

Springer Series in Astrophysics and Cosmology

Lin Liu

Algorithms for Satellite Orbital Dynamics

Translated by Shengpan Zhang

 南京大学出版社
NANJING UNIVERSITY PRESS

 Springer

Springer Series in Astrophysics and Cosmology

Series Editors

Cosimo Bambi, Department of Physics, Fudan University, Shanghai, China

Dipankar Bhattacharya, Inter-University Centre for Astronomy and Astrophysics, Pune, India

Yifu Cai, Department of Astronomy, University of Science and Technology of China, Hefei, China

Maurizio Falanga, (ISSI), International Space Science Institute, Bern, Bern, Switzerland

Paolo Pani, Department of Physics, Sapienza University of Rome, Rome, Italy

Renxin Xu, Department of Astronomy, Perkin University, Beijing, China

Naoki Yoshida, University of Tokyo, Tokyo, Chiba, Japan

Pengfei Chen, School of Astronomy and Space Science, Nanjing University, Nanjing, China

The series covers all areas of astrophysics and cosmology, including theory, observations, and instrumentation. It publishes monographs and edited volumes. All books are authored or edited by leading experts in the field and are primarily intended for researchers and graduate students.

Lin Liu

Algorithms for Satellite Orbital Dynamics

 南京大学出版社
NANJING UNIVERSITY PRESS

 Springer

Lin Liu
Department of Astronomy
Nanjing University
Nanjing, Jiangsu, China

Translated by
Shengpan Zhang
Department of Astronomy
York University
Toronto, ON, Canada

ISSN 2731-734X ISSN 2731-7358 (electronic)
Springer Series in Astrophysics and Cosmology
ISBN 978-981-19-4838-1 ISBN 978-981-19-4839-8 (eBook)
<https://doi.org/10.1007/978-981-19-4839-8>

Jointly published with Nanjing University Press
The print edition is not for sale in China (Mainland). Customers from China (Mainland) please order the print book from: Nanjing University Press.
ISBN of the Co-Publisher's edition: 9787305222276

© Nanjing University Press 2023

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publishers, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publishers nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publishers remain neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

“Orbital Dynamics” is an essential and fundamental part of Aerospace Dynamics (Aerospace System Engineering). It includes the launching of spacecraft, the entire system of orbital design, orbital observation and control, and the effective application of all aspects.

The Author teaches, researches, and works in this field for more than 50 years. Based on his abundant experience, this monograph summarizes the application part of Orbital Dynamics. The main content is about the orbital motion of spacecraft of different types in the Solar System (actually spacecraft are circling satellites with different center bodies). A systematic and effective calculation method is provided for orbital position telemetry, tracking, and orbital determination and prediction after a spacecraft is launched, and special orbital design, realization, and retainment, for all sorts of space projects.

This book describes the essential theory and analytical results but omits some details of the calculation principle and the process of formula derivation. The main purpose is to provide calculation methods and formulas so technicians and engineers with basic knowledge of orbital dynamics in the Aerospace industry can directly apply them in their work. The book is also beneficial for related professionals to obtain the necessary understanding of orbital dynamics.

References of this book include the author’s published academic research papers, teaching materials, and 11 books. The books published in Chinese are as follows.

1. Liu, L. and co-authors, Motion Theory of Earth’s Artificial Satellite, Science Press, Beijing, China, 1974.
2. Liu, L. and Zhao, D. Z., Orbital theory of Earth’s Artificial Satellite, teaching materials, Nanjing University, China, 1979.
3. Liu, L., Orbital Dynamics of Earth’s Artificial Satellite, Higher Education Press, Beijing, China, 1992.
4. Liu, L., Methods of Celestial Mechanics, Nanjing University Press, Nanjing, China, 1998.
5. Liu, L., Orbital Theory of Spacecraft, National Defense Industry Press, Beijing, China, 2000.

6. Liu, L. and co-authors, *Mathematical methods of Precision Orbital Determination for Earth's Artificial Satellite*, PLA Press, Beijing, 2002.
7. Liu, L., Hu, S. J., and Wang, X., *Introduction of Aerospace Dynamics*, Nanjing University Press, Nanjing, China, 2006.
8. Liu, L. and Wang, X., *Orbital Dynamics of a Moon Prober*, National Defense Industry Press, Beijing, China, 2006.
9. Liu, L. and Tang, J. S., *Satellite Orbital Theory and Application*, Electronic Industry Press, Beijing, China, 2015.
10. Liu, L., Hu, S. J., and co-authors, *Spacecraft Orbital Theory and Application*, Electronic Industry Press, Beijing, China, 2015.
11. Liu, L. and Hou, X. Y., *The Basics of Orbital Theory*, Higher Education Press, Beijing, China, 2018.

In the process of writing this monograph, two of the Author's students, Yanrong Wang and Gongyou Wu, who both work in the Chinese Aerospace Industry and have first-hand experience, provided valuable suggestions; another student, Zhitao Yang, reviewed Chap. 4. This monograph is also supported by Astronomy and Space Science Institute, Nanjing University research project (NSFC J1210039), and Jiangsu Brand Professional Construction Project (TAPP). The author is grateful for their help.

Nanjing, China

Lin Liu

Introduction

Orbital Dynamics in the Solar System

From a general point of view, the motion of any celestial body, natural or artificial, includes two different types of states. One is about the motion of its barycenter, and the other is about the motion of any part of the body with respect to its barycenter. The first type is called the orbital motion, which is the subject of Orbital Dynamics, and the second is called the attitude motion, which is the subject of Attitude Dynamics. Ancient astronomers already observed attitude motions such as Earth's precession and nutation, and the lunar physical libration. This book is about orbital dynamics in the Solar System. The primary content is about the characteristics of the barycenter motion of spacecraft for different purposes. Some parts of the book deal with the attitude motion, such as in the selection of a proper coordinate system when it is necessary to consider the precession and nutation due to the vibration of Earth's equator; also, the orbital motion of a spacecraft may be related to its attitude when a surface force on the spacecraft acts as an external force.

The Solar System is an extremely complicated dynamical system. In this system, besides the Sun which is the dominating body, there are eight major planets and a large number of asteroids, natural satellites, comets, and space debris. The primary subject of Celestial Mechanics is to study orbital motions of celestial bodies, big and small, in the Solar System and the evolution of their orbits. In the Space Era, the ever-increasing artificial bodies have been added to the Solar System. Although they can be regarded as small celestial bodies, the problem of their motions including the dynamical environment and the wide range of their usage is quite different from the natural celestial bodies. The existence of artificial bodies has expanded the research scope and content of the dynamics of the Solar System, and has made Celestial Mechanics closely linked to Aerospace Dynamics.

As mentioned above, in the Solar System there are numerous celestial bodies with relatively small masses, including both natural and artificial bodies. From the point of view of dynamics, the question is what kind of celestial body can be regarded as a small body. The answer is as follows. If the mass of a body is too small to influence the

related dynamical system as well as the motions of other celestial bodies in the system, then this body is regarded as a small body. Based on all artificial spacecraft launched from Earth, it is obvious that a spacecraft is a small body, its mass is relatively small therefore the motions of all other celestial bodies in the Solar System (including Earth) cannot be affected by a spacecraft.

In a dynamical system, all bodies are sources of gravitational forces. If a dynamical system is formed by a group of celestial objects, and each object is treated as a “particle” (for now we ignore details about non-particle gravitational forces and non-gravitational forces, which does not affect the content provided here), and all objects interact with each other gravitationally, therefore the problem of predicting the individual motions is called an N -body problem mathematically, and the system is called an N -body system, N is the total number of bodies in the system. If one of the objects is a small body whose mass can be omitted, and the motions of the other $N-1$ objects are defined, then the problem of predicting the motion of the small body is called a restricted N -body problem. The difference between an N -body problem and a restricted N -body problem is not merely in names but fundamental in research methods and concepts. When $N = 3$, this problem is the most famous “**restricted three-body problem**” in Celestial Mechanics. In the restricted three-body problem because the research goal is the motion of a small body whose mass can be ignored, the mathematical method and the motion property of the small body are significantly different from the general three-body problem. For example, in the general three-body problem, there are only 10 classical integrals, and there is no other dynamical information available. In the restricted three-body problem not only the motions of the two big bodies are defined but also the characteristics of the motion of the small body are also given. The available dynamical information of their motions is extremely important for studying the motions of natural small bodies, such as asteroids, and all sorts of spacecraft. The information is also closely related to the launch of deep-space spacecraft and to the formation of a specific orbit due to a specific purpose. Therefore, the Orbital Dynamics of deep-space exploration and the restricted N -body problem are inseparable.

The above-described restricted N -body model is built up in a gravitational system of particles, which is a classical model. In reality, under certain circumstances the motion of a small body is also affected by the irregular shape and the uneven mass distribution of the big bodies, the non-gravitational forces (such as the radiation force), and the post-Newtonian effect, etc. Because these forces do not change the basic principles and methods applied in the classical model, it is unnecessary to restrict ourselves in the classical definition, and these forces can be treated as external forces and their effects can be expressed mathematically.

Two Dynamical Systems in the Orbital Dynamics

In a dynamical system, the motion of a celestial body including spacecraft, whether big or small, is usually controlled by more than two external forces. But in the Solar System, where the Sun has been existing for more than 4.6 billion years and is the dominator, the motion of a celestial body is mainly determined by no more than two external forces. For the major planets, there is only one main source of force, i.e., the Sun, the other sources of forces are regarded as perturbations, i.e., small disturbances. For an asteroid, there is usually only one source, the Sun, or are two, the Sun and one of the major planets. For a natural satellite, the main force is from a related major planet. For most of the artificial Earth's satellites, it is Earth. If the satellite has a high Earth orbit (such as a lunar rover, its orbit needs to be changed during its mission) there are two sources of force, Earth and the Moon. For a deep spacecraft (to explore major planets or natural satellites) the external sources can be the Sun, or the Sun and a major planet, or a major planet and one of its satellites. In all the cases mentioned, besides the one or two main forces, other forces (including non-particle gravitational forces and non-gravitational forces) can be treated as perturbing forces. Therefore, from the perspective of orbital dynamics in the actual Solar System there are only two rational dynamical models for studying the motion of various spacecraft. One model is for the circular orbital problem (for artificial satellites including Earth's satellites and the Moon's satellites, etc.) with one main force (i.e., the central celestial body), which corresponds to **"the perturbed two-body problem"**. The other model has one or two primary forces as in the cases of most deep-space spacecraft, which corresponds to **"the perturbed restricted three-body-problem"** are regarded as perturbations, i.e., small disturbances.

Mathematical Models for Satellite Motion: The Perturbed Two-Body Problem [1–8]

As mentioned above, in the Solar System there is only one main force that controls the motions of major planets, asteroids, satellites, and artificial spacecraft (artificial Earth's satellites, the Moon's satellites, Mars's satellites, and other orbiting spacecraft). For planets, the main force is the Sun, for natural satellites, it is the related planet, and for spacecraft, it is the target celestial body. Compared to the main force, other forces are small, therefore generally the N -body system ($N \geq 3$) can be regarded as a "perturbed two-body system", which is mathematically called a perturbed two-body problem. For distinguishing the two bodies, the main external source is called the "central body", denoted as P_0 , and its mass is denoted as m_0 ; whereas the other body, which is the object to study, is denoted as p , and its mass as m . Our research object is the orbital motion of a celestial body, no matter it is a planet, a satellite, or a circling spacecraft, controlled by the gravitational force of the central body and a few other perturbing forces.

The orbital motion of the perturbed two-body problem can be presented by an ordinary differential equation, which is

$$\ddot{\vec{r}} = -\frac{G(m_0 + m)}{r^3}\vec{r} + \sum_{i=1}^k \vec{F}_i, \quad (1)$$

where G is the universal gravitational constant, $\vec{F}_i$ is the i -th perturbing acceleration, k (>1) is the number of perturbing sources. The origin of the coordinate system is located at the barycenter of the central body P_0 , $\vec{r} = \vec{r}(x, y, z)$ is the position vector of the moving body in the coordinate system. The initial values are given by

$$\vec{r}(t_0) = \vec{r}_0, \dot{\vec{r}}(t_0) = \dot{\vec{r}}_0. \quad (2)$$

By convention, we introduce a symbol μ defined by

$$\mu = G(m_0 + m). \quad (3)$$

Then Eq. (1) becomes

$$\ddot{\vec{r}} = -\frac{\mu}{r^3}\vec{r} + \sum_{i=1}^k \vec{F}_i. \quad (4)$$

For a small body p (representing any circling prober) with mass $m = 0$, we have $\mu = Gm_0$ which is the gravitational constant of the central body. In the motion problem of an artificial Earth satellite, the central body is Earth, therefore $\mu = Gm_0 = GE$, where E is Earth's mass, and GE equals $3.98603 \times 10^{14} \text{ (m}^3/\text{s}^2\text{)}$ is the Earth's gravitational constant. For a low Earth orbit, if the altitude of a satellite is about 300 km, then Earth's center gravitational acceleration (μ/r^2) would be about 9 m/s^2 . Denoting $\vec{F}_i$ ($i = 1, 2, \dots$) to the natural existing perturbing accelerations, then the largest among them is due to Earth's non-spherical part, which is only 10^{-3} of the acceleration of the central force. We can say that Eq. (4) is for a typical perturbed two-body problem, and the corresponding motion orbit is a slowly changing ellipse. If the weight of the satellite is 1 ton, and it has a constant thrust of 100 N (like a mobile platform), then the mechanical acceleration of the thruster is about 0.1 m/s^2 , which is 10^{-2} of the acceleration by the barycenter force, so this thrust can be also treated as a perturbation. Actually, when the Moon moves around Earth, Earth is the primary body, and the force from the Sun is a perturbation, which is about 2×10^{-2} of the barycenter force of Earth, greater than the mechanical thrusting acceleration on the satellite. Therefore, the perturbed two-body model can be applied to the motion of a satellite with a mobile platform.

The Two-Body Problem and Kepler Orbit

The reference model of the perturbed two-body problem is a simple two-body problem, which is expressed by an ordinary differential equation as

$$\ddot{\vec{r}} = -\frac{\mu}{r^3}\vec{r} \quad (5)$$

with the initial condition given by (2). This equation is completely solvable, and the solution of this equation is the well-known Kepler orbit. A Kepler orbit is a conic curve, i.e., an ellipse, a parabola, or a hyperbola, and can be presented as

$$r = \frac{p}{1+e \cos f} . \quad (6)$$

where f is the true anomaly, e is the eccentricity, and P is the semi-latus rectum given by

$$p = a(1 - e^2), e < 1 ; \quad (7)$$

$$p = 2q, \quad e = 1 ; \quad (8)$$

$$p = a(e^2 - 1), e > 1 . \quad (9)$$

The three curves are ellipse ($e < 1$), parabola ($e = 1$), and hyperbola ($e > 1$). In (7) and (9), a is denoted as the semi-major axis; in (8), q is denoted as the periapsis.

Another key integral in the two-body problem is the anomaly, which is a function of time t and is directly related to the position of the orbiting body. The relationship of the anomaly and time t has three forms for ellipse, parabola, and hyperbola that

$$E - e \sin E = n(t - \tau) = M , \quad (10)$$

$$2 \tan \frac{f}{2} + \frac{3}{2} \tan^3 \frac{f}{2} = 2\sqrt{\mu q}^{-\frac{3}{2}}(t - \tau) , \quad (11)$$

$$e \sinh E - E = n(t - \tau) = M . \quad (12)$$

The three formulas are the three forms of the famous **Kepler Equation**, and the motion of the body is called Kepler orbit. In the three equations τ is the time when the moving body is at the periapsis; f , E , M and are the true anomaly, eccentric anomaly, and mean anomaly, respectively; and n is the mean angular speed given by

$$n = \sqrt{\mu} a^{-\frac{3}{2}}. \quad (13)$$

As described above, the two-body problem and the Kepler orbit are indiscriminative and both contain three types of orbits, i.e., ellipse, parabola, and hyperbola. The focus of attention is usually on the ellipse because this is the primary form of celestial motion in the Solar System.

The Method of Solving the Perturbed Two-Body Problem

For solving the **perturbed two-body motion** Eq. (1) so far, we do not have a very efficient method. Summarized here is the widely used perturbation method in advanced science and engineering. The accepted **reference orbit** is a **Kepler orbit**, the actual orbit is a slowly changing Kepler orbit. The related motion, at any given time, can be presented by an instantaneous Kepler orbit (such as an **instantaneous ellipse orbit**). Specifically based on the reference model, by the method of the variation of arbitrary constants we first transfer the original equation to a small parameter equation; then construct a required analytical solution according to the analytical theory of ordinary differential equation (Poincare Theorem) as power series of a small parameter of the first-order, the second-order, or the higher-order form. In the method of the variation of arbitrary constants, the basic parameters are usually the six constant integrals in the complete solution of the two-body problem. Note that for a perturbed two-body problem the six constant integrals are no long constants. These basic parameters have definite geometrical meanings and are called Kepler orbital elements denoted by a set σ that

$$\sigma = (a, e, i, \Omega, \omega, M)^T, \quad (14)$$

where the superscript T means the transposition of a matrix. In the barycenter celestial coordinate system, the definitions of the orbital elements are a the semi-major axis, e the eccentricity, i the inclination, Ω the longitude of the ascending node, ω the argument of the periapsis, and M the mean anomaly. The first three elements, a , e , and i , are angular momentums, and the other three elements Ω , ω , and M are angular variables (Ω and ω are slowly changing variables, and M a fast-changing variable).

The perturbed two-body problem Eq. (1) can be converted into a system of equations of small parameters using the method of the variation of arbitrary constants, written as

$$\dot{\sigma} = f(\sigma, t, \varepsilon), \quad (15)$$

where ε is a small parameter related to the perturbing acceleration $\vec{F}_\varepsilon$. This system of equations has a few different forms and is discussed in depth in subsequent chapters. The initial value is denoted by

$$\sigma(t_0) = \sigma_0, \quad (16)$$

where σ_0 is for the initial values of the six orbital elements. Equations (15) has other forms, as discussed in related chapters.

The perturbed solution of orbital elements can be expressed as a small parameter power series by the classic perturbation method (or other improved perturbation methods) written as

$$\sigma(t) = \sigma^{(0)} + \Delta\sigma^{(1)} + \Delta\sigma^{(2)} + \cdots + \Delta\sigma^{(k)}, \quad (17)$$

where $\sigma^{(0)}$ is for the orbital elements of the reference orbit which is an unperturbed orbit. This classic method for the perturbation solution is still the best method in use, and is also applicable for both solutions of a varying ellipse and a varying hyperbola.

In the development of Celestial Mechanics and the Satellite Orbital Dynamics, researchers have tried different methods. One of them is the “**intermediary orbit**” method. The so-called intermediary orbit is an orbit including some influences of perturbing forces therefore is closer to the actual orbit than the non-perturbed orbit. One of the successful examples is the Moon’s intermediary orbital solution (the Hill problem). The intermediary orbital method is also applied in forming the orbital solution for the artificial satellite when the effect of the non-spherical gravitational force of Earth is included in the intermediary reference orbit. The intermediary orbit, in fact, is a changing ellipse including some perturbing forces, thus it does not have any essential improvement, so is not necessary to be called a non-Kepler orbit. Actually, neither the Hill solution of the Moon orbiting around Earth nor the intermediary orbit for an artificial Earth satellite can be directly applied in practice. The practical method is still based on adding remaining perturbing forces to the changing ellipse orbit. Therefore, at the present time, the Kepler orbit is still the most desired reference orbit in solving the perturbed two-body problem.

In dealing with an actual problem, the sixth orbital element of the perturbed orbit after using the method of the variation of arbitrary constants is neither τ (τ is the time when the moving satellite is at the periapsis) nor $M_0 = n\tau$, $n = \sqrt{\mu}a^{-\frac{3}{2}}$, but M , the mean anomaly, given by

$$M = n(t - \tau). \quad (18)$$

There are two reasons for using M . One is that τ and M_0 have no practical meanings in a perturbed motion, whereas M has a defined geometric meaning so is easy to use; the other is that M is a function of a and τ , therefore in the perturbed equation the operation of $\partial R / \partial a$ no longer deals with the problem of inexplicitly including a in

the perturbation function R (through M). With M as an independent element, the perturbed equation can be simplified.

The Perturbed Restricted Three-Body Problem in the Motion of Deep-Space Prober

The Restricted Three-Body Problem for Circular and Elliptical Motions [9–12]

In a three-body problem with $N = (2 + 1)$, there are two primary bodies and a small body. Because the small body has no influence on the motions of the two primary bodies, the motions of two primary bodies are defined by a simple two-body problem. Each of the big bodies moves in a circle or an ellipse around their common barycenter, but neither a parabola nor a hyperbola, realistically. This problem, therefore, is a restricted circular three-body problem or a restricted elliptical three-body problem. The motion of the third body (a small body) in this system is to be studied.

In the Sun-Earth-Moon three-body system the Moon's mass (m) by comparison is much smaller than the masses of the Sun and Earth, m_1 and m_2 , respectively, that $m = 0.012 m_2$, approximately; and the eccentricity of Earth's orbit around the Sun is only 0.017, thus the motion of the Moon in this system can be treated in a circular restricted three-body problem. Of course, an elliptical restricted three-body problem is closer to the real situation than a circular restricted three-body problem. This model is also applied to the motion of an asteroid located in the asteroid belt (between Mars's orbit and Jupiter's orbit, most asteroids are in this belt). The motion of an asteroid is due to mainly the gravitational forces from the Sun and Jupiter. Because the eccentricity of Jupiter's orbit is relatively small, the motion of an asteroid can be also treated as a circular restricted three-body problem.

The orbit of a deep-space spacecraft is more complicated. The whole process can be divided into a few segments. For example, after launch a Moon's prober has a near Earth orbit like an Earth's satellite; when it is near the Moon, it changes its orbit and moves around the Moon, between the two orbits the motion of the prober is in a typical restricted three-body system of Earth, the Moon, and the prober, which can be a circular or an elliptical restricted three-body problem. Another example is about a Mars's prober. In the early stage after launch, it moves like an Earth's satellite. During the time it leaves the Earth-Moon system and before it reaches the area of Mars's gravitational field there is a long cruising period controlled by the Sun's attraction, and the motion is decided by a perturbed two-body problem with the Sun as the central body. After it moves into Mars's gravitational field its motion then is provided by a typical restricted three-body problem of the Sun, Mars, and the prober. There are many more examples like these in the exploration of the Solar System.

The equations of the above-mentioned motion models, including the simplest circular restricted three-body problem, are unsolvable. There is only one solved problem which is the restricted three-body problem with two “motionless” main bodies. It is an approximate model when the motions of the two big bodies are much slower than the small body. This is called the problem of two stationary main bodies in the restricted three-body problem. The equation of this model is solved, but the model is too simple to be used in solving any actual dynamic problem for a spacecraft in the Solar System.

Models for the Restricted N -body Problem and the Perturbed Restricted Three-Body Problem [13, 14]

In the restricted N -body ($N \geq 3$) problem, there are n -big bodies and one small body. One example of this problem is the motion of an asteroid in the asteroid belt. In order to present the motion close to the real situation and to agree with the characteristics of the distribution of asteroids (such as the Kirkwood gaps), the gravitational forces should be included are not only the primary forces from the Sun and Jupiter but also the forces from Saturn and Mars, thus a restricted problem of $(4 + 1)$ bodies is formed. Another example is about a Moon’s prober. The motion of the prober is determined by gravitational forces from Earth, the Moon, and the Sun; therefore, it is a kind of restricted problem of $(3 + 1)$ bodies. In this kind of system, it does not matter how many big bodies there are (i.e., different values of N), the motion of a small body, which can be an asteroid or a Moon’s prober, is studied by assuming that the motions of the big bodies are defined. In the first example, if the force from the third body and the fourth body is not strong enough to cause obvious changes to the results of the original restricted three-body problem, then their forces can be treated as perturbations, and the N -body problem can be regarded as a perturbed restricted three-body problem. Similarly in the second example, the force of the third body can be treated as a perturbation. In reality, the motion of an asteroid in the asteroid belt or a Moon’s prober is studied in this way. In other words, this kind of problem is solved by the perturbation method using as much information as possible from a restricted three-body problem. This method is applied to design deep spacecraft orbits and some specific orbits for specific purposes (such as the Halo orbit). From the above two examples, although using the five-body problem model or the four-body problem model is seemingly more precise and more attractive, the corresponding four-body problem or three-body problem of the big bodies is still unsolved to the present day. The fact is that in the Solar System to study the motion of a natural celestial body or an artificial body the most commonly used models are the **perturbed restricted two-body problem** model and the **perturbed restricted three-body problem** model, especially the **perturbed circular restricted three-body problem** model.

The Restricted Problem of $(n + k)$ -Bodies [15, 16]

The $(n + k)$ -body problem is an N -body problem where $N = n + k$, and there are n -big bodies and k -small bodies with $n \geq 2$ and $k \geq 2$. If $k = 1$, the problem is then reduced to one of the above discussed two examples. This $(n + k)$ system is actually about motions of k -small bodies in an n -body system. The motions of these big bodies are defined, and the small bodies are attracted by the big bodies. The masses of the k small bodies are much smaller than those of the big bodies, thus the small bodies do not affect the motions of the big bodies, but if the distances between the small bodies are short, the gravitational forces between them should be considered. If the gravitational forces of the small bodies can be ignored, then the motion of each small body can be studied separately in an $(n + 1)$ -body restricted problem.

The type of $(n + k)$ -body problem exists in the Solar System. For example, to study the motions of two closely located asteroids in the asteroid belt, the gravitational forces between them need to be considered, then the Sun, Jupiter, and the two small asteroids make up a restricted problem of $(2 + 2)$ bodies. Another example is about launching two geosynchronous satellites at a fixed point high up above the equator. If the weight of each satellite is a few tons, and the distance between them is about a few hundred meters, then to obtain a high precision solution the influence between the two small satellites needs to be included. In this case, Earth, an ellipsoid, is treated as if there were two bodies, one is a sphere with evenly distributed mass, and the other is a “body” formed by the non-spherical part around Earth’s equator, thus there is a restricted problem of $(2 + 2)$ -bodies. Similar situations also appear in the launching of a few spacecraft at a specific point, the gravitational forces between them cannot be ignored, if there are two related big bodies, then the spacecraft and the two big bodies form a restricted problem of $(2 + k)$ -bodies.

General Restricted Three-Body Problem

In a restricted three-body problem under Newton’s gravitational forces if a big body has strong radiation, then its post-Newtonian effect (i.e., the post-Newtonian expansion) should be considered, and the system might be called a generalized restricted three-body problem. In this case, if the motion of the second big body is not affected by the radiation (rigorously speaking, the effect is small enough to be omitted), then the non-gravitational force on the small body should be included. The research of the generalized restricted three-body problem is not discussed in the book.

References

1. Brouwer D, Clemence GM (1961) *Methods of Celestial Mechanics*. Academic Press, New York and London
2. Beutler G (2005) *Methods of Celestial Mechanics*. Springer-Verlag Berlin, Heidelberg
3. Vinti JP (1998) *Orbital and Celestial Mechanics*. *AIAA Education Series*, American Institute of Aeronautics and Astronautics, Inc., Reston, Virginia
4. Boccaletti D, Pucacco G (1999) *Theory of Orbits*. Vol.1–2, Springer-Verlag, Berlin, Heidelberg
5. Kozai Y (1959) The motion of a close earth satellite. *Astron. J.*, 64 (9): 367–377
6. Liu L (1992) *Orbital Dynamics of Earth's Artificial Satellite*. Higher Education Press, Beijing
7. Liu L (2000) *Orbital Theory of Spacecraft*. National Defense Industry Press, Beijing
8. Liu L, Hu SJ, Wang X (2006) *Introduction of Aerospace Dynamics*. Nanjing University Press, Nanjing
9. Szebehely V (1967) *Theory of Orbit: The Restricted Problem of Three Bodies*. Academic Press, New York, London
10. Brown EW (1896) *An Introductory Treatise on Lunar Theory*. Cambridge University Press
11. Murray CD, Dermott SF (1999) *Solar System Dynamics*. Cambridge University Press, 1999
12. Gómez G et al (2001) *Dynamics and Mission Design near Libration Points*, Vol. 1–4. World Scientific, Singapore, New Jersey, London, Hong Kong
13. Hou XY, Liu L (2008) Dynamical characteristics of collinear Lagrangian points and the application in the deep space exploration, *Journal of Astronautics*, 2008, 29(3): 461–466
14. Liu L, Hou XY (2012) *Orbital Dynamics of Deep Spacecraft*. Electronic Industry Press, Beijing
15. Whipple AL, Szebehely V (1984) The Restricted Problem of $n + v$ Bodies. *Celest Mech* 32(2):137–144
16. Whipple AL (1984) Equilibrium Solutions of the Restricted Problem of 2+2 Bodies. *Celest Mech* 33(3):271–294

Contents

1	Selections and Transformations of Coordinate Systems	1
1.1	Time Systems and Julian Day [1, 2]	2
1.1.1	Selection of Standard Time	2
1.1.2	Time Reference Systems	3
1.1.3	Julian Day	5
1.2	Space Coordinate Systems [2–6]	6
1.3	Earth’s Coordinate Systems [2, 6–10]	10
1.3.1	The Realization of the Dynamical Reference System and J2000.0 Mean Equatorial Reference System	10
1.3.2	The Intermediate Equator and Three Related Datum Points	11
1.3.3	Three Geocentric Coordinate Systems	12
1.3.4	Transformation of the Earth-Fixed Coordinate System O -XYZ and the Geocentric Celestial Coordinate System O -xyz	14
1.3.5	Relationship Between the IAU 1980 Model and the IAU 2000 Model	23
1.3.6	The Complicity in the Selection of Coordinate System Due to the Wobble of Earth’s Equator	24
1.3.7	Coordinate Systems Related to Satellite Measurements, Attitudes, and Orbital Errors	24
1.4	The Moon’s Coordinate Systems	24
1.4.1	Definitions of the Three Selenocentric Coordinate Systems [6]	25
1.4.2	The Moon’s Physical Libration	26
1.4.3	Transformations Between the Three Selenocentric Coordinate Systems	30
1.5	Planets’ Coordinate Systems	34
1.5.1	Definitions of Three Mars-Centric Coordinate Systems	34

1.5.2	Mars's Precession Matrix	36
1.5.3	Transformation of the Mars-Centric Equatorial Coordinate System and the Mars-Fixed Coordinate System	37
1.5.4	Transformation of the Geocentric Coordinate System and the Mars-Centric Coordinate System	37
1.5.5	An Explanation of the Application of the IAU 2000 Orientation Models of Celestial Bodies	40
	References	40
2	The Complete Solution for the Two-Body Problem	43
2.1	Six Integrals of the Two-Body Problem	43
2.1.1	The Angular Momentum Integral (the Areal Integral)	45
2.1.2	The Orbital Integral in the Motion Plane and the Vis Viva Formula	45
2.1.3	The Sixth Motion Integral: Kepler's Equation	49
2.2	Basic Formulas of the Elliptical Orbital Motion	51
2.2.1	Geometric Relationships of the Orbital Elements in the Elliptical Motion	51
2.2.2	Expressions of the Position Vector $\vec{r}$ and Velocity $\dot{\vec{r}}$	51
2.2.3	Partial Derivatives of Some Variables with Respect to Orbital Elements	56
2.2.4	Derivatives of M , E , and F with Respect to Time t	59
2.3	Expansions of Variables in the Elliptical Orbital Motion	60
2.3.1	Expansions of $\sin kE$ and $\cos kE$	61
2.3.2	Expansions of E , r/a , and a/r	62
2.3.3	Expansions of $\sin F$ and $\cos F$	62
2.3.4	The Expansion of F	63
2.3.5	Expansions of $\left(\frac{r}{a}\right)^n \cos mf$ and $\left(\frac{r}{a}\right)^n \sin mf$	63
2.3.6	Expansions of $\left(\frac{a}{r}\right)p$, E , and $(F - M)$ in the Trigonometric Function of F	66
2.4	Transformations from the Orbital Elements to the Position Vector and Velocity and Vice Versa	67
2.4.1	Calculations of the Position Vector $\vec{r}(t)$ and Velocity $\dot{\vec{r}}(t)$ from Orbital Elements $\sigma(t)$	67
2.4.2	Calculations of the Orbital Elements $\sigma(t)$ from $\vec{r}(t)$ and $\dot{\vec{r}}(t)$	67
2.4.3	Calculations of Orbital Elements $\sigma(t_0)$ from Two Position Vectors $\vec{r}(t_1)$ and $\vec{r}(t_2)$	69
2.4.4	Method to Solve Kepler's Equation	70
2.5	Expressions and Calculations of Satellite Orbital Variables	71
2.5.1	Two Expressions of the Longitude of Satellite's Orbital Ascending Node	71

2.5.2	Expressions of Satellite's Position Measurements from a Ground-Based Tracking Station	72
2.5.3	Equatorial Coordinates of the Sub-Satellite Point	73
2.5.4	Satellite's Orbital Coordinate System	74
2.5.5	Expressions of Errors in Satellite Position	75
2.6	Parabolic Orbit and Hyperbolic Orbit	75
2.6.1	The Parabolic Orbit	76
2.6.2	The Hyperbolic Orbit	76
2.6.3	Formulas for Calculating the Position Vector and Velocity	79
	References	79
3	Analytical Methods of Constructing Solution of Perturbed Satellite Orbit	81
3.1	The Method of the Variation of Arbitrary Constants Applied to the Perturbed Two-Body Problem	81
3.2	Common Forms of Perturbed Motion Equation	84
3.2.1	Perturbed Motion Equations Formed by Accelerations of the (S, T, W) -Version and the (U, N, W) -Version	84
3.2.2	The Perturbation Motion Equations Formed by $\partial R/\partial \sigma$ -Version	87
3.2.3	Canonical Equations of Perturbation Motion	88
3.2.4	Singularities in the Perturbation Equations	88
3.3	Perturbation Method of Constructing Power Series Solution with a Small Parameter	94
3.3.1	Perturbation Equations with a Small Parameter	94
3.3.2	Existence of Power Series Solution with a Small Parameter	95
3.3.3	Construction of the Power Series Solution with a Small Parameter: The Perturbation Method	96
3.3.4	Secular Variations and Periodic Variations	99
3.4	An Improved Perturbation Method: The Method of Mean Orbital Elements	100
3.4.1	Introduction of the Method of Mean Orbital Elements	101
3.4.2	The Mean Values of Related Variables in an Elliptic Motion	103
3.4.3	Construction of Formal Solution: The Method of Mean Orbital Elements [3–8]	106
3.4.4	Example	109
3.4.5	Two Annotations About the Method of Mean Elements	111
3.5	The Method of Quasi-Mean Elements: The Structure of the Formal Solution	113

3.5.1	Small Divisors in Expressions of Perturbation Solutions	113
3.5.2	Configuration of Formal Solution: The Method of Quasi-Mean Elements	115
3.6	Methods of Constructing Non-singularity Solutions for a Perturbed Orbit	117
3.6.1	Configuration of the Non-singularity Perturbation Solutions of the First Type	118
3.6.2	Configuration of the Non-singularity Perturbation Solutions of the Second Type	124
	References	126
4	Analytical Non-singularity Perturbation Solutions for Extrapolation of Earth's Satellite Orbital Motion	129
4.1	The Complete Dynamic Model of Earth's Satellite Motion	129
4.1.1	Selection of Calculation Units in Satellite Orbit Dynamics	130
4.1.2	Analyses of Forces on Satellite's Orbital Motion	132
4.1.3	Further Analyses of the Forces Acting on a Satellite ...	136
4.2	The Perturbed Orbit Solution of the First-Order Due to Earth's Dynamical Form-Factor J_2 Term	137
4.2.1	The Perturbed Orbit Solution of the First-Order in Kepler Orbital Elements [1–5]	137
4.2.2	The Non-singularity Perturbation Solution of the First Type	146
4.2.3	The Non-singularity Perturbation Solution of the Second Type	157
4.3	The Perturbed Orbit Solution of the First-Order Due to Earth's Ellipticity $J_{2,2}$ Term	160
4.3.1	The Perturbed Orbit Solution of the First-Order in Kepler Orbital Elements	160
4.3.2	The Non-singularity Perturbation Solution of the First Type	165
4.3.3	The Non-singularity Perturbation Solution of the Second Type	167
4.4	Additional Perturbation of the Coordinate System for the First-Order Solution	168
4.4.1	The Cause of the Additional Perturbation of the Coordinate System [3, 8]	168
4.4.2	The Additional Perturbation Solution in Kepler Orbital Elements	170
4.4.3	The Non-singularity Additional Perturbation Solution of the First Type	173
4.4.4	The Non-singularity Additional Perturbation Solution of the Second Type	176

4.4.5	Selection of Coordinate System and Related Problems	176
4.5	The Perturbation Orbit Solution Due to the Higher-Order Zonal Harmonic Terms J_l ($l \geq 3$) of Earth's Non-spherical Gravitation	178
4.5.1	General Expression of the Perturbation Function of the Zonal Harmonic Terms J_l ($l \geq 3$)	178
4.5.2	The Perturbation Solution of the Zonal Harmonic J_l ($l \geq 3$) Terms	181
4.5.3	The Non-singularity Perturbation Solution of the First Type by the Zonal Harmonic Terms J_l ($l \geq 3$)	185
4.5.4	The Non-singularity Perturbation Solution of the Second Type by Zonal Harmonic Terms J_l ($l \geq 3$)	186
4.5.5	The Perturbation Solution of the Main Zonal Harmonic Terms J_3 and J_4 in Kepler Elements	188
4.6	The Perturbation Solution Due to the High-Order Tesseral Harmonic Terms $J_{l,m}$ ($l \geq 3, M = 1, 2, \dots, l$) of Earth's Non-spherical Gravitation	196
4.6.1	The General Expression of the Perturbation Function of the Tesseral Harmonic Terms $J_{l,m}$ ($l \geq 3, M = 1, 2, \dots, l$)	196
4.6.2	The Perturbation Solution Due to the Tesseral Harmonic Terms $J_{l,m}$ ($l \geq 3, m = 1 - l$)	197
4.6.3	The Non-singularity Perturbation Solution of the First Type Due to the Tesseral Harmonic $J_{l,m}$ ($l \geq 3, m = 1 - l$) Terms	200
4.6.4	The Non-Singularity Perturbation Solution of the Second Type Due to the Tesseral Harmonic $J_{l,m}$ ($l \geq 3, m = 1 - l$) Terms	200
4.6.5	The Perturbation Solution Due to the Tesseral Terms, $J_{3,m}$ ($M = 1, 2, 3$) and $J_{4,m}$ ($M = 1, 2, 3, 4$) in Kepler Elements	201
4.6.6	The Non-singularity Perturbation Solution of the First Type Due to the Tesseral Harmonic Terms $J_{3,m}$ ($M = 1, 2, 3$) and $J_{4,m}$ ($M = 1, 2, 3, 4$)	211
4.6.7	The Non-singularity Perturbation Solution of the Second Type Due to the Tesseral Harmonic Terms $J_{3,m}$ ($m = 1, 2, 3$) and $J_{4,m}$ ($m = 1, 2, 3, 4$)	211
4.7	The Perturbed Orbit Solution Due to the Gravitational Force of the Sun or the Moon	211
4.7.1	The Perturbation Function and Its Decomposition	212
4.7.2	The Perturbation Solution Due to the Gravity of the Sun or the Moon	215

4.8	The Perturbed Orbit Solution Due to Earth's Deformation	230
4.8.1	Expression of the Additional Potential of Tidal Deformation	231
4.8.2	Effect of the Main Term in the Additional Tidal Deformation Potential (the Second-Order Term of $l = 2$) on a Satellite Orbit	232
4.9	Post-Newtonian Effect on the Orbital Motion	235
4.9.1	The Post-Newtonian Effect	235
4.9.2	Perturbation Solution Due to the Post-Newtonian Effect	237
4.9.3	Other Post-Newtonian Effects on the Earth's Artificial Satellite Motion	239
4.10	Perturbed Orbit Solution Due to the Solar Radiation Pressure	239
4.10.1	Calculation of Radiation Pressure	239
4.10.2	Two States of Radiation Pressure Perturbation	242
4.10.3	The Perturbation Solution Due to Radiation Pressure	243
4.10.4	The Non-singularity Perturbation Solution of the First Type Due to the Radiation Pressure	252
4.10.5	The Non-singularity Perturbation Solution of the Second Type Due to the Radiation Pressure	254
4.11	Perturbed Orbit Solution Due to Atmospheric Drag	256
4.11.1	Damping Effect: Atmospheric Drag	256
4.11.2	Atmosphere Density Model	258
4.11.3	Atmospheric Rotation and the Expression of Atmospheric Drag	262
4.11.4	Structure of the Perturbed Solution Due to the Atmospheric Drag	265
4.11.5	The Non-singularity Perturbation Solution by the Atmospheric Drag	273
4.12	Orbital Variations Due to a Small Thruster	274
4.12.1	The Perturbation Solution Due to an (S, T, W) -Type Thrust	275
4.12.2	The Non-singularity Perturbation Solution Due to an (S, T, W) -Type Thrust	277
4.12.3	The Perturbation Solution by a U -type Thrust	281
4.12.4	The Non-singularity Perturbation Solution Due to a U -type Thrust	283
	References	285
5	Satellite Orbit Design and Orbit Lifespan Estimation	287
5.1	Sidereal Period and Nodal Period [1–3]	287
5.1.1	The Transformation Between the Sidereal Period T_s and the Nodal Period T_φ	288

5.1.2	The Anomalistic Period	292
5.2	Orbital Characteristics of Polar Orbit Satellite [2, 3]	293
5.2.1	Basic Theories	293
5.2.2	Preservation of Polar Orbit	295
5.3	Existence and Design of Sun-Synchronous Orbit [2–5]	296
5.3.1	Conditions of Forming a Sun-Synchronous Orbit	296
5.3.2	Sun-Synchronous Orbits for Different Celestial Bodies	299
5.4	Existence and Design of Frozen Orbit [2–5]	300
5.4.1	Basic State of Frozen Orbit	300
5.4.2	Basic Equations of a Possible Frozen Orbit	301
5.4.3	A Particular Solution of Eq. (5.40): The Frozen Orbit	302
5.4.4	Stability of Frozen Orbit	303
5.4.5	Frozen Orbit for Other Celestial Bodies	306
5.4.6	Characteristics and Applications of Satellite Orbit with a Critical Inclination	306
5.5	Existence and Design of Central Body Synchronous Orbit	309
5.5.1	Basic State of Central Body Synchronous Satellite Orbit [2, 3, 5]	309
5.5.2	Existence and Evolution of a Central Body Synchronous Satellite (Earth, Mars)	312
5.6	Estimation and Calculation of Satellite's Lifespan Due to the Mechanism of Gravitational Perturbation	316
5.6.1	Definition and Mechanism of a Low Orbit Satellite Lifespan Due to Gravitational Perturbations [6–10]	317
5.6.2	Overview of Low Orbit Satellite Lifespan for Earth, the Moon, Mars, and Venus	319
5.6.3	Evolution Characteristics and Lifespans of Orbit with a Large Eccentricity [2, 9]	323
5.6.4	Evolution Characteristics and Lifespans of High Earth Satellite Orbit [6, 10]	330
5.6.5	Key Points About Estimating Satellite Orbit Lifespan Due to Gravitational Perturbations	332
5.7	Estimation and Calculation of Satellite Orbit Lifespan in the Perturbed Mechanism of Atmospheric Drag	332
	References	334
6	Orbital Solutions of Satellites of the Moon, Mars, and Venus	335
6.1	Characteristics of Gravitational Fields of Earth, the Moon, Mars, and Venus	335
6.1.1	Basic Characteristics of Earth's Gravity Potential	335
6.1.2	Basic Characteristics of the Moon's Gravity Potential	337
6.1.3	Basic Characteristics of Mars's Gravity Potential	338

6.1.4	Basic Characteristics of Venus's Gravity Potential	339
6.2	Perturbed Orbital Solution of the Moon's Satellite	340
6.2.1	Selection of Coordinate System	341
6.2.2	Mathematical Model for the Perturbed Motion of the Moon's Satellite	342
6.2.3	The Numerical Solution for the High Precise Orbital Extrapolation	345
6.2.4	The Analytical Perturbation Solution of the Moon's Satellite Orbit	346
6.2.5	Additional Perturbation of Coordinate System [5, 6]	364
6.2.6	Applications of Analytical Orbital Solution in Orbital Design	366
6.3	Perturbed Orbital Solution of Mars's Satellite	371
6.3.1	Selection of Coordinate System	372
6.3.2	The Mathematical Model of Perturbed Motion for a Mars's Satellite	372
6.3.3	The Analytical Perturbation Solution of Mars's Satellite Orbit [7, 8]	374
6.4	Perturbed Orbital Solution of Venus's Satellite	383
6.4.1	The Perturbation Function of Venus's Non-Spherical Gravity Potential	384
6.4.2	The Structure and Results of the Analytical Perturbation Solution	386
	References	387
7	Orbital Motion and Calculation Method in the Restricted Three-Body Problem	389
7.1	Selection of Coordinate System and Motion Equation of a Small Body	389
7.1.1	The Motion Equation of a Small Body in the Barycenter Inertial Coordinate System	391
7.1.2	The Motion Equation of a Small Body in the Synodic Coordinate System	393
7.2	Jacobi Integral and Solution Existence of the Circular Restricted Three-Body Problem	395
7.2.1	Jacobi Integral in the Circular Restricted Three-Body Problem	395
7.2.2	Existence of Solution of the Circular Restricted Three-Body Problem	396

7.3	Calculation and Application of the Libration Point	
	Positions of the Circular Restricted Three-Body Problem	397
7.3.1	Conditions of Existence for Libration Solutions	398
7.3.2	The Positions of the Three Collinear Libration Points	399
7.3.3	Two Triangle Libration Points	401
7.3.4	Dynamical Characteristics of the Five Libration Points	402
7.3.5	Characteristics and Applications of the Stability of the Five Libration Points	410
7.3.6	Calculations and Applications of Libration Points in the Restricted Problem of (2+2) Bodies	419
7.4	Orbit Design for Formation Flying of Satellites and Companion-Flying in the Exploration of Asteroids	422
7.4.1	The Principle of Satellite Formation Flying	422
7.4.2	The Problem with the Eccentricity in Orbit Design of Formation Flying of Satellites	426
7.4.3	Extension of the Principle and Related Orbit Design Method of Satellite Formation Flying	428
7.4.4	Orbital Problem of Companion Flying in Asteroid Exploration	428
7.5	Geometric Characteristics of Libration Point Orbit and Orbit Determination	430
7.5.1	Geometric Characteristics of Libration Point Orbits	430
7.5.2	Analysis of Forces on a Prober's Motion in a Libration Orbit	430
7.5.3	Orbit Determination and Forecast Method of Libration Point Orbit	435
7.5.4	Orbit Determination of Libration Point Orbit and Precision Examination of Short-Arc Forecast [22]	436
7.5.5	Orbital Transformation Between the Two Coordinate Systems for a Libration Point Orbit Prober	439
	References	439
8	Numerical Method for Satellite Orbit Extrapolations	441
8.1	Basic Knowledge of Numerical Method in Solving the Motion Equation	442
8.1.1	Basic Principles of Numerical Method in Solving Motion Equation	442
8.1.2	Basic Concepts	443

8.2	Conventional Singer-Step Method: The Runge–Kutta Method	445
8.2.1	The Fourth-Order RK Method (RK4)	446
8.2.2	The Runge–Kutta-Fehlberg (RK4) Method	447
8.3	Linear Multistep Methods: Adams Method and Cowell Method	450
8.3.1	Adams Methods: Explicit Methods and Implicit Methods	450
8.3.2	Cowell’s Method and Størmer’s Method	453
8.3.3	Adams-Cowell Method	457
8.4	Key Issues in Applications of the Numerical Method in Orbital Dynamics	459
8.4.1	Selections of Variables and Corresponding Basic Equations	459
8.4.2	Singularity Problem	463
8.4.3	Homogenization of Step-Size	463
8.4.4	Control of the Along-Track Errors	464
8.5	Numerical Calculation of the Right-Side Function	465
8.5.1	The Perturbation Acceleration of the Zonal Harmonic Term $\vec{F}_1(J_l)$	467
8.5.2	The Perturbation Acceleration of the Tesseral Harmonic Term $\vec{F}(J_{l,m})$	468
8.5.3	The Recursive Formulas of Legendre Polynomials, $P_l(\mu)$ and the Associated Legendre Polynomials $P_{l,m}(\mu)$, and Their Derivatives [15, 16]	469
8.5.4	The Perturbation Acceleration of the Tidal Deformation $\vec{F}(k_2, J_{2,m})$	470
8.6	The Role of the Hamiltonian Method in the Orbital Evolution	471
	References	472
9	Formulation and Calculation of Initial Orbit Determination	473
9.1	Formulation of Orbit Determination	473
9.2	A Review of Initial Orbit Calculation in the Sense of the Two-Body Problem	475
9.2.1	Basic Conditions for Initial Orbit Determination	475
9.2.2	Construction of the Basic Equation for an Initial Orbit	476
9.3	Initial Orbit Determination for Perturbed Motion	478
9.3.1	Construction of the Basic Equation for Initial Orbit Determination	478
9.3.2	Initial Orbit Determination Using Angle Data Over a Short-Arc	480
9.3.3	Initial Orbit Determination Using (P, A, h) Data or Navigation Information	487

9.3.4	Examination of Orbit Determination Method Using Actual Measurements	488
9.3.5	Initial Orbit Determination When a Deep-Space Prober is on a Transfer Orbit	489
9.3.6	Initial Orbit Determination Using Space-Based Angle Measurements (α, δ)	494
9.3.7	A Brief Summary of Initial Orbit Determination	496
	References	497
10	Precise Orbit Determination	499
10.1	Precise Orbit Determination: Orbit Determination and Parameter Estimation	499
10.2	Theoretical Calculation of Measurement Variables	503
10.3	Calculation of Transformation Matrixes	508
10.3.1	Matrix $\frac{\partial Y}{\partial(\vec{r}, \dot{\vec{r}})}$	509
10.3.2	Matrix $\left(\frac{\partial(\vec{r}, \dot{\vec{r}})}{\partial \sigma}\right)$	510
10.3.3	State Transition Matrix Φ	514
10.4	Estimation of the State Variable: Calculation of Precise Orbit Determination	521
10.4.1	Certainty of Solution in the Orbit Determination	521
10.4.2	Process of Calculating Solution in the Orbit Determination	522
10.5	The Least Squares Estimator and Its Application in Precise Orbit Determination	523
10.5.1	Estimation Theory and a Few Commonly Used Optimal Estimation Methods	524
10.5.2	The Least Squares Estimator	526
10.5.3	Two Processes of the Least Squares Estimator	530
10.5.4	Least Squares Estimator with a Priori State Value	531
10.6	Orbit Determination by Ground-Based and Space-Based Joint Network and Autonomous Orbit Determination by Star-To-Star Measurements	532
10.6.1	Outline of Space-Based Network of Orbit Tracking and Determination	534
10.6.2	Basic Principles of the Orbit Determination of Ground-Based and Space-Based Joint Network	534
10.6.3	The Rank Deficiency in the Autonomous Orbit Determination by Start-To-Star Measurements	536
	References	539

Appendix A: Astronomical Constants	541
Appendix B: Formulas of Mean Ephemeris of Major Celestial Bodies in the Solar System	543
Appendix C: Orientation Models of Major Celestial bodies in the Solar System	549
References	553

About the Author

Lin Liu is a Chinese Astronomer currently a Distinguished Professor at the Astronomy and Space Science Institute, Nanjing University, China. He is an expert in Celestial Dynamics and Spacecraft Orbital Determination both in theory and application. He is regarded as the main founder of this field in China and a highly respected teacher over 50 years. His work has profound influence especially on the Chinese Aerospace Industry. He has been involved in many important Chinese Space programs such as the Shenzhou Spacecraft, the Moon Exploration, etc.

In recent decades, Professor Liu has worked on the Orbital Dynamics of deep space exploration. Many of his initial and ground-breaking research results are connected to the Chinese Aerospace Industry. He was in charge of several Aerospace research projects and research programs of the National Natural Science Foundation of China. He was a director of the Chinese Astronomy Society, a director of the Celestial Dynamics and Satellite Dynamics sections, and a director of the Chinese Aerospace Society. Currently, he is on the editorial board of the Chinese Astronomical Journal, a member of the Academic Committee of the Deep Space Exploration Joint Center, Ministry of Education, an external expert for the National Astronomical Observatory of Chinese Academy of Science, and a member of the Chinese Committee of COSPAR.

Professor Liu has more than 250 research publications in national and international journals and 11 monographs. There is a long list of awards Professor Liu has received including the National Science Congress Major Achievement Award in 1978, the Chinese Astronomical Society Zhang Yuzhe Award, and three times of the State Educational Commission awards, to name a few.

In April of 2016, for recognising Professor Liu's scientific achievements, the International Astronomy Union (IAU) named Asteroid 261936 Liu in his honor.

Chapter 1

Selections and Transformations of Coordinate Systems



The main content of orbital dynamics is about solving a dynamical problem. The first step of solving a specific dynamical problem is to select a proper spatial reference frame and a time reference system. A small body, most likely an artificial satellite or a specific spacecraft, moves in an orbit. The orbit then can be presented in a reference system, such as the Earth reference system, the Moon reference system, a planet's reference system, and the heliocentric reference system. This chapter introduces these reference systems and their relationships, and the formulas for mutual transformations.

According to the general relativity theory, a reference system is a 4-dimensional space–time system. In the Solar System, there are two important inertial reference systems. One system is centered at the barycenter of the Solar System, and its orientation is decided by remote quasars. This system is the barycenter reference system of the Solar System called the Barycentric Celestial Reference System (**BCRS**) and is related to all motions of celestial bodies in the Solar System. The other system is centered at Earth's barycenter and is called the Geocentric Celestial Reference System (**GCRS**). This system is related to all motions around and on Earth including the observers on the ground. The fourth demission of a reference system is the time variable called the **Coordinate Time**, whose variation is relative to the local gravitational field. Therefore, there are two Coordinate times for the two reference systems, the Barycentric Coordinate Time (**TCB**) for **BCRS**, and the Geocentric Coordinate Time (**TCG**) for **GCRS**. These are theoretical definitions, in the application, there can be some changes. In this chapter, the commonly used space coordinate systems and time systems are introduced and discussed.

1.1 Time Systems and Julian Day [1, 2]

As mentioned above the two coordinate times are **TCB** and **TCG**, but for making ephemerides and in motion equations the time variables are measured by the Barycentric Dynamical Time (**TDB**) and the Terrestrial Time (**TT**) for BCRS and GCRS, respectively. The terrestrial time was called Terrestrial Dynamical Time (**TDT**) but changed to **TT** after 1991. The difference between the two time systems, **TDB** and **TT**, is caused by the effect of relativity, and the transfer relationship can be defined by the theory of gravitation. In practical applications, their relationship is given by the International Astronomy Union (**IAU**) in 2000 as

$$\text{TDB} = \text{TT} + 0^{\text{s}}.001657 \sin g + 0^{\text{s}}.000022 \sin(L - L_J), \quad (1.1)$$

where g is the mean anomaly of Earth's orbit around the Sun, and $(L - L_J)$ is the difference of the Sun's mean ecliptic longitude and Jupiter's mean ecliptic longitude that

$$\begin{cases} g = 357^{\circ}.53 + 0^{\circ}.98560028t, \\ L - L_J = 246^{\circ}.00 + 0^{\circ}.90251792t, \end{cases} \quad (1.2)$$

$$t = \text{JD}(t) - 2451545.0. \quad (1.3)$$

In (1.3), $\text{JD}(t)$ is the related Julian Day Number of time t , the definition of Julian Day is given in Sect. 1.1.3. Formula (1.1) is valid between 1980 and 2050, and the error is less than $30 \mu\text{s}$ (10^{-6} s). Near Earth's surface with errors in the order of ms (10^{-3} s), there is approximately

$$\text{TDB} = \text{TT}. \quad (1.4)$$

Modern space and time reference systems accept IAU 2009 Astronomy Constant System (Appendix 1). In this system, the Astronomical Unit (AU) is provided by IAU 2012 resolution and is directly related to the unit of length "meter (m)". The value of AU is now given by

$$1\text{AU} = 1.49597870700 \times 10^{11} \text{ m}. \quad (1.5)$$

1.1.1 Selection of Standard Time

In the TT system, the time is realized by atomic time. The earliest atomic clock, which used the period of atomic oscillation as the standard to measure time, was built in 1949. In 1967, International System of Units (**SI**) defined the base unit of time (a

second) as the duration of 9,192,631,770 cycles of the radiation corresponding to the transition between two energy levels of the ground state of the cesium-133 atom. In 1997, the International Committee for Weights and Measures (**EIPM**) added that the preceding definition refers to a cesium atom at rest at a temperature of absolute zero. The atomic time (**TAI**, in French Temps Atomique International) uses the SI second (s) as a unit and the universal time of 1 January 1958, 0:00:00 as the starting epoch. Since 1971, TAI is provided by the International Bureau Weights and Measures (**BIPM**, in French) as a weighted average of the time kept by over 400 atomic clocks in over 50 national laboratories worldwide. The only difference between TT and TAI is the starting points that

$$TT = TAI + 32^s.184. \quad (1.6)$$

To the present day, TAI is the most accurate and uniform standard time, its accuracy is about 10^{-16} s, and its error would be less than 1 s over one billion years.

1.1.2 Time Reference Systems

To study the motions of celestial bodies including spacecraft, it is necessary to have a time system with a uniform time scale. For an observatory on the surface of Earth, it is also necessary to have a time system related to Earth's rotation. Before the atomic time became the standard time Earth's rotation was the time basis for the two time systems. But Earth's rotation is non-uniform and the accuracy of measurements of Earth's rotation is continuously improving. Therefore, there is a need to build a time system that uses the uniform unit scale TAI and is also related to Earth's rotation in a coordinative way.

Sidereal Time (ST). The definition of a “sidereal day” is the time interval for the March equinox at the upper transit between two successive returns. Therefore, **ST** is the angle measured along the celestial equator from the observer's meridian (at longitude λ) to the great circle that passes through the March equinox and both poles, and is given by time. Its value (S) equals the right ascension (α) of a star at the upper transit of the observatory, that

$$S = \alpha, \quad (1.7)$$

where S is the local sidereal time (**LST**), and the Greenwich sidereal time (S_G , **GST**) is given by

$$S_G = S - \lambda. \quad (1.8)$$

In fact, **GST** is the Greenwich apparent sidereal time and is different from the Greenwich mean sidereal time (**GMST**). Because the sidereal time is defined by

Earth's rotation, then the non-uniformness of Earth's rotation can be measured by the difference between the sidereal time and an unformed time.

Universal Time (UT). It is, like ST, a time measured according to Earth's rotation but chooses the mean solar day as its unit, therefore one second of UT is $1/86400$ mean solar day. The astronomically measured universal time, UT0, is relative to the instantaneous polar meridian. UT0 is affected by the motion of the pole, to provide the universal time with respect to the mean pole, denoted by UT1, a correction is needed that

$$UT1 = UT0 + \Delta\lambda, \quad (1.9)$$

where $\Delta\lambda$ is a correction of the polar shift.

UT1 is not a uniform time scale because of the non-uniformness of Earth's rotation. There are three types of Earth's rotation variation. The first is the slow long-term variation (the universal day increases 1.6 ms per 100 years); the second is the periodic variation (mainly the seasonal variation, about 0.001 s in a year, and some other smaller periodic variations); and the third is the irregular variations. These variations cannot be easily corrected, only the annual variation can be given by an empirical formula based on multi-year observations. If the annual variation is denoted by ΔT_s , then the adjusted universal time, UT2, is

$$UT2 = UT1 + \Delta T_s. \quad (1.10)$$

UT2 is a relatively uniform time scale, although it still includes the long-term variation of Earth's rotation and the irregular variations. The physical cause of the irregularity is unknown therefore there is no way to adjust.

For general requirements UT1 is commonly accepted as a united time system because its periodic variation ΔT_s is rather small, also it is directly related to Earth's instantaneous position. For high precision problems, even UT2 is not precise enough, and a more uniform time scale is required, thus it is necessary to introduce the atomic time TAI as the time basis.

As mentioned above, TAI is defined at 1 January 1958, 0 h, which is very close to the UA2's starting time, the difference is

$$(TAI - UT2)_{1958.0} = -0^s.0039. \quad (1.11)$$

TAI is defined in the geocentric coordinate system and is measured by the international time unit. Since 1984, the ephemeris time system (ET) has been formally replaced by TAI which became the uniform scale required by researchers in the field of Dynamics. The Terrestrial Dynamical Time (TDT), therefore, was introduced (renamed as terrestrial time TT in 1991). The epoch 1 January 1977, 0:00:00 by TAI corresponds to January 1^d.0003725, 1977 by TDT. This difference equals the difference between ET and TAI at that moment. With the definition of the beginning of TT, it is easy to use the TT system to replace the ET system.

Coordinated Universal Time (UTC). A uniform time system is appropriate for high precession ephemerides, which require a uniform scale for time intervals, but it cannot replace a time system related to the Earth's non-uniform rotation. Thus, the coordinated universal time (UTC) is introduced for solving this problem. To the present day there are many suggestions, discussions, and arguments, but without a definite conclusion. We still keep the sidereal time system and the universal time system, as each of them has its merits.

The difference between TAI and UT2 (or UT1), given in (1.11), is $0^s.0039$ at the beginning of 1 January 1958, it was near zero. Because Earth's rotation has a long-term slowness, the difference between TAI and UT2 increases then causes the problem. In order to keep UT (UT1 or UT2) as close as possible to TAI and still, use the uniform scale. In 1963, the international communities adopted a third-time system, which is the coordinated Universal Time (UTC). UTC is still a time system based on TAI but with leap seconds added at irregular intervals, such as 12 months or 18 months, to make it as close as possible to UT. Since 1972 it has been required that UTC must be kept within ± 0.9 s of UT1. Actual adjustments of leap seconds are given by the International Time Bureau based on observational information, which can be found on the EOP web page. Until 1 January 2017, the adjustment is 37^s , that

$$\text{TAI} = \text{UTC} + 37^s.$$

The transformation process from UTC to UT1 is that first to download the newest EOP (Earth Orientation Parameters) data (use the B data if the time is more than a month earlier than present, and use the A data for other times), then to calculate the adjustment ΔUT by interpolation, which gives UT1 as

$$\text{UT1} = \text{UTC} + \Delta\text{UT}. \quad (1.12)$$

According to the international convention, if a measurement is given at time t , the time system means UTC unless there is a special description.

1.1.3 Julian Day

Besides the time system, in solving the dynamical problem we often have to choose an epoch and deal with the problem related to the length of different types of one year. In Astronomy, there are several definitions of a year. One is the Besselian year, which has a length of a tropical year, i.e., 365.2421988 mean solar days. The epoch of a Besselian year is the moment when the Sun's mean ecliptic longitude is 280° . For example, the Besselian year 1950.0 does not mean 1 January 1950, 0:00:00 but is 31 December 1949, 22:09:42 (UT), which corresponds to the Julian day number (JDN) 2,433,282.4234. Another type of year is the Julian year, which has 365.25 mean solar days. The epoch of each Julian year is exactly the beginning of a year, for example, 1950.0 means 1 January 1950, 00:00:00. Obviously, it is easier to use

Table 1.1 Besselian Epoch, Julian Epoch, and Julian day number

Besselian Epoch	Julian Epoch	Julian day number
1900.0	1900.000858	2,415,020.3135
1950.0	1949.999790	2,433,282.4234
2000.0	1999.998722	2,451,544.5333
1989.999142	1900.0	2,415,020.0
1950.000210	1950.0	2,433,282.5
2000.001278	2000.0	2,451,545.0

Julian years than Besselian years. Therefore, since 1984 Besselian year has been replaced by the Julian year. Some correspondences of the Besselian epoch, Julian epoch, and Julian day number are listed in Table 1.1.

For convenience, the Modified Julian Date (MJD) is introduced and defined as

$$\text{MJD} = \text{JD} - 2400000.5. \quad (1.13)$$

As an example, $\text{JD}(1950.0)$ corresponds to $\text{MJD} = 33,282.0$. The lengths of a century of a Besselian year (tropical century) and a Julian year are 36,524.22 and 36,525 mean solar days, respectively.

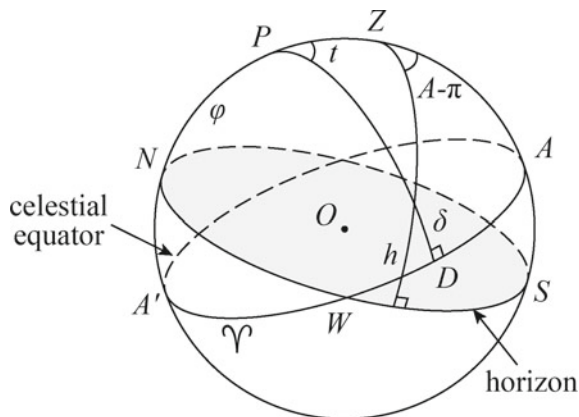
1.2 Space Coordinate Systems [2–6]

A coordinate system is actually a mathematical representation of a theoretical concept. A reference frame is the physical realization of a coordinate system, therefore, a reference system is an integrated system of a theoretical concept and a physical frame. Although the concept of a reference system is different from that of a coordinate system, in the practical application of most fields, as in this book, these two systems are interchangeable without misunderstanding.

To study the motions of celestial bodies in the Solar System, there are commonly accepted three types of the coordinate system, which are the horizontal coordinate system, the equatorial coordinate system, and the ecliptic coordinate system. These coordinates are applied to problems no matter from the point of view of Earth or other celestial bodies (such as the major planets or the Moon). For each space coordinate system, there are three key elements, the origin of the coordinate system, the fundamental plane, i.e., the xy -plane, and the primary direction (the direction of the x -axis). In this section, we introduce three coordinates with respect to Earth.

Horizontal system. A proper name for this system should be the topocentric horizontal coordinate system. In this system the origin is at the center of an observatory (or a sampling center), the fundamental plane is the local horizontal plane containing the origin and is tangential to the ellipsoid of Earth (the horizon), and the primary direction is towards the north (N) in the xy -plane. The direction of the z -axis is towards the zenith (Z) (Fig. 1.1).

Fig. 1.1 The horizontal system and the equatorial system



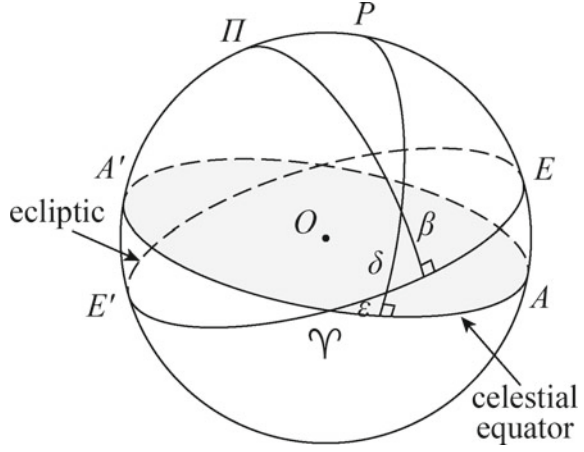
Equatorial system. There are two equatorial systems, one is the topocentric equatorial system with the center at the location of an observatory, and the other is the geocentric equatorial system with the center at the center of Earth. For both systems, Earth's equatorial plane is the reference plane, but for the topocentric system the reference plane is parallel to the equatorial plane, and in the celestial sphere the two planes converge into one, therefore the two systems are related only by a translation. The primary direction of both systems is towards the March equinox (Υ).

Ecliptic system. There are also two ecliptic systems, the geocentric ecliptic system with its origin at the center of Earth and the heliocentric ecliptic system with its origin at the center of the Sun. For both systems, the fundamental plane is the ecliptic plane of Earth's orbit around the Sun, and the primary direction is towards the March equinox (Υ).

The geometrical relationship of the horizontal system and the equatorial system is illustrated in Fig. 1.1, and that of the equatorial system and the ecliptic system in Fig. 1.2. The symbols in the figures are customarily used in Astronomy, therefore, are not explained here.

