u \to -u, \alpha \to -\alpha$, each solution of the boundary value problem (21.13), (21.16) satisfies the relations (21.15), as well as $-\alpha(-u + \pi/2) = \alpha(u + \pi/2)$ and $\Omega_2(-u + \pi/2) = -\Omega_2(u + \pi/2)$. Consequently, the solution is π–periodic.

21.4.2 Numerical Construction of Periodic Solutions

Numerical construction of the periodic solution (21.14) reduces to solving the boundary value problem (21.13), (21.16). For this purpose, the equation $\alpha(\pi/2) = 0$ is solved using *Newton's method*[¶] with respect to

[¶] We need to solve the vector equation

$$F(x, \eta) = 0 \tag{21.17}$$

with respect to x, where x is an unknown vector ($x \in R^n$), $F \in R^n$ and η is a parameter. Suppose that the approximate solution of (21.17) $x \approx x_0$ is known for a given η. Now expand the function F in a Teylor series in the vicinity of x_0

$$F(x_0 + \Delta x) = F(x_0) + \left(\frac{\partial F}{\partial x}\right)_{x=x_0} \Delta x + \mathrm{O}((\Delta x)^2) \tag{21.18}$$

where the symbol Δx denotes a small deviation from x_0. It is assumed that function F is continuous and may be differentiated a sufficient number of times. Neglect

$\Omega_2(0)$ at fixed η and the relationship $\alpha(0) = 0$. The value $\Omega_2(0) = \Omega_{20}$ is chosen as an initial approximation. The results of calculation for $\lambda = 1$ are shown in Fig.21.4 and in Fig.21.5. In these figures, initial

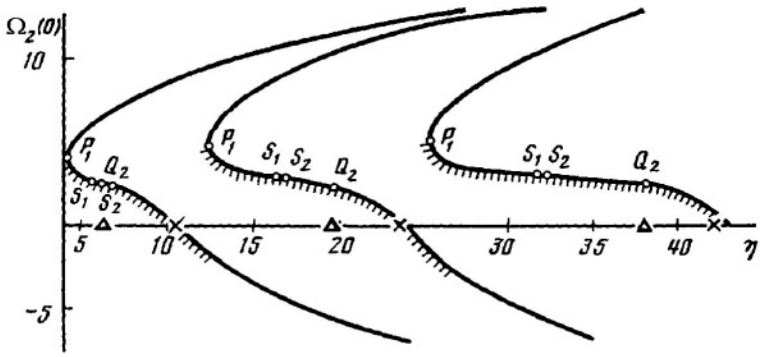

Figure 21.4. Initial conditions determining solutions of the boundary value problem (21.13), (21.16)

conditions $\Omega_2(0)$ and amplitudes $\mathrm{Am}_\alpha = \max\limits_{0\le u\le\pi} |\alpha(u)|$ are respectively plotted vs the magnetic parameter η in the range $4 \le \eta \le 45$. Each point on the plots in Fig.21.4 defines a π-periodic, planar, solution that satisfies the boundary conditions (21.16). An amplitude of a pe-

---

second-order and higher terms denoted in (21.18) by $O((\Delta x)^2)$ and substitute this in (21.17). Resolving this equation, we obtain an improved approximation x_1 to the solution x_0 of the equation (21.17)

$$x_1 = x_0 - \left(\frac{\partial F}{\partial x}\right)^{-1}_{x=x_0} F(x_0). \tag{21.19}$$

Using x_1 as the next approximation, we can employ this procedure again to increase the accuracy of the solution of (21.17). The fact of convergence and the rate of convergence depend on the initial approximation x_0 and on the properties of the function $F(x)$. This method is called Newton's method or the *Newton-Raphson* iteration method and it is a version of the *successive approximations method.*

This method can be used in the case where x_0 is a known solution of (21.17) for $\eta = \eta_0$ and we need to obtain its solution for $\eta_1 = \eta_0 + \Delta\eta$. Taking x_0 as a first approximation to the solution of the equation

$$F(x, \eta_1) = 0$$

and executing the Taylor expansion as in (21.18), we obtain formula (21.19) where $(\partial F/\partial x)^{-1}_{x=x_0}$ and $F(x_0)$ have to be calculated for $\eta = \eta_1$.

riodic solution describes maximum deviation of the axis of symmetry

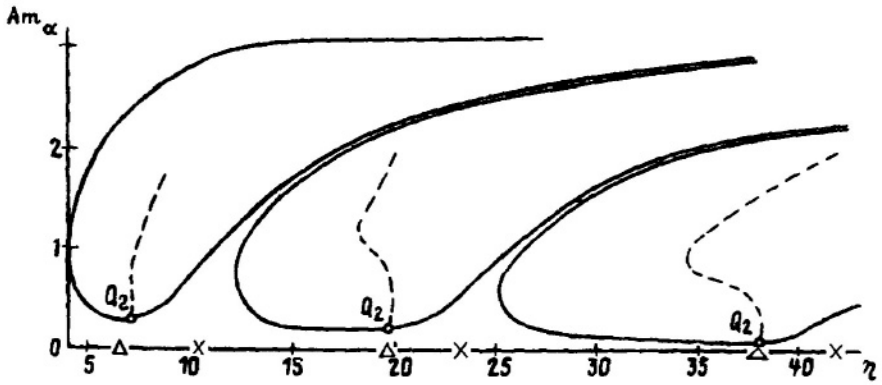

Figure 21.5. Amplitudes of solutions of the boundary value problem (21.13), (21.16)

of the SC from vector **H** over the period π. As η increases further, we obtain new curves and, consequently, new solutions. Roughly speaking, the difference between the solutions defined by different curves (called 'branches') is that the number of librations within the period π increases with η.

The plot in Fig.21.4 displays a number of separate branches. For initial conditions pertaining to the same 'relatively flat' part of a particular branch, the magnitude $J_p = \partial\alpha(\pi/2)/\partial\Omega_2(0)$ has the same sign. Relatively flat in this context means 'almost flat' and it is noted that the more distant a particular branch is from the origin, the closer this part of the curve approximates to being flat. After passing to the relatively flat part of a neighboring branch, the sign of J_p changes. The profiles of curves in the vicinity of points where J_p changes sign, that is, at *points of discontinuity*, were calculated using the method of parameter extension.

The relatively flat parts of the curves between points of discontinuity correspond to the solution presented in the form of a formal series (21.14) at $\eta \gg 1$. The discontinuities between the relatively flat parts are induced by bifurcation of the boundary value problem (21.13), (21.16) caused by resonances between the orbital motion of the SC and natural librations of the SC axis of symmetry in the orbital plane. Such bifurcation reflects a divergence of the series (21.14) that justifies calling this a 'formal series'.

Bifurcation magnitudes of the parameter η, that is abscissa η_k of points on the curves at which the tangents are vertical (see Fig.21.4), are determined for $\eta \gg 1$ by the approximate formula

$$\eta_k \approx p_a k^2 - r_0 + r_1 \lambda \tag{21.20}$$

where

$$p_a = \frac{\pi^2}{a_a^2} \approx 2.63,\ a_a = \int_0^{\pi/2} \sqrt{N_1}\,du,\ r_1 = -\frac{3}{a_a}\int_0^{\pi/2} \frac{1-5\sin^2 u}{N_1^{5/2}}\,du \approx 0.51,$$

$$r_0 = -\frac{1}{a_a}\int_0^{\pi/2}\left[\frac{1}{16N_1}(N_1')^2 + 3\frac{1-5\sin^2 u}{N_1^2}\right]\frac{du}{\sqrt{N_1}} \approx -0.49.$$

A few points determined using (21.20) are denoted by the symbol $\times$ on the abscissa-axis (Fig.21.4). The way to obtain the asymptotic formula (21.20) for determining the abscissae of the bifurcation points of periodic solutions is as follows. At $\eta \gg 1$ the equations (21.13) satisfy the conditions of the theorem for the existence of periodic solutions of second order differential equations involving a large parameter[||]. In accordance with this theorem, (21.13) has a unique π-periodic solution (21.14) satisfying the condition $\alpha(u) = \mathrm{O}(1/\eta)$ for all sufficiently large values of the parameter η, except in the vicinity of the resonant points η_k. Let us find these points.

The equations in deviations for the solution (21.14) have the form

$$\Delta\Omega_2' = \left[-\eta N_1 + 3\frac{1-\lambda}{N_1}(1-5\sin^2 u) + \mathrm{O}(1/\eta)\right]\Delta\alpha,$$
$$\Delta\alpha' = \Delta\Omega_2. \tag{21.21}$$

The values of parameter η, at which (21.21) with boundary conditions

$$\Delta\Omega_2(0) = \Delta\Omega_2(\pi), \qquad \Delta\alpha(0) = \Delta\alpha(\pi)$$

has a nontrivial solution, are the bifurcation ones.

[||]B.M.Levitan. *Expansion by Natural Functions.* Moscow, GITTT Publ., 1950.

Reducing (21.21) to a single equation and performing variable substitutions, called *Liouville's substitutions*[**], using formulae

$$\Delta\alpha = \frac{x}{N_1^{1/4}}, \qquad \tau = \int_0^u \sqrt{N_1(s)}ds$$

we get the equation and boundary conditions corresponding to (21.15)

$$\frac{d^2x}{d\tau^2} + \eta x = q_a x, \qquad x(0) = x(a_a) = 0. \tag{21.22}$$

Here

$$a_a = \int_0^{\pi/2} \sqrt{N_1}du,$$

$$q_a = \left[\frac{1-\lambda}{N_1^2}(1 - 5\sin^2 u) + \frac{1}{2N_1^{3/2}}\frac{d\sqrt{N_1}}{du} - \frac{3}{4N_1^2}\left(\frac{d\sqrt{N_1}}{du}\right)^2\right]\Bigg|_{u=u(\tau)}$$

and terms $O(1/\eta)$ are truncated. This is a problem of eigenvalues for a Schturm[††]-Liouville operator, with boundary values corresponding to the periodicity. The asymptotic formula (21.20) for the eigenvalues η_k obtained by the *method of successive approximations*[‡‡] is valid to an accuracy of $O(1/k)$.

[**]Liouville, Joseph (1809–1882). French mathematician. Member of the Paris Academy of Sciences. His main work was devoted to mathematical analysis. Also, he made contributions to the theory of elliptic functions and investigated the boundary value problem for linear differential equations of the second order (the Schturm-Liouville problem).

[††]Schturm, Jacques Charles François (1803–1855). French mathematician. Member of the Paris Academy of Sciences. He studied the boundary value problem for differential equations with applications in mathematical physics, as well as problems concerning eigenvalues and eigenfunctions.

[‡‡]In the framework of the method of successive approximations, a solution of equation (21.22) can be sought in the form

$$x(\tau) = c_1\cos\sqrt{\eta}\tau + c_2\sin\sqrt{\eta}\tau + \frac{1}{\sqrt{\eta}}\int_0^{\pi}\sin\sqrt{\eta}(\tau - s)q_a(s)x(s)ds, \tag{21.23}$$

where c_1 and c_2 are constants. The equality $c_1 = 0$ follows from (21.23) and

21.4.3 Investigation of Stability

The stability of solution (21.12) is next investigated in linear approximation, using the linearized equations

$$\Delta\Omega_2' = -[\eta N_1 \cos\alpha - 3(1-\lambda)(\tilde{\Lambda}_2 - \tilde{\Lambda}_1)]\Delta\alpha,\ \Delta\alpha' = \Delta\Omega_2; \tag{21.27}$$

$$\Delta\Omega_3' = -[\eta N_1 \cos\alpha + \Omega_2^2 - 3(1-\lambda)\tilde{\Lambda}_2^2)]\Delta\beta,\ \Delta\beta' = \Delta\Omega_3 \tag{21.28}$$

where symbol Δ denotes the deviation of some associated variable; $\tilde{\Lambda}_1$ and $\tilde{\Lambda}_2$ are calculated at $i = \pi/2$. These equations in deviations are

(21.22) so we can write

$$x(\tau) = c_2 \sin\sqrt{\eta}\tau + O(\frac{1}{\sqrt{\eta}}). \tag{21.24}$$

On using $x(a_a) = 0$ we obtain

$$x(a_a) \equiv c_2 \sin\sqrt{\eta}a_a + \frac{1}{\sqrt{\eta}} \sin\sqrt{\eta}a_a \int_0^{a_a} (\cos\sqrt{\eta}s) q_a(s) x(s) ds -$$

$$\frac{1}{\sqrt{\eta}} \cos\sqrt{\eta}a_a \int_0^{a_a} (\sin\sqrt{\eta}s) q_a(s) x(s) ds = 0. \tag{21.25}$$

On substituting (21.24) into (21.25) we get

$$\sin\sqrt{\eta}a_a[1 + O(\frac{1}{\sqrt{\eta}})] - \frac{1}{\sqrt{\eta}} \cos\sqrt{\eta}a_a[J_a + O(\frac{1}{\sqrt{\eta}})] = 0, \tag{21.26}$$

where $J_a = \int_0^{a_a} q_a(s) ds/2$ and, without losing generality, we can assume that $c_2 = 1$. For large $\sqrt{\eta}$ equation (21.26) clearly has a solution. The roots $\sqrt{\eta}a$ lie near points corresponding to the numbers πk, where k is an integer, that is $\sqrt{\eta}a = \pi k + \delta_k$ where δ_k is small. Substituting this expression in (21.26) we determine $\delta_k \approx J_a/\sqrt{\eta}$. Now the relationship

$$a\eta - \pi k\sqrt{\eta} - J_a \approx 0$$

can be written. Resolving this with respect to η we get

$$\eta = \frac{\pi^2 k^2}{a^2} + \frac{2J_a}{a} + O(\frac{1}{k}).$$

This expression determines the abscissae of points of bifurcation.

obtained through linearization of (21.11) in the vicinity of the planar solution (21.12). They are separable into two independent sets of equations. Here, equations (21.27) describe the motion of the SC axis of symmetry perturbed in the plane of the orbit and equations (21.28) describe the motion of the axis perturbed in the plane perpendicular to the orbit. Also equations (21.27) and (21.28) can be used to calculate the natural frequencies and amplitudes of small librations as functions of the initial conditions of motion and of the SC parameters.

The characteristic equations for (21.27) and (21.28) each have a similar form

$$\xi^2 - 2A_k\xi + 1 = 0 \qquad (k = 1, 2). \tag{21.29}$$

The coefficients A_1 and A_2 are determined through obtaining solutions of (21.27) and (21.28). If $|A_1| \leq 1$ and $|A_2| \leq 1$, then all the roots of (21.29) are situated on a circle $|\xi| = 1$, and necessary conditions for the stability of the investigated periodic solution are fulfilled. For brevity we say that planar stability is present if $|A_1| \leq 1$ and spatial stability if $|A_2| \leq 1$. If both conditions are fulfilled together, we say that the solution is stable. Graphs of coefficients A_1 and A_2 vs the parameter η are plotted in Fig.21.6 Segments of curves where the

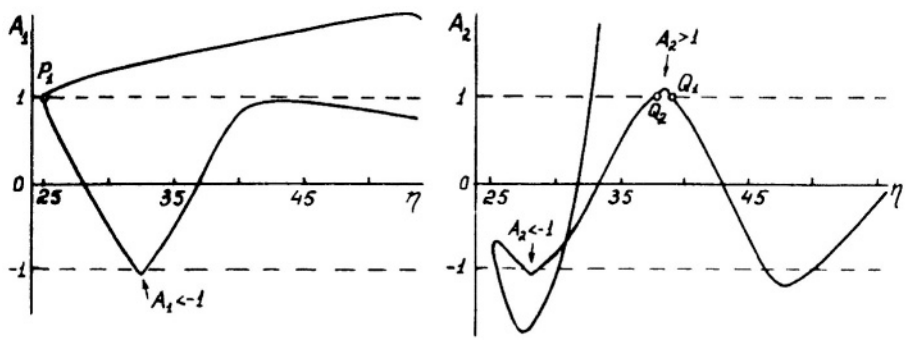

Figure 21.6. Coefficients A_1 and A_2 which determine planar and spatial stability of the planar π-periodic solution of equations (21.13) vs. η

necessary conditions for stability are satisfied are shown hatched in Fig. 21.4.

At $A_1 = -1$ equations (21.27) have a nontrivial 2π-periodic solution. If $\operatorname{rank}\|X(\pi) + E\| \neq 0$, where $X(\pi)$ is a *fundamental matrix* of (21.27), then this solution is unique up to a constant factor. The matrix $X(\pi)$ is constructed from columns, each of which is a solution

of the initial problem for period π. The initial conditions for each problem are given by consecutive columns (for the considered case these are $(1,0)^\top$ and $(0,1)^\top$) of the identity matrix E. Analysis of the fundamental matrix elements shows that, at the point S_1 (Fig.21.7) on one boundary of the interval of stability where $A_1 = -1$, the equations (21.29) have a nontrivial solution satisfying the conditions of

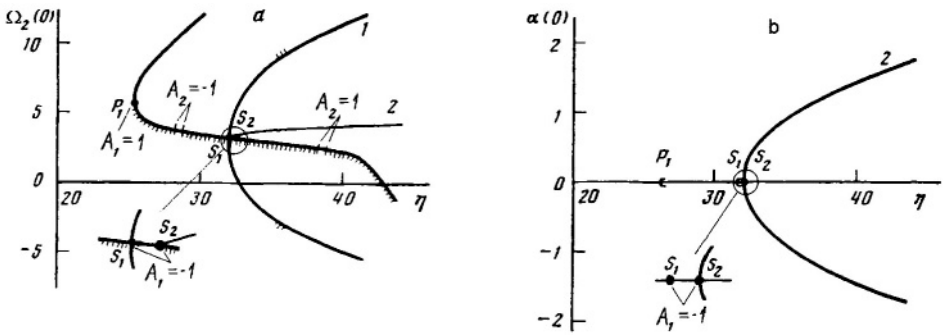

Figure 21.7. Initial conditions of π- and 2π-periodic solutions of (21.13)

symmetry

$$\Delta\Omega_2(-u) = \Delta\Omega_2(u), \qquad \Delta\alpha(-u) = -\Delta\alpha(u). \tag{21.30}$$

At the point S_2 on the other boundary of the interval of stability where $A_1 = -1$, a nontrivial solution of (21.29) satisfies the conditions

$$\Delta\Omega_2(-u) = -\Delta\Omega_2(u), \qquad \Delta\alpha(-u) = \Delta\alpha(u). \tag{21.31}$$

At the point S_1 the 2π–periodic solution of (21.13) emerges from the π–periodic solution (21.14). This solution satisfies the conditions (21.15) (i.e. it consists of an odd $\alpha(u)$ and an even $\Omega_2(u)$), and consequently, satisfies the boundary conditions

$$\alpha(0) = \alpha(\pi) = 0. \tag{21.32}$$

The initial conditions of the 2π-periodic solution are represented by curve *1* in Fig.21.7(*a*).

At the point S_2, the nonsymmetric 2π-periodic solution of (21.13) satisfying the boundary conditions

$$\Omega_2(2\pi) - \Omega_2(0) = 0, \qquad \alpha(2\pi) - \alpha(0) = 0 \tag{21.33}$$

appears. Initial conditions of this solution are represented by curves *2* in Fig.21.7. The plot of the initial conditions for the π-periodic solutions is taken from Fig.21.4. Stable solutions are indicated by hatching.

21.5 Spatial Motion of the SC Axis of Symmetry

The terms of (21.11) have period π with respect to u and obtaining the spatial π-periodic solutions of these equations for the general case is similar to solving a boundary value problem under the following boundary conditions

$$\Omega_2(\pi) - \Omega_2(0) = 0, \quad \Omega_3(\pi) - \Omega_3(0) = 0,$$
$$\alpha(\pi) - \alpha(0) = 0, \quad \beta(\pi) - \beta(0) = 0. \tag{21.34}$$

At $i = \pi/2$, $\Omega_1 = 0$ the boundary value problem (21.11), (21.34) has a planar solution (21.12), for which the coefficients A_1 and A_2 in equations (21.29) are not equal to unity. In this case, by Poincare's theorem (Appendix C), for small enough $|\Omega_1|$, $|i-\pi/2|$, the boundary value problem (21.11), (21.34) has a unique solution which analytically depends on W_1 and i. This solution passes, at $i = \pi/2$, $\Omega_1 = 0$, through the planar solution. Using the invariance of (21.11) with respect to the variable substitution

$$u \to -u, \ \Omega_3 \to -\Omega_3, \ \alpha \to -\alpha, \tag{21.35}$$

we reduce the boundary conditions (21.34) to

$$\Omega_3(0) = \Omega_3(\pi/2) = \alpha(0) = \alpha(\pi/2) = 0. \tag{21.36}$$

Any solution of the boundary value problem (21.11), (21.36) is π-periodic and satisfies (21.35). The solution obtained by solving the boundary value problem (21.11), (21.36) has a larger domain of definition than the solution of the boundary value problem (21.11), (21.34). This is because the solution of (21.11), (21.36) for sufficiently small $|\Omega_1|$, $|i - \pi/2|$ and for the generating, planar, π-periodic solution

(21.14) is defined everywhere, with the exception only of points P_1 and Q_2 (Fig.21.4). The abscissa of point Q_2 is determined by

$$\eta_k \approx p_a n^2 - r_0 + r_1 \lambda, \tag{21.37}$$

which is similar to the expression (21.20) for the planar case, with the difference that the values r_1 and r_0 pertinent here are calculated using

$$r_1 = -\frac{12}{a_a} \int_0^{\pi/2} \frac{\sin^2 u}{N_1^{5/2}} du \approx 1.25,$$

$$r_0 = -\frac{1}{a_a} \int_0^{\pi/2} \left[\frac{1}{16 N_1} (N_1')^2 - 12 \frac{\sin^2 u}{N_1^2} - 1 \right] \frac{du}{\sqrt{N_1}} \approx -4.25.$$

A number of points obtained using (21.37) are designated by the symbol 'Δ' along the abscissa-axis in Fig.21.4 and Fig.21.5.

If $\eta \gg 1$ then the spatial π-periodic solution of (21.9) can be constructed as a formal series

$$\begin{aligned} \Omega_2(u, \Omega_1, i) &= \Phi_2 + \sum_{k=1}^{\infty} \eta^{-k} \Omega_{2k}(u, \Omega_1, i), \\ \Omega_3(u, \Omega_1, i) &= \Phi_3 + \sum_{k=1}^{\infty} \eta^{-k} \Omega_{3k}(u, \Omega_1, i), \\ \alpha(u, \Omega_1, i) &= \sum_{k=1}^{\infty} \alpha_k(u, \Omega_1, i), \quad \beta(u, \Omega_1, i) = \sum_{k=1}^{\infty} \beta_k(u, \Omega_1, i) \end{aligned} \tag{21.38}$$

similar to (21.14). Owing to this method of construction, the solution when $i = \pi/2$, $\Omega_1 = 0$ transfers to the planar solution (21.16).

For numerical investigation of the spatial π–periodic solutions, the boundary value problem (21.11), (21.36) is reduced to the following equations

$$\begin{aligned} g_1(a_1, a_2, \eta, \Omega_1, i) &\equiv \alpha(\pi/2, a_1, a_2, \eta, \Omega_1, i) = 0, \\ g_2(a_1, a_2, \eta, \Omega_1, i) &\equiv \Omega_3(\pi/2, a_1, a_2, \eta, \Omega_1, i) = 0 \end{aligned} \tag{21.39}$$

where $a_1 \equiv \beta(0)$ and $a_2 \equiv \Omega_2(0)$. For given magnitudes of η, Ω_1, i, these equations form a closed set of two equations with respect to the unknown a_1 and a_2. Equations (21.39) describe a two-dimensional

surface in the space of the five parameters a_1, a_2, η, Ω_1, i. Numerical investigation of this surface reduces to intersecting it by a three-dimensional hyperplane $i = \pi/2$, $\Omega_1 = 0$. Plots of the functions a_1 and a_2 against η, which represent symmetrical, spatial π-periodic solutions are shown in Fig.21.8. The positive ray $a_1 = 0$, starting at P_1, in Fig.21.8(a) and the dashed curve in Fig.21.8(b), jointly represent the planar, π-periodic, solution, odd with respect to α, obtained in Section 21.4.2. A spatial solution extends the planar one at the point Q_2. Necessary conditions for stability are satisfied along the hatched intervals. Dashed lines in Fig.21.5 represent the amplitudes of particular spatial solutions.

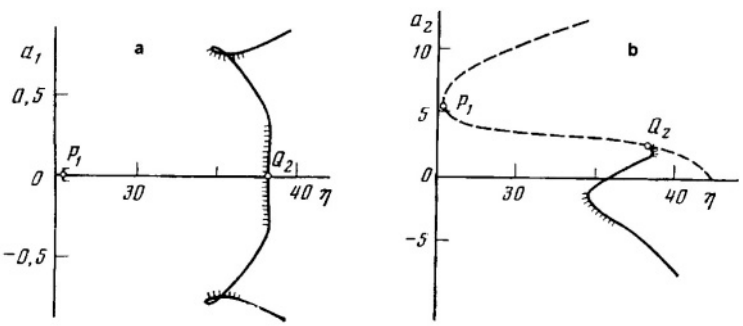

Figure 21.8. Initial conditions of the spatial periodic solutions at $i = \pi/2$ and $\Omega_1 = 0$

We now consider certain bifurcation curves along which new solutions emerge. Bifurcation here is determined by the two equations (21.39) and by

$$\frac{\partial(g_1, g_2)}{\partial(a_1, a_2)} = 0.$$

In the space of the five parameters $a_1, a_2, \Omega_1, i, \eta$, bifurcation can be represented by a two-dimensional hypersurface specified by these three equations. The intersection of this hypersurface by a hyperplane, for example $\Omega_1 = \text{const}$, results in bifurcation curves. Numerically obtained bifurcation curves in the plane (η, i) for $\Omega_1 = 0$ are shown in Fig.21.9. In the vicinity of these curves, the amplitude of periodic motion increases over the considered values of parameter η. The range chosen for this parameter can be examined typical for a SC provided

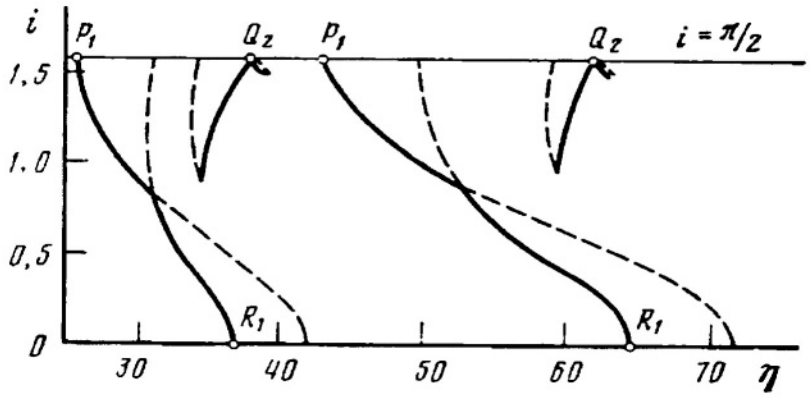

Figure 21.9. Bifurcation curves of the periodic solutions at $\Omega_1 = 0$

with a passive magnetic attitude control system. Solid lines show the bifurcation curves of solutions whose amplitudes are small at a fixed orbit inclination. Dashed lines show intervals of the bifurcation curves where the solutions have relatively large amplitudes. Horizontal line at $i = \pi/2$ corresponds to the polar orbit.

A trivial solution, which corresponds to the coincidence of the axis of orientation of the SC with vector **H** in an equatorial orbit, has the bifurcation points R_1, where non-trivial π-periodic solutions satisfying the boundary conditions (21.34) emerge (Fig.21.10).

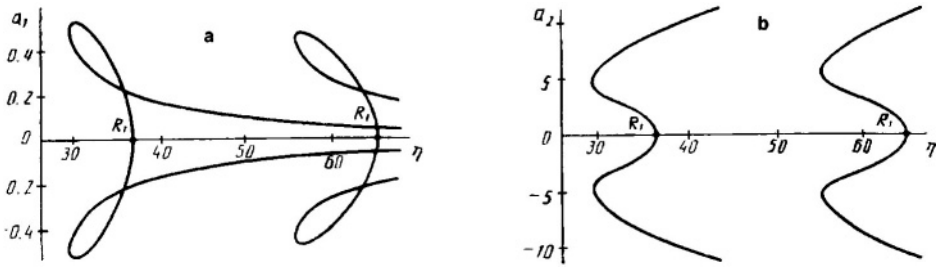

Figure 21.10. Initial conditions of the π-periodic solutions for an equatorial orbit

Bifurcation magnitudes of the parameters η, Ω_1, i can be obtained in a similar way to that already used to determine the bifurcation points of planar and spatial solutions for a polar orbit with $\eta \gg 1$.