The position of a celestial body in a space coordinate system can be presented by its coordinate vector. In the horizontal coordinate system, the position vector of a body is denoted by $\vec{\rho}$ with spherical coordinates $(\rho, A, h, \text{ or } E)$, where ρ is the distance between the origin of the system and the body, A is the azimuth (do not be confused with the equator AA') measured from the north point eastward along the horizontal circle (clockwise), and h is the altitude (i.e. the height angle E). In the equatorial coordinate system, this vector is denoted by $\vec{r}$ with spherical coordinates (r, α, δ) , where r is the same as ρ , α is the right ascension measured from the March equinox eastward along the equator (i.e., the arc ΥD on the equator AA'), and δ is the declination angle. In the ecliptic coordinate system, the position vector is denoted by $\vec{R}$ with spherical coordinates (R, λ, β) , where R is the same as ρ , λ is the ecliptic longitude measured from the March equinox eastward along the ecliptic, and β is the ecliptic latitude. The relationships of the coordinates are given by

Fig. 1.2 The equatorial system and the ecliptic system



$$\vec{\rho} = \rho \begin{pmatrix} \cos h \cos A \\ -\cos h \sin A \\ \sin h \end{pmatrix}, \quad \vec{r} = r \begin{pmatrix} \cos \delta \cos \alpha \\ \cos \delta \sin \alpha \\ \sin \delta \end{pmatrix}, \quad \vec{R} = R \begin{pmatrix} \cos \beta \cos \lambda \\ \cos \beta \sin \lambda \\ \sin \beta \end{pmatrix}. \quad (1.14)$$

The azimuth A sometimes is measured from the south point (S) eastward along the horizontal circle (anti-clockwise), then $\vec{\rho}$ is given by

$$\vec{\rho} = \rho \begin{pmatrix} \cos h \cos A \\ \cos h \sin A \\ \sin h \end{pmatrix}. \quad (1.15)$$

In the topocentric equatorial coordinate system and the geocentric ecliptic coordinate system, the position vectors of a body can be presented by $\vec{r}'$ and $\vec{R}'$, respectively, and the corresponding relationship is similar to that for $\vec{r}$ and $\vec{R}$, but r and R should be replaced by r' and R' , respectively, then α and δ are for the topocentric equatorial system, and λ and β for the geocentric ecliptic system.

The transformation relationships between these coordinate systems are simple and only involve translations and rotations, that

$$\vec{r}' = R_z(\pi - S)R_y\left(\frac{\pi}{2} - \varphi\right)\vec{\rho}, \quad (1.16)$$

$$\vec{r} = \vec{r}' + \vec{r}_A, \quad (1.17)$$

$$\vec{R}' = R_z(\varepsilon)\vec{r}, \quad (1.18)$$

$$\vec{R} = \vec{R}' + \vec{R}_E, \quad (1.19)$$

where $S = \alpha + t$ is the hour angle of the March equinox, which equals the sidereal time at the observatory (i.e. the arc $\Upsilon D + DA$ measured along the equator circle AA' in Fig. 1.1), φ is the astronomical latitude of the observatory, $\vec{r}_A$ is the position vector of the observatory from the Earth's center, ε is the obliquity, and R_E is the position vector of Earth's center in the heliocentric coordinate system.

The rotation matrices R_x , R_y , and R_z in (1.16) and (1.18) are given by

$$R_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{pmatrix}, \quad (1.20)$$

$$R_y(\theta) = \begin{pmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{pmatrix}, \quad (1.21)$$

$$R_z(\theta) = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (1.22)$$

In the dynamics of the Solar System for studying the motions of the major planets and asteroids, we use the heliocentric ecliptic coordinate system; whereas in the dynamics of artificial satellites, we use the equatorial coordinate system centered at the barycenter of the main body, such as the geocentric equatorial system, or the Moon-centric equatorial system, or the Mars-centric equatorial system, etc.

The coordinate systems used for artificial satellites are mainly the geocentric celestial coordinate system and the Earth-fixed equatorial coordinate system (see 1.3.3). The origins of both systems are obviously at the center of Earth, but their fundamental planes and primary directions are affected by Earth's precession, nutation, and polar motion, which make these space coordinate systems rather complicated. As we know that Earth is an ellipsoid with unevenly distributed mass. The gravitational forces from the Sun, the Moon, and other major planets act on Earth's non-spherical part and produce two phenomena. One is an effect of a rigid body translation force, resulting in an indirect perturbation of Earth's oblateness. The other is an effect of rotation torque of a rigid body due to Earth's gyro-like motion, producing precession and nutation. Because of precession and nutation, Earth's equatorial plane vibrates. Besides the two effects, Earth's internal motion and motions on the surface produce a slow shift of the rotational axis, i.e., the polar motion, which also influences the selection of the coordinate system. The equators of Mars and the Moon have similar variations. As a result, there are different types of equatorial coordinate systems. The properties of equatorial coordinate systems with respect to Earth, the Moon, and Mars, and the transformation between these systems are discussed in the following sections.

1.3 Earth's Coordinate Systems [2, 6–10]

1.3.1 *The Realization of the Dynamical Reference System and J2000.0 Mean Equatorial Reference System*

Current observational measurements, such as Ephemerides of the planets in the Solar System, are all provided in the International Celestial Reference System (ICRS). The realization of this reference system is a reference frame called International Celestial Reference Frame (ICRF). The origin of ICRS is at the barycenter of the Solar System, the fundamental plane and the primary direction of X -axis are decided by precise observations of a group of extragalactic radio sources to be as close as possible to the J2000.0 mean equatorial plane and the mean March equinox, respectively. Here J2000.0 means the Julia year 2000, 1 January 12:00:00. Because the radio sources are so distant, they are stationary to our technology, the coordinate system and its orientation of ICRF are relatively fixed in space, and free of the dynamics of the Solar System and Earth's precession and nutation, also unrelated to the traditional concepts of the equator, March equinox, and ecliptic. As a result, this system is closer to an inertia reference system than any other system.

Before the existence of ICRS and ICRF, the basic astronomy reference system is the Fifth Fundamental Catalog dynamic system (**FK5**) (strictly speaking, this is a reference system dynamically defined and includes the correction of sidereal kinematics). FK5 is built up on observations of bright stars and the IAU 1976 Astronomic constants. Its fundamental plane is the J2000.0 mean equatorial plane, and the direction of X -axis points to J2000.0 mean March equinox. Obviously, this system is related to the epoch. The present system ICRS is an improvement of FK5. The dynamical reference system is the **J2000.0 mean equatorial reference system**, usually called **J2000.0 mean equatorial coordinate system**. Specifically, the fundamental plane and the primary direction of the X -axis of ICRF are realized by observations of the Very-Long-Baseline Interferometry (VLBI) from hundreds of extragalactic radio sources. The deviation of its pole from the pole of the dynamical reference system FK5 is only about 20 milliarcseconds. In order to keep the continuity of the reference system, the fundamental plane and the primary direction of ICRF are kept as close as possible to these of FK5, which are the J2000.0 mean equatorial plane and the J2000.0 mean March equinox. The origin of ICRS (or the zero point, its definition is given in Sect. 1.3.2) is chosen to be the average right ascension of 23 radio sources thus being close to that of FK5. The relationship of the dynamical reference systems, ICRS and FK5, depends on three parameters, which are the deviations of the celestial pole, ξ_0 and η_0 , and the zero right ascension deviation $d\alpha_0$. Their values are

$$\begin{cases} \xi_0 = -0''.016617 \pm 0''.000010, \\ \eta_0 = -0''.006819 \pm 0''.000010, \\ d\alpha_0 = -0''.0146 \pm 0''.0005. \end{cases} \quad (1.23)$$

The relationship of ICRS and the J2000.0 mean equatorial coordinate system can be given by

$$\begin{cases} \vec{r}_{J2000.0} = B\vec{r}_{ICRS}, \\ B = R_x(-\eta_0)R_y(\xi_0)R_z(d\alpha_0), \end{cases} \quad (1.24)$$

where $\vec{r}_{J2000.0}$ and $\vec{r}_{ICRS}$ are for the same vector but in different coordinate systems, the constant matrix B is the **deviation matrix of reference frame** composed of the three small rotation angles.

The **J2000.0 mean equatorial coordinate system** is the commonly accepted geocentric celestial coordinate reference system (GCRS) in present Aerospace Dynamics (especially for Earth's satellites). If unnecessary, the above-given deviation matrix of the reference frame is not mentioned again.

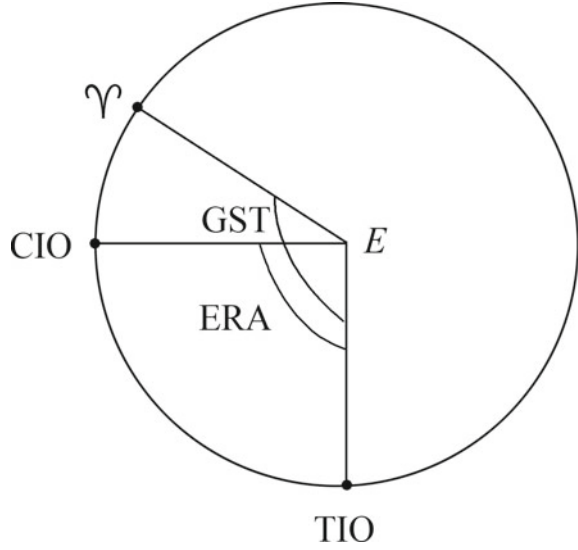
1.3.2 The Intermediate Equator and Three Related Datum Points

The **intermediate equator** is introduced to better describe the relationship between the **Celestial Reference System** (CRS) and the **Terrestrial Reference System**. The celestial axis is the extension of Earth's rotation axis, the points of intersection of the celestial axis and the celestial sphere are called celestial poles. Because of Earth's precession, the direction of Earth's rotation axis changes over time in CRS, which is instantaneous, therefore the celestial pole and the celestial equator are also instantaneous. For clarity IAU 2003 named the instantaneous celestial pole and the celestial equator as the **Celestial Intermediate Pole** (CIP) and the **Intermediate Equator**, respectively.

In order to take measurements in the celestial reference system, it is necessary to select a fixed point with respect to the celestial reference system on the intermediate equator as the origin, which is called the **Celestial Intermediate Origin** (CIO). Similarly, in the terrestrial system a point, which is fixed with respect to the system, is needed and called the **Terrestrial Intermediate Origin** (TIO). CIO is decided based on observations of a group of quasars and is close to the 0° right ascension, i.e., the **March equinox** on the International Celestial Reference Frame; whereas TIO is decided by a group of observatories on Earth, and is near the 0° longitude (the **prime meridian**, i.e., the **Greenwich meridian**) on the International Terrestrial Reference Frame. In Fig. 1.3 the intermediate equator is given by the circle, E is Earth's barycenter, and Υ is the March equinox.

In the celestial reference system, the intermediate equator tied with CIO is called the **Celestial Intermediate Equator**, TIO moves along the equator anti-clockwise, its period is a sidereal day. In the terrestrial reference system, the intermediate equator tied with TIO is called the **Terrestrial Intermediate Equator**, CIO moves along the equator over the same period of a sidereal day but clockwise. Both observations reflect

Fig. 1.3 Illustration of the intermediate equator



Earth's rotation, and the angle between CIO and TIO is called **Earth Rotation Angle (ERA)**.

1.3.3 Three Geocentric Coordinate Systems

(1) The geocentric celestial coordinate system O - xyz

This system is actually the above-mentioned epoch **J2000.0 mean equatorial reference system**, also called the geocentric celestial coordinate system. Its origin is Earth's barycenter, the xy -plane is the epoch J2000.0 mean equatorial plane, the direction of x -axis points to the epoch J2000.0 mean March equinox $\overline{\gamma}_0$, which is the intersection of the epoch J2000.0 mean equator and the epoch J2000.0 instantaneous ecliptic. This system, in a certain sense, is a "fixed system" (because it eliminates the rotation of the frame caused by the vibration of Earth's equator), thus the motion orbits of a celestial body (such as a satellite) at different times can be displayed in the same frame and the actual variation of the orbit can be compared. The geocentric celestial system is the adopted space coordinate system worldwide. It should be noticed that in this system the gravitation potential due to Earth's non-spherical part is variable.

(2) The Earth-fixed geocentric coordinate system O - XYZ

This system is the **Terrestrial Reference System (TRS)**, which is a space reference system rotating with Earth, commonly called the **Earth-fixed coordinate system**. In this system, the position of an observatory is fixed on the surface of Earth, except for some minor variations due to the tidal force or Earth's physical deformation force.

As mentioned above that the realization of ICRS requires ICRF. It is the same that the Terrestrial Reference Frame (TRF) is needed for the realization of TRS. TRF (used in navigation, survey, terrestrial physics, etc.) is defined by a group of fixed points on Earth's surface, whose positions are precisely determined in TRS. The first TRF is given by the **International Latitude Service**. Based on the observations over five years, 1900–1905, the International Latitude Service defined the **Conventional International Origin (CIO)**, which was the average direction of the third axis (z -axis), i.e., the mean direction of Earth's pole. It should be noticed that nowadays the abbreviation CIO is given to the Celestial Intermediate Origin (see Sect. 1.3.2), so is no longer for the Conventional International Origin.

In the Earth-fixed coordinate system, the origin of the frame is at the center of Earth, the xy -plane is close to the 1900.0 mean equatorial plane, and the direction of the x -axis points to the intersection of the Greenwich meridian and the equator, so can be called the Greenwich meridian direction [2]. Several Earth's gravitational models and the related reference ellipsoid are defined in this system, thus these models are self-consistent. If there is no specific explanation the Earth-fixed system in this book agrees with the World Geodetic System 84 (also known as **WGS 1984**). For this system there are

$$GE = 398600.4418(\text{km}^3/\text{s}^2),$$

$$a_e = 6378.137(\text{km}), \quad \frac{1}{f} = 298.257223563, \quad (1.25)$$

where GE is the geocentric gravitational constant, a_e and f are the equatorial radius and the flattening factor of the reference ellipsoid, respectively.

In the Earth-fixed frame, the position vector of an observatory is given by $\vec{R}_e(H, \lambda, \varphi)$. For the position vector, the relationship between the rectangular coordinates (X_e, Y_e, Z_e) and the spherical coordinates (H, λ, φ) is given by

$$\begin{cases} X_e = (N + H) \cos \phi \cos \lambda, \\ Y_e = (N + H) \cos \phi \sin \lambda, \\ Z_e = [N(1 - f)^2 + H] \sin \phi, \end{cases} \quad (1.26)$$

with

$$\begin{aligned} N &= a_e [\cos^2 \varphi + (1 - f)^2 \sin^2 \varphi]^{-\frac{1}{2}} \\ &= a_e \left[1 - 2f \left(1 - \frac{f}{2} \right) \sin^2 \varphi \right]^{-\frac{1}{2}}, \end{aligned} \quad (1.27)$$

where a_e and f are given in (1.25). The spherical coordinate H is the geodetic height of the observatory, and λ and φ are the geodetic longitude and latitude of the observatory, respectively. Their relationships with the rectangular coordinates are

given by

$$\tan \lambda = \frac{Y_e}{X_e}, \quad \sin^2 \varphi = \frac{Z_e}{[N(1-f)^2 + H]}. \quad (1.28)$$

(3) Geocentric ecliptic coordinate system $O-x'y'z'$

The origin of this system is also Earth's barycenter, and there is only a translation relationship between this system and the heliocentric ecliptic system. The $x'y'$ -plane is the epoch J2000.0 ecliptic plane, and the direction of the x' -axis is the same as in the celestial coordinate system $O-xyz$, which points to the epoch J2000.0 mean March equinox.

1.3.4 Transformation of the Earth-Fixed Coordinate System $O-XYZ$ and the Geocentric Celestial Coordinate System $O-xyz$

1.3.4.1 Transformation Relationship (I) by the IAU 1980 Model

Using $\vec{r}$ and $\vec{R}$ as the position vectors of a spacecraft in the geocentric celestial system $O-xyz$ and the Earth-fixed system $O-XYZ$, respectively, then the transformation relationship is given by

$$\vec{R} = (HG)\vec{r}. \quad (1.29)$$

The coordinate transformation matrix (HG) is given by four rotating matrices, that

$$(HG) = (EP)(ER)(NR)(PR), \quad (1.30)$$

where (PR) is the precession matrix, (NR) the nutation matrix, (ER) the Earth's rotation matrix, and (EP) the polar motion matrix, given by

$$(EP) = R_y(-x_p)R_x(-y_p), \quad (1.31)$$

$$(ER) = R_z(S_G), \quad (1.32)$$

$$\begin{aligned} (NR) &= R_x(-\Delta\varepsilon)R_y(\Delta\theta)R_z(-\Delta\mu) \\ &= R_x(-(\varepsilon + \Delta\varepsilon))R_z(-\Delta\psi)R_x(\varepsilon), \end{aligned} \quad (1.33)$$

$$(PR) = R_z(-z_A)R_y(\theta_A)R_z(-\zeta_A), \quad (1.34)$$

In (1.31) x_p and y_p are the components of the polar shift vector. The Greenwich sidereal time S_G in (1.32) is given by

$$S_G = \bar{S}_G + \Delta\mu, \quad (1.35)$$

where $\Delta\mu$ is the nutation of the right ascension; $\bar{S}_G$ is the J2000.0 Greenwich mean sidereal time given by

$$\bar{S}_G = 18^{\text{h}}.697374558 + 879000^{\text{h}}.051336907T + 0^{\text{s}}.093104T^2, \quad (1.36)$$

$$T = \frac{1}{36525.0} [\text{JD}(t) - \text{JD}(\text{J2000.0})]. \quad (1.37)$$

In these two formulas, t is the UT1 time, but for calculating other variables such as precession and nutation, t is the TDT time. Time T is measured from J2000.0 but uses a century as a unit.

The precession constants in (3.14) ζ_A , θ_A , and z_A are given by

$$\begin{cases} \zeta_A = 2306''.2181T + 0''.30188T^2, \\ \theta_A = 2004''.3109T - 0''.42665T^2, \\ z_A = 2306''.2181T + 1''.09468T^2, \end{cases} \quad (1.38)$$

where θ_A is the precession in declination, and μ (or m_A) is the precession in right ascension that

$$\begin{aligned} \mu &= \zeta_A + z_A \\ &= 4612''.4362T + 1''.39656T^2. \end{aligned} \quad (1.39)$$

In (1.33) ε is the mean obliquity. The nutation components in ecliptic longitude $\Delta\psi$ and in obliquity $\Delta\varepsilon$ can be calculated using the sequences provided by the IAU 1980 model, which has 106 terms with amplitudes greater than $0''.0001$. For the requirement of general orbital accuracy, only the terms with amplitudes greater than $0''.005$ need to be included, which are the first 20 terms. Because these terms are periodic (the shortest period term is due to the Moon's motion), so there is no accumulative effect, and the error caused by the terms with amplitude less than $0''.005$ is equivalent to the order of meter in ground-based positioning, and it is less than $0^{\text{s}}.001$ with respect to time. The first 20 terms are:

$$\begin{cases} \Delta\psi = \sum_{j=1}^{20} (A_{0j} + A_{1j}t) \sin\left(\sum_{i=1}^5 k_{ji}\alpha_i(t)\right), \\ \Delta\varepsilon = \sum_{j=1}^{20} (B_{0j} + B_{1j}t) \cos\left(\sum_{i=1}^5 k_{ji}\alpha_i(t)\right), \end{cases} \quad (1.40)$$

where the components of nutation in right ascension and in inclination, $\Delta\mu$ and $\Delta\theta$, respectively, are given by

$$\begin{cases} \Delta\mu = \Delta\psi \cos\varepsilon, \\ \Delta\theta = \Delta\psi \sin\varepsilon, \end{cases} \quad (1.41)$$

and the value of the obliquity ε is given by

$$\varepsilon = 23^\circ 26' 21''.448 - 46''.8150 t. \quad (1.42)$$

In (1.40) there are five basic arguments α_i ($i = 1, \dots, 5$) related to the positions of the Sun and the Moon, which are given by:

$$\begin{cases} \alpha_1 = 134^\circ 57' 46''.733 + (1325'' + 198^\circ 52' 02''.633)t + 31''.310t^2, \\ \alpha_2 = 357^\circ 31' 39''.804 + (99'' + 359^\circ 03' 01''.224)t - 0''.577t^2, \\ \alpha_3 = 93^\circ 16' 18''.877 + (1342'' + 82^\circ 01' 03''.137)t - 13''.257t^2, \\ \alpha_4 = 297^\circ 51' 01''.307 + (1236'' + 307^\circ 06' 41''.328)t - 6''.891t^2, \\ \alpha_5 = 125^\circ 02' 40''.280 - (5'' + 134^\circ 08' 10''.539)t + 7''.455t^2, \end{cases} \quad (1.43)$$

where $1'' = 360^\circ$. The first 20 terms of the nutation sequences are listed in Table 1.2. To reach the above-mentioned accuracy of the order of meter, the terms on the right side of (1.40) only A_{11} and B_{11} in Table 1.2 are needed, other terms of A_{1j} and B_{1j} can be omitted. Specifically, the number of terms needed depends on not only the required accuracy but also the capability of the software, such as the factor of functional expansion. The time t in (1.40)–(1.43) is the same as T , the century number given by (1.37), but is in TDT.

The formulas for calculating the rotational matrices $R_x(\theta)$, $R_y(\theta)$, and $R_z(\theta)$ are given by (1.20)–(1.22). Note that they are orthogonal matrices, that $R_x^T(\theta) = R_x^{-1}(\theta) = R_x(-\theta)$, $\dots$.

1.3.4.2 Transformation Relationship (II) by the IAU 2000 Model

By the IAU 2000 model, the transformation from the geocentric celestial reference system (GCRS) to the International Terrestrial Reference System (ITRS) is given by

$$[\text{ITRS}] = W(t)R(t)M(t)[\text{GCRS}], \quad (1.44)$$

where [GCRS] and [ITRS] correspond to the geocentric celestial coordinate system and the Earth-fixed coordinate system by the IAU 1980 model, respectively. Using $\vec{r}$ and $\vec{R}$ for the position vectors of a spacecraft in the two systems, respectively (the symbols are used in Sect. 1.3.4.1, for consistency we use the same symbols here), then the transformation relationship is given by

Table 1.2 The first 20 terms of the IAU 1980 Nutation sequence

j	Period (d)	k_{j1}	k_{j2}	k_{j3}	k_{j4}	k_{j5}	A_{0j} (0"0.0001)	A_{1j}	B_{0j} (0"0.0001)	B_{1j}
1	6798.4	0	0	0	0	1	−171,996	−174.2	92,025	8.9
2	182.6	0	0	2	−2	2	−13,187	−1.6	5736	−3.1
3	13.7	0	0	2	0	2	−2274	−0.2	977	−0.5
4	3399.2	0	0	0	0	2	2062	0.2	− 895	0.5
5	365.2	0	1	0	0	0	1426	−3.4	54	−0.1
6	27.6	1	0	0	0	0	712	0.1	−7	0.0
7	121.7	0	1	2	−2	2	−517	1.2	224	−0.6
8	13.6	0	0	2	0	1	−386	−0.4	200	0.0
9	9.1	1	0	2	0	2	−301	0.0	129	−0.1
10	365.3	0	−1	2	−2	2	217	− 0.5	−95	0.3
11	31.8	1	0	0	−2	0	− 158	0.0	−1	0.0
12	177.8	0	0	2	−2	1	129	0.1	−70	0.0
13	27.1	−1	0	2	0	2	123	0.0	−53	0.0
14	27.7	1	0	0	0	1	63	0.1	−33	0.0
15	14.8	0	0	0	2	0	63	0.0	−2	0.0
16	9.6	−1	0	2	2	2	−59	0.0	26	0.0
17	27.4	−1	0	0	0	1	−58	− 0.1	32	0.0
18	9.1	1	0	2	0	1	−51	0.0	27	0.0
19	205.9	2	0	0	−2	0	48	0.0	1	0.0
20	1305.5	−2	0	2	0	1	46	0.0	−24	0.0

$$\vec{R} = W(t)R(t)M(t)\vec{r}, \quad (1.45)$$

where $M(t)$ is the precession and nutation matrix, $R(t)$ is the Earth rotation matrix, and $W(t)$ is the polar shift matrix. Based on the transformation relationship of the March equinox the matrix $M(t)$ can be written as

$$M(t) = N(t)P(t)B, \quad (1.46)$$

where $N(t)$ is the nutation matrix, $P(t)$ is the precession matrix, and B is the deviation matrix of the reference frame defined in (1.24), which is a small constant matrix. When the **J2000.0 mean equatorial coordinate system** is directly used as the geocentric celestial coordinate system, the effect of B can be omitted then

$$M(t) = N(t)P(t). \quad (1.47)$$

The calculation methods for these matrices are given as follows.

(1) Calculations of precession and nutation

The 24th IAU general assembly (August 2000, Manchester) decided that from 1 January 2003, the IAU 2000 Precession-Nutation model formally replaces the IAU 1976 Precession Model and the IAU 1980 Nutation Model. For different accuracy requirements, the IAU 2000 model includes two versions, IAU 2000A and IAU 2000B with accuracies of 0.2 mas (milliarcsecond) and 1 mas, respectively.

In calculating the precession at a given epoch measured from J2000.0 the required three equatorial precession quantities ξ_A , z_A , and θ_A for transforming mean equatorial coordinate systems are given by

$$\begin{aligned}\xi_A &= 2''.650545 + 2306''.083227t + 0''.2988499t^2 + 0''.01801828t^3 \\ &\quad - 0''.000005971t^4 - 0''.0000003173t^5, \\ \theta_A &= 2004''.191903t - 0''.4294934t^2 - 0''.04182264t^3 \\ &\quad - 0''.000007089t^4 - 0''.0000001274t^5, \\ z_A &= 2''.650545 + 2306''.077181t + 1''.0927348t^2 + 0''.01826837t^3 \\ &\quad - 0''.000028596t^4 - 0''.0000002904t^5,\end{aligned}\tag{1.48}$$

where t is the Julian century number of the epoch time measured from J2000.0 (TT time) that

$$t = (\text{JD}(\text{TT}) - 2452545.0)/36525.\tag{1.49}$$

The nutation components in ecliptic longitude $\Delta\psi$ and in obliquity $\Delta\varepsilon$ can be calculated by the IAU 2000 model as

$$\begin{aligned}\Delta\psi &= \Delta\psi_p + \sum_{i=1}^{77} (A_i + A'_i t) \sin(\alpha_i) + (A''_i + A'''_i t) \cos(\alpha_i), \\ \Delta\varepsilon &= \Delta\varepsilon_p + \sum_{i=1}^{77} (B_i + B'_i t) \cos(\alpha_i) + (B''_i + B'''_i t) \sin(\alpha_i),\end{aligned}\tag{1.50}$$

where t is the same as given by (1.49), $\Delta\psi_p$ and $\Delta\varepsilon_p$ are long period variations of nutation, that

$$\Delta\psi_p = -0''.135 \times 10^{-3}, \quad \Delta\varepsilon_p = 0''.388 \times 10^{-3}.\tag{1.51}$$

The arguments α_i in (1.50) is a linear combination of five basic arguments, that

$$\alpha_i = \sum_{k=1}^5 n_{ik} F_k = n_{i1}l + n_{i2}l' + n_{i3}F + n_{i4}D + n_{i5}\Omega,\tag{1.52}$$

where n_{ik} are integers, and F_k are the five basic arguments related to the positions of the Sun and the Moon given by

$$F_1 \equiv l = 134^\circ.96340251 + 1717915923''.2178t + 31''.8792t^2 + 0''.051635t^3 - 0''.00024470t^4, \quad (1.53)$$

$$F_2 \equiv l' = 357^\circ.52910918 + 129596581''.0481t - 0''.5532t^2 + 0''.000136t^3 - 0''.00001149t^4, \quad (1.54)$$

$$F_3 \equiv F = 93^\circ.27209062 + 1739527262''.8478t - 12''.7512t^2 - 0''.001037t^3 + 0''.00000417t^4, \quad (1.55)$$

$$F_4 \equiv D = 297^\circ.85019547 + 1602961601''.2090t - 6''.3706t^2 + 0''.006593t^3 - 0''.00003169t^4, \quad (1.56)$$

$$F_5 \equiv \Omega = 125^\circ.04455501 - 6962890''.5431t + 7''.4722t^2 + 0''.007702t^3 - 0''.00005939t^4, \quad (1.57)$$

The five basic arguments are defined as: F_1 the Moon's mean anomaly, F_2 the Sun's mean anomaly, F_3 the angular distance of the Moon's mean ascending node, F_4 the mean angular distance between the Sun and the Moon, and F_5 the mean ecliptic longitude of the Moon's ascending node. Actually, F_k ($k = 1, \dots, 5$) are the same five basic arguments α_i ($i = 1, \dots, 5$) given by (1.43).

Table 1.3 gives the coefficients of the first 20 terms of the nutation sequence by the IAU 2000B model for comparing with the IAU 1980 model in Table 1.2.

(2) Calculations of the precession matrix $P(t)$

① The classical formula for three rotations of $P(t)$

$$P(t) = R_z(\zeta_A)R_y(-\theta_A)R_z(z_A), \quad (1.58)$$

where the three rotation angles are given by (1.48).

② The formula for four rotations of $P(t)$

$$P(t) = R_x(-\varepsilon_0)R_z(\psi_A)R_x(\omega_A)R_z(-\chi_A), \quad (1.59)$$

where the last three rotation angles are given as

$$\begin{aligned} \psi_A &= 5038''.481507t - 1''.0790069t^2 - 0''.00114045t^3 \\ &\quad + 0''.000132851t^4 - 0''.0000000951t^5 \\ \omega_A &= \varepsilon_0 - 0''.025754t + 0''.0512623t^2 - 0''.007725036t^3 \\ &\quad - 0''.000000467t^4 + 0''.0000003337t^5 \\ \chi_A &= 10''.556403t - 2''.3814292t^2 - 0.00121197t^3 \\ &\quad + 0''.000170663t^4 - 0''.0000000560t^5 \end{aligned} \quad (1.60)$$

and ε_0 is given by (1.63), and t in (1.60) is given by (1.49).

Table 1.3 The first 20 terms of the IAU 2000B Nutation sequence

i	n_1	n_2	n_3	n_4	n_5	Period (d)	A_i (mas)	A'_i (mas)	B_i (mas)	B'_i (mas)
1	0	0	0	0	1	-6798.383	- 17,206.4161	-17.4666	9205.2331	0.9086
2	0	0	2	-2	2	182.621	-1317.0906	-0.1675	573.0336	-0.3015
3	0	0	2	0	2	13.661	-227.6413	-0.0234	97.8459	-0.0485
4	0	0	0	0	2	-3399.192	207.4554	0.0207	-89.7492	0.0470
5	0	1	0	0	0	365.260	147.5877	-0.3633	7.3871	-0.0184
6	0	1	2	-2	2	121.749	-51.6821	0.1226	22.4386	-0.0677
7	1	0	0	0	0	27.555	71.1159	0.0073	-0.6750	0.0000
8	0	0	2	0	1	13.633	-38.7298	-0.0367	20.0728	0.0018
9	1	0	2	0	2	9.133	-30.1461	-0.0036	12.9025	-0.0063
10	0	-1	2	-2	2	365.225	21.5829	-0.0494	-9.5929	0.0299
11	0	0	2	-2	1	177.844	12.8227	0.0137	-6.8982	-0.0009
12	-1	0	2	0	2	27.093	12.3457	0.0011	-5.3311	0.0032
13	-1	0	0	2	0	31.812	15.6994	0.0010	-0.1235	0.0000
14	1	0	0	0	1	27.667	6.3110	0.0063	-3.3228	0.0000
15	-1	0	0	0	1	-27.443	-5.7976	-0.0063	3.1429	0.0000
16	-1	0	2	2	2	9.557	-5.9641	-0.0011	2.5543	-0.0011
17	1	0	2	0	1	9.121	-5.1613	-0.0042	2.6366	0.0000
18	-2	0	2	0	1	1305.479	4.5893	0.0050	-2.4236	-0.0010
19	0	0	0	2	0	14.765	6.3384	0.0011	-0.1220	0.0000
20	0	0	2	2	2	7.096	-3.8571	-0.0001	1.6452	-0.0011

(3) Calculations of the nutation matrix $N(t)$

$$\begin{aligned}
N(t) &= R_x(-\varepsilon_A) R_z(\Delta\psi) R_x(\varepsilon_A + \Delta\varepsilon) \\
&= R_z(\Delta\mu) R_y(-\Delta\theta) R_x(\Delta\varepsilon),
\end{aligned} \tag{1.61}$$

where $\Delta\mu = \Delta\Psi \cos\varepsilon_A$, and $\Delta\theta = \Delta\Psi \sin\varepsilon_A$; $\Delta\mu$ and $\Delta\theta$ are the components of the nutation in right ascension and declination, respectively; $\Delta\Psi$ is the nutation in ecliptic longitude; $\Delta\varepsilon$ is the nutation in obliquity; ε_A is the obliquity of the instantaneous mean equator and the ecliptic, which is called the mean obliquity and is given by

$$\begin{aligned}
\varepsilon_A &= \varepsilon_0 - 46''.836769t - 0''.05127t^2 + 0''.00200340t^3 \\
&\quad - 0''.000000576t^4 - 0''.0000000434t^5,
\end{aligned} \tag{1.62}$$

$$\varepsilon_0 = 84381''.406 = 23^\circ 26' 21''.406, \quad (1.63)$$

where the angle ε_0 is the mean obliquity at the epoch J2000.0.

(4) Calculations of Earth's rotation matrix $R(t)$

The calculation of $R(t)$ is different from the **IAU 1980** model due to the calculation of time t with respect to the Greenwich sidereal time S_G . The apparent Greenwich sidereal time (GST) in the **IAU 2000** model is strictly different from Earth's rotation angle (ERA) as shown in Fig. 1.3. The calculation formulas of $R(t)$ become

$$\begin{cases} R(t) = R_z(\text{GST}), \\ \text{GST} = \text{GMST} + \text{EE}, \end{cases} \quad (1.64)$$

where GMST and EE are the Greenwich mean sidereal time and the equation of equinoxes, respectively, and are given by

$$\begin{aligned} \text{GMST} = & \theta(\text{UT1}) + 0''.014506 + 4612''.156534t + 1''.3915817t^2 \\ & - 0''.00000044t^3 - 0''.000029956t^4 - 0.0000000368t^5, \end{aligned} \quad (1.65)$$

$$\text{EE} = \Delta\psi \cos \varepsilon_A - \sum_k C_k \sin \alpha_k - 0''.00000087t \sin \Omega. \quad (1.66)$$

In (1.65), $\theta(\text{UT1})$ is Earth's rotational angle (**ERA**), which is a linear function of UT1, given by

$$\theta(\text{UT1}) = 2\pi(0.7790572732640 + 1.00273781191135448d), \quad (1.67)$$

where d is the Julian day number at UT1 starting from epoch J2000.0 given by

$$d = \text{JD}(\text{UT1}) - 2451545.0. \quad (1.68)$$

The angles α_k and amplitudes C_k on the right side of (1.66) are given in Table 1.4.

In Table 1.4, α_k are given by the five basic arguments F_k , similar to α_i in (1.50), with $F_1 \equiv l$, $F_2 \equiv l'$, $F_3 \equiv F$, $F_4 \equiv D$, and $F_5 \equiv \Omega$. The values of F_k are given by (1.53)–(1.57).

As mentioned above, the Greenwich apparent sidereal time GST is measured from the apparent March equinox, so is different from Earth's rotational angle ERA (i.e., UT1), their relationship is given by

$$\text{GST} = \theta(\text{UT1}) - \text{EO}, \quad (1.69)$$

Table 1.4 Values of α_k and C_k

k	α_k	C_k	k	α_k	C_k
1	Ω	$-0''.00264073$	7	$2F + \Omega$	$-0''.00000198$
2	2Ω	$-0''.00006352$	8	3Ω	$+0''.00000172$
3	$2F - 2D + 3\Omega$	$-0''.00001175$	9	$l' + \Omega$	$+0''.00000141$
4	$2F - 2D + \Omega$	$-0''.00001121$	10	$l' - \Omega$	$+0''.00000126$
5	$2F - 2D + 2\Omega$	$+0''.00000455$	11	$l + \Omega$	$+0''.00000063$
6	$2F + 3\Omega$	$-0''.00000202$	12	$l - \Omega$	$+0''.00000063$

where EO is called **the equation of origins**, which is related to the components of precession and nutation in the intermediate equator. According to the IAU 2006/2000A precession-nutation model EO (for 1975–2025 with all terms greater than $0.5 \mu\text{as}$), is given by

$$\begin{aligned} \text{EO} = & -0''.014506 - 4612''.156534t - 1''.3915817t^2 \\ & + 0''.00000044t^3 - \Delta\psi \cos \varepsilon_A + \sum_k C_k \sin \alpha_k, \end{aligned} \quad (1.70)$$

where t is given by (1.49), α_k and C_k are given in Table 1.4.

(5) Calculations of the polar motion matrix $W(t)$

$$W(t) = R_x(-y_p)R_y(-x_p)R_z(s'), \quad (1.71)$$

where x_p and y_p are two components of the Celestial Intermediate Pole (CIP) in the ITRS, and s' is called the locator of Earth's Intermediate Origin (TIO) which gives TIO's position on the CIP equator and is a function of x_p and y_p that

$$s'(t) = \left(\frac{1}{2}\right) \int_{t_0}^t (x_p \dot{y}_p - \dot{x}_p y_p) dt. \quad (1.72)$$

The value of $s'(t)$ cannot be known in advance, but it is very small and can be approximately given by

$$s'(t) = -(4''.7 \times 10^{-5})t, \quad (1.73)$$

where t is the same as given in (1.49).

1.3.5 Relationship Between the IAU 1980 Model and the IAU 2000 Model

The differences between the IAU 1980 model and the new IAU 2000 model exist not only in the calculations of precession and nutation but also in the definitions of the geocentric system and celestial system, as well as in the usage of symbols. This section provides the relationships of related calculation formulas of the two models.

In the **IAU 2000** model, the transformation from GCRS to ITRS is given in (1.44), which is

$$[\text{ITRS}] = W(t)R(t)M(t)[\text{GCRS}], \quad (1.74)$$

where $M(t)$ is the precession-nutation matrix, $R(t)$ is the Earth rotation matrix, and $W(t)$ is the polar motion matrix. In the **IAU 1980** model, the position vector of a satellite is given by $\vec{R}$ and $\vec{r}$ in ITRS and GCRS, respectively, and the transformation formula (1.74) can be similarly given by

$$\begin{aligned} \vec{R} &= (HG)\vec{r}, \\ (HG) &= W(t)R(t)M(t), \end{aligned} \quad (1.75)$$

with relationships

$$M(t) = N(t)P(t) = (NR)(PR), \quad (1.76)$$

$$P(t) = (PR) = R_z(-z_A)R_y(\theta_A)R_z(-\zeta_A), \quad (1.77)$$

$$\begin{aligned} N(t) &= (NR) = R_x(-\Delta\varepsilon)R_y(\Delta\theta)R_z(-\Delta\mu) \\ &= R_x(-(\varepsilon + \Delta\varepsilon))R_z(-\Delta\psi)R_x(\varepsilon), \end{aligned} \quad (1.78)$$

$$R(t) = (ER) = R_z(S_G), \quad (1.79)$$

$$W(t) = (EP) = R_y(-x_p)R_x(-y_p). \quad (1.80)$$

It is not difficult to see that the transformation processes and calculation formulas for the two matrices, the precession-nutation matrix and the Earth's rotation matrix, are the same for both models, except for a slight difference in S_G and small differences in parameters. But there is a difference in the expression of the polar motion matrix $W(t)$, that according to (1.71) in the **IAU 2000** model there is

$$W(t) = R_x(-y_p)R_y(-x_p)R_z(s').$$

Comparing (1.71) and (1.80), we can see that the small rotation matrix $R_z(s')$ is not included in (1.80), also the orders of the other two rotation matrices $R_x(-y_p)$ and $R_y(-x_p)$ are reversed. Because the polar motion components, x_p and y_p , are very small, the effects of these differences are even smaller. As a result, the effect on $W(t)$ is minor, and there is no obvious difference in the transformation. The author confirms that by calculations using actual examples.

In the entire transformation between $\vec{R}$ and $\vec{r}$, there is no obvious difference in results between the IAU 2000 model and the IAU 1980 model. For example, using the first 20 terms in the nutation sequence to calculate Greenwich sidereal time, results by both the IAU 1980 model and the IAU 2000 model are in an error range of $0^s.001$ compared to what is given in Ephemeris, and the error in the position transformation is in the order of meter.

1.3.6 The Complicity in the Selection of Coordinate System Due to the Wobble of Earth's Equator

It is inevitable that the motion of artificial Earth's satellite involves both the Earth-fixed coordinate system and the geocentric celestial coordinate system. The wobble of Earth's equator causes difficulties and even confusion in the selection of the coordinate system. This problem is discussed in Sect. 4.4.5 in detail.

1.3.7 Coordinate Systems Related to Satellite Measurements, Attitudes, and Orbital Errors

To provide satellite orbital measurements, attitudes, and related errors, it is necessary to define a related coordinate system for specific requirements. Other calculations such as for locations of sub-satellite points (the node of the line connecting a satellite and Earth's center at Earth's surface) are also related to actual coordinate systems and satellite's orbits. This problem is discussed in Chap. 2.

1.4 The Moon's Coordinate Systems

For describing the motion of a Moon's prober there are three selenocentric coordinate systems, which are the Moon-fixed coordinate system, the selenocentric equatorial coordinate system, and the selenocentric ecliptic coordinate system.

1.4.1 Definitions of the Three Selenocentric Coordinate Systems [6]

Similar to Earth's equator, the lunar equator also wobbles, resulting in a physical libration. This motion causes different definitions of selenocentric equatorial coordinate systems, and affects the determinations of the lunar satellite orbit and the position of the sub-satellite point (the node of the line connecting a satellite and the Moon's center at the lunar surface).

(1) The Moon-fixed coordinate system $O\text{-}XYZ$

The origin of this coordinate system is at the center of the Moon. The direction of the Z -axis is the direction of the Moon's rotational axis. The XY -plane is the lunar equatorial plane, which is perpendicular to the Z -axis. The primary X -axis points to the node of the meridian of the lunar "Greenwich", i.e., Sinus Medii on the Moon, which is also the direction towards Earth's principal axis of inertia in the lunar equatorial plane. Obviously, in this system, the gravitational potential of the Moon is determined. There are different Moon's gravitational models and reference ellipsoids defined in this system, which are all self-consistent systems.

(2) The selenocentric equatorial coordinate system $O\text{-}xyz$

This system has two definitions.

One is the **Epoch J2000.0 selenocentric celestial coordinate system**. Its origin is at the center of the Moon, but the xy -plane is Earth's mean equatorial plane at Epoch J2000.0, the x -axis points to the J2000.0 mean March equinox Υ . By this definition, it is convenient to connect the Earth's coordinate system with the coordinate system of the target body, in this case, the Moon's coordinate system. Details are given in Sect. 1.5.

The other definition is similar to the epoch J2000.0 geocentric celestial coordinate system. The xy -plane is the lunar mean equatorial plane at epoch J2000.0, the x -axis points to the J2000.0 mean March equinox Υ , which is decided by the mean ecliptic longitude of the ascending node of the Moon's orbit around Earth, Ω_m (Fig. 1.4). The method used in dealing with a lunar prober is similar to Earth's satellite, and this coordinate system is necessary for studying the orbital motion of a lunar prober. In order to distinguish this system from the epoch J2000.0 selenocentric celestial coordinate system, this system is called the epoch J2000.0 selenocentric mean equatorial coordinate system or J2000.0 selenocentric equatorial coordinate system. Similar to the geocentric celestial coordinate system, this system can be regarded as a "fixed" selenocentric system, which is a system after eliminating the effect of the wobble of the Moon's equator on the coordinate axes. By this definition, the orbits of a lunar prober at different times can be compared in the same coordinate system. Also, in this system, the gravitational potential due to the non-spherical parts of the Moon varies.

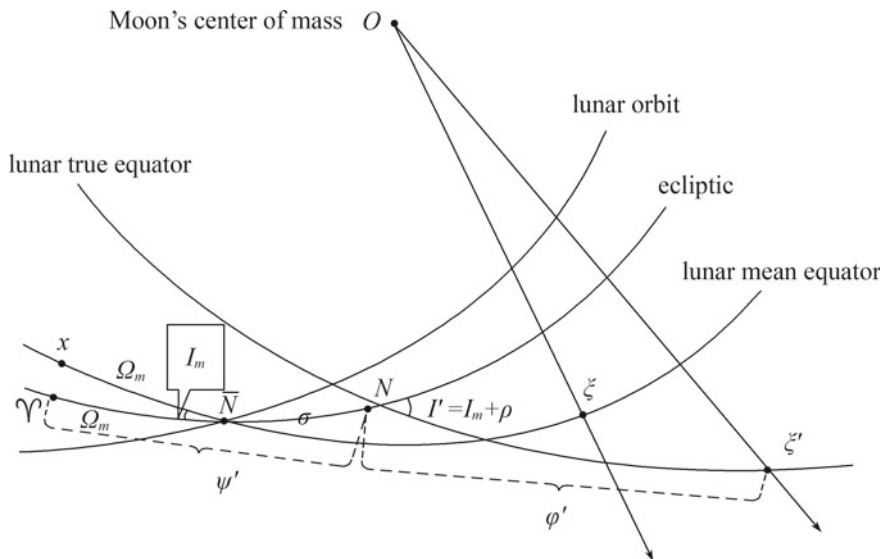


Fig. 1.4 The relationship between the lunar apparent equator and the lunar mean equator

(3) The selenocentric ecliptic coordinate system $O-x'y'z'$

The origin of this system is also at the Moon's center, and there is only a translation relationship between this system and the geocentric ecliptic system. The $x'y'$ -plane is the epoch J2000.0 ecliptic plane, and the x' -axis direction is the same as in the selenocentric celestial coordinate system $O-xyz$, which points to the epoch J2000.0 mean March equinox.

1.4.2 The Moon's Physical Libration

(1) Two expressions of the physical libration

The Moon's physical libration is a complicated fixed-point rotation problem similar to the precession and nutation of Earth's rotation. Over years there have been a few theories about this problem. Almost all theories express the physical libration by its three components in longitude τ , inclination ρ , and node σ . The lunar mean equator and the lunar true equator are then connected by the three components analytically.

According to Cassini's laws, the lunar orbit, the ecliptic, and the lunar mean equator intersect at the same point $\bar{N}$ (Fig. 1.4). Using the three components of the physical libration, the lunar true equator can be related to its mean equator, the relationship is given by

$$\begin{cases} \psi = \Omega_m, \\ I = I_m, \\ \phi = L_m - \Omega_m + \pi, \end{cases} \quad (1.81)$$

$$\begin{cases} \psi' = \Omega_m + \sigma, \\ I' = I_m + \rho, \\ \phi' = L_m - \Omega_m + \pi + (\tau - \sigma), \end{cases} \quad (1.82)$$

where I_m , Ω_m , and L_m are the lunar mean obliquity, the mean ecliptic longitude of the ascending node, and the lunar mean ecliptic longitude. The meanings of the symbols in (1.81) and (1.82) are shown in Fig. 1.4.

The American Jet Propulsion Laboratory of NASA (JPL) produces an ephemeris (DE405) which uses a different way to describe the lunar physical libration. DE405 provides daily values of three Euler angles, Ω' , i_s , and Λ (the meanings of the three angles are marked in Fig. 1.5) for the lunar physical libration. These values are used in calculations of the precise position of a lunar probe in the Moon-fixed coordinate system.

The relationship of the two expressions of the lunar physical libration, i.e., by (τ, ρ, σ) and by (Ω', i_s, Λ) , can be illustrated by Fig. 1.5, in which x_b is the direction of the X-axis in the Moon-fixed coordination system, i.e., the direction of ξ' in Fig. 1.4, and ε is Earth's mean obliquity.

As mentioned above based on the lunar rotation theory the physical libration is given by three parameters (τ, ρ, σ) , which can be calculated using analytical

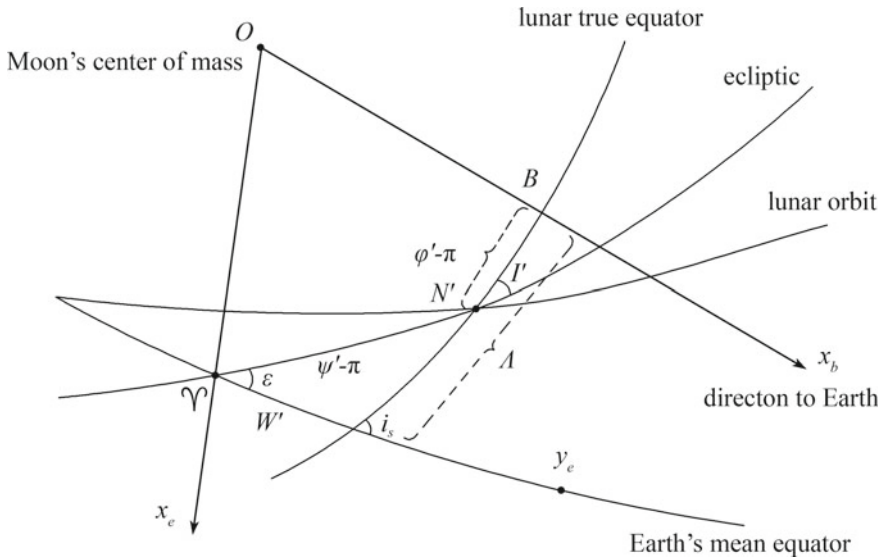


Fig. 1.5 Illustration of the Selenocentric coordinate system and the physical libration

sequences. These sequences are similar to Earth's nutation sequence, having a few hundred terms. The largest periodic term has an amplitude greater than 100 arc-seconds ($100''$), but there is no long periodic term, which exists in the vibration of Earth's equator (i.e., the term in the precession sequence over a period of about 26,000 years). Following the improvement of the lunar rotation theory, the differences in the results for the parameters by analytical calculation and by DE405 are as small as $1''$. But for high precision requests, it is inconvenient to use the analytical sequences, because a lot more terms should be included. To use the ephemeris DE405 seems simple and easy, but it may not be suitable for analyzing some particular problems. In this section, the two methods are discussed and compared. Results show that for dealing with orbital determinations or orbital predictions over a short arc (1–2 days or a bit longer) the physical libration can be given by the first four terms in Eckhardt's analytical formulas (1.84).

(2) Comparison of the Moon's physical libration given by the analytical method and DE405

First, we give two different expressions for the three parameters (τ , ρ , σ) in (1.82) for the analytical method. One is the Hayn expression, and the first three terms are given by [11]

$$\begin{cases} \tau = 59''.0 \sin l' - 12''.0 \sin l + 18''.0 \sin 2\omega_m, \\ \rho = -107''.0 \cos l + 37''.0 \cos(l + 2\omega_m) - 11''.0 \cos(2l + 2\omega_m), \\ l\sigma = -109''.0 \sin l + 37''.0 \sin(l + 2\omega_m) - 11''.0 \sin(2l + 2\omega_m). \end{cases} \quad (1.83)$$

The other is the Eckhardt expression, the first four terms are given by [12]

$$\begin{cases} \tau = 214''.170 + 90''.7 \sin l' + 17''.0 \sin(2l - 2F) - 16''.8 \sin l + 9''.9 \sin(2l - 2D), \\ \rho = -99''.1 \cos l + 24''.6 \cos(l - 2F) - 10''.5 \cos 2F - 80''.8 \sin F, \\ l\sigma = -101''.4 \sin l - 24''.6 \sin(l - 2F) - 10''.1 \sin 2F + 80''.8 \cos F. \end{cases} \quad (1.84)$$

In (1.84) the parameter τ has a constant term $214''.170$. In both formulas, l is the Moon's mean obliquity, also seen in (1.81), that $l = 1^\circ.542461 = 5552''.86$. The variables, l , l' , F , and D , are the Moon's mean anomaly, the Sun's mean anomaly, the angular distance of the Moon's mean ascending node (i.e., $F = l + \omega_m$, and ω_m is the argument of the perigee of the Moon's orbit), and the mean angular distance of the Sun and the Moon, given by

$$\begin{cases} l = 134^\circ.9633964 + 477198^\circ.8675055T, \\ l' = 357^\circ.5291092 + 35999^\circ.0502909T, \\ F = 93^\circ.27209062 + 483202^\circ.0174577T, \\ D = 297^\circ.85019547 + 445267^\circ.1114469T, \end{cases} \quad (1.85)$$

$$T = \frac{\text{JD}(t) - \text{JD}(\text{J2000.0})}{36525.0} \quad (1.86)$$

The variables F and D have also appeared in (1.55) and (1.56) for the calculations of Earth's nutation.

If a position vector in the Moon-fixed coordinate system $O\text{-}XYZ$ (i.e., the $O\text{-}\xi'\eta'\zeta$ coordinate system in Fig. 1.3) is given as $\vec{R}$, and in the selenocentric celestial coordinate system $O\text{-}x_e y_e z_e$ as $\vec{r}_e$. In the selenocentric celestial coordinate system, the $x_e y_e$ -plane is the J2000.0 Earth's mean equatorial plane. For comparing the analytical method and DE405 the relationship of $\vec{R}$ and $\vec{r}_e$ can be given by

$$\vec{R} = (M_1)\vec{r}_e = (M_2)\vec{r}_e, \quad (1.87)$$

where the transformation matrices are given by

$$(M_1) = R_z(\Lambda)R_x(i_s)R_z(\Omega'), \quad (1.88)$$

$$\begin{aligned} (M_2) &= R_z(\phi' - \pi)R_x(I')R_z(\psi' - \pi)R_x(\varepsilon) \\ &= R_z(\phi')R_x(-I')R_z(\psi')R_x(\varepsilon). \end{aligned} \quad (1.89)$$

The first line in (1.89) is given according to Fig. 1.5, whereas the second line is according to Fig. 1.4, they are actually the same. The value of M_1 is given by DE405, and the value of M_2 can be given by (1.83) or (1.84), the difference between M_1 and M_2 depends on how many terms are used to calculate (τ, ρ, σ) , such as in (1.83) there are three terms and in (1.84) there are four terms.

Applying the two methods to a fixed point on the surface of the Moon for three selected dates, 1 November 2003, 15 June 2004, and 1 January 2008, at 00:00:00, the position vector then was converted from the selenocentric celestial coordinate system to the Moon-fixed coordinate system. Results show that the difference is in the order of 1,000 m, and the differences of the corresponding elements in the matrices are in 10^{-3} . Using the analytical method or the ephemeris depends on the actual project.

As shown by the comparison, the analytical method with a limited number of terms cannot reach the accuracy requirement for some projects. But the effect of the physical libration on a Moon's prober is related to the non-spherical part of the Moon, which is small (the largest term J_2 is in the order of 10^{-4}), therefore, for certain accuracy requirements, the analytic method is acceptable. In the orbital determination or orbital prediction if the arc of an orbital extrapolation is 10^2 (the related time interval is about 1–2 days for a low orbit prober), then the analytic method using the Eckhardt four term expressions in (1.84), the errors can be less than 10 m.

In the development of orbital theory for a Moon's prober, or for understanding the orbital variation, or for the geometric position of the Moon's prober in the selenocentric coordinate system, we must use the selenocentric equatorial coordinate system not the selenocentric Earth's equatorial coordinate system (i.e., the above defined J2000.0 selenocentric celestial coordinate system). In this system, it is acceptable

to use the analytical expression of the parameters (τ, ρ, σ) to build the relationship between the selenocentric mean equatorial coordinate system $O\text{-}xyz$ and the Moon-fixed coordinate system $O\text{-}\xi'\eta'\zeta'$ (related to the apparent equator). The position vector of a prober, $\vec{R}$, in the Moon-fixed coordinate system can be given by the transformation relationship of the J2000.0 selenocentric mean equatorial coordinate system $O\text{-}xyz$ and the selenocentric celestial coordinate system $O\text{-}x_e y_e z_e$ (i.e., by the high precision values of Ω', i_s, A) that

$$\begin{cases} \vec{R} = (M_1)\vec{r}_e, & \vec{r}_e = (N)^T\vec{r}, \\ (N) = R_z(-\Omega_m)R_x(-I_m)R_z(\Omega_m)R_x(\varepsilon), \end{cases} \quad (1.90)$$

where $\vec{r}$ is the position vector in the selenocentric mean equatorial coordinate system given by the orbital determination or forecast. The transformation matrix (N) does not involve physical libration expression, the accuracy of the transformation only depends on the accuracy of the orbital determination or forecast of a Moon's prober.

1.4.3 Transformations Between the Three Selenocentric Coordinate Systems

- (1) The transformation between the Moon-fixed coordinate system $O\text{-}XYZ$ and the selenocentric equatorial coordinate system $O\text{-}xyz$

In dealing with an Earth's satellite in order to avoid the effect of precession and nutation it is often to use an orbital coordinate system of mixed forms [13–15]. But for a Moon's satellite its motion can be resolved entirely in the J2000.0 selenocentric mean equatorial coordinate system. As mentioned above, the xy -plane is the lunar mean equatorial plane at epoch J2000.0, the direction of the x -axis points to the J2000.0 mean March equinox which is decided by the mean ecliptic longitude of the ascending node of the lunar orbit around Earth, Ω_m . This coordinate system is used in the orbital determination and forecast by both the analytical and the numerical methods, and the effect of the physical libration on the system is treated as an additional perturbation of the coordinate system. The perturbation analytical expression has been given by the author, which is not complicated [16]. To provide the additional perturbation expression the first thing is to establish the transformation relationship between the Moon-fixed coordinate system and the selenocentric equatorial coordinate system.

Let $\vec{R}(X, Y, Z)$ and $\vec{r}(x, y, z)$ be the position vectors in the Moon-fixed coordinate system $O\text{-}\xi'\eta'\zeta'$ (corresponding to the apparent equator) and the epoch J2000.0 selenocentric mean equatorial coordinate system $O\text{-}xyz$, respectively, then the transformation relationships of $\vec{R}(X, Y, Z)$ and $\vec{r}(x, y, z)$ are given by

$$\begin{cases} \vec{r} = R_z(-\Omega_m)R_x(-I)R_z(-\sigma)R_x(I+\rho)R_z(-(\phi+\tau-\sigma))\vec{R} = (A)\vec{R}, \\ \vec{R} = R_z(\phi+\tau-\sigma)R_x(-(I+\rho))R_z(\sigma)R_x(I)R_z(\Omega_m)\vec{r} = (A)^T\vec{r}, \end{cases} \quad (1.91)$$

where $R_z(-\Omega_m)$, $R_x(-I)$, $R_z(-\sigma)$, $R_x(I+\rho)$, and $R_z(-(\phi+\tau-\sigma))$ are orthogonal matrices. For providing the additional perturbation of the coordinate system in building the Moon's satellite orbital solution it is necessary to have the analytical expressions of the transformation. Omitting the complicated process and keeping only the first order of τ , σ , and ρ , the results are

$$\vec{r} = (A)\vec{R}, \quad (A) = (a_{ij}), \quad (1.92)$$

$$\begin{cases} a_{11} = \cos(\varphi + \Omega_m) - (\tau - \sigma + \sigma \cos I) \sin(\varphi + \Omega_m) \\ \quad = \cos(\varphi + \Omega_m) - \tau \sin(\varphi + \Omega_m) + O(10^{-5}) \\ a_{12} = -\sin(\varphi + \Omega_m) - (\tau - \sigma + \sigma \cos I) \cos(\varphi + \Omega_m) \\ \quad = -\sin(\varphi + \Omega_m) - \tau \cos(\varphi + \Omega_m) + O(10^{-5}) \\ a_{13} = -\sigma \sin I \cos \Omega_m - \rho \sin \Omega_m = -I\sigma \cos \Omega_m - \rho \sin \Omega_m \\ a_{21} = \sin(\varphi + \Omega_m) + (\tau - \sigma + \sigma \cos I) \cos(\varphi + \Omega_m) \\ \quad = \sin(\varphi + \Omega_m) + \tau \cos(\varphi + \Omega_m) + O(10^{-5}) \\ a_{22} = \cos(\varphi + \Omega_m) - (\tau - \sigma + \sigma \cos I) \sin(\varphi + \Omega_m) \\ \quad = \cos(\varphi + \Omega_m) - \tau \sin(\varphi + \Omega_m) + O(10^{-5}) \\ a_{23} = -\sigma \sin I \sin \Omega_m + \rho \cos \Omega_m = -I\sigma \sin \Omega_m + \rho \cos \Omega_m \\ a_{31} = \sigma \sin I \cos \varphi - \rho \sin \varphi = I\sigma \cos \varphi - \rho \sin \varphi \\ a_{32} = -\sigma \sin I \sin \varphi - \rho \cos \varphi = -I\sigma \sin \varphi - \rho \cos \varphi \\ a_{33} = 1 \end{cases} \quad (1.93)$$

The formula (1.84) can be written as

$$\begin{cases} \tau = \tau_0 + \tau_1 \sin(l') + \tau_2 \sin(l) + \tau_3 \sin(2\omega_m) + \tau_4 \sin(2l - 2D) \\ \rho = \rho_1 \cos(l) + \rho_2 \cos(l + 2\omega_m) + \rho_3 \cos(2l + 2\omega_m) + \rho_4 \sin(l + \omega_m) \\ I\sigma = \sigma_1 \sin(l) + \sigma_2 \sin(l + 2\omega_m) + \sigma_3 \sin(2l + 2\omega_m) + \sigma_4 \cos(l + \omega_m) \end{cases} \quad (1.94)$$

With reasonable approximations, there are

$$\begin{aligned} \rho_1 = \sigma_1 = -99''.1, \quad \rho_2 = \sigma_2 = 24''.6, \quad \rho_3 = \sigma_3 = -10''.1, \\ \rho_4 = -\sigma_4 = -80''.8 \end{aligned} \quad (1.95)$$

Using the approximations

$$\cos \tau = 1 + O(10^{-6}), \quad \sin \tau = \tau [1 + O(10^{-6})],$$

the elements in the matrix (A) in (1.92) can be reduced to

$$\begin{cases} a_{11} = -\cos(L_m + \tau) \\ a_{12} = \sin(L_m + \tau) \\ a_{13} = -\sigma_1 \sin(L_m - \omega_m) - \sigma_2 \sin(L_m + \omega_m) - \sigma_3 \sin(2L_m - \Omega_m) - \sigma_4 \cos(L_m) \\ a_{21} = -\sin(L_m + \tau) \\ a_{22} = -\cos(L_m + \tau) \\ a_{23} = \sigma_1 \cos(L_m - \omega_m) + \sigma_2 \cos(L_m + \omega_m) + \sigma_3 \cos(2L_m - \Omega_m) - \sigma_4 \sin(L_m) \\ a_{31} = \sigma_1 \sin(\omega_m) - \sigma_2 \sin(\omega_m) - \sigma_3 \sin(l + \omega_m) - \sigma_4 \\ a_{32} = \sigma_1 \cos(\omega_m) + \sigma_2 \cos(\omega_m) + \sigma_3 \cos(l + \omega_m) + \sigma_4 \sin(2l + 2\omega_m) \\ a_{33} = 1 \end{cases} \quad (1.96)$$

- (2) The transformation between the selenocentric celestial coordinate system and the geocentric celestial coordinate system

For a Moon's satellite, the orbital measurements and orbital operation are performed by observatories stationed on Earth, therefore, the epoch geocentric celestial coordinate system, the epoch selenocentric celestial coordinate system, and the epoch selenocentric equatorial coordinate system are all involved at the same time. The transformations between these systems are related to the Moon's position in Earth's coordinate system. The Moon's coordinates can be given by a high precision ephemeris (such as JPL DE405), by a less precise analytical expression, or by a high precision semi-analytical ephemeris.

In the J2000.0 geocentric ecliptic coordinate system, the Moon's mean orbital elements ($\bar{\sigma}'$) are

$$\begin{cases} \bar{a} = 384747.981 \text{ km}, \\ \bar{e} = 0.054880, \\ \bar{i} = J = 5^\circ.1298, \\ \bar{\Omega} = 125^\circ.0446 - 1934^\circ.14t, \\ \bar{\omega} = 318^\circ.3087 + 6003^\circ.15t, \\ \bar{M} = 134^\circ.9634 + 13^\circ.0650d, \end{cases} \quad (1.97)$$

where t is the century number defined by (1.49), and d is the Julian day number starting from the epoch J2000.0.

The perturbed variation of the Moon's orbit is relatively large, the largest amplitude of the periodic terms is about 2×10^{-2} . Analytical formulas of the main periodic terms, which have an accuracy of 10^{-3} for $|\Delta \bar{r}'|/r'$, are given as follows.

The coordinates of the lunar center in the geocentric ecliptic coordinate system (λ, β, π) are given by

$$\lambda = 218^\circ.32 + 481267^\circ.883t + \sum_{j=1}^6 K_j \sin(\alpha_j), \quad (1.98)$$

$$\beta = \sum_{j=7}^{10} K_j \sin(\alpha_j), \quad (1.99)$$

$$\pi = 0^\circ.9508 + \sum_{j=11}^{14} K_j \cos(\alpha_{j-10}), \quad (1.100)$$

$$\begin{cases} K_1 = 6^\circ.29, & K_2 = -1^\circ.27, & K_3 = 0^\circ.66, & K_4 = 0^\circ.21, \\ K_5 = -0^\circ.19, & K_6 = -0^\circ.11, & K_7 = 5^\circ.13, & K_8 = 0^\circ.28, \\ K_9 = -0^\circ.28, & K_{10} = 0^\circ.17, & K_{11} = 0^\circ.0518, & K_{12} = 0^\circ.0095, \\ K_{13} = 0^\circ.0078, & K_{14} = 0^\circ.0028, \end{cases} \quad (1.101)$$

$$\begin{cases} \alpha_1 = 134^\circ.9 + 477198^\circ.85t, & \alpha_2 = 259^\circ.2 - 413335^\circ.38t, \\ \alpha_3 = 235^\circ.7 + 890534^\circ.23t, & \alpha_4 = 269^\circ.9 + 954397^\circ.70t, \\ \alpha_5 = 357^\circ.5 + 35999^\circ.05t, & \alpha_6 = 186^\circ.6 + 966404^\circ.05t, \\ \alpha_7 = 93^\circ.3 + 483202^\circ.13t, & \alpha_8 = 228^\circ.2 + 960400^\circ.87t, \\ \alpha_9 = 318^\circ.3 + 6003^\circ.18t, & \alpha_{10} = 217^\circ.6 - 407332^\circ.20t. \end{cases} \quad (1.102)$$

Then the position of the Moon's center in the geocentric equatorial coordinate system, $\vec{R} = (X, Y, Z)$, is given by

$$\vec{R} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = R \hat{R}, \quad (1.103)$$

$$R = 1/\sin \pi, \quad (1.104)$$

$$\hat{R} = \begin{pmatrix} \cos \beta \cos \lambda \\ 0.9175 \cos \beta \sin \lambda - 0.3978 \sin \beta \\ 0.3978 \cos \beta \sin \lambda + 0.9175 \sin \beta \end{pmatrix}. \quad (1.105)$$

This position vector is in the Earth's instantaneous mean equatorial coordinate system, and to give the position vector in the epoch J2000.0 geocentric mean equatorial coordinate system, $\vec{r}'$, requires a precession correction given by

$$\vec{r}' = (PR)^T \vec{R}. \quad (1.106)$$

The time t in the above formulas is the same as in (1.49).

1.5 Planets' Coordinate Systems

In 2009, the IAU Working Group on Cartographic Coordinates and Rotational Elements submitted recommendations (the IAU recommendations, as an abbreviation in this section) about the updated definitions of coordinate systems for planets and the Moon based on multi-year observations. Details about the IAU recommendations are given in Appendix 3. For both Earth and the Moon because there are already rigidly defined coordinate systems and related matrices for high precision transformation the IAU recommendations are for reference only.

The coordinate systems of a planet probe except for the planet-fixed system of the target planet can be divided into two types, the planet's celestial coordinate system for a circling probe and the geocentric celestial coordinate system for a probe to be launched from Earth. Each of the two systems is independent but both are related. This section introduces these systems using Mars as an example.

1.5.1 Definitions of Three Mars-Centric Coordinate Systems

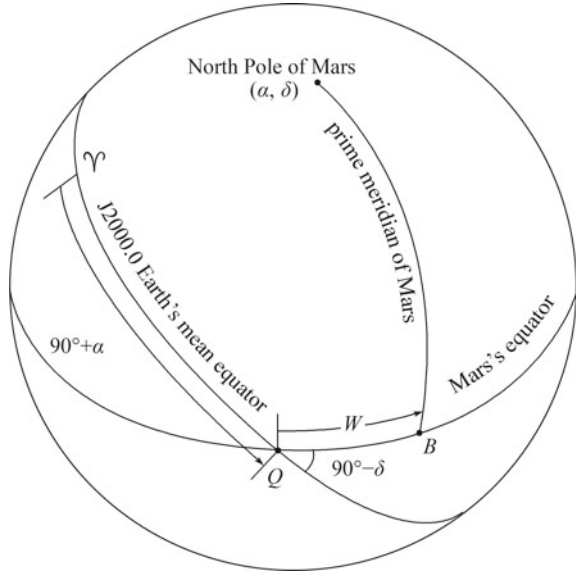
For a Mars's probe, the main systems are the **epoch J2000.0 Mars-centric equatorial coordinate system** and the **Mars-fixed coordinate system**. The Mars-fixed system is similar to the Earth-fixed system, and is related to the Mars's gravitational potential and the position of the sub-satellite point on the surface of Mars, etc. Therefore, it corresponds to a model of Mars's gravitational potential.

The **epoch J2000.0 Mars-centric equatorial coordinate system** involves the relationship with the geocentric system, similar to the Moon. The origin of all the Mars-centric systems is at the center of Mars's mass. The xy -plane is defined by the IAU 2000 celestial body orientation models for practical reason [6, 7]. The Mars's orientation model is shown in Fig. 1.6.

The IAU 2000 model provides the definitions of the **epoch J2000.0 Mars celestial coordinate system** and the **Mars-centric equatorial coordinate system**. The xy -plane of the J2000.0 Mars celestial coordinate system is the J2000.0 Earth's mean equatorial plane, the primary direction of x -axis points to the corresponding mean March equinox. For the Mars-centric equatorial coordinate system, the xy -plane is the J2000.0 Mars's mean equatorial plane, and the primary direction points to the intersection point Q of the J2000.0 Earth's mean equator and the J2000.0 Mars's mean equator in the Mars's orientation model. The point Q can be regarded as the "March equinox" in the Mars equatorial coordinate system (Fig. 1.6). The effect of nutation is not included in the model, because for Mars it is small (the largest amplitude is about $1''$) and has no accumulative effect. Therefore, for most cases, Mars's nutation can be ignored, and there is no difference between the apparent equator and the mean equator, both can be called Mars's equator.

The right ascension and declination of Mars's mean pole in the Mars's celestial coordinate system due to the precession given by the IAU 2000 Mars's orientation model (Fig. 1.6) are

Fig. 1.6 The IAU Mars's orientation model



$$\begin{aligned}\alpha &= 317^\circ.68143 - 0^\circ.1061T, \\ \delta &= 52^\circ.88650 - 0^\circ.0609T,\end{aligned}\tag{1.107}$$

where T is the Julian century number measured from J2000.0 as defined in (1.37). The values of α and δ are similar to the long-term (long-period) variations of Earth's mean pole, at epoch J2000.0 there are

$$\alpha_0 = 317^\circ.68143, \quad \delta_0 = 52^\circ.88650.\tag{1.108}$$

This is the direction of Mars's pole in the Mars's celestial coordinate system.

By this definition, Mars's rotational angle W (the arc QB in Fig. 1.6) is measured from Q eastward to B (the prime meridian as the Mars's Greenwich meridian) and can be regarded as the Mars's Greenwich sidereal time. Because the effect of Mars's nutation is omitted, the apparent sidereal time and the mean sidereal time are the same. By the IAU 2000 model Mars's rotation argument matrix is given by

$$\begin{aligned}R_{\text{IAU}}(t) &= R_z(W), \\ W &= 176^\circ.630 + 350^\circ.89198226d,\end{aligned}\tag{1.109}$$

where d is the Julian day number starting from J2000.0.

Because the above definitions are easily related to the Earth's coordinate system it is convenient for dealing with orbital problems of a Mars's probe including a launching orbit and an orbit around Mars.

1.5.2 Mars's Precession Matrix

Based on variation rules of Mars's mean pole in the IAU 2000 Mars's orientation model given in (1.107) the variation of Mars's mean equator due to precession is illustrated in Fig. 1.7, where α_0 and δ_0 are the right ascension and inclination of Mars's mean pole at epoch J2000.0, respectively, given in (1.108).

Figure 1.7 shows the geometric relationship of J2000.0 Earth's mean equator, J2000.0 Mars's mean equator, and Mars's instantaneous mean equator, where α and δ are the right ascension and inclination of Mars's instantaneous mean pole at time t , respectively, and are given by (1.107). Also in Fig. 1.7, Q and Q' are the epoch mean March equinox and the instantaneous mean March equinox in the Mars coordinate system, respectively. Assuming $\vec{r}$ and $\vec{r}'$ to be the position vectors of a Mars's probe in the epoch Mars's mean equatorial coordinate system (i.e., the Mars-centric equatorial coordinate system) and the instantaneous Mars's mean equatorial coordinate system, respectively, then the transformation relationship of the two vectors is

$$\vec{r}' = (PR)\vec{r}, \quad (1.110)$$

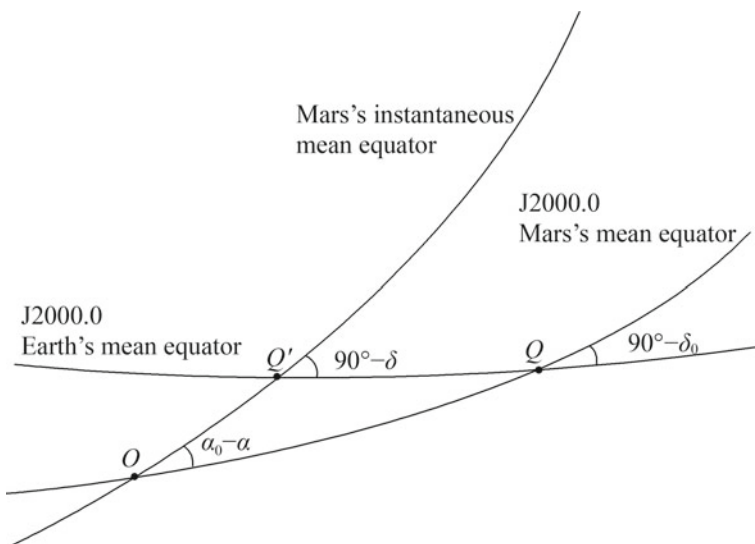


Fig. 1.7 Illustration of the variation of Mars's mean equator given by the IAU 2000 Mars's orientation model

where the transformation matrix (PR) is the Mars's precession matrix given by

$$\begin{aligned}(PR) &= R_x(90^\circ - \delta)R_z(-(\alpha_0 - \alpha))R_x(-(90^\circ - \delta_0)) \\ &= R_x(90^\circ - \delta)R_z(\alpha - \alpha_0)R_x(\delta_0 - 90^\circ).\end{aligned}\quad (1.111)$$

Note that the symbols used here are the same as in Earth's precession matrix.

1.5.3 Transformation of the Mars-Centric Equatorial Coordinate System and the Mars-Fixed Coordinate System

The orbital problem of a Mars's prober is involved in the transformation of the Mars-centric equatorial coordinate system and the Mars-fixed coordinate system. According to convention the position vector of a Mars's prober is denoted to $\vec{r}$ and $\vec{R}$ in the Mars-centric equatorial coordinate system and the Mars-fixed coordinate system, respectively, and if Mars's polar motion and nutation are ignored, then the transformation relationship of the two vectors is given by

$$\vec{R} = (MP)\vec{r}, \quad (1.112)$$

where the transformation matrix (MP) includes only two rotation matrices as

$$(MP) = (MR)(PR). \quad (1.113)$$

The matrix (MR) is the Mars rotation matrix that $(MR) = R_{\text{IAU}}(t) = R_z(W)$ with W given by (1.109).

1.5.4 Transformation of the Geocentric Coordinate System and the Mars-Centric Coordinate System

There are two types of physical variables involved in the transformation. One type includes the prober's position vector and velocity, and the other includes the position vectors of the Sun and major planets. To establish the transformation relationship the position vectors of Earth and Mars in the heliocentric ecliptic coordinate system denoted to $\vec{R}_E$, $\vec{R}_M$, are needed. For high precision requirements, their values can be found in ephemerides such as the JPL ephemeris (DE405). For general cases, the two vectors can be calculated using a simple analytical ephemeris, which requires mean orbital elements of Earth and Mars.

Earth's mean orbital elements $\bar{\sigma}$ at J2000.0 in the heliocentric ecliptic coordinate system are

$$\begin{cases} \bar{a} = 1.00000102(\text{AU}), \\ \bar{e} = 0.01670862 - 0.000042040T, \\ \bar{i} = 0^\circ.0, \\ \bar{\Omega} = 0^\circ.0, \\ \bar{\omega} = 102^\circ.937347 + 0^\circ.3225621T, \\ \bar{M} = 357^\circ.529100 + 0^\circ.98560028169d. \end{cases} \quad (1.114)$$

The Mars's mean orbital elements $\bar{\sigma}$ at J2000.0 in the heliocentric ecliptic coordinate system are

$$\begin{cases} \bar{a} = 1.52367934(\text{AU}), \\ \bar{e} = 0.09340062 + 0.000090484T, \\ \bar{i} = 1^\circ.849726 - 0^\circ.0006011T, \\ \bar{\Omega} = 49^\circ.558093 + 0^\circ.7720956T, \\ \bar{\omega} = 286^\circ.502141 + 1^\circ.068949T, \\ \bar{M} = 19^\circ.373041 + 0^\circ.52402068219d, \end{cases} \quad (1.115)$$

where time T is the Julian century number and d is the Julian day number, both start at J2000.0. The transformation from orbital elements to position vector and velocity, $\sigma(a, e, i, \Omega, \omega, M) \Rightarrow \vec{R}(X, Y, Z, \dot{X}, \dot{Y}, \dot{Z})$, is general knowledge, so is not presented here.

In the orbital problem of Mars's prober, there are two transformation relationships between the geocentric coordinate system and the Mars-centric coordinate system given as follows.

- (1) Transformations of the geocentric coordinate system and the Mars-centric coordinate system for a Mars prober

We use $\vec{R}_e, \vec{r}_e, \vec{R}_m, \vec{r}_m$ as the prober's position vectors and $\dot{\vec{R}}_e, \dot{\vec{r}}_e, \dot{\vec{R}}_m, \dot{\vec{r}}_m$ as its velocities in the Earth-fixed coordinate system, the geocentric celestial coordinate system, the Mars-fixed coordinate system, and the Mars-centric equatorial coordinate system, respectively. Do not confuse $\vec{R}_e, \vec{R}_m$ of the prober position vectors with above-mentioned position vectors $\vec{R}_E, \vec{R}_M$ of Earth and Mars in the heliocentric coordinate frame.

We need to transform the prober's position vector from a geocentric coordinate system to a Mars-centric coordinate system, which is $\vec{R}_e \Rightarrow \vec{r}_e \Rightarrow \vec{r}_m \Rightarrow \vec{R}_m$ or vice versa $\vec{R}_m \Rightarrow \vec{r}_m \Rightarrow \vec{r}_e \Rightarrow \vec{R}_e$.

From the geocentric system to the Mars-centric system, the transformations are

$$\vec{r}_e = (HG)^T \vec{R}_e, \quad (1.116)$$

$$\vec{r}'_e = R_x(\varepsilon)\vec{r}_e, \quad (1.117)$$

$$\vec{r}'_m = \left(\vec{r}'_e + \vec{R}_E\right) - \vec{R}_M, \quad (1.118)$$

$$\vec{r}_m = (ME)\vec{r}'_m, \quad (1.119)$$

$$\vec{R}_m = (MP)\vec{r}_m, \quad (1.120)$$

$$(HG) = (EP)(ER)(NR)(PR), \quad (1.121)$$

$$(ME) = R_x\left(\frac{\pi}{2} - \delta_0\right)R_z\left(\frac{\pi}{2} + \alpha_0\right)R_x(-\varepsilon), \quad (1.122)$$

$$(MP) = (MR)(PR), \quad (1.123)$$

where the transformation matrix (HG) is given by (1.30) or (1.75); α_0 and δ_0 in (ME) are given by (1.108); ε is the epoch J2000.0 mean obliquity given by (1.63); (PR) is the Mars's precession matrix given by (1.111), and the rotation matrix $(MR) = R_z(W)$ is given in (1.113) and (1.109).

The reverse transformations, $\vec{R}_m \Rightarrow \vec{r}_m \Rightarrow \vec{r}_e \Rightarrow \vec{R}_e$, are given as

$$\vec{r}_m = (MP)^T \vec{R}_m, \quad (1.124)$$

$$\vec{r}'_m = (ME)^T \vec{r}_m, \quad (1.125)$$

$$\vec{r}'_e = \vec{r}'_m + (\vec{r}_M - \vec{r}_E), \quad (1.126)$$

$$\vec{r}_e = R_x(-\varepsilon)\vec{r}'_e, \quad (1.127)$$

$$\vec{R}_e = (HG)\vec{r}_e, \quad (1.128)$$

where the matrices are previously provided.

(2) Transformations of coordinate systems related to the position vectors of the Sun and Earth

Because the orbit of a Mars's artificial satellite is perturbed by the Sun, major planets (such as Earth), and two of Mars's natural satellites the position vectors of these celestial bodies need to be provided. For the two Mars's natural satellites, Phobos and Deimos, as their orbits are located in the Mars-centric celestial coordinate system

their position vectors are known. The position vectors of the Sun and Earth in the Mars-centric equatorial coordinate system, denoted by $\vec{r}_S$ and $\vec{r}_E$, respectively, are given by

$$\vec{r}_S = (ME) \left(-\vec{R}_M \right), \quad (1.129)$$

$$\vec{r}_E = (ME) \left(\vec{R}_E - \vec{R}_M \right). \quad (1.130)$$

The transformations of the velocities of the prober and other celestial bodies are not provided here. It should be pointed out that the Mars's precession and nutation are small so the related transformation matrices can be regarded as constant matrices. For the transformations of velocity, only the variations of the rotational matrices of Earth and Mars are required, which are the angular speeds of their rotational angles.

1.5.5 An Explanation of the Application of the IAU 2000 Orientation Models of Celestial Bodies

The coordinate system selection discussed for Mars in this section can be applied to other major planets in the Solar System (such as Venus and Jupiter), with the help of the IAU 2000 celestial models of celestial bodies, the related details and numerical information are given in Appendix 3.

References

1. Liu JC, Zhu Z (2012) IAU resolutions about basic Astronomy and applications since 2000. Prog Astron 30(4):411–437, Science Press, Beijing
2. Huang C, Liu L (2015) Coordinate systems and applications in aerospace industry. Electronic Industry Press, Beijing
3. Xia YF, Huang TY (1995) Spherical astronomy. Nanjing University Press, Nanjing
4. Liu L (1992) Orbital dynamics of earth's artificial satellite. Higher Education Press, Beijing
5. Liu L (2000) Orbital theory of spacecraft. National Defense Industry Press, Beijing
6. Liu L, Hou XY (2012) Orbital dynamics of deep spacecraft. Electronic Industry Press, Beijing
7. Archinal BA, A'Hearn MF, Bowell E, Conrad A, Consolmagno GJ, Courtin R, Fukushima T, Hestroffer D, Hilton JL, Krasinsky GA, Neumann G, Oberst J, Seidelmann PK, Stooke P, Tholen DJ, Thomas PC, Williams IP (2011) Report of the IAU working group on cartographic coordinates and rotational elements: 2009. Celest Mech Dyn Astron 109:101–135
8. Seidelmann PK, Archinal BA, A'Hearn MF, Conrad A, Consolmagno GJ, Hestroffer D, Hilton JL, Krasinsky GA, Neumann G, Oberst J, Stooke P, Tedesco EF, Tholen DJ, Thomas PC, Williams IP (2007) Report of the IAU/IAG working group on cartographic coordinates and rotational elements: 2006. Celest Mech Dyn Astron 98:155–180
9. SPICE tutorial, the Navigation and Ancillary Information Facility team. ASA/JPL
10. Petit, G, Luzum B (eds.) (2010) IERS Conventions 2010, IERS Technical Note No. 36

11. Gappellari JO, Velez CE, Fuchs AJ (1976) Mathematical theory of goddard trajectory determination system. Goddard Space Flight Center, Greenbeit, Maryland. N76-24291-24302: 3-31-3-32
12. Eckhardt DH (1981) Theory of the Libration of the moon. *The Moon Planets* 25:3-49
13. Kozai Y (1960) Effect of precession and nutation on the orbital elements of a close earth satellite. *Astron J* 65(10):621-623
14. Lambeck K (1973) Precession, nutation and the choice of reference system for close earth satellite orbits. *Celest Mech Springer* 7(2):139-155
15. Kozai Y, Kinoshita H (1973) Effects of motion of the equatorial plane on the orbital elements of an earth satellite. *Celest Mech Springer* 7(3):356-366
16. Liu L, Wang X (2006) *Orbital dynamics of spacecraft for the moon's exploration*. National Defense Industry Press, Beijing

Chapter 2

The Complete Solution for the Two-Body Problem



The motions of satellites, including all sorts of spacecraft, in circling orbits correspond to perturbed dynamical systems. The most common mathematical model for a perturbed motion is a perturbed two-body problem whose reference model is the two-body problem. Therefore, the solution of the two-body problem is the foundation of the solution of a perturbed two-body problem. Although the two-body problem is introduced and discussed in numerous monographs and textbooks of Celestial Dynamics, [1–6] it is still necessary to review the fundamental concepts of the two-body problem in this book and to further introduce the algorithms for complete solutions and calculation formulas of elliptical orbits for different circumstances.

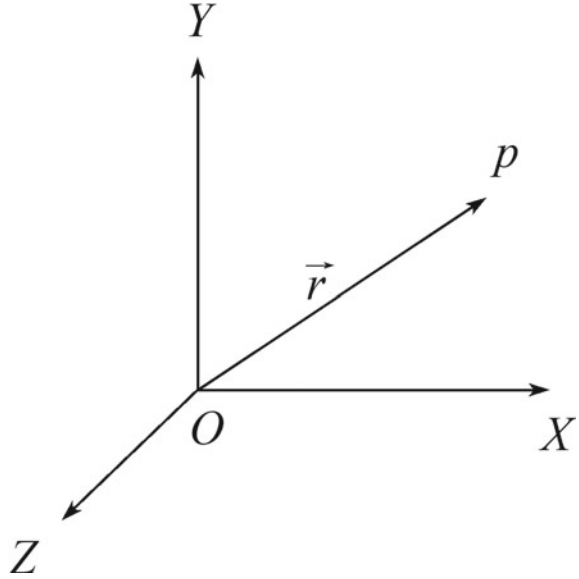
2.1 Six Integrals of the Two-Body Problem

In the two-body problem, the two bodies are denoted by P_0 for the center body and p for the less body, and their masses are denoted by m_0 and m , respectively. Both bodies are regarded as particles. The problem is to solve the motion of body p in respect to body P_0 . The motion Eq. (2.5) in the Introduction therefore can be presented as

$$\ddot{\vec{r}} = -\frac{G(m_0 + m)}{r^2}\hat{r}, \quad (2.1)$$

where $\hat{r} = \vec{r}/r$ is the unit vector in the direction from P_0 towards p . The coordinate system of motion is denoted by O -XYZ, and is centered at P_0 (Fig. 2.1). There are a few choices for the fundamental plane of the frame (the XY-plane). For the motion of Earth's artificial satellite, the XY-plane is Earth's equatorial plane. For a lunar satellite or a Mars's satellite, the choice is similar. For a planet or an asteroid in the Solar System, the XY-plane is often the heliocentric ecliptic plane. The primary direction of the coordinate system, which is the direction of X-axis, is almost always the direction towards the March equinox for motions in the Solar System, no matter

Fig. 2.1 The motion coordinate frame O - XYZ



p is a major planet or a satellite of any planet (including the Moon). As we know the equatorial plane, the ecliptic plane, and the March equinox are not fixed in the space, in this chapter we do not discuss how their variations affect the selection of a fundamental plane and the direction of a primary axis. Readers can find the details about this problem in other related chapters and in Chaps. 1 and 6 of Ref. [6].

For simplicity, the symbol μ is introduced as the standard gravitational parameter that

$$\mu = G(m_0 + m). \quad (2.2)$$

Equation (2.1) is an ordinary differential equation corresponding to a central-force problem. Theoretically, the complete solution of (2.1) can be presented by six independent constants of integration, usually given in the following form:

$$\begin{cases} \vec{r} = \vec{r}(t; C_1, \dots, C_6), \\ \dot{\vec{r}} = \dot{\vec{r}}(t; C_1, \dots, C_6). \end{cases}$$

Realistically it is difficult to obtain the solution directly from the motion equation, however, it is possible to solve the problem by finding the six independent integrals and understand the motion pattern of the two-body problem clearly through them.

2.1.1 The Angular Momentum Integral (the Areal Integral)

From the characteristics of the central-force, we can directly find the angular momentum integral of Eq. (2.1). Let $\vec{h} = \vec{r} \times \dot{\vec{r}}$ as the areal velocity then we have the angular momentum integral as

$$\vec{h} = \vec{r} \times \dot{\vec{r}} = h \hat{R}, \quad (2.3)$$

where $\vec{h}$ is a constant vector and the motion of p relative to P_0 is a planar motion, and $h = |\vec{r} \times \dot{\vec{r}}|$ is the areal speed constant. The unit vector denoted to $\hat{R}$ gives the direction of the areal velocity, which is the normal of the motion plane.

The geometric meaning of the angular momentum is often illustrated by an auxiliary celestial sphere shown in Fig. 2.2, where circles AA' and BB' are for the projections of the fundamental plane (the XY -plane) and the motion orbit on the auxiliary sphere, respectively. The prime body P_0 is at the origin of the coordinate frame, $\hat{R}$ is the normal direction of the orbital plane, and i is the included angle of the orbital plane and the XY -plane. If the primary direction (i.e., the direction of X -axis) points to the March equinox then Ω is the longitude of the ascending node (N) (N is the point when body p crosses AA' from the southern hemisphere to the northern hemisphere) measured from the X -direction. By the cosine rules of the Spherical Trigonometry (or using the coordinate rotation method), the unit vector $\hat{R}$ in the O - XYZ coordinate system is deduced as

$$\hat{R} = \frac{\vec{r} \times \dot{\vec{r}}}{h} = \begin{pmatrix} R_x \\ R_y \\ R_z \end{pmatrix} = \begin{pmatrix} \sin i \sin \Omega \\ -\sin i \cos \Omega \\ \cos i \end{pmatrix} \quad (2.4)$$

There are three integrals, h , i , and Ω , in the angular momentum integral (2.3) that h is twice the areal speed, i and Ω decide the position and orientation of the orbital plane in space. The constant h is often expressed in a different form and is discussed below.

2.1.2 The Orbital Integral in the Motion Plane and the Vis Viva Formula

Since the motion of body p is on a plane, and the plane is defined by (i, Ω) , the order of the original differential equation can be reduced and the motion problem can be discussed on the plane. We now introduce the polar coordinates (r, θ) of a moving body on the plane, then the radial component of the motion Eq. (2.1) is given by

One of the characteristics of the motion equation is that the independent variable t is not explicitly included in the equation. According to the basic knowledge of the ordinary differential equation, the orders of the equation can be reduced by one via separating the independent variable t . We now can discuss the variation of r with respect to θ first. Write $r' = dr/d\theta$ and $r'' = d^2r/d\theta^2$ by (2.7), we have

$$\begin{cases} \dot{r} = \frac{dr}{dt} = \frac{dr}{d\theta} \dot{\theta} = \frac{h}{r^2} r' \\ \ddot{r} = \frac{d\dot{r}}{dt} = \frac{d\dot{r}}{d\theta} \dot{\theta} = \frac{h^2}{r^2} \left(-\frac{2}{r^3} r'^2 + \frac{1}{r^2} r'' \right) \end{cases} \quad (2.10)$$

Substituting this relationship into (2.5) gives a differential equation of the second order for r with respect to θ . But this equation is still not easy to solve. We then introduce a new variable u for a variable transformation that

$$r = 1/u, \quad (2.11)$$

which results $\dot{r} = -hu'$, $\ddot{r} = -h^2u^2u''$. We then obtain an ordinary differential equation of the second order for u with respect to θ , as

$$u'' + u = \frac{\mu}{h^2}. \quad (2.12)$$

Integrating (2.12) leads to the orbital integral:

$$r = \frac{1}{u} = \frac{h^2/\mu}{1 + e \cos(\theta - \omega)}. \quad (2.13)$$

In which e and ω are two new constants of integration, and the expression of r is a conic curve. Under certain conditions it is an ellipse, the central body is at one of its foci (i.e., point O). Since this book is about satellite orbits, we assume that it is an ellipse (two other conic curves, hyperbola and parabola, are introduced briefly in the last section of the chapter). Write

$$p = a(1 - e^2) = h^2/\mu, \quad (2.14)$$

then (2.7) and (2.13) can be written as

$$r^2 \dot{\theta} = \sqrt{\mu p} = \sqrt{\mu a(1 - e^2)}, \quad (2.15)$$

$$r = \frac{a(1 - e^2)}{1 + e \cos(\theta - \omega)}. \quad (2.16)$$

The constant of integration h now is replaced by a , which is the semi-major axis; p is the semi-latus rectum; e is the eccentricity; ω is the argument of periapsis (perigee for an Earth's satellite) for the moving body, because at point P there is $\theta = \omega$, so the distance between the two bodies has its minimum, and point P is also called the periapsis point. Note that in the two-body problem both angles, θ and ω , start at the fixed ascending node N . For a perturbed elliptical problem, the ellipse changes over time, so does the ascending node, then the starting point of the angle ω is measured from a changing ascending node, whereas the angle θ is defined to start at a fixed point, that is the difference of how the two angles are measured.

Theoretically the last integral, which is related to the time t , can be obtained by substituting (2.13) $r = r(\theta)$ into (2.15), this problem is discussed in the following section. Here we first give some useful relationships in an elliptical motion. By (2.15) and (2.16) using simple calculations we have

$$v^2 = \dot{r}^2 + r^2 \dot{\theta}^2 = \mu \left(\frac{2}{r} - \frac{1}{a} \right). \quad (2.17)$$

The relationship (2.17) is given by a Latin name, vis viva formula, also called the orbital-energy-invariance formula.

Since it is an elliptical motion, the area covered by the radial of a moving body over a period T is the area of the ellipse, which is $\pi a^2 \sqrt{1 - e^2}$. Therefore, the value of h , i.e., two times the area speed, is given by

$$h = \sqrt{\mu a (1 - e^2)} = 2\pi a^2 \sqrt{(1 - e^2)} / T \quad (2.18)$$

Rearranging the variables, we have

$$\frac{a^3}{T^2} = \frac{\mu}{4\pi^2} \quad (2.19)$$

Introducing n as the mean angle speed, that $n = 2\pi / T$, then (2.19) becomes

$$n^2 a^3 = \mu. \quad (2.20)$$

These two formulas are exactly Kepler's third law derived from Newton's law of universal gravitation.

It should be mentioned that similar to the method of deriving the angular moment integral (2.3), the vis viva formula (2.17) can be also derived from (2.1) directly by multiplying $\dot{\vec{r}}$ on both sides, resulting

$$\dot{\vec{r}} \cdot \ddot{\vec{r}} = -\frac{\mu}{r^3} \vec{r} \cdot \dot{\vec{r}}, \quad \frac{1}{2} \left(\dot{\vec{r}}^2 \right) = \mu \frac{d}{dt} \left(\frac{1}{r} \right).$$

Then by integrating we have

$$v^2 = \mu \left(\frac{2}{r} + C \right),$$

which is exactly the vis viva formula. The vis viva formula actually is the energy constant for the two-body problem. The energy constant is one of the 10 classical integrals of an N -body ($N \geq 2$) problem. In the two-body problem, we have derived five independent integrals, the last one must be related to the independent variable time t (reflecting the orbital motion).

2.1.3 The Sixth Motion Integral: Kepler's Equation

For convenience, the sixth integral is not directly derived by integrating $d\theta/dt$ given in (2.15) but by integrating dr/dt given in (2.17). Rewriting Eq. (2.17) gives

$$\dot{r}^2 = \mu \left(\frac{2}{r} - \frac{1}{a} \right) - (r\dot{\theta})^2 = \mu \left(\frac{2}{r} - \frac{1}{a} \right) - \mu \frac{p}{r^2}.$$

Eliminating μ by (2.20) leads

$$ndt = \frac{rdr}{a\sqrt{a^2e^2 - (a-r)^2}}. \quad (2.21)$$

For an elliptical orbit, the values of the maximum and minimum of r are given by

$$r_{\max} = a(1+e), \quad r_{\min} = a(1-e), \quad (2.22)$$

which shows that $|a-r| \leq ae$. Therefore, we can introduce an auxiliary variable E defined by

$$a-r = ae \cos E$$

or

$$r = a(1 - e \cos E). \quad (2.23)$$

Substituting (2.23) into (2.21) yields

$$ndt = (1 - e \cos E)dE.$$

Integrating (2.23) gives the sixth integral as

$$E - e \sin E = n(t - \tau). \quad (2.24)$$

The Eq. (2.24) is called Kepler's equation with τ as the constant of integration. When $t = \tau$, there is $E = 0$, and $r = a(1 - e) = r_{\min}$, therefore τ is the moment when the moving body passes the periapsis.

Finally, we introduce two angles, f and M , defined as

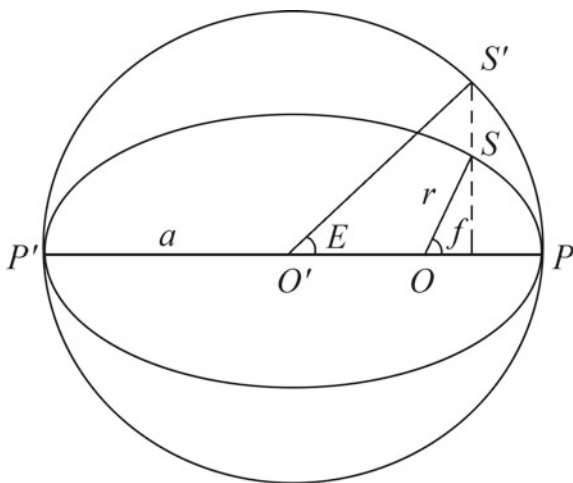
$$f = \theta - \omega, \quad M = n(t - \tau). \quad (2.25)$$

The three angles, f the true anomaly, M the mean anomaly, and E the eccentric anomaly, are all measured from the periapsis. The geometric meaning of E is shown in Fig. 2.3, where O is one of the foci of an ellipse (the origin of the coordinate system), O' is the center of the auxiliary circle. Obviously in the two-body problem, the areal integral (2.7) can be simplified to

$$r^2 \dot{f} = h. \quad (2.26)$$

The above given six independent integrals are a set of independent parameters, often called the orbital elements, which are used to describe the motion of a celestial body. When the initial condition is given the six elements are decided. The elements a and e give the size and shape of an orbit; i , Ω , and ω are parameters defining the orientations of the orbital plane and the major axis; and the sixth element τ is usually replaced by one of the three anomalies, most likely the mean anomaly M . The three anomalies are functions of time, so are not constants, also called time elements. The six elements a , e , i , Ω , ω , and M (f , E) have historic meanings in the development of the Celestial Mechanics, and for their importance are often called Kepler elements.

Fig. 2.3 The elliptical orbit and the auxiliary circle



2.2 Basic Formulas of the Elliptical Orbital Motion

As described above that the six elements define the entire motion in the two-body problem. But in some practical situations, it is not convenient to use the expressions of the six elements. Based on the requirements of actual work, we provide some useful formulas which deal with the six elements, time, anomalies, radial, speed, etc., and organize them as follows.

2.2.1 Geometric Relationships of the Orbital Elements in the Elliptical Motion

From Fig. 2.3 and Kepler's Eq. (2.24) it is not difficult to notice that the three anomalies are all in $[0, \pi]$ or $[\pi, 2\pi]$ at the same time. This is an important relationship, and their connections are given by

$$r = \frac{a(1 - e^2)}{1 + e \cos f} = a(1 - e \cos E), \quad (2.27)$$

$$E - e \sin E = M. \quad (2.28)$$

From the characteristics of the ellipse and in Fig. 2.3 there is $\overline{OO'} = ae$, which yields

$$r \cos f = a(\cos E - e), \quad (2.29)$$

$$r \sin f = a\sqrt{1 - e^2} \sin E, \quad (2.30)$$

$$\tan \frac{f}{2} = \sqrt{\frac{1 + e}{1 - e}} \tan \frac{E}{2}. \quad (2.31)$$

2.2.2 Expressions of the Position Vector $\vec{r}$ and Velocity $\dot{\vec{r}}$

As mentioned above Eq. (2.1) is a differential equation of the second order, whose solution can have a form as

$$\begin{cases} \vec{r} = \vec{r}(t; C_1, \dots, C_6), \\ \dot{\vec{r}} = \dot{\vec{r}}(t; C_1, \dots, C_6). \end{cases} \quad (2.32)$$

By the six integrals, the solution of Eqs. (2.32) now can be expressed explicitly. The constants $C_1, \dots, C_6$ are the six orbital elements, where C_6 is τ , or if choose M to replace τ , then t in (2.32) is included in M .

For the position vector, there is

$$\begin{aligned}\vec{r} &= r\hat{r} = r \cos f \hat{P} + r \sin f \hat{Q} \\ &= a(\cos E - e)\hat{P} + a\sqrt{1-e^2} \sin E \hat{Q},\end{aligned}\quad (2.33)$$

where $\hat{P}$ and $\hat{Q}$ are unit vectors in the periapsis direction and the semi-latus rectum direction, respectively. These two vectors can be expressed in the Cartesian coordinate system $O-XYZ$ by a rotation. If on the orbital plane the x -axis of the planar Cartesian coordinate system points to the periapsis, then the unit vector $\hat{P}_0$ in this direction has the form as

$$\hat{P}_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \quad (2.34)$$

Then in the $O-XYZ$ coordinate system, $\hat{P}$ is given by three rotation matrices that

$$\hat{P} = R_z(-\Omega)R_x(-i)R_z(-\omega)\hat{P}_0, \quad (2.35)$$

and the three rotation matrices are

$$R_z(-\omega) = \begin{pmatrix} \cos \omega & -\sin \omega & 0 \\ \sin \omega & \cos \omega & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (2.36)$$

$$R_x(-i) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos i & -\sin i \\ 0 & \sin i & \cos i \end{pmatrix}, \quad (2.37)$$

$$R_z(-\Omega) = \begin{pmatrix} \cos \Omega & -\sin \Omega & 0 \\ \sin \Omega & \cos \Omega & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (2.38)$$

The expression of $\hat{Q}$ can be given by replacing $R_z(-\omega)$ with $R_z(\alpha)$, where $\alpha = -(\omega + 90^\circ)$.

In some cases, the analytical formulas of $\hat{P}$ and $\hat{Q}$ are given by

$$\hat{P} = \begin{pmatrix} \cos \Omega \cos \omega - \sin \Omega \sin \omega \cos i \\ \sin \Omega \cos \omega + \cos \Omega \sin \omega \cos i \\ \sin \omega \sin i \end{pmatrix}, \quad (2.39)$$

$$\hat{Q} = \begin{pmatrix} -\cos \Omega \sin \omega - \sin \Omega \cos \omega \cos i \\ -\sin \Omega \sin \omega + \cos \Omega \cos \omega \cos i \\ \cos \omega \sin i \end{pmatrix}. \quad (2.40)$$

For $\dot{\vec{r}}$ based on the characteristics of the two-body problem, and by (2.33), we have

$$\dot{\vec{r}} = \frac{\partial \vec{r}}{\partial f} \frac{df}{dt} = \frac{\partial \vec{r}}{\partial E} \frac{dE}{dt}. \quad (2.41)$$

With f given by the areal constant (2.26) or $\dot{E}$ given by Kepler's Eq. (2.28), we have

$$\begin{aligned} \dot{\vec{r}} &= -\sqrt{\frac{\mu}{p}} \left[\sin f \hat{P} - (\cos f + e) \hat{Q} \right] \\ &= -\frac{\sqrt{\mu a}}{r} \left[\sin E \hat{P} - \sqrt{1 - e^2} \cos E \hat{Q} \right]. \end{aligned} \quad (2.42)$$

For some particular problems, the six elements need to be expressed as functions of initial values of $\vec{r}(t)$ and $\dot{\vec{r}}(t)$, i.e., at the time $t = t_0$, $\vec{r}(t_0) = \vec{r}_0$, $\dot{\vec{r}}(t_0) = \dot{\vec{r}}_0$, that

$$\begin{cases} \vec{r} = \vec{r}(t; t_0, \vec{r}_0, \dot{\vec{r}}_0), \\ \dot{\vec{r}} = \dot{\vec{r}}(t; t_0, \vec{r}_0, \dot{\vec{r}}_0). \end{cases} \quad (2.43)$$

The analytical formulas can be derived by (2.33) and (2.42). First, we express $\vec{r}_0$ and $\dot{\vec{r}}_0$ in directions of $\hat{P}$ and $\hat{Q}$, that

$$\begin{cases} \vec{r}_0 = a(\cos E_0 - e) \hat{P} + a\sqrt{1 - e^2} \sin E_0 \hat{Q}, \\ \dot{\vec{r}}_0 = \frac{\sqrt{\mu a}}{r_0} \left[(-\sin E_0) \hat{P} + \sqrt{1 - e^2} \cos E_0 \hat{Q} \right]. \end{cases} \quad (2.44)$$

Then $\hat{P}$ and $\hat{Q}$ can be obtained by solving (2.44). Substituting the solutions of $\hat{P}$ and $\hat{Q}$ into (2.33) and (2.42), after arranging we have $\vec{r}(t)$ and $\dot{\vec{r}}(t)$ presented by 'linear' combinations of $\vec{r}_0$ and $\dot{\vec{r}}_0$ as

$$\begin{cases} \vec{r} = F \vec{r}_0 + G \dot{\vec{r}}_0, \\ \dot{\vec{r}} = F' \dot{\vec{r}}_0 + G' \vec{r}_0, \end{cases} \quad (2.45)$$

where F , G , F' , and G' are still functions of $\vec{r}_0$ and $\dot{\vec{r}}_0$.

The formulas for F and G are

$$\begin{cases} F = 1 - \frac{a}{r_0} (1 - \cos \Delta E), \\ G = \Delta t - \frac{1}{n} (\Delta E - \sin \Delta E), \end{cases} \quad (2.46)$$

where $\Delta t = t - t_0$, $\Delta E = E - E_0$, and a and ΔE are calculated by

$$a = \left(\frac{2}{r_0} - \frac{v_0^2}{\mu} \right)^{-1}, \quad v_0^2 = \dot{r}_0^2, \quad (2.47)$$

$$\Delta E = n\Delta t + \left[\left(1 - \frac{r_0}{a} \right) \sin \Delta E - \frac{r_0 \dot{r}_0}{\sqrt{\mu a}} (1 - \cos \Delta E) \right], \quad (2.48)$$

$$n = \sqrt{\mu a}^{-\frac{3}{2}}, \quad r_0 \dot{r}_0 = \vec{r}_0 \cdot \dot{\vec{r}}_0. \quad (2.49)$$

Because

$$\left(1 - \frac{r_0}{a} \right) = O(e), \quad \frac{r_0 \dot{r}_0}{\sqrt{a}} = O(e). \quad (2.50)$$

Formula (2.48) is similar to Kepler's Equation. It is not difficult to calculate ΔE , especially when Δt is small that calculating ΔE is faster than solving Kepler's Equation (the method to solve Kepler's Equation is provided in Sect. 2.4.4).

The formulas of F' and G' can be given according to $F' = \dot{G}$ and $G' = \dot{F}$ that

$$\begin{cases} F' = 1 - \frac{a}{r_0} (1 - \cos \Delta E), \\ G' = -\frac{1}{r_0} \left(\frac{\sqrt{\mu a}}{r_0} \sin \Delta E \right). \end{cases} \quad (2.51)$$

When Δt is small we have

$$\begin{cases} F = 1 + O(\Delta t^2), \\ G = \Delta t [1 + O(\Delta t^2)]. \end{cases} \quad (2.52)$$

Based on the characteristics of F , G , F' , and G' , these formulas can be expressed by power series in Δt . The actual formulas are given in Refs. [5] and [6]. We now give a brief introduction to this method so readers can understand it better. Everyone who studies the ordinary differential equation knows that if the right side of Eq. (2.1) satisfies certain conditions (not provided here, but Eq. 2.1 in fact does meet those conditions), then the equation has a solution satisfying the initial conditions, and the solution can be expressed by power series in $\Delta t = t - t_0$, thus we have

$$\vec{r}(t) = \vec{r}_0 + \vec{r}_0^{(1)} \Delta t + \frac{1}{2!} \vec{r}_0^{(2)} \Delta t^2 + \cdots + \frac{1}{k!} \vec{r}_0^{(k)} \Delta t^k + \cdots, \quad (2.53)$$

where $\vec{r}_0^{(k)}$ is the value of the k th derivative of $\vec{r}(t)$ with respect to t at time t_0 , which is

$$\vec{r}_0^{(k)} = \left(\frac{d^k \vec{r}}{dt^k} \right)_{t=t_0}. \quad (2.54)$$

To give the complete analytical solution in power series we have to calculate the values of all orders of the derivative of $\vec{r}^{(k)}$ at time t_0 . Obviously, there is $\vec{r}_0^{(1)} = \dot{\vec{r}}_0$, then for any order greater than 2, the derivative of $\vec{r}_0^{(k)}$ ($k \geq 2$) can be given by $\vec{r}_0$ and $\dot{\vec{r}}_0$ as

$$\vec{r}_0^{(k)} = \vec{r}_0^{(k)}(t_0, \vec{r}_0, \dot{\vec{r}}_0), \quad k \geq 2. \quad (2.55)$$

Then the power series solution (2.53) can be organized as a function of $\vec{r}_0$ and $\dot{\vec{r}}_0$ that

$$\vec{r}(t) = F(\vec{r}_0, \dot{\vec{r}}_0, \Delta t) \vec{r}_0 + G(\vec{r}_0, \dot{\vec{r}}_0, \Delta t) \dot{\vec{r}}_0. \quad (2.56)$$

For the two-body problem presented by Eq. (2.1), F and G in (2.56) can be expressed by power series in Δt as provided below.

In practice, it is convenient to use normalized units, which means to use particular units of mass and length to make the gravitational constant $G = 1$, and $\mu = G(m_0 + m) = 1$. The normalized units are particularly useful when we analyze magnitudes of variables. Here we choose $(m_0 + m)$ as the unit of mass, and L as the unit of length, which can be the equatorial radius of the central body P_0 or other suitable lengths, then the corresponding unit of time is $\{L^3/[G(m_0 + m)]\}^{1/2}$. Using this normalized unit system, we have

$$\begin{cases} F = 1 - \left(\frac{1}{2}u_0\right)\Delta t^2 + \left(\frac{1}{2}u_0p_0\right)\Delta t^3 + \left(\frac{1}{8}u_0q_0 - \frac{1}{12}u_0^2 - \frac{5}{8}u_0p_0^2\right)\Delta t^4 + O(\Delta t^5), \\ G = \Delta t - \left(\frac{1}{6}u_0\right)\Delta t^3 + \left(\frac{1}{4}u_0p_0\right)\Delta t^4 + O(\Delta t^5), \end{cases} \quad (2.57)$$

where

$$u_0 = \frac{1}{r_0^3}, p_0 = \frac{\vec{r}_0 \dot{\vec{r}}_0}{r_0^2}, q_0 = \frac{v_0^2}{r_0^2}, \quad (2.58)$$

and

$$\begin{cases} F' = 1 - \left(\frac{1}{2}u_0\right)\Delta t^2 + (u_0p_0)\Delta t^3 + O(\Delta t^4), \\ G' = -(u_0)\Delta t + \left(\frac{3}{2}u_0p_0\right)\Delta t^2 + \left(\frac{1}{2}u_0q_0 - \frac{1}{3}u_0^2 - \frac{5}{2}u_0p_0^2\right)\Delta t^3 + O(\Delta t^4). \end{cases} \quad (2.59)$$

In the two-body problem if we assume r_0 to be approximately the semi-major axis of the orbit a , then in the normalized units system $u_0 = r_0^{-3} \approx n^2$, where n is the mean angular speed. We now can estimate the orders of magnitude of F and G that

$$\begin{cases} F = 1 + O(\Delta\tau^2), \\ G = \Delta t[1 + O(\Delta\tau^2)], \end{cases} \quad (2.60)$$

where $\Delta\tau = n\Delta t$ is the arc that the moving body passes in the time interval of Δt . This property of the estimated magnitudes is important initial information, its application is discussed in Chap. 6.

2.2.3 Partial Derivatives of Some Variables with Respect to Orbital Elements

To research the motion of a celestial body and to calculate its position, we must deal with the six orbital elements, $a, e, i, \Omega, \omega, M$, and also some functions of the elements. In these functions, the basic variables are E, f , and r , therefore, we need to know the partial derivatives of these variables with respect to each orbital element.

First, we analyze E, f , and r as functions of the six orbital elements. From (2.27)–(2.30) there are

$$\begin{cases} E = E(e, M), \\ f = f\left(e, E(e, M), \frac{a}{r}(e, M)\right) = f(e, M), \\ r = r(a, e, E(e, M)) = r(a, e, M). \end{cases}$$

Then according to the above given geometric relationships, we derive the related partial derivatives as

$$\begin{cases} \frac{\partial E}{\partial e} = \frac{a}{r} \sin E, & \frac{\partial E}{\partial M} = \frac{a}{r}, \\ \frac{\partial f}{\partial e} = \frac{1}{1 - e^2} \left(1 + \frac{p}{r}\right) \sin f, & \frac{\partial f}{\partial M} = \sqrt{1 - e^2} \left(\frac{a}{r}\right)^2, \\ \frac{\partial r}{\partial a} = \frac{r}{a}, \quad \frac{\partial r}{\partial e} = -a \cos f, & \frac{\partial r}{\partial M} = \frac{ae}{\sqrt{1 - e^2}} \sin f. \end{cases} \quad (2.61)$$

If choosing M as the independent element instead of E there are $M = M(e, E)$, $f = f(e, E)$, $r = r(a, e, E)$,

$$\begin{cases} \frac{\partial M}{\partial e} = -\sin E, & \frac{\partial M}{\partial E} = \left(\frac{r}{a}\right) = 1 - e \cos E, \\ \frac{\partial f}{\partial e} = \frac{\sin f}{1 - e^2}, & \frac{\partial f}{\partial E} = \left(\frac{a}{r}\right) \sqrt{1 - e^2}, \\ \frac{\partial r}{\partial a} = \left(\frac{r}{a}\right), \quad \frac{\partial r}{\partial e} = -a \cos E, & \frac{\partial r}{\partial E} = ae \sin E. \end{cases} \quad (2.62)$$

If choosing f as the independent element instead of M , there are $E = E(e, f)$, $M = M(e, E(e, f)) = M(e, f)$, $r = r(a, e, f)$,

$$\begin{cases} \frac{\partial E}{\partial e} = -\frac{\sin E}{1-e^2}, \quad \frac{\partial E}{\partial f} = \left(\frac{r}{a}\right)/\sqrt{1-e^2}, \\ \frac{\partial M}{\partial e} = -\left[1 + \left(\frac{r}{a}\right)/(1-e^2)\right]\left(\frac{r}{a}\right)\frac{\sin f}{\sqrt{1-e^2}}, \quad \frac{\partial M}{\partial f} = \left(\frac{r}{a}\right)^2/\sqrt{1-e^2}, \\ \frac{\partial r}{\partial a} = \left(\frac{r}{a}\right), \quad \frac{\partial r}{\partial e} = \frac{r}{1-e^2}(\cos E + e), \quad \frac{\partial r}{\partial f} = \left(\frac{r}{a}\right)^2\left(\frac{ae}{1-e^2}\right)\sin f. \end{cases} \quad (2.63)$$

In applications, we often meet the factor $\left(\frac{a}{r}\right)$. From $\partial r/\partial \sigma$, we can get $\partial\left(\frac{a}{r}\right)/\partial \sigma$ directly. Obviously, $\left(\frac{a}{r}\right)$ is a function of e and an anomaly, so we have

$$\frac{\partial\left(\frac{a}{r}\right)}{\partial e} = -\left(\frac{a}{r^2}\right)\left(\frac{\partial r}{\partial e}\right), \quad \frac{\partial\left(\frac{a}{r}\right)}{\partial \theta} = -\left(\frac{a}{r^2}\right)\left(\frac{\partial r}{\partial \theta}\right),$$

where θ refers to one of the three anomalies, M , E , and f .

When the eccentricity e is small, we often use a set of regrouped orbital elements instead of the original ones, which are

$$a, i, \Omega, \xi = e \cos \omega, \eta = e \sin \omega, \lambda = M + \omega.$$

Accordingly, f and E are replaced by $u = f + \omega$, and $v = E + \omega$, respectively. In deriving the related partial derivatives, the key is to analyze the connections between variables. From the definitions of the new elements, there are

$$e^2 = \xi^2 + \eta^2, \quad \omega = \arctan\left(\frac{\eta}{\xi}\right), \quad M = \lambda - \arctan\left(\frac{\eta}{\xi}\right), \quad (2.64)$$

which give

$$\begin{aligned} f &= f(e(\xi, \eta), M(\xi, \eta, \lambda)), \\ E &= E(e(\xi, \eta), M(\xi, \eta, \lambda)), \\ \frac{a}{r} &= \frac{a}{r}(e(\xi, \eta), M(\xi, \eta, \lambda)). \end{aligned}$$

From these relationships, we can derive the partial derivatives. For example,

$$\begin{cases} \frac{\partial}{\partial \xi}\left(\frac{a}{r}\right) = \frac{\partial}{\partial e}\left(\frac{a}{r}\right)\frac{\partial e}{\partial \xi} + \frac{\partial}{\partial M}\left(\frac{a}{r}\right)\frac{\partial M}{\partial \xi}, \\ \frac{\partial}{\partial \eta}\left(\frac{a}{r}\right) = \frac{\partial}{\partial e}\left(\frac{a}{r}\right)\frac{\partial e}{\partial \eta} + \frac{\partial}{\partial M}\left(\frac{a}{r}\right)\frac{\partial M}{\partial \eta}, \\ \frac{\partial}{\partial \lambda}\left(\frac{a}{r}\right) = \frac{\partial}{\partial M}\left(\frac{a}{r}\right)\frac{\partial M}{\partial \lambda}, \end{cases} \quad (2.65)$$

where $\frac{\partial}{\partial e}(\frac{a}{r})$ and $\frac{\partial}{\partial M}(\frac{a}{r})$ are given already, what left are $\frac{\partial e}{\partial \xi}$, $\frac{\partial M}{\partial \xi}$, $\dots$. We do not list the results here because most readers are able to finish the job.

In the orbital determination, we also need two sets of partial derivatives, $\frac{\partial \vec{r}}{\partial \sigma}$ and $\frac{\partial \dot{\vec{r}}}{\partial \sigma}$. From the expressions of $\vec{r}$ (2.33) and $\dot{\vec{r}}$ (2.42) if we take $\sigma(a, e, i, \Omega, \omega, M)$ as the basic variables, then there are two sets of partial derivatives, that one includes partial derivatives with respect to a, e , and M as $\frac{\partial(r, f, E)}{\partial(a, e, M)}$ and the other includes partial derivatives of unit vectors, $\hat{P}$ and $\hat{Q}$, with respect to the three angles, i, Ω , and ω , as $\frac{\partial \hat{P}}{\partial(i, \Omega, \omega)}$ and $\frac{\partial \hat{Q}}{\partial(i, \Omega, \omega)}$. These partial derivatives can be derived from (2.39) and (2.40) directly, but the results are complicated, it is better to use the vector rotation method, and the results are

$$\left\{ \begin{array}{l} \frac{\partial \vec{r}}{\partial a} = \frac{1}{a} \vec{r}, \\ \frac{\partial \vec{r}}{\partial e} = H \vec{r} + K \dot{\vec{r}}, \\ \frac{\partial \vec{r}}{\partial M} = \frac{1}{n} \dot{\vec{r}}, \\ \frac{\partial \vec{r}}{\partial i} = \hat{J}_N \times \vec{r} = \frac{z}{\sin i} \hat{R}, \hat{J}_N = \begin{pmatrix} \cos \Omega \\ \sin \Omega \\ 0 \end{pmatrix}, \\ \frac{\partial \vec{r}}{\partial \Omega} = \hat{J}_z \times \vec{r} = \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}, \hat{J}_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \\ \frac{\partial \vec{r}}{\partial \omega} = \hat{R} \times \vec{r} = \begin{pmatrix} z R_y - y R_z \\ x R_z - z R_x \\ y R_x - x R_y \end{pmatrix}, \end{array} \right. \quad (2.66)$$

$$\left\{ \begin{array}{l} \frac{\partial \dot{\vec{r}}}{\partial a} = -\frac{1}{2a} \dot{\vec{r}}, \\ \frac{\partial \dot{\vec{r}}}{\partial e} = H' \vec{r} + K' \dot{\vec{r}}, \\ \frac{\partial \dot{\vec{r}}}{\partial M} = -n \left(\frac{a}{r} \right)^3 \vec{r} = -\frac{\mu}{n} \left(\frac{\vec{r}}{r^3} \right), \\ \frac{\partial \dot{\vec{r}}}{\partial i} = \hat{J}_N \times \dot{\vec{r}} = \frac{\dot{z}}{\sin i} \hat{R}, \\ \frac{\partial \dot{\vec{r}}}{\partial \Omega} = \hat{J}_z \times \dot{\vec{r}} = \begin{pmatrix} -\dot{y} \\ \dot{x} \\ 0 \end{pmatrix}, \\ \frac{\partial \dot{\vec{r}}}{\partial \omega} = \hat{R} \times \dot{\vec{r}} = \begin{pmatrix} \dot{z} R_y - \dot{y} R_z \\ \dot{x} R_z - \dot{z} R_x \\ \dot{y} R_x - \dot{x} R_y \end{pmatrix}, \end{array} \right. \quad (2.67)$$

where $\hat{R}$ is the unit vector of the normal to the orbital plane given by (2.4), and it can be also given by

$$\hat{R} = \frac{1}{\sqrt{\mu p}} (\vec{r} \times \dot{\vec{r}}), \quad (2.68)$$

with

$$\sin i = \sqrt{1 - \cos^2 i} = \sqrt{1 - R_z^2}. \quad (2.69)$$

Also, H and K in (2.66), and H' and K' in (2.67) are given by

$$H = -\frac{a}{p}(\cos E + e), \quad K = \frac{\sin E}{n} \left(1 + \frac{r}{p}\right), \quad (2.70)$$

$$H' = \frac{\sqrt{\mu a} \sin E}{rp} \left[1 - \frac{a}{r} \left(1 + \frac{p}{r}\right)\right], \quad K' = \frac{a}{p} \cos E. \quad (2.71)$$

2.2.4 Derivatives of M , E , and F with Respect to Time t

The derivatives of the three variables with respect to time t can be given from their definitions with the areal integral (2.26) and Kepler's Eq. (2.28). The results are:

$$\frac{dM}{dt} = n, \quad \frac{dE}{dt} = n \left(\frac{a}{r}\right), \quad \frac{df}{dt} = n \sqrt{1 - e^2} \left(\frac{a}{r}\right)^2. \quad (2.72)$$

In this book, we often come across transformations between these variables. For convenience, the derivatives given in (2.72) are rearranged as

$$dM = ndt = \left(\frac{r}{a}\right) dE = \frac{1}{\sqrt{1 - e^2}} \left(\frac{r}{a}\right)^2 df, \quad (2.73)$$

$$dE = n \left(\frac{a}{r}\right) dt = \left(\frac{a}{r}\right) dM = \frac{1}{\sqrt{1 - e^2}} \left(\frac{r}{a}\right) df, \quad (2.74)$$

$$df = n \sqrt{1 - e^2} \left(\frac{a}{r}\right)^2 dt = \sqrt{1 - e^2} \left(\frac{a}{r}\right)^2 dM = \sqrt{1 - e^2} \left(\frac{a}{r}\right) dE, \quad (2.75)$$

$$dt = \frac{1}{n} dM = \frac{1}{n} \left(\frac{r}{a}\right) dE = \frac{1}{n \sqrt{1 - e^2}} \left(\frac{r}{a}\right)^2 df. \quad (2.76)$$

Note that the above derivatives are given assuming the six orbital elements to be constants. Strictly speaking, they are correct only for the two-body problem, whereas the previously given geometric relationships and related partial derivatives are not restricted to the two-body problem.

2.3 Expansions of Variables in the Elliptical Orbital Motion

To solve the elliptical motion equation, we often need to express some variables as explicit functions of time t through the mean anomaly M . This requirement is impossible to achieve because Kepler's Equation involves transcendental functions. In order to circumvent the impossibility, the method applied here is to expand variables like f , E , and $(\frac{a}{r})$ into series of trigonometric functions of M through two special functions, the Bessel function of the first kind and the hypergeometric function (also called hypergeometric series). Here we give a brief introduction of the two special functions, for a better understanding readers are suggested to refer to specialized books about special functions.

The Bessel function of the first kind $J_n(x)$ is a solution of a linear ordinary differential equation of the second order which is

$$x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} + (x^2 - n^2)y = 0$$

The series solution, which is the Bessel function of the first kind, is given by

$$J_n(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(n+k)!k!} \left(\frac{x}{2}\right)^{n+2k}, \quad (2.77)$$

where n is an integer ($n = 0, 1, 2, \dots$); x is an arbitrary real number; $k!$ is defined by

$$\begin{cases} k! = k(k-1) \cdots (k-(k-2)) \cdot 1, \\ 0! = 1. \end{cases} \quad (2.78)$$

$J_n(x)$ is also defined as a generating function of $e^{\frac{x}{2}(z-\frac{1}{z})}$ that

$$e^{\frac{x}{2}(z-\frac{1}{z})} = \sum_{n=-\infty}^{\infty} J_n(x) z^n, \quad (2.79)$$

where e is the base of the natural logarithm, and z can be a complex number. By (2.79) we get the integration form of $J_n(x)$ as

$$J_n(x) = \frac{1}{2\pi} \int_0^{2\pi} e^{\sqrt{-1}(x \sin \theta - n\theta)} d\theta = \frac{1}{2\pi} \int_0^{2\pi} \cos(x \sin \theta - n\theta) d\theta. \quad (2.80)$$

The characteristics of $J_n(x)$ can be obtained from the definition which are

$$\begin{cases} J_{-n}(x) = (-1)^n J_n(x), & J_n(-x) = (-1)^n J_n(x), & J_{-n}(-x) = J_n(x), \\ J_n(x) = \frac{x}{2n} [J_{n-1}(x) + J_{n+1}(x)], \\ \frac{d}{dx} J_n(x) = \frac{1}{2} [J_{n-1}(x) - J_{n+1}(x)]. \end{cases} \quad (2.81)$$

The hypergeometric function $F(a, b, c; x)$ is also a solution of a linear ordinary differential equation of the second order which is

$$(x^2 - x)y'' + [(a + b + 1)x - c]y' + aby = 0$$

where a, b , and c are constants, and the solution F takes the form as

$$\begin{aligned} F(a, b, c; x) &= 1 + \sum_{n=1}^{\infty} \frac{a(a+1) \cdots (a+n-1) \cdot b(b+1) \cdots (b+n-1)}{n! \cdot c(c+1) \cdots (c+n-1)} x^n \\ &= 1 + \frac{a \cdot b}{1 \cdot c} x + \frac{a(a+1) \cdot b(b+1)}{1 \cdot 2 \cdot c(c+1)} x^2 + \dots \end{aligned} \quad (2.82)$$

2.3.1 Expansions of $\sin kE$ and $\cos kE$

For these expansions, we only give the results. Details of derivation are given in Refs. [1] and [2] of this chapter. Results for $k > 1$ are

$$\sin kE = \sum_{n=1}^{\infty} \frac{k}{n} [J_{n-k}(ne) + J_{n+k}(ne)] \sin nM, \quad (2.83)$$

$$\cos kE = \sum_{n=1}^{\infty} \frac{k}{n} [J_{n-k}(ne) - J_{n+k}(ne)] \cos nM. \quad (2.84)$$

For $k = 1$,

$$\sin E = \frac{2}{e} \sum_{n=1}^{\infty} \frac{1}{n} J_n(ne) \sin nM, \quad (2.85)$$

$$\begin{aligned} \cos E &= -\frac{e}{2} + \sum_{n=1}^{\infty} \frac{1}{n} [J_{n-1}(ne) - J_{n+1}(ne)] \cos nM \\ &= -\frac{e}{2} + \sum_{n=1}^{\infty} \frac{2}{n^2} \frac{d}{de} [J_n(ne)] \cos nM. \end{aligned} \quad (2.86)$$

2.3.2 Expansions of E , r/a , and a/r

From

$$E = M + e \sin E, \quad \frac{r}{a} = 1 - e \cos E, \quad \frac{a}{r} = \frac{\partial E}{\partial M},$$

we have

$$E = M + 2 \sum_{n=1}^{\infty} \frac{1}{n} J_n(ne) \sin nM, \quad (2.87)$$

$$\frac{r}{a} = 1 + \frac{e^2}{2} - 2e \sum_{n=1}^{\infty} \frac{1}{n^2} \frac{d}{de} [J_n(ne)] \cos nM, \quad (2.88)$$

$$\frac{a}{r} = 1 + 2 \sum_{n=1}^{\infty} J_n(ne) \cos nM. \quad (2.89)$$

2.3.3 Expansions of $\sin F$ and $\cos F$

By the partial derivatives in (2.61), we have

$$\frac{\partial}{\partial M} \left(\frac{r}{a} \right) = \frac{\partial}{\partial M} \left(\frac{a}{r} \right)^{-1} = \frac{e}{\sqrt{1-e^2}} \sin f$$

which gives

$$\begin{aligned} \sin f &= \frac{\sqrt{1-e^2}}{e} \frac{\partial}{\partial M} \left(\frac{r}{a} \right) \\ &= 2\sqrt{1-e^2} \sum_{n=1}^{\infty} \frac{1}{n} \frac{d}{de} [J_n(ne)] \sin nM. \end{aligned} \quad (2.90)$$

And from the orbital Eq. (2.16), we have

$$\begin{aligned} \cos f &= \frac{1}{e} \left[-1 + (1-e^2) \frac{a}{r} \right] \\ &= -e + \frac{2}{e} (1-e^2) \sum_{n=1}^{\infty} J_n(ne) \cos nM. \end{aligned} \quad (2.91)$$

2.3.4 The Expansion of F

Using the expansions of $\sin f$ and $\cos f$ we have the expansion of f up to $O(e^4)$ that

$$f = M + (2e - \frac{1}{4}e^3 + \dots)\sin M + (\frac{5}{4}e^2 - \frac{11}{24}e^4 + \dots)\sin 2M \\ + (\frac{13}{12}e^3 - \dots)\sin 3M + (\frac{103}{96}e^4 - \dots)\sin 4M + \dots \quad (2.92)$$

2.3.5 Expansions of $(\frac{r}{a})^n \cos mf$ and $(\frac{r}{a})^n \sin mf$

The expansions are for the two functions with arbitrary integers, n and m , including 0. It is rather difficult to give a series of trigonometric functions of M for these two functions especially using general expressions. The method we use here is to expand the two functions directly into the Fourier series.

The basic form of the Fourier series for a function $F(f)$ is

$$\begin{cases} F(f) = \frac{a_0}{2} + \sum_{p=1}^{\infty} (a_p \cos pM + b_p \sin pM), \\ a_p = \int_0^{2\pi} F(f) \cos pM dM, \quad b_p = \int_0^{2\pi} F(f) \sin pM dM. \end{cases} \quad (2.93)$$

The function $(\frac{r}{a})^n \cos mf$ is an even function, so $b_p = 0$, and

$$a_p = \int_0^{2\pi} \left(\frac{r}{a}\right)^n [\cos(mf - pM) + \cos(mf + pM)] dM. \quad (2.94)$$

For the second part of the integrand, we let $p = -p$, that $p = -1, -2, \dots, -\infty$, resulting

$$\left(\frac{r}{a}\right)^n \cos mf = \sum_{p=-\infty}^{\infty} X_p^{n,m}(e) \cos pM, \quad (2.95)$$

where

$$X_p^{n,m}(e) = \frac{1}{2\pi} \int_0^{2\pi} \left(\frac{r}{a}\right)^n \cos(mf - pM) dM. \quad (2.96)$$

The function $(\frac{r}{a})^n \sin mf$ is an odd function, so $a_p = 0$, and

$$b_p = \frac{1}{2\pi} \int_0^{2\pi} \left(\frac{r}{a}\right)^n [\cos(mf - pM) - \cos(mf + pM)] dM. \quad (2.97)$$

Similarly, as done to (2.94), there is

$$\left(\frac{r}{a}\right)^n \sin mf = \sum_{p=-\infty}^{\infty} X_p^{n,m}(e) \sin pM. \quad (2.98)$$

As we can see that the coefficients in the expansions of the two functions in (2.95) and (2.98) are the same, thus the two functions can be expressed by one exponential function as

$$\begin{cases} \left(\frac{r}{a}\right)^n \exp(jmf) = \sum_{p=-\infty}^{\infty} X_p^{n,m}(e) \exp(jpM), \\ X_p^{n,m}(e) = \frac{1}{2\pi} \int_0^{2\pi} \left(\frac{r}{a}\right)^n \exp[j(mf - pM)] dM, \end{cases} \quad (2.99)$$

where $j = \sqrt{-1}$ is the imaginary unit. Because

$$\int_0^{2\pi} \left(\frac{r}{a}\right)^n \sin(mf - pM) dM = 0,$$

the coefficients $X_p^{n,m}(e)$ in (2.99) are exactly the same coefficients $X_p^{n,m}(e)$ in (2.96), and are called Hanse coefficients. Hansen coefficients are functions of eccentricity e , and cannot be expressed by elementary functions, but can be expressed by Bessel functions and hypergeometric functions. Here we only give the results (detailed derivations are given in Ref. [7]) that

$$X_p^{n,m}(e) = (1 + \beta^2)^{-(n+1)} \sum_{q=-\infty}^{\infty} J_q(pe) X_{p,q}^{n,m}, \quad (2.100)$$

$$\beta = \frac{1}{e} (1 - \sqrt{1 - e^2}) = \frac{e}{1 + \sqrt{1 - e^2}}, \quad (2.101)$$

$$X_{p,q}^{n,m} = \begin{cases} (-\beta)^{(p-m)-q} \binom{n-m+1}{p-m-q} F(p-q-n-1, -m-n-1, p-m-q+1, \beta^2), & (q \leq p-m); \\ (-\beta)^{q-(p-m)} \binom{n+m+1}{q-p+m} F(q-p-n-1, m-n-1, q-p+m+1, \beta^2), & (q \geq p-m). \end{cases} \quad (2.102)$$

$$\begin{cases} \binom{n}{m} = \frac{n!}{m!(n-m)!}, \binom{-n}{m} = (-1)^m \binom{n+m-1}{m}, \\ \binom{n}{-m} = \binom{-n}{-m} = 0, \binom{n}{0} = \binom{-n}{0} = 1. \end{cases} \quad (2.103)$$

From (2.96) we have

$$X_{-p}^{n,-m} = X_p^{n,m}. \quad (2.104)$$

Also from (2.100) and with $J_q(pe) = O(e^{|q|})$ we know

$$X_p^{n,m}(e) = O(e^{|m-p|}). \quad (2.105)$$

As we can notice that to give analytical expansions for $\left(\frac{r}{a}\right)^n \sin mf$ and $\left(\frac{r}{a}\right)^n \cos mf$ is rather troublesome. In order to solve this problem, based on practical applications the author derived the expressions for $X_p^{n,m}(e)$ up to $O(e^4)$ as follows [8].

$$\left\{ \begin{array}{l} X_m^{n,m}(e) = 1 + \frac{e^2}{4}(n^2 + n - 4m^2) + \frac{e^4}{64}[n^4 - 2n^3 - (1 + 8m^2)n^2 + 2n - m^2(9 - 16m^2)] \\ X_{m+1}^{n,m}(e) = -\frac{e}{2}(n - 2m) - \frac{e^3}{16}[n^3 - (1 + 2m)n^2 - (3 + 5m + 4m^2)n + m(2 + 10m + 8m^2)] \\ X_{m-1}^{n,m}(e) = -\frac{e}{2}(n + 2m) - \frac{e^3}{16}[n^3 - (1 - 2m)n^2 - (3 - 5m + 4m^2)n - m(2 - 10m + 8m^2)] \\ X_{m+2}^{n,m}(e) = \frac{e^2}{8}[n^2 - (4m + 3)n + m(4m + 5)] + \frac{e^4}{96}[n^4 - (6 + 4m)n^3 - (1 - 3m)n^2 \\ + (22 + 47m + 48m^2 + 16m^3)n - m(22 + 64m + 60m^2 + 16m^3)] \\ X_{m-2}^{n,m}(e) = \frac{e^2}{8}[n^2 + (4m - 3)n + m(4m - 5)] + \frac{e^4}{96}[n^4 - (6 - 4m)n^3 - (1 + 3m)n^2 \\ + (22 - 47m + 48m^2 - 16m^3)n + m(22 - 64m + 60m^2 - 16m^3)] \\ X_{m+3}^{n,m}(e) = -\frac{e^3}{48}[n^3 - (9 + 6m)n^2 + (17 + 33m + 12m^2)n - m(26 + 30m + 8m^2)] \\ X_{m-3}^{n,m}(e) = -\frac{e^3}{48}[n^3 - (9 - 6m)n^2 + (17 - 33m + 12m^2)n + m(26 - 30m + 8m^2)] \\ X_{m+4}^{n,m}(e) = \frac{e^4}{384}[n^4 - (18 + 8m)n^3 + (95 + 102m + 24m^2)n^2 \\ - (142 + 330m + 192m^2 + 32m^3)n + (206m + 283m^2 + 120m^3 + 16m^4)] \\ X_{m-4}^{n,m}(e) = \frac{e^4}{384}[n^4 - (18 - 8m)n^3 + (95 - 102m + 24m^2)n^2 \\ - (142 - 330m + 192m^2 - 32m^3)n - (206m - 283m^2 + 120m^3 - 16m^4)] \end{array} \right. \quad (2.106)$$

In the above-given expansions, all the coefficients are infinite power series in eccentricity e , and are convergent only when $e < e_1 = 0.6627 \dots$. The limit e_1 is called the Laplace limit.