Write equations in deviations for solution (21.38) by linearizing (21.9)

$$\Delta\Omega_2' - [p + \mathrm{O}(\eta^{-1})]\Delta\Omega_3 + [\eta N + \mathrm{O}(1)]\Delta\alpha = 0,$$
$$\Delta\Omega_3' + [p + \mathrm{O}(\eta^{-1})]\Delta\Omega_2 + [\eta N + \mathrm{O}(1)]\Delta\beta = 0,$$
$$\Delta\alpha' = \Delta\Omega_2, \quad \Delta\beta' = \Delta\Omega_3 \tag{21.40}$$

where $p = 2\Phi_1 - \lambda\Omega_1$ and only terms, which on transfer to a 'fast' time $\tau_f = u\sqrt{\eta}$ scale have an order not higher than $\mathrm{O}(\eta^{-1/2})$, are retained. Consider for (21.40) the boundary value problem

$$\Delta\Omega_3(0) = \Delta\Omega_3(\pi/2) = \Delta\alpha(0) = \Delta\alpha(\pi/2) = 0. \tag{21.41}$$

Determine the values η, Ω_1, i which provide this problem with a non-trivial solution. Introducing a new complex variable $z = \Delta\alpha + j\Delta\beta$, $(j^2 = -1)$, we transform (21.40) to

$$z'' + jpz' + \eta N z = 0. \tag{21.42}$$

On substituting in this equation new variables

$$\tau = \int_0^u \sqrt{N(s)}ds, \; z = \frac{x}{N^{1/4}\exp j\overline{p}/2}, \; \overline{p} = \int_0^u \sqrt{p(s)}ds$$

equation (21.42) attains the form

$$\frac{d^2x}{d\tau^2} + \eta x = (q_1 + jq_2)x, \tag{21.43}$$

where

$$q_1 = \frac{(\sqrt{N})''}{N^{3/2}} - \frac{3((\sqrt{N})')^2}{N^2} - \frac{p^2}{4N}, \qquad q_2 = \frac{p'}{N}.$$

Equation (21.43) is solved by the method of successive approximations. The conditions under which a non-trivial solution exists with regard to the transformed boundary conditions (21.41) are as follows

$$4\sin\left(a_s\sqrt{\eta} + \tfrac{1}{2}\overline{p}(\pi/2)\right)\sin\left(a_s\sqrt{\eta} - \tfrac{1}{2}\overline{p}(\pi/2)\right) -$$
$$-\frac{1}{2\sqrt{\eta}}\overline{q}_1(\pi/2)\sin(a_s\sqrt{\eta}) = \mathrm{O}(\eta^{-1}) \tag{21.44}$$

where

$$a_s = \int_0^{\pi/2}\sqrt{N}du, \quad \overline{q}_1(\pi/2) = \int_0^{\pi/2} q_1 du.$$

Solutions of (21.44) have the form

$$\eta = \frac{\pi^2 k^2}{a_s^2} \pm \frac{\pi \overline{p}(\pi/2) k}{a_s^2} + \frac{\overline{p}^2(\pi/2)}{4a_s^2} + \frac{\overline{q}_1(\pi/2)}{4a_s^2} + \mathrm{O}(k^{-1}). \qquad (21.45)$$

Here k is integer. Coefficients a_s, $\overline{p}(\pi/2)$, $\overline{q}_1(\pi/2)$ depend on the orbit inclination but coefficient $\overline{p}(\pi/2)$ depends also on the product $\lambda\Omega_1$. This allows us to gain insight into the general behaviour of the bifurcation curves presented in Fig. 21.9 in the framework of the approximations made. A detailed analysis of the spatial periodic motion of a SC was performed by Sarychev and Ovchinnikov*.

*V.A.Sarychev and M.Yu.Ovchinnikov. Motion of a Satellite Equipped with a Permanent Magnet around Its Center of Mass. *Cosmic Research*, 1986, Vol.24, Issue 4, pp.527–543.

Chapter 22

Motion of a SC under Damping

In order to develop a passive attitude control system it is necessary to provide both restoring and damping torques so that a SC can, as required, either attain a steady-state orientation or undergo forced motion. Methods to develop a restoring torque through applying different kinds of external force have already been considered in Chapter 14 for passive systems. Damping torques may be provided based on various physical effects, for example magnetic *hysteresis* (which appears in soft magnetic materials when they are re-magnetized). For damping purposes, a soft magnetic material can be employed featuring: high initial relative magnetic permeability (of the order of several thousand, or even tens of thousands); a low coercive force (significantly less than the intensity of the Earth's 're-magnetizing' field) and non-saturation in this latter field. A damper based on the hysteresis effect might typically be in the shape of a rod with the ratio (length to diameter) in the range 100 – 300. This ratio is a key-factor in optimizing the time-response of an attitude control system.

A number of investigations in the field of passive magnetic orientation were performed by Fishell, Stopkuchen, Mager, Ninomiya and Kamuller in the 1960s and 1970s within the framework of small SC projects with a mass of a few tens of kilograms (*Transit-1B, Transit-2A, Azur, ESRO-1A, ESRO-1B, EXOS-A, Magion* etc). A renewed interest in this topic is generated by the development of nano-SC with a mass of a few kilograms. A list of nano-SC equipped with atti-

tude control systems is presented in Table 22.1*. This shows that the majority of these nano-SC were provided with a passive attitude control system. Among several competing technologies, magnetic attitude control systems (MACS) occupied a favored position due to their simplicity and reliability. In contrast, a passive gravity-gradient attitude control system (GGACS) is complicated by the need to deploy a boom given that, to fabricate such a device appropriately scaled, is a difficult technological problem. Also, small flywheels, which provide an accuracy of orientation that is an order of magnitude higher than that achieved by GGACS, require precise technology and dedicated on-board resources.

To develop a passive magnetic system employing minimum SC resources a careful study of SC dynamics, rather than of the fabrication of mechanical elements, is required. Despite the inherent simplicity of such attitude control systems, the mathematical models associated with them are complicated. Simulation of the dynamics of a SC equipped with a permanent magnet and hysteresis rods requires the initial solving of two basic problems. The first is to obtain solutions from which the nominal motion of the SC is chosen. The second is to analyze the transient motion of a SC under the damping effects of hysteresis rods and to determine the parameters of these rods. A detailed review of research concerning the dynamics of SC with passive MACS is given in [27].

As an example of the application of such damping, we consider next the transient motion of a magnetically stabilized SC provided with a strong permanent magnet installed along the SC axis of symmetry, and with hysteresis rods fixed in the SC structure.

22.1 Equations of Motion

Let the SC be approximated by an axisymmetric rigid body (so that the two equatorial principal moments of inertia are equal to each other). Suppose that a strong permanent magnet with magnetic moment $\mathbf{m}_s$ is located along the Ox_1 axis of the BRS, and that a single hysteresis rod with unit vector **e** is inclined with respect to the plane

*M.Ovchinnikov. Attitude Control Systems for Nanosatellites, *ZAMM Z. angew. Math. Mech.*. Berlin, 2001, Vol.81, Suppl.4, pp.1027–1030.

Table 22.1. Nano-SC equipped with various attitude control systems

ACS aboard particular Nano-SC				main SC parameters			
passive GGACS	passive MACS	active MACS	fly-wheels	date of launch	orbit (incl.[degrs] & perigee ×apogee [km])	mass [kg]	dimensions (ϕ – diameter) [cm]
	MagIon-1			1978	83 & 406 ×768	15	30×30×16
	PACSAT			1990	98.7 & 800 ×800	13.34	15×15×15
	DOVE			1990	98.7 &800×800	12.92	15×15×15
	WeberSat			1990	98.7 &800×800	16.03	15×15×15
	LUSat			1990	98.7 &800×800	13.76	15×15×15
	ITAMSat			1993	98.6 &799×823	11.2	15×15×15
	EYESat			1993	98.6 &799×823	11.8	15×15×15
	UNAMSat-B			1996	98 &	10.7	15×15×15
		TUBSat-N		1998	78.9 & 400 ×776	8.5	32×32×10.4
		TUBSat-N1		1998	78.9 & 400 ×771	3	32×32×3.4
ASUSat-1				2000	100.23 & 753 ×809.1	5.9	ϕ34.3×25.4
	JAK			2000	100.23 & 753.3 ×807.1	0.174	10.2×7.6×2.5
	Thelma			2000	100.23 & 753.3×807.1	0.438	20.3×7.6×2.5
	Louise			2000	100.23 & 753.3×807.1	0.438	19.9×7.6×2.3
	StenSat			2000	100.23 & 753.3×807.1	0.233	10.2×7.6×2.5
			SNAP-1	2000	78.9 & 400 ×771	2.5	32×32×3.4
	Munin			2000	98.22 & 707×1965	6	21×21×21.5
REFLECTOR				2001	98 & 1050×1050	7	30×30×140 (inc.boom)
		NAVGOLD		2001	51 & 400×400	14	ϕ50×41
	SAPPHIRE				> 37 & < 1000	18.8	ϕ44×27
	EMERALD			2003	> 37 & < 1000	13.89	ϕ46×30
		ION-F			? & 360×360	10	ϕ46×13

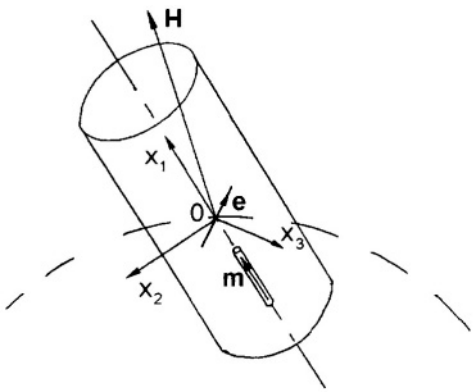

Figure 22.1. SC with a passive magnetic attitude control system

Ox_2x_3 (Fig.22.1). We take into account in the present analysis only the magnetic restoring torque developed by the permanent magnet; the magnetic damping torque produced by the hysteresis rod and the gravity-gradient torque. For an axisymmetric SC the equations of motion then have the form

$$
\begin{aligned}
A\frac{d\omega_1}{dt} &= \mu_0 H V_b \overline{W}(e_2 b_{13} - e_3 b_{12}),\\
B\frac{d\omega_2}{dt} &+ (A-B)\omega_3\omega_1\\
&= 3\omega_0^2(A-B)\xi_1\xi_3 + \mu_0 H V_b \overline{W}(e_3 b_{11} - e_1 b_{13}) - \mu_0 m H b_{13},\\
B\frac{d\omega_3}{dt} &+ (B-A)\omega_1\omega_2 \qquad (22.1)\\
&= 3\omega_0^2(B-A)\xi_1\xi_2 + \mu_0 H V_b \overline{W}(e_1 b_{12} - e_2 b_{11}) + \mu_0 m H b_{12},\\
\frac{\gamma}{dt} &= \overline{\omega}_1 - (\overline{\omega}_2\cos\gamma - \overline{\omega}_3\sin\gamma)\tan\beta,\\
\frac{d\alpha}{dt} &= \frac{1}{\cos\beta}(\overline{\omega}_2\cos\gamma - \overline{\omega}_3\sin\gamma),\\
\frac{d\beta}{dt} &= \overline{\omega}_2\sin\gamma + \overline{\omega}_3\cos\gamma
\end{aligned}
$$

where A and B are the axial and equatorial principal moments of inertia of the SC; V_b is the volume of the hysteresis rod; μ_0 the permeability of vacuum; $\overline{W}(\overline{H}_\tau, \overline{H}_\tau)$ is an hysteresis function which describes magnetization of the rod (it depends on the properties of the magnetic material from which the rod is made, on the ratio rod length to diam-

eter, on the running intensity $\overline{H}_\tau$ of the re-magnetizing field, and on the rate $\overline{H}_\tau$ of re-magnetization); e_1, e_2, e_3 are the projections of the unit vector **e** on the axes of the BRS. Other notations used were introduced in Chapter 21. It is noted that the interaction of the rod with the permanent magnet is not taken into account in (22.1). Also that, if several rods are installed on a SC, the terms corresponding to each rod should be appropriately inserted on the right-sides of equations (22.1).

To approximate an hysteresis loop, the *Rayleigh*† *model*

$$\overline{W}(\dot{\overline{H}}_\tau, \overline{H}_\tau) = (\mu_{in} + \alpha_R \overline{H}_{\tau m})\overline{H}_\tau - \frac{\alpha_R}{2}(\overline{H}_{\tau m}{}^2 - \overline{H}_\tau{}^2)\,\mathrm{sign}\,\frac{d\overline{H}_\tau}{dt}$$

is used, where μ_{in} is the initial magnetic permeability; α_R is the Rayleigh constant and $H_{\tau m}$ an amplitude of the re-magnetizing field. This model correctly describes the shape and area of the hysteresis loop in the vicinity of the origin $(\overline{H}_\tau = 0)$, since re-magnetization is symmetrical with respect to this origin.

Another hysteresis model called the *parallelogram model* is described by the formula

$$\overline{W}(\dot{\overline{H}}_\tau, \overline{H}_\tau) = \mu_{in}\left(\overline{H}_\tau - \frac{k_h}{2}\,\mathrm{sign}\,\frac{d\overline{H}_\tau}{dt}\right) \tag{22.2}$$

where $k_h/2$ is a constant and the product $\mu_{in}k_h/2$ is associated with the coercive force of the rod material. This model takes into account the dependence of the magnetization on the sign of the rate of re-magnetization, although the magnitude of this rate is not involved (analogous to the case of Coulomb friction). Since the hysteresis damping effect has itself the nature of Coulomb friction, this model has the advantage of describing the hysteresis effect for arbitrary variations of the magnetizing field. However, its accuracy is low. The *Rayleigh model* provides a quantitative advance in accuracy but, on the other hand, is subject to strong constraints with respect to the external re-magnetizing field.

---

†Rayleigh (Strutt), John William (1842–1919). English physicist. Member of the London Royal Society. His main work was devoted to the investigation of linear and non-linear oscillations and he was one of the founders of the theory of oscillations.

Introduce again the unitless parameters and variables

$$\lambda = \frac{A}{B}, \quad \eta = \frac{\mu_0 m H_0}{B\omega_0^2}, \quad \varepsilon = \frac{\mu_0 \alpha_R V_b H_0^3}{B\omega_0^2},$$
$$\nu = \frac{\mu_{in}}{\alpha_R H_0}, \quad \Omega_1 = \frac{\omega_1}{\omega_0}, \quad u = \omega_0(t - t_0),$$
$$\Omega_2 = \frac{1}{\omega_0}(\omega_2 \cos\gamma - \omega_3 \sin\gamma), \quad \Omega_3 = \frac{1}{\omega_0}(\omega_2 \sin\gamma + \omega_3 \cos\gamma)$$

and rewrite equations (22.1) in the following form

$$\begin{aligned}
\dot{\Omega}_1 &= \frac{\varepsilon\sqrt{1-e_1^2}}{\lambda} NW(\sin\alpha\cos z + \cos\alpha\sin\beta\sin z),\\
\dot{\gamma} &= \Omega_1 - \Omega_2\tan\beta - \Pi_2/\cos\beta,\\
\dot{\Omega}_2 &= -(\lambda\Omega_1 - \Omega_2\tan\beta - \Pi_2/\cos\beta)\Omega^3 - \eta N\sin\alpha\\
&\quad -3(1-\lambda)\Lambda_1(\Lambda_2\cos\beta + h_{32}\sin\beta)\\
&\quad +\varepsilon NW(\sqrt{1-e_1^2}\cos\alpha\cos\beta\sin z - e_1\sin\alpha), \qquad (22.3)\\
\dot{\Omega}_3 &= (\lambda\Omega_1 - \Omega_2\tan\beta - \Pi_2/\cos\beta)\Omega^2 - \eta N\cos\alpha\sin\beta\\
&\quad +3(1-\lambda)(\Lambda_2\cos\beta + h_{32}\sin\beta)(-\Lambda_2\sin\beta + h_{32}\cos\beta)\\
&\quad -\varepsilon NW(\sqrt{1-e_1^2}\cos\alpha\cos z + e_1\sin\beta)\cos\alpha,\\
\dot{\alpha} &= \Omega_2/\cos\beta + \Pi_2\tan\beta - \Phi_2, \quad \dot{\beta} = \Omega_3 - \Pi_1,
\end{aligned}$$

where $N = H/H_0$, $H_\tau = \overline{H}_\tau/H_0$ and

$$\begin{aligned}
\Pi_1 &= \Phi_1\sin\alpha + \Phi_3\cos\alpha, \qquad \Pi_2 = \Phi_1\cos\alpha - \Phi_3\sin\alpha,\\
\Lambda_1 &= h_{31}\sin\alpha + h_{33}\cos\alpha, \qquad \Lambda_2 = h_{31}\cos\alpha - h_{33}\sin\alpha,\\
z &= \gamma + \varphi, \qquad \sin\varphi = e_3/\sqrt{1-e_1^2}, \qquad \cos\varphi = e_2/\sqrt{1-e_1^2},\\
W(\dot{H}_\tau, H_\tau) &= (\nu + H_{\tau m})H_\tau - \frac{1}{2}({H_{\tau m}}^2 - {H_\tau}^2)\,\mathrm{sign}\,\frac{dH_\tau}{dt}.
\end{aligned}$$

In order now to apply averaging we introduce, so called, 'fast time'. which has a scale similar to the scale of natural SC librations.

22.2 Fast Time. Equations of Motion in Standard Form

We already supposed that the SC is equipped with a strong permanent magnet. This allows us to introduce the small unitless parameter

$\delta = 1/\sqrt{\eta} \ll 1$ and to use the fast time $\tau = (u - u_0)/\delta$, where u_0 is the magnitude of the argument of latitude u at the initial moment of fast time. We assume the spin rotation of the SC to be slow (i.e. $\Omega_1 \sim 1$). Denote by $(')$ differentiation with respect to τ and rewrite equations (22.3) in the form

$$\begin{aligned}
\Omega_1' &= \delta\frac{\varepsilon}{\lambda}\sqrt{1-e_1^2}NW(\sin\alpha\cos z + \cos\alpha\sin\beta\sin z),\\
\gamma' &= \alpha'\sin\beta + \delta(\Omega_1 - \Phi_2\sin\beta - \Pi_2\cos\beta),\\
\alpha''\cos\beta &= 2\alpha'\beta'\sin\beta - N\sin\alpha + \delta(-\lambda\Omega_1 + 2\Pi_2\cos\beta)\beta'\\
&+\delta^2[-3(1-\lambda)\Lambda_1(\Lambda_2\cos\beta + h_{32}\sin\beta)\\
&+\varepsilon NW(\sqrt{1-e_1^2}\cos\alpha\cos\beta\sin z + e_1\sin\alpha)\\
&+\frac{\partial\Pi_2}{\partial u}\sin\beta - \Phi_2\cos\beta - (\lambda\Omega_1 - \Phi_2\tan\beta - \frac{\Pi_2}{\cos\beta})\Pi_1], \qquad (22.4)\\
\beta'' &= (\alpha')^2\sin\beta\cos\beta - N\sin\alpha\sin\beta\\
&+\delta(\lambda\Omega_1 - 2\Pi_2\cos\beta - 2\Phi_2\sin\beta)\alpha'\beta\beta\\
&+\delta^2[3(1-\lambda)(\Lambda_2\cos\beta + h_{32}\sin\beta)(-\Lambda_2\sin\beta + h_{32}\cos\beta)\\
&-\varepsilon NW(\sqrt{1-e_1^2}\cos\beta\cos z + e_1\sin\beta)\cos\alpha\\
&-\frac{\partial\Pi_1}{\partial u} + (\lambda\Omega_1 - \Phi_2\tan\beta - \Pi_2\cos\beta)(\Pi_2\cos\beta - \Pi_2\sin\beta)].
\end{aligned}$$

At $\delta = 0$ equations (22.4) describe the motion of a rigid body for the Lagrange case, yielding the three first integrals

$$\begin{aligned}
\Omega_1 &= C_1,\\
(\alpha'\cos\beta)^2 + (\beta')^2 - N\cos\alpha\cos\beta &= C_2,\\
\alpha'\cos\alpha\sin\beta\cos\beta - \beta'\sin\alpha &= C_3,
\end{aligned}$$

where C_j are constants $(j = 1, 2, 3)$.

Also, for $\delta = 0$, equations (22.4) have the specific solution $\Omega_1 = \text{const}$, $\gamma = \text{const}$, $\alpha = \beta = 0$. Substitute in (22.4) $\alpha = \delta\bar{\alpha}$, $\beta = \delta\bar{\beta}$, where $\bar{\alpha}$ and $\bar{\beta}$ are new variables, and retain in the expansion the first non-zero terms in positive degree powers of the parameter δ. Thus, we obtain

$$\begin{aligned}
\Omega_1' &= \delta^2\frac{\varepsilon\sqrt{1-e_1^2}}{\lambda}NW(\bar{\alpha}\cos z + \bar{\beta}\sin z), \quad \gamma' = \delta(\Omega_1 - \Phi_1),\\
&\bar{\alpha}'' + N\bar{\alpha} = \delta G_1, \quad \bar{\beta}'' + N\bar{\beta} = \delta G_2, \qquad (22.5)
\end{aligned}$$

where the following notations are introduced

$$G_1 = (-\lambda\Omega_1 + 2\Phi_1)\overline{\beta}' - 3(1-\lambda)h_{31}h_{33} - \dot{\Phi}_2 - (\lambda\Omega_1 - \Phi_1)\Phi_3 + \\ + \varepsilon NW\sqrt{1-e_1^2}\sin z,$$
$$G_2 = -(-\lambda\Omega_1 + 2\Phi_1)\overline{\alpha}' - 3(1-\lambda)h_{31}h_{32} - \dot{\Phi}_3 - (\lambda\Omega_1 - \Phi_1)\Phi_1 - \\ - \varepsilon NW\sqrt{1-e_1^2}\cos z,$$
$$H'_\tau = N[e_2(\overline{\alpha}'\sin\gamma - \overline{\beta}'\cos\gamma) + e_3(\overline{\alpha}'\cos\gamma + \overline{\beta}'\sin\gamma)].$$

At $\delta = 0$ the two last equations in (22.4) have the solution

$$\overline{\alpha} = A_1\cos\psi_1, \qquad \overline{\alpha}' = -\rho A_1\sin\psi_1,$$
$$\overline{\beta} = A_2\cos\psi_2, \qquad \overline{\beta}' = -\rho A_2\sin\psi_2, \tag{22.6}$$

where $\psi_j = \rho\tau + \zeta_j$, $(j = 1, 2)$; A_j and ζ_j are constants of integration; $\rho = \sqrt{N}$.

When $\delta \neq 0$ then A_1, A_2, ζ_1, ζ_2 are not constants and we use A_1, A_2, ψ_1, ψ_2 through the transformation (22.6) as new variables (instead of the previously used variables $\overline{\alpha}$, $\overline{\beta}$, $\overline{\alpha}'$, $\overline{\beta}'$). Resolving the equations obtained on employing the transformation (22.6) with respect to the derivatives, we obtain the equations in *standard form* (as is required to allow the application of the *Krilov*[‡]*-Bogolyubov*[§] *averaging method*)

$$A'_j = -\delta(\dot{\rho}A_j\sin\psi_j + G_j)(\sin\psi_j)/\rho,$$
$$\psi'_j = \rho - \delta(\dot{\rho}\sin\psi_j + G_j/A_j)(\cos\psi_j)/\rho, \quad (j = 1, 2) \tag{22.7}$$
$$\Omega'_1 = \delta^2\varepsilon NW\sqrt{1-e_1^2}(A_1\cos\psi_1\cos z + A_2\cos\psi_2\sin z)/\lambda,$$
$$\gamma' = \delta(\Omega_1 - \Phi_1);$$
$$H'_\tau = \rho N\sqrt{1-e_1^2}(-A_1\sin\psi_1\sin z + A_2\sin\psi_2\cos z),$$

where the variables A_1, A_2, Ω_1, γ are called *slow variables* and ψ_1 and ψ_2 are *fast variables.*

[‡]Krilov, Nikolay Mitrofanovich (1879–1955), Russian mathematician. Member of the USSR Academy of Sciences. He developed interpolation techniques and methods for the approximated integration of non-linear differential equations.

[§]Bogolyubov, Nikolay Nikolaevich (1909–1992), Russian mathematician and physicist. Member of the USSR Academy of Sciences. He developed approximated methods and averaging techniques for mathematical analysis, mathematical physics and nonlinear mechanics.

The rates of variation of the fast variables ψ_1 and ψ_2 are the same and equal to ρ in an asymptotic sense. Following the Krilov-Bogolyubov method for the resonant case, introduce a new variable θ to denote the phase difference

$$\theta = \psi_2 - \psi_1. \tag{22.8}$$

Substituting ψ_2 from (22.8) in (22.7) we obtain

$$\begin{gathered}
A_1' = -\delta(\dot{\rho} A_1 \sin^2 \psi_1 + \overline{G}_1 \sin \psi_1)/\rho, \\
A_2' = -\delta[\dot{\rho} A_2 \sin^2(\psi_1 + \theta) + \overline{G}_2 \sin(\psi_1 + \theta)]/\rho, \\
\theta' = -\delta[\overline{G}_2 \cos(\psi_1 + \theta)/A_2 - \overline{G}_1 \cos \psi_1 / A_1]/\rho, \qquad (22.9) \\
\Omega_1' = \delta^2 \varepsilon \sqrt{1 - e_1^2} NW[A_1 \cos \psi_1 \cos z + A_2 \cos(\psi_1 + \theta) \sin z]/\lambda, \\
\gamma' = \delta(\Omega_1 - \Phi_1), \\
\psi_1' = \rho - \delta(\dot{\rho} \sin \psi_1 + G_1/A_1)(\cos \psi_1)/\rho,
\end{gathered}$$

where

$$\begin{gathered}
\overline{G}_1 = G_j\Big|_{\psi_2 = \psi_1 + \theta}, \quad (j = 1, 2) \\
\sin \psi_1^* = A_2 \cos z \sin \theta / \Phi \qquad \cos \psi_1^* = (-A_1 \sin z + A_2 \sin z \cos \theta / \Phi, \\
H_\tau' = \rho N \sqrt{1 - e_1^2} \Phi \sin(\psi_1 + \psi_1^*), \\
\Phi = \sqrt{A_1^2 \sin^2 z \ + A_2^2 \cos^2 z - A_1 A_2 \sin 2z \cos \theta}.
\end{gathered}$$

Now the averaging procedure may be carried out.

22.3 Averaging the Equations of Motion

Using the Krilov-Bogolyubov averaging method adapted for equations with discontinuous right-sides[¶], we seek a solution of (22.9) in the following form

$$y = \overline{y} + \delta Y(\overline{y}, u_0, \delta), \tag{22.10}$$

where $y = (\Omega_1, \gamma, A_1, A_2, \theta)$ is a vector of slow variables and $\overline{y} = (\overline{\Omega}_1, \overline{\gamma}, \overline{A}_1, \overline{A}_2, \overline{\theta})$ a vector of new slow variables still to be obtained; Y is an unknown vector-function.

[¶]V.A.Plotnikov and T.S.Zver'kova. Averaging Method for Equations of Standard Form with Discontinuous Right-Sides. *Differential Equations,* 1982, Vo1.18, N 6, pp.1076-1078.

To obtain the new slow variables $\overline{y}$ to a first approximation with regard to the small parameter δ, we write the vector equation

$$\overline{y}' = \delta \overline{Y}. \tag{22.11}$$

Here $\overline{Y}$ is a vector-function of the right-sides of equations (22.9) divided by δ and averaged over the fast variable ψ_1 (termed 'averaged over the fast phase ψ_1'), i.e. $\overline{Y} = \frac{1}{2\pi}\int_0^{2\pi} Y d\psi_1$. Later we will remove the bar from the notation for the new slow variables.

Equation $H'_\tau = 0$ has only two physically different solutions

$$\psi_1^{(0)} = -\psi_1^*, \quad \text{and} \quad \psi_1^{(0)} = 2\pi - \psi_1^* \tag{22.12}$$

if $N\Phi(1-e_1^2) \neq 0$. Otherwise, any value of the variable ψ_1 is a solution of the equation $H'_\tau = 0$ if either

$$e_1^2 = 1 \quad \text{or} \quad A_1 = A_2 = 0 \quad \text{or} \quad \tan z = \pm A_2/A_1.$$

We do not, as already mentioned, take into account in this analysis the effect of the strong permanent magnet on the hysteresis rod. It is, however, reasonable to assume that the configuration pertains where this rod lies in the equatorial plane of the permanent magnet. Also, that it is perpendicular to this magnet and crosses it through its center (i.e. $e_1 = 0$). In this case, the vector of intensity of the magnetic field of the permanent magnet is closely perpendicular to the equatorial plane. Thus, the component of this vector of intensity is at a minimum along the hysteresis rod. Consequently, the range of the external remagnitazing field variation is shifted through a minimum distance from the origin $H_\tau = 0$, and the effectiveness of the rod as a damper is only minimally decreased.

Averaging the right-sides of equations (22.9) over ψ_1, returning to the initially used independent variable u and noting that $\chi = (2\Phi_1 - \lambda\Omega_1)/2$, we obtain the following equations

$$\dot{A}_1 = -\frac{\dot{\rho}A_1}{2\rho} + \chi A_2 \cos\theta - \frac{\varepsilon\delta N^2 \sin z}{2\rho}[\nu A_2 \cos z \sin\theta$$
$$+\delta N\Phi(\frac{4}{3\pi}A_1 \sin z - A_2 \cos z(\frac{4}{3\pi}\cos\theta - \sin\theta))],$$
$$\dot{A}_2 = -\frac{\dot{\rho}A_2}{2\rho} - \chi A_1 \cos\theta \frac{\varepsilon\delta N^2 cosz}{2\rho}[\nu A_1 \sin z \sin\theta$$

$$+\delta N\Phi(A_1 \sin z(\sin\theta + \tfrac{4}{3\pi}\cos\theta) - \tfrac{4}{3\pi}A_2 \cos z)], \qquad (22.13)$$

$$\dot{\theta} = -\left\{\chi\left(\frac{A_2}{A_1} - \frac{A_1}{A_2}\right)\right.$$
$$+\frac{\varepsilon\delta N^2}{2\rho}\left[\nu\left(\cos 2z + \sin z\cos z\cos\theta\left(\frac{A_2}{A_1} - \frac{A_1}{A_2}\right)\right)\right.$$
$$+\delta N\Phi\left(\cos 2z + \sin z\cos z\left(\frac{4}{3\pi}\sin\theta\left(\frac{A_2}{A_1} + \frac{A_1}{A_2}\right)\right.\right.$$
$$\left.\left.\left.\left.+\cos\theta\left(\frac{A_2}{A_1} - \frac{A_1}{A_2}\right)\right)\right)\right]\right\},$$

$$\dot{\Omega}_1 = \frac{\varepsilon\delta^2 N^2}{2\lambda}[\nu((A_1^2 - A_2^2)\sin z\cos z - A_1A_2\cos\theta\cos 2z)$$
$$+\delta N\Phi((A_1^2 - A_2^2)\sin z\cos z - A_1A_2(\cos 2z\cos\theta + \tfrac{4}{3\pi}\sin\theta))],$$
$$\dot{z} = \Omega_1 - \Phi_1,$$

where we retain on the right-sides only terms up to the order $O(\delta)$, apart from in the second last equation where we keep terms up to the order $O(\delta^2)$.

22.4 Case of Two Mutually Orthogonal, Identical, Hysteresis Rods

Next we consider a case where two, mutually orthogonal, identical, hysteresis rods are installed on the SC. For two such rods, equations (22.13) have the form

$$\dot{A}_1 = -\frac{\dot{\rho}A_1}{2\rho} + \chi A_2\cos\theta - \frac{\bar{\varepsilon}N^3}{2\rho}[\tfrac{4}{3\pi}A_1(\Phi_{(1)}\sin^2 z + \Phi_{(2)}\cos^2 z)$$
$$-A_2\sin z\cos z(\Phi_{(1)} - \Phi_{(2)})(\tfrac{4}{3\pi}\cos\theta - \sin\theta)],$$
$$\dot{A}_2 = -\frac{\dot{\rho}A_2}{2\rho} - \chi A_1\cos\theta$$
$$+\frac{\bar{\varepsilon}N^3}{2\rho}[A_1\sin z\cos z(\Phi_{(1)} - \Phi_{(2)})(\sin\theta + \tfrac{4}{3\pi}\cos\theta)$$
$$-\tfrac{4}{3\pi}A_2(\Phi_{(1)}\cos^2 z + \Phi_{(2)}\sin^2 z)], \qquad (22.14)$$
$$\dot{\theta} = -\left\{\chi\left(\frac{A_2}{A_1} - \frac{A_1}{A_2}\right) + \frac{\bar{\varepsilon}N^3}{2\rho}(\Phi_{(1)} - \Phi_{(2)})\left[\cos 2z\right.\right.$$
$$\left.\left.+\sin z\cos z\left(\frac{4}{3\pi}\sin\theta\left(\frac{A_2}{A_1} + \frac{A_1}{A_2}\right) + \cos\theta\left(\frac{A_2}{A_1} - \frac{A_1}{A_2}\right)\right)\right]\right\},$$

$$\dot{\Omega}_1 = \frac{\bar{\varepsilon}N^3}{2\lambda}[(\Phi_{(1)} - \Phi_{(2)})[(A_1^2 - A_2^2)\sin z\cos z - A_1A_2\cos 2z\cos\theta] - \frac{4}{3\pi}A_1A_2(\Phi_{(1)} + \Phi_{(2)})\sin\theta],$$
$$\dot{z} = \Omega_1 - \Phi_1,$$

where

$$\Phi_{(1)} = \sqrt{A_1^2\sin^2 z + A_2^2\cos^2 z - A_1A_2\sin 2z\cos\theta},$$
$$\Phi_{(2)} = \sqrt{A_1^2\cos^2 z + A_2^2\sin^2 z + A_1A_2\sin 2z\cos\theta}, \qquad \bar{\varepsilon} = \varepsilon\delta^2.$$

It follows from equations (22.14) that two mutually orthogonal, identical, hysteresis rods installed in the equatorial plane of the magnet cause cancellation of those terms in the equations which are proportional to the parameter ν. These latter terms represent that part of the magnetization of the rods which is linearly dependent on the applied external magnetic field. Thus, they do not contribute to the damping but, rather, exert a perturbing effect. The potential to retain in the equations as a result of cancellation only those terms responsible for non-linear magnetization, provides a practical reason to install two hysteresis rods.

22.4.1 Case without Hysteresis Rods

If there are no hysteresis rods on the SC (i.e. $\bar{\varepsilon} = 0$), equations 22.13) have three first integrals. Two of these are

$$\rho(A_1^2 + A_2^2) = c_1 \quad \text{and} \quad 2\rho A_1^2A_2^2\sin\theta = c_2 \tag{22.15}$$

and they allow us to obtain an insight into SC motion. From the expression for c_1 it follows that the angle between the SC axis of symmetry and vector **H** has a maximum value over the equator and a minimum value over the poles. The ratio of these maximum and minimum deviations is equal to $\sqrt{2}$.

Substituting amplitudes A_1 and A_2 from (22.15) in the third equation of (22.13) when $\bar{\varepsilon} = 0$, yields the equation

$$\dot{\theta} = -\frac{2\chi\sin\theta\sqrt{c_1^2\sin^2\theta - c_2^2}}{c_2}.$$

Its solution has the form[||]

$$\operatorname{cotan}\theta = \sqrt{p^2 - 1}\sin(s + c_3), \tag{22.16}$$

where $p = c_1/c_2$, $s = 2\int_0^u \chi du$; and c_3 is a constant of integration. Write expressions for A_1, A_2 in the following way.

$$A_1^2 = \frac{1}{2\rho}[c_1 - \sqrt{c_1^2 - c_2^2}\cos(s + c_3)],$$
$$A_2^2 = \frac{1}{2\rho}[c_1 + \sqrt{c_1^2 - c_2^2}\cos(s + c_3)] \tag{22.17}$$

where $\Phi = \{[c_1+(c_1^2-c_2^2)^{1/2}\cos(s+c_3+2z)]/(2\rho)\}^{1/2}$. Use now c_1, c_2, c_3 as new slow variables instead of A_1, A_2, θ through formulae (22.16) and(22.17). Thus, we obtain

$$\dot{c}_1 = -\bar{\varepsilon}\frac{4}{3\pi}\frac{N^3}{\rho}c_1, \qquad \dot{c}_2 = -\frac{\bar{\varepsilon}c_2}{\rho\Phi}, \qquad \dot{c}_3 = \frac{\bar{\varepsilon}c_1\sin(s + c_3 + 2z)}{\rho\Phi(c_1^2 - c_2^2)^{1/2}}, \tag{22.18}$$

Note that the equation for c_1 may be considered separately from the others since it does not contain the fast variable z (remembering that $z = \gamma + \varphi$). This allows us to obtain a solution for c_1 in the form of a quadrature.

22.4.2 Case of Weak Damping

The case of weak damping corresponds to the inequality $\bar{\varepsilon} \ll 1$. To obtain a solution for (22.18) we need, generally, to average the right-sides of the equations over the fast variable z. As previously noted, we do not need to make this averaging in the case of the first equation, for which we can immediately write a solution in the form of the quadrature

$$c_1(u) = c_1(u_0)\exp\left(-\frac{4}{3\pi}\bar{\varepsilon}\int_{u_0}^{u}(1 + 3\sin^2 i\sin^2 u)^{5/4}du\right). \tag{22.19}$$

[||]M.L.Pivovarov. Concerning the Librations of a Magnetically Stabilized Satellite, Transactions of Russian Academy of Sciences: *Mechanics of Solids,* 1994, Issue 3, p.5.

Thus, the instantaneous amplitude Θ defined by $\sqrt{A_1^2 + A_2^2}$ of the deviation of the SC axis of symmetry from the vector **H** as a function of the argument of latitude u, is described by an expression derived from (22.15) and (22.19)

$$\Theta(u) = \frac{1}{\delta} \frac{\Theta(u_0)}{(1 + 3\sin^2 i \sin^2 u)^{1/8}} \times \exp\left(-\frac{2}{3\pi} \bar{\varepsilon} \int_{u_0}^{u} (1 + 3\sin^2 i \sin^2 u)^{5/4} du \right). \qquad (22.20)$$

In the framework of the Rayleigh model, the area S_h of an hysteresis loop is expressed in terms of the Rayleigh constant α_R and the amplitude H_{rm} of the re-magnetizing field $S_h = 16\alpha_R \overline{H}_{rm}^3/3$. For small $\overline{H}_{rm}$ the value of α_R is constant. Thus, the values of S_h and α_R may be determined using the results of experimental measurements. Parameter $\bar{\varepsilon}$ contained in (22.20) is determined from the expression $\bar{\varepsilon} = \alpha_R V_b H_0^2/m$. When measurement data for hysteresis loops are available, one can quantitatively describe the transient process using formulae (22.18) and (22.20).

It is easy to integrate averaged equations (22.14) and to calculate the quadratures (22.19), (22.20) using standard techniques. An example of the numerical computation of quadratures (22.19) and (22.20) is now provided. Assume that $\Theta(0)/\delta = 1$. Two identical rods are installed on a SC as described above and $\bar{\varepsilon} = 0.1$. A graph of c_1 plotted against u, when $c_1(0) = 1$ and $i = \pi/2$, is shown in Fig. 22.2. Also, a

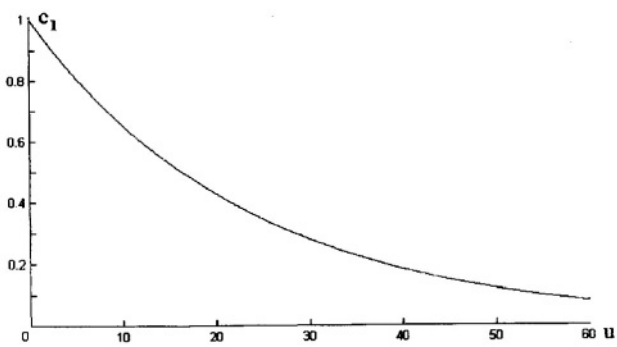

Figure 22.2. Graph of c_1 against u when $i = \pi/2$

graph of the instantaneous amplitude Θ divided by δ plotted against u when $i = \pi/2$ is shown in Fig. 22.3.

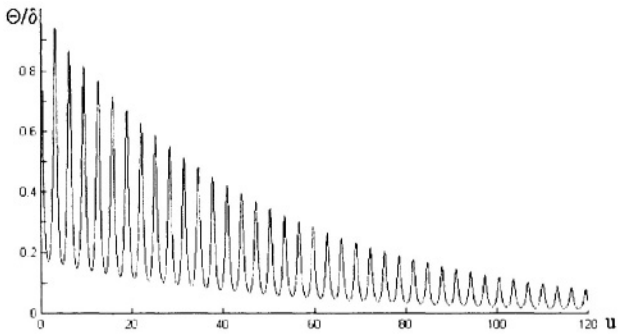

Figure 22.3. Graph of the instantaneous amplitude Θ against u when $i = \pi/2$

Two curves (*1* and *2*) representing instantaneous amplitudes calculated by two different methods in the framework of the parallelogram model (22.2) are plotted in Fig.22.4. Curve *1* which was obtained using an averaged equation similar to (22.19) represents the amplitude $\overline{\Theta} = \delta\sqrt{A_1^2 + A_2^2}$ with respect to time when $\sin i = 1, \delta = 0.17$ and $\bar{\varepsilon} = 0.06$. Actually, this amplitude $\overline{\Theta}$ is an osculating curve for fast librations of the SC axis of symmetry with frequency of order $\sqrt{\eta}$. Curve *2* was obtained by integrating the non-linear equations (22.3), while meeting the constraint $H_\tau = 0$ inherent in the hysteresis model. It describes the discrete function denoted by large black dots, which is calculated using the formula

$$\Theta_n = \max_{2\pi(n-1)<u<2\pi n} |\arccos[\cos\alpha(u)\cos\beta(u)]|$$

where Θ_n is the maximum deviation of the SC axis of symmetry from **H** during the n-th revolution around the Earth ($n = 1, 2, 3, \ldots$). The values of the parameters contained in the equations are $\sin i = 1$, $\eta = 36$, $\lambda = 0.2$. Parameters ε and ν when two identical rods are installed provide the value 0.06 for parameter $\bar{\varepsilon}$. Initial conditions of motion were chosen such that the initial amplitudes $\overline{\Theta}$ and Θ_1 were the same. Symbol x on curve *1* denotes the maximum value of $\overline{\Theta}$ during a particular revolution. The curves closely match until line *2* tends to the horizontal. This approximately corresponds to the attainment by

the SC of steady-state motion. The amplitude of steady-state motion is determined by several factors, including non-uniform variation of the vector **H**, the residual magnetization of the rods and residual, non-damped, spin rotation of the SC. Comparison of the sine-style curves

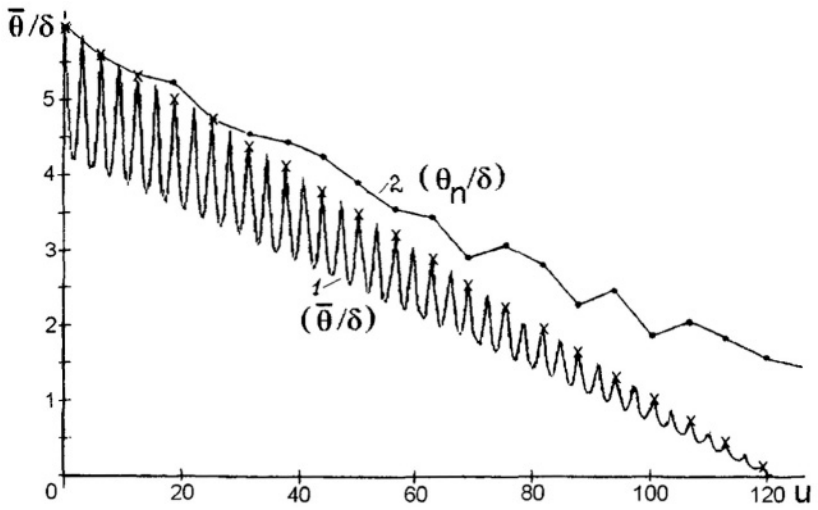

Figure 22.4. Results of integrating the averaged equation (curve *1*) and the non-linear equations (curve *2*) using the parallelogram model (22.2)

in Fig. 22.3 and in Fig. 22.4 shows that their behaviour depends on the hysteresis model which is employed (Rayleigh's model and the parallelogram model respectively).

This approach to investigating the motion of a SC with respect to the vector **H** was employed for the development of an ACS for the nano-SC *Munin*[**].

[**]M.Ovchinnikov, V.Pen'kov, O.Norberg and S.Barabash. Attitude Control System for the First Swedish Nanosatellite Munin. *Acta Astronautica,* 2000, Vol.46, Issue 2–6, pp.319-326.

Appendix A

Method of van der Pol

The idea behind this method is as follows. Let an equation

$$\ddot{z} + \omega^2 z = \varepsilon\varphi(z, \dot{z}) \tag{A.1}$$

be considered, where ω is a real constant and ε is a small parameter. We call this equation 'quasi-linear' and its librational solution a 'quasi-linear' oscillation. The corresponding equation obtained from (A.1) at $\varepsilon = 0$

$$\ddot{z} + \omega^2 z = 0$$

is called the *generating equation.* Function $\varphi(z, \dot{z})$ should be finite. A general solution of the generating equation has the form

$$z = x \cos y \tag{A.2}$$

where $y = \omega(t + t_0)$; x, which is called an 'amplitude' and t_0 are arbitrary constants; $y(t)$ is termed the 'phase'. Since ε is small we can expect that the general solution (A.2) can approximately provide a solution of (A.1) if amplitude x is not in fact a constant but, rather, slowly varies. The instantaneous character of the librations is totally described by instantaneous values of x and y. We may choose x and y as new variables and use the following transformation

$$z = x \cos y, \quad \dot{z} = -\omega x \sin y.$$

The latter expression is obtained by differentiating the solution (A.2) for constant x and t_0. (Actually, this is a transformation from the

Cartesian coordinates $(z, \dot{z}/\omega)$ to polar ones (x, z)). On substituting these coordinates in the original quasi-linear equation (A.1) and equating $\dot{z}$ from the first formula to the second, resolve the system obtained with respect to $\dot{x}$ and $\dot{y}$ to get

$$\dot{x} = \frac{\varepsilon}{\omega}\varphi_1(x, y), \quad \dot{y} = \omega - \frac{\varepsilon}{\omega x}\varphi_2(x, y). \tag{A.3}$$

These equations are totally equivalent to the original equation due to the non-singularity of the transformation. The first equation shows that the amplitude x varies slowly, due to its being proportional to the small parameter ε. Thus we can expect that, during one libration when the phase y increases by 2π, the amplitude x correspondingly increases weakly. Therefore, we can exchange the right sides of the transformed equations with their values averaged over the period 2π

$$\dot{\bar{x}} = \frac{\varepsilon}{\omega}\overline{\varphi}_1(\bar{x}), \quad \dot{\bar{y}} = \omega - \frac{\varepsilon}{\omega\bar{x}}\overline{\varphi}_2(\bar{x}) \tag{A.4}$$

where

$$\overline{\varphi}_1(x) = \frac{1}{2\pi}\int_0^{2\pi} \varphi(x\cos y, -\omega x \sin y)\sin y dy,$$

$$\overline{\varphi}_2(x) = \frac{1}{2\pi}\int_0^{2\pi} \varphi(x\cos y, -\omega x \sin y)\cos y dy.$$

These equations are called *truncated* equations or *equations of van der Pol.* The first of them depends only on the amplitude x and it can usually be integrated. In the interval t of order $\mathrm{O}(1/\varepsilon)$, the difference $|x(t) - \bar{x}(t)|$ is less than ε (as can be shown using the Krilov-Bogolyubov method).

Appendix B

Linear Equations with Periodic Coefficients

Consider linear differential equations with periodic coefficients in the form

$$\dot{\mathbf{x}} = P(t)\mathbf{x} \tag{B.1}$$

where $\mathbf{x} = (x_1, \ldots, x_n)$ is a vector of variables and $P(t)$ is a matrix with periodic coefficients $p_{jk}(t)$, $(j, k = 1, \ldots, n)$ of period T. Matrix $\mathbf{X}(t)$ is called a *fundamental matrix* since it is composed of n linear-independent solutions of (B.1). A general solution of (B.1) has the form

$$\mathbf{x}(t) = X(t)\mathbf{c} \tag{B.2}$$

where $\mathbf{c} = (c_1, \ldots, c_n)$ is a constant vector determined by the initial conditions. Without loosing generality, we may assume $X(0) = E$, where E is the *identity matrix.* Denoting by $\Delta(t)$ the determinant of the fundamental matrix X, that is, $\Delta(t) = \det X(t)$ and $\Delta(0) = 1$, we can write

$$\Delta(t) = \Delta(0) \exp \int_0^t (p_{11} + \ldots + p_{nn})dt = \exp \int_0^t (p_{11} + \ldots + p_{nn})dt.$$

Since $X(t)$ is a matrix of solutions then, due to the peridiocity of the matrix $P(t)$,

$$X(t + T) = X(t)A \tag{B.3}$$

where A is a matrix with constant coefficients a_{kj}. We assumed that the fundamental matrix $X(t)$ satisfies the initial condition $X(0) = E$.

Thus, at $t = 0$, we obtain from (B.3)

$$X(T) = X(0)A = EA = A.$$

We can, therefore, calculate the matrix A for a finite time interval.

Next, we show that a solution $\mathbf{x}(t)$ satisfying the relationship

$$\mathbf{x}(t+T) = \rho\mathbf{x}(t) \tag{B.4}$$

exists, where ρ is a constant to be determined. Since any solution of (B.1) has the form (B.2), a constant vector $\mathbf{b}$ satisfying the equality $\mathbf{x}(t) = X(t)\mathbf{b}$ and, consequently, also an equality $\mathbf{x}(t+T) = X(t+T)\mathbf{b}$, should exist. Using these last two equalities and (B.4) we obtain

$$X(t+T)\mathbf{b} = \rho X(t)\mathbf{b}.$$

Relation (B.3) allows us to rewrite this formula as follows

$$X(t)(A - \rho E)\mathbf{b} = 0.$$

In order that this equality will have a non-trivial solution with respect to $\mathbf{b}$ for any t, the constant ρ should satisfy the equation

$$\det(A - \rho E) = 0.$$

This is called the *characteristic equation.* For simplicity we assume that all the roots of this equation, which are the *eigenvalues* or *characteristic roots* of the matrix A and, consequently, also of the fundamental matrix $X(T)$, are distinct so that each root ρ_k corresponds to an individual $\mathbf{x}_k$. Hence, we obtain n independent solutions $\mathbf{x}_1, \ldots, \mathbf{x}_n$ satisfying (B.4). These solutions may be represented in the following form

$$\mathbf{x}_k = \exp(\alpha_k t)\ \boldsymbol{\varphi}_k(t),\ (k = 1, \ldots, n)$$

where $\boldsymbol{\varphi}_k(t)$ is a periodic vector (i.e. $\boldsymbol{\varphi}_k(t+T) = \boldsymbol{\varphi}_k(t)$) and α_k is a constant called the *characteristic index,* which is calculated using the formula

$$\alpha_k = \frac{1}{T}\ln\rho_k = \ln|\rho_k| + i\arg\rho_k.$$

The latter expression is valid for calculations where ρ is a complex number. Thus, we obtain a set of equalities

$$\begin{aligned}\mathbf{x}_k(t+T) &= \exp(\alpha_k(t+T))\ \boldsymbol{\varphi}_k(t+T)\\ &= \exp(\alpha_k t)\ \exp(\alpha_k T)\ \boldsymbol{\varphi}_k(t+T)\\ &= \exp(\alpha_k t)\ \exp(\tfrac{1}{T}\ln\rho_k)T\ \boldsymbol{\varphi}_k(t) = \rho_k\boldsymbol{\varphi}_k(t)\end{aligned}$$

which confirm (B.4).

Since the general solution (B.1) has the form (B.2), we can write

$$\mathbf{x}(t+T) = \sum_{k=1}^{n} c_k \mathbf{x}_k(t+T).$$

Due to (B.4) the components of this vector satisfy the equality $c_k\mathbf{x}_k(t+T) = \rho_k c_k \mathbf{x}_k(t)$. Equating the magnitudes of both sides we obtain the relation

$$|c_k\mathbf{x}_k(t+T)| = |\rho_k||c_k\mathbf{x}_k(t)|$$

which implies that, if *all* $|\rho_k| < 1$ and, associatively, all $\alpha_k < 0$ then, in a period T, the magnitudes of all the solutions will decrease. Therefore, a mapping point tends to zero with time. The conditions for stability of the linear differential equations with periodic coefficients are as follows.

- If all $|\rho_k| < 1$, then the trivial solution $\mathbf{x} = 0$ of (B.1) is asymptotically stable.
- If at least one root satisfies the inequality $|\rho_k| > 1$, then the solution $\mathbf{x} = 0$ is unstable.
- If all roots satisfy the inequality $|\rho_k| \leq 1$, then the solution $\mathbf{x} = 0$ is stable but not asymptotically stable.

Earlier it was assumed that all the roots of the characteristic equation have to be distinct. Now we may remark, that the first two conditions for stability listed above can be applied also in the case of multiple roots. However, the third condition requires distinctness of the roots with respect to elementary denominators of the matrix $X(T)$.

Using *Liouville's formula*

$$\Delta(t) = \Delta(0)\exp(\int_0^t (p_{11} + \ldots + p_{nn}))dt$$

we get for the last coefficient a_n of the characteristic equation

$$a_0\rho^n + a_1\rho^{n-1} + a_2\rho^{n-2} + \ldots + a_{n-1}\rho + a_n = 0$$

the set of expressions

$$a_n = (-1)^n \det A = (-1)^n \exp(\int_0^T (p_{11} + \ldots + p_{nn}))dt \qquad \text{(B.5)}$$

which allow us to calculate a_n employing coefficients of matrix P rather than through using the fundamental solution.

Consider now two useful cases. If the characteristic equation has a root $\rho = +1$, then $\mathbf{x}(t+T) = \mathbf{x}(t)$. This implies that (B.1) has a T-periodic solution. If, on the other hand. $\rho = -1$ then $\mathbf{x}(t+2T) = -\mathbf{x}(t+T) = \mathbf{x}(t)$ and, consequently, (B.1) has a 2T-periodic solution. Both cases were already met when we consider in Section 17.4.2 and in Section 21.4.3 applications involving the stability and periodic motion of, respectively, gravity-gradient stabilized and magnetically stabilized SC.

Appendix C

Poincare's Method

Consider the differential equation

$$\dot{x} = f(x, t, \varepsilon), \tag{C.1}$$

where vector x is of dimension n, t is a scalar argument, 'dot' denotes a derivative with respect to t and ε is a small parameter ($\varepsilon \ll 1$). Assume that a function f is analytical with respect to x and ε. Our task is to obtain a solution $x(t, \varepsilon)$ of (C.1) that satisfies the initial condition

$$x(t_0) = x_0.$$

Introduce a *generating equation*

$$\dot{z} = f(z, t, 0) \tag{C.2}$$

corresponding to (C.1) and assume that a solution of this equation is known. Let the same initial conditions $x(t_0) = z(0) = x_0$ be imposed on both equations. Substitute

$$x = z + y$$

in (C.1). Then, the new variable y has to satisfy the equation

$$\dot{y} = f(z + y, t, \varepsilon) - f(z, t, 0) \tag{C.3}$$

and also the initial condition

$$y(0) = 0.$$

Since the right side of (C.3) is an analytical function of y and ε, we may rewrite it in the following form

$$\dot{y} = Ay + \varepsilon \left(\frac{\partial f}{\partial \varepsilon} \right)_0 + B(y, \varepsilon, t). \tag{C.4}$$

Here $A = (\partial f / \partial x)_0$; $B(y, \varepsilon, t)$ is composed of non-linear terms (having order higher than one) with respect to y and ε. The subscripted index 'zero' in (C.4) and in A means that they are calculated for $x = z(t)$ and $\varepsilon = 0$, where $z(t)$ is a known solution of the generating equation (C.2).

Since we assumed that a solution of the initial problem for the generating equation is known, then terms A and $\partial f / \partial \varepsilon$ are also known functions of t. We now seek a solution of the linearized equation (C.4) for y in the form of the following expansion

$$y = \sum_{k=1}^{\infty} y_k \varepsilon^k.$$

On substituting this expansion in (C.4) and equating terms of the same order with respect to powers of ε, we obtain an infinite set of equations

$$\dot{y}_k = Ay_k + D_k, \quad (k = 1, 2, \ldots).$$

In the consecutive solving of these equations, starting from the first one, the functions D_k are known functions of t as they are functions of known variables $y_1, \ldots, y_{k-1}$. The variables y_k have to satisfy

$$y_k(t_0) = 0.$$

Using the above we formulate *Poincare's theorem:*

1. A solution y of the linearized equation (C.4) may be obtained by employing the operations of differentiation and integration in quadrature since the general integral of the generating equation (C.2) is known.

2. Solution x of the non-linear equation (C.1) is an analytical function of the parameter ε, i.e. for sufficiently small absolute values of ε, the set $y = \sum_{k=1}^{\infty} y_k \varepsilon^k$ converges and, consequently, represents an integral of the equation (C.1) expanded in powers of ε.

Bibliography

[1] M.B.Balk. *Elements of Space Flight Mechanics.* Moscow, Nauka Publ., 1965.

[2] R.H.Battin. *An Introduction to the Mathematics and Methods of Astrodynamics.* AIAA Inc. Publ., 1999.

[3] V.V.Beletsky. *Essays on the Motion of Celestials Bodies.* Birkhauser Verlag. Basel – Boston – Berlin, 2001.

[4] V.V.Beletsky. *Motion of an Artificial Satellite About its Center of Mass.* Jerusalem, 1966.