In some cases, the above expansions are not enough, and there are other kinds of expansions needed, which are given as follows.

2.3.6 Expansions of $\left(\frac{a}{r}\right)^p$, E , and $(F - M)$ in the Trigonometric Function of F

$$\left(\frac{a}{r}\right)^p = (1 - e^2)^{-p/2} \left[T_0(p, 0) + 2 \sum_{n=1}^{\infty} T_n(p, 0) \left(\frac{p}{n}\right) \beta^n \cos nf \right], \quad (2.107)$$

where p is an integer, positive or negative, β is defined by (2.101), and $T_n(p, q)$ is given by the hypergeometric function, [2] that

$$T_n(p, q) \equiv F\left(-p - q, p - q + 1, n + 1, -\frac{\beta^2}{1 - \beta^2}\right). \quad (2.108)$$

When $p = -1, -2$, there are

$$\begin{cases} T_n(-1, 0) = T_0(-1, 0) = 1, \\ T_n(-2, 0) = \frac{1}{n+1} \left(n + \frac{1}{\sqrt{1-e^2}}\right), \quad T_0(-2, 0) = \frac{1}{\sqrt{1-e^2}}. \end{cases} \quad (2.109)$$

Then by (2.107), we have the actual expansions

$$\left(\frac{r}{a}\right) = \sqrt{1 - e^2} \left[1 + 2 \sum_{n=1}^{\infty} (-1)^n \beta^n \cos nf \right], \quad (2.110)$$

$$\left(\frac{r}{a}\right)^2 = \sqrt{1 - e^2} \left[1 + 2 \sum_{n=1}^{\infty} (-1)^n \left(1 + n\sqrt{1 - e^2}\right) \beta^n \cos nf \right]. \quad (2.111)$$

Substituting (2.110) and (2.111) into the following equations

$$\frac{\partial E}{\partial f} = (1 - e^2)^{-1/2} \left(\frac{r}{a}\right), \quad \frac{\partial M}{\partial f} = (1 - e^2)^{-1/2} \left(\frac{r}{a}\right)^2,$$

then integrating, we have

$$E = f + 2 \sum_{n=1}^{\infty} \frac{1}{n} (-1)^n \beta^n \sin nf, \quad (2.112)$$

$$f - M = 2 \sum_{n=1}^{\infty} (-1)^{n+1} \left(\frac{1}{n} + \sqrt{1 - e^2}\right) \beta^n \sin nf. \quad (2.113)$$

2.4 Transformations from the Orbital Elements to the Position Vector and Velocity and Vice Versa

2.4.1 Calculations of the Position Vector $\vec{r}(t)$ and Velocity $\dot{\vec{r}}(t)$ from Orbital Elements $\sigma(t)$

To give the position and the velocity of a celestial body at a given time is the job of making an ephemeris. If in the given orbital elements, the time element is the true anomaly f or the eccentric anomaly E , then $\vec{r}(t)$ and $\dot{\vec{r}}(t)$ can be calculated using formulas (2.33) and (2.42) directly. In reality, the time element is often given by the mean anomaly M neither f nor E . Therefore, to calculate $\vec{r}(t)$ and $\dot{\vec{r}}(t)$ from given orbital elements we must solve Kepler's Eq. (2.24) to obtain E . The method of solving Kepler's equation is provided in Sect. 2.5.

2.4.2 Calculations of the Orbital Elements $\sigma(t)$ from $\vec{r}(t)$ and $\dot{\vec{r}}(t)$

(1) Calculations of a , e , and M

The two orbital elements, a and e , decide the size and the shape of an orbit, respectively, and M decides the position of a moving body along the orbit measured from the periapsis. Clearly, all the three elements are defined in the orbital plane. According to the vis viva formula (2.17), the expressions (2.23) and (2.33) for r , (2.42) for $\dot{\vec{r}}$, the dot product of $\vec{r}$ and $\dot{\vec{r}}$, and the solution of Kepler's equation the formulas of a , e , and M are derived as

$$\frac{1}{a} = \frac{2}{r} - \frac{v^2}{\mu}, \quad (2.114)$$

$$\begin{cases} e \cos E = 1 - \frac{r}{a}, \\ e \sin E = \frac{r \dot{r}}{\sqrt{\mu a}}, \end{cases} \quad (2.115)$$

$$M = E - e \sin E, \quad (2.116)$$

in which, r , v , and $r \dot{r}$ are given by

$$\begin{cases} r^2 = x^2 + y^2 + z^2, \\ v^2 = \dot{x}^2 + \dot{y}^2 + \dot{z}^2, \\ r \dot{r} = \vec{r} \cdot \dot{\vec{r}} = x \dot{x} + y \dot{y} + z \dot{z}. \end{cases} \quad (2.117)$$

(2) Calculations of the three orientation orbital elements, i , Ω , and ω

Previously in Sects. 2.1 and 2.2, we give the relationships of the three orbital elements with the three unit vectors, $\hat{P}$, $\hat{Q}$, $\hat{R}$, which are defined by $\vec{r}$ and $\dot{\vec{r}}$. Then from (2.33) and (2.42), we have

$$\hat{P} = \frac{\cos E}{r} \vec{r} - \sqrt{\frac{a}{\mu}} (\sin E) \dot{\vec{r}}, \quad (2.118)$$

$$\sqrt{1-e^2} \hat{Q} = \frac{\sin E}{r} \vec{r} + \sqrt{\frac{a}{\mu}} (\cos E - e) \dot{\vec{r}}, \quad (2.119)$$

and by (2.3), we have

$$\hat{R} = \frac{(\vec{r} \times \dot{\vec{r}})}{\sqrt{\mu a (1-e^2)}}. \quad (2.120)$$

The three orbital elements, i , Ω , and ω therefore can be calculated by P_z , Q_z , R_x , R_y , and R_z . Using (2.39), (2.40), and (2.4), we have

$$\left\{ \begin{array}{l} P_z = \sin i \sin \omega, \quad Q_z = \sin i \cos \omega \\ \begin{pmatrix} R_x \\ -R_y \\ R_z \end{pmatrix} = \begin{pmatrix} \sin i \sin \Omega \\ \sin i \cos \Omega \\ \cos i \end{pmatrix} \end{array} \right. , \quad (2.121)$$

which give

$$\omega = \arctan\left(\frac{P_z}{Q_z}\right), \quad (2.122)$$

$$\Omega = \arctan\left(\frac{R_x}{(-R_y)}\right), \quad (2.123)$$

$$i = \arccos R_z. \quad (2.124)$$

It should be noted that ω and Ω are like E , each of them needs two values of trigonometric functions to decide which quadrant it belongs to, whereas i can be given by $\cos i$. It is easy to understand.

2.4.3 Calculations of Orbital Elements $\sigma(t_0)$ from Two Position Vectors $\vec{r}(t_1)$ and $\vec{r}(t_2)$

There are different ways to calculate $\sigma_0 = \sigma(t_0)$ from $\vec{r}_1 = \vec{r}(t_1)$ and $\vec{r}_2 = \vec{r}(t_2)$, where $t_0 \in [t_1, t_2]$. There are many methods for this conversion, here we introduce a typical one.

First, we use $\vec{r}_1$ and $\vec{r}_2$ to calculate $\vec{r}_0$ and $\dot{\vec{r}}_0$, then to calculate σ_0 by the above-given method. Based on (2.45) we have

$$\begin{cases} \vec{r}_1 = F_1 \vec{r}_0 + G_1 \dot{\vec{r}}_0, \\ \vec{r}_2 = F_2 \vec{r}_0 + G_2 \dot{\vec{r}}_0, \end{cases} \quad (2.125)$$

where F_1, G_1, F_2 , and G_2 are given by (2.46) or (2.57) with the corresponding Δt given by

$$\Delta t_1 = t_1 - t_0, \quad \Delta t_2 = t_2 - t_0. \quad (2.126)$$

Solving (2.125) yields

$$\begin{cases} \vec{r}_0 = \frac{G_2 \vec{r}_1 - G_1 \vec{r}_2}{F_1 G_2 - F_2 G_1}, \\ \dot{\vec{r}}_0 = \frac{F_2 \vec{r}_1 - F_1 \vec{r}_2}{F_2 G_1 - F_1 G_2}. \end{cases} \quad (2.127)$$

Final solutions of $\vec{r}_0$ and $\dot{\vec{r}}_0$ are actually given by the iterative method. The initial values of F_j and G_j ($j = 1, 2, 3, \dots$) can be given by

$$\begin{cases} F = 1 - \left(\frac{1}{2}u_0\right)\Delta t^2, \\ G = \Delta t - \left(\frac{1}{6}u_0\right)\Delta t^3, \end{cases} \quad (2.128)$$

where

$$u_0 = \frac{1}{r_0^3}, \quad r_0 = \frac{1}{2}(r_1 + r_2). \quad (2.129)$$

In practice, F and G can be given by the expansions in Δt as (2.57) or by formulas (2.46)–(2.49).

After $\vec{r}_0$ and $\dot{\vec{r}}_0$ are known we then use the method described in Sect. 2.4.2 to calculate the set of orbital elements σ_0 .

2.4.4 Method to Solve Kepler's Equation

According to the definition of mean anomaly M , Kepler's Eq. (2.24) can be written as

$$E - e \sin E = M. \quad (2.130)$$

When $e = 0$, there is $E = M$ and the orbit is a circle. If $e = 1$, this equation becomes the Barker equation [9, 10]. For an elliptical orbit there is $0 < e < 1$, and the corresponding Kepler's equation is actually a transcendental equation.

There are various ways to solve this equation. When the eccentricity e is small, the eccentric anomaly E can be directly calculated from the given e and M using the expansion series (2.87). In reality, there are different eccentricities and different requirements of precision. It requires a commonly accepted and relatively simple method to solve Kepler's equation. When e is not near 1 the general iterative method and the simple Newton's method (also called the differential corrector) are the ideal approximate methods by comparison.

(1) Simple iterative method

Since $e < 1$ the following iterative method is convergent,

$$E_{k+1} = M + e \sin E_k, \quad k = 0, 1, \dots \quad (2.131)$$

(2) Newton's method

Assuming

$$f(E) = (E - e \sin E) - M, \quad (2.132)$$

based on

$$\begin{cases} f(E_{k+1}) = f(E_k) + f'(E_k)[E_{k+1} - E_k] + \dots, \\ f'(E_k) = 1 - e \cos E_k, \end{cases} \quad (2.133)$$

we have

$$E_{k+1} = E_k - \frac{f(E_k)}{f'(E_k)}, \quad k = 0, 1, \quad (2.134)$$

Both methods need an initial value of E_0 . Generally, we can choose $E_0 = M$. Detailed discussions of choosing E_0 are given in Ref. [9]. Readers are also suggested to refer to related works about the care of e near 1. It should be noted that for any approximate method, when using an iterative process there is a criterion of convergence, i.e., $\Delta E_k = E_{k+1} - E_k \leq \epsilon$ for a given value ϵ . The value of ϵ should

be selected based on not only the precision requirement but also the effective word size of a computer to prevent an infinite loop. For a relatively large e , in order to save computing time, it is better to avoid using the mean anomaly as the sixth orbital element. When we use the numerical method to solve motion equations, in which the basic variables are the orbital elements, it is possible to avoid M . This problem is discussed in Chap. 6.

2.5 Expressions and Calculations of Satellite Orbital Variables

In Chap. 1, we point out that a space coordinate system is defined based on specific requirements of satellite measurement, satellite attitude, orbital determination, and other factors. Any coordinate system is involved in the orbit of a satellite's motion. In the next chapter, we discuss the problem of how to deal with a varying satellite orbit. We come to a conclusion that at any moment a satellite's orbit can be regarded as an instantaneous ellipse (or a parabola, or a hyperbola), and the geometric relationships between the orbit and the position and velocity of a satellite abide by the law of the two-body problem. Based on this conclusion it is easy to establish the necessary transformation relationships in the expressions of the satellite's measurements, attitude, and errors of orbital determination.

2.5.1 *Two Expressions of the Longitude of Satellite's Orbital Ascending Node*

During the process of launching a satellite for necessity a transitional "Earth-fixed coordinate system" is often accepted for defining the longitude of satellite's orbital ascending node. This expression may cause misunderstanding because in a true Earth-fixed coordinate system a satellite's orbit does not appear as a simple ellipse. The so-called "Earth-fixed coordinate system" is actually a "revised" geocentric celestial coordinate system by changing the direction of the X-axis from the March equinox direction to the Greenwich meridian direction (see details in Sect. 1.3.3). From the point of view of launching a satellite from the ground, it is understandable. In this coordinate system, the longitude of a satellite's orbital ascending node is its geographic longitude denoted to Ω_G , and there is a simple relationship between Ω_G and Ω that

$$\Omega_G = \Omega - S_G, \quad (2.135)$$

where S_G is the Greenwich sidereal time as discussed in Sects. 1.3.4–1.3.5.

Note that by the conventional usage, it is easy to mix up the longitude of an observatory with the longitude of an orbital ascending node. It is obvious that the longitude of an observatory is measured from the Greenwich meridian, but the longitude of an ascending node can be the right ascension in the equatorial coordinate system or the ecliptic longitude in the ecliptic coordinate system. By understanding the different definitions of longitude, confusion can be avoided.

2.5.2 Expressions of Satellite's Position Measurements from a Ground-Based Tracking Station

The measurements of satellite position from a ground-based tracking station are often gathered by two types of mounting, the horizontal mounting and the parallactic mounting. The horizontal mounting provides measurements with respect to the horizontal coordinate system, whereas measurements by the parallactic mounting to the equatorial coordinate system (both are centered at the tracking station). Usually, the position vector is denoted by $\vec{\rho}$ in the horizontal coordinate system, and by $\vec{r}'$ in the equatorial coordinate system that

$$\vec{\rho} = \rho \begin{pmatrix} \cos h \cos A \\ -\cos h \sin A \\ \sin h \end{pmatrix}, \quad \vec{r}' = r' \begin{pmatrix} \cos \delta \cos \alpha \\ \cos \delta \sin \alpha \\ \sin \delta \end{pmatrix}, \quad (2.136)$$

where ρ and r' are the distance between a satellite and the origin of the coordinate system; A and h are the azimuth and the height angles of a satellite in the horizontal coordinate system, respectively; α and δ are the right ascension and the declination, respectively. These variables are presented in Sect. 1.2.

In dealing with the orbital determination, it is necessary to present the measurements in the geocentric celestial coordinate system. To convert $\vec{\rho}$ in the horizontal coordinate system to $\vec{r}$ in the geocentric celestial coordinate system needs a rotation and a translation, whereas to convert $\vec{r}'$ in the equatorial coordinate system to $\vec{r}$ only needs a translation, thus

$$\vec{r}' = R_z(\pi - S)R_y\left(\frac{\pi}{2} - \varphi\right)\vec{\rho}, \quad (2.137)$$

$$\vec{r} = \vec{r}' + \vec{r}_A, \quad (2.138)$$

where φ is the astronomical latitude of the tracking station, $S = \alpha + t$ is the hour angle of the March equinox, i.e., the sidereal time at the station (see Fig. 1.1), and $\vec{r}_A$ is the position vector of the station in the geocentric coordinate system.

In the process of tracking a satellite, we need to convert the predicted satellite orbital state (i.e., its position vector in the geocentric celestial equatorial coordinate system) to the predicted observable variables, usually the coordinates of (A, h) in the horizontal coordinate system, or (α, δ) in the equatorial coordinate system. The transformation relationships are the reverse of (2.137) and (2.138) that

$$\vec{r}' = \vec{r} - \vec{r}_A, \quad (2.139)$$

$$\vec{\rho} = R_y\left(\varphi - \frac{\pi}{2}\right)R_z(S - \pi)\vec{r}'. \quad (2.140)$$

The measurements from a ground-based tracking station are related to the position of the station in the Earth-fixed coordinate system. If the position vector of the tracking station in the Earth-fixed geocentric coordinate system is denoted to $\vec{R}_e(H, \lambda, \varphi)$, where H is the height, λ and φ are the longitude and latitude, respectively, then the relationship between $\vec{R}_e$ and $\vec{r}_A$ (the position vector of the station in the geocentric celestial coordinate system) given in (1.29) is

$$\vec{r}_A = (HG)^T \vec{R}_e. \quad (2.141)$$

The transformation matrix (HG) includes four rotational matrices, which are the precession matrix, the nutation matrix, Earth's rotation matrix, and Earth's polar motion matrix. By using the IAU 1980 model and the IAU 2000 model, we have

$$(HG) = (EP)(ER)(NR)(PR), \quad (2.142)$$

$$(HG) = W(t)R(t)M(t). \quad (2.143)$$

In (2.143), the matrix $M(t)$ includes two matrices, the precession matrix and the nutation matrix, details about these matrices are given in Sect. 1.3.5. The relationship between $\vec{R}_e(H, \lambda, \varphi)$ and $\vec{R}_e(X_e, Y_e, Z_e)$, i.e., the relationship between the spherical coordinates and the Cartesian coordinates is given in Sect. 1.3.3 (1.26)–(1.28).

2.5.3 Equatorial Coordinates of the Sub-Satellite Point

The sub-satellite point is the node of the line connecting Earth's center and a satellite at the surface of Earth's reference ellipsoid. The position of this point is defined in the Earth's coordinate system (i.e., the Earth-fixed coordinate system described in Chap. 1), the corresponding spherical coordinates are the geodetic longitude and latitude (λ, φ) .

In the geocentric celestial coordinate system, the position vector of a satellite can be given as $\vec{r}(x, y, z) = \vec{r}(\sigma)$, where σ is a set of six orbital elements $(a, e, i, \Omega,$

ω, M). The same position vector in the Earth-fixed coordinate system is denoted by $\vec{R}(X, Y, Z)$, their relationship is given by

$$\vec{R}(X, Y, Z) = (HG)\vec{r}(x, y, z), \quad (2.144)$$

where the transformation matrix (HG) is the same as in (2.141).

Now we can give the coordinates of the sub-satellite point, (λ, φ') , by the following relationship

$$\vec{R}(t) = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} R \cos \varphi' \cos \lambda \\ R \cos \varphi' \sin \lambda \\ R \sin \varphi' \end{pmatrix}, \quad (2.145)$$

where R is the distance between the satellite and Earth's center, that

$$\begin{cases} \lambda = \arctan\left(\frac{Y}{X}\right) \\ \varphi' = \arctan\left(\frac{Z}{R}\right) = \arctan\left(\frac{Z}{\sqrt{X^2 + Y^2}}\right) \end{cases}. \quad (2.146)$$

If we need the geodetic latitude of the sub-satellite point φ , we can use formula (1.26) and let height $H = 0$, to give

$$\begin{aligned} \tan \varphi &= \frac{1}{(1-f)^2} \tan \varphi' \\ &= \frac{1}{(1-f)^2} \left(\frac{Z}{\sqrt{X^2 + Y^2}} \right). \end{aligned} \quad (2.147)$$

Note that here f is the flattening factor of Earth's reference ellipsoid.

2.5.4 Satellite's Orbital Coordinate System

For some particular measurements related to a satellite's attitude, we use the satellite's orbital coordinate system, in which the fundamental plane, i.e., the xy -plane, is the satellite's orbital plane, and the primary direction of the x -axis points to the satellite. What we need to do is to convert the geocentric celestial coordinate system to the satellite's orbital coordinate system. This can be obtained by rotating the geocentric celestial coordinate system three times. We denote the rotation matrix to (GD) that

$$(GD) = R_z(u)R_x(i)R_z(\Omega), \quad (2.148)$$

where $u = f + \omega$ is the latitude angle of the satellite, f is its true anomaly. We then translate the origin from the center of Earth to the center of the satellite etc. which depends on the actual requirement of a specific aerospace project.

2.5.5 Expressions of Errors in Satellite Position

Both orbital determination and orbital forecasting deal with how to express errors in satellite position. The basic expression of an error is in the form of $\Delta \vec{r} = (\Delta x, \Delta y, \Delta z)^T$, but more directly, an error is expressed as $\Delta \vec{r} = (\Delta r, \Delta t, \Delta w)^T$ by the radial, transverse, and normal components, defined as

$$\Delta r = \Delta \vec{r} \cdot \hat{r}, \quad \Delta t = \Delta \vec{r} \cdot \hat{t}, \quad \Delta w = \Delta \vec{r} \cdot \hat{w}. \quad (2.149)$$

Unit vectors in the three directions are given by

$$\begin{cases} \hat{r} = \cos u \hat{P} + \sin u \hat{Q}, \\ \hat{t} = -\sin u \hat{P} + \cos u \hat{Q}, \\ \hat{w} = \hat{r} \times \hat{t}, \end{cases} \quad (2.150)$$

where $u = f + \omega$ is the latitude angle of the satellite; $\hat{P}$ and $\hat{Q}$ are the unit vectors in the periapsis direction and the direction of the semi-latus rectum given in (2.39) and (2.40), respectively.

If the eccentricity of an orbit is not large, then the transverse direction and the tangential direction are close, we can assume that an error in the transverse direction, Δt , represents the orbital tracking error, which is the most important component of an error in satellite's position measurements.

2.6 Parabolic Orbit and Hyperbolic Orbit

Although the most important content of orbital dynamics is about the elliptical orbit and its variation, in reality, there are cases about motions of the natural body or artificial body (particularly the deep-space prober) involving the parabolic or the hyperbolic motions, especially the hyperbolic motion. Therefore, from a practical perspective, it is necessary to briefly introduce these two types of orbit.

2.6.1 The Parabolic Orbit

For a parabolic orbit, we have $e = 1$ and $a \rightarrow \infty$, then the areal constant of integration (2.15) and the orbital integration (2.13) become

$$r^2 \dot{\theta} = \sqrt{\mu p}, \quad p = 2q, \quad (2.151)$$

$$r = \frac{p}{1 + \cos(\theta - \omega)}. \quad (2.152)$$

The focus of the parabola is at the center of the prime body, ω is the angular distance of the periapsis, p is the semi-latus rectum, q is the distance of the periapsis. The true anomaly f now is given by

$$f = \theta - \omega, \quad (2.153)$$

then (2.151) and (2.152) become

$$r^2 \dot{f} = \sqrt{2\mu q}, \quad (2.154)$$

$$r = \frac{p}{1 + \cos f} = q \sec^2 \frac{f}{2}. \quad (2.155)$$

Substituting (2.155) into (2.154) then by integration we have

$$2 \tan \frac{f}{2} + \frac{3}{2} \tan^3 \frac{f}{2} = \sqrt{2\mu q}^{\frac{3}{2}} (t - \tau), \quad (2.156)$$

where τ is the last constant of integration, similar to the elliptical motion it is also the moment when the moving body passes the periapsis. As we can see that for a parabolic orbit because $e = 1$, there are only five orbital elements, i.e., i , Ω , q , ω , and τ .

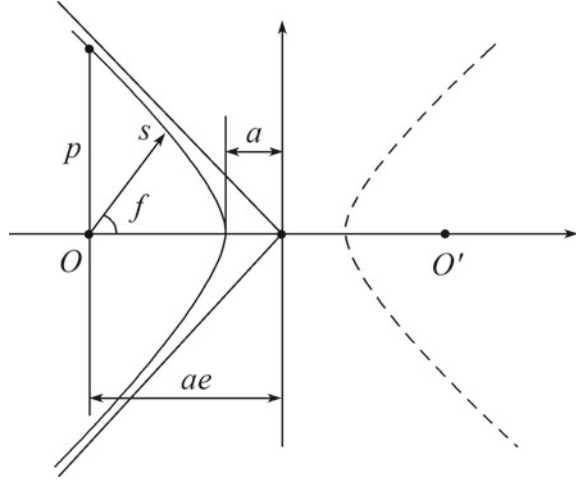
2.6.2 The Hyperbolic Orbit

For a hyperbolic orbit we have $e > 1$, then (2.15) and (2.13) become

$$r^2 \dot{\theta} = \sqrt{\mu p} \quad (2.157)$$

$$r = \frac{p}{1 + e \cos f}, \quad (2.158)$$

Fig. 2.4 A hyperbolic orbit of a small celestial body s with respect to a big body P_0 located at O (i.e., the focus)



and

$$p = a(e^2 - 1) \quad (2.159)$$

$$f = \theta - \omega. \quad (2.160)$$

where p is the semi-latus rectum, the geometric meanings of p and a are shown in Fig. 2.4; f is the true anomaly; ω is the angular distance of the periapsis; the distance of the periapsis is

$$r_p = a(e - 1). \quad (2.161)$$

The vis viva formula (2.17) for a hyperbola takes the form as

$$v^2 = \dot{r}^2 + r^2\dot{\theta}^2 = \mu\left(\frac{2}{r} + \frac{1}{a}\right). \quad (2.162)$$

Similar to the elliptical motion from (2.162) with (2.157), we eliminate $\dot{\theta}$ to give

$$\begin{cases} nadt = \frac{rdr}{\sqrt{(r+a)^2 - a^2e^2}} \\ n = \sqrt{\mu}a^{-3/2} \end{cases} \quad (2.163)$$

We then introduce an auxiliary variable E that

$$r = a(e \cosh E - 1). \quad (2.164)$$

Substituting (2.164) into (2.163), then by integration, we have the sixth constant of integration

$$e \sinh E - E = n(t - \tau) = M, \quad (2.165)$$

where τ is the sixth integral, which is also the moment when the moving body passes the periapsis. Although in this case, E has a different meaning from the eccentric anomaly E in the elliptical motion, the geometric relationships between f , E , and M are similar to those in the elliptical motion, that

$$\begin{cases} r \cos f = a(e - \cosh E), \\ r \sin f = a\sqrt{e^2 - 1} \sinh E, \end{cases} \quad (2.166)$$

$$\tan \frac{f}{2} = \sqrt{\frac{e+1}{e-1}} \tanh \frac{E}{2}. \quad (2.167)$$

The orbital Eq. (2.158) shows that when $r \rightarrow \infty$ there is $1 + e \cos f = 0$, therefore

$$-\pi + \arccos\left(\frac{1}{e}\right) \leq f \leq \pi - \arccos\left(\frac{1}{e}\right). \quad (2.168)$$

Equation (2.165) is similar to Kepler's equation in the elliptical motion, but because $e > 1$ we cannot get the solution by the general iterative method. The solution of E can be derived from given e and M using simple Newton's iterative method. We may give E an initial value of $E^{(0)}$, then the correction formula is

$$\begin{cases} \Delta E = \frac{M - (e \sinh E^{(0)} - E^{(0)})}{e \cosh E^{(0)} - 1}, \\ E^{(1)} = E^{(0)} + \Delta E - \frac{1}{2} \frac{e \sinh E^{(0)}}{(e \cosh E^{(0)} - 1)} \Delta E^2. \end{cases} \quad (2.169)$$

Generally, the iteration by (2.169) only needs to do once to give

$$E^{(1)} = E^{(0)} + \Delta E. \quad (2.170)$$

For example

Find E for given $e = 1.5$ and $M = \pi/4 = 0.785398163$. Starting with $E^{(0)} = M$, the process is as follows

$$\begin{aligned} E^{(1)} &= 1.056738913, \\ E^{(2)} &= 1.018032116, \\ E^{(3)} &= 1.016994172, \\ E^{(4)} &= 1.016993449. \end{aligned}$$

$E^{(4)}$ gives $e \sinh E - E = 0.785398163$, which agrees with the value of M by nine significant figures. This is just a simple example, of course we can use computers and other fast iterative methods to solve this problem.

2.6.3 Formulas for Calculating the Position Vector and Velocity

The position vector of a moving body in a parabolic orbit or a hyperbolic orbit, $\vec{r}$, is given by the same formula as in an elliptical orbit which is

$$\vec{r} = r \cos f \hat{P} + r \sin f \hat{Q}, \quad (2.171)$$

where $\hat{P}$ and $\hat{Q}$ are unit vectors of the periapsis direction and the semi-latus rectum direction, their expressions are similar to these for an elliptical orbit, given in (2.39) and (2.40), respectively.

The expressions of $\vec{r}$ for the parabolic orbit and the hyperbolic orbit are slightly different as given in the following formulas, (2.172) and (2.173), respectively

$$\begin{cases} \dot{\vec{r}} = \sqrt{\frac{\mu}{p}} [(-\sin f) \hat{P} + (\cos f + 1) \hat{Q}], \\ p = 2q; \end{cases} \quad (2.172)$$

$$\begin{cases} \dot{\vec{r}} = \sqrt{\frac{\mu}{p}} [(-\sin f) \hat{P} + (\cos f + e) \hat{Q}], \\ p = a(e^2 - 1). \end{cases} \quad (2.173)$$

References

1. Smart WM (1953) Celestial mechanics. University of Glasgow
2. Brouwer D, Clemence GM (1961) Methods of celestial mechanics. Academic Press, New York and London
3. Yi ZH (1994) Fundamentals of celestial mechanics, Nanjing University Press
4. Liu L (1998) Methods of celestial mechanics. Nanjing University Press
5. Beutler G (2005) Methods of celestial mechanics, Springer-Verlag, Berlin, Heideberg
6. Liu L, Hou XY (2018) The basic of orbital theory. Higher Education Press, Beijing
7. Giacaglia GE (1976) A note on Hansen's coefficients in satellite theory. Celest Mech 14:515–523
8. Liu L, Hu SJ, Wang X (2006) Introduction of aerospace dynamics. Nanjing University Press
9. Taff LG, Brennan TA (1989) On solving Kepler's equation. Celest Mech Kluwer 46:163–176
10. Battin RH (1999) An introduction to the mathematics and methods of astrodynamics. AIAA Inc Reston, Virginia

Chapter 3

Analytical Methods of Constructing Solution of Perturbed Satellite Orbit



To solve a dynamical problem, a mathematical model is required. For the orbital dynamics of satellites or spacecraft in the Solar System, the mathematical model in most cases can be summarized by a perturbed two-body problem as mentioned in the Introduction. The motion equation of a perturbed two-body problem is given by

$$\ddot{\vec{r}} = \vec{F}_0(r) + \vec{F}_\varepsilon(\vec{r}, \dot{\vec{r}}, t; \varepsilon), \quad (3.1)$$

$$\vec{F}_0(\vec{r}) = -\frac{\mu}{r^2} \left(\frac{\vec{r}}{r} \right), \quad (3.2)$$

where $\mu = G(m_0 + m)$, m_0 and m are the masses of a central celestial body and a moving body, respectively, and for a small body like a spacecraft its mass can be ignored, i.e., $m = 0$; $\vec{F}_\varepsilon$ is the perturbing acceleration, depending on a particular perturbing source, and ε is a small parameter representing the relative quantity of the perturbing force comparing to the central force. In this chapter, we discuss the method and specific treatment of solving the satellite motion based on Eq. (3.1).

3.1 The Method of the Variation of Arbitrary Constants Applied to the Perturbed Two-Body Problem

We start with the two-body problem (i.e., without perturbation), that there is $\vec{F}_\varepsilon = 0$, and the corresponding motion equation is

$$\ddot{\vec{r}} = -\frac{\mu}{r^2} \left(\frac{\vec{r}}{r} \right). \quad (3.3)$$

The solution of this equation is given in Chap. 2, and can be presented as

$$\vec{r} = \vec{f}(c_1, c_2, \dots, c_6, t), \quad (3.4)$$

$$\dot{\vec{r}} = \vec{g}(c_1, c_2, \dots, c_6, t), \quad (3.5)$$

where

$$\dot{\vec{r}} = \frac{\partial \vec{r}}{\partial t} = \frac{\partial \vec{f}}{\partial t}. \quad (3.6)$$

The six integrals, $c_1, c_2, \dots$ and c_6 , are the six orbital elements a, e, i, Ω, ω , and τ for an elliptic orbital motion. The sixth orbital element τ is the moment when the moving body passes the periapsis. If we use the mean anomaly M to replace τ , then the time variable t is not explicitly included in the solution of (3.4) and (3.5). In practice, the solution does not depend on M directly but on the true anomaly f or the eccentric anomaly E .

For the perturbed Eq. (3.1), since $\vec{F}_\varepsilon \neq 0$, naturally the solution of (3.4) and (3.5) no longer satisfies the equation. To make the solution still valid to Eq. (3.1) we assume that the six orbital elements, $c_1, c_2, \dots$ and c_6 , are not constants but functions of time t . Based on this principle we derive a system of differential equations for the original orbital constants. This method is called the method of the variation of arbitrary constants for solving ordinary differential equations. The new system of differential equations of c_j ($j = 1, 2, \dots, 6$) is called the “perturbed motion equation system”. Although the purpose of this monograph is to provide algorithms of solving practical orbital problems, for a better understanding of the algorithms, it is necessary to introduce this most important method in the field of orbital dynamics.

First, we differentiate (3.4) at t to give

$$\frac{d\vec{r}}{dt} = \frac{\partial \vec{f}}{\partial t} + \sum_{j=1}^6 \frac{\partial \vec{f}}{\partial c_j} \frac{dc_j}{dt}. \quad (3.7)$$

Because the solution (3.5) should also satisfy the perturbed equation we have

$$\frac{d\vec{r}}{dt} = \frac{\partial \vec{f}}{\partial t} = \vec{g}(c_1, \dots, c_6; t). \quad (3.8)$$

We then differentiate (3.8) at t and the derivative should agree with the perturbed Eq. (3.1) that

$$\frac{d^2\vec{r}}{dt^2} = \frac{\partial \vec{g}}{\partial t} + \sum_{j=1}^6 \frac{\partial \vec{g}}{\partial c_j} \frac{dc_j}{dt} = \vec{F}_0 + \vec{F}_\varepsilon. \quad (3.9)$$

Since $\partial \vec{g}/\partial t = \vec{F}_0$ for satisfying the perturbed equation the two conditions of the method of the variation of arbitrary constants must be

$$\begin{cases} \sum_{j=1}^6 \frac{\partial \vec{f}}{\partial c_j} \frac{dc_j}{dt} = 0 \\ \sum_{j=1}^6 \frac{\partial \vec{g}}{\partial c_j} \frac{dc_j}{dt} = \vec{F}_\varepsilon \end{cases} \quad (3.10)$$

The equation in (3.10) form a system of algebraic equations of six variables $\frac{dc_j}{dt}$. The coefficients $\frac{\partial \vec{f}}{\partial c_j}$ and $\frac{\partial \vec{g}}{\partial c_j}$ are functions of c_j and t , and are given in Chap. 2. In principle, the explicit forms of the six variables $\frac{dc_j}{dt}$ can be given by solving Eq. (3.10) that

$$\frac{dc_j}{dt} = f(c_1, \dots, c_6, t; \varepsilon), \quad j = 1, 2, \dots, 6. \quad (3.11)$$

Equation (3.11) are the needed perturbed motion equations. In reality, the coefficients in (3.10) are complicated, detailed processes of deriving them can be found in some books of Celestial Mechanics [1, 2]. There are mainly two types of deriving methods. The first type of the method is applied to conservative forces using the partial derivatives of the disturbing function R , $(\partial R/\partial \sigma)$, to replace the perturbing acceleration $\vec{F}_\varepsilon$. The second type is directly using the three components of the acceleration $\vec{F}_\varepsilon$ (such as the components in the radial, transversal, and normal directions). The author gives a simple method of the second type to derive the perturbed equations, the details are given in references [3] and [4]. The results are presented in the following section.

Before providing the perturbed motion equations, there is an important concept of the method of the variation of arbitrary constants needed to be emphasized. For a perturbed motion, the solution for unperturbed motion equations in the form of (3.4) and (3.5) (i.e., (2.33) and (2.42)) is still applicable, because this solution gives a rigorous relationship between the instantaneous position vector and instantaneous velocity for a perturbed motion. The difference is for an unperturbed motion c_j ($j = 1, 2, \dots, 6$) are constants, whereas for a perturbed motion $c_j = c_j(t)$ are functions of time t . In other words, a perturbed orbit can be regarded as a varying ellipse (or a conic curve), therefore all geometric relationships of an elliptic motion and partial derivatives described in Chap. 2 are all applicable. But note that the derivatives at t are no longer valid, especially the areal integral that for a perturbed motion it should be

$$\begin{cases} r^2 \dot{\theta} = \sqrt{\mu a(1 - e^2)}, \\ \dot{\theta} = \dot{f} + \dot{\omega} + \dot{\Omega} \cos i. \end{cases} \quad (3.12)$$

Only for an unperturbed motion, there are $\dot{\omega} = 0$, $\dot{\Omega} = 0$, and the areal integral is back to that for a two-body problem as

$$r^2 \dot{f} = h = \sqrt{\mu a(1 - e^2)}.$$

3.2 Common Forms of Perturbed Motion Equation

3.2.1 *Perturbed Motion Equations Formed by Accelerations of the (S, T, W)-Version and the (U, N, W)-Version*

The perturbation forces are mostly conservative forces, but some are not, in either case, we can establish the perturbation motion equations by the components of perturbation accelerations. We usually decompose the perturbation acceleration $\vec{F}_\varepsilon$ in the radial direction, the transversal direction, and the normal direction (normal to the orbital plane), which are denoted to S , T , and W , respectively; or in the tangential direction, the prime normal direction, and the second normal direction (i.e., normal to the orbital plane), which are denoted to U , N , and W , respectively. The converting relationships between (S, T) and (U, N) are given as

$$\begin{cases} S = \frac{e \sin f}{\sqrt{1 + 2e \cos f + e^2}} U - \frac{1 + e \cos f}{\sqrt{1 + 2e \cos f + e^2}} N \\ T = \frac{1 + e \cos f}{\sqrt{1 + 2e \cos f + e^2}} U + \frac{e \sin f}{\sqrt{1 + 2e \cos f + e^2}} N \end{cases}. \quad (3.13)$$

By (3.13), it is easy to transform the motion equations from the (S, T, W) -version to the (U, N, W) -version.

The perturbation motion equations formed directly by the perturbation accelerations are called Gauss perturbation motion equations. The actual expressions are as follows.

(1) The (S, T, W) -version of the perturbation equations

$$\left\{ \begin{array}{l} \frac{da}{dt} = \frac{2}{n\sqrt{1-e^2}} [Se \sin f + T(1 + e \cos f)] \\ \frac{de}{dt} = \frac{\sqrt{1-e^2}}{na} [S \sin f + T(\cos f + \cos E)] \\ \frac{di}{dt} = \frac{r \cos u}{na^2 \sqrt{1-e^2}} W \\ \frac{d\Omega}{dt} = \frac{r \sin u}{na^2 \sqrt{1-e^2} \sin i} W \\ \frac{d\omega}{dt} = \frac{\sqrt{1-e^2}}{nae} \left[-S \cos f + T \left(1 + \frac{r}{p} \right) \sin f \right] - \cos i \frac{d\Omega}{dt} \\ \frac{dM}{dt} = n - \frac{1-e^2}{nae} \left[-S \left(\cos f - 2e \frac{r}{p} \right) + T \left(1 + \frac{r}{p} \right) \sin f \right] \end{array} \right. \quad (3.14)$$

where $u = f + \omega$, $p = a(1 - e^2)$, f and E are the true anomaly and the eccentric anomaly, respectively.

(2) The (U, N, W) -version of the perturbation equations

$$\left\{ \begin{array}{l} \frac{da}{dt} = \frac{2}{n\sqrt{1-e^2}} (1 + 2e \cos f + e^2)^{1/2} U \\ \frac{de}{dt} = \frac{\sqrt{1-e^2}}{na} (1 + 2e \cos f + e^2)^{-1/2} \left[2(\cos f + e)U - (\sqrt{1-e^2} \sin E)N \right] \\ \frac{d\omega}{dt} = \frac{\sqrt{1-e^2}}{nae} (1 + 2e \cos f + e^2)^{-1/2} [(2 \sin f)U + (\cos E + e)N] - \cos i \frac{d\Omega}{dt} \\ \frac{dM}{dt} = n - \frac{1-e^2}{nae} (1 + 2e \cos f + e^2)^{-1/2} \left[\left(2 \sin f + \frac{2e^2}{\sqrt{1-e^2}} \sin E \right) U \right. \\ \left. + (\cos E - e)N \right] \end{array} \right. \quad (3.15)$$

The expressions of $\frac{di}{dt}$ and $\frac{d\Omega}{dt}$ are the same as given in (3.14).

The three components of the perturbation acceleration S , T , and W , depending on the perturbing sources. If we know the source is a conservative force and the disturbing function R , then the three components, S , T , and W , are given by

$$S = \frac{\partial R}{\partial r}, \quad T = \frac{1}{r} \frac{\partial R}{\partial \theta} = \frac{1}{r} \frac{\partial R}{\partial u}, \quad W = \frac{\partial R}{\partial z} = \frac{1}{r \sin u} \frac{\partial R}{\partial i}. \quad (3.16)$$

If we know the three components of the perturbation acceleration $\vec{F}_\varepsilon$ in the rectangular coordinate system, $\vec{F}_{1x}$, $\vec{F}_{1y}$, $\vec{F}_{1z}$, then we can obtain S , T , and W by

$$\begin{pmatrix} S \\ T \\ W \end{pmatrix} = (ZH) \begin{pmatrix} \vec{F}_{1x} \\ \vec{F}_{1y} \\ \vec{F}_{1z} \end{pmatrix}, \quad (3.17)$$

where the transformation matrix (ZH) is given by three rotation matrices, that $(ZH) = R_z(u)R_x(i)R_z(\Omega)$ which is

$$\begin{aligned} (ZH) &= R_z(u)R_x(i)R_z(\Omega) = \begin{pmatrix} l_1 m_1 n_1 \\ l_2 m_2 n_2 \\ l_3 m_3 n_3 \end{pmatrix} \\ &= \begin{pmatrix} \cos u \cos \Omega + \sin u(-\sin \Omega \cos i) & \cos u \sin \Omega + \sin u(\cos \Omega \cos i) & \sin u \sin i \\ -\sin u \cos \Omega + \cos u(-\sin \Omega \cos i) & -\sin u \sin \Omega + \cos u(\cos \Omega \cos i) & \cos u \sin i \\ \sin \Omega \sin i & -\cos \Omega \sin i & \cos i \end{pmatrix}. \end{aligned} \quad (3.18)$$

Now the transformation formulas from $\vec{F}_\varepsilon$ to (S, T, W) can be written as

$$S = \vec{F}_\varepsilon \cdot \hat{r}, \quad T = \vec{F}_\varepsilon \cdot \hat{t}, \quad W = \vec{F}_\varepsilon \cdot \hat{w}, \quad (3.19)$$

where $\hat{r}$, $\hat{t}$, $\hat{w}$ are vector units in the radial direction, the transversal, direction, and the normal direction to the orbital plane, that

$$\hat{r} = \begin{pmatrix} \cos u(\cos \Omega) + \sin u(-\cos i \sin \Omega) \\ \cos u(\sin \Omega) + \sin u(\cos i \cos \Omega) \\ \sin u(\sin i) \end{pmatrix} = \cos u \hat{P}^* + \sin u \hat{Q}^*, \quad (3.20)$$

$$\hat{t} = -\sin u \hat{P}^* + \cos u \hat{Q}^*, \quad (3.21)$$

$$\hat{w} = \begin{pmatrix} \sin i \sin \Omega \\ -\sin i \cos \Omega \\ \cos i \end{pmatrix} = \hat{P}^* \times \hat{Q}^*. \quad (3.22)$$

Note that here $\hat{P}^*$ and $\hat{Q}^*$ have different directions from $\hat{P}$ and $\hat{Q}$ in Chap. 2 (see (2.39) and (2.40)), that

$$\hat{P}^* = \begin{pmatrix} \cos \Omega \\ \sin \Omega \\ 0 \end{pmatrix}, \quad \hat{Q}^* = \begin{pmatrix} -\cos i \sin \Omega \\ \cos i \cos \Omega \\ \sin i \end{pmatrix}. \quad (3.23)$$

3.2.2 The Perturbation Motion Equations Formed by $\partial R/\partial\sigma$ -Version

When the perturbing force is conservative the perturbation acceleration $\vec{F}_\varepsilon$ can be given by

$$\vec{F}_\varepsilon = \text{grad}(R), \quad (3.24)$$

where R is the perturbation function, generally $R = R(\vec{r}, t; \varepsilon)$, the actual expression depends on the perturbing source.

The $\partial R/\partial\sigma$ -version of the perturbation motion equations can be given by transforming the (S, T, W) -version of the perturbation equations. Omitting the process, we give the results as

$$\left\{ \begin{array}{l} \frac{da}{dt} = \frac{2}{na} \frac{\partial R}{\partial M} \\ \frac{de}{dt} = \frac{1-e^2}{na^2e} \frac{\partial R}{\partial M} - \frac{\sqrt{1-e^2}}{na^2e} \frac{\partial R}{\partial \omega} \\ \frac{di}{dt} = \frac{1}{na^2\sqrt{1-e^2}\sin i} \left(\cos i \frac{\partial R}{\partial \omega} - \frac{\partial R}{\partial \Omega} \right) \\ \frac{d\Omega}{dt} = \frac{1}{na^2\sqrt{1-e^2}\sin i} \frac{\partial R}{\partial i} \\ \frac{d\omega}{dt} = \frac{\sqrt{1-e^2}}{na^2e} \frac{\partial R}{\partial e} - \cos i \frac{d\Omega}{dt} \\ \frac{dM}{dt} = n - \frac{1-e^2}{na^2e} \frac{\partial R}{\partial e} - \frac{2}{na} \frac{\partial R}{\partial a} \end{array} \right. \quad (3.25)$$

For a given perturbing function R directly using the method of the variation of arbitrary constants can derive (3.25), which are called the Lagrange perturbation equations. But direct derivations are often complicated, from a practical point of view, it is unnecessary to further discuss the process in this book.

The system of perturbation Eq. (3.25) has an obvious characteristic. The right sides of the first three equations involve only $\partial R/\partial(\Omega, \omega, M)$, whereas the other three equations involve only $\partial R/\partial(a, e, i)$, showing a kind of “symmetry”. This characteristic is related to the difference between the three angular variables (Ω, ω , and M) and the three angular momentum variables (a, e , and i). Among the three angular variables, the mean anomaly is a fast-varying variable, its varying speed depends on the mean angular speed n given by $n = \sqrt{\mu}a^{-3/2}$.

3.2.3 Canonical Equations of Perturbation Motion

For a Hamiltonian system, it is easy to establish the corresponding perturbation equations by the analytical dynamic method. For canonical conjugate variables, such as Delaunay variables, L , G , H , l , g , and h , the corresponding perturbation equations have a simple conjugate symmetric form that

$$\begin{cases} \frac{dL}{dt} = \frac{\partial F}{\partial l}, & \frac{dl}{dt} = -\frac{\partial F}{\partial L}, \\ \frac{dG}{dt} = \frac{\partial F}{\partial g}, & \frac{dg}{dt} = -\frac{\partial F}{\partial G}, \\ \frac{dH}{dt} = \frac{\partial F}{\partial h}, & \frac{dh}{dt} = -\frac{\partial F}{\partial H}, \end{cases} \quad (3.26)$$

where F is the Hamiltonian function, which is different from the commonly used Hamiltonian function K (here we use K to separate it from the variable H) by a factor of -1 , that

$$F = -K = \frac{\mu^2}{2L^2} + R. \quad (3.27)$$

Thus, the system of Eq. (3.26) is also different from the commonly used formulas by a factor of -1 . The variables, L , G , and H , are momentums (angular momentums), which are equivalent to the generalized momentum p ; the variables l , g , and h are angular variables, equivalent to the generalized coordinate q . The relationships between these variables and the elliptical orbital elements are

$$\begin{cases} L = \sqrt{\mu a}, & l = M \\ G = \sqrt{\mu a(1 - e^2)}, & g = \omega \\ H = \sqrt{\mu a(1 - e^2)} \cos i, & h = \Omega \end{cases} \quad (3.28)$$

From the system of Eq. (3.26) and the relationships (3.28) we can derive the system of Lagrange perturbation motion Eq. (3.25). This connection shows that although Hamiltonian dynamics mainly belongs to the theoretical research field, in order to solve practical problems, it is necessary to know the principle concepts and the relationships between the commonly used variables.

3.2.4 Singularities in the Perturbation Equations

In the systems of perturbation motion Eqs. (3.14), (3.15), or (3.25) we can see that the right sides of $\frac{d\omega}{dt}$ and $\frac{dM}{dt}$ have a factor of $\frac{1}{e}$, and the right sides of $\frac{d\Omega}{dt}$ and $\frac{di}{dt}$ have a factor of $\frac{1}{\sin i}$, showing that $e = 0$ and $\sin i = 0$ (i.e., $i = 0$ or 180°) are the singularities of the perturbation motion equations. In the next chapter we give the

perturbation motion solutions, and when $e \approx 0$ and $i \approx 0$ or 180° , the solutions are invalid. But the corresponding motion, such as a circular motion of $e = 0$, actually exists. Therefore, this kind of singularity is not essential, their appearances are due to the improper variables used in the equations. When $e = 0$, ω cannot be defined, nor M as M is related to ω ; when $i \approx 0$ or 180° , Ω cannot be defined nor ω , which is related to Ω . The inappropriately chosen variables must cause problems, but the singularities can be eliminated by choosing proper variables.

(1) Perturbation equations for any eccentricity ($0 \leq e < 1$)

To eliminate the singularity $e = 0$, we introduce the following set of variables

$$a, i, \Omega, \xi = e \cos \omega, \quad \eta = e \sin \omega, \quad \lambda = M + \omega. \quad (3.29)$$

Using this set of variables, the motion solution has no singularity when $e = 0$, and ξ, η , and λ are all meaningful.

By the definitions in (3.29) and the following relationships

$$\begin{cases} \frac{d\xi}{dt} = \cos \omega \frac{de}{dt} - e \sin \omega \frac{d\omega}{dt} \\ \frac{d\eta}{dt} = \sin \omega \frac{de}{dt} + e \cos \omega \frac{d\omega}{dt} \\ \frac{d\lambda}{dt} = \frac{dM}{dt} + \frac{d\omega}{dt} \end{cases} \quad (3.30)$$

we can derive the non-singularity perturbation equations by the new variables as follows.

① The $\partial R/\partial \sigma$ -version perturbation equations

$$\begin{cases} \frac{da}{dt} = \frac{2}{na} \frac{\partial R}{\partial \lambda} \\ \frac{di}{dt} = \frac{1}{na^2 \sqrt{1-e^2} \sin i} \left(\cos i \left(\xi \frac{\partial R}{\partial \eta} - \eta \frac{\partial R}{\partial \xi} + \frac{\partial R}{\partial \lambda} \right) - \frac{\partial R}{\partial \Omega} \right) \\ \frac{d\Omega}{dt} = \frac{1}{na^2 \sqrt{1-e^2} \sin i} \frac{\partial R}{\partial i} \\ \frac{d\xi}{dt} = -\frac{\sqrt{1-e^2}}{na^2} \frac{\partial R}{\partial \eta} - \xi \frac{\sqrt{1-e^2}}{na^2 (1 + \sqrt{1-e^2})} \frac{\partial R}{\partial \lambda} + \eta \cos i \frac{d\Omega}{dt} \\ \frac{d\eta}{dt} = \frac{\sqrt{1-e^2}}{na^2} \frac{\partial R}{\partial \xi} - \eta \frac{\sqrt{1-e^2}}{na^2 (1 + \sqrt{1-e^2})} \frac{\partial R}{\partial \lambda} - \xi \cos i \frac{d\Omega}{dt} \\ \frac{d\lambda}{dt} = n - \frac{2}{na} \frac{\partial R}{\partial a} + \frac{\sqrt{1-e^2}}{na^2 (1 + \sqrt{1-e^2})} \left(\xi \frac{\partial R}{\partial \xi} + \eta \frac{\partial R}{\partial \eta} \right) - \cos i \frac{d\Omega}{dt} \end{cases} \quad (3.31)$$

② The (S, T, W) -version perturbation equations

$$\left\{ \begin{aligned} \frac{da}{dt} &= \frac{2}{n\sqrt{1-e^2}} \left[S \left(\xi \sin u - \eta \cos u \right) + T \left(\frac{p}{r} \right) \right] \\ \frac{d\xi}{dt} &= \frac{\sqrt{1-e^2}}{na} \left\{ S \sin u + \frac{T}{\sqrt{1-e^2}} \left[\left(\cos \tilde{u} + \sqrt{1-e^2} \cos u \right) \right. \right. \\ &\quad \left. \left. - \frac{\xi}{1+\sqrt{1-e^2}} \left(\xi \cos \tilde{u} + \eta \sin \tilde{u} \right) \right] \right\} \\ &\quad + \eta \cos i \frac{d\Omega}{dt} \\ \frac{d\eta}{dt} &= -\frac{\sqrt{1-e^2}}{na} \left\{ S \cos u - \frac{T}{\sqrt{1-e^2}} \left[\left(\sin \tilde{u} + \sqrt{1-e^2} \sin u \right) \right. \right. \\ &\quad \left. \left. - \frac{\eta}{1+\sqrt{1-e^2}} \left(\xi \cos \tilde{u} + \eta \sin \tilde{u} \right) \right] \right\} \\ &\quad - \xi \cos i \frac{d\Omega}{dt} \\ \frac{d\lambda}{dt} &= n - \frac{\sqrt{1-e^2}}{na} \left\{ 2S\sqrt{1-e^2} \left(\frac{r}{p} \right) + \frac{1}{1+\sqrt{1-e^2}} \left[S \left(\xi \cos u + \eta \sin u \right) \right. \right. \\ &\quad \left. \left. - T \left(1 + \frac{r}{p} \right) \left(\xi \sin u - \eta \cos u \right) \right] \right\} - \cos i \frac{d\Omega}{dt} \end{aligned} \right. \quad (3.32)$$

where $\tilde{u} = E + \omega$, di/dt and $d\Omega/dt$ are the same as in (3.14).

In the above described converting we use the relationship

$$1 - \sqrt{1-e^2} = e^2 / (1 + \sqrt{1-e^2}) \quad (3.33)$$

to eliminate the factor $1/e$, which becomes $1/\sqrt{\xi^2 + \eta^2}$.

There is one thing about the definitions of ξ and η that should be mentioned. The author wrote a book titled *The Method of Celestial Dynamics* published by Nanjing University Press in 1998 [4]. In the book, ξ and η are defined as

$$\xi = e \cos \omega, \quad \eta = -e \sin \omega. \quad (3.34)$$

Several years ago, the author and his colleagues decided to use the definitions given in (3.29), which are formally accepted in related works. If readers insist to use the original definitions, the only thing that needs to do is to replace every η by $(-\eta)$ in this book.

- (2) Perturbation equations for any eccentricity ($0 \leq e < 1$) and any inclination ($0 \leq i < 180^\circ$)

For eliminating the singularity $i = 0$, we choose a set of variables as follows

$$a, e, h = \sin \frac{i}{2} \cos \Omega, \quad k = \sin \frac{i}{2} \sin \Omega, \quad \tilde{\omega} = \omega + \Omega, \quad M. \quad (3.35)$$

This is a set of non-singularity variables at $i = 0$. Obviously when $i = 0$, $h, k, \tilde{\omega}$ are meaningful. The case of $i = 180^\circ$ usually does not exist.

It is possible that both $e = 0$ and $i = 0$ happen, then the non-singularity variables for this situation are chosen as

$$\begin{cases} a, & h = \sin \frac{i}{2} \cos \Omega, & k = \sin \frac{i}{2} \sin \Omega \\ \xi = e \cos \tilde{\omega}, & \eta = e \sin \tilde{\omega}, & \lambda = M + \tilde{\omega} \end{cases}, \quad (3.36)$$

where

$$\tilde{\omega} = \omega + \Omega. \quad (3.37)$$

The corresponding perturbation equations are

$$\begin{cases} \frac{dh}{dt} = \frac{1}{2} \cos \frac{i}{2} \cos \Omega \frac{di}{dt} - \sin \frac{i}{2} \sin \Omega \frac{d\Omega}{dt} \\ \frac{dk}{dt} = \frac{1}{2} \cos \frac{i}{2} \sin \Omega \frac{di}{dt} + \sin \frac{i}{2} \cos \Omega \frac{d\Omega}{dt} \\ \frac{d\xi}{dt} = \cos(\omega + \Omega) \frac{de}{dt} - e \sin(\omega + \Omega) \left(\frac{d\omega}{dt} + \frac{d\Omega}{dt} \right) \\ \frac{d\eta}{dt} = \sin(\omega + \Omega) \frac{de}{dt} + e \cos(\omega + \Omega) \left(\frac{d\omega}{dt} + \frac{d\Omega}{dt} \right) \\ \frac{d\lambda}{dt} = \frac{dM}{dt} + \frac{d\omega}{dt} + \frac{d\Omega}{dt} \end{cases} \quad (3.38)$$

For applications, we provide the non-singularity perturbation equations in the (S, T, W) -version as

$$\frac{da}{dt} = \frac{2}{n\sqrt{1-e^2}} \left[S(\xi \sin u - \eta \cos u) + T \left(\frac{p}{r} \right) \right], \quad (3.39)$$

$$\begin{aligned} \frac{d\xi}{dt} = \frac{\sqrt{1-e^2}}{na} \left\{ S \sin u + T \left[(\cos \tilde{u} + \cos u) - \frac{\eta}{\sqrt{1-e^2}(1+\sqrt{1-e^2})} \right. \right. \\ \left. \left. (\xi \sin \tilde{u} - \eta \cos \tilde{u}) \right] + W \left(\frac{r}{p} \right) \left[-\frac{\eta}{\cos(i/2)} (h \sin u - k \cos u) \right] \right\} \end{aligned} \quad (3.40)$$

$$\begin{aligned} \frac{d\eta}{dt} = \frac{\sqrt{1-e^2}}{na} \left\{ -S \cos u + T \left[(\sin \tilde{u} + \sin u) + \frac{\xi}{\sqrt{1-e^2}(1+\sqrt{1-e^2})} \right. \right. \\ \left. \left. (\xi \sin \tilde{u} - \eta \cos \tilde{u}) \right] + W \left(\frac{r}{p} \right) \left[\frac{\xi}{\cos(i/2)} (h \sin u - k \cos u) \right] \right\} \end{aligned} \quad (3.41)$$

$$\frac{dh}{dt} = \frac{\sqrt{1-e^2}}{2na \cos(i/2)} W\left(\frac{r}{p}\right) [\cos u - h(h \cos u + k \sin u)], \quad (3.42)$$

$$\frac{dk}{dt} = \frac{\sqrt{1-e^2}}{2na \cos(i/2)} W\left(\frac{r}{p}\right) [\sin u - k(h \cos u + k \sin u)], \quad (3.43)$$

$$\begin{aligned} \frac{d\lambda}{dt} = n - \frac{\sqrt{1-e^2}}{na} \left\{ 2S\sqrt{1-e^2} \left(\frac{r}{p}\right) + \frac{1}{1+\sqrt{1-e^2}} \left[S\left(\xi \cos u + \eta \sin u\right) \right. \right. \\ \left. \left. - T\left(1 + \frac{r}{p}\right) \left(\xi \sin u - \eta \cos u\right) \right] - W\left(\frac{r}{p}\right) \left(\frac{h \sin u - k \cos u}{\cos(i/2)}\right) \right\}. \end{aligned} \quad (3.44)$$

The intermediate variables:

$$n, e^2, p, \sin^2 i, \tilde{u} = E + \tilde{\omega}, u = f + \tilde{\omega}, \dots$$

on the right sides of the above equations are given by

$$n = \sqrt{\mu} a^{-3/2}, \quad (3.45)$$

$$e^2 = \xi^2 + \eta^2, \quad (3.46)$$

$$p = a(1 - e^2), \quad (3.47)$$

$$\sin^2 \frac{i}{2} = h^2 + k^2, \quad \cos i = 1 - 2(h^2 + k^2), \quad \cos \frac{i}{2} = [1 - (h^2 + k^2)]^{1/2}, \quad (3.48)$$

$$\tilde{u} - \lambda = \xi \sin \tilde{u} - \eta \cos \tilde{u}, \quad (3.49)$$

$$\frac{a}{r} = [1 - (\xi \cos \tilde{u} + \eta \sin \tilde{u})]^{-1}, \quad (3.50)$$

$$\begin{aligned} \sin u &= \left(\frac{a}{r}\right) \left[(\sin \tilde{u} - \eta) - \frac{\xi}{1 + \sqrt{1-e^2}} (\xi \sin \tilde{u} - \eta \cos \tilde{u}) \right] \\ \cos u &= \left(\frac{a}{r}\right) \left[(\cos \tilde{u} - \xi) + \frac{\eta}{1 + \sqrt{1-e^2}} (\xi \sin \tilde{u} - \eta \cos \tilde{u}) \right] \end{aligned} \quad (3.51)$$

The accelerations, S , T , and W , on the right sides of the motion equations can be converted from acceleration $\vec{F}_\varepsilon$, that

$$S = \vec{F}_\varepsilon \cdot \hat{r}, \quad T = \vec{F}_\varepsilon \cdot \hat{t}, \quad W = \vec{F}_\varepsilon \cdot \hat{w}, \quad (3.52)$$

$$\hat{r} = \cos u \hat{P}^* + \sin u \hat{Q}^*, \quad (3.53)$$

$$\hat{t} = -\sin u \hat{P}^* + \cos u \hat{Q}^*, \quad (3.54)$$

$$\hat{w} = \hat{r} \times \hat{t}, \quad (3.55)$$

where unit vectors $\hat{P}^*$, $\hat{Q}^*$ are given by

$$\hat{P}^* = \begin{pmatrix} 1 - 2k^2 \\ 2hk \\ -2k \cos \frac{i}{2} \end{pmatrix}, \quad (3.56)$$

$$\hat{Q}^* = \begin{pmatrix} 2hk \\ 1 - 2h^2 \\ 2h \cos \frac{i}{2} \end{pmatrix}, \quad (3.57)$$

Also, there are

$$\hat{w} = \begin{pmatrix} 2k \cos \frac{i}{2} \\ -2h \cos \frac{i}{2} \\ \cos i \end{pmatrix}, \quad (3.58)$$

$$\vec{r} = r \hat{r}, \quad (3.59)$$

$$\dot{\vec{r}} = \sqrt{\frac{\mu}{p}} [-(\sin u + \eta) \hat{P}^* + (\cos u + \xi) \hat{Q}^*]. \quad (3.60)$$

(3) Non-singularity canonical conjugate variables

The problems about a small e and a small i also appear in the canonical motion equations when the variables are not chosen properly, therefore we also need a set of non-singularity variables. Here is a set of variables for eliminating $e = 0, (\tilde{L}, \tilde{G}, \tilde{H}, \tilde{l}, \tilde{g}, \tilde{h})$, that

$$\begin{cases} \tilde{L} = L, & \tilde{l} = l + g \\ \tilde{G} = \sqrt{2(L - G)} \cos g, & \tilde{g} = \sqrt{2(L - G)} \sin g, \\ \tilde{H} = H, & \tilde{h} = h \end{cases} \quad (3.61)$$

where

$$\sqrt{2(L - G)} = \left[L\sqrt{2/(L + G)} \right] e \quad (3.62)$$

This set of variables corresponds to the set given by (3.29). The variables L , G , H , l , g , and h in (3.61) and (3.62) are the original Delaunay variables.

We are not going to give the perturbation equations in terms of the non-singularity conjugate variables. The reason is in the process of solving the perturbation equations if we use the canonical conjugate variables, the method is usually the transformation method, often for theoretical researchers, readers who are interested in this topic may refer to references [3–6].

3.3 Perturbation Method of Constructing Power Series Solution with a Small Parameter

3.3.1 Perturbation Equations with a Small Parameter

The original perturbed equation is

$$\ddot{\vec{r}} = \vec{F}_0 + \vec{F}_\varepsilon. \quad (3.63)$$

After applying the method of the variation of arbitrary constants as described in Sect. 3.1, the problem of solving Eq. (3.63) becomes the problem of solving the perturbation equation system

$$\frac{d\sigma}{dt} = f_\varepsilon(\sigma, t, \varepsilon), \quad (3.64)$$

where σ is a vector with six dimensions, and the six components can be the instantaneous orbital elements, the corresponding six canonical conjugate variables, or the non-singularity variables introduced in the previous section. The function on the right side f_ε is a function with six dimensions, the absolute value of each component has the same order of ε that

$$|(f_\varepsilon)_i| = O(\varepsilon) \ll 1, \quad i = 1, 2, \dots, 6. \quad (3.65)$$

The solution of the perturbed motion equations consists of two parts, which are

$$\vec{r} = \vec{r}(\sigma, t), \quad \dot{\vec{r}} = \dot{\vec{r}}(\sigma, t), \quad (3.66)$$

$$\sigma(t) = \sigma(\sigma_0, t_0; t, \varepsilon), \quad \sigma(t_0) = \sigma_0, \quad (3.67)$$

where the expressions of $\vec{r}$ and $\dot{\vec{r}}$ are given in (2.33) and (2.42) in Chap. 2, corresponding to a motion on an instantaneous ellipse; and σ_0 is a set of initial values of

the orbital elements at time t_0 . The remaining problem is how to solve the equations of a small parameter (3.64) to obtain the perturbation solution $\sigma(t)$.

Although Eq. (3.64) is a system of complicated non-linear equations, because the functions on the right side are small (equivalent to ε), it is not difficult to solve the equations and to obtain solutions with a small parameter by the very well-developed perturbation method. In order to give readers an in-depth understanding of the principle of the perturbation method, also to prepare for studying different methods applied in Aerospace Dynamics which are discussed in subsequent chapters we must discuss the existence of the power series solution with a small parameter first, then provide the basic procedure of how to construct the power series solution.

3.3.2 Existence of Power Series Solution with a Small Parameter

One of the basic problems in the analytical theory of an ordinary differential equation is the existence of a solution of a given equation. Closely related to Celestial Dynamics there is a basic theorem called the Poincare theorem. The description of this theorem is as follows.

Giving a system of perturbation equations with a small parameter

$$\frac{dx_i}{dt} = X_i(x_1, x_2, \dots, x_n, t; \varepsilon), \quad i = 1, 2, \dots, n \quad (3.68)$$

in which the right-side functions X_i are continuous at t when $0 \leq t < t_1$ and can be expanded into convergent series in powers of ε , then the solution of the system of equations, $x_i = f_i(t, \varepsilon)$, can be expanded into convergent series in powers of ε for $0 \leq t < t_1$ if ε is sufficiently small, and t_1 satisfies the following discriminant

$$\frac{\alpha \varepsilon}{(1 + \alpha \varepsilon)^2} \exp(n \alpha M t_1) < \frac{1}{4}, \quad (3.69)$$

where M is the maximum of X_i on the time interval, and α is a real number related to x_i .

The convergent discriminant (3.69) can be regarded as the convergent range of the series solution. For a moving celestial body, it means that the series solution is valid when the body is on the arc s as long as

$$s \sim \frac{1}{\varepsilon}. \quad (3.70)$$

Here we use the arc $s = nt_1$ because the time t involves different time scales and is related to the moving speed of the body, so it does not have a “united” quantitative

meaning; whereas the arc s clearly reflects the scale of the continuity of orbital motion. Obviously, the smaller ε is the larger the convergent range of the power series solution is. In the convergent range, we then can construct the corresponding series solution into powers of ε . Although the solution may only reflect the characteristics of the moving body in a limited range it is enough for solving actual problems. The global configuration of the motion is beyond the scope of this monograph.