[5] V.V.Beletsky. *Motion of a Satellite About its Center of Mass in the Gravitational Field.* Moscow, MSU Publ., 1975.

[6] V.V.Beletsky and A.A.Khentov. *Rotating Motion of a Magnetized Satellite,* Moscow, Nauka Publ., 1985.

[7] D.Brouwer and G.M.Clemence. *Methods of Celestial Mechanics.* New York and London, Academic, 1961.

[8] A.E.Bryson, Jr. *Control of Spacecraft and Aircraft.* Pronceton Univ. Press, 1994.

[9] V.A.Chobotov. *Spacecraft Attitude Dynamics and Control.* Malabar, Florida, Krieger Publishing Company, 1991.

[10] Cosmonautica: Encyclopedia /Editor-in-chief V.P.Glushko. Moscow, Sovetskaya Entsiklopedia Publ., 1985.

[11] Yu.A.Ermilov, E.E.Ivanov and S.V.Pantjushin. *SC Rendezvous Control.* Moscow, Nauka Publ., 1977.

[12] P.Fortescue and J.Stark. *Spacecraft System Engineering.* Second Ed., New York, Wiley, 1995.

[13] F.J.Hale. *Introduction to Space Flight.* NJ, Prentice-Hall, 1996.

[14] P.C.Hughes. *Spacecraft Attitude Dynamics.* New York, Wiley, 1986.

[15] T.R.Kane, P.W.Likins and D.A.Levinson. *Spacecraft Dynamics.* McGraw-Hill Book Company, 1983.

[16] A.P.Markeev. *The Libration Points in Celestial Mechanics and Cosmodynamics.* Moscow, Nauka Publ., 1978.

[17] D.R.Merkin. *Introduction to the Theory of Motion Stability.* Moscow, Nauka Publ., 1976.

[18] F.R.Moulton. *Introduction to Celestial Mechanics.* Second Ed., New York, Dover Publ. Inc., 1970.

[19] D.E.Okhotsimsky and Yu.G.Sikharulidze. *Foundations of Spaceflight Mechanics.* Moscow. Nauka Publ., 1990.

[20] J.E.Prussing and B.A.Conway. *Orbital Mechanics.* London. Oxford Univ. Press, 1993.

[21] B.V.Rauschenbakh and E.N.Tokar. *Attitude Control of SC.* Moscow, Nauka Publ., 1974.

[22] B.V.Rauschenbakh and M.Yu.Ovchinnikov. *Lectures on Spaceflight Dynamics,* Moscow, MIPT Publ., 1997.

[23] A.Razigraev. *Principles of Spacecraft Control,* Moscow, Mashinostroenie Publ., 1990.

[24] F.P.J.Rimrott. *Introductory Attitude Dynamics* (Mechanical engineering Series). New York, Springer-Verlag Publ. Inc., 1989.

[25] A.E.Roy. *Orbital Motion.* Bristol, Adam Hilger Ltd., 1978.

[26] V.V.Sarychev. *Problems of Satellite Orientation.* Survey of Science and Technology, Series: Space Exploration, Moscow. VINITI Publ., 1978. Vol.11.

[27] V.V.Sarychev and M.Yu.Ovchinnikov. *Magnetic Attitude Control Systems for the Artifical Satellites.* Survey of Science and Technology, Series: Space Exploration, Moscow, VINITI Publ., 1985, Vol.23.

[28] V.Szebehely. *Theory of Orbits.* New York, Academic, 1971.

[29] L.G.Taff. *Celestial Mechanics. A Computational Guide for the Practitioner.* New York, Publ. J. Wiley & Sons, 1985.

[30] M.V.Tarasenko. *The Military Aspects of the Soviet Cosmonautics: A Side View.* Moscow, Nikol Publ., 1992.

[31] D.A.Vallado. *Fundamentals of Astrodynamics and Applications.* Second Ed., Microcosm Press & Kluwer Academic Publ., 2001.

[32] J.R.Wertz. *Spacecraft Attitude Determination and Control.* Kluwer Academic Publ., 1978.

[33] B.Wie. *Space Vehicle Dynamics and Control.* AIAA, Inc. Publ., 1998.

[34] W.E.Wiesel. *Spaceflight Dynamics.* Second Ed., New York, McGrow-Hill, 1996.

Addendum

The Space Environment

A.1 Solar Activity and Near Earth Space

An overview is here provided of various aspects of solar activity and its role in creating the environmental conditions in which SC function when close to the Earth.

A.1.1 The Sun and the Solar Wind

The Sun is a glowing sphere with a visible radius close to 700 000 km and an equatorial rotation period of 25.3 days (~ 27 days when viewed from the moving Earth). Most of the Sun's light originates in a ~ 500 km thick layer called the *Photosphere* with an emission temperature of ~ 6 000 K. Above this is the *Chromosphere* with thickness ~ 2 500 km and temperature ~ 10 000 K. Beyond that extends the *Corona,* which is readily observed during a total solar eclipse. Spectral measurements indicate that the mean energy of particles in the corona corresponds to a temperature of $1 - 2 \cdot 10^6$ K. The corona has no well defined boundary but fades into the *Solar Wind.*

As a result of the large difference in gas pressure between the corona and interstellar space, plasma is driven outward from the Sun despite the restraining influence of solar gravity. This 'solar wind' as it streams past the Earth at a mean Sun-Earth distance of $1.5 \cdot 10^8$ km, or 1 *Astronomical Unit* (AU), consists of protons and electrons ($6.6\,\mathrm{cm}^{-3}$ and $7.1\,\mathrm{cm}^{-3}$ respectively), with a small amount of ionized helium ($0.25\,\mathrm{cm}^{-3}$) and a few ions of heavier elements. The solar wind carries

mass away from the Sun at a rate of $1.6 \cdot 10^9$ kg/s. It has a typical flow speed of ~ 400 km/s at 1 AU but this parameter can vary over a wide range (from ~ 200 km/s to more than 1 000 km/s).

The expanding solar wind drags the solar magnetic field outward with it, forming which is called the 'frozen-in' *Interplanetary Magnetic Field* (IMF). The rotation of the Sun gives to this magnetic field a spiral form (the garden hose effect). The angle between the magnetic field lines and a line drawn from the Sun to an observer at the Earth is about 45°. For refinements to this model see Hundhausen (1995).

A.1.2 Co-rotating Interaction Regions

The solar wind at low *heliographic* latitudes is composed of alternating high speed and low speed flows that co-rotate with the Sun. The high-speed streams originate in structures called *Coronal Holes* that extend equator-ward from the magnetic poles of the Sun and are characterized by open magnetic field lines. The slow speed streams originate in the outer portions of dense coronal streamers that tend to straddle the solar magnetic equator.

At low solar latitudes, the Sun's rotation (~ 13.3° per day) causes high speed plasma to be aligned in the same radial direction, and to travel out behind, plasma exiting from solar regions located relatively to westward (since as viewed from the Earth the Sun rotates from east to west). With increasing distance from the Sun, the high speed streams steepen and overtake the slower plasma ahead, thereby producing compressive *Corotating Interaction Regions* (CIRs) at the leading edges of the streams. A CIR thus comprises a region of high pressure such that its leading edge is a forward wave that propagates into the slow plasma ahead while its trailing edge is a reverse wave that propagates in the backward direction. These waves commonly steepen into forward and reverse shocks at a distance from the Sun of ~ 1 – 2 AU. Ions can be accelerated at both of these shocks. However, the strongest acceleration occurs at the reverse shock, resulting in MeV ions propagating Sun-wards through the high speed stream.

Since slow plasma is accelerated as it encounters the forward wave and fast plasma is decelerated as it encounters the reverse wave, a CIR acts to limit the steepening of a high-speed stream. Also, it transfers momentum and energy from a fast stream to relatively slowly moving

plasma ahead so that speed differences reduce with increasing distance from the Sun. Studies using data recorded out of the ecliptic plane aboard the *Ulysses* SC show that CIRs are systematically tilted due to the tilt of the solar magnetic dipole relative to the solar rotation axis (see Chapter 21). Individual CIRs may survive over multiple solar rotations.

A SC encountering a CIR in interplanetary space can record signatures of its associated shock accelerated particles. For example, on the occasion of the first elliptical orbit of Mars by the *Phobos-2* Mission to Mars and its Moons in 1989 (average distance from the Sun at encounter ~ 1.5 AU), data recorded by the onboard instruments SLED and LET showed that the close planetary environment was characterized by the presence of energetic particles associated with a CIR transit (McKenna-Lawlor *et al.* 1991). Similarly, data obtained by the energetic particle experiment EPONA on ESA's *Giotto* Mission during a flyby of P/Grigg-Skjellerup in 1992 (then at ~ 1 AU from the Sun) showed that this comet was immersed in a CIR (Kirsch *et al.* 1997). The *SOHO* SC orbit at L_1*, the location of *IMP-8* at ~ 1 AU and the solar-polar orbit of *Ulysses,* currently provide convenient complementary locations for long term monitoring of CIRs.

A.1.3 The Solar Cycle

The Sun undergoes a well known cycle of variability showing a main periodicity of approximately eleven years (this period is not constant but varies between about 9.5 and 12.5 years). At the beginning of a cycle the solar magnetic field resembles a dipole, the axis of which is aligned with the Sun's rotation axis. During the following 5 – 6 years this configuration dissipates but, during the last part of the cycle a dipole configuration is restored, but with a polarity that is opposite to the previous one. Restoration of the original configuration requires approximately 22 years (the *Magnetic Cycle*).

Sunspots constitute solar regions characterized by unusually strong, localized, solar magnetic fields (0.1 – 0.4 T). These features were used historically to indicate levels of solar variability since the sunspot num-

* L_1 is the first Lagrangian libration point based on the three body dynamical problem located at 0.99 AU on the Sun-Earth line, just ahead of, and in synchronous orbit with, the Earth as it moves around the Sun.

ber is highest at around the middle of a cycle. More recently the mean daily flux of solar radio emissions at 10.7 cm wavelength (unit: $10^{-22}\,\mathrm{W/m^2Hz}$), which shows well defined peaks and troughs at solar maximum and solar minimum, has been adopted as a representative activity indicator (the *F*10.7 index).

The flux of cosmic radiation arriving at the Earth from the galaxy (see also Section A. 1.10) is modulated by the sunspot cycle. This is related to the fact that complex magnetic fields in space associated with enhanced levels of solar activity act, as solar maximum, to shield the Earth from the influx of energetic charged galactic cosmic rays. The annual mean of cosmic ray intensity displays, in consequence, a systematic variation over a solar cycle, with its maximum phase coincident with the years of minimum sunspot number. A signature of the 22–year magnetic cycle can also be identified in galactic cosmic ray intensities as well as a modulation associated with the rotation of the Sun. Data recorded aboard NASA's *Pioneer-10* and *Voyager-1* and *Voyager-2* SC clearly show classical cosmic ray modulation as well as solar rotation signatures in records obtained as they traveled outwards through the *Heliosphere*[†] during several solar cycles. *IMP-8* provided 1 AU baseline data for these 'deep space' missions and it was noted, for example, that disturbances originating on the Sun marking the rise of a new solar cycle (No. 23) in April/May 1998 were reflected at *Voyager-1* and *–2* in September/October 1998 when these SC were respectively at distances of 56 and 72 AU from the Sun. Such observations illustrate the long range influence of solar processes within the heliosphere.

[†]The solar wind as it flows out through the solar system pushes against the plasma, gas and fields of the *Interstellar Medium* (ISM), thereby forming a large 'bubble' in interstellar space called the heliosphere. This structure acts to shield the solar system from interstellar plasma and magnetic fields, as well as from most of the cosmic rays and dust in our local galactic neighborhood. The boundary where the Sun's atmosphere merges with interstellar space is called the *Heliopause.* The position of the heliopause depends both on the strength of the solar wind and on the properties of the local interstellar medium. Observations made by the *Ulysses* SC indicate that the solar wind from the Sun's poles displays a higher speed than is the case in the ecliptic. Thus, the heliopause should be further from the Sun in the polar direction. (See also Section A.4).

A.1.4 Solar Flares

At its most basic level, a *Solar Flare* consists of a rapid brightening in the solar atmosphere associated with an explosive release of energy that was previously stored in solar magnetic fields. Large flares radiate throughout the electromagnetic spectrum (producing gamma rays, X-rays, visible emissions and km long radio waves).

Energetic particles associated with *Impulsive Solar Flares* are interpreted to be accelerated stochastically by resonant wave-particle interactions in the low corona. The particles concerned are characterized by being electron rich, and they have a $^3He/^4He$ ratio which can attain a value of up to $\sim 10^4$. Also, they display enhancements in heavy elements by about a factor of ten relative to normal coronal abundances (Reames, 1999).

The amount of energy involved in the flare process varies over a wide range from some 10^{19} J for the smallest flaring phenomenon on the Sun to about 10^{25} J for the most important flare events. At the present time it is not clear how the energy is distributed between different kinds of emission. In only one instance to date has an in-depth estimate been made of the various kinds of energy released — that was in the case of a typical (rather small), flare event that occurred on 5 September, 1973 during the *Skylab* Mission. The availability in relation to this event of electromagnetic observations recorded over more than ten decades in wavelength (from less than 1 angstrom to greater than 1 meter), allowed Canfield *et al.* (1980) to estimate (within the limitations of certain inherent errors and assumptions), that the total radiated energy was more than 10^{22} J. The mechanical energy of large scale mass motions (including contributions from cool eruptive material; an emission front; a shock wave with a coronal transient; magnetic field convection and flare core motions), was in parallel estimated by Webb *et al.* (1980) to exceed 10^{24} J.

A.1.5 Coronal Mass Ejections

Coronal Mass Ejections (CMEs) comprise major ejections of material $(10^{13} - 10^{14}$ kg) from the Sun's outer atmosphere in the form of a large cloud of charged particles. These events are often associated with solar flares but can occur independently. CMEs have speeds in the general range 50 km/s to > 1200 km/s.

Strong interplanetary shocks driven by fast CMEs are effective in accelerating solar wind ions in the high solar corona up to energies of several hundred MeV. It is believed that the largest and most energetic *Solar Energetic Particle* (SEP) events recorded at the Earth are *Gradual Events* ('Gradual' implies particle profiles that slowly rise) composed of particles accelerated in the Sun-Earth space in association with shocks driven by fast CMEs (Reames, 1999). Measurements of the ionization rates of Fe in such events suggest that ambient (unheated) coronal material provides the seed population for the constituent particles, even those that are shock accelerated up to energies of ~ 600 MeV/amu (Tylka *et al*., 1995).

If the shocks that generate 'gradual' solar energetic particle (SEP) events are sufficiently energetic and their propagation direction appropriate, these particles can reach the Earth. A transient disturbance in the Earth's magnetic field provides in more than 90 % of these cases a clear indication of the arrival of a travelling shock.

At solar minimum CIRs are dominant in the interplanetary medium. However, as a new cycle develops, signatures of energetic particles associated with solar flares and CMEs gradually come to provide the main signatures of solar activity. During the Cruise Phase (July 1988 – March 1989) of the *Phobos-2* Mission to Mars, the interplanetary medium was in course of changing over from solar minimum dominated to solar maximum dominated conditions (solar cycle 22 began in September, 1986). Energetic particle measurements illustrating the change-over from one characteristic set of particle signatures to another recorded by the onboard SLED instrument are available in McKenna-Lawlor *et al.* (1991).

Fig.A.1 shows an example of SLED data recorded at Mars from 1 – 26 March, 1989 in two energy channels (202 – 609 keV and > 30 MeV respectively) when the activity of the new cycle was already established. In the interval concerned, the SC was in a circular orbit $(\mathrm{T} \sim 8^h)$ about the planet at an altitude of 6 330 km (it is noted that on 27 March telemetry contact with *Phobos-2* was lost). These long duration measurements of high energy particle radiation illustrate the hazardous environmental conditions to which space hardware and 'man in space' can be exposed when the Sun is active (see also Section A.3).

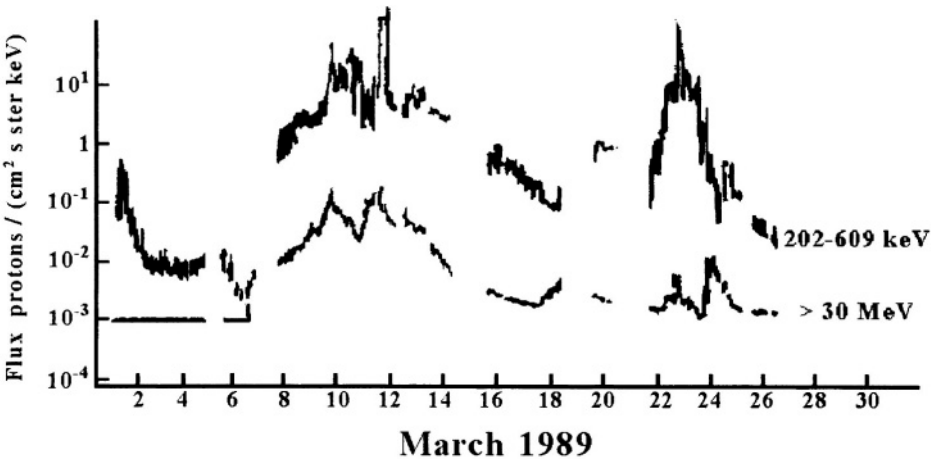

Figure A.1. Solar related particle enhancements recorded in circular orbit about Mars in two energy channels

A.1.6 Prediction of Proton Events

Two models are presently available for predicting long term solar proton fluences[‡], namely the King Model (King, 1974) and the (Jet Propulsion Laboratory) JPL Model of Feynman *et al.* (1993). The King model was constructed using data obtained exclusively during the active years of solar cycle 20 (maximum ~ 1969) and was, for a long time, the standard model used to predict mission integrated solar proton fluences. Several assumptions made by King were later considered questionable and this problem was addressed by the JPL group in a series of studies that, in the final version of their model (JPL-91), analyzed data from three solar cycles (19, 20 and 21). Using the exact dates of solar maximum as the zero reference year for each individual cycle, it was shown that the active phase lasts for 7 years in a particular 11 year period. The years of high fluence begin some 2.5 years prior to the zero reference date and end 4.5 years after that date. An asymmetry in the event frequency and intensity

[‡]*Particle Flux* is defined by the quotient dN/dt, where dN is the increment of particle number in the time interval dt (measured in sec^{-1}). *Particle Fluence* (φ) is defined by the quotient dN/da, where dN is the number of particles incident on a sphere of cross-sectional area da (measured in m^{-2})

thus exists with respect to the peak in solar activity. Only data collected during the seven active years of the cycles studied were used in JPL-91 for radiation level prediction. At the present time this model (Feynman *et al.* 1993) is considered to be the most reliable (in the statistical sense) for the prediction of solar fluences for space mission analysis (Tranquille and Daly, 1992). It is presently used at both ESA and NASA in this regard. It was noted by Goswami *et al.* (1988) that the variability observed in solar particle fluences during the three cycles considered at JPL precludes the possibility of improving the statistical predictability of solar particles until the physics underlying the processes concerned is better understood and this comment is still true.

A.1.7 Numerical Modeling

Data based on the continuous monitoring of various aspects of solar activity can be input to numerical models to predict the arrival at the Earth of individual shocks and energetic particles. Predictions of this kind are routinely made in 'real time' by forecasters at the *Space Environment Center* (SEC) of the *National Oceanographic and Atmospheric Administration,* (NOAA) in Boulder, Colorado, based on continuously updating global information on solar circumstances. The forecasts produced are made continuously available by SEC to a wide spectrum of users.

Three numerical models often used at SEC to provide 'real time' predictions of shock arrivals at the Earth, were tested against the measured arrival times of a sample of flare associated, CME generated, shocks. These shocks were detected in solar wind plasma, interplanetary magnetic field and energetic particle data recorded aboard *SOHO, ACE, WIND, INTERBALL-1* and *IMP-8* (McKenna-Lawlor *et al.* 2002, and references therein). Fig.A.2 shows a typical Gradual SEP recorded by the LION instrument aboard *SOHO* on Days 158-164. 2000. Shock passage was close to noon on Day 160. An account of LION is provided in McKenna-Lawlor *et al.* 1997.

The numerical models tested were the *Shock Time of Arrival Model* (STOA); the *Interplanetary Shock Propagation Model* (ISPM) and the *Hakamada-Akasofu-Fry Solar Wind Model* (HAFv.2). Both STOA and ISPM require as input data the initial coronal shock velocity (de-

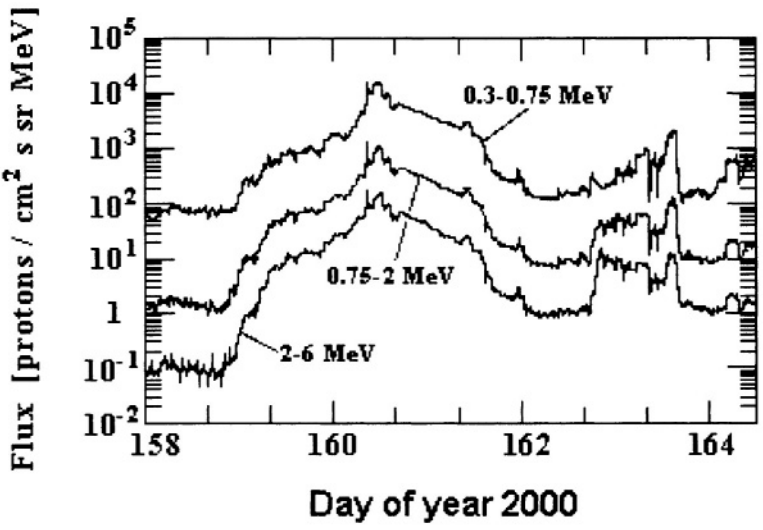

Figure A.2. Gradual SEP recorded aboard *SOHO* at L_1

rived from solar radio burst and coronagraph data) and the input energy duration (derived from solar X-ray measurements), as well as the location of the active source on the Sun. HAFv.2 generally uses the same observational inputs but differs from STOA and ISPM in the way the background solar wind is treated. STOA utilizes the observed solar wind speed at L_1 while ISPM employs a fixed internal model with a representative speed of 360 km/s at 1 AU. HAFv.2, models the inhomogeneous ambient solar wind that affects the propagation of disturbances *en route* from the Sun to the Earth. Realistic inner boundary conditions determine the modeled background solar wind flow and IMF topology. These data are derived from synoptic solar source surface maps of the radial magnetic field and from calculations of the magnetic flux divergence and solar wind velocity close to the Sun.

STOA and ISPM each predict whether a shock will arrive at the Earth (which is for these calculations equivalent to its arrival at L_1) and, if so, when. They also provide a measure of shock strength in terms, respectively, of the concerned magneto-acoustic Mach number and a shock strength index. HAFv.2 can predict the solar wind speed, density, dynamic pressure and interplanetary magnetic field vector

as functions of time at different locations in the heliosphere. The temporal profile of the predicted dynamic pressure at L_1 is used to compute a 'Shock Search Index'. The predicted shock arrival time is generated when this index exceeds an empirical threshold. The shock is deemed to be significant when the predicted post-shock dynamic pressure is greater than its pre-shock value.

For the (limited) sample investigated, STOA provided the smallest values of the (predicted minus observed) arrival times and showed a typical precision better than about 8^h. The ratio of the error estimate for each model to the standard deviations of the observations were 0.60, 1.15 and 1.02 for STOA, ISPM and HAFv.2 respectively. The testing of larger statistical samples and refining of the models is an ongoing activity but the results provide confidence that STOA, in particular, satisfactorily simulated what transpired.

A.1.8 The Earth's Magnetosphere

In the absence of interplanetary plasma, the geomagnetic field would extend in all directions. The Earth is, however, located within the expanding solar wind which acts to compress and confine the geomagnetic field to produce a 'tear shaped' cavity around which the solar wind flows. This cavity is called the *Magnetosphere.* Fig.A.3 illustrates schematically the locations of its characteristic boundaries, which are briefly described below.

A collisionless *Shock Wave* is formed upstream of the magnetosphere due to the encounter between the supersonic solar wind and the 'obstacle' to its flow presented by the Earth's magnetic field. In this region, wave-particle interactions take over the role played by particle collisions in a collisional shock. Beyond it, the shocked solar wind forms a layer of turbulent plasma (the *Magnetosheath)* between the *Bow Shock* and the *Magnetopause.* This latter surface constitutes the boundary between the solar wind and the Earth's magnetosphere. Across the magnetopause, the magnetic field usually undergoes a sharp change in both strength and direction so that, by Ampere's Law, electrical current flows at this interface (the *Chapman-Ferraro Current*).

The solar wind does not readily penetrate the magnetopause but, rather, compresses the field around the Earth to form the magnetospheric cavity, the Sun-ward boundary of which is located where the

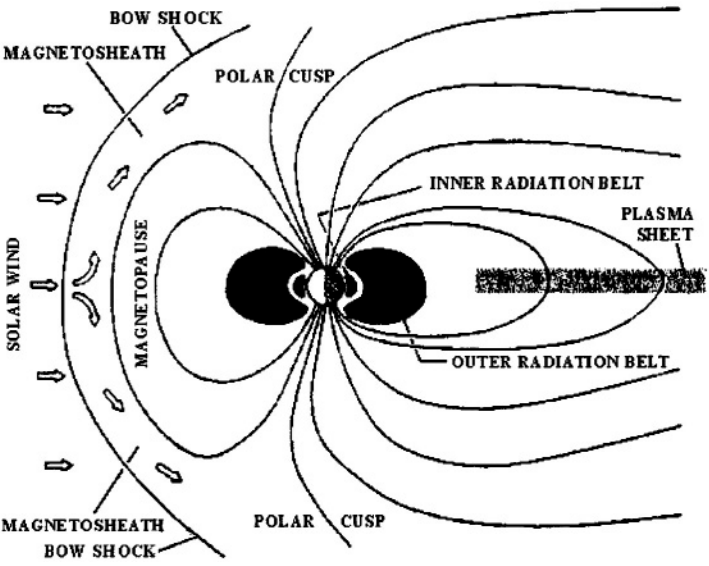

Figure A.3. Schematic diagram of plasma regions of the Earth's Magnetosphere viewed in the noon-midnight meridian plane

outward pressure exerted by the compressed geomagnetic field and the magnetospheric plasma pressure are mutually balanced by the pressure of the solar wind plasma. A complex process of inter-connection between the magnetic field of the solar wind and the geomagnetic field (not presently well understood) acts to stretch out the magnetic field in the anti-Sun direction to form a long *Magnetotail.* Under quiet conditions, the Sun-ward side magnetopause is located at $\sim 10\,R_E$ from the Earth ($1\,R_E$ is ~ 6 400 km) while the magnetotail extends anti-Sunward to a distance that may be in excess of $1\,000\,R_E$.

The magnetosphere effectively shields the Earth from most of the direct solar wind because charged particles do not readily travel across a magnetic field. In consequence, most of the shocked solar wind particles flow around the magnetosphere within the magnetosheath. Some solar wind plasma can, however, enter the magnetosphere. For instance, between the Sun-ward magnetic field and the tail-ward magnetic field are located two funnel-shaped regions called the *Polar Cusps* which the solar wind can access by following the local magnetic field lines towards the Earth.

Also, solar wind plasma is found in the magnetotail, although the

method of its entry to this region is a subject of some debate. In this regime the *Plasma Sheet* is sub-divided by a thin *Neutral Sheet,* where magnetic fields from the northern and southern hemispheres of the Earth effectively cancel to create a magnetically 'neutral' region. In the upper (northern) region of the plasma sheet the magnetic field is directed towards the Earth. In the lower (southern region) the field is directed away from the Earth. As long as the impingement of the solar wind on the magnetosphere is reasonably steady, the plasma sheet remains in equilibrium. When, however, coupling between the solar wind and the magnetosphere is enhanced due to solar circumstances, this balance is disturbed. Details of the processes involved are not well known. However it is generally envisaged that, when the IMF is oriented anti-parallel to the geomagnetic field lines, magnetic merging can occur between the interplanetary and geomagnetic fields at the dayside magnetopause and energy is associatively transferred from the solar wind to the magnetosphere. An associated increase in magnetic pressure in the magnetotail then causes the plasma sheet to become 'pinched' until a neutral point is formed. Here, energy stored in the magnetotail is explosively released by a re-connection process and plasma particles are resultingly accelerated so that they travel, on either side of the neutral point, towards the Earth and down the magnetotail respectively. These re-connection events are thought to occur at about $10 - 20\, R_E$ to tail-ward. The accelerated particles moving Earth-wards along the magnetic field lines bombard the upper atmosphere around the poles in regions called the *Auroral Ovals,* thereby stimulating auroral displays. Complementary, fast plasma flows down the tail showing speeds of several hundreds of km/s have been identified out to about $200\, R_E$ in observations made aboard the *ISEE-3* and *Geotail* SC.

On the occasion of an Earth gravity assist manoeuvre executed by ESA's *Giotto* SC on 2 July 1990, conditions were unusually quiet in the magnetosphere, thereby enabling a rare "snapshot" of this region to be recorded both in magnetic field and energetic particle measurements. These data, which show characteristic signatures of key magnetospheric regimes, including the bow shock; the magnetopause (inbound and outbound) and the midnight cusp are available in Glassmeier *et al.* (1991) — magnetic data, and in McKenna-Lawlor *et al.* (2001a) — energetic particle data.

All of the planets hitherto explored have magnetospheres. For planets that have no internal magnetic dynamo, the solar wind induces a magnetosphere through its interaction with the upper atmosphere and ionosphere. It is thus possible to distinguish between *intrinsic* and *induced* magnetospheres (Russell, 1991). Spacecraft are profoundly affected by disturbances in the terrestrial magnetosphere and display operational anomalies, and even failure, under particular perturbing circumstances (see Section A.2). Study of magnetospheres in other parts of the solar system where different boundary conditions and scale sizes pertain offer new scientific insights into magnetospheric processes, but also potentially constitute hazardous environments for investigating SC.

A.1.9 Particle Populations in the Magnetosphere

Key populations making up the near Earth energetic particle fluxes include Galactic Cosmic Rays (GCRs); particles trapped in the geomagnetic field forming the Van Allen Belts, and energetic particles associated with solar activity. Fig.A.4 and Fig.A.5 present typical energy spectra of these various populations.

A.1.10 Galactic Cosmic Rays

Galactic Cosmic Rays (GCRs) originate in our galaxy and they pervade the solar system. These particles constitute nuclei from all the known stable atoms as well as electrons (composition: 83 % protons; 13 % He nuclei; 3 % electrons and 1 % higher Z particles — with peaks at C, O, Ne, Mg, Si and Fe).

The omnidirectional flux of GCRs > 100 MeV/n is about four particles/$(\mathrm{cm}^2 s)$ at the time of minimum of the (approximately) eleven year solar cycle and about two particles/$(\mathrm{cm}^2 s)$ at the time of solar maximum. This modulation (see also Section A. 1.3) affects particles with momentum up to ~ 1.7 *GeV/c,* where *c* is the velocity of light. The annual GCR dose rates are 10 rads and 5 rads respectively at solar minimum and solar maximum (taking into account a slight related change in composition). The higher Z population is of particular interest with regard to the radiation damage produced by these particles (see Section A.2.1).

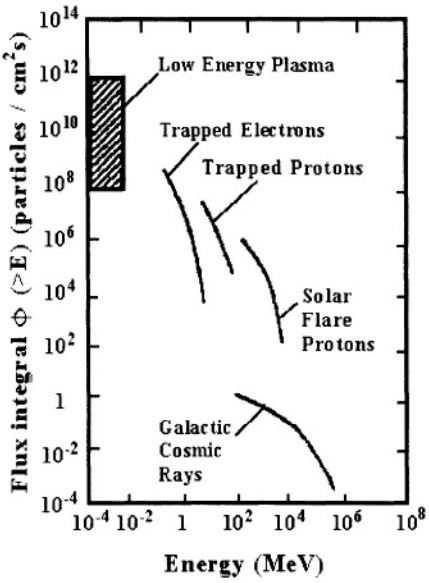

Figure A.4. Typical energy spectra of major populations of charged particles in the Earth's Magnetosphere

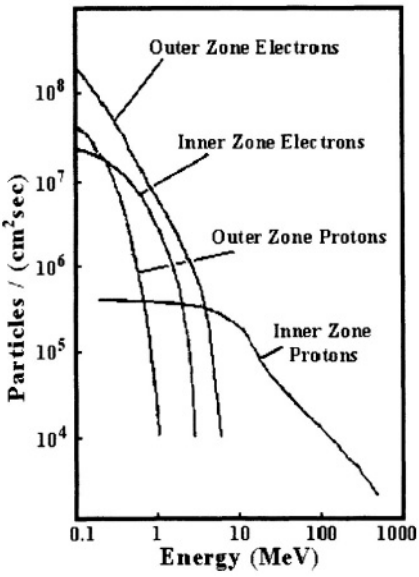

Figure A.5. Typical energy spectra of electrons and ions trapped in the inner and outer Radiation Belts

A.1.11 The Van Allen Belts

Early measurements indicated that energetic particles are trapped in two distinct belts in the geomagnetic field, the innermost centered at about $1.5\,R_E$ and the outermost at about $3.5\,R_E$. These are now called the *Van Allen Radiation Belts* to commemorate associated pioneering work by James Van Allen, in particular with regard to data recorded aboard *Explorer-1* and *Explorer-3* (Van Allen *et al.* 1958). Measurements made aboard NASA's Department of Defense SC *CRRES,* which was launched in 1990 at a time of solar maximum, showed that the region between the inner and outer belts may become filled, as a result of solar activity, with energetic protons that can remain trapped in this 'slot' region for many months (Dyer *et al.* 1995).

The *Inner Radiation Belt,* which in the equatorial plane is located at roughly $1.1 - 3.3\,R_E$, contains primarily high energy protons. Particles with energies > 10 MeV and maximum flux density $> 10^4 \mathrm{cm}^{-2}\ \mathrm{s}^{-1}$, are concentrated at radial distances of $\sim 1.6 - 1.9\,R_E$.

The source of these protons is the in-flight radioactive decay of neutrons generated in nuclear interactions produced by galactic cosmic rays in the upper atmosphere (*Cosmic Ray Albedo Neutron Decay* — GRAND). It is noted that this source would produce a radiation belt around a magnetized planet even if the solar wind did not exist. Extended measurements made aboard the *SAMPEX* SC showed that the inner belt is rather stable, while displaying a modulation over the general time scale of the solar cycle.

A population of particles located within the inner belt composed of heavy nuclei (mainly oxygen, with some nitrogen, helium and a small amount of carbon), is interpreted to have been produced by cosmic rays of interstellar origin (the *Anomalous Component*), see Cummings *et al.* (1993). Measurements of a representative atomic species (20 MeV/n Oxygen) made aboard *SAMPEX* over approximately seven years, indicate that the particle loss rates associatively observed are in conformity with estimates of the loss rates predicted to occur in a drift averaged model atmosphere under both solar minimum and solar maximum conditions (Li *et al.* 2001a).

The *Outer Radiation Belt* contains mainly electrons with energies ranging from several hundred keV to a few MeV. It is much more dynamic than the inner belt and exhibits changes on solar cycle, semiannual, solar rotation and diurnal time scales. Semiannual variations in the relativistic electron population were found to depend on a substantial, semiannual, solar wind speed increase associated with a pre-cursor solar wind density enhancement and, in particular, with a southward turn of the IMF (Blake *et al.* 1997). The outer belt is most intense, and extends further towards lower L values, at around the equinoxes (Baker *et al.* 1999). Also, it was long recognized that geomagnetic activity shows a semiannual variation with its greatest activity near the equinoxes (Chapman and Bartels, 1940).

The seasonal variation of geomagnetic activity was overviewed by Cliver *et al.* 2000 who cited three possible causes namely: the *Axial Effect* (that is the variation of the Earth in heliographic latitude and the associated increase in solar wind speed at higher heliographic latitudes); the *Equinoctical Effect* (that is the varying angle of the Earth's dipole with respect to the flow direction of the solar wind, which reflects a varying efficiency in coupling) and the *Russell-McPherron Effect* (that is the larger z-component of the IMF near the equinoxes in

GSM coordinates due to the then pertaining tilt of the dipole axis with respect to the heliographic equatorial plane, see Russell and McPherron, 1973).

Studies by Li *et al.* (2001a) using *SAMPEX* data indicate that, in the inner magnetosphere, the equinoctial effect is the primary contributor to the semiannual variation of relativistic electrons, and also of the accompanying variations in geomagnetic activity. A lesser role is played by the axial and Russell-McPheron effects. The semiannual variation of MeV electrons at geo-stationay orbit was attributed mostly to the semannual variation of solar wind velocity. Overall, it was inferred that the magnetosphere is strongly controlled by the solar wind.

An account of sporadic, rather than periodic, variability in the high energy, outer belt, electron fluxes is contained in Section A.1.16.

A.1.12 Particle Motion in the Geomagnetic Field

If an electrically charged particle is injected into the geomagnetic field with some initial velocity at an angle to the field vector, it will follow a helical trajectory about, and along, the magnetic field line (provided that its energy does not exceed the energy threshold for trapping). When it spirals towards a region of higher field intensity the particle experiences an induced electric field that accelerates its circular (cyclotron) motion transverse to the magnetic field line. An increase in the transverse kinetic energy of the particle resultingly occurs (due to conservation of energy) at the expense of its motion parallel to the field line, and a point may eventually be reached at which all the parallel kinetic energy has been converted into transverse energy. In this situation, the particle reverses its parallel motion and spirals back in the opposite direction. In this way in the Earth's field, a charged particle can bounce back and forth between two "mirror points" lying on opposite sides of the geomagnetic equator.

Since the terrestrial magnetic field is not uniform and its field lines are curved, the particle will, in addition, drift in longitude around the Earth. Particles with charge $q < 0$ drift eastward while ions with $q > 0$ drift westward (thus electrons drift eastward and protons and other ions westward, see Fig.A.6). A charged particle can remain trapped within a fixed 'drift shell' for a very long time. Early studies showed

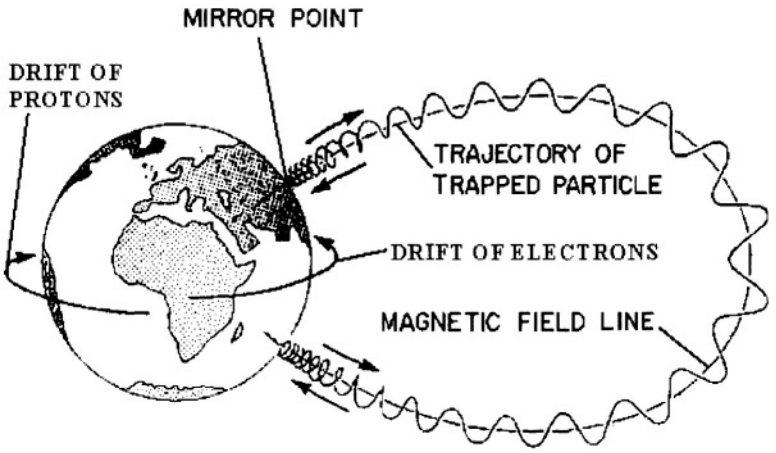

Figure A.6. Motion of a charged particle in the geomagnetic field

that geo-magnetically trapped ions are composed mainly of H^+, He^+ and O^+.

Radiation belt fluxes are usually mapped in a geomagnetic coordinate system (B, L). Here B is the magnetic field intensity at the point of measurement and L is a parameter defining the drift shell (B, L) of a trapped particle with pitch angle equal to 90° at the point of observation. For particles mirroring at this point, B and L are adiabatic invariants, uniquely determined in terms of the first and second adiabatic invariants of motion μ and I (McIlwain, 1961). For a particle not mirroring at the point of observation (i.e. having a pitch angle $\alpha \neq 90°$), the magnetic field intensity at the mirror point is $B_m = B/\sin^2\alpha$, and L_m is the L value determined by I_m, which in turn is obtained by tracing the field line passing through the point of observation to its conjugate mirror points. Both B_m and L_m are adiabatic invariants. Since the geomagnetic field distribution undergoes secular changes (Chapter 7), the (B, L) values of a fixed point in space change with epoch. At L values below about 3, this has important consequences for particle fluxes trapped at altitudes below about 2000 km.

A.1.13 The South Atlantic Anomaly

Since the center of the Earth's magnetic dipole is displaced with respect to the center of the Earth, particles trapped on the geomagnetic field lines in the inner radiation belt penetrate to lower altitudes over the South Atlantic Ocean than anywhere else on Earth. Two influences are involved in this effect. These are (a) the trapping of particles in the field of a titled dipole with significant contributions from the quadrupole $(n = 2)$ and octupole $(n = 3)$ terms and (b) the input of higher order terms (5,4), (6,5) and (4.3) which are associated with local, crustal, magnetic anomalies (Roederer, 1972).

The secular variation of the main dipole coefficient I_1^0 with time, at present a decrease in magnitude of about $1.6 \cdot 10^3$ nT per century, influences energetic protons trapped on low-L drift shells. These particles, whose characteristic lifetimes within the shells are of the order of centuries, are subject, due to conservation conditions, to a secular change in drift shell position and energy such that the L parameter of a given particle decreases as I_1^0 decreases. Also, a secular change in the expansion coefficients causes a slow variation in the position of the 'magnetic center' which drifts westwards from the Earth's center at a rate of between 2 and 3 km/year. The eccentricity of the main dipole thus increases with time. In consequence, particle shells are gradually dipping deeper into the atmosphere over the South Atlantic. The enhanced concentration of particles at this general location is referred to as the *South Atlantic Anomaly* (SAA). Fig.A.7 shows modeled proton fluxes with energies greater than 100 MeV in the SAA region. These concentrated fluxes extend above Brazil and South Africa where they constitute a radiation hazard to SC.

A.1.14 Magnetic Storms

A current, generated by the longitudinal drift of energetic (10–200 keV) charged particles trapped on field lines between $L \sim 2 - 7$ circulates clockwise around the Earth as viewed from the north (see also Section A.1.12 and Fig.A.6.). This current is termed the *Ring Current.* The global strength of the ring current can be monitored by ground based magnetometers at middle or equatorial latitudes.

A *Magnetic Storm* was defined by Gonzales *et al.* (1994) as 'an interval of time when a sufficiently intense and long-lasting interplane-

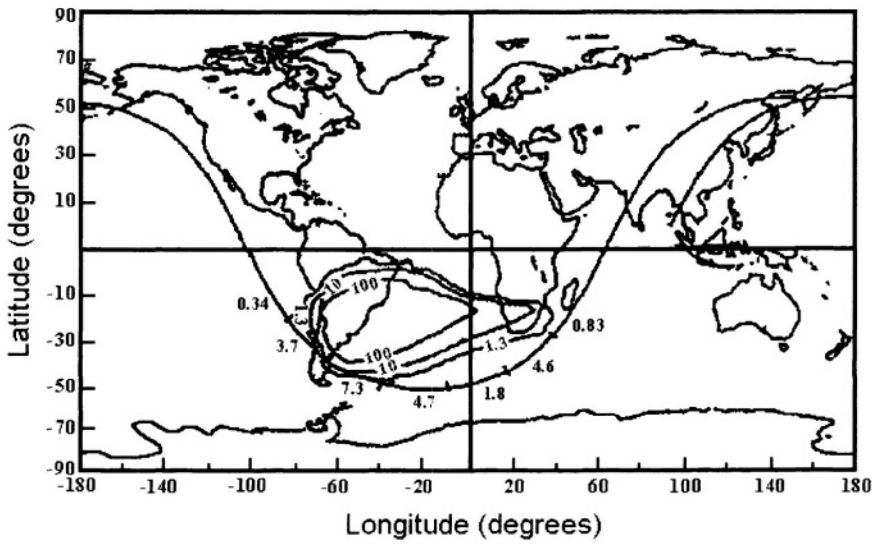

Figure A.7. The ground track during orbit 9 of the *STS-48* mission with respect to SAA flux contours. Isolines are in the unit $[protons/(cm^2 s)]$ with $E > 100$ MeV

tary convection electric field leads, through substantial energization in the magnetosphere-ionosphere system, to an intensified ring current strong enough to exceed some key threshold of a quantifying storm time index' (the electric field concerned derives from the solar wind velocity V and from the southward IMF component B_z). An index that is frequently used in quantifying the effect is the *Disturbance Storm Time Index* (D_{st}), obtained from hourly scalings of horizontal magnetic variations recorded by a network of, near equatorial, geomagnetic observatories. The characteristic accompaniments of elevated solar wind velocity and a strong, southward directed B_z component, suggests that enhanced energy transfer from the solar wind/IMF to the magnetosphere during these events includes a process of magnetic field merging.

Storms are typically divided into three distinct phases according to their signatures in D_{st}. During the *Initial Phase,* D_{st} is positively enhanced by amounts reaching peak values of approximately 100, 50 and 30 nT, which are respectively described as Intense, Moderate and Small (typically Sub-Storm) events. This Initial Phase can last from minutes to hours. During the *Main Phase* of an Intense Event (typi-

cally 30 minutes to several hours), D_{st} can attain negative values of hundreds of nT as the ring current builds up. Thereafter, the *Recovery Phase* can endure from tens of hours to more than a week as D_{st} gradually returns to its normal level.

Geomagnetic disturbances are also often described in terms of the K_p *Index,* which is the mean value of the disturbance level in the two horizontal magnetic components recorded at thirteen, selected, sub-auroral stations. The name K_p originates from *planetarische kennziffer* (planetary index).

Sub-storms are episodic events that release energy stored in the magnetosphere and magnetotail into the high latitude ionosphere, associatively creating intense auroral displays. They occur in association with most levels of magnetic activity but are more frequent and intense during magnetic storms. A 'classical' hypothesis that the ring current is enhanced via the injection of energized plasma sheet particles from the magnetotail to the inner magnetosphere during sub-storms is presently challenged and a more modern view is that the sub-storm expansion phase acts as the energy dissipation term, while the southward IMF constitutes the input term, in the energy balance equation (see for example Siscoe and Petschek, 1997 and McPherron *et al.* 1997).

The largest magnetic storms are often related to the occurrence of a CME accompanied by a fast shock and this combination is called an *Interplanetary Coronal Mass Ejection* (ICME), see Dryer (1994). Moreover, for geo-effectiveness, the characteristic configuration of a large, southward directed B_z component should endure for several hours. Not surprisingly, CMEs will usually not produce significant magnetic storms when the southward directed B_z is small; if it is directed northward or if the ICME source on the Sun is unfavorably far from the central meridian. Very energetic flares located at around central meridian can, however, be associated with major magnetic storms, as defined by their D_{st} and K_p indices (Gonzalez *et al.,* 1994, 1999).

Two different processes have been suggested for the formation of the storm-time ring current. These are (a) the injection of plasma into the inner magnetosphere during the expansion phase of magnetospheric sub-storms and (b) increased convective transport of charged particles from the night-side plasma sheet into the inner magneto-

sphere ($L < 4$), consequent on an intensification of the Earth's dawn-dusk convection electric field during extended periods of strong southward IMF. At the present time there is some evidence that favors the enhanced convection over the plasma injection model and it is increasingly inferred that sub-storms play a significant, albeit not a primary, role in the growth of the storm-time ring current.

A.1.15 External Magnetic Field Models

The geomagnetic field can be represented, to a first order approximation, by a magnetic dipole with its axis inclined to the Earth's spin axis by ~ 11.5° with some contributions from quadrupole, octupole and certain higher order terms (see Section A.1.13 and Chapter 21). The internal multipoles can be neglected beyond $L = 3 - 4$, and, in this region of space, even a centered dipole model provides a satisfactory approximation of the internal field. Beyond $L = 4$, the effect of external currents starts to play a role and this regime is called the *External Magnetic Field.* Three important current systems are present in the undisturbed outer magnetosphere namely, currents at the magnetospheric boundary and in the neutral sheet and the ring current (which flows in the equatorial, minimum induction, surface). During magnetic storms and sub-storms there are substantial changes in these current systems while field-aligned currents flow both out from, and into, the underlying ionosphere.

The external field has been investigated over several decades using SC and, based on these data, the average strength and shape of the field as it is configured during periods of enhanced magnetic activity has been modeled by several authors. In this regard, a frequently used series of empirical, global, magnetic field models was developed by N.Tsyganenko (for convenience the representative model described in Tsyganenko (1989) is generally referred to as T-89). This series is based on the *International Geomagnetic Reference Field* (IGRF) — which represents the internal contribution to the field — suitably chosen for the epoch concerned and modified by including the effects of the various magnetospheric current systems mentioned above.

To illustrate these models let (IGRF-2000) be represented by B_1 and the modified field, as determined in T-89, be represented by B_2. Denote $(B_2 - B_1)$ by ΔB. In Fig.A.8, $\Delta B/B_1$ is plotted against alti-

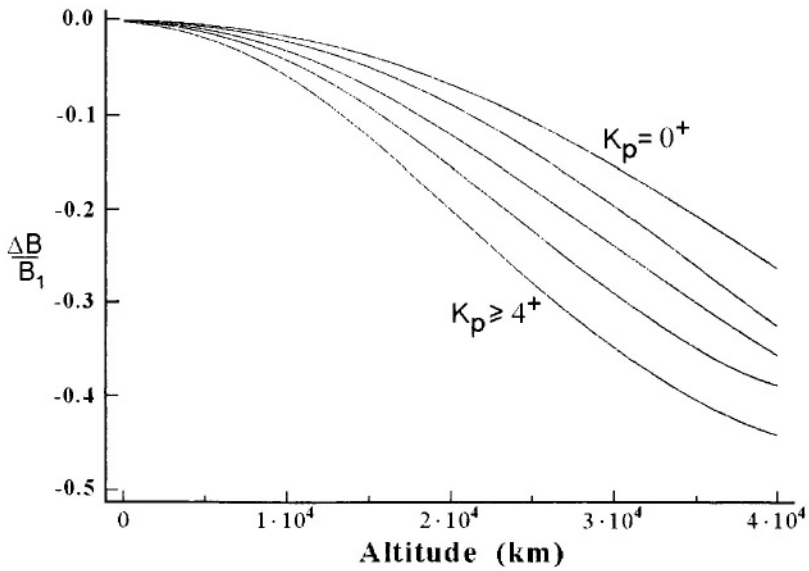

Figure A.8. Relative disturbances of the geomagnetic field with altitude at the geographic coordinates of the SAA (plotted using *SPENVIS* software)

tude at the geographic longitude and latitude of the SAA assuming five different environmental situations. The lowest curve in the plot corresponds to the most disturbed magnetic conditions (where $K_p \geq 4^+$) and the highest corresponds to quiet time conditions. These plots indicate that, at the altitude of geo-synchronous orbit, there is a difference of up to 40 % between the disturbed and undisturbed magnetic fields. Alternative external magnetic models include the Olson and Pfitzer tilt dependent model (Olson and Pfitzer, 1977) and the Olson and Pfitzer dynamic model (Pfitzer *et al.* 1988).

Lemaire *et al.* (1995) compared the modeled magnetic field obtained using T-89, as well as the two Olson and Pfitzer models mentioned above, with *in-situ* measurements made under disturbed magnetic conditions aboard the *CRRES* SC. These comparisons showed that none of the models was adequate to reproduce the measured fields over all parts of the orbits investigated. More recently, Zhou *et al.* (1997) tested the T-96-01 model against data from the *POLAR* SC measured at high altitudes above the polar cap and, again, found major discrepancies between the modeled and observed fields. More work on modeling the external field is thus presently required.

A.1.16 Relativistic Electrons

In addition to the 'periodic' variability described in Section A.1.11, dramatic day to day enhancements, showing slow as well as rapid buildups, are frequently found in the fluxes of the outer belt electrons. While in general (although not exclusively) these events are associated with geomagnetic storms, the magnitude of a particular enhancement can vary over a wide range for any given storm strength, attaining in many instances an increase of several orders of magnitude. The electrons can also display sudden 'drop-outs', such that the fluence decreases by 2 – 3 decades within 24 hours. These latter events may be due to magnetic changes in which the electron population is temporarily shifted out of geo-synchronous orbit. Alternatively, they may be related to sudden electron losses made either to the magnetopause or to the upper atmosphere.

Two classes of process have emerged to account for the observed enhancements in the outer belt, namely those that rely on internal acceleration or re-circulation mechanisms and those that rely on increased radial transport into the region. A general account of the several mechanisms proposed in these two categories is contained in Friedel *et al.* (2001). If it is supposed, as is popularly the case, that the variable high energy electrons originate in a population of lower energy electrons (from a few tens to a few hundred keV) already present in the outer radiation belt or, alternatively, injected into this region, the process whereby these electrons are subsequently accelerated to relativistic energies remains to be explained.

One possibility is that the enhanced fluxes of relativistic (MeV) electrons are driven by a drift-resonant interaction with the enhanced ULF wave activity observed during storm times. To test this suggestion, solar wind variations, relativistic electron fluxes and Pc5 ULF wave power data measured over a six month interval using two *STRV* micro-SC were inter-compared (*Pc5 pulsations* comprise strong, compressed, variations in the magnetic field magnitude, with typical periods of about ten minutes, which are commonly observed in the outer magnetosphere at low geomagnetic latitudes). These comparisons provide evidence that the largest relativistic MeV electron flux increases occurred in association with Pc5 wave power that was sustained over a number of days in response to, long lived, high speed solar wind

streams (Mathie and Mann, 2000). Also, it was noted in this study that the maximum available Pc5 energy was located at the outer edge of the radiation belt and decreased strongly with decreasing *L*. Although these results are encouraging, it is still appropriate in view of the great complexities involved, to seek to determine on a case by case basis for a large number of events the comparative roles of the various mechanisms that may, in principle, operate to produce individual relativistic electron enhancements, so as to elucidate which process is the most important under particular circumstances, and why (Friedel *et al.* 2001).

Theoretical aspects of relativistic electron dynamics have, meanwhile, yielded methods that can be used in predicting high energy electron enhancements. One such method is to take a simple physical diffusion model and modulate its diffusion coefficients using appropriate solar wind parameters. This procedure has already provided evidence that the processes concerned are predictable (i.e. non-chaotic) and based on a. low-dimension, solar wind input function (Li *et al.* 2001b). Another, method based on a statistical approach, looks for relationships between various input parameters (e.g. solar wind state, the D_{st} index, ULF wave activity) and the output relativistic electron fluxes. In this connection, a statistical study by O'Brien *et al.* (2001) yielded useful operational thresholds for predicting the effectiveness of a given storm in producing relativistic electron flux buildups at geo-synchronous orbit.

In addition, a statistical link identified between daily electron fluence and average solar wind speed prompted Baker *et al.* (1990) to suggest that the daily > 2 MeV electron fluence at geo-synchronous orbit can be predicted through employing a linear filter technique that uses average solar wind speed as its input. This method has since been refined to result in the *Relativistic Electron Forecast Model* (REFM) presently utilized at NOAA's Space Environment Center (SEC).

In this kind of forecasting, to help account for short-term 'drop outs' (see above) and longer term drifts, a flux offset is employed to adjust the output. This involves comparing previous selected measurements with relatively recent values to compute the offset. The number of days of previous measurements varies from one to twenty, and the predictions are made with lead times that range from one to three days, or (less reliably) eight days. A complementary method

based on forecasts that use a 'Quiet Sun' *Wang-Sheely Model*[§] rather than recorded solar wind data, is also used to provide electron fluence forecasts that cover up to 8 days.

Bombardment by high energy relativistic electrons can lead to anomalies in the operation of SC, or even to complete SC failure, see for example Regan *et al.* (1983) and Baker *et al.* (1987), as well as Section A.2.4. Against this background, and having regard in addition to the threat these high energy particles pose to astronauts on the International Space Station (Section A.3.9), relativistic electrons are presently routinely monitored by *NOAA* SC at geo-synchronous orbit.

A.1.17 Thermospheric Heating

The *Thermosphere* (at atmospheric altitudes between 120 – 600 km) is subject to large temperature variations (approximate range 800 – 1200 K) over a typical solar cycle as a result of the local absorption of *Extreme Ultraviolet Radiation* (EUV). At altitudes between 500 – 800 km the atmospheric density between solar maximum and solar minimum is increased by approximately two orders of magnitude. The expansion of the atmosphere in response to heating has a significant effect on SC lifetimes. The dominant effects in this connection are the level of solar activity in a particular cycle and the value of the ballistic coefficient of the SC concerned. A family of plots relating SC lifetime as a function of altitude with the solar cycle phase for various representative ballistic coefficients, was provided by Tribble *et al.* (1999)

A.1.18 Atmospheric Drag

When the Earth's atmosphere expands in response to solar activity, the drag experienced by SC in near Earth orbits increases. As already

[§]The Wang-Sheely empirical model determines the solar wind speed at the Earth from the divergence of magnetic flux tubes in a synoptic, map-based, potential field source, surface model of the coronal magnetic field. According to this model, fast solar wind emanates from regions of small magnetic field divergence while slow solar wind comes from high divergence regions.

shown in Chapter 6, a drag perturbing force that decreases SC energy causes the orbit to contract.

One of the most difficult tasks in evaluating and modeling the drag effect lies in determining appropriate values for the atmospheric density. There is at present a plethora of atmospheric models, divided between those that are *Static* and those that are *Time Varying*. Static models reflect latitudinal and longitudinal density variations. Attempts to model a symmetrical atmosphere to conform to every region on the Earth's surface are, however, frustrated by the presence of geographical features (for example mountain ranges) while the Earth's equatorial bulge (Chapter 7), in itself introduces a density anomaly.

Time varying models require that many short term, and long term, variations in the distribution of density in the Earth's atmosphere be taken into account. These are primarily associated with the changing position of the Earth in space and with levels of solar activity. Table A.1 lists some of the major contributors to the variations concerned. Other contributors include the effects of oceanic and atmospheric tides and winds and the influence of differential atmospheric rotation.