3.3.3 Construction of the Power Series Solution with a Small Parameter: The Perturbation Method

If we chose τ or $M_0 = -n\tau$ as the sixth orbital element, then the six components of function f_ε on the right side of (3.64) are all in the order of $O(\varepsilon)$, satisfying (3.65). But usually, the sixth orbital element is the mean anomaly M then the sixth component of f_ε contains $n = \sqrt{\mu}a^{-3/2} = O(\varepsilon^0)$, and Eq. (3.64) has the form as

$$\frac{d\sigma}{dt} = f_0(a) + f_1(\sigma, t, \varepsilon), \quad (3.71)$$

where

$$\begin{cases} f_0(a) = \delta n \\ \delta = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = (0 \ 0 \ 0 \ 0 \ 0 \ 1)^T \end{cases} \quad (3.72)$$

$$|f_{li}(\sigma, t, \varepsilon)| = O(\varepsilon), \quad i = 1, 2, \dots, 6. \quad (3.73)$$

Note that we now use f_1 instead of f_ε to relate the function to its magnitude, which means $f_1 = O(\varepsilon)$, $f_2 = O(\varepsilon^2)$, etc. If we add a term $f_2 = O(\varepsilon^2)$ on the right side of Eq. (3.71), it would not affect the following discussion.

The series solution of (3.71) in powers of ε has the form

$$\begin{aligned} \sigma(t) &= \sigma^{(0)}(t) + \Delta\sigma^{(1)}(t, \varepsilon) + \Delta\sigma^{(2)}(t, \varepsilon^2) + \dots + \Delta\sigma^{(l)}(t, \varepsilon^l) + \dots \\ \Delta\sigma^{(l)}(t, \varepsilon^l) &= \varepsilon^l \beta_l(t), \quad l = 1, 2, \dots \end{aligned} \quad (3.74)$$

where $\sigma^{(0)}(t)$ is the unperturbed motion solution that when $\varepsilon = 0$

$$\sigma^{(0)}(t) = \sigma_0 + \delta n_0(t - t_0) \quad (3.75)$$

and $\beta_l(t)$ is the perturbed term. The actual forms of (3.75) are

$$\begin{aligned} a^{(0)}(t) &= a_0, \quad e^{(0)}(t) = e_0, \quad i^{(0)}(t) = i_0, \\ \Omega^{(0)}(t) &= \Omega_0, \quad \omega^{(0)}(t) = \omega_0, \\ M^{(0)}(t) &= M_0 + n_0(t - t_0), \end{aligned} \quad (3.76)$$

where $\sigma_0(a_0, e_0, i_0, \Omega_0, \omega_0, M_0)$ is for the orbital elements at the epoch time t_0 , i.e., the starting time. We can see that the power series solution of ε (3.74), $\sigma(t)$, is actually the expansion of the reference orbit of the unperturbed motion equations at $\sigma^{(0)}(t)$; $\Delta\sigma^{(l)}(t, \varepsilon^l)$ is the l th variation term of the perturbation or abbreviated as the l th order perturbation term. Substituting the formatted solution (3.74) into Eq. (3.71) we have

$$\begin{aligned} & \frac{d}{dt} [\sigma^{(0)} + \Delta\sigma^{(1)} + \Delta\sigma^{(2)} + \cdots + \Delta\sigma^{(l)} + \cdots] \\ &= f_0(a) + \frac{\partial f_0}{\partial a} [\Delta a^{(1)} + \Delta a^{(2)} + \cdots] + \frac{1}{2} \frac{\partial^2 f_0}{\partial a^2} [\Delta a^{(1)} + \cdots]^2 + \cdots \\ &+ f_1(\sigma, t, \varepsilon) + \sum_{j=1}^6 \frac{\partial f_1}{\partial \sigma_j} [\Delta\sigma_j^{(1)} + \Delta\sigma_j^{(2)} + \cdots] \\ &+ \frac{1}{2} \sum_{j=1}^6 \sum_{k=1}^6 \frac{\partial^2 f_1}{\partial \sigma_j \partial \sigma_k} [\Delta\sigma_j^{(1)} + \cdots] [\Delta\sigma_k^{(1)} + \cdots] \\ &+ \cdots \end{aligned} \quad (3.77)$$

The orbital elements on the right side of (3.77) take the reference orbit values of $\sigma^{(0)}(t)$. If the power series on the right side of (3.77) is convergent (as discussed previously), then by comparing the coefficients in the same order (ε^l) on both sides we have

$$\begin{cases} \sigma^{(0)}(t) = \sigma_0 + \delta n_0(t - t_0) \\ \Delta\sigma^{(1)}(t) = \int_{t_0}^t \left[\delta \frac{\partial n}{\partial a} \Delta a^{(1)} + f_1(\sigma, t, \varepsilon)_1 \right]_{\sigma^{(0)}} dt \\ \Delta\sigma^{(2)}(t) = \int_{t_0}^t \left[\delta \left(\frac{\partial n}{\partial a} \Delta a^{(2)} + \frac{1}{2} \frac{\partial^2 n}{\partial a^2} (\Delta a^{(1)})^2 \right) + \sum_j \frac{\partial f_1}{\partial \sigma_j} \Delta\sigma_j^{(1)} \right]_{\sigma^{(0)}} dt \\ \dots \end{cases} \quad (3.78)$$

This is an efficient recursive procedure from a perturbation term of lower order to the next higher order term. With the actual analytical expression of $f_1(\sigma, t, \varepsilon)$ we can obtain each perturbation term in (3.74), and complete the construction of the power series solution with a small parameter for (3.71). This method of constructing series solutions is called the perturbation method.

To demonstrate this method, we use the following example.

Example: Use the perturbation method to solve the differential equation of the second order with a small parameter:

$$\ddot{x} + \omega^2 x = -\varepsilon x^3, \quad \varepsilon \ll 1, \quad (3.79)$$

where $\omega > 0$ is a real number.

Solution: When $\varepsilon = 0$, the unperturbed motion equation is

$$\ddot{x} + \omega^2 x = 0. \quad (3.80)$$

Its solution is

$$\begin{cases} x = a \cos(\omega t + M_0) \\ \dot{x} = -\omega a \sin(\omega t + M_0) \end{cases} \quad (3.81)$$

Assuming the initial time $t_0 = 0$, and the two constants of integration, a and M_0 are the two unperturbed elements.

When $\varepsilon \neq 0$, the two conditions obtained by the method of the variation of arbitrary constants are

$$\begin{cases} \frac{\partial x}{\partial a} \dot{a} + \frac{\partial x}{\partial M_0} \dot{M}_0 = 0 \\ \frac{\partial \dot{x}}{\partial a} \dot{a} + \frac{\partial \dot{x}}{\partial M_0} \dot{M}_0 = -\varepsilon x^3 \end{cases} \quad (3.82)$$

which yield the perturbation equations

$$\begin{cases} \dot{a} = \frac{\varepsilon}{\omega} a^3 \left(\frac{1}{4} \sin 2M + \frac{1}{8} \sin 4M \right) = (f_1)_a \\ \dot{M} = \omega + \dot{M}_0 \\ \quad = \omega + \frac{\varepsilon}{\omega} a^2 \left(\frac{3}{8} + \frac{1}{2} \cos 2M + \frac{1}{8} \cos 4M \right) = \omega + (f_1)_M \end{cases} \quad (3.83)$$

where $M = M_0 + \omega t$. Assuming the Eq. (3.83) have a set of series solutions as

$$\begin{cases} \sigma(t) = \sigma^{(0)}(t) + \Delta \sigma^{(1)}(t) + \dots \\ \sigma^{(0)}(t) = \begin{pmatrix} a_0 \\ M_0 + \omega t \end{pmatrix} \end{cases} \quad (3.84)$$

Since $\omega = \text{const.}$, there is

$$\Delta \sigma^{(1)}(t) = \int_0^t [f_1(\sigma, t, \varepsilon)]_{\sigma^{(0)}} dt. \quad (3.85)$$

By integrating we have

$$\begin{cases} \Delta a^{(1)}(t) = \frac{\varepsilon}{\omega^2} a^3 \left(-\frac{1}{8} \cos 2M - \frac{1}{32} \cos 4M \right)_0^t \\ \Delta M^{(1)}(t) = \frac{\varepsilon}{\omega^2} a^2 \left(\frac{3}{8} \omega t + \frac{1}{4} \sin 2M + \frac{1}{32} \sin 4M \right)_0^t \end{cases} \quad (3.86)$$

The perturbation terms of the second order are given by

$$\begin{cases} \Delta a^{(2)}(t) = \int_0^t \left[\frac{\partial(f_1)_a}{\partial a} \Delta a^{(1)} + \frac{\partial(f_1)_a}{\partial M} \Delta M^{(1)} \right]_{\sigma^{(0)}} dt \\ \Delta M^{(2)}(t) = \int_0^t \left[\frac{\partial(f_1)_M}{\partial a} \Delta a^{(1)} + \frac{\partial(f_1)_M}{\partial M} \Delta M^{(1)} \right]_{\sigma^{(0)}} dt \end{cases} \quad (3.87)$$

Substituting $\Delta \sigma^{(1)}$ in (3.86) into (3.87), then integrating, we have $\Delta a^{(2)}$ and $\Delta M^{(2)}$. Note that there are terms in $\Delta M^{(1)}$ containing ωt , therefore, to obtain $\Delta \sigma^{(2)}(t)$ we come across integrations like

$$\int_0^t \begin{pmatrix} \sin kM \\ \cos kM \end{pmatrix} \omega t dt, \quad k = 0, 1, \dots \quad (3.88)$$

This is a problem we should pay attention to when we deal with the perturbation equations in the orbital dynamics.

3.3.4 Secular Variations and Periodic Variations

If the disturbing force is conservative, then in a certain time interval the three variables, a , e , and i , usually only have periodic variations, whereas the variables of Ω and ω have secular variations, but their variations are much slower than the variation of the mean anomaly M . It is because the variations of Ω and ω are caused by the disturbing force, whereas M is directly related to the position of the moving body on its orbit around the center body. According to these characteristics usually a , e , and i , are called “invariant” variables, Ω and ω slow changing variables, and M (or E or f) a fast changing variable. Also because of these characteristics the perturbed terms of any order, $\Delta \sigma^{(1)}$, $\Delta \sigma^{(2)}$, $\dots$, generally contain three kinds of terms with different properties, the secular term, the long-period term, and the short-period term. The secular terms are linear functions or polynomials of $(t-t_0)$ whose coefficients are functions of a , e , and i ; the long-period terms are trigonometric functions of Ω and ω ; the short-period terms are periodic functions of M (also trigonometric functions). The short-period terms may transfer to long-period terms if they contain commensurable factors (it is discussed in the next chapter). Besides the three typical terms, there are mixed terms, such as terms in the forms of $(t-t_0) \sin(At+B)$ and $(t-t_0) \cos(At+B)$.

$t_0) \cos (At + B)$, which are called Poisson terms, the above-mentioned integration (3.88) can yield this kind of mixed term.

In the procedure of constructing a series solution by the perturbation method and in the example given in the previous section, we find that even if the perturbing force is conservative this method may still produce terms like $\varepsilon(t - t_0)$, $\varepsilon^2(t - t_0)^2$, etc., which are secular terms in polynomial form, and the actual long-period terms related to Ω and ω can become secular terms or Poisson terms. Also because we chose the unperturbed solution of $\sigma^{(0)}(t)$ as the reference orbit the results may include perturbed terms like

$$\int_{t_0}^t \cos \omega_0 dt = \cos \omega_0(t - t_0). \quad (3.89)$$

Then according to (3.78) in the constructing process, the secular terms like $(t - t_0)$, $(t - t_0)^2$, $\dots$ or Poisson terms can appear. These terms can be avoided if we let ω as $\bar{\omega} = \bar{\omega}_0 + \dot{\omega}(t - t_0)$ to replace $\omega^{(0)}(t) = \omega_0$ when we integrate (3.89), then (3.89) becomes

$$\int_{t_0}^t \cos \bar{\omega} dt = \left(\frac{\sin \bar{\omega}}{\dot{\omega}} \right)_{t_0}^t \quad (3.90)$$

This is a long-period term, and the kind of perturbation term, $(t - t_0)$, $(t - t_0)^2$ is removed.

From the qualitative point of view when the perturbing force is conservative the variables a , e , and i , usually do not have secular variations, but by the classical perturbation method to construct the perturbation solution these variables may have secular terms. This kind of solution definitely “distorts” the nature of orbital variation. Even from the quantitative point of view, it may not be serious for a short arc, but for a long arc, the difference between the long-period term and the secular term can be obvious and the accuracy of the solution suffers. Therefore, in some actual situations (particularly problems related to artificial Earth’s satellites on fast varying orbits) the analytical solution given by the classical perturbation method based on the unperturbed motion has its limitations from an either qualitative or quantitative point of view. It is necessary to improve it. In the following sections, we provide the improved methods.

3.4 An Improved Perturbation Method: The Method of Mean Orbital Elements

In Sect. 3.3.4, we show the defect in the classical perturbation method when we deal with satellite orbital motions using the unperturbed orbit as the reference orbit, particularly in some practical situations. Since the launch of the first satellite in the

previous century, many improved methods have invented. The foundation of improvement is the average method used in the non-linear dynamics and the transformation method used in celestial mechanics (i.e., the canonical transformation based on the Hamiltonian dynamics for a system of conservative force).

In this book, we introduce a few improved methods. For the motion of a satellite, the reference orbit is no longer the original unperturbed orbit given by $\sigma^{(0)}(t_0)$, but an ellipse with “long-term” precession, and the corresponding orbital elements are the mean orbital elements $\bar{\sigma}(t)$, which include secular variations, or the quasi-mean orbital elements $\bar{\sigma}'(t)$, which include both secular and long-period variations. The solutions given by the improved methods are not complicated, in which the periodic terms are simple trigonometric functions. The principle of these methods is to separate the perturbation terms by their characteristics thus avoiding the problem in the classical perturbation method. When using the improved methods to construct perturbation solutions all the classical perturbation theorems are obeyed, therefore these methods are different from the intermediate orbit of general definition, which often produces complicated structures of orbital solution.

3.4.1 Introduction of the Method of Mean Orbital Elements

The original work of using the mean orbital element method for constructing perturbed orbital solution of satellite motion was completed by Kozai [8]. Although his results were not perfect, his works are important contributions to the early research in Earth’s artificial satellite orbital dynamics and related fields. Based on Kozai’s method of mean orbital elements the author developed an improved method of mean orbital elements. In this chapter, we provide the improvements, which include the definition of referential solutions, the rigorous treatment of the mean orbital elements (separating different variations by their definitions), and the strict derivations of the long-period variations of the semi-major axis a , $a_l^{(1)}(t)$, $a_l^{(2)}(t)$ etc. These improvements make the original mean orbital elements method perfect [3–5].

The idea of using the mean orbit as the reference orbit came from the average method used in the non-linear dynamics, and was an “intuition” of the transformation method. From the practical point of view, we do not discuss the transformation method, readers who are interested in this method may refer to the related contents in references [3–7].

We start with the six orbital elements

$$\sigma = (a, e, i, \Omega, \omega, M)^T \quad (3.91)$$

and the corresponding perturbed motion equations with the initial condition are given by

$$\begin{cases} \frac{d\sigma}{dt} = f_0(a) + f_\varepsilon(\sigma, t, \varepsilon) \\ \sigma_0 = \sigma(t_0) \end{cases} \quad (3.92)$$

where

$$f_0(a) = \delta n, \quad n = \sqrt{\mu} a^{-3/2}, \quad (3.93)$$

$$\delta = (0 \ 0 \ 0 \ 0 \ 0 \ 1)^T. \quad (3.94)$$

The introduction of notation δ is for separating the five elements a, e, i, Ω , and ω from M , because the variation of M includes the unperturbed part (zero-order); and the perturbation function is given by $f_\varepsilon(\sigma, t, \varepsilon)$.

Now we separate the perturbed variations of the orbital elements, $\Delta\sigma^{(1)}, \Delta\sigma^{(2)}, \dots$ according to the varying types into the secular variation, the long-period variation, and the short-period variation (for definitions of the three types of variation see Sect. 3.3.4). The three types of variation are expressed as $\sigma_1(t - t_0), \dots, \Delta\sigma_l^{(1)}, \dots, \Delta\sigma_s^{(1)}, \dots$, and the series solution of (3.92) in powers of ε can be given as

$$\sigma(t) = \bar{\sigma}(t) + \sigma_l^{(1)} + \dots + \sigma_s^{(1)} + \dots, \quad (3.95)$$

where

$$\bar{\sigma}(t) = \bar{\sigma}^{(0)}(t) + \sigma_1(t - t_0) + \sigma_2(t - t_0) + \dots, \quad (3.96)$$

$$\bar{\sigma}^{(0)}(t) = \bar{\sigma}_0 + \delta \bar{n}(t - t_0), \quad (3.97)$$

$$\bar{\sigma}_0 = \bar{\sigma}(t_0) = \sigma_0 - [\sigma_l^{(1)}(t_0) + \dots + \sigma_s^{(1)}(t_0) + \dots]. \quad (3.98)$$

This set of orbital elements is a recombinant of the original convergent power series. The original perturbed variations $\Delta\sigma^{(1)}, \Delta\sigma^{(2)}, \dots$ are now arranged by the types of variation and presented by perturbed forms, meaning

$$\Delta\sigma_l^{(1)}(t) = \sigma_l^{(1)}(t) - \sigma_l^{(1)}(t_0), \quad \Delta\sigma_s^{(1)}(t) = \sigma_s^{(1)}(t) - \sigma_s^{(1)}(t_0), \quad \dots$$

In (3.95), there are only $\sigma_l^{(1)}(t), \dots, \sigma_s^{(1)}(t), \dots$, because $\sigma_l^{(1)}(t_0), \dots, \sigma_s^{(1)}(t_0), \dots$ are eliminated from σ_0 in (3.98). By this arrangement $\bar{\sigma}(t)$ contains only the secular variations, which is why $\bar{\sigma}(t)$ is called the set of mean orbital elements or abbreviated as mean elements, and the method is called the method of mean orbital elements.

The method of mean elements accepts $\bar{\sigma}(t)$ as the reference solution. Obviously, $\bar{\sigma}(t)$ still represents an elliptic orbit, not a fixed ellipse but an ellipse with perturbed secular variations. This reference orbit by definition is not the intermediate orbit of the usual sense. Under a conservative disturbing force, the reference orbit has a secular precession, and original geometric relationships for an elliptic motion are

still valid. We can say that the method of mean elements is a perturbation method based on a perturbed two-body problem so can be called an improved perturbation method.

3.4.2 The Mean Values of Related Variables in an Elliptic Motion

In the perturbation functions, the characteristics of variation of some terms are not easy to define, such as $\left(\frac{a}{r}\right)$ and $\cos f$. These are periodic functions of the true anomaly f , but after they are integrated over a motion period, the accumulated effects are not zero (unless the eccentricity $e = 0$). In order to separate different types of variation of a function, we need to know the average value of a function over time t .

For a function $F(t)$, we define its average value $\bar{F}$ over a motion period T as

$$\bar{F} = \frac{1}{T} \int_0^T F(t) dt. \quad (3.99)$$

Assuming F_s and F_c to be the periodic term and the non-periodic term, respectively, we have

$$F_c = \bar{F}, F_s = F - \bar{F}. \quad (3.100)$$

By the average method, we can extract the periodic term from the function, and the function now is separated into two terms as

$$F(t) = F_c + F_s. \quad (3.101)$$

This procedure shows that when we integrate $F(t)$ in (3.99), no matter what kind of method we use, it cannot affect the results of the separation given by (3.100) and (3.101), the function $F(t)$ is strictly decomposed into the periodic part and the non-periodic part. Although the purpose of using this method is to solve the perturbed orbital motion, we can still use the relationships of an elliptic motion to calculate the integral of (3.99). Of course when there is a perturbation the motion period T and the orbital elements of the ellipse all have slow variations, but these variations do not change the basic characteristics of the periodic term F_s . This process of decomposition is not only rigorous but also retains the dynamical nature of these terms.

There are various functions related to perturbing forces in the perturbed elliptic motion, the functions which are needed to be decomposed by the average method are represented mainly by four types, which are

$$\left(\frac{a}{r}\right)^p \sin qf, \quad \left(\frac{a}{r}\right)^p \cos qf,$$

$$\left(\frac{a}{r}\right)^p (f - M) \sin qf, \quad \left(\frac{a}{r}\right)^p (f - M) \cos qf, \quad p, q = 0, 1, 2, \dots$$

We now give examples to show the basic method of providing the average values and to discuss the characteristics of the average values. At the same time, we review the applications of the previously given elliptic relationships.

(1) $\sin f$ and $\cos f$ (i.e., when $p = 0$ and $q = 1$)

$$\overline{\sin f} = \frac{1}{T} \int_0^T \sin f dt = \frac{1}{2\pi} \int_0^{2\pi} \sin f dM \quad (3.102)$$

$$= \frac{1}{2\pi} \int_0^{2\pi} \sin f \left(\frac{r}{a}\right) dE = \frac{1}{2\pi} \int_0^{2\pi} \sqrt{1 - e^2} \sin E dE = 0$$

$$\overline{\cos f} = \frac{1}{T} \int_0^T \cos f dt \quad (3.103)$$

$$= \frac{1}{2\pi} \int_0^{2\pi} \cos f \left(\frac{r}{a}\right) dE = \frac{1}{2\pi} \int_0^{2\pi} (\cos E - e) dE = -e$$

(2) $\left(\frac{a}{r}\right)^3$ (i.e., when $p = 3$ and $q = 0$)

$$\begin{aligned} \overline{\left(\frac{a}{r}\right)^3} &= \frac{1}{T} \int_0^T \left(\frac{a}{r}\right)^3 dt = \frac{1}{2\pi} (1 - e^2)^{-1/2} \int_0^{2\pi} \left(\frac{a}{r}\right) df \\ &= \frac{1}{2\pi} (1 - e^2)^{-3/2} \int_0^{2\pi} (1 + e \cos f) df = (1 - e^2)^{-3/2} \end{aligned} \quad (3.104)$$

(3) $\left(\frac{r}{a}\right)$ (i.e., when $p = -1$ and $q = 0$)

$$\begin{aligned} \overline{\left(\frac{r}{a}\right)} &= \frac{1}{T} \int_0^T \left(\frac{r}{a}\right) dt = \frac{1}{2\pi} \int_0^{2\pi} \left(\frac{r}{a}\right)^2 dE \\ &= \frac{1}{2\pi} \int_0^{2\pi} (1 - e \cos E)^2 dE = 1 + \frac{1}{2} e^2 \end{aligned} \quad (3.105)$$

The above three examples show that the basic method of deriving the averages is by the transformations and the geometric relationships between time t , the eccentric anomaly E , and the true anomaly f . It also shows that it is important to know the difference between an average over time t and an average over the angle f . For example, the average of $\cos f$ over f is zero but is $-e$ over time t , which reflects the uneven motion around an ellipse.

We now give the general expressions of the averages for the four types of functions, which may be useful for readers:

$$\overline{\left(\frac{a}{r}\right)^p \sin qf} = 0 \quad (p, q = 0, 1, 2, \dots), \quad (3.106)$$

$$\overline{\cos qf} = \left(1 + q\sqrt{1 - e^2}\right) \left(\frac{-e}{1 + \sqrt{1 - e^2}}\right)^q \quad (q = 0, 1, 2, \dots), \quad (3.107)$$

$$\overline{\left(\frac{a}{r}\right) \cos qf} = \left(\frac{-e}{1 + \sqrt{1 - e^2}}\right)^q \quad (q = 0, 1, 2, \dots), \quad (3.108)$$

$$\overline{\left(\frac{a}{r}\right)^p \cos qf} = \begin{cases} 0 & (p \geq 2, q \geq p - 1) \\ -(1 - e^2)^{-(p-3/2)} \sum_{n(2)=q}^{(p-2)-\delta} \binom{p-2}{n} & \\ \left(\binom{n}{\frac{1}{2}(n-q)}\right) \left(\frac{e}{2}\right)^n & (p \geq 2, q < p - 1), \end{cases} \quad (3.109)$$

$$\overline{\left(\frac{a}{r}\right)^p (f - M) \cos qf} = 0 \quad (p \geq 0, q \geq 0), \quad (3.110)$$

$$\overline{\left(\frac{a}{r}\right)^2 (f - M) \sin qf} = -\frac{1}{q} \frac{\overline{\cos qf}}{\sqrt{1 - e^2}} \quad (q \geq 1), \quad (3.111)$$

$$\begin{aligned} \overline{\left(\frac{a}{r}\right)^p (f - M) \sin qf} &= (1 - e^2)^{-(p-3/2)} \sum_{n=0}^{p-2} \sum_{m=0}^n \binom{p-2}{n} \binom{n}{m} \left(\frac{e}{2}\right)^n \\ &\quad \times \left(-\frac{\cos(q + n - 2m)f}{q + n - 2m}\right)_{2m \neq q+n} \quad (p \geq 3, q \geq 1) \end{aligned} \quad (3.112)$$

In (3.111) and (3.112) the case of $p = 0$ or $p = 1$ rarely occurs, so is not discussed here. In the above procedure, we need two expressions related to the trigonometric functions which are

$$\begin{aligned} \sin^n f &= \frac{1}{2^n} \sum_{m=0}^n (-1)^{\frac{1}{2}(2m+n-\delta_1)} \binom{n}{m} \left[\begin{aligned} &(1 - \delta_1) \cos(n - 2m)f \\ &+ \delta_1 \sin(n - 2m)f \end{aligned} \right] \\ &= \frac{1}{2^n} \sum_{m=0}^{\frac{1}{2}(n-\delta_1)} 2^{\delta_2} (-1)^{\frac{1}{2}(2m+n-\delta_1)} \binom{n}{m} \left[\begin{aligned} &(1 - \delta_1) \cos(n - 2m)f \\ &+ \delta_1 \sin(n - 2m)f \end{aligned} \right], \end{aligned} \quad (3.113)$$

$$\begin{aligned} \cos^n f &= \frac{1}{2^n} \sum_{m=0}^n \binom{n}{m} \cos(n - 2m)f \\ &= \frac{1}{2^n} \sum_{m=0}^{\frac{1}{2}(n-\delta_1)} 2^{\delta_2} \binom{n}{m} \cos(n - 2m)f \end{aligned} \quad (3.114)$$

in which δ , δ_1 , and δ_2 are defined as

$$\delta = \frac{1}{2} [1 - (-1)^{p-q}], \quad (3.115)$$

$$\delta_1 = \frac{1}{2}[1 - (-1)^n], \quad (3.116)$$

$$\delta_2 = \begin{cases} 0, & n - 2m = 0 \\ 1, & n - 2m \neq 0 \end{cases} \quad (3.117)$$

3.4.3 Construction of Formal Solution: The Method of Mean Orbital Elements [3–8]

In most cases, the disturbing function on the right side of (3.92) can be expanded into power series of ε , that

$$f_\varepsilon(\sigma, t, \varepsilon) = f_1(\sigma, t, \varepsilon) + f_2(\sigma, t, \varepsilon^2) + \cdots + f_N(\sigma, t, \varepsilon^N) + \cdots \quad (3.118)$$

$$f_N = O(\varepsilon^N)$$

In order to use the method of mean elements, it is necessary to separate the perturbation variations into secular terms and periodic terms. We apply the same method described above to decompose the perturbation function $f_N(\sigma, t, \varepsilon^N)$, $N = 1, 2, \dots$ into three parts of secular, long-periodic, and short-periodic terms that

$$f_N = f_{Nc} + f_{Nl} + f_{Ns}, \quad N = 1, 2, \dots \quad (3.119)$$

On the right side the subscripts, “c”, “l”, and “s”, after N represent the secular, long-periodic, and short-periodic terms, respectively. According to the characteristics of the orbital elements, we know that f_{Nc} is only related to a , e , and i ; the period of f_{Nl} depends on the variations of Ω and ω , or the commensurable term (are discussed late); the period of f_{Ns} depends on the fast-varying variable M . To apply the method of mean elements requires

$$f_{lI} = 0. \quad (3.120)$$

In the orbital dynamics for a circling motion of a spacecraft, this condition is often satisfied.

Now substituting the formal solution (3.95) into Eq. (3.92), and expanding the function on the right side at $\bar{\sigma}(t)$, we have

$$\begin{aligned}
& \frac{d}{dt} \left[\bar{\sigma}^{(0)}(t) + \sigma_1(t - t_0) + \sigma_2(t - t_0) + \cdots + \sigma_l^{(1)}(t) + \cdots + \sigma_s^{(1)}(t) + \cdots \right] \\
&= f_0(\bar{a}) + \frac{\partial f_0}{\partial a} \left[a_l^{(1)}(t) + a_l^{(2)}(t) + \cdots + a_s^{(1)}(t) + a_s^{(2)}(t) + \cdots \right] \\
&\quad + \frac{1}{2} \frac{\partial^2 f_0}{\partial a^2} \left[a_l^{(1)}(t) + \cdots + a_s^{(1)}(t) + \cdots \right]^2 + \cdots \\
&+ f_1(\bar{\sigma}, t, \varepsilon) + \sum_{j=1}^6 \frac{\partial f_1}{\partial \sigma_j} \left[\sigma_l^{(1)}(t) + \sigma_l^{(2)}(t) + \cdots + \sigma_s^{(1)}(t) + \sigma_s^{(2)}(t) + \cdots \right]_j + \cdots \\
&+ \sum_{k=1}^6 \sum_{j=1}^6 \frac{\partial^2 f_1}{\partial \sigma_j \partial \sigma_k} \left[\sigma_l^{(1)}(t) + \cdots + \sigma_s^{(1)}(t) + \cdots \right]_j \left[\sigma_l^{(1)}(t) + \cdots + \sigma_s^{(1)}(t) + \cdots \right]_k \\
&\quad + \cdots + f_2(\bar{\sigma}, t, \varepsilon^2) + \sum_{j=1}^6 \frac{\partial f_2}{\partial \sigma_j} \left[\sigma_l^{(1)}(t) + \cdots + \sigma_s^{(1)}(t) + \cdots \right]_j \\
&\quad + \cdots + f_N(\bar{\sigma}, t, \varepsilon^N) + \cdots
\end{aligned} \tag{3.121}$$

In (3.121) the orbital elements of σ on the right side are the elements of the reference solution $\bar{\sigma}(t)$, and f_0 is given by (3.93). If the series (3.95) is convergent (its convergence is described in Sect. 3.3.2), we then compare the terms with the same orders (ε^N) on both sides, and integrate to give

$$\bar{\sigma}^{(0)}(t) = \int_{t_0}^t f_0(\bar{a}) dt = \bar{\sigma}_0 + \delta \bar{n}(t - t_0), \tag{3.122}$$

$$\sigma_1(t - t_0) = \int_{t_0}^t [f_{1c}]_{\bar{\sigma}} dt, \tag{3.123}$$

$$\sigma_s^{(1)}(t) = \int^t \left[\delta \frac{\partial n}{\partial a} a_s^{(1)}(t) + f_{1s} \right]_{\bar{\sigma}} dt, \tag{3.124}$$

$$\sigma_2(t - t_0) = \int_{t_0}^t \left[\delta \frac{1}{2} \frac{\partial^2 n}{\partial a^2} (a_s^{(1)})_c^2 + \left(\sum_j \frac{\partial f_1}{\partial \sigma_j} (\sigma_l^{(1)} + \sigma_s^{(1)})_j \right)_c + f_{2c} \right]_{\bar{\sigma}} dt, \tag{3.125}$$

$$\sigma_l^{(1)}(t) = \int^t \left[\delta \frac{\partial n}{\partial a} a_l^{(2)} + \delta \frac{1}{2} \frac{\partial^2 n}{\partial a^2} (a_s^{(1)})_l^2 + \left(\sum_j \frac{\partial f_1}{\partial \sigma_j} (\sigma_l^{(1)} + \sigma_s^{(1)})_j \right)_l + f_{2l} \right]_{\bar{\sigma}} dt, \tag{3.126}$$

$$\sigma_s^{(2)}(t) = \int^t \left[\delta \frac{\partial n}{\partial a} a_s^{(2)} + \delta \frac{1}{2} \frac{\partial^2 n}{\partial a^2} (a_s^{(1)})_s^2 + \left(\sum_j \frac{\partial f_1}{\partial \sigma_j} (\sigma_l^{(1)} + \sigma_s^{(1)})_j \right)_s + f_{2s} \right]_{\bar{\sigma}} dt, \tag{3.127}$$

In the integrands, the subscripts of a function in parentheses, “c”, “l”, and “s”, represent the secular, long-period, and short-period parts of the function. For example, let A as a function that

$$\begin{aligned}
A &= \cos f + \cos(f + \omega) \\
&= \cos f + \cos f \cos \omega - \sin f \sin \omega
\end{aligned}$$

by the average method, we decompose A into three parts that

$$\begin{aligned}
A &= (A)_c + (A)_l + (A)_s, \\
(A)_c &= -e, \\
(A)_l &= -e \cos \omega, \\
(A)_s &= (\cos f + e) + (\cos f + e) \cos \omega - \sin f \sin \omega.
\end{aligned}$$

It is easy to prove that if $f_{1l} = 0$ then $a_l^{(1)}(t) = 0$ (details are given in the next section). Therefore, as long as $f_{1l} = 0$, the above recursive procedure related to the method of mean elements is valid, i.e., from a low order to obtain the next higher order. There are a few points that need to be explained as follows.

- (1) For a conservative disturbing force, there are no secular variations in a , e , and i , and the secular terms of Ω , ω , and M are linear functions of $(t - t_0)$. The reason is that the integrands of the secular terms, $\sigma_1, \sigma_2, \dots$ are functions of a , e , and i , during integration, the values of a , e , and i , are $\bar{a} = \bar{a}_0, \bar{e} = \bar{e}_0, \bar{i} = \bar{i}_0$. If the disturbing force is a dissipative force, the configuration of the solution is more complicated, such as the secular terms are no longer linear functions of $(t - t_0)$. Generally, a dissipative force is relatively small in the same order of ε^2 , i.e., a small variable in the second order, or even smaller, therefore it does not affect the configuration of the series solution. We do not repeat this explanation in similar problems.
- (2) Unlike the classical perturbation method, the reference solution $\bar{\sigma}(t)$ of the mean elements is actually formed during the recursive procedure, but this process does not change the configuration of the solution. For example, when the disturbing force is conservative, we have

$$\int \cos \bar{\omega} dt = \frac{\sin \bar{\omega}}{(\omega_1 + \omega_2 + \dots)}$$

where $\omega_1, \omega_2, \dots$ are the coefficients of secular terms of ω , and are functions of $\bar{a}, \bar{e}, \bar{i}$. During integration, we do not have to know their actual forms, which are only needed when we calculate the solution.

- (3) The variations of the long-period terms depend on ω and Ω . Take ω as an example, because $\bar{\omega} = \bar{\omega}_0 + \omega_1(t - t_0) + \dots$, where $\omega_1 = O(\varepsilon)$ is a small quantity of the first order, it varies much slower than the mean anomaly M as $\bar{M} = \bar{M}_0 + \bar{n}_0(t - t_0) + M_1(t - t_0) + \dots$ and $\bar{n}_0 = O(\varepsilon^0)$. If $f_{2l} = \varepsilon^2 \cos \omega$, we have

$$\int_t f_{2l} dt = \int^t \varepsilon^2 \cos \bar{\omega} dt = \frac{\varepsilon^2 \sin \bar{\omega}}{(\omega_1 + \dots)} = A \sin \bar{\omega},$$

where $A = O(\varepsilon)$. The result of the above integration is a long-period term of the first order not a long-period term of the second order. This is the reducing-order phenomenon when integrating a long-period term, and it is why the left side of (3.126) is written as $\sigma_l^{(1)}(t)$. In fact for a limited time interval $\sigma_l^{(1)}(t)$ corresponds to $\Delta \sigma_l^{(1)}(t) = \sigma_l^{(1)}(t) - \sigma_l^{(1)}(t_0)$, which is equivalent to a secular term of the second order $\sigma_2(t - t_0)$ (this property of the long-period term is mentioned late in related problems). In the classical method, the form of $\sigma_2(t - t_0)$ appears in the result of a long-period variable as mentioned in Sect. 3.3.4, whereas in the method of mean elements it appears in the form of $\sigma_l^{(1)}(t)$, i.e., it keeps its periodic nature so is reasonable. By the recursive procedure, the term $\sigma_l^{(2)}(t)$ is given by integrating f_{3l} , etc. When calculate $\sigma_l^{(1)}(t)$ by (3.126), we find another problem, that in the integrand on the right side we need to know not only $\sigma_l^{(2)}(t)$ but also $a_l^{(2)}(t)$ for the sixth element M . Careful examinations can show that this problem does not influence the configuration of the solution. The author solved this problem and at the same time improved the original work by Kozai as mentioned in Sect. 3.4.1. Details about this problem are given in Sect. 3.5.

3.4.4 Example

The same example of the equations of a small parameter in Sect. 3.3.3 is used here to show how to construct a perturbation solution using the method of mean elements, and at the same time to compare this method with the classical method. The original Eq. (3.79) is

$$\ddot{x} + \omega^2 x = -\varepsilon x^3,$$

where $\varepsilon \ll 1$ and $\omega > 0$, both are real constants. The unperturbed solution given in Sect. 3.3.3 is

$$\begin{cases} x = a \cos(\omega t + M_0) \\ \dot{x} = -\omega a \sin(\omega t + M_0) \\ M = M_0 + \omega t \end{cases} \quad (3.128)$$

The corresponding perturbation equation system is

$$\begin{cases} \dot{a} = \frac{\varepsilon}{\omega} a^3 \left(\frac{1}{4} \sin 2M + \frac{1}{8} \sin 4M \right) \\ \dot{M} = \omega + \frac{\varepsilon}{\omega} a^2 \left(\frac{3}{8} + \frac{1}{2} \cos 2M + \frac{1}{8} \cos 4M \right) \end{cases} \quad (3.129)$$

Now we use the method of mean elements to solve the equations. The equations can be arranged according to this method as

$$\begin{cases} \dot{a} = (f_{1s})_a \\ \dot{M} = (f_0)_M + (f_{1c})_M + (f_{1s})_M \end{cases} \quad (3.130)$$

where

$$(f_{1s})_a = \frac{\varepsilon}{\omega} a^3 \left(\frac{1}{4} \sin 2M + \frac{1}{8} \sin 4M \right), \quad (3.131)$$

$$\begin{cases} (f_0)_M = \omega = \text{const} \\ (f_{1c})_M = \frac{\varepsilon}{\omega} a^2 \left(\frac{3}{8} \right) \\ (f_{1s})_M = \frac{\varepsilon}{\omega} a^2 \left(\frac{1}{2} \cos 2M + \frac{1}{8} \cos 4M \right) \end{cases} \quad (3.132)$$

Following the procedure of constructing the power series of ε by the method of mean elements (3.122)–(3.127) we have

$$\begin{cases} \bar{a}^{(0)}(t) = \bar{a}_0 \\ \bar{M}^{(0)}(t) = \bar{M}_0 + \omega(t - t_0) \end{cases} \quad (3.133)$$

With (3.133) we integrate (3.123) and (3.124) to give

$$\begin{cases} a_1(t - t_0) = 0 \\ M_1(t - t_0) = \frac{\varepsilon}{\omega^2} \bar{a}_0^2 \left(\frac{3}{8} \right) \omega(t - t_0) \end{cases} \quad (3.134)$$

$$\begin{cases} a_s^{(1)}(t) = \frac{\varepsilon}{\omega^2(1+M_1/\omega+\dots)} \bar{a}_0^3 \left(-\frac{1}{8} \cos 2\bar{M} - \frac{1}{32} \cos 4\bar{M} \right) \\ M_s^{(1)}(t) = \frac{\varepsilon}{\omega^2(1+M_1/\omega+\dots)} \bar{a}_0^2 \left(\frac{1}{4} \sin 2\bar{M} + \frac{1}{32} \sin 4\bar{M} \right) \end{cases} \quad (3.135)$$

$$\begin{cases} a_2(t - t_0) = \int_{t_0}^t \left[\frac{\partial(f_{1s})_a}{\partial a} a_s^{(1)}(t) + \frac{\partial(f_{1s})_a}{\partial M} M_s^{(1)}(t) \right]_c dt = 0 \\ M_2(t - t_0) = \int_{t_0}^t \left[\frac{\partial(f_{1s})_M}{\partial a} a_s^{(1)}(t) + \frac{\partial(f_{1s})_M}{\partial M} M_s^{(1)}(t) \right]_c dt = \left(\frac{\varepsilon}{\omega^2} \bar{a}_0^2 \right)^2 \left(-\frac{51}{256} \right) \omega(t - t_0) \end{cases} \quad (3.136)$$

Since $M_1/\omega = O(\varepsilon)$, for a solution with an accuracy of the first order, we can reduce the denominator on the right sides of (3.135) $\omega^2(1 + M_1/\omega + \dots)$ to ω^2 .

The above expressions of the derived perturbation terms of all orders are simpler than those by the classical method given in the previous section, furthermore, there is no integration in the form of (3.89), therefore there are no Poisson terms (or mixed terms). The solution now has the form as

$$\begin{cases} a(t) = \bar{a}_0 + a_s^{(1)}(t) + \dots \\ M(t) = \bar{M}_0 + \omega(t - t_0) + (M_1 + M_2 + \dots)(t - t_0) + M_s^{(1)}(t) + \dots \end{cases} \quad (3.137)$$

3.4.5 Two Annotations About the Method of Mean Elements

In the above Sect. 3.4.3, we describe the method of mean elements based on Kozai's theory. There are two details that need to be further explained. One is about the general property $a_l^{(1)}(t) = 0$, and the other is about the derivation of $a_l^{(2)}(t)$.

(1) Prove $a_l^{(1)}(t) = 0$ if $f_{ll} = 0$

A conservative system is a Hamiltonian system. In a perturbed two-body problem for a gravitational perturbation, the corresponding Hamiltonian function is

$$H = \frac{1}{2}v^2 - V, \quad (3.138)$$

where V includes both the gravity potential of the center body and all other perturbing potentials. Because the perturbation forces are conservative, we have

$$V = \frac{\mu}{r} + R(r, \varepsilon) \quad (3.139)$$

with $\mu = G(m_0 + m)$, and R is the perturbation function. For a perturbed two-body the vis viva integral is valid that

$$v^2 = \mu \left(\frac{2}{r} - \frac{1}{a} \right).$$

Substituting this into (3.138) leads to

$$H = -\frac{\mu}{2a} - R. \quad (3.140)$$

This dynamical system has an integral (the energy integral) which is

$$\frac{\mu}{2a} + R = C. \quad (3.141)$$

We now expand this integral at the reference solution $\bar{\sigma}(t)$, and separate terms according to their characteristics. The results are

“Constant terms”:

$$\frac{\mu}{2a} + (R_{1c} + R_{2c} + \cdots) + \cdots = C. \quad (3.142)$$

Long-period terms of the first-order:

$$\frac{\partial}{\partial a} \left(\frac{\mu}{2a} \right) \left[a_l^{(1)}(t) \right] + R_{ll} = 0. \quad (3.143)$$

Short-period terms of the first-order:

$$\frac{\partial}{\partial a} \left(\frac{\mu}{2a} \right) [a_s^{(1)}(t)] + R_{1s} = 0, \quad (3.144)$$

.

Because $f_{1l} = 0$ and f_{1l} is given by $\partial R_{1l} / \partial \sigma$, we must have $R_{1l} = 0$, which leads to

$$\frac{\partial}{\partial a} \left(\frac{\mu}{2a} \right) [a_l^{(1)}(t)] = 0.$$

Obviously $\frac{\partial}{\partial a} \left(\frac{\mu}{2a} \right) \neq 0$, therefore there is $a_l^{(1)}(t) = 0$.

- (2) Derivation of $a_l^{(2)}(t)$ for constructing a perturbation solution due to the dynamical form-factor of Earth's non-spherical gravity, J_2 term

One of the improvements to the Kozai theory initiated by the author is to provide a perturbation solution due to Earth's dynamical form-factor J_2 term. Based on Kozai's method and following the rigorous definition of the mean elements, the author derives the long-period term of the second-order for the semi-major axis a , $a_l^{(2)}(t)$, by which the author completes the configuration of the perturbation solution of a satellite orbit, therefore making the method of mean elements perfect. The actual derivation of $a_l^{(2)}(t)$ is described as follows.

In the perturbation solution of the sixth element M , we need $a_l^{(2)}(t)$ to derive $M_l^{(1)}(t)$. By the procedure of constructing the perturbation solution given in (3.121), we need the expansion up to the third order, which is extremely complicated. Now we use the similar method as described in the annotation (1), still chose $\bar{\sigma}(t)$ as the reference solution and from the energy integral to derive $a_l^{(2)}(t)$. The result is

$$\begin{aligned} a_l^{(2)}(t) = & \left(\frac{3J_2}{2p^2} \right)^2 a \sqrt{1-e^2} \\ & \times \left\{ \left[-\frac{1}{6} \sin^2 i (4 - 5 \sin^2 i) \overline{\cos 2f} + e^2 \sin^2 i \left(\frac{17}{12} - \frac{19}{8} \sin^2 i \right) \right] \cos 2\omega \right. \\ & \left. + \frac{e^4}{1-e^2} \left[\frac{7}{3} \sin^2 i \left(1 - \frac{3}{2} \sin^2 i \right) \cos 2\omega + \frac{1}{32} \sin^4 i \cos 4\omega \right] \right\} \end{aligned} \quad (3.145)$$

where

$$\overline{\cos 2f} = \frac{1 + 2\sqrt{1-e^2}}{(1 + \sqrt{1-e^2})^2} e^2 \quad (3.146)$$

The author worked on the improvement on the method of the mean elements and completed the above-described work during the time between 1963 and 1975. In 1975 the author submitted the completed analytical solution of satellite orbital motion to the Aerospace Application Department, China. With the sponsorship of the Department, the author wrote a textbook titled “Orbit theory of Earth’s artificial satellite dynamics”, which was printed by Nanjing University for university students in 1979 (Liu, L. and Zhao, D. Z.). The revised textbook was formally published in 1992 titled *Orbital dynamics of Earth’s artificial satellite* by Higher Education Press, China [3].

3.5 The Method of Quasi-Mean Elements: The Structure of the Formal Solution

3.5.1 Small Divisors in Expressions of Perturbation Solutions

When we construct the perturbation solution by the method of mean elements for the gravitational perturbation due to the dynamical form-factor of the central body, the J_2 term, the long-period term of the first-order has a form like

$$\sigma_l^{(1)}(t) = \int^t f_{2l} dt = \int^t J_2^2 \cos \bar{\omega} dt = \frac{J_2^2}{(\omega_1 + \dots)} \sin \bar{\omega}, \quad (3.147)$$

where the variation rate of the first-order secular term ω_1 is given by

$$\omega_1 = \frac{3J_2}{2p^2} n \left(2 - \frac{5}{2} \sin^2 i\right). \quad (3.148)$$

If the inclination of a satellite’s orbit $i = 63^\circ 26'$, then we have $(2 - \frac{5}{2} \sin^2 i) \approx 0$, which results in a small divisor in (3.147). This angle is called the critical inclination, and the corresponding problem of a small divisor is called the commensurable problem. The small divisor makes the perturbation solution constructed by the method of mean elements invalid formally.

Another kind of small divisor appears in periodic terms when the perturbation comes from the ellipticity of Earth’s equator, the $J_{2,2}$ term, or from a their-body gravity force. For example, for a medium Earth orbit satellite the $a_s^{(2)}(t, J_{2,2}; \alpha)$ due to the $J_{2,2}$ term taking the following form (only retaining the terms of $O(e)$)

$$a_s^{(2)}(t) = \frac{3J_{2,2}}{2a} \left\{ (1 + \cos i)^2 \left[-\frac{e}{2(1 - 2\alpha)} \cos(M + 2\omega + 2\Omega_{2,2}) \right] \right. \\ \left. + 2 \sin^2 i \left[\frac{3e}{2(1 - 2\alpha)} \cos(M + 2\Omega_{2,2}) \right] \right\} \quad (3.149)$$

where α is defined by

$$\alpha = n'/n, \quad (3.150)$$

and n and n' are the mean angular speeds of the satellite and Earth's rotation angular speeds, respectively. When $\alpha = n'/n \approx 1/2$, the commensurable problem occurs in a factor of $1/(1 - 2\alpha)$. Another example is for a geostationary Earth orbit (GEO) satellite, there is a commensurable factor $1/(1 - \alpha)$ corresponding to $\alpha = n'/n \approx 1$.

The above-mentioned commensurable problems make the perturbation solution provided by the method of mean elements invalid formally, but it does not mean that the orbital motion of the satellite is substantively abnormal. The commensurable problem caused by small divisors in the periodic terms of the perturbation solution is a theoretical problem in the field of Celestial Dynamics, the corresponding orbits do have some differences compared to the orbits without this problem. In this book we do not give a further discussion about this problem, our concern is on how to construct practical perturbation solutions for different applications.

In fact, if we only need to construct a power series solution with a small parameter, this commensurable singularity of a small divisor can be eliminated by proper methods. For example, if we construct a solution by power series in a small parameter $\varepsilon^{1/2}$, then the small divisors do not appear [11, 12]. Practically the procession by this method is too complicated to use. In order to solve the problem of singularity in a relatively simple way the author develops an improved method of mean elements, which can eliminate singularities and at the same time reserve all the advantages of the original method of the mean elements. The principle of the improvement is to treat the long-period term which has a small divisor, the usual long-period term, and the secular term together. We may call the new mean elements as the quasi-mean elements, by which the perturbation solution takes the following form

$$\sigma(t) = \bar{\sigma}(t) + \sigma_s^{(1)}(t) + \sigma_s^{(2)}(t) + \cdots, \quad (3.151)$$

$$\begin{cases} \bar{\sigma}(t) = \bar{\sigma}_0 + (\delta\bar{n}_0 + \sigma_1 + \sigma_2 + \cdots)(t - t_0) + \Delta\sigma_l^{(1)}(t) + \cdots \\ \bar{\sigma}_0 = \sigma_0 - [\sigma_s^{(1)}(t_0) + \sigma_s^{(2)}(t_0) + \cdots] \\ \Delta\sigma_l^{(1)} = \sigma_l^{(1)}(t) - \sigma_l^{(1)}(t_0) \end{cases} \quad (3.152)$$

For the sake of simplicity, the set of quasi-mean elements defined by (3.152) is still denoted to $\bar{\sigma}(t)$, and all the other notations have the same meanings as before. Using the quasi-mean elements as the reference orbit, we can construct the power series solution with a small parameter. We now call this method the method of quasi-mean elements. The author developed this method first in 1973 for solving the problem of critical inclination dealing with the perturbation of Earth's dynamical form-factor J_2 term, then published it in 1974 [9].

By this method, the treatment for the long-period term (including short-period terms with commensurable singularities) solves the problem caused by small divisors

in the classical perturbation method and in the method of mean elements. At the same time, this method keeps the main characteristics of the mean elements $\bar{\sigma}(t)$ as well as the basic principles of the method of mean elements for improving the classical perturbation method. The new set of $\bar{\sigma}(t)$ includes not only the secular term but also the long-period term, which is similar to the secular term to a certain extent, for that we call it a set of quasi-mean elements. In the next section, we give the entire procedure of constructing a power series solution with a small parameter by the method of quasi-mean elements.

3.5.2 Configuration of Formal Solution: The Method of Quasi-Mean Elements

By the same notation $\sigma = (a, e, i, \Omega, \omega, M)^T$ the corresponding perturbation equation with the initial condition is given by

$$\begin{cases} \frac{d\sigma}{dt} = f_0(a) + f_1(\sigma, t, \varepsilon) + f_2(\sigma, t, \varepsilon^2) + \cdots \\ \sigma_0 = \sigma(t_0) \end{cases} \quad (3.153)$$

where

$$f_0(a) = \delta n, \quad n = \sqrt{\mu} a^{-3/2}. \quad (3.154)$$

Again, we use $\delta = (0 \ 0 \ 0 \ 0 \ 0 \ 1)^T$ for separating the five elements a, e, i, Ω , and ω from M .

Like the method of mean elements, the function $f_N(\sigma, t, \varepsilon^N)$, $N = 1, 2, \dots$ is decomposed into three parts that

$$f_N = f_{Nc} + f_{Nl} + f_{Ns}, \quad N = 1, 2, \dots \quad (3.155)$$

Here on the right side the subscripts after N , “ c ”, “ l ”, and “ s ”, represent the secular, long-periodic, and short-periodic terms, respectively, meaning f_{Nc} is only related to a, e , and i ; the period of f_{Nl} depends on the variations of Ω and ω , or the commensurable term (given in Chap. 4); the period of f_{Ns} depends on the fast-varying variable M . Since the method of quasi-mean elements is based on the method of mean elements, it holds the same requirement of the original method of mean elements, which is

$$f_{ll} = 0. \quad (3.156)$$

Now we substitute the formal solution (3.151) into the Eq. (3.153), and expand the function on the right side at the quasi-mean elements $\bar{\sigma}(t)$ to give

$$\begin{aligned}
& \frac{d}{dt} [\bar{\sigma}^{(0)}(t) + \sigma_1(t - t_0) + \sigma_2(t - t_0) + \cdots + \Delta\sigma_l^{(1)}(t) + \cdots + \sigma_s^{(1)}(t) + \sigma_s^{(2)}(t) \\
& \quad + \cdots] \\
& = f_0(\bar{a}) + \frac{\partial f_0}{\partial a} [a_s^{(1)}(t) + a_s^{(2)}(t) + \cdots] + \frac{1}{2} \frac{\partial^2 f_0}{\partial a^2} [a_s^{(1)}(t) + \cdots]^2 + \cdots \\
& + f_1(\bar{\sigma}, t, \varepsilon) + \sum_{j=1}^6 \frac{\partial f_1}{\partial \sigma_j} [\sigma_s^{(1)}(t) + \sigma_s^{(2)}(t) + \cdots]_j \\
& \quad + \sum_{k=1}^6 \sum_{j=1}^6 \frac{\partial^2 f_1}{\partial \sigma_j \partial \sigma_k} [\sigma_s^{(1)}(t) + \cdots]_j [\sigma_s^{(1)}(t) + \cdots]_k + \cdots \\
& + f_2(\bar{\sigma}, t, \varepsilon^2) + \sum_{j=1}^6 \frac{\partial f_2}{\partial \sigma_j} [\sigma_s^{(1)}(t) + \cdots]_j + \cdots + f_N(\bar{\sigma}, t, \varepsilon^N) + \cdots + \cdots
\end{aligned} \tag{3.157}$$

On the right side σ is the reference solution $\bar{\sigma}(t)$. We then compare the coefficients of the terms with the same (ε^N) on both sides in the range of convergence of power series solution with a small parameter (3.151), and integrate to give

$$\bar{\sigma}^{(0)}(t) = \int_{t_0}^t f_0(\bar{a}) dt = \bar{\sigma}_0 + \delta \bar{n}(t - t_0), \tag{3.158}$$

$$\sigma_1(t - t_0) = \int_{t_0}^t [f_{1c}]_{\bar{\sigma}} dt, \tag{3.159}$$

$$\sigma_s^{(1)}(t) = \int^t \left[\delta \frac{\partial n}{\partial a} a_s^{(1)}(t) + f_{1s} \right]_{\bar{\sigma}} dt, \tag{3.160}$$

$$\sigma_2(t - t_0) = \int_{t_0}^t \left[\delta \frac{1}{2} \frac{\partial^2 n}{\partial a^2} (a_s^{(1)})_c^2 + \left(\sum_j \frac{\partial f_1}{\partial \sigma_j} (\sigma_s^{(1)})_j \right)_c + f_{2c} \right]_{\bar{\sigma}} dt, \tag{3.161}$$

$$\begin{aligned}
\Delta\sigma_l^{(1)}(t) &= \sigma_l^{(1)}(t) - \sigma_l^{(1)}(t_0) \\
&= \int_{t_0}^t \left[\delta \frac{1}{2} \frac{\partial^2 n}{\partial a^2} (a_s^{(1)})_l^2 + \left(\sum_j \frac{\partial f_1}{\partial \sigma_j} (\sigma_s^{(1)})_j \right)_l + f_{2l} \right]_{\bar{\sigma}} dt,
\end{aligned} \tag{3.162}$$

$$\begin{aligned}
\sigma_s^{(2)}(t) &= \int^t \left[\delta \frac{\partial n}{\partial a} a_s^{(2)} + \delta \frac{1}{2} \frac{\partial^2 n}{\partial a^2} (a_s^{(1)})_s^2 + \left(\sum_j \frac{\partial f_1}{\partial \sigma_j} (\sigma_s^{(1)})_j \right)_s + f_{2s} \right]_{\bar{\sigma}} dt \\
&\quad \dots \dots,
\end{aligned} \tag{3.163}$$

In the integrands, the subscripts of a function in parentheses, “ c ”, “ l ”, and “ s ”, represent the secular, long-period, and short-period parts of the function, respectively.

After eliminating short-period terms $\sigma_s^{(1)}(t)$, $\sigma_s^{(2)}(t)$, $\dots$ the set of quasi-mean elements $\bar{\sigma}(t)$ satisfies the following equations

$$\begin{cases} \frac{d\bar{\sigma}}{dt} = f_0(\bar{a}) + [f_\varepsilon(\bar{\sigma}, t, \varepsilon)]_{c,l} \\ \bar{\sigma}_0 = \bar{\sigma}(t_0) \end{cases} \quad (3.164)$$

Considering the slow variation of the right-side function we may numerically integrate it by large step lengths to get $\bar{\sigma}(t)$, and call this method a semi-analytical method. This kind of method is used successfully in studying some dynamic evolution problems in the Solar System. But for constructing the orbital solution for a small body, specifically a circling spacecraft (i.e., a satellite) the analytical solution by the method of quasi-mean elements $\bar{\sigma}(t)$ is easy to be constructed and is easy to calculate. In fact, to obtain the analytical power series solution the main trouble is in the procedure of extracting the short-period terms $\sigma_s^{(1)}(t)$, $\sigma_s^{(2)}(t)$, $\dots$ After separating the short period terms the system of quasi-mean elements is quite simple with respect to the entire dynamical system. If we use the numerical method instead of the analytical method to solve the equation system (3.153) we do not gain in-depth knowledge about the characteristics of the orbital variations provided by the analytical solution. Regarding the efficiency of calculation, the numerical method is not as fast as the analytical method, because a spacecraft orbital problem is totally different from a problem of dynamic evolution in the Solar System.

3.6 Methods of Constructing Non-singularity Solutions for a Perturbed Orbit

As described in Sect. 3.2.4, when we use Kepler orbital elements the geometric singularities of $e = 0$ and $i = 0$ or 180° appear in the perturbation equations, therefore in the vicinities of these singularities the motion solution is invalid, but the corresponding motion is normal. For example, when $e = 0$ the corresponding circular motion exists. The geometric singularities are “similar” to the commensurable singularities, both are caused by small divisors due to improper choices of basic variables. Thus, these problems can be solved by changing variables. Surely it would be more beneficial if we have a general method to eliminate all sorts of singularity. In this section, we provide two types of methods for constructing non-singularity perturbation solutions which are the follows:

- (1) The construction method of simultaneously eliminating a small e and commensurable singularities and
- (2) The construction method of simultaneously eliminating a small e , a small i , and commensurable singularities.

As mentioned above the author submitted a report to the Aerospace Application Department, China, which included the complete analytical solution of satellite orbit in 1975, then published a research paper titled “A calculation method for a perturbed Earth’s artificial satellite” in the Chinese Journal of Astronomy in the same year [8]. After a series of improvements in practical application over multi-years, the final method is described systematically in two books [9, 10]. In this section, we introduce the outlines of the two methods for constructing non-singularity perturbation solutions due to Earth’s dynamical form-factor J_2 term.

3.6.1 Configuration of the Non-singularity Perturbation Solutions of the First Type

(1) Basic variables and basic equations

In order to eliminate the small e singularity, we choose the non-singularity variables of the first-type as

$$a, i, \Omega, \xi = e \cos \omega, \eta = e \sin \omega, \lambda = M + \omega, \quad (3.165)$$

and use the same notation σ for the above six variables in the order of (3.165). The corresponding perturbation equations in terms of $(\partial R / \partial \sigma)$ are given by (3.31) in Sect. 3.2.

As an example, we use the perturbation function due to the main part of Earth’s non-spherical gravity potential, the J_2 term, given by

$$R = \frac{J_2}{2a^3} \left(\frac{a}{r} \right)^3 \left[\left(1 - \frac{3}{2} \sin^2 i \right) + \frac{3}{2} \sin^2 i \cos 2(f + \omega) \right] \quad (3.166)$$

This function can be decomposed into two parts using the average method that

$$R = R_1 = R_{1c} + R_{1s}, \quad (3.167)$$

$$R_{1c} = \frac{J_2}{2a^3} \left(1 - \frac{3}{2} \sin^2 i \right) (1 - e^2)^{-3/2}, \quad (3.168)$$

$$R_{1s} = \frac{J_2}{2a^3} \left\{ \left(1 - \frac{3}{2} \sin^2 i \right) \left[\left(\frac{a}{r} \right)^3 - (1 - e^2)^{-3/2} \right] + \frac{3}{2} \sin^2 i \cos 2(f + \omega) \right\}, \quad (3.169)$$

where

$$e^2 = \xi^2 + \eta^2, \omega = \arctan\left(\frac{\eta}{\xi}\right), M = \lambda - \arctan\left(\frac{\eta}{\xi}\right), f = f(M(\xi, \eta, \lambda), e(\xi, \eta)), \quad (3.170)$$

Substituting (3.170) into (3.31) leads

$$\dot{\sigma} = f_0(a) + f_1(\sigma, \varepsilon), \quad (3.171)$$

$$f_0(a) = \delta n, \quad n_0 = a_0^{-3/2}, \quad (3.172)$$

$$f_1(\sigma, \varepsilon) = f_{1c}(a, e(\xi, \eta), i) + f_{1s}(a, e(\xi, \eta), i, \omega(\xi, \eta), f(\lambda, \xi, \eta)), \quad (3.173)$$

where f_{1c} and f_{1s} correspond to R_{1c} and R_{1s} , respectively.

It should be noted that for the expressions in the above formulas, $a_l^{(2)}(t)$ in (3.145), and $a_s^{(2)}(t, J_{2,2}; \alpha)$ in (3.149) we use the format of normalized dimensionless units, in which the gravitational constant of the central body $\mu = Gm_0 = 1$. The selection of the normalized dimensionless units is discussed in Chap. 4.

(2) Configuration of a non-singularity perturbation solution

As described in Sect. 3.5.1 when we construct a perturbation solution of the first-order, the difference between the method of quasi-mean elements and the method of mean elements is about the long-period terms (including the short-period terms with commensurable singularities). As a result, the terms in (3.125), (3.126), (3.127) derived by the method of mean elements

$$\sum_j \frac{\partial(f_{1c} + f_{1s})}{\partial \sigma_j} (\sigma_l^{(1)} + \sigma_s^{(1)})_j$$

become

$$\sum_j \frac{\partial(f_{1c} + f_{1s})}{\partial \sigma_j} (\sigma_s^{(1)})_j. \quad (3.174)$$

These terms are the main part of the calculation in the process of constructing a first-order perturbation solution (including the first-order short-period terms, the first- and second-order secular terms, and the first-order long-period terms).

If we use the non-singularity variables and the perturbation Eq. (3.31) directly to construct a solution, the deriving procedure is rather tedious. The question is whether we can use the results given by using the original Kepler orbital elements to obtain the results by using the non-singularity elements. The answer is given as follows.

For any given function $\tilde{F} = \tilde{F}(a, e, i, \Omega, \xi, \eta, \lambda)$ if it is expressed by the original Kepler orbital elements then it becomes $F = F(a, e, i, \Omega, \omega, M)$. Obviously, there is $\tilde{F} = F$. The question is if we can prove the following relationship:

$$\begin{aligned}
\sum_j \frac{\partial \tilde{F}}{\partial \sigma_j} (\sigma_s^{(1)})_j &= \frac{\partial \tilde{F}}{\partial a} a_s^{(1)} + \frac{\partial \tilde{F}}{\partial i} i_s^{(1)} + \frac{\partial \tilde{F}}{\partial \Omega} \Omega_s^{(1)} + \frac{\partial \tilde{F}}{\partial \xi} \xi_s^{(1)} + \frac{\partial \tilde{F}}{\partial \eta} \eta_s^{(1)} + \frac{\partial \tilde{F}}{\partial \lambda} \lambda_s^{(1)} \\
&= \frac{\partial F}{\partial a} a_s^{(1)} + \frac{\partial F}{\partial i} i_s^{(1)} + \frac{\partial F}{\partial \Omega} \Omega_s^{(1)} + \frac{\partial F}{\partial e} e_s^{(1)} + \frac{\partial F}{\partial \omega} \omega_s^{(1)} + \frac{\partial F}{\partial M} M_s^{(1)} \\
&= \sum_j \frac{\partial F}{\partial \sigma_j} (\sigma_s^{(1)})_j
\end{aligned} \tag{3.175}$$

Note that the notation σ in $\tilde{F} = \tilde{F}(a, i, \Omega, \xi, \eta, \lambda)$ and that in $F = F(a, e, i, \Omega, \omega, M)$ represents different variables.

To prove (3.175) we only need to check the last three terms. First by the definitions of the non-singularity variables, ξ , η , and λ in (3.29) we have

$$\begin{cases} \frac{d\xi}{dt} = \cos \omega \frac{de}{dt} - e \sin \omega \frac{d\omega}{dt} \\ \frac{d\eta}{dt} = \sin \omega \frac{de}{dt} + e \cos \omega \frac{d\omega}{dt} \\ \frac{d\lambda}{dt} = \frac{dM}{dt} + \frac{d\omega}{dt} \end{cases} \tag{3.176}$$

which yield

$$\begin{cases} \xi_s^{(1)}(t) = \cos \omega \cdot e_s^{(1)}(t) - e \sin \omega \cdot \omega_s^{(1)}(t) \\ \eta_s^{(1)}(t) = \sin \omega \cdot e_s^{(1)}(t) + e \cos \omega \cdot \omega_s^{(1)}(t) \\ \lambda_s^{(1)}(t) = M_s^{(1)}(t) + \omega_s^{(1)}(t) \end{cases} \tag{3.177}$$

From (3.170) we have

$$\begin{cases} \frac{\partial \tilde{F}}{\partial \xi} = \frac{\partial F}{\partial e} \left(\frac{\xi}{e} \right) + \frac{\partial F}{\partial \omega} \left(-\frac{\eta}{e^2} \right) + \frac{\partial F}{\partial M} \left(\frac{\eta}{e^2} \right) \\ \frac{\partial \tilde{F}}{\partial \eta} = \frac{\partial F}{\partial e} \left(-\frac{\eta}{e} \right) + \frac{\partial F}{\partial \omega} \left(-\frac{\xi}{e^2} \right) + \frac{\partial F}{\partial M} \left(\frac{\xi}{e^2} \right) \\ \frac{\partial \tilde{F}}{\partial \lambda} = \frac{\partial F}{\partial M} \end{cases} \tag{3.178}$$

Substituting (3.176) and (3.178) into the sum of $\sum_j \frac{\partial \tilde{F}}{\partial \sigma_j} (\sigma_s^{(1)})_j$ in (3.175) leads to the conclusion.

$$\sum_j \frac{\partial \tilde{F}}{\partial \sigma_j} (\sigma_s^{(1)})_j = \sum_j \frac{\partial F}{\partial \sigma_j} (\sigma_s^{(1)})_j.$$

Based on this relationship the parts for the four variables, a , i , Ω , and λ , in the sum $\sum_j \frac{\partial \tilde{F}}{\partial \sigma_j} (\sigma_s^{(1)})_j$ can be directly given by the results derived using the original Kepler orbital elements. For the other two variables, ξ and η , by their definitions we have

$$\left(\tilde{f}_{1s} \right)_{\xi} = \cos \omega (f_{1s})_e - e \sin \omega (f_{1s})_{\omega} \quad (3.179)$$

$$\left(\tilde{f}_{1s} \right)_{\eta} = \sin \omega (f_{1s})_e + e \cos \omega (f_{1s})_{\omega} \quad (3.180)$$

$$\begin{aligned} \sum_j \frac{\partial \left(\tilde{f}_{1s} \right)_{\xi}}{\partial \sigma_j} (\sigma_s^{(1)})_j &= \cos \omega \sum_j \frac{\partial (f_{1s})_e}{\partial \sigma_j} (\sigma_s^{(1)})_j - e \sin \omega \sum_j \frac{\partial (f_{1s})_{\omega}}{\partial \sigma_j} (\sigma_s^{(1)})_j \\ &\quad - [\sin \omega (f_{1s})_e + e \cos \omega (f_{1s})_{\omega}] \omega_s^{(1)} - [\sin \omega (f_{1s})_{\omega}] e_s^{(1)} \end{aligned} \quad (3.181)$$

$$\begin{aligned} \sum_j \frac{\partial \left(\tilde{f}_{1s} \right)_{\eta}}{\partial \sigma_j} (\sigma_s^{(1)})_j &= \sin \omega \sum_j \frac{\partial (f_{1s})_e}{\partial \sigma_j} (\sigma_s^{(1)})_j + e \cos \omega \sum_j \frac{\partial (f_{1s})_{\omega}}{\partial \sigma_j} (\sigma_s^{(1)})_j \\ &\quad + [\cos \omega (f_{1s})_e - e \sin \omega (f_{1s})_{\omega}] \omega_s^{(1)} + [\cos \omega (f_{1s})_{\omega}] e_s^{(1)} \end{aligned} \quad (3.182)$$

with the results of the following integrations

$$\frac{1}{2\pi} \int_0^{2\pi} [(f_{1s})_e \omega_s^{(1)} + (f_{1s})_{\omega} e_s^{(1)}] dM = 0 \quad (3.183)$$

$$\frac{1}{2\pi} \int_0^{2\pi} [(f_{1s})_{\omega} \omega_s^{(1)}] dM = 0 \quad (3.184)$$

we have

$$\begin{aligned} &\left(\sum_j \frac{\partial \left(\tilde{f}_{1s} \right)_{\xi}}{\partial \sigma_j} (\sigma_s^{(1)})_j \right)_{c,l} \\ &= \frac{\xi}{e} \left(\sum_j \frac{\partial (f_{1s})_e}{\partial \sigma_j} (\sigma_s^{(1)})_j \right)_{c,l} - \eta \left(\sum_j \frac{\partial (f_{1s})_{\omega}}{\partial \sigma_j} (\sigma_s^{(1)})_j \right)_{c,l} \end{aligned} \quad (3.185)$$

$$\begin{aligned}
& \left(\sum_j \frac{\partial(\tilde{f}_{1s})_\eta}{\partial\sigma_j} (\sigma_s^{(1)})_j \right)_{c,l} \\
&= \frac{\eta}{e} \left(\sum_j \frac{\partial(f_{1s})_e}{\partial\sigma_j} (\sigma_s^{(1)})_j \right)_{c,l} + \xi \left(\sum_j \frac{\partial(f_{1s})_\omega}{\partial\sigma_j} (\sigma_s^{(1)})_j \right)_{c,l}
\end{aligned} \tag{3.186}$$

We now conclude that for ξ and η the related terms in the sum $\sum_j \frac{\partial\tilde{F}}{\partial\sigma_j} (\sigma_s^{(1)})_j$ can also be given directly by the results of the original Kepler orbital elements.