Among numerous mathematical models of atmospheric density developed since the late 1950s are:

- The Committee for Space Research (COSPAR) International Reference Atmosphere Models (CIRA) produced by The Committee for Space Research — COSPAR (Rees, 1988 and Rees *et al.* 1990). The CIRA-90 version covers altitudes from 25 – 2500 km.

- Jacchia Models. The 1977 version covers from 70 – 2500 km. (Jacchia, 977).

- Mass Spectrometer and Incoherent Scatter data (MSIS). The 1986 version covers from ground level to 2000 km (Hedin, 1986).

- The Russian Density Model (COST). The 1999 version covers altitudes from 120 – 1500 km (Vallado, 2001).

A study by Gaposchkin and Coster (1987) concludes that no one density model is universally applicable. Thus, when choosing a model for a particular application, the accuracy needed (which is a function

Table A.1. Significant contributors to density in the Earth's atmosphere

Effect	**Origin**
The Diurnal Variation	A pronounced atmospheric bulge (density maximum) which lags behind the solar direction (where the atmosphere is warmest). Variations over a year in the Sun's declination cause the location of this feature to depend in a complex way on latitude, local time and on the time of year. A complementary minimum in density in the Earth's atmosphere is approximately juxtaposed to the bulge.
Twenty Seven Day Variation	The rotation period of the Sun.
Semiannual/Seasonal Variations	Temporal variations in the Sun-Earth distance/ temporal variations in the Sun's declination.
Eleven Year Variation	Temporal changes in the electromagnetic and particle radiation impacting the Earth which also varies from one solar cycle to another.
Cyclic variation	An eleven year cycle that lags behind the sunspot cycle by a few years and varies from one cycle to another.
Magnetic Storm Variations	Changes in the geomagnetic field associated with solar activity.
Irregular Short Period Variations	Various effects such as transient geomagnetic disturbances.

of the computational requirements) relative to a reasonable speed of calculation for the task concerned, should be taken into account.

A dramatic case of orbit decay due to atmospheric drag was provided by NASA's *Skylab* orbital station, which was launched to study the active Sun in 1973. The high level of activity under investigation resulted in greatly enhanced levels of drag due to atmospheric heating. In consequence, *SKYLAB* re-entered the Earth's atmosphere in 1979 considerably in advance of its expected demise, and before an attempt to deliver, using the *Shuttle,* a propulsion module to boost its altitude could be made.

A.1.19 Solar Radiation Pressure

The total solar energy striking $1\,\mathrm{m}^2/\mathrm{s}$ at the Earth's mean distance from the Sun is traditionally referred to as the *Solar Constant* and it has the value[¶] $I = 1\,353 \pm 20\,\mathrm{W/m}^2$. Due to effects relating to the motion of the Earth around the Sun, Wertz (1978) provided a time varying formula instead of a constant value to account for variations recorded at the Earth over the course of a year

$$I(t) = \frac{1358}{1.004 + 0.0334\cos(D_{aphelion})}\ [\mathrm{W/m}^2]$$

where $D_{aphelion}$ is the number of days from when the Sun is at aphelion (this point is usually reached on July 4, but with some variation).

Solar Radiation Pressure is a significant source of attitude and trajectory errors for high altitude (> 1 000 km) and for interplanetary SC. It depends on the reflective characteristics of the surfaces on which this radiation impinges; on the influence of surface shading during different parts of the orbit and on general aspects of orbital position and orbital inclination with respect to the Sun. Since incident photons

[¶] Detailed studies of the *Solar Constant* were made using data from the Solar Maximum Mission. These measurements show variations in the energy received from the Sun of about 0.01% that lasted from days to weeks and were correlated with the passage of sunspots across the solar disk. Superimposed on these variations, a long term decrease of about 0.018% per year was interpreted to be related to a cycle of activity on the Sun with a period longer than the 22 year magnetic cycle (Wilson *et al.* 1984). It is noted that a change in the solar constant of 1% would change the average temperature of the Earth by some one or two degrees *C.*

exert a net force on each segment of a SC, the resulting solar radiation force may be considered to pass through a 'center of pressure' (CP). If this is not coincident with the center of mass (CM), a solar radiation torque is produced.

When radiant energy E falls on a surface, that surface is subjected to a force per unit area. The incoming photons during an interval of time Δt have an equivalent mass m, so the equivalent increment of momentum ΔH may be expressed as

$$\Delta H = \frac{IS\Delta t}{c}$$

where S is the surface area. The magnitude of the force F acting on the surface is then

$$F = \frac{\Delta H}{\Delta t} = \frac{IS}{c} = PS$$

where $P = I/c$ is the radiation pressure impinging on an energy absorbing surface. For a totally absorbing surface (black body) $P = 4.5 \cdot 10^{-6}\,\mathrm{N/m^2}$. For a perfectly reflecting surface (mirror) $P = 9.0 \cdot 10^{-6}\,\mathrm{N/m^2}$.

At the surface of a SC, radiant energy may be partially absorbed and partially reflected. Reflection can occur in two ways namely, the incoming energy may be reflected specularly (so that the angle of incidence with respect to the normal to the surface is equal to the angle of reflection), or it may be reflected diffusely. In the latter case, the radiation leaves the surface so that the intensity in any direction is proportional to the cosine of the angle between the surface normal and the direction of reflected radiation.

As in the case of the acceleration due to drag (which is analyzed in Chapter 6), the acceleration due to solar radiation pressure alters the energy, and consequently the semi-major axis, of a SC orbiting the Earth. During that part of the orbit where the SC approaches the Sun, the energy and semi-major axis decrease whereas, when the SC moves away from the Sun, both of these values increase. For an orbit that is completely in sunlight, the change in the semi-major axis over one orbital revolution is periodic and averages to zero. If a SC goes into Earth eclipse averaging is not maintained and a correction should be employed when making orbital determinations. Since the disturbing force is proportional to the area to mass ratio of each particular SC, those structures with a large value of this ratio are subject

to a relatively large perturbation. In this connection *ECHO-1,* the first communications SC (launched in 1960) which comprised a sphere 30.5 meters in diameter made of mylar polyester film designed to act as a passive communications reflector, also provided a suitable object for studying the effect of solar radiation pressure. Early observations showed a decay in perigee height of approximately 3.5 km/day from the initial 1 852 km altitude circular orbit of this SC.

A.1.20 Solar Sailing

The very small force that results from solar radiation pressure can be used to propel a SC using a specularly reflecting sail. Technical studies have demonstrated in this regard that, by building a SC with a large area to mass ratio and locating the center of mass near to, or in the plane of, such a sail, guided energy transfer can be used to navigate the SC to a given target plane. Also, it was demonstrated that rather sophisticated manoeuvres are associatively possible.

For applications involving inner planet encounters, the considerable problem must be solved of constructing a SC that can withstand severe ambient heat while, at the same time, accommodating a light weight sail with a large surface area. The advantages of such a difficult development lie in the saving of propellant that could be achieved, and in the fact that sensitive onboard instruments would not be contaminated by thruster plume emissions. An early study of how to optimize solar radiation pressure for space flight was carried out by Van der Ha and Modi (1979) and designs based on this technology have been proposed by many authors for missions to relatively nearby targets (e.g. Mercury and Halley's Comet). Solar sailing has not yet been put into practice because, in particular, of the difficult problem of deploying the very large sails required.

A.1.21 Perturbing Effects on Orbiting SC

A summary of the various perturbing influences that can be exerted, under the same environmental conditions, on a particular orbiting SC is provided in Fig.A.9 (following Fortescue and Stark, 1995). In this figure, for each effect, the logarithm of the disturbing acceleration normalized to one gram of the SC mass is plotted as a function of altitude.

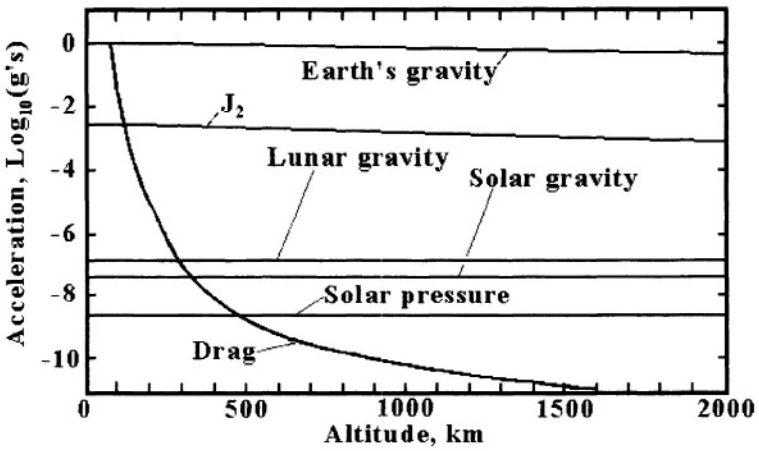

Figure A.9. Overview of perturbing influences on an orbiting SC (reproduced by courtesy of J. Wiley)

Conclusions from these comparisons are as follows:
— Below ~ 100 km, re-entry conditions prevail and atmospheric drag predominates.
— At altitudes below ~ 200 km, atmospheric drag is a significant factor in determining the life of a SC.
— Above 200 km, the effect of the Earth's oblateness (harmonic J_2) is second only to that of the gravitational field.

The relative magnitudes of solar radiation pressure and atmospheric drag are, according to Fig.A.9, nearly equal at 500 km, and this approximate level is frequently quoted in the literature as a transition altitude between the dominance of drag and solar radiation pressure effects. It should be remembered, however, that drag depends on the level of solar activity and curves representing other environmental circumstances would show different results. Further, only the magnitudes of the perturbing accelerations are taken into account in the comparisons presented. To determine in practice the mutual effects of individual perturbations on a particular SC, equations involving the accelerations concerned must be examined individually and detailed calculations made.

A.1.22 Disturbance Free SC and their Applications

SC that are controlled to follow a trajectory determined only by external gravitational forces (a geodesic) are referred to as *drag-free* SC. The key supporting technologies required are gravitational reference sensors and micro-thrusters. A gravitational reference sensor consists of a proof-mass freely floating in an evacuated cavity near the SC's center of mass. The housing of this cavity is attached to the SC and a sophisticated measurement system (based on a capacitance bridge arrangement) to determine the position of the proof-mass with respect to the housing is provided. If, for example, solar radiation pressure impels the SC towards the proof-mass, the associated change in position is sensed and the micro-thrusters activate to perform a correction. By this means, the influences of both solar radiation pressure and drag can be counteracted so that the proof-mass tends to follow the orbit created by the gravitational field only. In order to validate that the proof-mass is following such a geodesic, its trajectory must be compared with that of a suitable reference object. Drag-free control was first demonstrated in 1972 aboard the *Triad-1* SC of the U.S. Navy's Navigation Satellite System TRANSIT.

Fundamental physics experiments in space typically involve the measurement of very small effects. In consequence, exploitation of the full sensitivities of the technologically advanced detectors used in these investigations can only be achieved by eliminating disturbing environmental factors, as the following examples illustrate.

It can be inferred, based on Einstein's theory of *General Relativity*[||] ,

[||]Whereas Newton perceived gravity as an attractive force acting simultaneously at a distance across space to affect all bodies, Einstein in his *Theory of General Relativity* proposed that space, time and matter are not independent of each other, as Newton had assumed. Rather, he postulated that gravity acts not as a force but as a field that warps space and time around massive bodies such as stars and planets. The elliptical orbits of the planets around the Sun are, thus, straight lines through curved space-time.

Although Newton's and Einstein's theories are very different, their results with regard to the material presented in the preceding chapters of this book are identical. However, General Relativity predicts effects the detection, or non detection, of which at sufficiently high levels of accuracy can be used to test the validity of the theory itself. With present day understanding of the near Earth environment and the application of sophisticated technological advances, the problem of providing

that a rotating massive body drags space and time around with it (the *Frame Dragging Effect*). In consequence, a gyroscope in a polar Earth orbit at 650 km is predicted to turn with the Earth through an angle which, after one year, amounts to 42 milliarc-seconds (there are 3600 arc-seconds in a single angular degree). Further, according to Einstein's theory the rotating Earth causes a warping of space-time. In consequence, a change in the spin direction of an orbiting gyroscope is predicted to occur as a result of its motion through space-time curvature. This is called the *Geodetic Effect.* In the particular case of the orbit described above, a rotation in the orbit plane of 6,600 milliarc-seconds per year is expected. The philosophy of the *Gravity Probe B* Mission (GPB) is to fly a set of gyroscopes in the 'test orbit' mentioned above to check the validity of the Frame Dragging prediction to an accuracy of 1% and of the Geodetic Effect prediction to one part in *10 000.*

Calculations show that, in the test orbit, drag from residual atmospheric gases coupled with the effect of solar radiation pressure would produce spurious accelerations that approach 10^{-7} g. However, through use of a proof-mass, the mean acceleration on the gyroscopes can be reduced to 10^{-10} g. By rolling the SC (one revolution every ten minutes) about a guide star, disturbing forces due to residual gases, as well as errors in the gyroscope and telescope readouts can also be significantly reduced. The use of super-conducting magnetic shielding eliminates the possibility of a disturbing Lorentz force arising from motion through the geomagnetic field.

Galactic cosmic rays, solar protons and geo-magnetically trapped protons with energies in excess of ~ 70 MeV could penetrate the SC structure, depositing energy (heat), momentum and electrical charge on the proof mass. Using proton flux models (Section A.2.2) and the GEANT/CERN Monte Carlo radiation transport code (Brun *et al.* 1984)** an attempt was made by the GPB Team to quantify these

---

experimental evidence for important, but hitherto untested, predictions of General Relativity can now be addressed using space borne experiments potentially with profound implications for modern physics.

**In cases where more than just the absorbed radiation dose should be known, the powerful software tool *GEANT* is used which employs, for a specific geometry, a particle-transport method that takes into account such effects as back-scatter, the production of secondary particles etc. Its uses are varied and include, in addition to determining the charge deposition on proof-masses, the determination of

effects. It is, however, difficult to assess the absolute accuracy of the *GEANT* simulations, while the particle models themselves may only be accurate to within a factor of about two at the energies concerned. It can, nevertheless, be deduced that electrostatic charging of the proof-mass can present a particular problem. In these circumstances, a fault tolerant proton detector with the capability to provide ongoing monitoring of energetic particles in the range 50 – 500 MeV was included on the SC (McKenna-Lawlor *et al.* 2001b), so that the measured data can be interpreted relative to ambient energetic particle measurements.

Another representative experiment in fundamental physics is provided by proposals to test the *Equivalence Principle* (EP) in space (the EP asserts that all objects, irrespective of their composition, experience the same acceleration due to the force of gravity). Although no practical limit at which the EP violation should occur has been established, a goal to test it to one part in 10^{18} is presently deemed scientifically desirable and achievable. In addition to employing drag free technology, this requires the development of an analytical method to estimate the effects of various local orbital disturbances (in particular that due to the Earth's oblateness) on the measurement sensitivity. The approach adopted in this regard is to extend Hill's equations to appropriately account for the influence of the Earth's shape, as well as to assess the spectral characteristics of gravity-gradient disturbances, on the SC. Gravity gradient parametric excitation of the proof-masses is also taken into account.

Other gravity related missions include a planned series of 'separated SC' missions. These are based on the philosophy that the movement of massive bodies in the galaxy produces disturbances in the local gravity field of the solar system that are detectable through the minute shift associatively produced in the distance between a pair of SC in an orbit high above the Earth. By means of such experiments it is hoped to be able to study aspects of the universe that are undetectable using conventional visible light and radio astronomy techniques.

These brief accounts of new generation experiments underline the need to develop appropriate techniques to neutralize various disturbing features of the terrestrial environment in which the associated

---

background particle levels in detectors, the responsiveness of detectors to thermal heating, the shielding effectiveness of honey-comb structures etc.

measurements are made

A.2 The Environment and SC Performance

Conditions on the Sun as well as in the solar wind, magnetosphere, thermosphere and ionosphere influence the performance and reliability of SC. For example a SC can malfunction due to the fact that its constituent electronics suffer radiation damage. Also, SC can display anomalous behavior and failure following plasma induced charging. These, and other, on-board responses to the ambient environment are next discussed and outline suggestions provided in respect of mitigating strategies that can be employed.

A.2.1 Electronics and Energetic Particle Radiation

An accurate knowledge of the Earth's particle radiation environment is of special interest to design engineers and SC users due to increasing requirements to mount in advanced scientific experiments highly sophisticated electronic devices that are sensitive to particle radiation. Over optimistic models could lead to failure whereas too pessimistic models would lead to the use of unnecessarily heavy SC shielding or the choice of an over costly mission trajectory.

When a single particle causes a malfunction in common electronic components (e.g. in random access memories, micro-processors, HEXFET power transistors etc.), the effect produced is called a *Single Event Effect* (SEE). These are of three main types *Single Event Upsets* (SEUs); *Single Event Latch Ups* (SELs) and *Single Event Burnouts* (SEBs).

Single Event Upset (SEU); An SEU is a soft error consisting of bit-flips, with no preference between 1 to 0 and 0 to 1 transitions. SEUs are generated by galactic cosmic rays, as well as by the heavy ions produced in flare events. Particularly potent are ions in the CNO group and those with maximum energies of 2 MeV/n or higher. Protons or other ions can also cause SEUs indirectly when they create

particles featuring high *Linear Energy Transfer* (LET) via nuclear reactions near to, or within, the sensitive volume of an electronic device. SEUs can corrupt a data register, thereby causing a control parameter to be altered, but they do not damage the component concerned or interfere with its subsequent operation.

Single Event Latch Up (SEL); Single high Z-particles can 'turn on' part of an integrated circuit to produce a, so-called, Latch Up. Once in this state, the electronic system is no longer controlled in the way the designer had planned and it must be powered down and switched on again to enable control of the circuit to be regained. This unwanted configuration may draw enough power through the chip to damage it, and even to associatively effect the power supply. Latch Ups can, in addition, be produced due to the application of an over voltage, or as a result of a sudden flash of X-rays (Gamma dot events).

Single Event Burn Out (SEB); In this case the device fails permanently as a result of radiation damage. SEBs include burnouts of n channel power MOSFETs, gate rupture, frozen bits and noise in charge-coupled devices. An SEB may be triggered in a power MOSFET biased in the OFF state when a heavy ion passing through it deposits enough charge to turn the device on. SEB susceptibility decreases with increasing temperature. A power MOSFET may undergo *Single Event Gate Rupture* (SEGR) due to localized dielectric breakdown in the gate oxide, thereby producing an SEB. Destruction of bipolar junction transistors can also occur. *Single Event Dielectric Rupture* events (SEDRs) occur in CMOS and are similar to the SEGRs observed in power MOSFETs.

Fig.A.10 presents a world map of SC locations where SEUs were recorded in memories located onboard a polar orbiting satellite (at 700 km). These data show a strong clustering of events in the South Atlantic Anomaly and a preferential distribution of the remaining upsets at high latitudes. This pattern reflects the fact that upsets in the SAA are generated by energetic particles trapped in the geomagnetic field whereas high latitude upsets are produced by cosmic rays and solar related events.

SC components can be manufactured to withstand high total doses of radiation. *Survivability* is the ability of a space system to perform its intended function in a stressed environment. *Hardness* is an attribute defining the environmental stress level which a system can survive. For

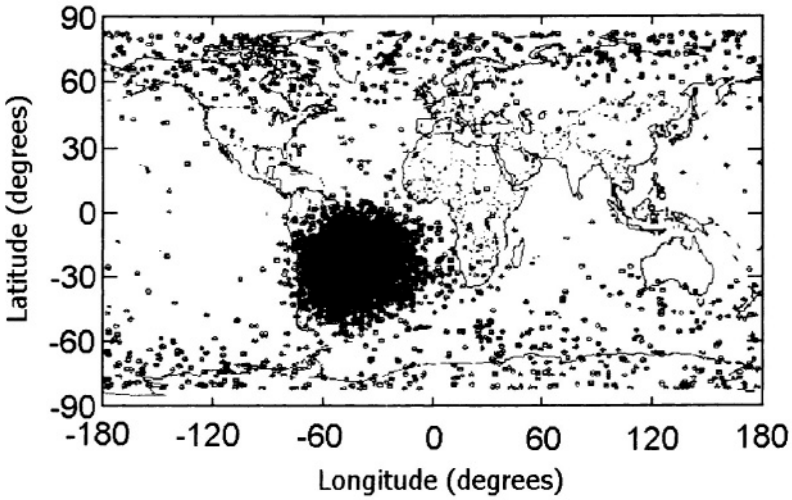

Figure A.10. A World Map of locations where SEUs were recorded

example, a space system that can survive the absorption of 10^7 rads of total dose is said to have a 'hardness' of that amount. In designing SC electronics a *Radiation Hardness Assurance Control Plan* (RHACP) should be followed to suit the mission profile. Various engineering techniques can be employed to enhance the survivability of a particular design (see for example Nordin and Kong, 1999).

A.2.2 Models of near Earth Energetic Particles

The first empirical models of the Earth's radiation environment were designed at the American Aerospace Corporation and later at NASA (Vette, 1991a). The latest proton models in this series are AP-8 MIN and AP-8 MAX which correspond, respectively, to minimum and maximum conditions of solar activity. Complementary electron models are AE-8 MIN and AE-8 MAX (Vette, 1977, 1991b). A series of trapped radiation belt models based on these NASA models was developed using data from Soviet SC by the Institute for Nuclear Physics of the Moscow State University (INP/MSU). The first of these was called the Cosmos Model-82. Later INP models are described by Getselev *et al.* (1991).

Since these models were built, the SAA has shifted considerably to the west due to the secular variation of the geomagnetic field distri-

bution (Chapter 21). This, and other, effects were analyzed and the performances of the Soviet and NASA models inter-compared in ESA sponsored studies carried out by Lemaire *et al.* (1991, 1995). It was associatively found that both the NASA and INP trapped radiation models produce similar fluxes in the bulk of the Earth's radiation belts. There are, however, significant differences between the predictions of these models in regions of steep gradients.

A number of new radiation belt models were more recently developed for ESA based on data from several SC experiments (including AZUR/EI-88, SAMPEX/PET, UARS/PEM, CRRES/MEA and ISEE-1/WIM). These new models were inter-compared, where this was appropriate, and also directly compared with the standard NASA models by Lemaire *et al.* (1995).

A.2.3 Radiation Models for Mission Evaluations

To allow estimates to be made of the total dose expected to be absorbed by a new SC and its payload elements in different orbits and at different times of the solar cycle, access to all the models is conveniently provided by *The SPace ENVironment Information System Project* (SPENVIS). This is a software tool prepared for ESA that allows, among other possibilities, the generation of a desired SC trajectory on a geomagnetic coordinate grid to facilitate the calculation of

— trapped proton and electron fluxes and solar proton fluxes;
— radiation doses (ionizing and non-ionizing);
— damage equivalent fluxes in Si and GaAs solar panels;
— LET spectra and single event upset rates;
— trapped proton flux anisotropy
— atmospheric and ionospheric densities and temperatures.

The way radiation environment models are used in planning new ESA missions is next described since this procedure is broadly representative of how new mission planning is carried out in all of the world's space agencies.

The initial stage of the analysis involves constructing a general model of the radiation environment that will be experienced aboard the planned SC. In this connection, for solar cells, the software tools *EQFRUX* and *EQFRUX.GA* (for silicon and gallium arsenide cells

respectively), provide damage equivalent 1 MeV electron fluxes. This assists both in sizing the SC solar array required and in determining an appropriate cover glass thickness. Since cell performance progressively declines in the space environment, it is necessary to ensure in advance that the power provided by the solar cells flown will, at the end of life (EOL), still be sufficient to operate the SC.

The software tool *SHIELDOSE* which provides curves displaying dose as a function of shielding thickness, allows minimum standards for electronic components to be established. Based on these curves, suitable components can be selected and items identified that require shielding additional to that provided by the SC. Components that it is essential to fly which have particular dose requirements are then subjected to an intensive shielding analysis. The software tool *ESABASE/DOSRAD,* which uses ray-tracing and sectoring analysis techniques to obtain an evaluation of the total dose at onboard locations of interest, is associatively employed. This approach is appropriate where spot shielding can be introduced. However, it does not take into account the effects of electron and bremsstrahlung angular scatter; secondary radiation or background particle levels. In cases where more rigorous analysis is required, the GEANT Monte-Carlo particle transport method can be employed (see also Section A.1.22).

To evaluate trapped particle spectra and solar proton spectra with respect to an arbitrary set of orbital parameters, the *UNIRAD SW* package is used. First, an orbit generator and a geomagnetic field model are availed of to convert the orbital parameter set to, idealized, magnetic dipole co-ordinates. These co-ordinates are then employed in concert with trapped particle models to obtain omni-directional fluxes along the SC trajectory and to compute orbital average particle flux spectra.

UNIRAD also incorporates the King and JPL Models (Section A.1.6) to allow (within their inherent individual limitations) the solar proton spectrum for a particular mission to be calculated (while duly taking into account the effect of geomagnetic shielding). Finally, the *CREME* package provides a comprehensive set of cosmic ray and flare ion LET and energy spectra (while also incorporating the contributions of geomagnetic and material shielding). When combined with radiation test data, upset cross-sections and LET thresholds for electronic components, this software can be used to estimate the SEU

and SEL rates required for component selection. A trapped proton spectrum can be included with the cosmic ray LET spectrum to also give SEU and SEL rates for the environment. The incorporation of interplanetary weather indices (which in some cases include solar proton fluences), allows both 'worst case' and 'nominal case' predictions to be made.

Mission objectives determine the type of orbit selected. The environment in GEO (which is generally used by communications SC) is already well understood. In this regime missions are generally long lived, thereby averaging out the effect of transient events. The predicted mission dose is consequently both straightforward to estimate and inherently reliable. On the other hand, missions flown at low altitudes (for example for Earth Observations) where the particle fluxes change rapidly and the secular drift of the SAA is appreciable, require the environmental models to be very carefully interpreted for each individual application.

Highly Eccentric Orbits (HEO) which are favored for astronomical investigations demand, since they evolve considerably during a typical mission lifetime, that studies be made of many orbits with respect to their individual peak particle flux spectra, mission average spectra, total mission dose and solar cell damage fluences. Also, since orbits of this kind extend beyond the regions covered by AP-8 and AE-8 (namely along the magnetotail and into interplanetary space), models based on data from relevant environments (e.g. from ISEE) should be associatively used.

Lunar missions and interplanetary missions usually involve Earth Flybys. The short duration of the associated radiation belt passage makes the absorbed dose sensitive to the pertaining state of the belts. Since transient conditions are not catered for by AP-8 and AE-8, a model based on *CRRES* data (CRRESPRO) which was designed for use during both quiet and active proton belt conditions can be suitably employed.

A.2.4 SC Charging

The plasma environment of an orbiting SC depends on where that SC is located, and also on its own gaseous products. In the Earth's environment, plasma and extreme ultra violet (EUV) radiation are

the major sources of SC charging currents. In its particular ambience, a SC can accumulate charge until an equilibrium condition is reached in which the net current is zero (i.e. when the electron and ion currents from the plasma to the spacecraft, the secondary electron current, the back-scattered electron current and the photo-electron current mutually cancel). EUV created photo-electric emissions are usually dominant in geo-synchronous orbits. The density of the ambient plasma also affects the speed of SC charging. In this connection a tenuous plasma (less than 1 particle/cm^3) will charge a SC and its surfaces more slowly than a 'dense' plasma (characterized by thousands of particles per cubic centimeter).

Terrestrial space plasma environments can be roughly divided into two regimes, namely those around Geo-Synchronous Orbit (GEO) at ~ 35 000 km and those around Low Earth Orbit (LEO) at ~ 100 – 1 000 km. At geo-synchronous orbit the plasma is hot, tenuous, highly variable and responds strongly to individual solar events. In LEO, the plasma is dense, cool and rather stable.

In GEO, the equilibrium potential of a SC is typically of the order of a few volts positive, as a result of the balance between photoemission and charged particle collection. However, during magnetic storms, electrons impinge on the SC with energies of several tens of keV. If the SC then enters eclipse so that the photoelectron current vanishes, its surface can rapidly charge up to a negative potential of the same order of magnitude as the energies of the impacting electrons. Such charging can occur differentially (i.e. different parts of the surface can attain significantly different potentials), thereby leading to arcing. This can cause physical damage to SC materials, degradation of solar arrays and generation of electromagnetic interference (onboard noise). Operational anomalies, or even overall failure, can ensue in these circumstances.

Data secured by the research SC *SCATHA* (Spacecraft Charging at High Altitudes) which was launched in 1979, supplemented by complementary data from other experiments, allowed the development in the U.S. of a computer code for the prediction and analysis of charging effects (the NASCAP–NASA Charging Analyzer Program). NASCAP is able to deal with complex SC shapes and with a range of possible materials (Katz *et al.* 1983). Also the code is applicable under a number of different conditions (e.g. it takes into account that, if a

SC is located near local noon, evening, midnight or dawn the ambient conditions experienced will be significantly different and result in a markedly different charging history). In addition, comprehensive guidelines for designers concerning the assessment and control of SC charging in GEO were drawn up by Purvis *et al.* (1984).

In LEO, low inclination orbits provide a relatively benign environment as far as SC charging is concerned. In this regime, a spacecraft can generally be regarded as closely coupled to the dense surrounding plasma and surface potentials of, at most, a few volts are typically displayed. In polar orbits at high latitudes although the plasma densities are comparable, intense fluxes of precipitating energetic electrons (5 – 10 keV) associated with the production of the aurorae can, in addition, be present. When these precipitating particles impinge on a SC differential surface charging can result.

In addition, SC motion relative to the background plasma plays a role since, in LEO, a SC moves with a velocity that is in excess of the velocities of the ambient thermal ions but lower than the velocities of the ambient electrons. A plasma wake characterized by ion depletion is accordingly formed behind the SC and, in the case of a structure where significant ram induced ion current collection occurs at the same time that a large ion-free wake is created, problems due to differential charging can become significant. In a worst case scenario, if a large, energetic, precipitating electron current impinges on wake related surfaces at a time when photo-emission is blocked due to the SC being in an eclipsed condition, a catastrophic discharge can occur. The consequences of such discharges range from transient operational anomalies to overall, permanent, failure. Also, physical damage to materials and deleterious effects such as sputtering and the attraction to the SC surface of chemically active species can ensue.

Investigations of charging in polar orbit at altitudes of about 840 km and inclinations around 99° were carried out using SC of the Defense Meteorological Satellite Program (DMSP). The data secured suggest that vehicle size is an important factor with reference to charging. General overviews by Stevens and Kirkpatrick (1986) and by Martin *et al.* (1990) discuss key circumstances that influence: environmental interactions between a SC and its surroundings; the various interactive mechanisms involved; their consequences for particular SC and techniques that can overcome the adverse effects concerned (see also

Section A.3.1). It is noted that most of the deleterious effects are increased if a high voltage power system is included in the SC design. Also, that damage to surfaces and cables produced due to impinging micrometeoroids and space debris can have a significant influence on arcing thresholds.

The computational codes developed in the U.S. for GEO (see above) were later modified and extended to make them applicable to LEO (NASCAP/ LEO). Also, a 3-D code (POLAR) was developed to predict the ion and electron structure in those plasma wakes generated by large SC (Katz *et al.* 1985). For ESA, the Spacecraft/Plasma Interactions and Electromagnetic Effects Program (ESPIRE) was developed to provide a design resource in relation to the various problem areas related to SC – plasma interactions. A feature of this development was the emphasis placed on validating experimentally the design related engineering tools produced. ESPIRE comprises a suite of programs individually capable of analyzing, on the basis of a specific, but restricted, set of assumptions a particular aspect of the SC – plasma interaction. The programs include LEOPOLD (to determine the LEO and Polar Orbit environment characteristics); SOLARC (analysis of solar array power loss); EQUIPOT (analysis of material charging); PICOCHARGE (analysis of SC charging) and SAPPHIRE (analysis of SC ram and wake flows). Although the individual codes allow only part of the overall problem to be addressed, use of the entire suite allows a more complete picture of a particular situation to be built up. ESPIRE was constructed to facilitate future additions to the programs as a result of the continuing development of engineering software tools.

Internal charging due to the penetration of dielectric material by high energy electrons, poses a particular threat to SC. Such high energy electrons, for example those present in the Van Allen Belts, can penetrate a SC skin and establish a negative potential on internal dielectric materials and on floating conductors. A fluence of $10^{10} - 10^{11}$ electrons/cm^2 with energies $E > 100$ keV over a period sufficiently long to dominate the relevant dielectric leakage rate, can build up a charge at (say) a circuit board sufficient to produce local arcing (thereby creating a pulse lasting of the order of tens of nanoseconds). Metalized areas on a printed circuit board can also charge up sufficiently to produce arcing. In addition, internal charging affects cable wrap, wire insulation, electrical connectors, feed-throughs etc.

An internal discharge is deemed to be more damaging than an external discharge since it occurs in close proximity to sensitive electronic circuitry (Leach and Alexander, 1995). Studies made in GEO aboard The Combined Release and Radiation Effects Satellite (*CRRES*) indicate that "most environmentally induced spacecraft anomalies result from deep dielectric charging and resulting discharge pulses, and not from surface insulator charging or single event upsets", Gussenhoven *et al.* (1996).

SC anomalies show an association with the high speed solar wind streams that impinge on the Earth in the declining phase of the sunspot cycle, although there is also some association with periods of elevated solar activity. The key factor is solar wind speed and the consequent enhancement of energetic electrons in the outer belt. Anomalies recorded on the *Meteosat-3* SC between 1988 and 1995 show a peak in 1994 of about 160 events whereas, during the solar maximum years, there were only about 50 events per year [††](a new solar cycle started in 1996). A dramatic situation occurred in May 1998 when *EQUATOR–S* failed on May 1, *POLAR* displayed an anomaly in its function on May 6 and *GALAXY-4* suffered failure on May 9. Baker *et al.* (1998) suggested that these breakdowns were due to dielectric break-down and material damage produced by ambient relativistic electrons. This interpretation is supported by the fact that the *SAMPEX* and *GOES* SC detected such electrons in the radiation belts at the relevant times. See also the analysis of Cyamukungu *et al.* (1999).

The ESA DDC code DICTAT was used to interpret internal charging data obtained aboard *CRRES* by the Internal Discharge Monitor, supplemented by data from the Russian *GLONAS* and *HORIZONT* SC. Complementary charging codes (COULOMB and ECO-M) made available by the Russian Scientific Production Association of Applied Mechanics were also used at ESA in evaluating these data sets. DICTAT calculates the electron current that, on passing through a conductive shield, is deposited inside a dielectric. From the deposited current, the maximum electric field within the dielectric is found. This field can then be compared with the breakdown field to determine if the material concerned is at risk of undergoing an electrostatic discharge.

[††]Private communication from M.Hapgood.

NOAA presently makes routine electron flux measurements in polar orbit using its Polar Orbiting Operational Environmental Satellite (*POES*). Its Geo-stationary Operational Environmental Satellite *(GOES)* records complementary measurements in geo-synchronous orbit. Alerts are issued by NOAA/SEC when the > 2 MeV electron flux at *GOES* exceeds 10^3 particle flux units (pfu) during more than three consecutive five-minute periods. An even more significant threshold is flagged when the cumulative fluence exceeds 10^9 pfu over a 72-hour period. Data recorded by *GOES* show that there is a close correlation between *Phantom commands* (changes in the operation of a SC not instigated by ground command) and occasions when electrons in the > 2 MeV range penetrate the skin of a particular SC. See also Section A.3.2.

A.2.5 SC Contamination

The exposure of SC materials to the vacuum of space gives rise to out-gassing. This can lead to an alteration in the surface properties of the out-gassing material itself as well as to the contamination of surrounding surfaces where the released substances are deposited. Also, thruster emissions can cause degradation of the thermal and/or optical characteristics of those SC surfaces on which they impinge. In particular, deposits on solar cells reduce their electric power output. Also, the release of charged particles during active charge control modifies the ambient neutral environment, thereby affecting local plasma interactions. Further, SC charging by charge exchange causes changes in surface potential, with the associated risk of arc discharges.

The *Shuttle* (which has typical apogees in the range 330 – 350 km) is immersed during its flight in a multi-species gas cloud, the shape of which is governed by interactions with the ambient neutral atmosphere and with the space plasma environment. Contributors to this environment are: releases in the form of particulate matter and gases from the onboard Reaction Control System (RCS) and from the Orbital Manoeuvring System (OMS); products of engine firings; cabin gas leaks; gasses from the waste management system; water releases; material out-gassing and ablated surface coatings. At certain angles, sunlight is scattered from this induced atmosphere producing a glare that interferes with the observations of onboard optical instruments.

Engine firings while producing their own particular contaminant cloud, are also associated with a light flash that illuminates the Shuttle and contributes to the production of surface glow phenomena and optical pollution. Particulate contamination is enhanced when the exhaust plume of the RCS impacts directly on *Shuttle* surfaces.

A.2.6 Sputtering from SC Surfaces

In general, atmospheric constituents in the path of an Earth orbiting SC impinge on its surface with kinetic energies that vary in accordance with particle mass and SC velocity. Excited impacting species have, in addition to kinetic energy, excitation and ionization potential energies extending up to ~ 25 eV / particle. The recombination of such atoms at a SC surface to form diatomic molecules involves a release of energy and further energy is released during the formation of surface compounds. The majority of the collisional energies involved in gas-surface interactions in LEO have values that are similar to the surface binding energies of atoms in a solid. Thus, depending on the degree of surface localization, it is possible for bond breaking to occur, resulting in the ejection of atoms and molecules. The particular case of the effect of chemically reactive atomic oxygen on a SC in LEO is discussed in Section A.2.7.

The results of early sputtering calculations for the most abundant atmospheric constituents indicate (Laher and Magill, 1988) that significant surface degradation should be observable in the cases of cadmium, thallium, zinc, lead and indium over exposure times in LEO of the order of months. For magnesium, tin, silver and gold measurable mass loss would occur during exposure times of the order of years.

A.2.7 Corrosive Oxygen

In the diffusive equilibrium conditions prevailing above 110 km in the Earth's atmosphere, the *scale height* (i.e. the height interval over which the density of a particular species drops to $1/e$ of its value) of each atmospheric species varies in inverse proportion to its molecular weight. The density of light constituents thus decreases less rapidly with altitude than is the case for heavy constituents. Atomic oxygen which is formed when solar ultraviolet radiation dissociates molecular

oxygen, is in consequence, the dominant component of the Earth's atmosphere between ~ 200 – 600 km. A solar cycle related variation results in an increase by one or two orders of magnitude in the number density (atoms/m^3) of atomic oxygen found to be present at an altitude of ~ 300 km at solar maximum relative to solar minimum, under quiet geo-magnetic conditions.

Atomic oxygen impingement on a SC depends on: the ambient density; on the relative SC velocity (which, in the case of an easterly orbit, is ~ 7.2 km/s at an altitude of 400 km) and on the orientation of particular SC surfaces (the worst case fluence is on a ram surface at solar maximum). Atomic oxygen is in a highly reactive state and produces erosion of SC surfaces through oxidation. Further, as the SC goes in and out of Earth eclipse, the associated 'thermal cycling' effect removes outer oxidized products from the surface, thereby exposing deeper layers to the erosion process. The effects produced can, in the case of some materials, weaken the strength of elements of the structure and also produce changes in the surface properties that adversely influence SC thermal control. Accounts of extensive studies of atomic oxygen erosion are contained, for example, in Stevens (1990), Caledonia and Krech (1990), Tennyson (1993) and Tribble (1995).

Investigations of surface damage using retrieved space hardware were carried out on the *Long Duration Exposure Facility* (LDEF), which was returned to Earth in 1990. This vehicle spent ~ 5.7 years in space in a nearly circular orbit with an initial altitude of 476 km and with $i = 28.4°$. LDEF experienced the effects of one half of a solar cycle, being deployed during a solar minimum and recovered (when close to re-entry) at a solar maximum. Under these circumstances, about 57% of the overall exposure of LDEF to atomic oxygen occurred during the last six months of its flight and it was observed that there was more corrosive damage at its forward than at its rear extremity. Another specimen used to study erosion was provided by *Eureka* (European Retrievable Carrier), which was brought back from space in 1993 after 326 days in a nearly circular orbit with an initial altitude around 500 km and $i = 28.5°$. The above investigations, supplemented by studies of pieces of Surveyor III (which were brought back to Earth from the surface of the moon) and curated parts of the *Solar Maximum Mission* SC and of the Hubble Space Telescope, were used to facilitate the development of software tools for the corrosive

damage assessment of new SC. Also, these studies stimulated the production of a new generation of materials that have the capability to better withstand the space environment than those previously flown.

Mitigating design strategies for SC in hazardous environments include; choice of suitable structural materials; configuring the power system so that the exposure of high voltage surface areas is minimized; choosing the operating voltages so that they remain below possible arcing thresholds; arranging SC grounding to maintain sensitive structures near the plasma- electrical potential and adopting appropriate technical solutions to maintain SC potential.

A.2.8 Thermal Problems

To provide a successful thermal design for a SC in Earth orbit, it is necessary to take into account both external and internal influences. In the former case, direct solar, Earth reflected solar (albedo) and Earth-emitted infra-red fluxes jointly contribute to SC surface heating. The heating produced varies with SC shape, reflectivity and orientation during different in-orbit events (for example during occultations, periods of oblique illumination etc.). Onboard electronic subsystems and scientific instruments meanwhile generate internal heat. Thermal control can be achieved using either standard *passive* or *active* techniques.

To develop a thermal design for a new SC, all the pertaining key requirements and onboard constraints must first be identified. Also, the SC attitude and orientation relative to the Earth and the Sun during all phases of the mission must be known. Once the thermal boundary conditions, temperature limits for individual onboard items and orbital information is established, a model can be developed that allows SC temperatures to be predicted as a function of time. Thermal tests at both component, subsystem and system level may, thereafter, be conducted to verify the model and demonstrate the proper operation of particular items under stressed environmental conditions.

To guard against changes in the thermal characteristics of SC surfaces due to their degradation during the mission, materials with stable thermo-optical properties and controlled out-gassing behavior should be selected. A current trend to reduce the size of SC components to support the development of small, micro and nano-SC, with the associated necessity to concentrate functions and subsystems in a substan-

tially reduced volume, carries the consequence of high density power dissipation. Highly conductive materials, small-scale heat pipes and pumped fluid loops can assist in providing suitable technical solutions in this regard.

A.2.9 Ambient Electric and Magnetic Fields

Studies conducted for ESA by Martin *et al.* (1990) of SC motion through ambient electric and magnetic fields in Earth orbit have high lighted several important effects and their consequences, namely:

1. The component of SC velocity perpendicular to the magnetic field gives rise to a Lorentz force $(q\mathbf{V} \times \mathbf{B})$ which leads to:
   - Generation of currents causing power dissipation;
   - Changes in SC charging;
   - Structural currents;
   - Electromagnetic interference;
   - Orbital drag;
   - Changes in the SC magnetic moment;
   - Eddy currents;
   - Change in the potential reference for plasma physics experiments;
2. Interactions of SC currents with the magnetic field produce $\mathbf{I}\times\mathbf{B}$ forces and torques causing:
   - Orbital drag;
   - Forces on structural elements;
3. Interactions of a charged SC with an ambient electric field produce qE Coulomb forces, again giving rise to orbital drag and forces on structural elements.
4. Interaction of SC magnetic and electric dipole moments with external fields creates magnetic and electric forces leading to:

- Forces generated in a non-uniform magnetic field causing orbital drag and forces on structural elements;
- Generation of $\mathbf{m} \times \mathbf{H}$ and $\mathbf{p} \times \mathbf{E}$ torques, potentially causing attitude control and pointing errors and deflection of structural elements.

SC with residual net dipole moments (due for example to uncompensated current flows or onboard permeable materials), are subject to magnetic torques. Magnetic cleanliness requirements are thus routinely imposed with respect to the construction and assembly of SC, their subsystems and payload. Recommended practice applies to: material selection; components (for example mounting items with large dipole moments in pairs to produce magnetic moment cancellation); the minimization of current loops etc. Thereafter, the SC dipole moment per unit mass is determined and appropriate magnetic compensation provided.

A.3 Overview of In-Orbit Disturbances

The planets of the solar system are immersed in the dynamic outer atmosphere of the Sun. Thus, in addition to orbital perturbations due, in particular, to the Earth's oblateness; lunisolar gravitational attraction; solar radiation pressure and atmospheric drag, SC are subject to numerous disturbances due to solar activity, the deleterious effects of which are of concern to SC designers and operators as well as to the 'Man in Space Program'. An overview of the onboard problems experienced aboard a SC in orbit close to the Earth due to solar activity is provided in Table A.2. Column 1 indicates the groups within the space community (space segment) concerned with particular, disruptive, in-orbit effects. Column 2 specifies the nature of these problems and Column 3 their cause. Column 4 summarizes the technical consequences ensuing. Column 5 indicates the magnetospheric locations where specific SC problems occur. The various onboard effects concerned and their causes are briefly reviewed below and referenced to preceding information. Also, mention is made of circumstances leading to orbit decay as well as to problems that can be anticipated during interplanetary missions.

Table A.2. In-orbit problems and their origin

Concerned Space Segment	Onboard Effect	Cause of the Effect	Onboard consequence	Magnetospheric location/space weather events
SC designers and SC operators.	Internal charging leading to electrostatic discharges.	Accumulation of charge — mainly electrons ($E >$ 0.5 MeV).	In-orbit anomalies. Phantom commands. Component failure.	Radiation belts and SAA. / Variations in flux due to magnetic storms; substorms; mangetopause compression; fast solar wind streams. SEP events.
	Total Ionizing Dose.	Accumulated radiation damage due to all ionizing radiation — mainly electrons ($E >$ 0.5 MeV); protons ($E >$ 1 MeV) and heavy ions — mainly He^+ and O^+ ($E >$ 1 MeV/n).	Reduction in solar cell power. Reduction in SC lifetime.	All magnetospheric locations but mainly in the radiation belts. / Variations in flux due to magnetic storms; substorms; mangetopause compression; fast solar wind streams. Cosmic Rays. SEP events.
	Single event effects.	Protons ($E >$ 50 MeV). Ions ($E >$ 10 MeV/n). Electrons ($E >$ 500 keV).	Corruption of memory. Component failure.	All locations but mainly the radiation belts. / Cosmic rays (variations due to the solar cycle). SEP events.
	Solar cell degradation and displacement damage.	Electrons ($E >$ 100 keV) and protons ($E \sim$ 1 – 10 MeV).	Reduction in power.	All locations. / Cosmic rays (variations due to solar cycle). SEP events. Radiation belts (as above). Ring current and magnetic storms.

continued on the next page

Concerned Space Segment	Onboard Effect	Cause of the Effect	Onboard consequence	Magnetospheric location/space weather events
	Surface charging resulting in electrostatic discharges.	Solar illumination. Thermal electrons and ions ($E \sim 1 - 50$ eV) and spectrum up to plasma-sheet energies. Electrons and ions ($E \sim 1$ eV – 100 keV). Changing plasma density.	In-orbit anomalies. Phantom commands. Component failure.	All locations, especially the plasma-sheet and outside the plasmapause during early morning; SAA; auroral precipitation region. / Changes in solar UV during SC eclipse. Variations in keV plasma, especially during magnetic storms and substorms. Plasma boundary crossings.
	Surface material degradation, sputtering and erosion.	As for total dose but including protons ($E \sim 0.1 - 1$ MeV) and atomic oxygen.	SC heating through loss of reflectivity.	All locations. / Cosmic rays, SEP events, radiation belts (as above), ring current and magnetic storms.
	Sensor interference and degradation.	As for total dose.	Increase in sensor noise.	Solar wind, magnetosphere, lower ionosphere. / SEP events, cosmic rays, radiation belts (as above), ring current and magnetic storms.
SC operators. GPS navigation. Altimetry. Remote sensing.	Increased atmospheric drag.	Atmospheric expansion.	Loss of stability Early reentry in LEO. Additional use of fuel.	Thermosphere. / Enhanced solar EUV during flares. Joule heating and particle precipitation during substorms and storms.

continued on the next page

Concerned Space Segment	**Onboard Effect**	**Cause of the Effect**	**Onboard consequence**	**Magnetospheric location/space weather events**
	In-orbit anomalies. Phantom commands. Mode switching.	Internal electrostatic discharge. Surface electrostatic discharge. Single event effects.	Partial or complete loss of operations.	All locations, but especially the radiation belts, SAA, Auroral Region, morning magnetosphere. / Variations due to magnetic storms, substorms, magnetopause compression, fast solar wind streams. SEP events.
	Loss of attitude control.	Magnetic field reversals for SC with magnetorquers.	Loss of directional coverage for broadcast signals.	Magnetopause compression inside GEO related to CMEs and periods of increased solar wind pressure.
		Spurious signals in star sensors.	As above.	Solar wind and magnetosphere. SEP events.
	Loss of signal phase and amplitude lock.	Ionospheric conditions. Radio interference.	Errors in navigation. Loss of usable data.	Mainly equatorial and auroral ionosphere. / Enhanced solar EUV during flares. Joule heating and particle precipitation during substorms, magnetic storms and plasma instabilities. SEP events.
Man in Space Program.	Increase in ionizing radiation.	Electrons and ions ($E > 10$ MeV/n).	Increased radiation dose for astronauts.	Low altitude portion of the radiation belts and SAA region. / Variations due to magnetic storms, substorms, magnetopause compression, fast solar wind streams. SEP Events. Galactic cosmic rays.

A.3.1 Surface and Internal Charging

Comprehensive studies of the various environmental interactions that occur between a SC and its surroundings were conducted by Stevens and Kilpatrick (1986) and by Martin *et al.* (1990). These investigations allow the various deleterious situations that occur in a plasma to be grouped under six general headings, These six situations, together with their individual consequences, are listed for convenience below. See also the account of influences of the space environment on a SC contained in Sections A.2.4 – A.2.7.

1. The accumulation of charged particles on a SC surface gives rise, due to associated electrostatic fields and induced currents, to :

   (a) Coulomb force on components and materials causing mechanical forces and torques on the SC structure.

   (b) Electrostatic fields which affect materials, electronics and scientific instruments of the payload as a result of charge–induced conductivity, piezoelectric effects, Kerr (polarization) cell effects, deflection of electrons in detection equipment and changes in the resistivity of materials and components.

   (c) Enhanced contamination (leading to changes in surface thermal and optical properties).

   (d) Surface arc discharges giving rise to electromagnetic interference, electronics damage, logic upsets, material degradation, optical emissions, contamination, forces and torques and an enhanced local plasma density. Also, induced electrical currents can lead to:

   (e) Plasma power leakage between the SC and the plasma, causing power loss and fluctuations, material degradation, contamination and electromagnetic interference.

   (f) Sustained discharges between SC components through the conducting plasma, causing short circuits in the power distribution system, electromagnetic interference, materials degradation and contamination, optical emissions and mechanical forces and torques.

2. Naturally occurring ambient plasma oscillations can couple to the SC systems leading to: the electrical oscillation of surface potentials causing interference with grounding to the plasma, electromagnetic interference and temporal fluctuations in SC potential.

3. SC motion through the plasma, which is hypersonic with respect to the ambient ions, creates a ram-wake structure and local plasma density variations that lead to:

   (a) Radio frequency refraction, causing a degradation in communications links.

   (b) An influence on charging phenomena, causing changes in SC potential.

   (c) Plasma wave generation, causing electromagnetic interference.

4. The impact of low energy thermal ions results in:

   (a) Interactions with chemically volatile ions, leading to chemical reactions which cause changes in the surface properties and material degradation.

   (b) Material-ion collisions which lead to: deposition and sputtering – causing material degradation, change of surface properties and contamination; production of secondary electron emission – causing a change in SC potential.

5. The impact of low energy thermal electrons results in material–electron collisions leading to:

   (a) Secondary electron emission, causing a change in SC potential.

   (b) Back-scattered electrons, causing a cancellation of the charging effects of incoming electrons.

6. The interaction of electromagnetic waves with the plasma can (due to the absorption, dispersion, scattering or reflection of these waves), result in disturbances to communications signals.

Also, the impingement of energetic (> 10 keV)) charged particles on SC surfaces results in internal and bulk charging (Section A.2.4). The onboard consequences include those already noted above under (1a-1c) and, in addition, the following effects:

- Internal discharges, including arcs on circuit boards and inside cables, Lichtenberg pattern arcs causing electromagnetic interference, electronic damage, logic upsets, material degradation, optical emissions, contamination and forces and torques on the SC.

- Sustained discharges, causing electromagnetic interference, material degradation, contamination, optical emissions and mechanical forces and torques.

It should be noted against the above background that the docking of two SC in a plasma, or contact in a plasma between other possible pairs of objects that may be at different potentials, can lead to current flow (charge redistribution), thereby causing arc discharges. This has important safety implications for astronauts engaged in extravehicular-activities.

A.3.2 Phantom Commands

Disturbances due to electrical transients produced in association with surface and internal charging can appear to SC systems as directions from the ground (Phantom Commands). These transients often occur in the local time period between dawn/midnight and midnight/dawn during magnetic storms. This is because, during such SC transitions, the photo-electric effect is abruptly rendered either absent or present, thereby potentially triggering a discharge. Also, thruster firings can produce changes in the local plasma environment that trigger arc discharges. These various unpredicted events can result in loosing control of SC power and propulsion systems as well as of individual scientific instruments. Induced *Mode switching* in which a SC puts itself into Standby Mode, or otherwise changes its expected operational procedure, can also occur.

A.3.3 Total Ionizing Dose and Single Event Effects

Total Ionizing Dose (TID) is used in estimating the effect produced on electronic components by electrons and protons. TID is measured in terms of the *absorbed dose,* which is a measure of the energy absorbed by matter. Absorbed dose may be quantified using the *gray* (Gy), where 1 Gy = 100 rads = 1 J/kg.

Long term, cumulative, damage due to particle radiation can cause electronic devices aboard a SC to display: increased leakage (with related changes in power consumption); threshold effects, timing changes, increased detector background noise etc. Single Event Effects (SEEs) concern instant failure mechanisms (see Section A.2.1). The TID failure rate can thus be described in terms of a *mean time to failure* whereas SEEs are expressed in terms of a *random* failure rate.

To compute the radiation dose for the components to be flown on a new SC, the procedure described in Section A.2.3 can be adopted.

A.3.4 Solar Cell Degradation and Displacement Damage

Solar cells on Earth orbiting SC undergo bombardment by energetic particles. These particles pass through protective coverings and produce both ionization damage and *Displacement Damage* (crystal lattice defects) in the underlying cell. In consequence, cell performance is increasingly degraded over the life time of a mission. Such degradation is made manifest by a reduction in both the voltage and current output (with associated implications for the survival of the SC).

Solar cells are usually made from either silicon or gallium arsenide (the latter provides enhanced efficiency but at an increased production cost). Individual cells can be arranged in series or in parallel to produce a desired output. Solar cell strings can be suitably arranged to minimize power loss from a complete array in case of single cell failure. Various mitigating procedures are adopted by commercial manufacturers to combat the effects of solar cell degradation. See also the procedure for designing an array for a new mission described in Section A.2.3.

A.3.5 Loss of Attitude Control / Orientation

The sun-ward boundary of the magnetopause is usually located at $\sim 10\,R_E$ from the Earth's center. Variations in the pressure of the incoming solar wind can, however, cause this boundary to become severely compressed. Thus, under conditions characterized by high solar wind velocity and density and a strongly southward B_z component of the IMF, the magnetopause may be displaced until it is inside GEO orbit $(\text{at } 6.6\,R_E)$. A GEO SC on the sun-ward side of the Earth can, under these conditions, be outside the magnetopause for periods of from minutes to hours. When a SC in GEO crosses the magnetopause and enters the magnetosheath (located between the bow shock and the magnetopause, see Fig.A.3), its on-board sensors register a drop to near zero in the measured magnetic field. Also, the sign of the field varies erratically in this regime. Under these circumstances, SC attitude control torques can inadvertently be applied in the wrong direction. The design of the attitude control system should, overall, be such as to enable the SC to withstand unwanted magnetic field torques generated during disturbed conditions.

During SEPs, photonic devices such as certain star trackers and CCDs experience a noise floor increase. Such noise in onboard star trackers can result in SC orientation problems.

A.3.6 Loss of Signal Phase and Amplitude Lock

The ionosphere at times becomes highly irregular causing SC signals traversing this disturbed medium to scintillate at the receiver. Strong geomagnetic storms can lead to such scintillation in the auroral zones. Scintillation effects may cause the phase tracking loops in the receivers to temporarily loose lock, thereby introducing discontinuities in the phase derived biased pseudo ranges and significant signal loss ensues (Heroux and Kleusberg, 1989). Not only does this interfere with telemetry (the up / down link) but the Global Positioning System (GPS) can associatively loose tracking contact with a sufficient number of SC to inhibit precise location finding.

The disturbances described above also introduce changes in the time it takes signals to traverse the ionosphere. These delays are proportional to the *Total Electron Content* (TEC) along the signal path and inversely proportional to the square of the frequency of the

signal. The pseudo-range error from variations in TEC can be up to about 16 m at the zenith and about three times that distance close to the horizon. Changes occur over times as short as ten minutes at high latitudes and on diurnal time scales at mid-latitudes. These effects also reduce the accuracy and reliability of the GPS.

A.3.7 Solar Radio Interference

The Sun is in itself a, highly variable, broad-band, radio source. Sometimes when a SC is tracked, the Sun may be within a side-lobe, or even in the main beam, of the tracking antenna. If it then happens that the Sun produces a large radio burst, the signal from the SC can be overwhelmed. Such interference events occur most frequently during the years around solar maximum.