The above results show that for the perturbation of Earth's dynamical form-factor J_2 term the non-singularity solution of the first-order can be given by recombining the solution given by the original Kepler orbital elements. For other perturbing sources (if they have the same order as J_2 term) it requires careful exams, particularly about whether (3.186) is valid.

(3) Expressions of the non-singularity perturbation solution

Sections 3.5.1 and 3.5.2 provide the method of constructing non-singularity solutions of the first order using the method of quasi-mean elements for Earth's dynamical form-factor J_2 term. Compared to the method of mean elements for Kepler orbital elements, there are two points that need further explanation.

① When we use the method of quasi-mean variables for constructing non-singularity solutions the quasi-mean elements, $\bar{\sigma}$, are given by extracting the short-period terms, $\sigma_s^{(1)}(t)$, $\sigma_s^{(2)}(t)$, $\dots$, from the original elements $(a, e, i, \Omega, \omega, M)$. Therefore, during the procedure of constructing the solution, the terms of the second-order given by the original mean element method

$$\sum_j \frac{\partial(f_{1c} + f_{1s})}{\partial\sigma_j} (\sigma_s^{(1)} + \sigma_l^{(1)})_j$$

become

$$\sum_j \frac{\partial(f_{1c} + f_{1s})}{\partial\sigma_j} (\sigma_s^{(1)})_j.$$

The difference between them shows that if we directly accept the perturbation solution given by the original Kepler elements for the non-singularity solution, we must remove the following part

$$\sum_j \frac{\partial(f_{1c} + f_{1s})}{\partial\sigma_j} (\sigma_l^{(1)})_j$$

② The final non-singularity solution can be expressed completely by the non-singularity variables $(a, i, \Omega, \xi, \eta, \lambda)$. The author provided this kind of solution in his previous works published in references [3–5], but the formulas were complicated, which caused difficulties in the corresponding software designs and affected the calculation efficiency. After comparing and analyzing the results, we conclude that from the perspective of application efficiency it is not necessary to keep the seemly perfect formulas, we can retain the Kepler orbital elements in the perturbation solutions according to their rigorous geometric definitions. Therefore, the non-singularity solutions of the first type can be expressed by the following relationships.

The secular terms (including the long-period terms) for gravitational perturbations are constructed as

$$\Delta a(t) = 0, \quad (3.187)$$

$$\Delta i(t) = \Delta i_l(t), \quad (3.188)$$

$$\Delta \Omega(t) = \Delta \Omega_c(t) + \Delta \Omega_l(t), \quad (3.189)$$

$$\Delta \xi(t) = \cos \omega[\Delta e(t)] - \sin \omega[e \Delta \omega(t)], \quad (3.190)$$

$$\Delta \eta(t) = \sin \omega[\Delta e(t)] + \cos \omega[e \Delta \omega(t)], \quad (3.191)$$

$$\Delta \lambda(t) = \bar{n}(t - t_0) + [\Delta M(t) + \Delta \omega(t)], \quad (3.192)$$

where $\bar{n} = \bar{a}^{-3/2} = \bar{a}_0^{-3/2}$, and $\Delta \sigma(t)$ includes the secular and the long-period variation terms.

The expressions of (3.190) and (3.192) are linear, which contain the first-order secular perturbation part of Earth's dynamical form-factor J_2 term. If we need the second-order secular perturbation part of the J_2 term, we still need the “complete” expressions of a non-singularity solution to include the secular part caused by the secular variation in ω which are

$$\Delta \xi(t) = \bar{\xi}_0 \cos \Delta \omega(t) - \bar{\eta}_0 \sin \Delta \omega(t), \quad (3.193)$$

$$\Delta \eta(t) = \bar{\eta}_0 \cos \Delta \omega(t) + \bar{\xi}_0 \sin \Delta \omega(t), \quad (3.194)$$

where the definitions of $\bar{\xi}_0$ and $\bar{\eta}_0$ are the same as given above.

The short-period terms are given by

$$a_s(t) = a_s^{(1)}(t) + a_s^{(2)}(t), \quad (3.195)$$

$$\dot{i}_s^{(1)}(t) = \dot{i}_s^{(1)}(t), \quad (3.196)$$

$$\Omega_s^{(1)}(t) = \Omega_s^{(1)}(t), \quad (3.197)$$

$$\xi_s^{(1)}(t) = \cos \omega [e_s^{(1)}(t)] - \sin \omega [e\omega_s^{(1)}(t)], \quad (3.198)$$

$$\eta_s^{(1)}(t) = \sin \omega [e_s^{(1)}(t)] + \cos \omega [e\omega_s^{(1)}(t)], \quad (3.199)$$

$$\lambda_s^{(1)}(t) = [M_s^{(1)}(t)] + [\omega_s^{(1)}(t)]. \quad (3.200)$$

In the perturbation solutions (3.188)–(3.200) the terms $\Delta e(t)$, $\dots$, $\Delta M(t)$, and $a_s^{(1)}(t)$, $\dots$, $M_s^{(1)}(t)$ are composed of related solutions given by Kepler elements, the actual formulas are provided in the next chapter. By the above method the non-singularity solutions given by recombining the original solutions of Kepler elements no longer have the factor of $(1/e)$.

(4) Calculation of the non-singularity perturbation solutions

In the non-singularity perturbation solutions (3.151), (3.152) obtained by the method of quasi-mean elements, each term can be given by solutions of Kepler elements composed in certain forms, i.e., (3.187)–(3.200). In the procedure of calculation, the Kepler elements act as intermediate variables, they are transferred from the non-singularity variables by (3.170) and are not affected by a small eccentricity e . The corresponding solutions of Kepler elements are given in the next chapter.

3.6.2 Configuration of the Non-singularity Perturbation Solutions of the Second Type

(1) Basic variables and the configuration of the non-singularity solution

In order to solve the problem related to a small e , a small i , and commensurable small divisors, we choose the set of non-singularity variables as defined by (3.36) which are

$$\begin{aligned} a, \quad \xi &= e \cos \tilde{\omega}, \quad \eta = e \sin \tilde{\omega}, \\ h &= \sin \frac{i}{2} \cos \Omega, \quad k = \sin \frac{i}{2} \sin \Omega, \quad \lambda = M + \tilde{\omega}, \end{aligned} \quad (3.201)$$

where $\tilde{\omega} = \omega + \Omega$. By this set of variables, the corresponding perturbation equations are given by (3.39)–(3.44). We can use the same method and the same principle

described in Sect. 3.6.1 for the non-singularity solution of the first type to construct the perturbation solution of the second type, details are not repeated here.

(2) Expressions of the non-singularity perturbation solution of the second type

We still use Earth's non-spherical dynamical form-factor J_2 term as the perturbing source. The method of constructing the perturbation solutions is the same as for the non-singularity variables of the first type, which is to recombine the terms in the solutions derived using Kepler elements. The analytical non-singularity solution still takes the form as

$$\sigma(t) = \bar{\sigma} + \sigma_s^{(1)}(t) + \sigma_s^{(2)}(t) + \cdots, \quad (3.202)$$

where $\bar{\sigma}(t)$ is the quasi-mean elements after eliminating the short-period terms defined by (3.152).

The secular terms (including the long-period terms) in the perturbation solution for the six non-singularity variables $\sigma = (a, \xi, \eta, h, k, \lambda)$ are constructed by

$$\Delta a(t) = 0 \quad (3.203)$$

$$\Delta \xi(t) = \cos \tilde{\omega}[\Delta e(t)] - \sin \tilde{\omega}[e \Delta \omega(t) + e \Delta \Omega(t)] \quad (3.204)$$

$$\Delta \eta(t) = \sin \tilde{\omega}[\Delta e(t)] + \cos \tilde{\omega}[e \Delta \omega(t) + e \Delta \Omega(t)] \quad (3.205)$$

$$\Delta h(t) = \frac{1}{2} \cos \frac{i}{2} \cos \Omega[\Delta i(t)] - \sin \Omega \left[\sin \frac{i}{2} \Delta \Omega(t) \right] \quad (3.206)$$

$$\Delta k(t) = \frac{1}{2} \cos \frac{i}{2} \sin \Omega[\Delta i(t)] + \cos \Omega \left[\sin \frac{i}{2} \Delta \Omega(t) \right] \quad (3.207)$$

$$\Delta \lambda(t) = \bar{n}(t - t_0) + [\Delta M(t) + \Delta \omega(t) + \Delta \Omega(t)] \quad (3.208)$$

where $\bar{n} = \bar{a}^{-3/2} = \bar{a}_0^{-3/2}$, and $\Delta \sigma(t)$ includes the secular terms and the long-period terms. The short-period terms are constructed by

$$a_s(t) = a_s^{(1)}(t) + a_s^{(2)}(t), \quad (3.209)$$

$$\xi_s^{(1)}(t) = \cos \tilde{\omega}[e_s^{(1)}(t)] - \sin \tilde{\omega}[e \omega_s^{(1)}(t) + e \Omega_s^{(1)}(t)], \quad (3.210)$$

$$\eta_s^{(1)}(t) = \sin \tilde{\omega}[e_s^{(1)}(t)] + \cos \tilde{\omega}[e \omega_s^{(1)}(t) + e \Omega_s^{(1)}(t)], \quad (3.211)$$

$$h_s^{(1)}(t) = \frac{1}{2} \cos \frac{i}{2} \cos \Omega [i_s^{(1)}(t)] - \sin \Omega \left[\sin \frac{i}{2} \Omega_s^{(1)}(t) \right], \quad (3.212)$$

$$k_s^{(1)}(t) = \frac{1}{2} \cos \frac{i}{2} \sin \Omega [i_s^{(1)}(t)] + \cos \Omega \left[\sin \frac{i}{2} \Omega_s^{(1)}(t) \right], \quad (3.213)$$

$$\lambda_s^{(1)}(t) = \Omega_s^{(1)}(t) + \omega_s^{(1)}(t) + M_s^{(1)}(t). \quad (3.214)$$

In the above formulas (3.203)–(3.214) the terms of $\Delta e(t)$, $\Delta i(t)$, $\Delta \omega(t)$, $\Delta \Omega(t)$, $\Delta M(t)$, and

$$a_s^{(1)}(t), e_s^{(1)}(t), i_s^{(1)}(t), \omega_s^{(1)}(t), \Omega_s^{(1)}(t), M_s^{(1)}(t),$$

and $a_s^{(2)}(t)$, are all provided by the solutions of Kepler elements. The final solution has neither singularities caused by the critical inclination nor those of $(1/e)$ and $(1/\sin i)$.

As we can see that in the expression (3.200) $\lambda_s^{(1)}(t) = [M_s^{(1)}(t)] + [\omega_s^{(1)}(t)]$, there is no problem with a small e . Similarly, we can give expressions for the terms in $\lambda_s^{(1)}(t)$ (3.214), which no longer has singularities by a small e and a small i , details are given in the next chapter.

Although the above discussion is focused on the potential of Earth's dynamical form-factor J_2 term, the method and the principles can be applied to other perturbation sources. Also when we use the perturbation solution for orbital extrapolation the basic variables must be non-singularity variables, but in the process, the variables are Kepler elements, $a(t)$, $e(t)$, $\dots$ or the mean orbital elements $\bar{a}(t)$, $\bar{e}(t)$, $\dots$ which are transferred from the non-singularity variables (the instantaneous variables $\sigma(t)$ or mean variables $\bar{\sigma}(t)$).

References

1. Smart WM (1953) Celestial mechanics. University of Glasgow
2. Brouwer D, Clemence GM (1961) Methods of Celestial Mechanics. Academic Press, New York and London
3. Liu L (1992) Orbital dynamics of Earth's artificial satellite. Higher Education Press, Beijing
4. Liu L (1998) Methods of celestial mechanics. Nanjing University Press
5. Liu L (2000) Orbital theory of spacecraft. National Defense Industry Press, Beijing
6. Liu L, Tang JS (2015) Orbital theory of satellites and applications. Electronic Industry Press, Beijing
7. Liu L, Hou XY (2018) The basic of orbital theory. Higher Education Press, Beijing
8. Kozai Y (1959) The motion of a close Earth satellite. Astron J 64(9):367–377
9. Liu L (1974) A solution of the motion of an artificial satellite in the vicinity of the critical inclination. ACTA Astronomica Sinica 15(2):230–240, and China Astron. Astrophys (1977) 1(1):31–42
10. Liu L (1975) A method of calculating the perturbation of artificial satellites. ACTA Astronomica Sinica 16(1):5–80, and China Astron Astrophys 1977 1(1):63–78

11. Garfinkel B (1966) Formal solution in the problem of small divisors. *Astron J* 71(8):657–669
12. Garfinkel B, Jupp AH, Williams CA (1971) A recursive Von Zeipel algorithm for the ideal resonance problem. *Astron J* 76(2):157–166

Chapter 4

Analytical Non-singularity Perturbation Solutions for Extrapolation of Earth's Satellite Orbital Motion



4.1 The Complete Dynamic Model of Earth's Satellite Motion

In the Introduction of this book, we describe that in the Solar System all motions of major planets and asteroids, and motions of natural and artificial satellites (such as Earth's artificial satellites, the Moon's satellites, Mars's satellites, and other orbiting spacecraft) have only one main external force. For an Earth's artificial satellite, the main external force is the central gravity attraction of Earth which is assumed to be a sphere of evenly distributed mass. The other external forces by comparison are much smaller and can be treated as perturbations. Earth and a satellite, therefore, form a perturbed two-body system, and the corresponding mathematical problem is the perturbed two-body problem. Earth as the source of the main external force is called the "central body", denoted by P_0 , and its mass is denoted by m_0 . The Earth's satellite is denoted by p , and its mass is m . The object of our research is the orbital motion of the artificial satellite determined by the central body and a few perturbing forces.

The perturbed two-body orbital motion can be presented by an ordinary differential equation with initial conditions, which are

$$\begin{cases} \ddot{\vec{r}} = -\frac{G(m_0+m)}{r^3}\vec{r} + \sum_{i=1}^N \vec{F}_i, \\ \vec{r}(t_0) = \vec{r}_0, \quad \dot{\vec{r}}(t_0) = \dot{\vec{r}}_0, \end{cases} \quad (4.1)$$

where G is the universal gravitational constant, and $\vec{F}_i$ ($i = 1, 2, \dots, N$) are the perturbation accelerations. The origin of the coordinate system is at the barycenter of the central body P_0 , and $\vec{r} = \vec{r}(x, y, z)$ is the position vector of a small moving body in the frame.

By convention, the symbol μ is defined by

$$\mu = G(m_0 + m). \quad (4.2)$$

For the motion of the small artificial body (including different circling spacecraft) p , its mass can be omitted, i.e., $m = 0$, then the motion Eq. (4.1) has the form as

$$\ddot{\vec{r}} = -\frac{\mu}{r^3}\vec{r} + \sum_{i=1}^N \vec{F}_i, \quad (4.3)$$

where $\mu = Gm_0$ is the gravitational constant of the central body. For an Earth's artificial satellite, the central celestial body is Earth then $\mu = Gm_0 = GE$, where GE is Earth's gravitational constant.

4.1.1 Selection of Calculation Units in Satellite Orbit Dynamics

There are three basic calculation units related to a dynamical problem, which are the unit of length $[L]$, the unit of mass $[M]$, and the unit of time $[T]$. For orbital dynamics of satellite motion, the three units are defined by the following references,

$$\begin{aligned} [L] &= a_e \text{ (the equatorial radius of the central body's reference ellipsoid),} \\ [M] &= m_0 \text{ (the mass of the central body),} \\ [T] &= \left(\frac{a_e^3}{Gm_0} \right)^{\frac{1}{2}}. \end{aligned} \quad (4.4)$$

The unit of time $[T]$ is a derived unit in order to make the gravitational constant $G = 1$, then we can have $\mu = Gm_0 = 1$. By these definitions, we have the normalized units for all physical variables. It is convenient to estimate the order of magnitude of a variable and to compare related variables using the normalized units, at the same time the expressions of formulas can be simplified.

In the satellite orbital dynamics, we also deal with orbital motions of major planets around the Sun. For the Sun-Earth system, the unit of length $[L]$ is the average distance between the Sun and Earth, i.e., the Astronomical unit, AU, and the corresponding time unit now is decided by $[T] = \{1 \text{ AU}^3/[G(m_0 + m)]\}^{1/2}$, where m is Earth's mass; whereas for the Sun – (Earth + Moon) system m is the sum of Earth's mass and the Moon's mass. The gravitational constants for the Sun, Earth, Mars, Venus, and the Moon and other related quantities are given as follows.

The gravitational constants of the Sun, Earth, Mars, Venus, and the Moon are:

$$\begin{aligned} GS &= 1.32712440041 \times 10^{11} \text{ (km}^3/\text{s}^2\text{)}, \\ GE &= 398600.4418 \text{ (km}^3/\text{s}^2\text{)}, \quad \text{in WGS84 system,} \\ GM &= 42828.3719 \text{ (km}^3/\text{s}^2\text{)}, \quad \text{in GMM-2B system,} \end{aligned}$$

$$\begin{aligned}
 GV &= 324858.64 \text{ (km}^3/\text{s}^2\text{)}, & \text{in GVM-1 system,} \\
 GL &= 4902.800269 \text{ (km}^3/\text{s}^2\text{)}, & \text{in LP75G system.}
 \end{aligned} \tag{4.5}$$

For the Sun-Earth system, the time unit is

$$[T] = 58^{\text{d}}.1323535673601. \tag{4.6a}$$

For the Sun – (Earth + Moon) system, the time unit is

$$[T] = 58^{\text{d}}.13235249375701. \tag{4.6b}$$

For Earth in WGS84 system (World Geodetic System), there are (note that GM in (4.7) is GE in (4.5))

$$\begin{aligned}
 a_e &= 6378.137 \text{ (km)}, \\
 GM &= 398600.4418 \text{ (km}^3/\text{s}^2\text{)}, \\
 [T] &= 13^{\text{m}}.4468520637382.
 \end{aligned} \tag{4.7}$$

The corresponding Earth's flattening factor $\varepsilon = 0.00335281$, and the dynamical form-factor (which includes the shape and the mass distribution of Earth) $J_2 = 1.082636022 \times 10^{-3}$.

For Mars in the American Goddard Mars Model (GMM-2B), there are

$$\begin{aligned}
 a_e &= 3397.0 \text{ (km)}, \\
 GM &= 42828.3719 \text{ (km}^3/\text{s}^2\text{)}, \\
 [T] &= 15^{\text{m}}.945064755181.
 \end{aligned} \tag{4.8}$$

The corresponding Mars's flattening factor $\varepsilon = 0.005231844$, and the dynamical form-factor $J_2 = 1.955453679 \times 10^{-3}$.

For Venus in the American and JPLd11993 Venus Model (GVM-1), there are (note that here GM is GV in (4.5))

$$\begin{aligned}
 a_e &= 6051.8130 \text{ (km)}, \\
 GM &= 324858.64 \text{ (km}^3/\text{s}^2\text{)}, \\
 [T] &= 13^{\text{m}}.766698101341069.
 \end{aligned} \tag{4.9}$$

The corresponding Venus's flattening factor $\varepsilon = 0.0$, and the dynamical form-factor $J_2 = 4.45749887 \times 10^{-6}$.

For the Moon in the American JPL gravity model LP75G, there are (note that GM in (4.10) is GL in (4.5))

$$a_e = 1738.0 \text{ (km)},$$

$$\begin{aligned} GM &= 4902.800269 \text{ (km}^3/\text{s}^2\text{)}, \\ [T] &= 17^{\text{m}}.246513279967907. \end{aligned} \quad (4.10)$$

The corresponding Moon's flattening factor $\varepsilon = 0.000178366$, and the dynamical form-factor $J_2 = 2.034284544 \times 10^{-4}$.

4.1.2 Analyses of Forces on Satellite's Orbital Motion

We take the motion of an Earth's artificial satellite as an example. The main external force in the motion Eq. (4.3) denoted by $\vec{F}_0$ is given by

$$\vec{F}_0 = -\frac{\mu}{r^3} \vec{r} = -\frac{1}{r^3} \vec{r}. \quad (4.11)$$

The related acceleration denoted by g due to the attraction of Earth (as a particle) is given by

$$g = \frac{GE}{r^2} = \frac{1}{r^2}. \quad (4.12)$$

The values of g (m/s²) at altitudes 200, 300, 600, and 1200 km are

$$g = \begin{cases} 9.798285, & h = 0.0 \text{ km}, \\ 9.211534, & h = 200.0 \text{ km}, \\ 8.937728, & h = 300.0 \text{ km}, \\ 8.185756, & h = 600.0 \text{ km}, \\ 7.537636, & h = 1200.0 \text{ km}. \end{cases} \quad (4.13)$$

These altitudes are for a low Earth orbit satellite. For this kind of satellite, there are other external perturbing forces $\vec{F}_i (i = 1, N)$ including attractions from Earth's non-spherical part and a third body as a particle, Earth's deformation, atmospheric drag, the radiation pressure of sunlight, and post-Newtonian effect, etc. We give some simple quantitative analyses about these perturbing forces (with respect to the central gravitational force) as follows.

- (1) The perturbing magnitude of the gravity potential due to Earth's non-spherical dynamical form-factor J_2 term

For a low Earth orbit satellite, the largest perturbation is from the gravity force of Earth's dynamical form-factor J_2 term. The corresponding perturbation acceleration is $F_1(J_2)$ on the right side of the motion Eq. (4.3). In the altitude range given in (4.13), the relative magnitude of the J_2 term acceleration compared to the gravitational acceleration due to Earth's mass as a particle is

$$\varepsilon_1 = \frac{F_1}{g} = O\left(\frac{J_2}{r^2}\right) = O(10^{-3}). \quad (4.14)$$

- (2) The perturbing magnitude of the gravity potential due to Earth's ellipticity $J_{2,2}$ term

$$\varepsilon_2 = \frac{F_2}{g} = O\left(\frac{J_{2,2}}{r^2}\right) = O(10^{-6}). \quad (4.15)$$

- (3) The perturbing magnitude of the gravity potential due to Earth's non-spherical high-order zonal harmonic terms $J_l (l \geq 3)$

$$\varepsilon_3 = \frac{F_3}{g} = O\left(\frac{J_l}{r^l}\right) = O(10^{-6}). \quad (4.16)$$

- (4) The perturbing magnitude of the gravity potential due to Earth's non-spherical high-order tesseral harmonic terms $J_{l,m} (l \geq 3, m = 1 - l)$

$$\varepsilon_4 = \frac{F_4}{g} = O\left(\frac{J_{l,m}}{r^l}\right) = O(10^{-6}). \quad (4.17)$$

- (5) The perturbing magnitude of the gravity potential of the Sun or the Moon

$$\varepsilon_5 = \frac{F_5}{g} = m' \left(\frac{r}{r'}\right)^3 = O(10^{-7}), \quad (4.18)$$

where m' and r' are the mass and the average distance to Earth's center in normalized units, respectively, for the Sun or the Moon. The magnitudes of the perturbation forces of the Sun and the Moon are in the same order.

- (6) The perturbing magnitude of the force due to Earth's tidal deformation

$$\varepsilon_6 = \frac{F_6}{g} = O\left(\left(\frac{k_2}{r^3}\right)m' \left(\frac{r}{r'}\right)^3\right) = O\left(\left(\frac{k_2}{r^3}\right)\varepsilon_5\right) = O(10^{-8}), \quad (4.19)$$

where the parameter k_2 is the Love number of the second-order for Earth's tidal deformation, it is about 0.30.

- (7) The perturbing magnitude of the radiation pressure of sunlight

Both the radiation pressure of sunlight and the atmospheric drag act on a surface, so the effects of these perturbations are related to the efficient area-to-mass ratio of a satellite. In the calculation of the Earth's system, by the normalized units, we have

$$1 \text{ (m}^2/\text{kg)} = 1.4686 \times 10^{11}. \quad (4.20)$$

Usually, the area-to-mass ratio of a satellite is about $1 \text{ m}^2/100 \text{ kg}$, which is 10^9 by the normalized units, then the magnitude of the radiation pressure of sunlight on a satellite is about

$$\varepsilon_7 = \frac{F_7}{g} = \kappa \left(\frac{S}{m} \right) \rho_{\odot} r^2 = O(10^{-8} - 10^{-7}), \quad (4.21)$$

where $\kappa = 1 + \eta$, η is the reflection coefficient of the satellite's surface, $\eta = 1$ corresponds to total reflection, and $\eta = 0$ to complete absorption, usually $0 < \eta < 1$; S/m is the satellite's efficient area-to-mass ratio, and $\rho_{\odot}$ is the intensity of the radiation pressure near Earth, that

$$\rho_{\odot} = 4.5606 \times 10^{-6} \text{ (N m}^{-2}\text{)} = 0.3169 \times 10^{-17}. \quad (4.22)$$

The last number is the normalized dimensionless value.

(8) The perturbing magnitude of the atmospheric drag

In the atmosphere above the altitude of $h = 200 \text{ km}$, the acceleration due to the atmospheric drag is often expressed as

$$D = \frac{1}{2} \left(C_D \frac{S}{M} \right) \rho v^2. \quad (4.23)$$

The perturbing magnitude of the corresponding atmospheric drag is

$$\varepsilon_8 = \frac{F_8}{g} = \frac{1}{2} \left(\frac{C_D S}{m} \right) \rho v^2 / \left(\frac{1}{r^2} \right) \approx \left(\frac{C_D S}{m} \rho \right) \frac{r}{2}, \quad (4.24)$$

where the coefficient of the atmospheric drag C_D takes a value of 2.2 at altitudes above $h = 200 \text{ km}$. Similar to the radiation pressure perturbation we assume the satellite's equivalent area-to-mass ratio to be 10^9 . Usually, the unit of atmospheric density is (kg/m^3) which can be transferred to the normalized unit that

$$1 \text{ (kg}/\text{m}^3\text{)} = 0.4343 \times 10^{-4}. \quad (4.25)$$

According to the International standard atmosphere model, in a normal condition, the atmospheric densities at altitudes 200, 400, and 600 km are

$$\bar{\rho} \text{ (kg}/\text{m}^3\text{)} = \begin{cases} 0.291 \times 10^{-9}, & h = 200.0 \text{ km}, \\ 0.428 \times 10^{-11}, & h = 400.0 \text{ km}, \\ 0.186 \times 10^{-12}, & h = 600.0 \text{ km}. \end{cases} \quad (4.26)$$

Then the corresponding accelerations due to the atmospheric drag, $D \text{ (m/s}^2\text{)}$, at these altitudes are

$$D = \begin{cases} 1.736 \times 10^{-4}, & h = 200.0 \text{ km}, \\ 2.517 \times 10^{-6}, & h = 400.0 \text{ km}, \\ 1.062 \times 10^{-7}, & h = 600.0 \text{ km}. \end{cases} \quad (4.27)$$

The corresponding perturbing magnitudes of the atmospheric drag are given by

$$\varepsilon_8 = \frac{F_8}{g} = \begin{cases} 1.9 \times 10^{-5}, & h = 200.0 \text{ km}, \\ 2.9 \times 10^{-7}, & h = 400.0 \text{ km}, \\ 1.3 \times 10^{-8}, & h = 600.0 \text{ km}. \end{cases} \quad (4.28)$$

(9) The perturbing magnitudes of the post-Newtonian effect

In the orbital motion of Earth's satellite, the post-Newtonian effect appears mainly in the Schwarzschild solution, also in the geodetic precession, Earth's rotation, and Earth's dynamical form-factor J_2 term. The corresponding post-Newtonian accelerations $\vec{A}_{PN}$ can be written as

$$\vec{A}_{PN} = \vec{A}_1 + \vec{A}_2 + \vec{A}_3 + \vec{A}_4.$$

We do not discuss the actual expressions of these effects here. For a low Earth orbit satellite, the perturbing accelerations of the four post-Newtonian effects relative to the central acceleration are

$$10^{-9}, 10^{-11}, 10^{-12}, 10^{-12}.$$

Therefore, the perturbing magnitude of the post-Newtonian effect for a low Earth orbit satellite is

$$\varepsilon_9 = \frac{F_9}{g} = O(10^{-9}). \quad (4.29)$$

In summary, for a model including complete forces the orbital motion equation of a low Earth orbit satellite (4.3) has the following form

$$\ddot{\vec{r}} = -\frac{\vec{r}}{r^3} + \sum_{i=1}^9 \vec{F}_i. \quad (4.30)$$

Using Kepler orbital elements σ ($a, e, i, \Omega, \omega, M$) as the state variables, the corresponding orbital perturbation equation system is

$$\frac{d\sigma}{dt} = f_0(a) + \sum_{j=1}^9 f_j(\sigma, t, \varepsilon), \quad (4.31)$$

where

$$\begin{cases} f_0(a) = \delta n, & n = a^{-\frac{3}{2}}, \\ \delta = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = (0 \ 0 \ 0 \ 0 \ 0 \ 1)^T, \end{cases} \quad (4.32)$$

$$|f_j(\sigma, t, \varepsilon)| = O(\varepsilon), \quad j = 1, 2, \dots, 9. \quad (4.33)$$

If we arrange f_j according to the above analysis of the perturbing magnitude, then we have $f_1 = O(\varepsilon)$, $f_2 = O(\varepsilon^2)$, $f_3 = O(\varepsilon^3)$, $\dots$, and the perturbation motion equation system (4.31) has the form as

$$\dot{\sigma} = f_0(a) + f_1(\sigma; J_2) + f_2(\sigma, t; \varepsilon^2) + f_3(\sigma, t; \varepsilon^3). \quad (4.34)$$

Based on the motion of a low Earth orbit satellite, from the given perturbing magnitudes it is clear that Earth's non-spherical form-factor J_2 term has the largest magnitude, thus we make 10^{-3} as the magnitude of the small perturbation parameter ε , i.e., the first-order quantity, and all the other perturbing terms have quantities of the second-order, third-order, or higher-order. For further understanding of the influences of these perturbation factors on the satellite's orbital motion, we may treat them as terms of the second-order of ε . In reality, the question of how many individual perturbations should be included depends on each situation, but for the theoretical analysis it makes no difference, and usually, the perturbation equation system of orbital motion is given in the form as

$$\begin{cases} \dot{\sigma} = f_0(a) + f_1(\sigma; \varepsilon) + f_2(\sigma, t; \varepsilon^2), \\ \varepsilon = O(J_2). \end{cases} \quad (4.35)$$

4.1.3 Further Analyses of the Forces Acting on a Satellite

The above analysis is based on a low Earth orbit satellite, the analytical method of perturbation forces and details of treatment are also suitable for a medium or a high Earth orbit satellite, as well as other celestial bodies (the Moon, Mars, natural satellites, asteroids, etc.), and all sorts of circling spacecraft.

For a high Earth orbit satellite, the magnitude of Earth's dynamical form-factor term J_2 decreases but the gravitational forces of the Sun and the Moon increase, and the influence of atmosphere drag reduces or disappears. These variations change the magnitudes of the perturbations and influence the decision of which forces to be included in the motion equation, but have no fundamental effect on the structure of the perturbation solution. We do not repeat the analysis for other types of the satellite

orbit, the perturbation motion equation system can always be given in the following form

$$\dot{\sigma} = f_0(a) + f_1(\sigma; \varepsilon) + f_2(\sigma, t; \varepsilon^2) + \cdots, \quad (4.36)$$

How to choose the small parameter ε depends on the actual dynamic situation. For example, in the process of constructing the perturbation solution for a low Moon orbit satellite, the small parameter ε is not the Moon's dynamical form-factor J_2 term but artificially decided that

$$\varepsilon = 10^{-2}. \quad (4.37)$$

Details about the selection of a small parameter is discussed in the subsequent chapters.

4.2 The Perturbed Orbit Solution of the First-Order Due to Earth's Dynamical Form-Factor J_2 Term

4.2.1 The Perturbed Orbit Solution of the First-Order in Kepler Orbital Elements [1–5]

In the geocentric equatorial coordinate system, the gravitational acceleration $\vec{F}$ of a satellite by the non-spherical Earth's dynamical form-factor J_2 term is presented by the gradient of the potential $V(\vec{r})$ that

$$\begin{aligned} \vec{F} &= \text{grad } V(\vec{r}), \\ V(\vec{r}) &= \frac{GE}{r} \left[1 - J_2 \left(\frac{a_e}{r} \right)^2 P_2(\sin \varphi) \right], \end{aligned} \quad (4.38)$$

where $\vec{r}$ is the position vector of a satellite; GE and a_e are Earth's gravitational constant and the radius of Earth's equator, respectively; r is the distance between the satellite and Earth's center, and φ is the latitude of the satellite. Note that the perturbation solution of the satellite's motion should be presented in J2000.0 geocentric celestial coordinate system, the geocentric equatorial coordinate system we use now for (4.38) is an **Earth-fixed coordinate system**. The difference between the two coordinate systems is discussed in detail in Sect. 4.4.

In (4.38), $P_2(\sin \varphi)$ is Legendre polynomial of the second-order, that

$$P_2(\sin \varphi) = \frac{3}{2} \sin^2 \varphi - \frac{1}{2}, \quad (4.39)$$

and the coefficient J_2 is Earth's dynamical form-factor, its value reflects the deviation of Earth's shape and mass distribution from a mass evenly distributed sphere, usually, it has the same magnitude as Earth's geometric flattening factor (only related to Earth's shape), therefore J_2 is the small parameter ε of perturbation, that

$$J_2 = O(10^{-3}). \quad (4.40)$$

If we use the normalized calculation units described in Sect. 4.1.1, by which Earth's mass E is the unit of mass, and Earth's equatorial radius a_e is the unit of length, and $(a_e^3/G E)^{1/2}$ is the unit of time, then the gravitational constant $G = 1$ and $\mu = GE = 1$. In this unit system, all variables are dimensionless therefore it is easy for formula expressions and quantitative analyses. Using the normalized units, the potential in (4.38) can be written as

$$V(\vec{r}) = V_0 + V_1, \quad (4.41)$$

$$V_0 = \frac{1}{r}, \quad V_1 = -\frac{J_2}{r^3} \left(\frac{3}{2} \sin^2 \varphi - \frac{1}{2} \right). \quad (4.42)$$

Usually, V_1 is the perturbation function and is denoted by R , i.e., $V_1 = R$. Obviously, this perturbed system is a conservative system.

4.2.1.1 Decomposition of the Perturbation Function R

Substituting the formula of spherical trigonometry $\sin \varphi = \sin i \sin(f + \omega)$ into (4.42) we have

$$R = \frac{J_2}{2a^3} \left(\frac{a}{r} \right)^3 \left[\left(1 - \frac{3}{2} \sin^2 i \right) + \frac{3}{2} \sin^2 i \cos 2(f + \omega) \right]. \quad (4.43)$$

Then using the flowing averages

$$\overline{\left(\frac{a}{r} \right)^3} = (1 - e^2)^{-3/2}, \quad \overline{\left(\frac{a}{r} \right)^3 \cos 2f} = 0, \quad \overline{\left(\frac{a}{r} \right)^3 \sin 2f} = 0, \quad (4.44)$$

we can decompose R into the secular part R_{1c} and the periodic part R_{1s} that

$$R_{1c} = \frac{J_2}{2a^3} \left(1 - \frac{3}{2} \sin^2 i \right) (1 - e^2)^{-3/2} \quad (4.45)$$

$$R_{1s} = \frac{J_2}{2a^3} \left\{ \left(1 - \frac{3}{2} \sin^2 i \right) \left[\left(\frac{a}{r} \right)^3 - (1 - e^2)^{-3/2} \right] + \frac{3}{2} \left(\frac{a}{r} \right)^3 \sin^2 i \cos 2(f + \omega) \right\} \quad (4.46)$$

4.2.1.2 The Perturbation Motion Equation System

Substituting R of (4.43) into the $(\partial R/\partial \sigma)$ -version perturbation motion equation system gives

$$\dot{\sigma} = f_0(a) + f_{1c}(a, e, i) + f_{1s}(a, e, i, \omega, M), \quad (4.47)$$

$$f_0(a) = \delta n, \quad n = a^{-3/2}, \quad (4.48)$$

$$\delta = (0 \ 0 \ 0 \ 0 \ 0 \ 1)^T, \quad (4.49)$$

where σ is the set of Kepler orbital elements, i.e.,

$$\sigma = (a \ e \ i \ \Omega \ \omega \ M)^T. \quad (4.50)$$

The expressions of each component of f_{1c} and f_{1s} are given by

$$(f_{1c})_a = 0, \quad (4.51)$$

$$(f_{1c})_e = 0, \quad (4.52)$$

$$(f_{1c})_i = 0, \quad (4.53)$$

$$(f_{1c})_\Omega = -\frac{3J_2}{2p^2}n \cos i, \quad (4.54)$$

$$(f_{1c})_\omega = \frac{3J_2}{2p^2}n \left(2 - \frac{5}{2} \sin^2 i\right), \quad (4.55)$$

$$(f_{1c})_M = \frac{3J_2}{2p^2}n \left(1 - \frac{3}{2} \sin^2 i\right) \sqrt{1 - e^2}, \quad (4.56)$$

$$\begin{aligned} (f_{1s})_a &= \frac{3J_2}{a}n\sqrt{1 - e^2}\left(\frac{a}{r}\right)^4 \left\{ -\frac{e \sin f}{1 - e^2} \right. \\ &\quad \times \left[\left(1 - \frac{3}{2} \sin^2 i\right) + \frac{3}{2} \sin^2 i \cos 2(f + \omega) \right] - \sin^2 i \left(\frac{a}{r}\right) \sin 2(f + \omega) \Big\}, \end{aligned} \quad (4.57)$$

$$\begin{aligned} (f_{1s})_e &= \frac{3J_2}{2a^2e}n\sqrt{1 - e^2}\left(\frac{a}{r}\right)^4 \left\{ -e \sin f \left[\left(1 - \frac{3}{2} \sin^2 i\right) + \frac{3}{2} \sin^2 i \cos 2(f + \omega) \right] \right. \\ &\quad \left. - (1 - e^2) \sin^2 i \left(\frac{a}{r}\right) \sin 2(f + \omega) + \sin^2 i \left(\frac{r}{a}\right) \sin 2(f + \omega) \right\} \end{aligned} \quad (4.58)$$

$$(f_{1s})_i = -\frac{3J_2}{2a^2\sqrt{1-e^2}}n \sin i \cos i \left(\frac{a}{r}\right)^3 \sin 2(f + \omega), \quad (4.59)$$

$$(f_{1s})_\Omega = -\frac{3J_2}{2a^2\sqrt{1-e^2}}n \cos i \left\{ \left[\left(\frac{a}{r}\right)^3 - (1-e^2)^{-3/2} \right] - \left(\frac{a}{r}\right)^3 \cos 2(f + \omega) \right\}, \quad (4.60)$$

$$\begin{aligned} (f_{1s})_\omega = & -\frac{3J_2}{2a^2\sqrt{1-e^2}}n \cos^2 i \left\{ \left[\left(\frac{a}{r}\right)^3 - (1-e^2)^{-3/2} \right] - \left(\frac{a}{r}\right)^3 \cos 2(f + \omega) \right\} \\ & + \frac{3J_2}{2a^2e}n \sqrt{1-e^2} \left\{ \left(1 - \frac{3}{2} \sin^2 i\right) \left[\left(\frac{a}{r}\right)^4 \cos f - e(1-e^2)^{-5/2} \right] \right. \\ & + \frac{3}{2} \sin^2 i \left(\frac{a}{r}\right)^4 \cos f \cos 2(f + \omega) \\ & \left. - \frac{\sin^2 i}{1-e^2} \left(\frac{a}{r}\right)^3 (2 + e \cos f) \sin f \sin 2(f + \omega) \right\}, \end{aligned} \quad (4.61)$$

$$\begin{aligned} (f_{1s})_M = & \left(-\frac{3n}{2a}\right)a_s^{(1)}(t) + \frac{3J_2}{a^2}n \left\{ \left(1 - \frac{3}{2} \sin^2 i\right) \left[\left(\frac{a}{r}\right)^3 - (1-e^2)^{-3/2} \right] + \frac{3}{2} \sin^2 i \left(\frac{a}{r}\right)^3 \cos 2(f + \omega) \right\} \\ & - \frac{3J_2}{2a^2e}n(1-e^2) \left\{ \left(1 - \frac{3}{2} \sin^2 i\right) \left[\left(\frac{a}{r}\right)^4 \cos f - e(1-e^2)^{-5/2} \right] + \frac{3}{2} \sin^2 i \left(\frac{a}{r}\right)^4 \cos f \cos 2(f + \omega) \right. \\ & \left. - \frac{\sin^2 i}{1-e^2} \left(\frac{a}{r}\right)^3 (2 + e \cos f) \sin f \sin 2(f + \omega) \right\}. \end{aligned} \quad (4.62)$$

4.2.1.3 The Perturbation Solution

The power series solution of a small parameter can be constructed effectively using the method of mean elements as

$$\sigma(t) = \bar{\sigma}_0 + \delta \bar{n}(t - t_0) + (\sigma_1 + \sigma_2 + \dots)(t - t_0) + \sigma_l^{(1)}(t) + \sigma_s^{(2)}(t) + \dots \quad (4.63)$$

(1) The first-order secular terms $\sigma_1(t - t_0)$

$$a_1(t - t_0) = 0, \quad e_1(t - t_0) = 0, \quad i_1(t - t_0) = 0, \quad (4.64)$$

$$\Omega_1(t - t_0) = -\frac{3J_2}{2p^2} \cos in(t - t_0), \quad (4.65)$$

$$\omega_1(t - t_0) = \frac{3J_2}{2p^2} \left(2 - \frac{5}{2} \sin^2 i\right) n(t - t_0), \quad (4.66)$$

$$M_1(t - t_0) = \frac{3J_2}{2p^2} \left(1 - \frac{3}{2} \sin^2 i\right) \sqrt{1-e^2} n(t - t_0), \quad (4.67)$$

where $p = a(1 - e^2)$. The elements a , e , and i on the right sides of the above formulas are mean elements, that

$$\bar{a} = \bar{a}_0, \quad \bar{e} = \bar{e}_0, \quad \bar{i} = \bar{i}_0. \quad (4.68)$$

(2) The first-order short-period terms $\sigma_s^{(1)}(t)$ and $a_s^{(2)}(t)$

$$a_s^{(1)}(t) = \frac{3J_2}{2a} \left\{ \frac{2}{3} \left(1 - \frac{3}{2} \sin^2 i \right) \left[\left(\frac{a}{r} \right)^3 - (1 - e^2)^{-3/2} \right] + \sin^2 i \left(\frac{a}{r} \right)^3 \cos 2(f + \omega) \right\}, \quad (4.69)$$

$$\begin{aligned} e_s^{(1)}(t) &= \frac{1 - e^2}{e} \left[\frac{1}{2a} a_s^{(1)}(t) - (\tan i) i_s^{(1)}(t) \right] \\ &= \left(\frac{J_2}{2a^2} \right) \left(\frac{1 - e^2}{e} \right) \left\{ \left(1 - \frac{3}{2} \sin^2 i \right) \left[\left(\frac{a}{r} \right)^3 - (1 - e^2)^{-3/2} \right] + \frac{3}{2} \sin^2 i \left(\frac{a}{r} \right)^3 \cos 2(f + \omega) \right. \\ &\quad \left. - \frac{3}{2} \sin^2 i (1 - e^2)^{-2} \left[e \cos(f + 2\omega) + \cos 2(f + \omega) + \frac{1}{3} e \cos(3f + 2\omega) \right] \right\} \\ &= \left(\frac{3J_2}{2p^2} \right) \left\{ \frac{1}{3} \left(1 - \frac{3}{2} \sin^2 i \right) \left[e \left(\frac{1}{1 + \sqrt{1 - e^2}} + \sqrt{1 - e^2} \right) + \cos f (3(1 + e \cos f) + (e \cos f)^2) \right] \right. \\ &\quad \left. + \frac{1}{2} \sin^2 i \left[(e + \cos f (3(1 + e \cos f) + (e \cos f)^2)) \cos 2(f + \omega) \right. \right. \\ &\quad \left. \left. - (1 - e^2) \left(\cos(f + 2\omega) + \frac{1}{3} \cos(3f + 2\omega) \right) \right] \right\} \end{aligned} \quad (4.70)$$

$$i_s^{(1)}(t) = \left(\frac{3J_2}{2p^2} \right) \sin 2i \left\{ \frac{e}{4} \cos(f + 2\omega) + \frac{1}{4} \cos(2f + 2\omega) + \frac{e}{12} \cos(3f + 2\omega) \right\}, \quad (4.71)$$

$$\begin{aligned} \Omega_s^{(1)}(t) &= - \left(\frac{3J_2}{2p^2} \right) \cos i \{ (f - M + e \sin f) \\ &\quad - \frac{1}{2} [e \sin(f + 2\omega) + \sin(2f + 2\omega) + \frac{e}{3} \sin(3f + 2\omega)] \}, \end{aligned} \quad (4.72)$$

$$\begin{aligned} \omega_s^{(1)}(t) &= - \cos i \Omega_s^{(1)}(t) + [\omega_s^{(1)}(t)]_1 \\ &\quad [\omega_s^{(1)}(t)]_1 = \frac{1}{e} \left(\frac{3J_2}{2p^2} \right) \\ &\quad \times \left\{ \left(1 - \frac{3}{2} \sin^2 i \right) \left[(f - M + e \sin f) e + \left(1 - \frac{e^2}{4} \right) \sin f + \frac{e}{2} \sin 2f + \frac{e^2}{12} \sin 3f \right] \right. \\ &\quad \left. + \sin^2 i \left[- \left(\frac{1}{4} - \frac{7}{16} e^2 \right) \sin(f + 2\omega) + \frac{3}{4} e \sin 2(f + \omega) \right. \right. \\ &\quad \left. \left. + \left(\frac{7}{12} + \frac{11}{48} e^2 \right) \sin(3f + 2\omega) + \frac{3}{8} e \sin(4f + 2\omega) \right. \right. \\ &\quad \left. \left. + \frac{e^2}{16} (\sin(5f + 2\omega) + \sin(f - 2\omega)) \right] \right\}. \end{aligned} \quad (4.73a)$$

The complete first-order short-period term of ω is

$$\begin{aligned}
\omega_s^{(1)}(t) = & \frac{1}{e} \left(\frac{3J_2}{2p^2} \right) \left\{ \left(2 - \frac{5}{2} \sin^2 i \right) (f - M + e \sin f) e \right. \\
& + \left(1 - \frac{3}{2} \sin^2 i \right) \left[\left(1 - \frac{e^2}{4} \right) \sin f + \frac{e}{2} \sin 2f + \frac{e^2}{12} \sin 3f \right] \\
& - \frac{e}{2} \left[e \sin(f + 2\omega) + \sin 2(f + \omega) + \frac{e}{3} \sin(3f + 2\omega) \right] \\
& + \sin^2 i \left[\frac{e^2}{16} \sin(f - 2\omega) - \left(\frac{1}{4} - \frac{15}{16} e^2 \right) \sin(f + 2\omega) + \frac{5}{4} e \sin 2(f + \omega) \right. \\
& \left. \left. + \left(\frac{7}{12} + \frac{19}{48} e^2 \right) \sin(3f + 2\omega) + \frac{3}{8} e \sin(4f + 2\omega) + \frac{e^2}{16} \sin(5f + 2\omega) \right] \right\}, \\
& (4.73b)
\end{aligned}$$

$$\begin{aligned}
M_s^{(1)}(t) = & -\sqrt{1 - e^2} [\omega_s^{(1)}(t)]_1 \\
& + \left(\frac{3J_2}{2p^2} \right) \sqrt{1 - e^2} \left\{ \left(1 - \frac{3}{2} \sin^2 i \right) (f - M + e \sin f) \right. \\
& \left. + \sin^2 i \left[\frac{3}{4} e \sin(f + 2\omega) + \frac{3}{4} \sin(2f + 2\omega) + \frac{1}{4} e \sin(3f + 2\omega) \right] \right\}. \\
& (4.74)
\end{aligned}$$

If the required accuracy is the first-order power series perturbation solution, then the above formulas of $\bar{\sigma}(t)$ need only the following forms

$$\begin{cases} \bar{a} = \bar{a}_0, & \bar{e} = \bar{e}_0, & \bar{i} = \bar{i}_0 \\ \bar{\omega} = \bar{\omega}_0 + \omega_1(t - t_0), & \bar{M} = \bar{M}_0 + (\bar{n} + M_1)(t - t_0) \end{cases} \quad (4.75)$$

We also need the term of $a_s^{(2)}(t)$ which has two forms given as follows:

$$\begin{aligned}
\left[a_s^{(2)}(t) \right]_1 = & -\left(\frac{2}{a} \right) \left\{ a_s^{(1)}(t) + \left(\frac{3J_2}{2a} \right) (1-e^2)^{-3/2} \left(1 - \frac{3}{2} \sin^2 i \right) \right\} a_s^{(1)} \\
& + \left(\frac{3J_2}{2a} \right) \left\{ \left[2e(1-e^2)^{-5/2} \left(1 - \frac{3}{2} \sin^2 i \right) \right] \left(e_s^{(1)} - \overline{e_s^{(1)}} \right) \right. \\
& \quad \left. - \left[(1-e^2)^{-3/2} \sin 2i \right] \left(i_s^{(1)} - \overline{i_s^{(1)}} \right) \right\} \\
& + \left(\frac{3J_2}{2a} \right) \left\{ 2 \left(1 - \frac{3}{2} \sin^2 i \right) \left[\left(\frac{a}{r} \right)^4 \cos f - e(1-e^2)^{-5/2} \right] \right. \\
& \quad \left. - \frac{4}{1-e^2} \sin^2 i \left(\frac{a}{r} \right)^3 \left(1 + \frac{e}{2} \cos f \right) \sin f \sin 2(f+\omega) \right. \\
& \quad \left. + 3 \sin^2 i \left(\frac{a}{r} \right)^4 \cos f \cos 2(f+\omega) \right\} \left(e_s^{(1)} + e_l^{(1)} \right) \\
& + \left(\frac{3J_2}{2a} \right) \left\{ \sin 2i \left[\left(\frac{a}{r} \right)^3 \cos 2(f+\omega) - \left(\left(\frac{a}{r} \right)^3 - (1-e^2)^{-3/2} \right) \right] \left(i_s^{(1)} + i_l^{(1)} \right) \right. \\
& \quad \left. + \left(\frac{3J_2}{2a} \right) \left\{ -2 \sin^2 i \left(\frac{a}{r} \right)^3 \sin 2(f+\omega) \right\} \left(\omega_s^{(1)} + \omega_l^{(1)} \right) \right. \\
& \quad \left. + \left(\frac{3J_2}{2a} \right) \left\{ -\frac{e}{\sqrt{1-e^2}} \left(\frac{a}{r} \right)^4 \sin f \left[2 \left(1 - \frac{3}{2} \sin^2 i \right) + 3 \sin^2 i \cos 2(f+\omega) \right] \right. \right. \\
& \quad \left. \left. - 2\sqrt{1-e^2} \sin^2 i \left(\frac{a}{r} \right)^5 \sin 2(f+\omega) \right\} \left(M_s^{(1)} + M_l^{(1)} \right) \right. \\
& \quad \left. - \left\{ \left[aD \left(\frac{a_s^{(1)}}{a} \right) \right]_c + \left[aD \left(\frac{a_s^{(1)}}{a} \right) \right]_l \right\}, \tag{4.76}
\end{aligned}$$

$$\begin{aligned}
& \left\{ \left[aD \left(\frac{a_s^{(1)}}{a} \right) \right]_c + \left[aD \left(\frac{a_s^{(1)}}{a} \right) \right]_l \right\} \\
& = \left(\frac{3J_2}{2p^2} \right)^2 a \sqrt{1-e^2} \left\{ \left(1 - \frac{3}{2} \sin^2 i \right)^2 \left[\left(\frac{16}{9} + \frac{19}{9} e^2 \right) + \frac{1}{1-e^2} \left(\frac{35}{18} e^4 \right) + \frac{2}{9} \sqrt{1-e^2} \right] \right. \\
& \quad \left. + \sin^2 i \left(1 + \frac{2}{3} e^2 \right) + \sin^4 i \left[-\left(\frac{5}{6} - \frac{25}{24} e^2 \right) + \frac{1}{1-e^2} \left(\frac{35}{16} e^4 \right) \right] \right. \\
& \quad \left. + \sin^2 i \left[-\frac{2}{3} \left(2 - \frac{5}{2} \sin^2 i \right) \frac{\cos 2f}{e^2} + \left(\frac{5}{6} - \frac{7}{4} \sin^2 i \right) + \frac{e^2}{1-e^2} \left(\frac{7}{3} - \frac{7}{2} \sin^2 i \right) \right] e^2 \cos 2\omega \right. \\
& \quad \left. + \sin^4 i \left[\frac{1}{32(1-e^2)} \right] e^4 \cos 4\omega \right\}, \tag{4.77}
\end{aligned}$$

$$\begin{aligned}
[a_s^{(2)}(t)]_2 = & -\left(\frac{1}{a}\right)\left\{2a_s^{(1)}(t) + \left(\frac{3J_2}{2a}\right)(1-e^2)^{-3/2}\left(1 - \frac{3}{2}\sin^2 i\right)\right\}a_s^{(1)} \\
& - \left(\frac{3J_2}{2a}\right)\left\{\left[2(1-e^2)^{-3/2}tgi\left(2 - \frac{5}{2}\sin^2 i\right)\right]\right\}\left(i_s^{(1)} - \overline{i_s^{(1)}}\right) \\
& + \left(\frac{3J_2}{2a}\right)\left\{2\left(1 - \frac{3}{2}\sin^2 i\right)\left[\left(\frac{a}{r}\right)^4 \cos f - e(1-e^2)^{-5/2}\right]\right. \\
& - \frac{4}{1-e^2}\sin^2 i\left(\frac{a}{r}\right)^3\left(1 + \frac{e}{2}\cos f\right)\sin f \sin 2(f+\omega) \\
& + 3\sin^2 i\left(\frac{a}{r}\right)^4 \cos f \cos 2(f+\omega)\left\}(e_s^{(1)} + e_l^{(1)}) \right. \\
& + \left(\frac{3J_2}{2a}\right)\left\{\sin 2i\left[\left(\frac{a}{r}\right)^3 \cos 2(f+\omega) - \left(\left(\frac{a}{r}\right)^3 - (1-e^2)^{-3/2}\right)\right]\right\}\left(i_s^{(1)} + i_l^{(1)}\right) \\
& + \left(\frac{3J_2}{2a}\right)\left\{-2\sin^2 i\left(\frac{a}{r}\right)^3 \sin 2(f+\omega)\right\}\left(\omega_s^{(1)} + \omega_l^{(1)}\right) \\
& + \left(\frac{3J_2}{2a}\right)\left\{-\frac{e}{\sqrt{1-e^2}}\left(\frac{a}{r}\right)^4 \sin f\left[2\left(1 - \frac{3}{2}\sin^2 i\right) + 3\sin^2 i \cos 2(f+\omega)\right]\right. \\
& - 2\sqrt{1-e^2}\sin^2 i\left(\frac{a}{r}\right)^5 \sin 2(f+\omega)\left\}(M_s^{(1)} + M_l^{(1)}) \right. \\
& - \left\{\left[aD\left(\frac{a_s^{(1)}}{a}\right)\right]_c + \left[aD\left(\frac{a_s^{(1)}}{a}\right)\right]_l\right\}. \tag{4.78}
\end{aligned}$$

The term $(e_s^{(1)} - \overline{e_s^{(1)}})$ in the first form $[a_s^{(2)}(t)]_1$ is treated separately, whereas in the second form $[a_s^{(2)}(t)]_2$, this term does not appear because it is expressed by $a_s^{(1)}(t)$ and $(i_s^{(1)} - \overline{i_s^{(1)}})$.

(3) The second-order secular terms $\sigma_2(t - t_0)$

$$\begin{cases} a_2(t - t_0) = 0, \\ e_2(t - t_0) = 0, \\ i_2(t - t_0) = 0, \end{cases} \tag{4.79}$$

$$\begin{aligned}
\Omega_2(t - t_0) = & -\left(\frac{3J_2}{2p^2}\right)^2 \cos i \\
& \times \left[\left(\frac{3}{2} + \frac{1}{6}e^2 + \sqrt{1-e^2}\right) - \sin^2 i\left(\frac{5}{3} - \frac{5}{24}e^2 + \frac{3}{2}\sqrt{1-e^2}\right)\right]n(t - t_0) \tag{4.80}
\end{aligned}$$

$$\begin{aligned}
\omega_2(t - t_0) = & \left(\frac{3J_2}{2p^2}\right)^2 \left[\left(4 + \frac{7}{12}e^2 + 2\sqrt{1-e^2}\right) - \sin^2 i\left(\frac{103}{12} + \frac{3}{8}e^2 + \frac{11}{2}\sqrt{1-e^2}\right)\right. \\
& \left.+ \sin^4 i\left(\frac{215}{48} - \frac{15}{32}e^2 + \frac{15}{4}\sqrt{1-e^2}\right)\right]n(t - t_0) \tag{4.81}
\end{aligned}$$

$$\begin{aligned}
M_2(t - t_0) = & \left(\frac{3J_2}{2p^2} \right)^2 \sqrt{1 - e^2} \left[\frac{1}{2} \left(1 - \frac{3}{2} \sin^2 i \right)^2 \sqrt{1 - e^2} + \left(\frac{5}{2} + \frac{10}{3} e^2 \right) \right. \\
& - \sin^2 i \left(\frac{19}{3} + \frac{26}{3} e^2 \right) + \sin^4 i \left(\frac{233}{48} + \frac{103}{12} e^2 \right) \\
& \left. + \frac{e^4}{1 - e^2} \left(\frac{35}{12} - \frac{35}{4} \sin^2 i + \frac{315}{32} \sin^4 i \right) \right] n(t - t_0).
\end{aligned} \quad (4.82)$$

(4) The first-order long-period terms $\sigma_l^{(1)}(t)$

$$a_l^{(1)}(t) = 0, \quad (4.83)$$

$$\begin{aligned}
e_l^{(1)}(t) = & - \left(\frac{1 - e^2}{e} \tan i \right) i_l^{(1)}(t) \\
= & \left(\frac{3J_2}{2p^2} \right) \frac{2 \sin^2 i}{4 - 5 \sin^2 i} \left(\frac{7}{24} - \frac{5}{16} \sin^2 i \right) (1 - e^2) e \cos 2\omega
\end{aligned} \quad (4.84)$$

$$i_l^{(1)}(t) = - \left(\frac{3J_2}{2p^2} \right) \frac{\sin 2i}{4 - 5 \sin^2 i} \left(\frac{7}{24} - \frac{5}{16} \sin^2 i \right) e^2 \cos 2\omega, \quad (4.85)$$

$$\Omega_l^{(1)}(t) = - \left(\frac{3J_2}{2p^2} \right) \frac{\cos i}{(4 - 5 \sin^2 i)^2} \left(\frac{7}{3} - 5 \sin^2 i + \frac{25}{8} \sin^4 i \right) e^2 \sin 2\omega, \quad (4.86)$$

$$\begin{aligned}
\omega_l^{(1)}(t) = & - \left(\frac{3J_2}{2p^2} \right) \frac{1}{(4 - 5 \sin^2 i)^2} \left[\sin^2 i \left(\frac{25}{3} - \frac{245}{12} \sin^2 i + \frac{25}{2} \sin^4 i \right) \right. \\
& \left. - e^2 \left(\frac{7}{3} - \frac{17}{2} \sin^2 i + \frac{65}{6} \sin^4 i - \frac{75}{16} \sin^6 i \right) \right] \sin 2\omega,
\end{aligned} \quad (4.87)$$

$$M_l^{(1)}(t) = \left(\frac{3J_2}{2p^2} \right) \frac{\sqrt{1 - e^2}}{4 - 5 \sin^2 i} \sin^2 i \left[\left(\frac{25}{12} - \frac{5}{2} \sin^2 i \right) - e^2 \left(\frac{7}{12} - \frac{5}{8} \sin^2 i \right) \right] \sin 2\omega. \quad (4.88)$$

It should be pointed out that the original complete formulas for long-period terms and short-period terms derived by the method of mean orbital elements, $\langle \sigma_l^{(1)}(t) \rangle$ and $\langle \sigma_s^{(1)}(t) \rangle$, respectively, are given by

$$\langle \sigma_s^{(1)}(t) \rangle = \sigma_s^{(1)}(t) - \overline{\sigma_s^{(1)}}(t), \quad \langle \sigma_l^{(1)}(t) \rangle = \sigma_l^{(1)}(t) + \overline{\sigma_s^{(1)}}(t), \quad (4.89)$$

where the average value of $\sigma_s^{(1)}(t)$, $\overline{\sigma_s^{(1)}}(t) \neq 0$. According to the principle of the method of mean orbital elements both $\sigma_l^{(1)}(t)$ and $\sigma_s^{(1)}(t)$ are derived by the same mean elements $\overline{\sigma}(t)$, and

$$\langle \sigma_s^{(1)}(t) \rangle + \langle \sigma_l^{(1)}(t) \rangle = \sigma_s^{(1)}(t) + \sigma_l^{(1)}(t).$$

Thus we can use the above given $\sigma_s^{(1)}(t)$ and $\sigma_l^{(1)}(t)$ without introducing $\overline{\sigma_s^{(1)}}(t)$. But when we use the method of quasi-mean elements the short-period terms and long-period terms must be separated rigorously, the formulas are given in the following related sections.

4.2.2 The Non-singularity Perturbation Solution of the First Type

In Sect. 3.6, we provide the methods of eliminating singularities in the perturbation solution. The singularities include geometric singularities (i.e., $e = 0$ and $i = 0$ given by improperly selected state variables) and commensurable singularities. There are two types of non-singularity perturbation solutions as follows.

- (1) The non-singularity perturbation solution of the first type eliminates a small e and commensurable singularities.
- (2) The non-singularity perturbation solution of the second type eliminates all singularities including a small e , a small i , and commensurable singularities.

Usually, the inclination of a low Earth orbit or a medium Earth orbit satellite cannot be near zero, so we only need to construct the non-singularity solution of the first type. In Sect. 4.2.3, we discuss the non-singularity solution of the second type.

4.2.2.1 Basic Variables

For eliminating singularities due to a small e , we use the first type of non-singularity variables which are

$$a, \quad i, \quad \Omega, \quad \xi = e \cos \omega, \quad \eta = e \sin \omega, \quad \lambda = M + \omega. \quad (4.90)$$

The six variables in the order of (4.90) are denoted by the vector σ . The corresponding perturbation motion equation system is expressed in the $(\partial R / \partial \sigma)$ -version.

We use Earth's dynamical form-factor term J_2 as the perturbation source, whose gravity potential is

$$R = \frac{J_2}{2a^3} \left(\frac{a}{r} \right)^3 \left[\left(1 - \frac{3}{2} \sin^2 i \right) + \frac{3}{2} \sin^2 i \cos 2(f + \omega) \right]. \quad (4.91)$$

By the average method, we separate R into two parts,

$$R = R_1 = R_{1c} + R_{1s}, \quad (4.92)$$

$$R_{1c} = \frac{J_2}{2a^3} \left(1 - \frac{3}{2} \sin^2 i \right) (1 - e^2)^{-3/2}, \quad (4.93)$$

$$R_{1s} = \frac{J_2}{2a^3} \left\{ \left(1 - \frac{3}{2} \sin^2 i \right) \left[\left(\frac{a}{r} \right)^3 - (1 - e^2)^{-3/2} \right] + \frac{3}{2} \sin^2 i \cos 2(f + \omega) \right\}, \quad (4.94)$$

where

$$e^2 = \xi^2 + \eta^2, \quad \omega = \arctan\left(\frac{\eta}{\xi}\right), \quad M = \lambda - \arctan\left(\frac{\eta}{\xi}\right),$$

$$f = f(M(\xi, \eta, \lambda), e(\xi, \eta)). \quad (4.95)$$

The corresponding perturbation equation system is

$$\dot{\sigma} = f_0(a) + f_1(\sigma, \varepsilon), \quad (4.96)$$

$$f_0(a) = \delta n, \quad n = a^{-3/2}, \quad (4.97)$$

$$f_1(\sigma, \varepsilon) = f_{1c}(a, e(\xi, \eta), i) + f_{1s}(a, e(\xi, \eta), i, \omega(\xi, \eta), f(\lambda, \xi, \eta)), \quad (4.98)$$

where f_{1c} and f_{1s} correspond to R_{1c} and R_{1s} , respectively. To be consistent all the formulas are given in the dimensionless forms.

4.2.2.2 Construction of the Non-singularity Perturbation Solution [3–7]

The method we use is the method of quasi-mean elements, by which the power series solution of a small parameter has the following form

$$\sigma(t) = \bar{\sigma}_0 + \left[\delta \bar{n}(t - t_0) + (\sigma_1 + \sigma_2 + \cdots)(t - t_0) + \Delta \sigma_l^{(1)}(t) + \cdots \right] + \sigma_s^{(1)}(t) + \cdots, \quad (4.99)$$

Details are given in Sects. 3.6 and 3.7. Because the way to deal with the long-period term by the method of quasi-mean elements is different from that of the method of mean elements, in the process of constructing the perturbation solution of the first order the following sum originally by the method of mean elements

$$\sum_j \frac{\partial(f_{1c} + f_{1s})}{\partial \sigma_j} \left(\sigma_l^{(1)} + \sigma_s^{(1)} \right)_j,$$

now becomes

$$\sum_j \frac{\partial(f_{1c} + f_{1s})}{\partial \sigma_j} (\sigma_s^{(1)})_j. \quad (4.100)$$

The calculation involved in the sum is the main part of the configuration of the perturbation solution of the first order (including the short-period terms of the first order, the secular terms of the first order and the second order, and the long-period terms of the first order). In Chap. 3, we prove that the final results can be given by recombining the solutions derived from Kepler orbital elements. Details are given in Sect. 3.7.1.