A.3.8 Orbit Decay

Atmospheric heating, stimulated in particular by solar cycle related enhanced EUV radiation, causes the Earth's atmosphere to expand. The resulting increased densities at LEO significantly increase the number of microscopic collisions between a particular SC (with its characteristic *Ballistic Parameter*) and ambient gas particles, thereby increasing the pertaining drag and inducing premature orbit decay, Tribble *et al.* (1999). Atmospheric drag is the largest single source of position prediction uncertainty for most SC at altitudes below ~ 1 000 km and can result in a one-day prediction uncertainty of several tens of km in estimated SC positions. Under conditions of strong, solar related, disturbances, drag can alter the orbit sufficiently that ground contact with a SC is temporarily lost.

Due to solar radiation pressure, all the orbital elements of an Earth orbiting SC in sunlight show annual periodic variations. These variations are rendered more complicated if the SC is eclipsed or undergoes partial surface shadowing. Induced changes in perigee height can significantly affect the lifetime of a high altitude (above ~ 1 000 km) orbiting SC with a large area to mass ratio (see Section A. 1.19).

A SC injected into a geo-stationary orbit is subject, in particular, to perturbations due to the Earth's oblateness, luni-solar gravitation and solar radiation pressure. The onboard propulsion system can make

corrective manoeuvres (station keeping) with a tolerance designed to ensure that the SC remains within the ground antenna beam-width. Once the propellant to enable station keeping is exhausted, the orbital inclination of the SC will grow so that it drifts in longitude, thereby potentially posing a hazard to adjacent SC. There is now a legal requirement to design such SC so that, at the end of their useful lifetimes, they can be suitably moved away to a location where they cannot cause a collision.

A.3.9 Biological Effects

There are a number of parameters that influence astronaut exposure to energetic particle radiation in near Earth orbits. These include: the SC structure; SC altitude and inclination: the orbit phase: the start time and duration (where applicable) of *Extravehicular Activity* (EVAs); the status of the outer electron belts and of interplanetary flux; the phase of the sunspot cycle and the pertaining geomagnetic field conditions.

Low inclination, high altitude flights undertaken during solar minimum are subject to higher dose rates than high inclination, low altitude flights conducted during solar maximum (since the influx of GCRs is less under those conditions characterizing solar maximum, see Section A.1.3). At high altitudes, the area encompassed by the SAA is larger and its concentration of protons enhanced. Although, however, the trajectories of high inclination flights pass through maximum intensity regions within the SAA, less time is spent there than is the case for low inclination flights at the same altitude. Thus, the crew is subjected to less related net exposure in the former case.

At solar maximum, increased solar activity leads to heating and expansion of the atmosphere. This expansion results in a loss of protons from the inner radiation belt due to their interaction with the atmosphere. On the other hand, new belts of trapped particles can form during active periods (Section A.1.11).

The US National Council on Radiation Protection (NCRP) for space-flight activities has drawn up recommendations for organ specific exposure limits for astronauts that include for annual exposure: 50 rem[‡‡] for blood forming organs (200 rem for eyes and 300 rem for

[‡‡]The *Roentgen Equivalent Man* (rem) is a unit used to derive *Equivalent Dose.*

skin). The annual exposure limit for blood forming organs adopted for astronauts in NASA's *ALARA Policy* (As Low As Reasonably Achievable), is ten times higher than that adopted for terrestrial radiation workers. ALARA involves providing real time, onboard, radiological support for flight crews and motivates a general philosophy of minimizing the exposure of personnel to space weather events.

The mission plan for the construction and maintenance of the *International Space Station* (ISS) requires astronauts and cosmonauts to work in space suits outside their SC in six-hour shifts for a total time estimated to exceed 1 500 hours. These EVAs are scheduled, during the construction phase of the station, to take place at a time of high solar activity. One or two EVAs per month are, thereafter, planned (for general maintenance) throughout the lifetime of the *ISS*.

The high inclination orbit (51.6°) traversed by the *ISS* has the consequence that, for a portion of nearly every day, some fraction of the orbit lies within the outer radiation belt. In the worst case this fraction amounts to ~ 20%. During relativistic electron events (Section A.1.16), the intensity of the ambient relativistic electron population increases by a few orders of magnitude. These events occur, on average, approximately once per month and can last for several days. Further, during *Solar Proton Events* (SPEs) the high latitude zones to which solar energetic particles have access show a tendency to widen over the polar latitudes traversed by the *ISS*. This tendency becomes more pronounced when SPEs intensify. Internal doses on the *ISS* are calculated to be of the order of 30 mrads/day. In EVA they are expected to be about ten times higher.

To minimize astronaut exposure during EVAs, risk management procedures, based on the analysis of continuously updating environmental information, are implemented. In this connection, a constellation of strategically placed research and operational SC are coordinated to provide 'real time' information on the radiation environment at the *ISS* orbit. These SC carry out various tasks that include; monitoring the Sun and its corona at multiple wavelengths (to pro-

---

This relates the absorbed dose in human tissue to the effective biological damage of the radiation. Not all radiation has the same biological effect, even for the same amount of absorbed dose. To determine equivalent dose, the absorbed dose (in rad) is multiplied by a quality factor (Q) that is unique to the type of incident radiation.

vide inputs on the strengths of X-ray events and on the generation of CMEs); monitoring SPEs and monitoring the solar wind (to allow the size and shape of SPE access zones relevant to the ISS to be computed). Meanwhile, *NOAA* SC in low altitude orbits monitor relativistic electron fluxes in the outer radiation belt. Relativistic electron flux predictions based on the REFM forecast model (Section A.1.16) are also made available by *NOAA*. An implementation plan to swiftly channel all of this information to the radiation risk managers has been put in place and associated response protocols defined.

A.3.10 Interplanetary Conditions

On interplanetary flights, a flight crew will be subjected to galactic cosmic radiation (GCR) and to solar proton events (SPEs). One percent of GCRs consists of heavy charged particles (mostly carbon, oxygen and iron nuclei) and these constituents, which are known as *High Z and Energy* (HZE) particles, are likely to cause significantly more than one percent of the biological damage produced. Iron is the most important particle biologically because of its relatively high abundance, and also because of the large quantities of energy it deposits per unit path length in tissue (its *Linear Energy Transfer* (LET)). High LET radiation has an elevated *Radiobiological Effectiveness* (RBE). It is noted with reference to SPIEs that, between the *Apollo-16* and *–17* missions in 1972, a solar proton enhancement produced radiation levels of sufficient energy that an astronaut outside of the Earth's magnetosphere would have absorbed a lethal dose within ten hours of the start of this event.

Rough estimates indicate that crews en-route to Mars would receive radiation to their blood forming organs at a rate of 20 rem per year during solar maximum and at 49 rem per year during solar minimum (Case *et al.* 1993). Although the occurrence of GCRs is reduced at the time of solar maximum, missions flown at that time would undergo an increased incidence of SPEs. Energetic particle radiation recorded in a circular orbit ~ 6 000 km above Mars during the *Phobos-2* Mission at a time of enhanced solar activity, see Fig.A.1, shows the presence of significant fluxes of protons with $E > 30$ MeV between March 7-27, 1989. The interaction of such energetic particle radiation with SC shielding, as well as the biological effects estimated to char-

acterize a range of possible mission profiles of different durations need to be studied in depth before an appropriate SC design and a suitable trajectory for a manned flight to Mars can be decided.

In the light of our present understanding of hazardous, solar related, 'near Earth' problems, a preliminary evaluation of risks to SC and their power systems in planetary environments can also be made. In the case of a 'low orbit' at Mars, many of the concerns applicable in LEO at the Earth transpire to be again relevant, namely aerodynamic drag, solar radiation pressure, SC charging and magnetic torques. It is noted, however, that drag varies much less with solar activity in the atmospheres of Mars and Venus than is the case at the Earth. That is because their thermospheres are mainly composed of carbon dioxide, so that they lose heat much more effectively than does the atomic oxygen that dominates the terrestrial thermosphere. Thus, the temperature fluctuations due to solar activity and the main species scale height change are both significantly smaller at these planets than at the Earth.

A.4 Beyond the Solar System

The *Voyager-1* and *-2* Missions were extended in 1989 after some twelve years of space flight. The objective was that these SC would continue to explore the region beyond the outer planets and reach the limits of the Sun's sphere of influence — and maybe beyond. In this connection, the re-named Voyager Interstellar Mission (*VIM*) was commanded to continue to provide particles and fields measurements characterizing the outer solar system. *Voyager-1* is exiting the solar system at a speed of about 3.5 AU per year some 35° out of the ecliptic plane to the north and *Voyager-2* is exiting at a speed of about 3.3 AU per year 48° out of the ecliptic plane to the south.

According to our present understanding, at some distance from the Sun, the supersonic solar wind flow is held back from further expansion by the interstellar wind. As a result of this interaction a *Termination Shock* is formed where the solar wind slows from supersonic to subsonic speed. In this regime, large changes in plasma flow direction and magnetic field orientation occur. In the context of planned *VIM* exploration, study of this region will constitute the 'Termination Shock

Phase' of the mission. Exit from this region will initiate SC entry to the heliosheath and the start of a 'Heliosheath Exploration Phase'. Here, the ambient environment is still dominated by the Sun's magnetic field and by the solar wind. This second phase will end with SC passage through the heliopause, a boundary that defines the outermost extent of both the Sun's magnetic field and the solar wind. Transition of the heliopause boundary will begin the third 'Interstellar Exploration Phase', where the SC finally operates in an interstellar wind dominated environment (the Interstellar Medium (ISM)). It is presently estimated that the location of the terminator shock is at 85 ± 5 AU and the heliopause boundary at between 120 and 150 AU from the Sun.

Meanwhile, a dedicated mission called the *Interstellar Probe Mission* (*IPM*), is independently under design with a minimum goal to reach 200 AU within 15 years after its launch, and with enough onboard consumables to reach 400 AU. The payload is planned to measure the detailed properties of the plasma, neutral atoms, energetic particles, magnetic fields and dust in the outer heliosphere and in the local *ISM*. This mission could constitute a defining step in a yet more ambitious program to explore the nearby galactic neighborhood. It is still in a concept definition phase and requires the realization of several challenging technical developments in low power avionics, advanced power systems and phased array Ka-band telecommunications. Propulsion methods under consideration include a solar sail with a diameter of 400 m and nuclear electric propulsion.

The *IPM* initiative is the latest among a series of such proposals dating back to 1977. However, this time it is listed in NASA's strategic planning documentation as a 'potential mission beyond 2007'. It is described here in that it is representative of the age old human quest for knowledge through extending our confining frontiers, while providing in addition a logical step in regard to what could be attempted next.

List of References for Addendum

Baker, D.N., Belian, R.D. and Higbie, P.R., Deep dielectric charging effects due to high energy electrons in the Earth's outer magnetosphere, *J. Electrost.* **20**, 3-19, 1987.

Baker, D.N., McPherron, R., Cayton, T., and Klebesadel, R., Linear prediction filter analysis of relativistic electron properties at 6.6 RE, *J. Geophys. Res.* **95**, 15133-15140, 1990.

Baker, D.N., Allen, J.H., Kanekal, S.G. and Reames, D. G., Disturbed space environment may have been related to Pager Satellite failure, *EOS Transactions AGU,* **40**, 477-483, 1998.

Baker, D.N., Kanekal, S.G., Pulkkinen, T.I. and Blake, J.B., Equinoctical and solstital averages of magnetospheric relativsitic electrons: A strong semiannual modulation, *Geophys. Res., Lett.* **26**, 3193–3196, 1999.

Blake, J.B., Baker, D.N., Turner, N., Oglivie, K.W. and Lepping, K.P., Correlation of changes in the outer-zone relativistic electron population with upstream solar wind and magnetic field measurements, *Geophys. Res. Lett.* **24**, 927-929, 1997.

Brun, R., Bruyant, F., Maire, M. McPherson, A.C. and Zanarini, P., GEANT 3 User's Guide DDEE, 84-1, Publ. CERN, 1984.

Caledonia, G.E. and Krech, R.H., *Studies of the interaction of* 8 *km/s ions in Materials Degradation in Low Earth Orbit (LEO).* Ed. by V. Srinivasan and B.A. Banks, Publ. Minerals Metals and Materials Society, Warendale, PA, 145-153, 1990.

Canfield, R.C., Cheng, C-C, Dere, K.P., Dulk, G.A., McLean, D.J., Robinson Jr. R.D., Schmahl, E.J. and Schoolman, S.A., Radiative output of the 5 September 1973 flare, Appendix A., Solar Flares, *A Monogram from Skylab Solar Workshop II,* Ed. Sturrock, P.A. (Publ.

Colorado Univ. Press) 451-469, 1980.

Case, C., Cheung, B., Fry, M., Garshnek.V., Money, K., Smith, G.J., Uri, J. and Wimisdoerffer, Human Factors and Physiological Aspects, in International Exploration of Mars: a Mission Whose Time Has Come, IAA Cosmic Studies, Ed. G. Morgenthaler and Smith, G.J. *Acta Astronautica,* **31**, 73-85, 1993.

Chapman, S. and Bartels, J., *Geomagnetism,* Chapter 11, Publ. Oxford Univ. Press, N.York, 1940.

Cliver, E.W, Kamide, Y. and Ling, A.G., Mountains versus Valleys, Semiannual variations of geomagnetic activity, *J. Geophys. Res.* **105**, 2413-2424, 2000.

Cummings, J.R., Cummings, J.C., Mewoldt, R.A., Selesnick, R.S., Stone, E.C. and von Rosenvinge, T.T., New evidence for Anomalous Cosmic Rays trapped in the magnetosphere, *Geophys. Res. Lett.* **20**, 2003-2006 , 1993.

Cyamukungu, M., Lippens, C., Adams, L., Nickson, R., Boeder, C., Pierrard, V., Daly, E., Gregoire, Ch. and Lemaire, J., Magnetic storm acceleration of radiation belt electrons observed by the Scintillating Fiber Detector (SFD) onboard EQUATOR-S, *Ann Geophys.* **17**, 1622-1625, 1999.

Dryer, M., Interplanetary studies: Propagation of disturbances between the Sun and the magnetosphere, *Space Sci. Rev.* **67** (3/4), 363-419, 1994.

Dyer, C.S., Sims, A.J., Trascott, A,J., Peerles, C., and Underwood, C., Temporal variation in the new proton belt created in March 1991 observed using the CREAM and CREDO experiments, *Adv. Sp. Res.* **17**, No.3, 159-152, 1995.

Feynman. J., Armstrong, T. P., Dao Gibner, L.V. and Silverman, S.A.. Solar proton events during solar cycles 19,20 and 21, *Solar Phys.* **125**, No.2, 385-401, 1990.

Fortescue, P. and Stark, J., *Spacecraft Systems Engineering,* 2nd Edition, Publ. Wiley New York, 1995.

Friedel, R.H.W., Cayton, D., O'Brien, P. and Bourdaire, S., Relativistic electron dynamics in the inner magnetosphere, Paper IAA-01-IAA 6.3.03, *52nd International Astronautical Congress* (Toulouse, France), 2001.

Gaposchkin, E.M. and Coster, A.J., Evaluation of Recent Atmospheric Density Models, Paper AAS-87-557, *Proc. of 1987 AAS/AIAA*

Astrodynamics Conference, San Diego, CA., Publ. AAS Publications Office, 1987.

Getselev, I.V., Gusev, A.N., Darchieva, I..A., Kabashova, N.A., Morozova, T.I., Pavlov, A.V., Panasyuk, M.I., Pugacheva, G. I., Rejzman, S. Ya., Savun, O.I., Sosnovets, E.N., Tverskaya, N.V., Timofeyev, G.A. and Yushkov, B.I., Model of Spatial Energetic Distribution of Charged Particle (Protons and Electrons) Fluxes in the Earth's Radiation Belts, INP MSU Preprint MGU-91-37/241 (in Russian) 1991.

Glassmeier, K-H., Neubauer, F.M., Brecht, G., Marschall, H., Acuna, M.H., Burlaga, L.F., Mariani, F., Musmann, G., Ness, N.F., Wallis, M.K., Ungstrup, E. and Schmidt, H.U., Giotto's Mission to Planet Earth, *Geophys. Res. Lett.* **18**, (9), 1663-1666, 1991.

Gonzalez, W.D., Joselyn, J.A., Kamide, Y., Kroehl, H.W., Rostoker, G., Tsurutani, B.T. and Vasyliunas, V.M., What is a Geomagnetic Storm, *J. Geophys. Res.* **99**, 5771-5792, 1994.

Gonzalez, W.D., Tsurutani, B.T. and Clua de Gonzalez, A., Interplanetary origin of geomagnetic storms, *Space Sci. Rev.* **88**, 529-562, 1999.

Goswami, D.V., McGuire, R.E., Reeedy, R.C., Lal, D. and Jha, R., Solar flare protons and alpha particles during the last three solar cycles, *J. Geophys. Res.* **93**, 7195-7205, 1988.

Gussenhoven, M.S., Muller, E.G. and Brautigan, D.H., Quasistatic Model of Outer Zone Electrons. Improved understanding of the Earth's radiation belts from the CRRES satellite, *IEEE Trans. on Nuclear Science,* **43**, No.2, 353-368, 1996.

Hedin, A.E., MSIS-86 Thermospheric Model, *J. Geophys. Res.* **92**, 4749-4662, 1986.

Herox, P. and Kleusberg, A., GPS precise relative positioning and the ionosphere in auroral regions,*Proc. 5th International. Symposium on Satellite Positioning,* Las Cruces, NM., 13-17, 1989.

Hundhausen, A.J. *The Solar Wind in Introduction to Space Physics,* Ed. Kivelson, M. and Russell, C., (Publ. Cambridge Univ. Press) 91-128, 1995.

Jacchia, L.G., Thrmospheric Temperature, Density and Composition New Models, Special Report 375, Cambridge M.A. Smithsonian Astrophysical Observatory, 1977.

Katz, I., Stannard, P.R., Gedeon, L., Roche, J.C., Rubin, J.G., and Tautz, M.F., NASCAP simulations of SC charging of the SCATHA

satellite, ESA SP-198, 109-114, 1983.

Katz, I.,Parks, D.E. and Wright, K.H., A model of the plasma wake generated by a large object, *IEEE Trans. Nucl. Sci.,* NS-32, 4092-4096, 1985.

King, J.H., Solar Proton Fluences for 1977-1983 Space Mission, *J. Spacecraft,* **11**, 401-408, 1974.

Kirsch, E., Daly, P.W., McKenna-Lawlor, S., Daly, P., Korth, A., Neubauer, F.M. and Coates, A.J., Observation of interplanetary particles in a Co- rotating Interaction Region and of energetic water group ions from comet Grigg-Skjellerup, *Planet. Space Sci.* **45**, No.9, 1105-1117, 1997.

Laher, R.R.and Megill, L.R., Ablation of materials in the low-Earth-orbit environment, *Planet. Space Sci.* **36**, 1497-1508, 1988.

Leach, R.D. and Alexander, M.B., Failures and Anomalies Attributed to Spacecraft Charging , NASA document RP-1375, Publ. by the Marshall Space Flight Center, Alabama, 1995.

Lemaire, J., Roth, M. Wisemberg, J., Domange, P., Fonteyn, D., Lesceux, J.M., Loh, G., Ferrante, G., Garres, C., Bordes, J., McKenna-Lawlor, S. and Vette, J.I., TREND, Development study of improved models of the Earth's radiation environment, ESTEC Report 9011, 1991.

Lemaire, J., Johnstone, A.D., Heynderickx, D., Rodgers, D.J., Szita, S. and Pierrard, W., TREND–2, Trapped Radiation Environment Model Development, ESTEC Report 9828, 1995.

Li., X., Baker, D.N. Kanekal, S.G., Looper, M. and Temerin, M., Long Term measurements of Radiation Belts by SAMPEX and Their Variations, *Geophys. Res., Lett.* **28** (20) 3827-2830, 2001a.

Li., X., Temerin, M., Baker, D.N., Reeves, G.D. and Larson, D., Quantitative prediction of radiation belt electrons at geostationary orbit on the basis of solar wind measurements, *Geophys. Res. Lett.* **28**, 1837-1890, 2001b.

Martin, A.R., Rodgers, D.J., Kessel, R.L., Johnstone, A.D., Coates, A.J. Maehlum, R.N, Svenes, K. and Friedrich, M., Spacecraft Plasma Interactions and Electromagnetic Effects in LEO and POLAR orbits, ESA Report, CR(P) No.3025, 1990.

Mathie, R.A. and Mann, J.R., A correlation between extended intervals of ULF wave power and storm-time geo-synchronous relativistic electron flux enhancements, *Geophys. Res. Lett.* **27**, 3261-3264,

2000.

McIlwain, C.E., Coordinates for mapping the distribution of magnetically trapped particles, *J. Geophys. Res.* **66**, 3681-3691,1961.

McKenna-Lawlor, S.M.P., Afonin, V.V., Gringauz, K.I., Keppler, E., Kirsch, E., Richter, A.K., Witte, M., O'Sullivan, D., Thompson, A., Kecskemety, K., Somogyi, A., Szabo, L., Varga, A. and Marsden, R., Interplanetary variability in particle fluxes recorded by the low energy charged particle detector SLED (~ 30 keV - > 30 MeV) during the Cruise Phase of the Phobos Mission to Mars and its Moons, *Ann. Geophyicae* **9**, 348-356, 1991.

McKenna-Lawlor, S.M.P., Elendt, I., Rusznyak, P., Kunnow, H., Muller- Mellin, M. and Witte, M., The Lion instrument on SOHO and its scientific objectives, *Ann. Geophys.* **15**, 1-4, 1997.

McKenna-Lawlor, S., Kirsch, E., Heynderickx, D. and Lemaire, J., Energetic particle data recorded onboard Giotto during the first encounter of an observing spacecraft coming from deep space with planet Earth, *Planet. and Space Sci.* **49**, 1365-1369, 2001a.

McKenna-Lawlor, S., Rusznyak, P., Buchman, S., Shestople, P., and Thatcher, J., A high energy proton monitor for the Gravity Probe-B Mission, Paper IAF-01-QA.04, *52nd Congress of the International Astronautical Federation,* Toulouse (France), 2001b.

McKenna-Lawlor, S.M.P, Dryer,M., Smith, Z., Kecskemety,K., Fry ,C.D., Deehr, C.S., Berdichevsky, D., Kudela, K. and Zastenker, G., Arrival times of flare/halo CME associated shocks at the Earth: comparison of the predictions of three numerical models with these observations, *Ann. Geophys.* **20**, 917-935, 2002.

McPherron, R.I., The role of substorms in the generation of magnetic storms, in Magnetic Storms Ed. Tsurutani, B.T., Gonzales, W.D., Kamide, Y. and Arballo, J.K., *AGU Geophys. Monograph* **98**, 131-147, 1987.

Nord, P. and Kong, M.K., Hardness and Survivability Requirements, in *Space Mission Analysis and Design* (3rd Edition), Eds. Wertz, J.R,. and Larson, W.J. (Editors). Publ. Space Technology Library, Microcosm Press and Kluwer Academic Publications, 221-240, 1999.

O'Brien, T.P., McPherron, R.L., Sorntette, D., Reeves, G.D., Fridel, R. and Singer, H.J., Which magnetic storms produce relativistic electrons at geosynchronous orbit, *J. Geophys. Res.* **107**, 15533-

15544, 2001.

Olson, W.P. and Pfitzer, K.A. Magnetospheric Magnetic Field Modeling, Annual Scientific Report AFOSR Publ. McDonnell Douglas Astronautics Co., Huntington Beach, CA, 1977.

Pfitzer, K.A., Olson, W.P. and Mogstad, T., A Time Dependent, Source Driven, Magnetospheric Magnetic Field Model (Abstract), *EOS* **69**, p.426, 1988.

Purvis, C.K., Garrett, H.B., Whittlesey, A.C. and Stevens, N.J. Design guidelines for assessing and controlling spacecraft charging effects, NASA Technical Paper TP-2361, 1984.

Reames, D. V., Particle acceleration in the Sun and in the Heliosphere, *Space Sci. Rev.* **90**, 413-491, 1999.

Rees, D. (Ed), COSPAR International Reference Atmosphere, Part I Thermosphere Models, *Adv. in Space Res.* **8**, Nos 5-6 1988.

Rees, D. Barnett, J.J. and Labitzke, K. (Eds), COSPAR International Reference Atmosphere Part II Middle Atmosphere Models, *Adv. in Space Res.* **10**, No.12, 520 pp., 1990.

Regan, J.B. Meycrott, R.E., Gaines, E.E., Nightingale, R.W. , Filbert, P.C. and Imhof, W.I., Space Charging Currents and their effects on Spacecraft Systems, *IEEE Trans. Electr. Insul.* EI-18 345-365, 1983.

Roederer, J.G., Geomagnetic field distortions and their effects on radiation belt particles, *Rev. Geophys. and Space Phys.* **30**, No.2, 599-630, 1972.

Russell, C.T., Planetary Magnetospheres, *Science Progress,* **75**, 93-105, 1991.

Russell, C.T. and McPherron, R.L., Semiannual variation of geomagnetic activity, *J. Geophys. Res.* **78**, 92-108, 1973.

Sciscoe, G. L. and Petschek, H.E., On storm weakening during substorm expansion phase, *Ann. Geophys.* **15**, 221-216, 1997.

Stevens, J.N., Method for estimating atomic oxygen surface erosion in space environments. *J. Spacecraft and Rockets,* **27**, No.1, 93-95, 1990.

Stevens, N. J. and Kirkpatrick, M.E., Spacecraft environment interaction investigation, Airforce Geophysical Laboratory Publication AFGL-TR—96-0214, 1986.

Tennyson, R.C., Atomic oxygen and its effect on materials, in *The Behavior of Systems in the Space Environment,* Ed. Dewitt, R.N.

Publ. Kluwer Academic, Amsterdam, 233-357, 1993.

Tranquille, C. and Daly, E.J. An evaluation of solar proton event models for ESA Mission, *ESA Journal,* 16(3). 275-287, 1992.

Tribble, A.C., *The Space Environment: Implications for Spacecraft-Design,* Publ. Princeton University Press, 1995.

Tribble, A.C., Gorney, D.J., Blake, J.B., Koons, H.C., Schulz, M., Vampolar, A.M., Walterscheid, R.L. and Wertz, J., The Space Environment and Survivability, in *Space Mission Analysis and Design* (3rd Edition), Ed. Wertz, J.R. and Larson, W.J. Publ. Space Technology Library, Microcosm Press and Kluwer Academic Publications, Chapter 8, p.210, 1999.

Tsyganenko, N. A., A magnetospheric magnetic field model with a warped tail current sheet, *Planet. Space Sci.* **37**, 5-20, 1989.

Tylka, A.J., Boberg, P.R., Adams, J.R. Beahm, L.P., Dietrich, W.F. and Kleis, T., The mean ionic charge state of solar energetic Fe ions above 200 MeV per nucleon, *Astrophys. J.,* **444**, L109-L113, 1995.

Vallado, D. A., *Russian GOST Atmosphere in Fundamentals of Astrodynamics and Applications,* 2nd Edition, Publ. Microcosm Press, and Kluwer Academic Publ. 874-879, 2001.

Van Allen, J. A., Ludwig, G.H., Ray, E.C. and McIlwain, C.E., Observations of high intensity radiation by satellites 1958 Alpha and Gamma, *Jet Propul.,* **28**, 588-592, 1958.

Van der Ha, J.C. and Modi, V.J., On the maximization of orbital momentum and energy using solar radiation pressure, *J. Aeronautical Sci.,* **27**, 63-84, 1979.

Vette, J.I., A comparison of trapped proton models AP8MIN and AP8MAX with their compressed versions AP8MIC and AP8MAC, NSSDC-WDC-A, File No.13489, 1977.

Vette, J.I., The NASA National Space Science Data Center: Trapped Radiation Environment Model Program (1964-1991), NSSDC-WDC-A, R & S, 91-29, 1991a.

Vette, J.I., The AE-8 Trapped Electron Model Environment, NSSDC-WDC-A, R & S, 91-24, 1991b.

Webb, D.F., Cheng, C-C., Dulk, G.A., Edberg, S.J., Martin, S.F., McKenna-Lawlor, S. and McLean, D.J., Mechanical energy output of the 5 September 1973 flare, Appendix B., Solar Flares, *A Monogram from Skylab Solar Workshop II,* Ed. Sturrock, P.A. (Publ. Colorado

Univ. Press) 471-499, 1980.

Wertz, J.R., *Spacecraft Attitude Determination and Control,* Publ. Dordrecht, D. Reidel Publ. Co. 1978.

Wilson, R.C., Hudson, H. and Woodard, M., The inconstant solar constant, *Sky and Telescope,* **67** (6), 501-503, 1984.

Zhou, X-W, Russell, C.T., Lee, G. and Tsyganenko, N., Comparison of observed and model magnetic fields at high altitudes above the polar cap: POLAR initial results, *Geophys. Res. Lett.* **24**, 1451-1454, 1997.

Index of Scientists Cited in Footnotes

Index

List of Figures

List of Tables