4.2.2.3 Expressions of the Non-Singularity Perturbation Solution

Theoretically for a gravitational perturbation, the secular terms (including the long-period terms) can be formed by the following expressions

$$\Delta a(t) = 0, \quad (4.101)$$

$$\Delta i(t) = \Delta i_l(t), \quad (4.102)$$

$$\Delta \Omega(t) = \Delta \Omega_c(t) + \Delta \Omega_l(t), \quad (4.103)$$

$$\Delta \xi(t) = \cos \omega[\Delta e(t)] - \sin \omega[e \Delta \omega(t)], \quad (4.104)$$

$$\Delta \eta(t) = \sin \omega[\Delta e(t)] + \cos \omega[e \Delta \omega(t)], \quad (4.105)$$

$$\Delta \lambda(t) = \bar{n}(t - t_0) + [\Delta M(t) + \Delta \omega(t)], \quad (4.106)$$

where $\bar{n} = \bar{a}^{-3/2} = \bar{a}_0^{-3/2}$ and $\Delta \sigma(t)$ includes secular terms and long-period terms.

The expressions (4.104) and (4.105) are linear which include the secular part of J_2 perturbation of the first order. By definition, the perturbation solution of the first order must include secular terms of the second order. Because the secular variation of ω affects the secular variations of other elements, it is better to use the “complete” expressions given by the original non-singularity solution of the first type, which are

$$\Delta \xi(t) = \bar{\xi}_0 \cos \Delta \omega(t) - \bar{\eta}_0 \sin \Delta \omega(t), \quad (4.107)$$

$$\Delta \eta(t) = \bar{\eta}_0 \cos \Delta \omega(t) + \bar{\xi}_0 \sin \Delta \omega(t), \quad (4.108)$$

where $\bar{\xi}_0$ and $\bar{\eta}_0$ have the same definitions as previously given.

The short-period terms are expressed as

$$a_s(t) = a_s^{(1)}(t) + a_s^{(2)}(t), \quad (4.109)$$

$$i_s^{(1)}(t) = i_s^{(1)}(t), \quad (4.110)$$

$$\Omega_s^{(1)}(t) = \Omega_s^{(1)}(t), \quad (4.111)$$

$$\xi_s^{(1)}(t) = \cos \omega [e_s^{(1)}(t)] - \sin \omega [e_s^{(1)}(t)], \quad (4.112)$$

$$\eta_s^{(1)}(t) = \sin \omega [e_s^{(1)}(t)] + \cos \omega [e_s^{(1)}(t)], \quad (4.113)$$

$$\lambda_s^{(1)}(t) = [M_s^{(1)}(t)] + [\omega_s^{(1)}(t)]. \quad (4.114)$$

In the above formulas (4.102)–(4.124) the terms $\Delta e(t)$, $\Delta i(t)$, $\Delta \omega(t)$, $\Delta \Omega(t)$, $\Delta M(t)$, and $a_s^{(1)}(t)$, $e_s^{(1)}(t)$, $i_s^{(1)}(t)$, $\omega_s^{(1)}(t)$, $\Omega_s^{(1)}(t)$, $M_s^{(1)}(t)$ are given by the results from the perturbation solution of Kepler elements, the actual formulas are provided in Sect. 4.2.2.4. By this method, the perturbation solution formed by rearranging the original solution of Kepler elements no longer has the singularity of $(1/e)$ type.

4.2.2.4 The Non-singularity Perturbation Solution Expressed in Kepler Elements

① σ_1 and σ_2 for $\sigma_1(t - t_0)$ and $\sigma_2(t - t_0)$ in $\Delta\sigma(t)$

$$a_1 = 0, \quad e_1 = 0, \quad i_1 = 0, \quad (4.115)$$

$$\Omega_1 = -\frac{3J_2}{2p^2} n \cos i, \quad (4.116)$$

$$\omega_1 = \frac{3J_2}{2p^2} n \left(2 - \frac{5}{2} \sin^2 i\right), \quad (4.117)$$

$$M_1 = \frac{3J_2}{2p^2} n \left(1 - \frac{3}{2} \sin^2 i\right) \sqrt{1 - e^2}, \quad (4.118)$$

$$a_2 = 0, \quad e_2 = 0, \quad i_2 = 0, \quad (4.119)$$

$$\Omega_2 = -\left(\frac{3J_2}{2p^2}\right)^2 n \cos i \left[\left(\frac{3}{2} + \frac{1}{6}e^2 + \sqrt{1 - e^2}\right) - \left(\frac{5}{3} - \frac{5}{24}e^2 + \frac{3}{2}\sqrt{1 - e^2}\right) \sin^2 i \right], \quad (4.120)$$

$$\omega_2 = \left(\frac{3J_2}{2p^2}\right)^2 n \left[\left(4 + \frac{7}{12}e^2 + 2\sqrt{1-e^2}\right) - \left(\frac{103}{12} + \frac{3}{8}e^2 + \frac{11}{2}\sqrt{1-e^2}\right) \sin^2 i \right. \\ \left. + \left(\frac{215}{48} - \frac{15}{32}e^2 + \frac{15}{4}\sqrt{1-e^2}\right) \sin^4 i \right], \quad (4.121)$$

$$M_2 = \left(\frac{3J_2}{2p^2}\right)^2 n \sqrt{1-e^2} \left[\frac{1}{2} \left(1 - \frac{3}{2} \sin^2 i\right)^2 \sqrt{1-e^2} + \left(\frac{5}{2} + \frac{10}{3}e^2\right) \right. \\ \left. - \left(\frac{19}{3} + \frac{26}{3}e^2\right) \sin^2 i + \left(\frac{233}{48} + \frac{103}{12}e^2\right) \sin^4 i \right. \\ \left. + \frac{e^4}{1-e^2} \left(\frac{35}{12} - \frac{35}{4} \sin^2 i + \frac{315}{32} \sin^4 i\right) \right]. \quad (4.122)$$

② Long-period terms in $\Delta\sigma(t)$, $\Delta\sigma_l^{(1)}(t) = \sigma_l^{(1)}(t) - \sigma_l^{(1)}(t_0)$

$$\Delta a_l^{(1)}(t) = 0, \quad (4.123)$$

$$\Delta e_l^{(1)}(t) = \left(\frac{3J_2}{2p^2}\right) (1-e^2) e \sin^2 i \left[\frac{1}{3} \left(2 - \frac{5}{2} \sin^2 i\right) F_2(e) + \left(\frac{7}{12} - \frac{5}{8} \sin^2 i\right) \right] G_{2s}, \quad (4.124)$$

$$\Delta i_l^{(1)}(t) = -\left(\frac{3J_2}{2p^2}\right) e^2 \sin i \cos i \left[\frac{1}{3} \left(2 - \frac{5}{2} \sin^2 i\right) F_2(e) + \left(\frac{7}{12} - \frac{5}{8} \sin^2 i\right) \right] G_{2s}, \quad (4.125)$$

$$\Delta \Omega_l^{(1)}(t) = \left(\frac{3J_2}{2p^2}\right) e^2 \cos i \left[\left(\frac{2}{3} - \frac{5}{3} \sin^2 i\right) F_2(e) + \left(\frac{7}{12} - \frac{5}{4} \sin^2 i\right) \right] G_{2c}, \quad (4.126)$$

$$\Delta \omega_l^{(1)}(t) = -\left(\frac{3J_2}{2p^2}\right) \left\{ \sin^2 i (4 - 5 \sin^2 i) \left(\frac{1}{8} + \frac{1-e^2}{6} F_2(e)\right) (e) \right. \\ \left. + e^2 \left(\frac{2}{3} - \frac{11}{3} \sin^2 i + \frac{10}{3} \sin^4 i\right) F_2(e) \right. \\ \left. - \sin^2 i \left(\frac{25}{12} - \frac{5}{2} \sin^2 i\right) + e^2 \left(\frac{7}{12} - \frac{79}{24} \sin^2 i + \frac{45}{16} \sin^4 i\right) \right\} G_{2c}, \quad (4.127)$$

$$\Delta M_l^{(1)}(t) = -\left(\frac{3J_2}{2p^2}\right) \sqrt{1-e^2} \sin^2 i \left\{ \left[-\left(4 - 5 \sin^2 i\right) \left(\frac{e^2}{4} + \frac{1+2e^2}{6}\right) F_2(e) \right. \right. \\ \left. \left. + \left(\frac{19}{12} - \frac{15}{8} \sin^2 i\right) + e^2 \left(\frac{2}{3} - 2 \sin^2 i\right) \right] G_{2c} \right. \\ \left. + \frac{e^4}{1-e^2} \left[\frac{7}{2} \left(1 - \frac{3}{2} \sin^2 i\right) G_{2c} + \left(\frac{3}{64} \sin^2 i\right) G_{4c} \right] \right\}, \quad (4.128)$$

$$F_2(e) = \frac{\overline{\cos 2f}}{e^2} = \frac{1 + 2\sqrt{1-e^2}}{(1 + \sqrt{1-e^2})^2} = \frac{3}{4} + \frac{1}{8}e^2 + \frac{3}{64}e^4 + \dots, \quad (4.129)$$

$$\begin{cases} G2s = \frac{(\cos 2\bar{\omega} - \cos 2\bar{\omega}_0)}{2(2 - \frac{5}{2}\sin^2 i)} \rightarrow -\sin 2\bar{\omega}_0 \left(\frac{3J_2}{2p^2} \right) n(t - t_0) \\ G2c = \frac{-(\sin 2\bar{\omega} - \sin 2\bar{\omega}_0)}{2(2 - \frac{5}{2}\sin^2 i)} \rightarrow -\cos 2\bar{\omega}_0 \left(\frac{3J_2}{2p^2} \right) n(t - t_0). \\ G4c = \frac{-(\sin 4\bar{\omega} - \sin 4\bar{\omega}_0)}{4(2 - \frac{5}{2}\sin^2 i)} \rightarrow -\cos 4\bar{\omega}_0 \left(\frac{3J_2}{2p^2} \right) n(t - t_0) \end{cases} \quad (4.130)$$

In the above formulas, the elements a , e , i , and n are quasi-mean elements $\bar{a}_0, \bar{e}_0, \bar{i}_0, \bar{n}_0 = \sqrt{\mu \bar{a}_0^{(-3/2)}}$, and $p = a(1 - e^2)$ is $p_0 = \bar{a}_0(1 - \bar{e}_0^2)$. Also, in (4.130) $\bar{\omega}_0$ is the quasi-mean argument of perigee at time t_0 , which is $\bar{\omega}_0(t_0)$. The formulas on the right sides of (4.130) have two expressions for $G2s$, $G2c$, and $G4c$, when the inclination of the orbit is near the critical inclination $i_c = 63.4^\circ$, the three terms take the second expressions, otherwise take the first ones, by this way the commensurable small divisor problem can be avoided.

③ Short-period terms $\sigma_s^{(1)}(t)$ and $a_s^{(2)}(t)$

$$a_s^{(1)}(t) = \frac{3J_2}{2a} \left\{ \frac{2}{3} \left(1 - \frac{3}{2}\sin^2 i \right) \left[\left(\frac{a}{r} \right)^3 - (1 - e^2)^{-3/2} \right] + \sin^2 i \left(\frac{a}{r} \right)^3 \cos 2(f + \omega) \right\}, \quad (4.131)$$

$$\begin{aligned} e_s^{(1)}(t) = & \left(\frac{3J_2}{2p^2} \right) \left\{ \frac{1}{3} \left(1 - \frac{3}{2}\sin^2 i \right) \left[e \left(\frac{2 + \sqrt{1-e^2} - e^2}{1 + \sqrt{1-e^2}} \right) + \cos f (3(1 + e \cos f) + (e \cos f)^2) \right] \right. \\ & + \frac{1}{2} \sin^2 i \left[(e + \cos f (3(1 + e \cos f) + (e \cos f)^2)) \cos 2(f + \omega) \right. \\ & \left. \left. - (1 - e^2) \left(\cos(f + 2\omega) + \frac{1}{3} \cos(3f + 2\omega) \right) \right] \right\}, \end{aligned} \quad (4.132)$$

$$i_s^{(1)}(t) = \left(\frac{3J_2}{2p^2} \right) \sin 2i \left\{ \frac{e}{4} \cos(f + 2\omega) + \frac{1}{4} \cos(2f + 2\omega) + \frac{e}{12} \cos(3f + 2\omega) \right\}, \quad (4.133)$$

$$\begin{aligned} \Omega_s^{(1)}(t) = & - \left(\frac{3J_2}{2p^2} \right) \cos i \{ (f - M + e \sin f) \\ & - \frac{1}{2} [e \sin(f + 2\omega) + \sin(2f + 2\omega) + \frac{e}{3} \sin(3f + 2\omega)] \}, \\ e\omega_s^{(1)}(t) = & \left(\frac{3J_2}{2p^2} \right) \left\{ \left(2 - \frac{5}{2}\sin^2 i \right) (f - M + e \sin f) e \right. \end{aligned} \quad (4.134)$$

$$\begin{aligned}
& + \left(1 - \frac{3}{2} \sin^2 i\right) \left[\left(1 - \frac{e^2}{4}\right) \sin f + \frac{e}{2} \sin 2f + \frac{e^2}{12} \sin 3f \right] \\
& - \frac{e}{2} \left[e \sin(f + 2\omega) + \sin 2(f + \omega) + \frac{e}{3} \sin(3f + 2\omega) \right] \\
& + \sin^2 i \left[\frac{e^2}{16} \sin(f - 2\omega) - \left(\frac{1}{4} - \frac{15}{16} e^2\right) \sin(f + 2\omega) + \frac{5}{4} e \sin 2(f + \omega) \right. \\
& \left. + \left(\frac{7}{12} + \frac{19}{48} e^2\right) \sin(3f + 2\omega) + \frac{3}{8} e \sin(4f + 2\omega) + \frac{e^2}{16} \sin(5f + 2\omega) \right] \Bigg\}, \\
& \hspace{25em} (4.135)
\end{aligned}$$

$$\begin{aligned}
\lambda_s^{(1)}(t) &= M_s^{(1)}(t) + \omega_s^{(1)}(t) \\
&= -\cos i \Omega_s^{(1)}(t) \\
&+ \left(\frac{3J_2}{2p^2}\right) \sqrt{1 - e^2} \left\{ \left(1 - \frac{3}{2} \sin^2 i\right) (f - M + e \sin f) \right. \\
&+ \sin^2 i \left[\frac{3}{4} e \sin(f + 2\omega) + \frac{3}{4} \sin(2f + 2\omega) + \frac{1}{4} e \sin(3f + 2\omega) \right] \Bigg\} \\
&+ \frac{1}{1 + \sqrt{1 - e^2}} \left(\frac{3J_2}{2p^2}\right) \left\{ \left(1 - \frac{3}{2} \sin^2 i\right) [(f - M + e \sin f) e^2 \right. \\
&+ \left(1 - \frac{e^2}{4}\right) e \sin f + \frac{e^2}{2} \sin 2f + \frac{e^3}{12} \sin 3f] \\
&+ \sin^2 i \left[-\left(\frac{1}{4} - \frac{7}{16} e^2\right) e \sin(f + 2\omega) + \frac{3}{4} e^2 \sin 2(f + \omega) \right. \\
&+ \left(\frac{7}{12} + \frac{11}{48} e^2\right) e \sin(3f + 2\omega) + \frac{3}{8} e^2 \sin(4f + 2\omega) \\
&\left. \left. + \frac{e^3}{16} (\sin(5f + 2\omega) + \sin(f - 2\omega)) \right] \right\}. \\
& \hspace{25em} (4.136)
\end{aligned}$$

The complete expression of $\sigma_s^{(1)}(t)$ should be $\sigma_s^{(1)}(t) - \overline{\sigma_s^{(1)}(t)}$, and $\overline{\sigma_s^{(1)}(t)}$ are given by

$$\left\{ \begin{aligned}
\overline{a_s^{(1)}}(t) &= 0, \\
\overline{e_s^{(1)}}(t) &= \frac{1}{6} \left(\frac{3J_2}{2p^2} \right) \sin^2 i F_2(e) (1 - e^2) e \cos 2\omega, \\
\overline{i_s^{(1)}}(t) &= -\frac{1}{6} \left(\frac{3J_2}{2p^2} \right) \sin i \cos i F_2(e) e^2 \cos 2\omega, \\
\overline{\Omega_s^{(1)}}(t) &= -\frac{1}{6} \left(\frac{3J_2}{2p^2} \right) \cos i F_2(e) e^2 \sin 2\omega, \\
\overline{\omega_s^{(1)}}(t) &= \frac{1}{6} \left(\frac{3J_2}{2p^2} \right) \left[\sin^2 i \left(\frac{3}{4} + (1 - e^2) F_2(e) \right) + \cos^2 i (e^2 F_2(e)) \right] \sin 2\omega, \\
\overline{\lambda_s^{(1)}}(t) &= \frac{1}{6} \left(\frac{3J_2}{2p^2} \right) \left[\sin^2 i \left(\frac{1}{1 + \sqrt{1 - e^2}} \right) \left(\frac{3}{4} + \left(1 + \frac{e^2}{2} \right) F_2(e) \right) \right. \\
&\quad \left. + \left(1 - \frac{5}{2} \sin^2 i \right) F_2(e) \right] e^2 \sin 2\omega.
\end{aligned} \right. \quad (4.137)$$

About $a_s^{(2)}(t)$ we use the transformed form of (4.67):

$$\begin{aligned}
[a_s^{(2)}(t)]_1 = & -\left(\frac{2}{a}\right) \left\{ a_s^{(1)}(t) + \left(\frac{3J_2}{2a}\right) (1 - e^2)^{-3/2} \left(1 - \frac{3}{2} \sin^2 i\right) \right\} a_s^{(1)} \\
& + \left(\frac{3J_2}{2a}\right) \left\{ \left(1 - \frac{3}{2} \sin^2 i\right) \left[2\left(\frac{a}{r}\right)^4 \cos f - e(1 - e^2)^{-5/2} \right] \right. \\
& \quad - \frac{4}{1 - e^2} \sin^2 i \left(\frac{a}{r}\right)^3 \left(1 + \frac{e}{2} \cos f\right) \sin f \sin 2(f + \omega) \\
& \quad \left. + 3 \sin^2 i \left(\frac{a}{r}\right)^4 \cos f \cos 2(f + \omega) \right\} \left(e_s^{(1)} - \overline{e_s^{(1)}} \right) \\
& + \left(\frac{3J_2}{2a}\right) \left\{ \sin 2i \left(\frac{a}{r}\right)^3 [\cos 2(f + \omega) - 1] \left(i_s^{(1)} - \overline{i_s^{(1)}} \right) \right. \\
& \quad \left. + \left(\frac{3J_2}{2a}\right) \left\{ -2 \sin^2 i \left(\frac{a}{r}\right)^3 \sin 2(f + \omega) \right\} \left(\omega_s^{(1)} - \overline{\omega_s^{(1)}} \right) \right. \\
& \quad + \left(\frac{3J_2}{2a}\right) \left\{ -\frac{e}{\sqrt{1 - e^2}} \left(\frac{a}{r}\right)^4 \sin f \left[2\left(1 - \frac{3}{2} \sin^2 i\right) + 3 \sin^2 i \cos 2(f + \omega) \right] \right. \\
& \quad \left. \left. - 2\sqrt{1 - e^2} \sin^2 i \left(\frac{a}{r}\right)^5 \sin 2(f + \omega) \right\} \left(M_s^{(1)} - \overline{M_s^{(1)}} \right) \right. \\
& \quad \left. - \left[\left[aD \left(\frac{a_s^{(1)}}{a} \right) \right]_c + \left[aD \left(\frac{a_s^{(1)}}{a} \right) \right]_l \right] \right\}. \quad (4.138)
\end{aligned}$$

In (4.138), there are two terms that have $\left(\omega_s^{(1)} - \overline{\omega_s^{(1)}} \right)$ and $\left(M_s^{(1)} - \overline{M_s^{(1)}} \right)$. These two terms have the factor $(1/e)$ and need to be treated as follows.

$$\begin{aligned}
& \left(\frac{3J_2}{2a} \right) \left\{ -2 \sin^2 i \left(\frac{a}{r} \right)^3 \sin 2(f + \omega) \right\} \left(\omega_s^{(1)} - \overline{\omega_s^{(1)}} \right) \\
& + \left(\frac{3J_2}{2a} \right) \left\{ -\frac{e}{\sqrt{1-e^2}} \left(\frac{a}{r} \right)^4 \sin f \left[2 \left(1 - \frac{3}{2} \sin^2 i \right) + 3 \sin^2 i \cos 2(f + \omega) \right] \right. \\
& \quad \left. - 2\sqrt{1-e^2} \sin^2 i \left(\frac{a}{r} \right)^5 \sin 2(f + \omega) \right\} \left(M_s^{(1)} - \overline{M_s^{(1)}} \right) \\
& = \left(\frac{3J_2}{2a} \right) \left\{ -\frac{1}{\sqrt{1-e^2}} \left(\frac{a}{r} \right)^4 \sin f \left[2 \left(1 - \frac{3}{2} \sin^2 i \right) + 3 \sin^2 i \cos 2(f + \omega) \right] \right\} \left(eM_s^{(1)} - e\overline{M_s^{(1)}} \right) \\
& + \left(\frac{3J_2}{2a} \right) \left\{ -2 \sin^2 i \left(\frac{a}{r} \right)^3 \sin 2(f + \omega) \right\} \left(\lambda_s^{(1)} - \overline{\lambda_s^{(1)}} \right) \\
& + \left(\frac{3J_2}{2a} \right) \left\{ 2 \sin^2 i \left(\frac{a}{r} \right)^3 \sin 2(f + \omega) \right\} \left[1 - \sqrt{1-e^2} \left(\frac{a}{r} \right)^2 \right] \left(M_s^{(1)} - \overline{M_s^{(1)}} \right).
\end{aligned}$$

The final expression of $a_s^{(2)}(t)$ is now given by

$$\begin{aligned}
[a_s^{(2)}(t)]_1 = & -\left(\frac{2}{a} \right) \left\{ a_s^{(1)}(t) + \left(\frac{3J_2}{2a} \right) (1-e^2)^{-3/2} \left(1 - \frac{3}{2} \sin^2 i \right) \right\} a_s^{(1)} \\
& + \left(\frac{3J_2}{2a} \right) \left\{ \left(1 - \frac{3}{2} \sin^2 i \right) \left[2 \left(\frac{a}{r} \right)^4 \cos f - e(1-e^2)^{-5/2} \right] \right. \\
& \quad - \frac{4}{1-e^2} \sin^2 i \left(\frac{a}{r} \right)^3 \left(1 + \frac{e}{2} \cos f \right) \sin f \sin 2(f + \omega) \\
& \quad \left. + 3 \sin^2 i \left(\frac{a}{r} \right)^4 \cos f \cos 2(f + \omega) \right\} \left(e_s^{(1)} - \overline{e_s^{(1)}} \right) \\
& + \left(\frac{3J_2}{2a} \right) \left\{ \sin 2i \left(\frac{a}{r} \right)^3 [\cos 2(f + \omega) - 1] \right\} \left(i_s^{(1)} - \overline{i_s^{(1)}} \right) \\
& + \left(\frac{3J_2}{2a} \right) \left\{ -2 \sin^2 i \left(\frac{a}{r} \right)^3 \sin 2(f + \omega) \right\} \left(\lambda_s^{(1)} - \overline{\lambda_s^{(1)}} \right) \\
& + \left(\frac{3J_2}{2a} \right) \left\{ \frac{2}{(1-e^2)^2} \sin^2 i \left(\frac{a}{r} \right)^3 \sin 2(f + \omega) \left[e \left(\frac{-1-2\sqrt{1-e^2}}{1+\sqrt{1-e^2}} + e^2 \right) - \sqrt{1-e^2} (2 \cos f + e \cos^2 f) \right] \right. \\
& \quad \left. - \frac{1}{\sqrt{1-e^2}} \left(\frac{a}{r} \right)^4 \sin f \left[2 \left(1 - \frac{3}{2} \sin^2 i \right) + 3 \sin^2 i \cos 2(f + \omega) \right] \right\} \left(eM_s^{(1)} - e\overline{M_s^{(1)}} \right) \\
& - \left\{ \left[aD \left(\frac{a_s^{(1)}}{a} \right) \right]_c + \left[aD \left(\frac{a_s^{(1)}}{a} \right) \right]_l \right\}, \tag{4.139}
\end{aligned}$$

where $\left\{ \left[aD \left(\frac{a_s^{(1)}}{a} \right) \right]_c + \left[aD \left(\frac{a_s^{(1)}}{a} \right) \right]_l \right\}$ is given by (4.77). We can see that in (4.139) the factor $(1/e)$ is eliminated.

If necessary $a_s^{(2)}(t)$ can be expressed directly using variables for the non-singularity elements, which does not bring trouble to the calculation. The formula is

$$\begin{aligned}
a_s^{(2)}(t) = & -\left(\frac{2}{a}\right)\left\{a_s^{(1)}(t) + \left(\frac{3J_2}{2a}\right)(1-e^2)^{-3/2}\left(1 - \frac{3}{2}\sin^2 i\right)\right\}a_s^{(1)} \\
& -\left(\frac{3J_2}{2a}\right)\left\{\sin 2i\left(\frac{a}{r}\right)^3(1-\cos 2u)\right\}\left(i_s^{(1)} - \overline{i_s^{(1)}}\right) \\
& -\left(\frac{3J_2}{2a}\right)\left\{\frac{1}{\sqrt{1-e^2}}\left(\frac{a}{r}\right)^4(\xi \sin u - \eta \cos u)\left[2 - 3\sin^2 i(1-\cos 2u)\right]\right. \\
& \quad \left.+ \sqrt{1-e^2}\left(\frac{a}{r}\right)^5 \sin^2 i \sin 2u\right\}\left(\lambda_s^{(1)} - \overline{\lambda_s^{(1)}}\right) \\
& +\left(\frac{3J_2}{a}\right)\left\{\left(\frac{a}{r}\right)^4\left[1 - \frac{3}{2}\sin^2 i(1-\cos 2u)\right]\left[\cos u + F_4(e)(\eta^2 \cos u - \xi \eta \sin u)\right]\right. \\
& \quad \left.+ (1-e^2)^{-3/2}\left(\frac{a}{r}\right)^3 \sin^2 i \sin 2u\left[-F_5(e)\eta - 2\sin u - \frac{1}{2}(\xi \sin 2u - \eta \cos 2u)\right]\right. \\
& \quad \left.+ \frac{1}{2}F_1(e)\xi\left(4(\xi \sin u - \eta \cos u) + (\xi^2 - \eta^2)\sin 2u - 2\xi\eta \cos 2u\right)\right\}\xi_s^{(1)} \\
& +\left(\frac{3J_2}{a}\right)\left\{\left(\frac{a}{r}\right)^4\left[1 - \frac{3}{2}\sin^2 i(1-\cos 2u)\right]\left[\sin u + F_4(e)(\xi^2 \sin u - \xi \eta \cos u)\right]\right. \\
& \quad \left.+ (1-e^2)^{-3/2}\left(\frac{a}{r}\right)^3 \sin^2 i \sin 2u\left[F_5(e)\xi + 2\cos u + \frac{1}{2}(\xi \cos 2u + \eta \sin 2u)\right.\right. \\
& \quad \left.\left.- \frac{1}{2}F_1(e)\eta\left(4(\xi \sin u - \eta \cos u) + (\xi^2 - \eta^2)\sin 2u - 2\xi\eta \cos 2u\right)\right]\right\}\eta_s^{(1)} \\
& -\left(\frac{3J_2}{2p^2}\right)^2 a\sqrt{1-e^2}\left\{\left(1 - \frac{3}{2}\sin^2 i\right)^2\left[\left(\frac{16}{9} + \frac{19}{9}e^2\right) + \frac{1}{1-e^2}\left(\frac{35}{18}e^4\right) + \frac{2}{9}\sqrt{1-e^2}\right]\right. \\
& \quad \left.+ \sin^2 i\left(1 + \frac{2}{3}e^2\right) + \sin^4 i\left[-\left(\frac{5}{6} - \frac{25}{24}e^2\right) + \frac{1}{1-e^2}\left(\frac{35}{16}e^4\right)\right]\right. \\
& \quad \left.+ \sin^2 i\left[-\frac{2}{3}\left(2 - \frac{5}{2}\sin^2 i\right)F_2(e) + \left(\frac{5}{6} - \frac{7}{4}\sin^2 i\right) + \frac{e^2}{1-e^2}\left(\frac{7}{3} - \frac{7}{2}\sin^2 i\right)\right](\xi^2 - \eta^2)\right. \\
& \quad \left.+ \sin^4 i\left[\frac{1}{32(1-e^2)}(\xi^4 - 6\xi^2\eta^2 + \eta^4)\right]\right\}
\end{aligned} \tag{4.140}$$

In the above formulas for the short-period terms, the orbit elements on the right sides are $\overline{\sigma}(t)$.

In the short-period terms, the non-singularity variables ξ and η can also be given by Kepler elements according to their definitions, that $\xi = e \cos \omega$, $\eta = e \sin \omega$, and $u = f + \omega$, where f is the true anomaly. In the process of calculation, $\left(\frac{a}{r}\right)$ and f are given by

$$\left(\frac{a}{r}\right) = (1 - e \cos E)^{-1}, \tag{4.141}$$

$$\sin f = \frac{a}{r}\sqrt{1-e^2}\sin E, \quad \cos f = \frac{a}{r}(\cos E - e). \tag{4.142}$$

The calculations involve the eccentric anomaly E , which is given by solving Kepler's equation

$$E = M + e \sin E. \quad (4.143)$$

The auxiliary functions in the above formulas, $F_1(e)$, $F_2(e)$, $\dots$, $F_5(e)$ are:

$$\left\{ \begin{array}{l} F_1(e) = \frac{1}{(1+\sqrt{1-e^2})} \\ F_2(e) = \frac{\cos 2f}{e^2} = \frac{1+2\sqrt{1-e^2}}{(1+\sqrt{1-e^2})^2} = \frac{3}{4} + \frac{1}{8}e^2 + \frac{3}{64}e^4 + \dots \\ F_3(e) = \frac{1}{e^4}(\cos 2f - \frac{3}{4}e^2) = \frac{1+3\sqrt{1-e^2}}{4(1+\sqrt{1-e^2})^3} = \frac{1}{8} + \frac{3}{64}e^2 + \frac{3}{128}e^4 + \dots \\ F_4(e) = \frac{1}{\sqrt{1-e^2}(1+\sqrt{1-e^2})} = \frac{1}{\sqrt{1-e^2}}F_1(e) \\ F_5(e) = \frac{5+3\sqrt{1-e^2}-2e^2}{2(1+\sqrt{1-e^2})} = \frac{1}{2}(5+3\sqrt{1-e^2}-2e^2)F_1(e) \end{array} \right. \quad (4.144)$$

4.2.2.5 Calculations of the Non-singularity Perturbation Solution

The non-singularity perturbation solution constructed using the quasi-mean elements has the form as

$$\sigma(t) = \bar{\sigma} + \sigma_s^{(1)}(t) + \sigma_s^{(2)}(t) + \dots \quad (4.145)$$

The quasi-mean elements after eliminating the short-period terms denoted by $\bar{\sigma}(t)$ is

$$\left\{ \begin{array}{l} \bar{\sigma}(t) = \bar{\sigma}_0 + (\delta \bar{n}_0 + \sigma_1 + \sigma_2 + \dots)(t - t_0) + \Delta\sigma_l^{(1)}(t) + \dots, \\ \bar{\sigma}_0 = \sigma_0 - [\sigma_s^{(1)}(t_0) + \sigma_s^{(2)}(t_0) + \dots], \\ \Delta\sigma_l^{(1)}(t) = \sigma_l^{(1)}(t) - \sigma_l^{(1)}(t_0), \end{array} \right. \quad (4.146)$$

where δ is previously defined in (4.49). By this method, the terms $\Delta\sigma_l^{(1)}(t)$ and $\sigma_2(t - t_0)$ are treated as the same type of terms. In reality, when the arc of orbital extrapolation is not too long, there is no obvious difference between the long-period terms and the secular terms. The idea of the method of quasi-mean elements agrees with this fact. The formulas for calculation are given in Sects. 4.2.2.3 and 4.2.2.4.

4.2.3 *The Non-singularity Perturbation Solution of the Second Type*

4.2.3.1 Basic Variables and the Configuration of the Perturbation Solution

The singularities of a small e and a small i may occur together for high Earth orbit satellites. To deal with this kind of situation we need the second type of non-singularity variables, which are defined as

$$\begin{aligned} a, \quad \xi &= e \cos \tilde{\omega}, \quad \eta = e \sin \tilde{\omega}, \\ h &= \sin \frac{i}{2} \cos \Omega, \quad k = \sin \frac{i}{2} \sin \Omega, \quad \lambda = M + \tilde{\omega}, \end{aligned} \quad (4.147)$$

where $\tilde{\omega} = \omega + \Omega$. The method is similar to that for the first type of non-singularity solution, which is using the quasi-mean variables to construct the solution. The three types of singularities by a small e , a small i , and commensurable small divisors can be eliminated simultaneously in the solution.

4.2.3.2 Expressions of the Non-singularity Perturbation Solution of the Second Type

We still use the J_2 perturbation as an example, and the same method to construct the perturbation solution as that in Sect. 4.2.2.2 except for the second type of non-singularity variables. The solution is expressed as

$$\sigma(t) = \bar{\sigma} + \sigma_s^{(1)}(t) + \sigma_s^{(2)}(t) + \dots$$

where $\bar{\sigma}(t)$ is for the quasi-mean variables after the short-period terms are eliminated and have the same structure as that in (4.146).

- (1) The forms of the secular terms (including the long-period terms) of the elements $\sigma = (a, \xi, \eta, h, k, \lambda)$

$$\Delta a(t) = 0, \quad (4.148)$$

$$\Delta \xi(t) = \cos \tilde{\omega}[\Delta e(t)] - e \sin \tilde{\omega}[\Delta \omega(t) + \Delta \Omega(t)], \quad (4.149)$$

$$\Delta \eta(t) = \sin \tilde{\omega}[\Delta e(t)] + e \cos \tilde{\omega}[\Delta \omega(t) + \Delta \Omega(t)], \quad (4.150)$$

$$\Delta h(t) = \frac{1}{2} \left(\cos \frac{i}{2} \cos \Omega \right) [\Delta i(t)] - \left(\sin \frac{i}{2} \sin \Omega \right) [\Delta \Omega(t)], \quad (4.151)$$

$$\Delta k(t) = \frac{1}{2} \left(\cos \frac{i}{2} \sin \Omega \right) [\Delta i(t)] + \left(\sin \frac{i}{2} \cos \Omega \right) [\Delta \Omega(t)], \quad (4.152)$$

$$\Delta \lambda(t) = \bar{n}(t - t_0) + [\Delta M(t) + \Delta \omega(t) + \Delta \Omega(t)], \quad (4.153)$$

where $\bar{n} = \bar{a}^{-3/2} = \bar{a}_0^{-3/2}$, and $\Delta \sigma(t)$ includes the secular term and the long-period term.

(2) The forms of the short-period terms

$$a_s(t) = a_s^{(1)}(t) + a_s^{(2)}(t), \quad (4.154)$$

$$\xi_s^{(1)}(t) = \cos \tilde{\omega} [e_s^{(1)}(t)] - e \sin \tilde{\omega} [\omega_s^{(1)}(t) + \Omega_s^{(1)}(t)], \quad (4.155)$$

$$\eta_s^{(1)}(t) = \sin \tilde{\omega} [e_s^{(1)}(t)] + e \cos \tilde{\omega} [\omega_s^{(1)}(t) + \Omega_s^{(1)}(t)], \quad (4.156)$$

$$h_s^{(1)}(t) = \frac{1}{2} \cos \frac{i}{2} \cos \Omega [i_s^{(1)}(t)] - \sin \Omega \left[\sin \frac{i}{2} \Omega_s^{(1)}(t) \right], \quad (4.157)$$

$$k_s^{(1)}(t) = \frac{1}{2} \cos \frac{i}{2} \sin \Omega [i_s^{(1)}(t)] + \cos \Omega \left[\sin \frac{i}{2} \Omega_s^{(1)}(t) \right], \quad (4.158)$$

$$\lambda_s^{(1)}(t) = [M_s^{(1)}(t)]_1 + (1 - \sqrt{1 - e^2}) [\omega_s^{(1)}(t)]_1 + (1 - \cos i) [\Omega_s^{(1)}(t)]. \quad (4.159)$$

In the above formulas (4.149)–(4.159) the terms $\Delta e(t)$, $\Delta i(t)$, $\Delta \omega(t)$, $\Delta \Omega(t)$, $\Delta M(t)$, and $a_s^{(1)}(t)$, $e_s^{(1)}(t)$, $i_s^{(1)}(t)$, $\omega_s^{(1)}(t)$, $\Omega_s^{(1)}(t)$, $M_s^{(1)}(t)$ are results from the perturbation solution of Kepler elements. The complete expression of $\sigma_s^{(1)}(t)$ is $\sigma_s^{(1)}(t) - \overline{\sigma_s^{(1)}(t)}$. By this method the factors $(1/e)$ and $(1/\sin i)$ are eliminated in the solution. Similarly, as done in the first type of non-singularity solution, the result for $\lambda_s^{(1)}(t)$, after factors of a small e and a small i are eliminated, is derived as

$$\begin{aligned}
\lambda_s^{(1)}(t) &= (1 - \cos i) [\Omega_s^{(1)}(t)] + \left(1 - \sqrt{1 - e^2}\right) [\omega_s^{(1)}(t)]_1 + [M_s^{(1)}(t)]_1 \\
&\quad - \left[\overline{\Omega_s^{(1)}(t)} + \overline{\omega_s^{(1)}(t)} + \overline{M_s^{(1)}(t)} \right] \\
&= \frac{1}{1 + \cos i} [\sin^2 i \Omega_s^{(1)}(t)] \\
&\quad + \frac{3J_2}{2p^2} \left\{ \left(2 - \frac{5}{2} \sin^2 i\right) [(f - M) + e \sin f] \right. \\
&\quad + F_1(e) \left(1 - \frac{3}{2} \sin^2 i\right) \left[\left(1 - \frac{e^2}{4}\right) e \sin f + \frac{e^2}{2} \sin 2f + \frac{e^3}{12} \sin 3f \right] \\
&\quad - \left(1 - \frac{5}{2} \sin^2 i\right) \left[\frac{1}{2} e \sin(f + 2\omega) + \frac{1}{2} \sin(2f + 2\omega) + \frac{1}{6} e \sin(3f + 2\omega) \right] \\
&\quad - F_1(e) \sin^2 i \left[\left(\frac{1}{4} + \frac{5}{16} e^2\right) e \sin(f + 2\omega) - \left(\frac{7}{12} - \frac{1}{48} e^2\right) e \sin(3f + 2\omega) \right. \\
&\quad \left. \left. - \frac{3}{8} e^2 \sin(4f + 2\omega) - \frac{e^3}{16} (\sin(5f + 2\omega) + \sin(f - 2\omega)) \right] \right\} \\
&\quad - \left[\overline{\Omega_s^{(1)}(t)} + \overline{\omega_s^{(1)}(t)} + \overline{M_s^{(1)}(t)} \right]
\end{aligned} \tag{4.160}$$

where the three terms, $\overline{\Omega_s^{(1)}(t)}$, $\overline{\omega_s^{(1)}(t)}$, $\overline{M_s^{(1)}(t)}$, have neither $(1/e)$ nor $(1/\sin i)$, as provided in (4.137).

When we use the perturbation solution for orbital extrapolation the basic state variables must be the non-singularity variables, but in the process of calculating the variables are Kepler elements $a(t)$, $e(t)$, $\dots$, or $\bar{a}(t)$, $\bar{e}(t)$, $\dots$ which are transformed from the non-singularity variables (the instantaneous variables $\sigma(t)$ or mean variables $\bar{\sigma}(t)$).

The perturbation solution provided in this section no longer has the factors of commensurable small divisors, such as the critical inclination, similarly, the factors of $(1/e)$ and $(1/\sin i)$ are also eliminated. The solution of $a_s^{(2)}(t)$ is the same as in (4.139) since it does not have the small i problem.

$$\begin{aligned}
a_s^{(2)}(t) = \frac{3(J_{2,2})}{2a} & \left\{ (1 + \cos i)^2 \left[-\frac{e}{2(1-2\alpha)} \cos(M + 2\omega + 2\Omega_{2,2}) \right. \right. \\
& + \frac{1}{1-\alpha} \cos(2M + 2\omega + 2\Omega_{2,2}) + \frac{7e}{2(1-2\alpha/3)} \cos(3M + 2\omega + 2\Omega_{2,2}) \Big] \\
& + 2 \sin^2 i \left[\frac{3e}{2(1-2\alpha)} \cos(M + 2\Omega_{2,2}) + \frac{3e}{2(1+2\alpha)} \cos(M - 2\Omega_{2,2}) \right] \\
& + (1 - \cos i)^2 \left[-\frac{e}{2(1+2\alpha)} \cos(M + 2\omega - 2\Omega_{2,2}) \right. \\
& + \frac{1}{1+\alpha} \cos(2M + 2\omega - 2\Omega_{2,2}) + \frac{7e}{2(1+2\alpha/3)} \cos(3\Omega + 2\omega - 2\Omega_{2,2}) \Big] \Big\}, \quad (4.162)
\end{aligned}$$

$$\begin{aligned}
e_s^{(2)}(t) = \frac{3(J_{2,2})}{4a^2} & \times \left\{ (1 + \cos i)^2 \left[\frac{1}{2(1-2\alpha)} \cos(M + 2\omega + 2\Omega_{2,2}) - \frac{e}{2(1-\alpha)} \cos(2M + 2\omega + 2\Omega_{2,2}) \right. \right. \\
& + \frac{7}{6(1-2\alpha/3)} \cos(3M + 2\omega + 2\Omega_{2,2}) + \frac{17e}{4(1-\alpha/2)} \cos(4M + 2\omega + 2\Omega_{2,2}) \Big] \\
& + 2 \sin^2 i \left[\frac{3}{2(1-2\alpha)} \cos(M + 2\Omega_{2,2}) + \frac{9e}{4(1-\alpha)} \cos(2M + 2\Omega_{2,2}) \right. \\
& + \frac{3}{2(1+2\alpha)} \cos(M - 2\Omega_{2,2}) + \frac{9e}{4(1+\alpha)} \cos(2M - 2\Omega_{2,2}) \Big] \\
& + (1 - \cos i)^2 \left[\frac{1}{2(1+2\alpha)} \cos(M + 2\omega - 2\Omega_{2,2}) - \frac{e}{2(1+\alpha)} \cos(2M + 2\omega - 2\Omega_{2,2}) \right. \\
& + \frac{7}{6(1+2\alpha/3)} \cos(3M + 2\omega - 2\Omega_{2,2}) + \frac{17e}{4(1+\alpha/2)} \cos(4M + 2\omega - 2\Omega_{2,2}) \Big] \Big\}, \quad (4.163)
\end{aligned}$$

$$\begin{aligned}
i_s^{(2)}(t) = \frac{3(J_{2,2})}{2a^2} \sin i & \left\{ (1 + \cos i) \left[\frac{e}{2(1-2\alpha)} \cos(M + 2\omega + 2\Omega_{2,2}) \right. \right. \\
& - \frac{1}{2(1-\alpha)} \cos(2M + 2\omega + 2\Omega_{2,2}) - \frac{7e}{6(1-2\alpha/3)} \cos(3M + 2\omega + 2\Omega_{2,2}) \Big] \\
& + 2 \left[-\frac{3e}{2(1-2\alpha)} \cos(M + 2\Omega_{2,2}) + \frac{1}{2\alpha} \cos(2\Omega_{2,2}) + \frac{3e}{2(1+2\alpha)} \cos(M - 2\Omega_{2,2}) \right] \\
& + (1 - \cos i) \left[-\frac{e}{2(1+2\alpha)} \cos(M + 2\omega - 2\Omega_{2,2}) \right. \\
& + \frac{1}{2(1+\alpha)} \cos(2M + 2\omega - 2\Omega_{2,2}) + \frac{7e}{6(1+2\alpha/3)} \cos(3M + 2\omega - 2\Omega_{2,2}) \Big] \Big\}, \quad (4.164)
\end{aligned}$$

$$\begin{aligned}
\Omega_s^{(2)}(t) = \frac{3(J_{2,2})}{2a^2} & \left\{ (1 + \cos i) \left[\frac{e}{2(1-2\alpha)} \sin(M + 2\omega + 2\Omega_{2,2}) \right. \right. \\
& - \frac{1}{2(1-\alpha)} \sin(2M + 2\omega + 2\Omega_{2,2}) - \frac{7e}{6(1-2\alpha/3)} \sin(3M + 2\omega + 2\Omega_{2,2}) \Big] \\
& + 2 \cos i \left[\frac{3e}{2(1-2\alpha)} \sin(M + 2\Omega_{2,2}) - \frac{1}{2\alpha} \sin(2\Omega_{2,2}) + \frac{3e}{2(1+2\alpha)} \sin(M - 2\Omega_{2,2}) \right]
\end{aligned}$$

$$\begin{aligned}
& + (1 - \cos i) \left[-\frac{e}{2(1+2\alpha)} \sin(M+2\omega-2\Omega_{2,2}) \right. \\
& + \frac{1}{2(1+\alpha)} \sin(2M+2\omega-2\Omega_{2,2}) + \frac{7e}{6(1+2\alpha/3)} \sin(3M+2\omega-2\Omega_{2,2}) \left. \right] \Big\}, \\
\end{aligned} \tag{4.165}$$

$$\begin{aligned}
& \omega_s^{(2)}(t) = -\cos i \Omega_s^{(2)}(t), \\
& + \frac{3(J_{2,2})}{4a^2} \left(\frac{1}{e} \right) \left\{ (1 + \cos i)^2 \left[-\frac{1}{2(1-2\alpha)} \sin(M+2\omega+2\Omega_{2,2}) \right. \right. \\
& - \frac{5e}{2(1-\alpha)} \sin(2M+2\omega+2\Omega_{2,2}) + \frac{7}{6(1-2\alpha/3)} \sin(3M+2\omega+2\Omega_{2,2}) \\
& + \frac{17e}{4(1-\alpha/2)} \sin(4M+2\omega+2\Omega_{2,2}) \left. \right] \\
& + 2 \sin^2 i \left[\frac{9e}{4(1-\alpha)} \sin(2M+2\Omega_{2,2}) + \frac{3}{2(1-2\alpha)} \sin(M+2\Omega_{2,2}) \right. \\
& - \frac{3e}{2\alpha} \sin(2\Omega_{2,2}) + \frac{3}{2(1+2\alpha)} \sin(M-2\Omega_{2,2}) + \frac{9e}{4(1+\alpha)} \sin(2M-2\Omega_{2,2}) \left. \right] \\
& + (1 - \cos i)^2 \left[-\frac{1}{2(1+2\alpha)} \sin(M+2\omega-2\Omega_{2,2}) - \frac{5e}{2(1+\alpha)} \sin(2M+2\omega-2\Omega_{2,2}) \right. \\
& + \frac{7}{6(1+2\alpha/3)} \sin(3M+2\omega-2\Omega_{2,2}) + \frac{17e}{4(1+\alpha/2)} \sin(4M+2\omega-2\Omega_{2,2}) \left. \right] \Big\}, \\
\end{aligned} \tag{4.166}$$

$$\begin{aligned}
M_s^{(2)}(t) = & - \left(1 - \frac{e^2}{2} \right) \left(\omega_s^{(2)}(t) + \cos i \Omega_s^{(2)}(t) \right) \\
& + \frac{9(J_{2,2})}{4a^2} \left\{ (1 + \cos i)^2 \left[-\frac{e}{(1-2\alpha)} \left(1 - \frac{1}{2(1-2\alpha)} \right) \sin(M+2\omega+2\Omega_{2,2}) \right. \right. \\
& + \frac{1}{(1-\alpha)} \left(1 - \frac{1}{2(1-2\alpha)} \right) \sin(2M+2\omega+2\Omega_{2,2}) \\
& + \frac{7e}{3(1-2\alpha/3)} \left(1 - \frac{1}{2(1-2\alpha/3)} \right) \sin(3M+2\omega+2\Omega_{2,2}) \left. \right] \\
& + 2 \sin^2 i \left[\frac{3e}{(1-2\alpha)} \left(1 - \frac{1}{2(1-2\alpha)} \right) \sin(M+2\Omega_{2,2}) - \frac{1}{\alpha} \sin(2\Omega_{2,2}) \right. \\
& + \frac{3e}{(1+2\alpha)} \left(1 - \frac{1}{2(1+2\alpha)} \right) \sin(M-2\Omega_{2,2}) \left. \right] \\
& + (1 - \cos i)^2 \left[\frac{-e}{(1+2\alpha)} \left(1 - \frac{1}{2(1+2\alpha)} \right) \sin(M+2\omega-2\Omega_{2,2}) \right. \\
& + \frac{1}{(1+\alpha)} \left(1 - \frac{1}{2(1+\alpha)} \right) \sin(2M+2\omega-2\Omega_{2,2}) \\
& + \frac{7e}{3(1+2\alpha/3)} \left(1 - \frac{1}{2(1+2\alpha/3)} \right) \sin(3M+2\omega-2\Omega_{2,2}) \left. \right] \Big\}, \\
\end{aligned} \tag{4.167}$$

where the angles, Ω , ω , and M , are the quasi-mean orbital elements, $\overline{\Omega}(t)$, $\overline{\omega}(t)$, $\overline{M}(t)$, at time t .

In the above formulas, the term $\Omega_{2,2}$ is given by

$$\begin{cases} \Omega_{2,2} = (\bar{\Omega} - \theta_{2,2}) - n_e(t - t_0), \\ \theta_{2,2} = S_G + \lambda_{2,2}, \end{cases} \quad (4.168)$$

where $\theta_{2,2}$ is the “local sidereal time” of the “symmetric axis” direction of Earth's equator (whose longitude is $\lambda_{2,2}$) at time t_0 ; n_e is the angular speed of Earth's rotation, S_G is the Greenwich sidereal time, by the IAU 1980 model, there are

$$\begin{cases} S_G = \bar{S}_G + \Delta\mu, \\ \bar{S}_G = 18^h.6973746 + 879000^h.0513367t + 0^s.093104t^2, \\ t = \frac{1}{36525.0}[\text{JD}(t) - \text{JD}(2000.0)], \end{cases} \quad (4.169)$$

where $\Delta\mu$ is the nutation of the right ascension given in Chap. 1, (1.65) and (1.66); the angular speed of Earth's rotation can be given by $n_e \approx \dot{S}_G = 360^\circ.98564745/\text{d}$ approximately.

According to the IAU 2000 model the Greenwich sidereal time S_G is given by

$$\text{GAST} = \text{GMST} + \text{EE}, \quad (4.170)$$

where GAST corresponds to the true sidereal time S_G in the IAU 1980 model, GMST and EE are the Greenwich mean sidereal time and the equation of equinoxes (in the IAU 1980 model they are the Greenwich mean sidereal time $\bar{S}_G$ and the nutation of right ascension $\Delta\mu$), respectively. The formulas for calculations GMST and EE in the IAU 2000 model are given in Chap. 1 (1.65) and (1.66).

4.3.1.2 Commensurable Terms in $J_{2,2}$ Term Perturbation

There are two main types of the commensurable state in the medium Earth orbit (MEO) and the high Earth orbit (HEO), such as the geostationary Earth orbit (GEO).

(1) The state of MEO

The resonance parts in the motion of an MEO satellite (a semi-diurnal period orbital motion) due to the perturbation of $J_{2,2}$ term in the accuracy of $O(e)$ are

$$\begin{aligned} a_s^{(2)}(t) &= \frac{3J_{2,2}}{2a} \left\{ (1 + \cos i)^2 \left[-\frac{e}{2(1 - 2\alpha)} \cos(M + 2\omega + 2\Omega_{2,2}) \right] \right. \\ &\quad \left. + 2 \sin^2 i \left[\frac{3e}{2(1 - 2\alpha)} \cos(M + 2\Omega_{2,2}) \right] \right\} \\ e_s^{(2)}(t) &= \frac{3(J_{2,2})}{4a^2} \left\{ (1 + \cos i)^2 \left[\frac{1}{2(1 - 2\alpha)} \cos(M + 2\omega + 2\Omega_{2,2}) \right] \right. \end{aligned} \quad (4.171)$$

$$+2 \sin^2 i \left[\frac{3}{2(1-2\alpha)} \cos(M + 2\Omega_{2,2}) \right] \Big\} \quad (4.172)$$

$$i_s^{(2)}(t) = \frac{3(J_{2,2})}{2a^2} \sin i \left\{ (1 + \cos i) \left[\frac{e}{2(1-2\alpha)} \cos(M + 2\omega + 2\Omega_{2,2}) \right] \right. \\ \left. + 2 \left[-\frac{3e}{2(1-2\alpha)} \cos(M + 2\Omega_{2,2}) \right] \right\} \quad (4.173)$$

$$\Omega_s^{(2)}(t) = \frac{3(J_{2,2})}{2a^2} \left\{ (1 + \cos i) \left[\frac{e}{2(1-2\alpha)} \sin(M + 2\omega + 2\Omega_{2,2}) \right] \right. \\ \left. + 2 \cos i \left[\frac{3e}{2(1-2\alpha)} \sin(M + 2\Omega_{2,2}) \right] \right\} \quad (4.174)$$

$$\omega_s^{(2)}(t) = -\cos i \Omega_s^{(2)}(t) \\ + \frac{3(J_{2,2})}{4a^2} \left(\frac{1}{e} \right) \left\{ (1 + \cos i)^2 \left[-\frac{1}{2(1-2\alpha)} \sin(M + 2\omega + 2\Omega_{2,2}) \right] \right. \\ \left. + 2 \sin^2 i \left[\frac{3}{2(1-2\alpha)} \sin(M + 2\Omega_{2,2}) \right] \right\}, \quad (4.175)$$

$$M_s^{(2)}(t) = -\left(1 - \frac{e^2}{2}\right) \left(\omega_s^{(2)}(t) + \cos i \Omega_s^{(2)}(t) \right) \\ + \frac{9(J_{2,2})}{4a^2} \left\{ (1 + \cos i)^2 \left[-\frac{e}{(1-2\alpha)} \left(1 - \frac{1}{2(1-2\alpha)}\right) \sin(M + 2\omega + 2\Omega_{2,2}) \right] \right. \\ \left. + 2 \sin^2 i \left[\frac{3e}{(1-2\alpha)} \left(1 - \frac{1}{2(1-2\alpha)}\right) \sin(M + 2\Omega_{2,2}) \right] \right\}, \quad (4.176)$$

where $\Omega_{2,2}$ is given by (4.168).

In the above formulas, there is a factor of $\frac{1}{1-2\alpha}$, where

$$\alpha = n'/n, \quad (4.177)$$

n and n' are the mean angular speeds of the satellite and Earth's rotation, respectively. When

$$\alpha = n'/n \approx \frac{1}{2}, \quad (4.178)$$

the commensurable small divisor problem occurs, which corresponds to a semi-diurnal orbital motion of a medium Earth satellite.

(2) The state of HEO

For an HEO satellite such as a GEO satellite, the commensurable problem occurs when $\alpha = n'/n \approx 1$. In this state the related short-period terms $\sigma_s^{(2)}(t, J_{2,2})$ become

commensurable long-period terms, which are denoted by $\sigma_l^{(1)}(t, J_{2,2})$ and have the following forms:

$$a_l^{(1)}(t) = \frac{3(J_{2,2})}{2a} \left[(1 + \cos i)^2 \frac{\cos(2M + 2\omega + 2\Omega_{2,2})}{(n - n') + (\Omega_1 + \omega_1 + M_1)} \right] n \quad (4.179)$$

$$e_l^{(1)}(t) = \frac{3(J_{2,2})}{8a^2} e \left[\begin{aligned} & -(1 + \cos i)^2 \frac{\cos(2M + 2\omega + 2\Omega_{2,2})}{(n - n') + (\Omega_1 + \omega_1 + M_1)} \\ & + 9 \sin^2 i \frac{\cos(2M + 2\Omega_{2,2})}{(n - n') + (\Omega_1 + M_1)} \end{aligned} \right] n, \quad (4.180)$$

$$i_l^{(1)}(t) = \frac{3(J_{2,2})}{4a^2} \left[-\sin i (1 + \cos i) \frac{\cos(2M + 2\omega + 2\Omega_{2,2})}{(n - n') + (\Omega_1 + \omega_1 + M_1)} \right] n, \quad (4.181)$$

$$\Omega_l^{(1)}(t) = \frac{3(J_{2,2})}{4a^2} \left[-(1 + \cos i) \frac{\sin(2M + 2\omega + 2\Omega_{2,2})}{(n - n') + (\Omega_1 + \omega_1 + M_1)} \right] n, \quad (4.182)$$

$$\begin{aligned} \omega_l^{(1)}(t) = & -\cos i \left[\Omega_l^{(1)}(t) \right] \\ & + \frac{3(J_{2,2})}{4a^2} \left[\begin{aligned} & \frac{5}{2} (1 + \cos i)^2 \frac{\sin(2M + 2\omega + 2\Omega_{2,2})}{(n - n') + (\Omega_1 + \omega_1 + M_1)} \\ & + \frac{9}{2} \sin^2 i \frac{\sin(2M + 2\Omega_{2,2})}{(n - n') + (\Omega_1 + M_1)} \end{aligned} \right] n, \end{aligned} \quad (4.183)$$

$$\begin{aligned} M_l^{(1)}(t) = & -\frac{3(J_{2,2})}{4a^2} \left[\begin{aligned} & \frac{5}{2} (1 + \cos i)^2 \frac{\sin(2M + 2\omega + 2\Omega_{2,2})}{(n - n') + (\Omega_1 + \omega_1 + M_1)} \\ & + \frac{9}{2} \sin^2 i \frac{\sin(2M + 2\Omega_{2,2})}{(n - n') + (\Omega_1 + M_1)} \end{aligned} \right] n \\ & + \frac{9(J_{2,2})}{4a^2} \left[(1 + \cos i)^2 \frac{\sin(2M + 2\omega + 2\Omega_{2,2})}{(n - n') + (\Omega_1 + \omega_1 + M_1)} \right] n, \end{aligned} \quad (4.184)$$

where $\Omega_{2,2}$ is given by (4.168).

4.3.2 The Non-singularity Perturbation Solution of the First Type

The perturbation due to the $J_{2,2}$ term only has a short periodic effect, the corresponding short-period terms are formed by

$$a_s^{(2)}(t) = a_s^{(2)}(t), \quad (4.185)$$

$$i_s^{(2)}(t) = i_s^{(2)}(t), \quad (4.186)$$

$$\Omega_s^{(2)}(t) = \Omega_s^{(2)}(t), \quad (4.187)$$

$$\xi_s^{(2)}(t) = \cos \omega [e_s^{(2)}(t)] - \sin \omega [e \omega_s^{(2)}(t)], \quad (4.188)$$

$$\eta_s^{(2)}(t) = \sin \omega [e_s^{(2)}(t)] + \cos \omega [e \omega_s^{(2)}(t)]. \quad (4.189)$$

For the sixth element λ , the short period term is given by $\lambda_s^{(2)}(t) = [M_s^{(2)}(t)] + [\omega_s^{(2)}(t)]$, and the actual formula is

$$\begin{aligned} \lambda_s^{(2)}(t) = & -\cos i \Omega_s^{(2)}(t) \\ & + \frac{3(J_{2,2})}{4a^2} \left(\frac{e}{2} \right) \left\{ (1 + \cos i)^2 \left[-\frac{1}{2(1-2\alpha)} \sin(M + 2\omega + 2\Omega_{2,2}) \right. \right. \\ & \quad \left. \left. + \frac{7}{6(1-2\alpha/3)} \sin(3M + 2\omega + 2\Omega_{2,2}) \right] \right. \\ & + 2 \sin^2 i \left[\frac{3}{2(1-2\alpha)} \sin(M + 2\Omega_{2,2}) + \frac{3}{2(1+2\alpha)} \sin(M - 2\Omega_{2,2}) \right] \\ & \quad \left. + (1 - \cos i)^2 \left[-\frac{1}{2(1+2\alpha)} \sin(M + 2\omega - 2\Omega_{2,2}) \right. \right. \\ & \quad \left. \left. + \frac{7}{6(1+2\alpha/3)} \sin(3M + 2\omega - 2\Omega_{2,2}) \right] \right\} \\ & + \frac{9(J_{2,2})}{4a^2} \left\{ (1 + \cos i)^2 \left[-\frac{e}{(1-2\alpha)} \left(1 - \frac{1}{2(1-2\alpha)} \right) \sin(M + 2\omega + 2\Omega_{2,2}) \right. \right. \\ & \quad \left. \left. + \frac{1}{(1-\alpha)} \left(1 - \frac{1}{2(1-2\alpha)} \right) \sin(2M + 2\omega + 2\Omega_{2,2}) \right. \right. \\ & \quad \left. \left. + \frac{7e}{3(1-2\alpha/3)} \left(1 - \frac{1}{2(1-2\alpha/3)} \right) \sin(3M + 2\omega + 2\Omega_{2,2}) \right] \right. \\ & + 2 \sin^2 i \left[\frac{3e}{(1-2\alpha)} \left(1 - \frac{1}{2(1-2\alpha)} \right) \sin(M + 2\Omega_{2,2}) - \frac{1}{\alpha} \sin(2\Omega_{2,2}) \right. \\ & \quad \left. \left. + \frac{3e}{(1+2\alpha)} \left(1 - \frac{1}{2(1+2\alpha)} \right) \sin(M - 2\Omega_{2,2}) \right] \right. \\ & + (1 - \cos i)^2 \left[-\frac{e}{(1+2\alpha)} \left(1 - \frac{1}{2(1+2\alpha)} \right) \sin(M + 2\omega - 2\Omega_{2,2}) \right. \\ & \quad \left. \left. + \frac{1}{(1+\alpha)} \left(1 - \frac{1}{2(1+\alpha)} \right) \sin(2M + 2\omega - 2\Omega_{2,2}) \right. \right. \\ & \quad \left. \left. + \frac{7e}{3(1+2\alpha/3)} \left(1 - \frac{1}{2(1+2\alpha/3)} \right) \sin(3M + 2\omega - 2\Omega_{2,2}) \right] \right\}. \quad (4.190) \end{aligned}$$

In the above formulas the terms $a_s^{(2)}(t)$, $e_s^{(2)}(t)$, $i_s^{(2)}(t)$, $\Omega_s^{(2)}(t)$, $\omega_s^{(2)}(t)$, and $M_s^{(2)}(t)$ are given by the results of perturbation solutions of Kepler orbital elements. The recombined solution is the non-singularity solution without the factor of $(1/e)$.

4.3.3 The Non-singularity Perturbation Solution of the Second Type

The non-singularity variables of the second type are defined by (4.147). Using the same method of quasi-mean elements, we construct the perturbation solution to eliminate the singularities of a small e , a small i , and commensurable small divisors.

It is the same as in the above section, the perturbation due to the $J_{2,2}$ term is short periodic, the corresponding short-period terms are expressed as

$$a_s^{(2)}(t) = a_s^{(2)}(t), \quad (4.191)$$

$$\xi_s^{(2)}(t) = \cos \tilde{\omega} [e_s^{(2)}(t)] - \sin \tilde{\omega} [e\omega_s^{(2)}(t) + e\Omega_s^{(2)}(t)], \quad (4.192)$$

$$\eta_s^{(2)}(t) = \sin \tilde{\omega} [e_s^{(2)}(t)] + \cos \tilde{\omega} [e\omega_s^{(2)}(t) + e\Omega_s^{(2)}(t)], \quad (4.193)$$

$$h_s^{(2)}(t) = \frac{1}{2} \cos \frac{i}{2} \cos \Omega [i_s^{(2)}(t)] - \sin \Omega \left[\sin \frac{i}{2} \Omega_s^{(2)}(t) \right], \quad (4.194)$$

$$k_s^{(2)}(t) = \frac{1}{2} \cos \frac{i}{2} \sin \Omega [i_s^{(2)}(t)] + \cos \Omega \left[\sin \frac{i}{2} \Omega_s^{(2)}(t) \right]. \quad (4.195)$$

For the sixth element λ , the short period term is given by $\lambda_s^{(2)}(t) = M_s^{(2)}(t) + \omega_s^{(2)}(t) + \Omega_s^{(2)}(t)$, and the actual formula is

$$\begin{aligned} \lambda_s^{(2)}(t) = & (1 - \cos i) \Omega_s^{(2)}(t) \\ & + \frac{3(J_{2,2})}{4a^2} \left(\frac{e}{2} \right) \left\{ (1 + \cos i)^2 \left[-\frac{1}{2(1-2\alpha)} \sin(M + 2\omega + 2\Omega_{2,2}) \right. \right. \\ & \quad \left. \left. + \frac{7}{6(1-2\alpha/3)} \sin(3M + 2\omega + 2\Omega_{2,2}) \right] \right. \\ & + 2 \sin^2 i \left[\frac{3}{2(1-2\alpha)} \sin(M + 2\Omega_{2,2}) + \frac{3}{2(1+2\alpha)} \sin(M - 2\Omega_{2,2}) \right] \\ & \quad \left. + (1 - \cos i)^2 \left[-\frac{1}{2(1+2\alpha)} \sin(M + 2\omega - 2\Omega_{2,2}) \right. \right. \\ & \quad \left. \left. + \frac{7}{6(1+2\alpha/3)} \sin(3M + 2\omega - 2\Omega_{2,2}) \right] \right\} \\ & + \frac{9(J_{2,2})}{4a^2} \left\{ (1 + \cos i)^2 \left[-\frac{e}{(1-2\alpha)} \left(1 - \frac{1}{2(1-2\alpha)} \right) \sin(M + 2\omega + 2\Omega_{2,2}) \right. \right. \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{(1-\alpha)} \left(1 - \frac{1}{2(1-2\alpha)} \right) \sin(2M + 2\omega + 2\Omega_{2,2}) \\
& + \frac{7e}{3(1-2\alpha/3)} \left(1 - \frac{1}{2(1-2\alpha/3)} \right) \sin(3M + 2\omega + 2\Omega_{2,2}) \Big] \\
& + 2 \sin^2 i \left[\frac{3e}{(1-2\alpha)} \left(1 - \frac{1}{2(1-2\alpha)} \right) \sin(M + 2\Omega_{2,2}) - \frac{1}{\alpha} \sin(2\Omega_{2,2}) \right. \\
& \quad \left. + \frac{3e}{(1+2\alpha)} \left(1 - \frac{1}{2(1+2\alpha)} \right) \sin(M - 2\Omega_{2,2}) \right] \\
& + (1 - \cos i)^2 \left[-\frac{e}{(1+2\alpha)} \left(1 - \frac{1}{2(1+2\alpha)} \right) \sin(M + 2\omega - 2\Omega_{2,2}) \right. \\
& \quad \left. + \frac{1}{(1+\alpha)} \left(1 - \frac{1}{2(1+\alpha)} \right) \sin(2M + 2\omega - 2\Omega_{2,2}) \right. \\
& \quad \left. + \frac{7e}{3(1+2\alpha/3)} \left(1 - \frac{1}{2(1+2\alpha/3)} \right) \sin(3M + 2\omega - 2\Omega_{2,2}) \right] \Big\}. \quad (4.196)
\end{aligned}$$

On the right sides of the above formulas the terms $a_s^{(2)}(t)$, $e_s^{(2)}(t)$, $i_s^{(2)}(t)$, $\Omega_s^{(2)}(t)$, $\omega_s^{(2)}(t)$, and $M_s^{(2)}(t)$ are given by the results of perturbation solutions of Kepler orbital elements. The recombined solution is the non-singularity solution which has neither the factor $(1/e)$ nor $(1/\sin i)$.

4.4 Additional Perturbation of the Coordinate System for the First-Order Solution

4.4.1 The Cause of the Additional Perturbation of the Coordinate System [3, 8]

For an Earth's artificial satellite (especially a medium or a low Earth orbit satellite) the most important perturbation comes from the gravity of Earth's non-spherical part, as discussed in the above sections. The gravitational potential is defined in the Earth-fixed coordinate system (related to a "true" equator). When we analyze the satellite's orbital motion in the geocentric celestial coordinate system (i.e., the geocentric mean equatorial coordinate system) we must consider the influence of the change of Earth's gravity potential due to the change of coordinate system.

In the geocentric mean equatorial coordinate system, the gravity perturbation of Earth's non-spherical part $\vec{F}$ is given by the gradient of the gravity potential $V(\vec{r})$, that

$$\vec{F} = \text{grad } V(\vec{r}).$$

Earth's gravity potential is actually defined in the Earth-fixed coordinate system, denoted to $V(\vec{R})$, which takes the form in the normalized units as

$$V(\vec{R}) = \frac{GE}{R} \left\{ 1 - J_2 \left(\frac{a_e}{R} \right)^2 P_2(\sin \varphi) + \left(\frac{a_e}{R} \right)^2 P_{2,2}(\sin \varphi) [C_{2,2} \cos 2\lambda_G + S_{2,2} \sin 2\lambda_G] \right\}, \quad (4.198)$$

where R , φ , and λ_G are the three spherical components of $\vec{R}$ in the Earth-fixed coordinate system.

It is acceptable to use the IAU 1980 model for dealing with the problem of the additional perturbation of the coordinate system because it is convenient and agrees with the practical accuracy requirement. The difference between the IAU 1980 model and the IAU 2000 model would not influence the results due to the additional perturbation of the coordinate system, analytically or numerically.

We use $\vec{r}$ and $\vec{R}$ as the position vectors in the geocentric celestial coordinate system O -xyz and the Earth-fixed coordinate system O -XYZ, respectively. The transformation relationship of the two vectors is given in Sect. 1.3.4 of Chap. 1 that

$$\vec{R} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = (HG)\vec{r} = (HG) \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad (4.199)$$

where the transformation matrix (HG) is given by

$$(HG) = (EP)(ER)(NR)(PR). \quad (4.200)$$

Here (PR) is the precession matrix, (NR) is the nutation matrix, (ER) is Earth's rotation matrix, and (PR) is the polar motion matrix. From the analytical and the practical points of view, we only need to provide the solution for the first-order accuracy, therefore we have the following relationships:

$$X = (x \cos S_G + y \sin S_G) - [(\mu + \Delta\mu)y + (\theta_A + \Delta\theta)z] \cos S_G + [(\mu + \Delta\mu)x - (\Delta\varepsilon)z] \sin S_G + (x_p z), \quad (4.201)$$

$$Y = (y \cos S_G - x \sin S_G) + [(\mu + \Delta\mu)x - (\Delta\varepsilon)z] \cos S_G + [(\mu + \Delta\mu)y + (\theta_A + \Delta\theta)z] \sin S_G + (-y_p z), \quad (4.202)$$

$$Z = z + [(\theta_A + \Delta\theta)x + (\Delta\varepsilon)y] + [-x_p(x \cos S_G + y \sin S_G) + y_p(y \cos S_G - x \sin S_G)]. \quad (4.203)$$

In the epoch geocentric celestial coordinate system, the oscillation of the equator changes Earth's potential, as a result, the potential of the J_2 term changes. The variation $\Delta V(J_2)$ has a magnitude as

$$\Delta V(J_2) = O(J_2\theta) = J_2 \times 2004''T. \quad (4.204)$$

If the time interval of $\Delta t = 10$ to 20 years (measured from the epoch 2000) the value of $\Delta V(J_2)$ can reach 10^{-6} . The second-order precession and nutation terms are equivalent to the third-order perturbation, therefore can be omitted for the first-order perturbation solution. By the transformation relationships (4.201)–(4.203) the additional perturbation showed in the potential of the J_2 term, denoted to $\Delta V(J_2)$, is given by

$$\begin{aligned} \Delta V(J_2) = & \left(\frac{3J_2}{2a^3} \right) \left(\frac{a}{r} \right)^3 \sin i \{ \cos i [(\theta_A + \Delta\theta) \sin \Omega - \Delta\varepsilon \cos \Omega] \} \\ & - \left(\frac{3J_2}{2a^3} \right) \left(\frac{a}{r} \right)^3 \sin i \{ [(\theta_A + \Delta\theta) \cos \Omega + \Delta\varepsilon \sin \Omega] \sin 2u \\ & + \cos i [(\theta_A + \Delta\theta) \sin \Omega - \Delta\varepsilon \cos \Omega] \cos 2u \}, \end{aligned} \quad (4.205)$$

where r is the distance between the satellite and Earth's center, a , i , Ω , and $u = f + \omega$ are orbital elements; θ_A , $\Delta\theta$, and $\Delta\varepsilon$ are precession and nutation momentums, that

$$\begin{cases} \theta_A = 2004.3109''t, & \Delta\theta = -6.84'' \sin \Omega', \\ \Delta\varepsilon = 9.20'' \cos \Omega', \end{cases} \quad (4.206)$$

where Ω' is the mean ecliptic longitude of the ascending node of the Moon's orbit, t is given in Sect. 4.4.2 (4.214).

The choice of the coordinate system leads to the change in Earth's gravity potential, which then affects the satellite's orbital motion. This kind of perturbation is called the additional perturbation of the coordinate system.

4.4.2 The Additional Perturbation Solution in Kepler Orbital Elements

The additional perturbation of the coordinate system cannot be ignored, because it produces long-period terms of mixed type, and the magnitudes of these terms increase over increasing time intervals from the epoch. Usually, the long-period terms of mixed type are treated as long-period terms of the first order denoted by $\sigma_l^{(1)}(t)$. The related long-period terms of $\Delta\sigma_l^{(1)}(t) = \sigma_l^{(1)}(t) - \sigma_l^{(1)}(t_0)$ and the short-period terms of $\sigma_s^{(2)}(t)$ are given as follows.

The terms of $\sigma_l^{(1)}(t)$ in $\Delta\sigma_l^{(1)}(t) = \sigma_l^{(1)}(t) - \sigma_l^{(1)}(t_0)$ are

$$a_l^{(1)}(t) = 0, \quad (4.207)$$

$$e_l^{(1)}(t) = 0, \quad (4.208)$$

$$i_l^{(1)}(t) = I, \quad (4.209)$$

$$\Omega_l^{(1)}(t) = (\cot i)Q, \quad (4.210)$$

$$\omega_l^{(1)}(t) = (-\sec i)\Omega_l^{(1)}, \quad (4.211)$$

$$M_l^{(1)}(t) = 0, \quad (4.212)$$

where

$$\begin{aligned} I &= (2004''.3t) \sin \Omega - 8''.0 \cos(\Omega - \Omega') - 1''.2 \cos(\Omega + \Omega'), \\ Q &= (2004''.3t) \cos \Omega + 8''.0 \sin(\Omega - \Omega') + 1''.2 \sin(\Omega + \Omega'), \end{aligned} \quad (4.213)$$

$$t = \frac{\text{JD}(t_0) - \text{JD}(J2000.0)}{36525}, \quad (4.214)$$

In the above formulas, t_0 is the epoch of an orbit determination or an orbit extrapolation. In a few days, the variations of precession and nutation are very small and can be ignored, then the mixed long-period terms can be treated as general long-period terms.

The short-period terms $\sigma_s^{(2)}(t)$ are:

$$a_s^{(2)}(t) = \frac{3J_2}{a} \sin i \left\{ \cos i A(\theta) \left[\left(\frac{a}{r} \right)^3 - (1 - e^2)^{-3/2} - \left(\frac{a}{r} \right)^3 \cos 2u \right] - B(\theta) \left[\left(\frac{a}{r} \right)^3 \sin 2u \right] \right\}, \quad (4.215)$$

$$\begin{aligned} e_s^{(2)}(t) &= \frac{3J_2}{2p^2} \sin i \left\{ \cos i A(\theta) \left[\left(3 + \frac{3}{4}e^2 \right) \cos f + \frac{3}{2}e \cos 2f + \frac{1}{4}e^2 \cos 3f \right. \right. \\ &\quad - \frac{e^2}{8} \cos(f - 2\omega) - \left(\frac{1}{2} + \frac{11}{8}e^2 \right) \cos(f + 2\omega) - \frac{5}{2}e \cos(2f + 2\omega) \\ &\quad \left. - \left(\frac{7}{6} + \frac{17}{24}e^2 \right) \cos(3f + 2\omega) - \frac{3}{4}e \cos(4f + 2\omega) - \frac{e^2}{8} \cos(5f + 2\omega) \right] \\ &\quad - B(\theta) \left[-\frac{e^2}{8} \sin(f - 2\omega) + \left(\frac{1}{2} + \frac{11}{8}e^2 \right) \sin(f + 2\omega) + \frac{5}{2}e \sin(2f + 2\omega) \right. \\ &\quad \left. + \left(\frac{7}{6} + \frac{17}{24}e^2 \right) \sin(3f + 2\omega) + \frac{3}{4}e \sin(4f + 2\omega) + \frac{e^2}{8} \sin(5f + 2\omega) \right] \right\} \quad (4.216) \end{aligned}$$

$$\begin{aligned} i_s^{(2)}(t) &= \frac{3J_2}{4p^2} \left\{ -\cos 2i A(\theta) \left[e \cos(f + 2\omega) + \cos(2f + 2\omega) + \frac{e}{3} \cos(3f + 2\omega) \right] \right. \\ &\quad - \cos i B(\theta) [2(f - M) + 2e \sin f + e \sin(f + 2\omega) \\ &\quad \left. + \sin(2f + 2\omega) + \frac{e}{3} \sin(3f + 2\omega) \right] \left. \right\} \quad (4.217) \end{aligned}$$

$$\begin{aligned} \Omega_s^{(2)}(t) = \frac{3J_2}{2p^2} \left(\frac{1}{\sin i} \right) & \left\{ \cos 2i A(\theta) \left[(f - M) + e \sin f - \frac{e}{2} \sin(f + 2\omega) \right. \right. \\ & \left. \left. - \frac{1}{2} \sin(2f + 2\omega) - \frac{e}{6} \sin(3f + 2\omega) \right] \right. \\ & \left. + \cos i B(\theta) \left[\frac{e}{2} \cos(f + 2\omega) + \frac{1}{2} \cos(2f + 2\omega) + \frac{e}{6} \cos(3f + 2\omega) \right] \right\} \end{aligned} \quad (4.218)$$

$$\begin{aligned} \omega_s^{(2)}(t) = -\cos i \Omega_s^{(2)}(t) & + \frac{3J_2}{2p^2} \left(\frac{\sin i}{e} \right) \left\{ \cos i A(\theta) \left[3e(f - M) + \left(3 + \frac{9}{4}e^2 \right) \sin f + \frac{3}{2}e \sin 2f + \frac{1}{4}e^2 \sin 3f \right. \right. \\ & - \frac{e^2}{8} \sin(f - 2\omega) + \left(\frac{1}{2} - \frac{7}{8}e^2 \right) \sin(f + 2\omega) - \frac{3}{2}e \sin(2f + 2\omega) \\ & - \left(\frac{7}{6} + \frac{11}{24}e^2 \right) \sin(3f + 2\omega) - \frac{3}{4}e \sin(4f + 2\omega) - \frac{e^2}{8} \sin(5f + 2\omega) \left. \right] \\ & - B(\theta) \left[\frac{e^2}{8} \cos(f - 2\omega) + \left(\frac{1}{2} - \frac{7}{8}e^2 \right) \cos(f + 2\omega) - \frac{3}{2}e \cos(2f + 2\omega) \right. \\ & \left. \left. - \left(\frac{7}{6} + \frac{11}{24}e^2 \right) \cos(3f + 2\omega) - \frac{3}{4}e \cos(4f + 2\omega) - \frac{e^2}{8} \cos(5f + 2\omega) \right] \right\} \end{aligned} \quad (4.219)$$

$$\begin{aligned} M_s^{(2)}(t) = -\sqrt{1 - e^2} & \left[\omega_s^{(2)}(t) + \cos i \Omega_s^{(2)}(t) \right] \\ & + \frac{9J_2}{2p^2} \sin i \sqrt{1 - e^2} \left\{ \cos i A(\theta) \left[(f - M) + e \sin f - \frac{e}{2} \sin(f + 2\omega) \right. \right. \\ & \left. \left. - \frac{1}{2} \sin(2f + 2\omega) - \frac{e}{6} \sin(3f + 2\omega) \right] \right. \\ & \left. + B(\theta) \left[\frac{e}{2} \cos(f + 2\omega) + \frac{1}{2} \cos(2f + 2\omega) + \frac{e}{6} \cos(3f + 2\omega) \right] \right\}. \end{aligned} \quad (4.220)$$

where $A(\theta)$ and $B(\theta)$ involve precession and nutation, and are equal to I and Q in the long-period terms (4.213). They are given by

$$\begin{cases} A(\theta) = (\theta_A + \Delta\theta) \sin \Omega - (\Delta\varepsilon) \cos \Omega, \\ B(\theta) = (\theta_A + \Delta\theta) \cos \Omega + (\Delta\varepsilon) \sin \Omega, \end{cases} \quad (4.221)$$

$$\theta_A = 2004''.3109t, \quad \Delta\theta = -6''.84 \sin \Omega', \quad \Delta\varepsilon = 9''.20 \cos \Omega. \quad (4.222)$$

Here t and Ω' are the same variables as in (4.213) and (4.206), respectively.

4.4.3 The Non-singularity Additional Perturbation Solution of the First Type

4.4.3.1 Forms of the Perturbation Solution

By the same method applied to the mixed terms in the solution provided by Kepler elements, the perturbation solution due to the additional perturbation with the non-singularity elements $\sigma = (a, i, \Omega, \xi, \eta, \lambda)$ is also treated as long-period terms of the first order that $\Delta\sigma_l^{(1)}(t) = [\sigma_l^{(1)}(t) - \sigma_l^{(1)}(t_0)]$. Denote the long-period term by $\Delta\sigma(t) = \Delta\sigma_l(t)$ then the long-period terms of the six non-singularity variables are

$$\Delta a(t) = 0, \quad (4.223)$$

$$\Delta i(t) = \Delta i_l(t), \quad (4.224)$$

$$\Delta \Omega(t) = \Delta \Omega_l(t), \quad (4.225)$$

$$\Delta \xi(t) = \cos \omega [\Delta e_l(t)] - \sin \omega [e \Delta \omega_l(t)], \quad (4.226)$$

$$\Delta \eta(t) = \sin \omega [\Delta e_l(t)] + \cos \omega [e \Delta \omega_l(t)], \quad (4.227)$$

$$\Delta \lambda(t) = [\Delta M_l(t) + \Delta \omega_l(t)], \quad (4.228)$$

where $\Delta e_l(t) = 0$, $\bar{n} = \bar{a}^{-3/2} = \bar{a}_0^{-3/2}$.

4.4.3.2 The Perturbation Solution by the Quasi-Mean Elements $\Delta\sigma_l(t)$ Expressed by Kepler Orbital Elements

$$\Delta a_l(t) = 0, \quad (4.229)$$

$$\Delta e_l(t) = 0, \quad (4.230)$$

$$\Delta i_l(t) = \Omega_1(J_2)[B(t) - B(t_0)], \quad (4.231)$$

$$\Delta \Omega_l(t) = -\Omega_1(J_2) \left(\frac{\cos 2i}{\sin i \cos i} \right) [A(t) - A(t_0)], \quad (4.232)$$

$$\Delta \omega_l(t) = -\cos i \cdot \Delta \Omega(t) - \Omega_1(J_2)(3 \sin i)[A(t) - A(t_0)], \quad (4.233)$$

$$\Delta M(t) = -\Omega_1(J_2) \left(3\sqrt{1-e^2} \sin i \right) [A(t) - A(t_0)], \quad (4.234)$$

where $\Omega_1(J_2)$, $[A(t) - A(t_0)]$, and $[B(t) - B(t_0)]$ are given by

$$\Omega_1(J_2) = -\left(\frac{3J_2}{2p^2} \right) n \cos i, \quad (4.235)$$

$$A(t) - A(t_0) = (\theta_A + \Delta\theta) [\sin \overline{\Omega} - \sin \overline{\Omega}_0] - (\Delta\varepsilon) [\cos \overline{\varpi} - \cos \overline{\Omega}_0], \quad (4.236)$$

$$B(t) - B(t_0) = (\theta_A + \Delta\theta) [\cos \overline{\Omega} - \cos \overline{\Omega}_0] + (\Delta\varepsilon) [\sin \overline{\Omega} - \sin \overline{\Omega}_0], \quad (4.237)$$

$$\theta_A = 2004''.3109\bar{t}, \quad \Delta\theta = -6''.84 \sin \Omega', \quad \Delta\varepsilon = 9''.20 \cos \Omega'. \quad (4.238)$$

In (4.238), both the time $\bar{t}$ and the time for Ω' can be the intermediate time of the extrapolation or the initial time t_0 .

The $\Delta\Omega_j(t)$ given by (4.232) has two singularities, $i = 0$ and $i = 90^\circ$. When $i \approx 90^\circ$ the satellite has a polar orbit, then $\Delta\Omega_j(t)$ is calculated using a different formula which is

$$\Delta\Omega_l = \left(\frac{3J_2}{2p^2} \right) n \left(\frac{\cos 2i}{\sin i} \right) [A(t) - A(t_0)]. \quad (4.239)$$

The accuracy can be controlled by a computer program.

4.4.3.3 Corrections in the J_2 Perturbation Solution

For the non-singularity perturbation solution by the method of quasi-mean elements, the reference orbit is given by the quasi-mean elements $\bar{\sigma}$, that

$$\bar{\sigma}(t) = \bar{\sigma}_0 + (\delta n + \sigma_1 + \sigma_2)(t - t_0) + \Delta\sigma_l^{(1)} + \dots, \quad (4.240)$$

where $\bar{\sigma}_0 = \bar{\sigma}(t_0)$ is for the initial quasi-mean elements. The orbital elements in the secular terms are all quasi-mean elements $\bar{\sigma}$. For example, the formulas of the first-order secular term $\sigma_1(t - t_0)$ for Ω and λ are

$$\Omega_1 = -\left(\frac{3J_2}{2p^2} \right) n \cos i, \quad (4.241)$$

$$\lambda_1 = \left(\frac{3J_2}{2p^2} \right) n \left[\left(2 - \frac{5}{2} \sin^2 i \right) + \sqrt{1-e^2} \left(1 - \frac{3}{2} \sin^2 i \right) \right]. \quad (4.242)$$

In the process of deriving these two formulas, the integrands have $e = (\xi^2 + \eta^2)^{1/2}$ and i , which should be

$$\bar{e} = \bar{e}_0 + \left[e_l^{(1)}(t) - e_l^{(1)}(t_0) \right], \quad (4.243)$$

$$\bar{i} = \bar{i}_0 + \left[i_l^{(1)}(t) - i_l^{(1)}(t_0) \right]. \quad (4.244)$$

In reality, when we derive the perturbation solution due to the gravity of Earth's non-spherical part for Ω and λ , we do not have to use these relationships. The reason is that the long-period terms of i and e are only related to the argument of perigee ω , therefore, must have a factor e (see (4.84) and (4.85)). Then when e is small (for a low Earth orbit usually $e \leq 0.1$) we can use $\bar{i}_0$ and $\bar{e}_0$ to replace i and e in the integrands, and do not need (4.243) and (4.244). But the long-period term of the inclination i due to the additional perturbation is related to Ω , which does not have the factor e , then we must consider this problem.

Based on the original perturbation method we now provide the supplements for the long-period terms for Ω and λ , which are equivalent to the supplements for the secular terms of the second-order, denoted by $\Delta\Omega_2(t)$ and $\Delta\lambda_2(t)$, respectively. The results are

$$\Delta\Omega_2(t) = \left(\frac{3J_2}{4p^2} n \sin i \right) (\theta_A \cos \bar{\Omega}) \Omega_1(t - t_0)^2, \quad (4.245)$$

$$\Delta\lambda_2(t) = -8 \cos i \left(\frac{3J_2}{4p^2} n \sin i \right) (\theta_A \cos \bar{\Omega}) \Omega_1(t - t_0)^2, \quad (4.246)$$

where Ω_1 is given by (4.241) and $\bar{\Omega}$ is given by

$$\bar{\Omega} = \bar{\Omega}_0 + \frac{1}{2} \Omega_1(t - t_0). \quad (4.247)$$

If the arc of extrapolation is not long, then $\bar{\Omega}$ can be replaced by $\bar{\Omega}_0$.

The above results show that the perturbation solution including the additional long-period terms related to the quasi-mean elements is the complete solution in a certain sense. The errors in Ω and λ do not increase with the increasing time interval between the calculation epoch t_0 and the standard epoch J2000.0, therefore, meet the accuracy for a first-order perturbation solution.

4.4.4 The Non-singularity Additional Perturbation Solution of the Second Type

The method is the same as for the non-singularity perturbation solution of the first type, that the long-period terms due to the additional perturbation $\Delta\sigma_l^{(1)}(t) = [\sigma_l^{(1)}(t) - \sigma_l^{(1)}(t_0)]$ for the six non-singularity elements $\sigma = (a, \xi, \eta, h, k, \lambda)$ are

$$\Delta a(t) = 0, \quad (4.248)$$

$$\Delta \xi(t) = \cos \tilde{\omega}[\Delta e_l(t)] - \sin \tilde{\omega}[e \Delta \omega_l(t) + e \Delta \Omega_l(t)], \quad (4.249)$$

$$\Delta \eta(t) = \sin \tilde{\omega}[\Delta e_l(t)] + \cos \tilde{\omega}[e \Delta \omega_l(t) + e \Delta \Omega_l(t)], \quad (4.250)$$

$$\Delta h(t) = \frac{1}{2} \left(\cos \frac{i}{2} \cos \Omega \right) [\Delta i_l(t)] - (\sin \Omega) \left[\sin \frac{i}{2} \Delta \Omega_l(t) \right], \quad (4.251)$$

$$\Delta k(t) = \frac{1}{2} \left(\cos \frac{i}{2} \sin \Omega \right) [\Delta i_l(t)] + (\cos \Omega) \left[\sin \frac{i}{2} \Delta \Omega_l(t) \right], \quad (4.252)$$

$$\Delta \lambda(t) = [\Delta M(t) + \Delta \omega(t) + \Delta \Omega(t)]. \quad (4.253)$$

The corresponding perturbation solution by the quasi-mean elements corresponding to Kepler orbital elements is given by (4.229)–(4.234).

4.4.5 Selection of Coordinate System and Related Problems

The above analyses, actual calculations, and treatments show that the analytical solution due to the additional perturbation of the coordinate system is not only simple but also meets certain accuracy requirements of orbital extrapolation. If the standard epoch changes every 50 years, based on the variation of the space coordinate system due to the oscillation of Earth's equator we only need to consider the additional perturbation term $\Delta V(J_2)$, which is not complicated. Realistically for the orbital determination and forecast (including the transformation of the instantaneous elements and the mean elements, and applications), we can accept a unified coordinate system, which is the epoch (it is **J2000.0** for present days) **geocentric celestial coordinate system**. By this we do not need a mixed type of **orbital coordinate systems**, therefore, to avoid the “chaos” caused by switching systems, also to simplify the transformation between the instantaneous elements and the mean elements when using both the analytical and numerical methods for a project.

For some particular fields in order to keep the continuity of historical work, the transformation of coordinate systems is unavoidable. For example, using the American TLE elements data, we need to deal with the transformation of coordinate systems. It can be done without much trouble as long as the relationships between different frames are clear. The transformations between geocentric frames are given as follows.

- (1) Transformation between the J2000.0 geocentric celestial coordinate system and the J2000.0 orbital coordinate system

The J2000.0 orbital coordinate system uses the true equatorial plane at time t as the coordinate xy -plane, and the J2000.0 mean March equinox direction as the direction of the x -axis.

Assuming $\vec{r}_c$ and $\vec{r}_2$ to be the position vectors of a point in the J2000.0 geocentric celestial coordinate system and the J2000.0 orbital coordinate system, respectively, as discussed in Chap. 1, the transformation is that

$$\vec{r}_2 = R_z(\mu + \Delta\mu)(NR)(PR)\vec{r}_c, \quad (4.254)$$

where $\mu + \Delta\mu$ is the sum of precession and nutation of the right ascension; the matrices (PR) and (NR) are the precession matrix and nutation matrix, respectively, given in Sect. 1.3.4.

- (2) Transformation between the J2000.0 orbital coordinate system and the J1950.0 orbital coordinate system

The J1950.0 orbital coordinate system uses the true equatorial plane at time t as the coordinate xy -plane, and the J1950.0 mean March equinox direction as the direction of the x -axis.

Assuming $\vec{r}_1$ and $\vec{r}_2$ to be the position vectors of a point in the J1950.0 orbital coordinate system and the J2000.0 orbital coordinate system, respectively, the transformation between them is simple, that

$$\vec{r}_2 = R_z(\mu_{50})\vec{r}_1, \quad (4.255)$$

where μ_{50} is the precession of the right ascension from J1950.0 to J2000.0. By the IAU 1980 model (1.39) of Chap. 1, there is

$$\mu_{50} = \zeta_A + z_A = 4612''.4362t + 1''.39656t^2, \quad (4.256)$$

where time t is given by

$$t = \frac{1}{36525.0}[\text{JD}(\text{J2000.0}) - \text{JD}(\text{J1950.0})]. \quad (4.257)$$

- (3) Transformation between the J2000.0 geocentric celestial coordinate system and the orbital coordinate system at time t

The orbital coordinate system at time t uses the true equatorial plane at time t as the coordinate xy -plane and the mean March equinox direction at time t as the direction

of the x -axis. This is the coordinate system currently used by American TLE elements data. Assuming $\vec{r}_t$ to be the position vector in this coordinate system, then the transformation between this system and the J2000.0 geocentric celestial coordinate system is given by

$$\vec{r}_t = R_z(\Delta\mu)(NR)(PR)\vec{r}_c. \quad (4.258)$$

The above transformations for the instantaneous position of the same point in different coordinate systems (only involve the precession and nutation of Earth's equator) can be also applied to the velocity at that point. By the transformation relationships of the position and velocity, we then can derive the transformation relationships of orbital elements according to the relationships between the orbital elements with the position vector and velocity.

In summary, the variety of coordinate systems and the complication of orbital element definitions have historical reasons and different requirements for a project. When we refer to information about satellite orbits, we must understand the related coordinate systems and the correct definitions of orbital variables to avoid man-made errors or even mistakes. The author suggests that to use a commonly used coordinate system, we should always consider what is generally applicable, rather than be restricted by specific reasons.

4.5 The Perturbation Orbit Solution Due to the Higher-Order Zonal Harmonic Terms J_l ($l \geq 3$) of Earth's Non-spherical Gravitation

4.5.1 General Expression of the Perturbation Function of the Zonal Harmonic Terms J_l ($l \geq 3$)

The perturbation function using the dimensionless variables is given by

$$\begin{cases} R_2 = \sum_{l \geq 3} R_l, \\ R_l = C_l \left(\frac{1}{r} \right)^{l+1} P_l(\mu) = -J_l \left(\frac{1}{r} \right)^{l+1} P_l(\mu), \end{cases} \quad (4.259)$$

where $\mu = \sin \varphi = \sin i \sin u$, and $u = f + \omega$. By definition there is

$$P_l(\mu) = \frac{1}{2^l l!} \frac{d^l}{d\mu^l} [(\mu^2 - 1)^l] = \frac{1}{2^l l!} \frac{d^l}{d\mu^l} [(-1)^l (1 - \mu^2)^l]. \quad (4.260)$$

Omitting the differentiating process, the result is

$$\begin{aligned}
P_l(\mu) &= \frac{1}{2^l l!} \sum_{m(2)=\delta_1}^l (-1)^{l+\frac{1}{2}(l+m)} \left(\frac{1}{2} \binom{l}{l+m} \right) [(l+m)(l+m-1) \cdots (m+1)] \mu^m \\
&= \frac{1}{2^l} \sum_{m(2)=\delta_1}^l (-1)^{\frac{1}{2}(l-m)} \left(\frac{1}{2} \binom{l}{l-m} \right) \binom{l+m}{l} \mu^m
\end{aligned} \tag{4.261}$$

where

$$\delta_1 = \frac{1}{2} [1 - (-1)^l] = \begin{cases} 1, & l \text{ is odd,} \\ 0, & l \text{ is even.} \end{cases} \tag{4.262}$$

In (4.261), the subscript $m(2) = \delta_1$ means if l is an odd number, $m = 1, 3, \dots, l$; and if l is an even number then $m = 0, 2, 4, \dots, l$. Substituting $P_l(\mu)$ into (4.259) and using the expression of $\sin^m u$ in multiple u forms, we have the general perturbation function of Earth's non-spherical potential of the l th-order zonal harmonic term in orbital elements that

$$\begin{aligned}
R_l &= \frac{(-J_l)}{a^{l+1}} \sum_{m(2)=\delta_1}^l \sum_{q=0}^{\frac{1}{2}(m-\delta_1)} (-1)^{\frac{1}{2}(l+2q-\delta_1)} \frac{2^{\delta_m}}{2^{l+m}} \left(\frac{1}{2} \binom{l}{l-m} \right) \binom{l+m}{l} \binom{m}{q} \sin^m i \\
&\quad \times \left[(1 - \delta_1) \left(\frac{a}{r} \right)^{l+1} \cos(m-2q)u + \delta_1 \left(\frac{a}{r} \right)^{l+1} \sin(m-2q)u \right],
\end{aligned} \tag{4.263}$$

where

$$\delta_m = \begin{cases} 0, & m-2q = 0, \\ 1, & m-2q \neq 0. \end{cases} \tag{4.264}$$

In the two sums of (4.263) for a given l the trigonometric functions $\cos(m-2q)u$ and $\sin(m-2q)u$ are calculated twice. To simplify the calculation, it is better to rearrange the terms by $(m-2q)u$ to make each trigonometric function only be calculated once. Since $(m-2q)$ is restricted by the value of l , we introduce p as

$$m - 2q = l - 2p,$$

then use p instead of m that

$$m = (l - 2p) + 2q, \quad (l - m) = 2p - 2q,$$

where p takes the value on $[0, (l - \delta_1)/2]$, and q depends on $\delta_1 \leq m = l - 2p + 2q \leq l$, i.e., on $[0, p]$, then (4.263) becomes

$$\begin{aligned}
R_l &= \frac{(-J_l)}{a^{l+1}} \sum_{p=0}^{\frac{1}{2}(l-\delta_1)} \sum_{q=0}^p (-1)^{\frac{1}{2}(l+2q-\delta_1)} \cdot 2^{-(2l-2p+2q-\delta_2)} \\
&\times \binom{l}{p-q} \binom{2l-2p+2q}{l} \binom{l-2p+2q}{q} (\sin i)^{l-2p+2q} \\
&\times \left[(1-\delta_1) \left(\frac{a}{r}\right)^{l+1} \cos(l-2p)u + \delta_1 \left(\frac{a}{r}\right)^{l+1} \sin(l-2p)u \right]
\end{aligned} \tag{4.265}$$

and δ_2 is given by

$$\delta_2 = \begin{cases} 0, & l-2p = 0, \\ 1, & l-2p \neq 0. \end{cases} \tag{4.266}$$

In the perturbation function (4.265) the argument in the trigonometric function $\sin(1-2p)u$ now only depends on p .

To separate the secular, short-period, and long-period parts of the perturbation function we need the averages of the following two terms:

$$\overline{\left(\frac{a}{r}\right)^{l+1} \sin(l-2p)f} = 0, \tag{4.267}$$

$$\overline{\left(\frac{a}{r}\right)^{l+1} \cos(l-2p)f} = \delta_3 (1-e^2)^{-(l-1/2)} \sum_{\alpha(2)=l-2p}^{l-2} \binom{l-1}{\alpha} \binom{n}{\frac{1}{2}(\alpha-l+2p)} \left(\frac{e}{2}\right)^\alpha, \tag{4.268}$$

$$\delta_3 = \begin{cases} 0, & p = 0, \\ 1, & p \neq 0, \end{cases} \tag{4.269}$$

where $\alpha(2)$, like $m(2)$, increases by a step length of 2. Then the secular, long-period, and short-period parts of R_2 for $J_l (l \geq 3)$ are

$$R_{2c} = \sum_{l(2) \geq 4} \frac{(-J_l)}{a^{l+1}} K_{l+1}(e) \sum_{q=0}^{l/2} (-1)^{\frac{1}{2}(l+2q)} \cdot 2^{-(l+2q)} \binom{l}{\frac{l}{2}-q} \binom{l+2q}{l} \binom{2q}{q} (\sin i)^{2q}, \tag{4.270}$$

$$\begin{aligned}
R_{2l} &= \sum_{l \geq 3} \frac{(-J_l)}{a^{l+1}} \sum_{p=1}^{\frac{1}{2}(l-2+\delta_1)} K_{l+1}^p(e) \left[\sum_{q=0}^p (-1)^{\frac{1}{2}(l+2q-\delta_1)} \cdot 2^{-(2l-2p+2q-1)} \right. \\
&\times \binom{l}{p-q} \binom{2l-2p+2q}{l} \binom{l-2p+2q}{q} (\sin i)^{l-2p+2q} \left. \right] \\
&\times [(1-\delta_1) \cos(l-2p)\omega + \delta_1 \sin(l-2p)\omega],
\end{aligned} \tag{4.271}$$

$$R_{2S} = R_2 - (R_{2c} + R_{2l}). \tag{4.272}$$

In (4.270) the sub of the first sum $l(2) \geq 4$ increases from 4 by a step-length of 2, i.e., $l = 4, 6, \dots$. The two auxiliaries $K_{l+1}(e)$, $K_{l+1}^p(e)$ are defined by

$$K_{l+1}(e) = \overline{\left(\frac{a}{r}\right)^{l+1}}, \quad (4.273)$$

$$K_{l+1}^p(e) = \overline{\left(\frac{a}{r}\right)^{l+1} \cos(l-2p)f}, \quad (4.274)$$

whose calculations are given by (4.268), and when $l-2p=0$ $K_{l+1}(e)$ is the same as $K_{l+1}^p(e)$, thus both can be calculated by one calculation program.

The expressions of R_{2c} and R_{2l} show that

- (1) Only the perturbation functions related to even-order zonal harmonic terms, $J_4, J_6, \dots$ have “secular” terms which are functions of a , e , and i only, because the corresponding values of p make $l-2p=0$; whereas for the odd-order zonal harmonic terms, the values of p cannot make $l-2p=0$.
- (2) The long-period terms of the even-order zonal harmonic terms involve

$$\cos 2\omega, \cos 4\omega, \dots, \cos(l-2)\omega,$$

whereas the long-period terms of the odd-order zonal harmonic terms involve

$$\sin \omega, \sin 3\omega, \dots, \sin(l-2)\omega.$$

These are general properties of the zonal harmonic terms in the perturbation function of Earth’s non-spherical potential, including the J_2 term.

4.5.2 The Perturbation Solution of the Zonal Harmonic $J_l(l \geq 3)$ Terms

- (1) Coefficients of secular terms, σ_2

$$a_2 = 0, \quad (4.275)$$

$$e_2 = 0, \quad (4.276)$$

$$i_2 = 0, \quad (4.277)$$

$$\begin{aligned}\Omega_2 &= n \cos i \sum_{l(2) \geq 4} \left(\frac{-J_l}{p_0^l} \right) \sum_{q=1}^{l/2} (-1)^{(l+2q)/2} \left(\frac{1}{2} \right)^{(l+2q)} \\ &\quad \times 2q \binom{l}{l/2-q} \binom{l/2+q}{l} \binom{2q}{q} (\sin i)^{(2q-2)} K_1(e),\end{aligned}\quad (4.278)$$

$$\begin{aligned}\omega_2 &= -\cos i \Omega_2 \\ &+ n \sum_{l(2) \geq 4} \left(\frac{-J_l}{p_0^l} \right) \sum_{q=0}^{l/2} (-1)^{(l+2q)/2} \left(\frac{1}{2} \right)^{(l+2q)} \\ &\quad \times \binom{l}{l/2-q} \binom{l/2+q}{l} \binom{2q}{q} (\sin i)^{2q} \\ &\quad \times [(2l-1)K_1(e) + (1-e^2)K_2(e)],\end{aligned}\quad (4.279)$$

$$\begin{aligned}M_2 &= -\sqrt{1-e^2}(\omega_2 + \cos i \Omega_2) \\ &+ n \sqrt{1-e^2} \sum_{l(2) \geq 4} \left(\frac{-J_l}{p_0^l} \right) \sum_{q=0}^{l/2} (-1)^{(l+2q)/2} \left(\frac{1}{2} \right)^{(l+2q)} 2(l+1) \\ &\quad \times \binom{l}{l/2-q} \binom{l/2+q}{l} \binom{2q}{q} (\sin i)^{2q} K_1(e),\end{aligned}\quad (4.280)$$

where

$$K_1(e) = \sum_{\alpha(2)=0}^{l-2} \binom{l-1}{\alpha} \binom{\alpha}{\alpha/2} \left(\frac{1}{2} \right)^\alpha e^\alpha, \quad (4.281)$$

$$K_2(e) = \sum_{\alpha(2)=2}^{l-2} \binom{l-1}{\alpha} \binom{\alpha}{\alpha/2} \alpha \left(\frac{1}{2} \right)^\alpha e^{\alpha-2}, \quad (4.282)$$

and $\alpha(2)$ has the same meaning in (4.268). The elements a , e , i , and n , p_0 are for $\bar{a}$, $\bar{e}$, $\bar{i}$, and $\bar{n} = \bar{a}_0^{-3/2}$, $\bar{p}_0 = \bar{a}_0(1 - \bar{e}_0^2)$.

(2) The long-period term $\sigma_l^{(1)}(t)$

The direct parts of the long-period term are

$$a_l^{(1)}(t) = 0, \quad (4.283)$$

$$e_l^{(1)}(t) = -\left(\frac{1-e^2}{e} \tan i \right) i_l^{(1)}(t)$$

$$\begin{aligned}
&= -(1 - e^2) \sum_{l \geq 3} \left(\frac{-J_l}{p_0^l} \right)^{\frac{1}{2}(l-2+\delta_1)} \sum_{p=1}^p \left[\sum_{q=1}^p (-1)^{(l+2q-\delta_1)/2} \left(\frac{1}{2} \right)^{(2l-2p+2q-1)} \right. \\
&\quad \times \binom{l}{p-q} \binom{2l-2p+2q}{l} \binom{l-2p+2q}{q} (\sin i)^{(l-2p+2q)} \left. \right] \frac{1}{e} K_3(e) I(\omega),
\end{aligned} \tag{4.284}$$

$$\begin{aligned}
i_l^{(1)}(t) &= \cos i \sum_{l \geq 3} \left(\frac{-J_l}{p_0^l} \right)^{\frac{1}{2}(l-2+\delta_1)} \sum_{p=1}^p \left[\sum_{q=1}^p (-1)^{(l+2q-\delta_1)/2} \left(\frac{1}{2} \right)^{(2l-2p+2q-1)} \right. \\
&\quad \times \binom{l}{p-q} \binom{2l-2p+2q}{l} \binom{l-2p+2q}{q} (\sin i)^{(l-2p+2q-1)} \left. \right] K_3(e) I(\omega),
\end{aligned} \tag{4.285}$$

$$\begin{aligned}
\Omega_l^{(1)}(t) &= \cos i \sum_{l \geq 3} \left(\frac{-J_l}{p_0^l} \right)^{\frac{1}{2}(l-2+\delta_1)} \sum_{p=1}^p \left[\sum_{q=1}^p (-1)^{(l+2q-\delta_1)/2} \left(\frac{1}{2} \right)^{(2l-2p+2q-1)} (l-2p+2q) \right. \\
&\quad \times \binom{l}{p-q} \binom{2l-2p+2q}{l} \binom{l-2p+2q}{q} (\sin i)^{(l-2p+2q-2)} \left. \right] K_3(e) H(\omega),
\end{aligned} \tag{4.286}$$

$$\begin{aligned}
\omega_l^{(1)}(t) &= -\cos i \Omega_l^{(1)}(t) \\
&\quad + \sum_{l \geq 3} \left(\frac{-J_l}{p_0^l} \right)^{\frac{1}{2}(l-2+\delta_1)} \sum_{p=1}^p \left[\sum_{q=1}^p (-1)^{(l+2q-\delta_1)/2} \left(\frac{1}{2} \right)^{(2l-2p+2q-1)} \right. \\
&\quad \times \binom{l}{p-q} \binom{2l-2p+2q}{l} \binom{l-2p+2q}{q} (\sin i)^{(l-2p+2q)} \left. \right] \\
&\quad \times [(2l-1)K_3(e) + (1-e^2)K_4(e)] H(\omega),
\end{aligned} \tag{4.287}$$

$$\begin{aligned}
M_l^{(1)}(t) &= -\sqrt{1-e^2} [\omega_l^{(1)}(t) + \cos i \Omega_l^{(1)}(t)] \\
&\quad + \sqrt{1-e^2} \sum_{l \geq 3} \left(\frac{-J_l}{p_0^l} \right)^{\frac{1}{2}(l-2+\delta_1)} \sum_{p=1}^p \left[\sum_{q=1}^p (-1)^{(l+2q-\delta_1)/2} \left(\frac{1}{2} \right)^{(2l-2p+2q-1)} 2(l+1) \right. \\
&\quad \times \binom{l}{p-q} \binom{2l-2p+2q}{l} \binom{l-2p+2q}{q} (\sin i)^{(l-2p+2q)} \left. \right] K_3(e) H(\omega).
\end{aligned} \tag{4.288}$$

For Ω , ω , and M , there are also indirect parts of their long-period terms which are

$$\begin{aligned}
\Omega_l^{(1)}(t) &= \frac{5 \cos i}{(2-5 \sin^2 i/2)} \sum_{l \geq 3} \left(\frac{-J_l}{p_0^l} \right)^{\frac{1}{2}(l-2+\delta_1)} \sum_{p=1}^p \left[\sum_{q=1}^p (-1)^{(l+2q-\delta_1)/2} \left(\frac{1}{2} \right)^{(2l-2p+2q-1)} \right. \\
&\quad \times \binom{l}{p-q} \binom{2l-2p+2q}{l} \binom{l-2p+2q}{q} (\sin i)^{(l-2p+2q)} \left. \right] K_3(e) H(\omega),
\end{aligned} \tag{4.289}$$

$$\begin{aligned} \omega_l^{(1)}(t) = & -\frac{(13 - 15 \sin^2 i)}{(2 - 5 \sin^2 i/2)} \sum_{l \geq 3} \left(\frac{-J_l}{p_0^l} \right)^{\frac{1}{2}(l-2+\delta_1)} \sum_{p=1}^p \left[\sum_{q=1}^p (-1)^{(l+2q-\delta_1)/2} \left(\frac{1}{2} \right)^{(2l-2p+2q-1)} \right. \\ & \times \binom{l}{p-q} \binom{2l-2p+2q}{l} \binom{l-2p+2q}{q} (\sin i)^{(l-2p+2q)} \left. \right] K_3(e) H(\omega), \end{aligned} \quad (4.290)$$

$$\begin{aligned} M_l^{(1)}(t) = & -3\sqrt{1-e^2} \sum_{l \geq 3} \left(\frac{-J_l}{p_0^l} \right)^{\frac{1}{2}(l-2+\delta_1)} \sum_{p=1}^p \left[\sum_{q=1}^p (-1)^{(l+2q-\delta_1)/2} \left(\frac{1}{2} \right)^{(2l-2p+2q-1)} \right. \\ & \times \binom{l}{p-q} \binom{2l-2p+2q}{l} \binom{l-2p+2q}{q} (\sin i)^{(l-2p+2q)} \left. \right] K_3(e) H(\omega). \end{aligned} \quad (4.291)$$

The new terms in the above long-period terms are

$$\begin{cases} K_3(e) = \sum_{\alpha(2)=l-2p}^{l-2} \binom{l-1}{\alpha} \binom{\alpha}{(\alpha-l+2p)/2} \left(\frac{1}{2} \right)^\alpha e^\alpha, \\ K_4(e) = \sum_{\alpha(2)=l-2p}^{l-2} \binom{l-1}{\alpha} \binom{\alpha}{(\alpha-l+2p)/2} \alpha \left(\frac{1}{2} \right)^\alpha e^{\alpha-2}, \end{cases} \quad (4.292)$$

$$\begin{cases} I(\omega) = \left(\frac{n}{\omega_1} \right) [(1 - \delta_1) \cos(l-2p)\omega - \delta_1 \sin(l-2p)\omega,] \\ H(\omega) = \left(\frac{n}{\omega_1} \right) \left[(1 - \delta_1) \frac{1}{l-2p} \sin(l-2p)\omega - \delta_1 \frac{1}{(l-2p)} \cos(l-2p)\omega \right]. \end{cases} \quad (4.293)$$

The elements a , e , i , and n , p_0 in the above formulas are the same as in σ_2 ; and ω is the mean element $\bar{\omega}(t)$ with the corresponding ω_1 , which is the coefficient of the first-order secular term given by (4.117).

(3) The short-period term $a_s^{(2)}(t)$

For the first-order solution, only the short-period term of the second-order for a is needed, which is given by

$$a_s^{(2)}(t) = \frac{2}{n^2 a} R_{2s} = 2a^2 R_{2s}, \quad (4.294)$$

where R_{2s} is given by (4.272), in which the terms $K_{l+1}(e)$ and $K_{l+1}^p(e)$ are given by (4.273) and (4.274) but can also be given by

$$K_{l+1}(e) = \overline{\left(\frac{a}{r} \right)^{l+1}} = (1 - e^2)^{-(l-\frac{1}{2})} K_1(e), \quad (4.295)$$

$$K_{l+1}^p(e) = \overline{\left(\frac{a}{r} \right)^{l+1} \cos(l-2p)f} = \delta_3 (1 - e^2)^{-(l-\frac{1}{2})} K_3(e), \quad (4.296)$$

where δ_3 is defined by (4.269).

4.5.3 The Non-singularity Perturbation Solution of the First Type by the Zonal Harmonic Terms $J_l (l \geq 3)$

The non-singularity perturbation solution of the first type is given using the quasi-mean orbital elements. The structures of the secular and long-period terms can be expressed by Kepler orbital elements as the follows.

$$\Delta a(t) = 0, \quad (4.297)$$

$$\Delta i(t) = \Delta i_l(t), \quad (4.298)$$

$$\Delta \Omega = \Omega_2(t - t_0) + \Delta \Omega_l(t), \quad (4.299)$$

$$\Delta \xi(t) = \cos \omega [\Delta e_l(t)] - \sin \omega [e(\omega_2(t - t_0) + \Delta \omega_l(t))], \quad (4.300)$$

$$\Delta \eta(t) = \sin \omega [\Delta e_l(t)] + \cos \omega [e(\omega_2(t - t_0) + \Delta \omega_l(t))], \quad (4.301)$$

$$\Delta \lambda(t) = [M_2(t - t_0) + \omega_2(t - t_0)] + [\Delta M_l(t) + \Delta \omega_l(t)], \quad (4.302)$$

where Ω_2 , ω_2 , and M_2 in the secular terms $\Delta \sigma(t)$ are given in (4.278)–(4.280), and

$$\begin{aligned} M_2 + \omega_2 = & -\cos i \Omega_2 \\ & + n \sqrt{1 - e^2} \sum_{l(2) \geq 4} \left(\frac{-J_l}{p_0^l} \right) \sum_{q=0}^{l/2} (-1)^{(l+2q)/2} \left(\frac{1}{2} \right)^{(l+2q)} \\ & \times 2(l+1) \binom{l}{l/2-q} \binom{l/2+q}{l} \binom{2q}{q} (\sin i)^{2q} K_1(e) \\ & + \left(\frac{e^2}{1 + \sqrt{1 - e^2}} \right) n \sum_{l(2) \geq 4} \left(\frac{-J_l}{p_0^l} \right) \sum_{q=0}^{l/2} (-1)^{(l+2q)/2} \left(\frac{1}{2} \right)^{(l+2q)} \\ & \times \binom{l}{l/2-q} \binom{l/2+q}{l} \binom{2q}{q} (\sin i)^{2q} [(2l-1)K_1(e) + (1-e^2)K_2(e)], \end{aligned} \quad (4.303)$$

where $\Delta \sigma_l(t) = \sigma_l^{(1)}(t) - \sigma_l^{(1)}(t_0)$ is the direct part of the long-period term, and $\sigma_l^{(1)}(t)$ is given by (4.283)–(4.288). Also, there is

$$\begin{aligned}
M_l^{(1)} + \omega_l^{(1)}(t) = & -\cos i \cdot \Omega_l^{(1)}(t) \\
& + \sqrt{1-e^2} \sum_{l \geq 3} \left(\frac{-J_l}{p_0^l} \right)^{\frac{1}{2}(l-2+\delta_1)} \sum_{p=1}^l \left[\sum_{q=1}^p (-1)^{(l+2q-\delta_1)/2} \left(\frac{1}{2} \right)^{(2l-2p+2q-1)} 2(l+1) \right. \\
& \times \left(\frac{l}{p-q} \right) \binom{2l-2p+2q}{l} \binom{l-2p+2q}{q} (\sin i)^{(l-2p+2q)} \left. \right] K_3(e) H(\omega) \\
& + \left(\frac{e^2}{1+\sqrt{1-e^2}} \right) \sum_{l \geq 3} \left(\frac{-J_l}{p_0^l} \right)^{\frac{1}{2}(l-2+\delta_1)} \sum_{p=1}^l \left[\sum_{q=1}^p (-1)^{(l+2q-\delta_1)/2} \left(\frac{1}{2} \right)^{(2l-2p+2q-1)} \right. \\
& \times \left(\frac{l}{p-q} \right) \binom{2l-2p+2q}{l} \binom{l-2p+2q}{q} (\sin i)^{(l-2p+2q)} \left. \right] \\
& \times \left[(2l-1)K_3(e) + (1-e^2)K_4(e) \right] H(\omega). \tag{4.304}
\end{aligned}$$

Considering the commensurable small divisors caused by the critical inclination, the corresponding $I(\omega)$ and $H(\omega)$ have two formulas to calculate given by

$$\begin{aligned}
I(\omega) = & \left(\frac{n}{\omega_1} \right) [(1-\delta_1) \cos(l-2p)\omega - \delta_1 \sin(l-2p)\omega]_{t_0}^t \\
= & -[(1-\delta_1) \sin(l-2p)\omega + \delta_1 \cos(l-2p)\omega](l-2p)n(t-t_0), \tag{4.305}
\end{aligned}$$

$$\begin{aligned}
H(\omega) = & \left(\frac{n}{\omega_1} \right) \left[(1-\delta_1) \frac{1}{l-2p} \sin(l-2p)\omega - \delta_1 \frac{1}{(l-2p)} \cos(l-2p)\omega \right]_{t_0}^t \\
= & [(1-\delta_1) \cos(l-2p)\omega + \delta_1 \sin(l-2p)\omega]n(t-t_0), \tag{4.306}
\end{aligned}$$

The first formulas in (4.305) and (4.306) are for the normal case and the second formulas are for the case of critical inclination.

The short-period term $a_s^{(2)}(t)$ is the same as given by (4.294).

4.5.4 The Non-singularity Perturbation Solution of the Second Type by Zonal Harmonic Terms $J_l (l \geq 3)$

Similar to the non-singularity perturbation solution of the first type, structures of the perturbation terms can be expressed by Kepler orbital elements as follows.

$$\Delta a(t) = 0, \tag{4.307}$$

$$\begin{aligned}
\Delta \xi(t) = & \cos \tilde{\omega} [\Delta e_l(t)] \\
& - e \sin \tilde{\omega} [e(\omega_2(t-t_0) + \Delta \omega_l(t)) + e(\Omega_2(t-t_0) + \Delta \Omega_l(t))] \tag{4.308}
\end{aligned}$$

$$\begin{aligned}\Delta\eta(t) &= \sin \tilde{\omega}[\Delta e_l(t)] \\ &\quad + e \cos \tilde{\omega}[e(\omega_2(t - t_0) + \Delta\omega_l(t)) + e(\Omega_2(t - t_0) + \Delta\Omega_l(t))] \quad (4.309)\end{aligned}$$

$$\Delta h(t) = \frac{1}{2} \left(\cos \frac{i}{2} \cos \Omega \right) [\Delta i_l(t)] - \left(\sin \frac{i}{2} \sin \Omega \right) [\Omega_2(t - t_0) + \Delta\Omega_l(t)] \quad (4.310)$$

$$\Delta k(t) = \frac{1}{2} \left(\cos \frac{i}{2} \sin \Omega \right) [\Delta i_l(t)] + \left(\sin \frac{i}{2} \cos \Omega \right) [\Omega_2(t - t_0) + \Delta\Omega_l(t)] \quad (4.311)$$

$$\Delta\lambda(t) = [M_2(t - t_0) + \omega_2(t - t_0) + \Omega_2(t - t_0)] + [\Delta M_l(t) + \Delta\omega_l(t) + \Delta\Omega_l(t)] \quad (4.312)$$

where Ω_2 , ω_2 , and M_2 are given by (4.278)–(4.280), and

$$\begin{aligned}M_2 + \omega_2 + \Omega_2 &= \frac{\sin^2 i}{1 + \cos i} \Omega_2 + n \sqrt{1 - e^2} \sum_{l(2) \geq 4} \left(\frac{-J_l}{p_0^l} \right) \sum_{q=0}^{l/2} (-1)^{(l+2q)/2} \left(\frac{1}{2} \right)^{(l+2q)} 2(l+1) \\ &\quad \times \binom{l}{l/2-q} \binom{l/2+q}{l} \binom{2q}{q} (\sin i)^{2q} K_1(e) \\ &\quad + \left(\frac{e^2}{1 + \sqrt{1 - e^2}} \right) n \sum_{l(2) \geq 4} \left(\frac{-J_l}{p_0^l} \right) \sum_{q=0}^{l/2} (-1)^{(l+2q)/2} \left(\frac{1}{2} \right)^{(l+2q)} \\ &\quad \times \binom{l}{l/2-q} \binom{l/2+q}{l} \binom{2q}{q} (\sin i)^{2q} [(2l-1)K_1(e) + (1-e^2)K_2(e)] \quad (4.313)\end{aligned}$$

Similarly, $\Delta\sigma_l(t) = \sigma_l^{(1)}(t) - \sigma_l^{(1)}(t_0)$ is the direct part of the long-period term $\sigma_l^{(1)}(t)$ given by (4.283)–(4.288), and

$$\begin{aligned}M_l^{(1)} + \omega_l^{(1)} + \Omega_l^{(1)} &= \frac{\sin^2 i}{1 + \cos i} \cdot \Omega_l^{(1)}(t) \\ &\quad + \sqrt{1 - e^2} \sum_{l \geq 3} \left(\frac{-J_l}{p_0^l} \right)^{\frac{1}{2}(l-2+\delta_1)} \sum_{p=1}^{l/2} \left[\sum_{q=1}^p (-1)^{(l+2q-\delta_1)/2} \left(\frac{1}{2} \right)^{(2l-2p+2q-1)} 2(l+1) \right. \\ &\quad \times \binom{l}{p-q} \binom{2l-2p+2q}{l} \binom{l-2p+2q}{q} (\sin i)^{(l-2p+2q)} \left. \right] K_3(e) H(\omega) \\ &\quad + \left(\frac{e^2}{1 + \sqrt{1 - e^2}} \right) \sum_{l \geq 3} \left(\frac{-J_l}{p_0^l} \right)^{\frac{1}{2}(l-2+\delta_1)} \sum_{p=1}^{l/2} \left[\sum_{q=1}^p (-1)^{(l+2q-\delta_1)/2} \left(\frac{1}{2} \right)^{(2l-2p+2q-1)} \right. \\ &\quad \times \binom{l}{p-q} \binom{2l-2p+2q}{l} \binom{l-2p+2q}{q} (\sin i)^{(l-2p+2q)} \left. \right] \\ &\quad \times [(2l-1)K_3(e) + (1-e^2)K_4(e)] H(\omega). \quad (4.314)\end{aligned}$$

Again, the commensurable small divisor is caused by the critical inclination, the corresponding $I(\omega)$ and $H(\omega)$ have two formulas to calculate given by (4.305) and (4.306), here we provide them again,

$$\begin{aligned} I(\omega) &= \left(\frac{n}{\omega_1} \right) [(1 - \delta_1) \cos(l - 2p)\omega - \delta_1 \sin(l - 2p)\omega]_{t_0}^t \\ &= -[(1 - \delta_1) \sin(l - 2p)\omega + \delta_1 \cos(l - 2p)\omega](l - 2p)n(t - t_0), \end{aligned} \quad (4.315)$$

$$\begin{aligned} H(\omega) &= \left(\frac{n}{\omega_1} \right) \left[(1 - \delta_1) \frac{1}{l - 2p} \sin(l - 2p)\omega - \delta_1 \frac{1}{(l - 2p)} \cos(l - 2p)\omega \right]_{t_0}^t \\ &= [(1 - \delta_1) \cos(l - 2p)\omega + \delta_1 \sin(l - 2p)\omega]n(t - t_0), \end{aligned} \quad (4.316)$$

in which the first formulas are for the normal case and the second formulas for the case of critical inclination.

The short-period term $a_s^{(2)}(t)$ is the same as given by (4.294).

4.5.5 The Perturbation Solution of the Main Zonal Harmonic Terms J_3 and J_4 in Kepler Elements

4.5.5.1 The Perturbation Function of the Main Zonal Harmonic Terms

For a low Earth orbit (LEO) satellite the dynamical form-factor J_2 and the zonal harmonic terms, J_3 and J_4 , make the main part of Earth's non-spherical perturbation. For some aerospace projects, the effect of the three terms is regarded as the effect of the whole zonal harmonic terms. In this section, we give the perturbation solution caused by the three terms.

The perturbation function of the J_2 , J_3 , and J_4 terms is

$$\begin{aligned} R &= -\frac{J_2}{r^3} P_2(\sin \varphi) - \frac{J_3}{r^4} P_3(\sin \varphi) - \frac{J_4}{r^5} P_4(\sin \varphi) \\ &= -\frac{J_2}{r^3} \left(\frac{3}{2} \sin^2 \varphi - \frac{1}{2} \right) - \frac{J_3}{r^4} \left(\frac{5}{2} \sin^3 \varphi - \frac{3}{2} \sin \varphi \right) - \frac{J_4}{r^5} \left(\frac{35}{8} \sin^4 \varphi - \frac{15}{4} \sin^2 \varphi + \frac{3}{8} \right). \end{aligned}$$

By the relationship of $\sin \varphi = \sin i \sin(f + \omega)$, the perturbation function R can be expressed in the satellite's orbital elements that

$$\begin{aligned}
R = & \frac{3J_2}{2a^3} \left(\frac{a}{r}\right)^3 \left[\frac{1}{3} \left(1 - \frac{3}{2} \sin^2 i\right) + \frac{1}{2} \sin^2 i \cos 2(f + \omega) \right] \\
& + \frac{J_3}{a^4} \left(\frac{a}{r}\right)^4 \sin i \left[\frac{3}{4} \left(2 - \frac{5}{2} \sin^2 i\right) \sin(f + \omega) + \frac{5}{8} \sin^2 i \sin 3(f + \omega) \right] \\
& - \frac{35J_4}{8a^5} \left(\frac{a}{r}\right)^5 \left[\left(\frac{3}{35} - \frac{3}{7} \sin^2 i + \frac{3}{8} \sin^4 i\right) \right. \\
& \left. + \left(\frac{3}{7} - \frac{1}{2} \sin^2 i\right) \sin^2 i \cos 2(f + \omega) + \frac{1}{8} \sin^4 i \cos 4(f + \omega) \right].
\end{aligned} \tag{4.317}$$

The J_2 term is regarded as the small parameter of the first order, then J_3 and J_4 have magnitudes of the second order, therefore R can be expressed as

$$R = R_1(J_2) + R_2(J_3, J_4).$$

We then decompose R into the secular parts R_{1c} and R_{2c} , the long-period parts R_{1l} and R_{2l} , and the short-period parts R_{1s} and R_{2s} . The parts corresponding to J_3 and J_4 in $R_2(J_3, J_4)$ are

$$R_{2c} = -\frac{35J_4}{8a^3} \left(\frac{3}{35} - \frac{3}{7} \sin^2 i + \frac{3}{8} \sin^4 i\right) \left(1 + \frac{3}{2}e^2\right) (1 - e^2)^{-7/2}, \tag{4.318}$$

$$\begin{aligned}
R_{2l} = & \frac{3J_3}{4a^4} \sin i \left(2 - \frac{5}{2} \sin^2 i\right) (1 - e^2)^{-5/2} e \sin \omega \\
& - \frac{35J_4}{8a^5} \sin^2 i \left(\frac{9}{28} - \frac{3}{8} \sin^2 i\right) (1 - e^2)^{-7/2} e^2 \cos 2\omega,
\end{aligned} \tag{4.319}$$

$$\begin{aligned}
R_{2s} = & \frac{J_3}{a^4} \left\{ \frac{3}{4} \sin i \left(2 - \frac{5}{2} \sin^2 i\right) \left[\left(\frac{a}{r}\right)^4 \sin(f + \omega) - (1 - e^2)^{-5/2} e \sin \omega \right] \right. \\
& \left. + \frac{5}{8} \sin^3 i \left(\frac{a}{r}\right)^4 \sin 3(f + \omega) \right\} \\
& - \frac{35J_4}{8a^5} \left\{ \left(\frac{3}{35} - \frac{3}{7} \sin^2 i + \frac{3}{8} \sin^4 i\right) \left[\left(\frac{a}{r}\right)^5 - \left(1 + \frac{3}{2}e^2\right) (1 - e^2)^{-7/2} \right] \right. \\
& + \sin^2 i \left(\frac{3}{7} - \frac{1}{2} \sin^2 i\right) \left[\left(\frac{a}{r}\right)^5 \cos 2(f + \omega) - \frac{3}{4} (1 - e^2)^{-7/2} e^2 \cos 2\omega \right] \\
& \left. + \frac{1}{8} \sin^4 i \left(\frac{a}{r}\right)^5 \cos 4(f + \omega) \right\}.
\end{aligned} \tag{4.320}$$

4.5.5.2 The Perturbation Motion Equation System of the Main Zonal Harmonic Terms

The set of the six orbital elements is given by σ that

$$\sigma = (a, e, i, \Omega, \omega, M)^T.$$

The corresponding perturbation motion equation system with the initial condition is given by

$$\begin{cases} \frac{d\sigma}{dt} = f_0(a) + f_1(\sigma, t, J_2) + f_2(\sigma, t, J_3, J_4), \\ \sigma(t_0) = \sigma_0, \end{cases} \quad (4.321)$$

where $f_0 = (0, 0, 0, 0, 0, n)^T$.

Adapting the normalized units of dimensionless variables, the right-side function of (4.321), $f_2(\sigma, t, J_3, J_4)$ has the forms as

$$\begin{cases} f_{2c} = (0, 0, 0, (f_{2c})_\Omega, (f_{2c})_\omega, (f_{2c})_M)^T, \\ f_{2l} = (0, (f_{2l})_e, (f_{2l})_i, (f_{2l})_\Omega, (f_{2l})_\omega, (f_{2l})_M)^T, \\ f_{2s} = ((f_{2s})_a, (f_{2s})_e, (f_{2s})_i, (f_{2s})_\Omega, (f_{2s})_\omega, (f_{2s})_M)^T, \end{cases} \quad (4.322)$$

$$(f_{2c})_\Omega = \frac{35J_4}{8p^4}n \cos i \left[\left(\frac{6}{7} + \frac{9}{7}e^2 \right) - \sin^2 i \left(\frac{3}{2} + \frac{9}{4}e^2 \right) \right], \quad (4.323)$$

$$(f_{2c})_\omega = -\frac{35J_4}{8p^4}n \left[\left(\frac{12}{7} + \frac{27}{14}e^2 \right) - \sin^2 i \left(\frac{93}{14} + \frac{27}{4}e^2 \right) + \sin^4 i \left(\frac{21}{4} + \frac{81}{16}e^2 \right) \right] \quad (4.324)$$

$$(f_{2c})_M = -\frac{35J_4}{8p^4}n \sqrt{1-e^2} \left[e^2 \left(\frac{9}{14} - \frac{45}{14} \sin^2 i + \frac{45}{16} \sin^4 i \right) \right] \quad (4.325)$$

$$\begin{aligned} (f_{2l})_e &= -\frac{1-e^2}{e} \tan i (f_{2l})_i \\ &= -\frac{J_3}{p^3}n \left[\sin i \left(\frac{3}{2} - \frac{15}{8} \sin^2 i \right) \right] (1-e^2) \cos \omega \\ &\quad - \frac{35J_4}{8p^4}n \left[\sin^2 i \left(\frac{9}{14} - \frac{3}{4} \sin^2 i \right) \right] (1-e^2) e \sin 2\omega \end{aligned} \quad (4.326)$$

$$\begin{aligned} (f_{2l})_i &= \frac{J_3}{p^3}n \cos i \left[\left(\frac{3}{2} - \frac{15}{8} \sin^2 i \right) \right] e \cos \omega \\ &\quad + \frac{35J_4}{8p^4}n \cos i \left[\sin i \left(\frac{9}{14} - \frac{3}{4} \sin^2 i \right) \right] e^2 \sin 2\omega \end{aligned} \quad (4.327)$$

$$\begin{aligned} (f_{2l})_\Omega &= \frac{J_3}{p^3}n \cos i \left[\frac{1}{\sin i} \left(\frac{3}{2} - \frac{45}{8} \sin^2 i \right) \right] e \sin \omega \\ &\quad - \frac{35J_4}{8p^4}n \cos i \left[\left(\frac{9}{14} - \frac{3}{2} \sin^2 i \right) \right] e^2 \cos 2\omega \end{aligned} \quad (4.328)$$

$$\begin{aligned}
(f_{2l})_{\omega} = & \frac{J_3}{p^3} n \frac{1}{e \sin i} \left[\frac{3}{4} \sin^2 i \left(2 - \frac{5}{2} \sin^2 i \right) - \frac{3}{8} e^2 \left(4 - 35 \sin^2 i + 35 \sin^4 i \right) \right] \sin \omega \\
& - \frac{35 J_4}{8 p^4} n \left[\sin^2 i \left(\frac{9}{14} - \frac{3}{4} \sin^2 i \right) - e^2 \left(\frac{9}{14} - \frac{15}{4} \sin^2 i + \frac{27}{8} \sin^2 i \right) \right] \cos 2\omega
\end{aligned} \quad (4.329)$$

$$\begin{aligned}
(f_{2l})_M = & -\frac{J_3}{p^3} n \frac{\sqrt{1-e^2}}{e \sin i} \left[\frac{3}{4} \sin^2 i \left(2 - \frac{5}{2} \sin^2 i \right) (1 - 4e^2) \right] \sin \omega \\
& + \frac{35 J_4}{8 p^4} n \sqrt{1-e^2} \left[\sin^2 i \left(\frac{9}{14} - \frac{3}{4} \sin^2 i \right) \left(1 - \frac{5}{2} e^2 \right) \right] \cos 2\omega
\end{aligned} \quad (4.330)$$

where p is the semi-latus rectum that $p = a(1-e^2)$.

Usually, when we analyze qualitatively or calculate quantitatively the variation of satellite orbital motion, the only short-period variation of the second-order needed is that of the semi-major axis a , the others can be omitted. Therefore, we do not provide the formula for $f_{2s}(\sigma, t, J_3, J_4)$, and the formula of $a_s^{(2)}(t)$ is given in (4.350).

4.5.5.3 Perturbation Solution of the First Order Due to J_2, J_3 , and J_4

The configuration of perturbation solution in the quasi-mean orbital elements is

$$\sigma(t) = \bar{\sigma}(t) + \sigma_s^{(1)}, \quad (4.331)$$

$$\bar{\sigma}(t) = \bar{\sigma}^{(0)}(t) + \sigma_1(t - t_0) + \sigma_2(t - t_0) + \Delta\sigma_l^{(1)}(t), \quad (4.332)$$

$$\begin{cases} \bar{\sigma}^{(0)}(t) = \bar{\sigma}_0 + \delta\bar{n}(t - t_0), \\ \bar{\sigma}_0 = \bar{\sigma}(t_0) = \sigma_0 - [\sigma_s^{(1)}(t_0) + \sigma_s^{(2)}(t_0)], \end{cases} \quad (4.333)$$

where $\sigma_s^{(2)}(t_0)$ only contains $a_s^{(2)}(t_0)$. The terms of J_3 and J_4 , and their mix-terms with J_2 are given as follows.

(1) $\sigma_2(\sigma, t, J_4)$

The J_3 term does not have a direct secular effect, the effects of the J_4 term are

$$a_2 = 0, \quad e_2 = 0, \quad i_2 = 0, \quad (4.334)$$

$$\Omega_2 = -\left(\frac{35}{8}\right) \frac{(-J_4)}{p^4} n \cos i \left[\left(\frac{6}{7} + \frac{9}{7} e^2 \right) - \sin^2 i \left(\frac{3}{2} + \frac{9}{4} e^2 \right) \right], \quad (4.335)$$

$$\omega_2 = \left(\frac{35}{8}\right) \frac{(-J_4)}{p^4} n \left\{ \left(\frac{12}{7} + \frac{27}{14} e^2 \right) - \sin^2 i \left(\frac{93}{14} + \frac{27}{4} e^2 \right) + \sin^4 i \left(\frac{21}{4} + \frac{81}{16} e^2 \right) \right\}, \quad (4.336)$$

$$M_2 = \left(\frac{35}{8}\right) \frac{(-J_4)}{p^4} n e^2 \sqrt{1-e^2} \left\{ \frac{9}{14} - \frac{45}{14} \sin^2 i + \frac{45}{16} \sin^4 i \right\}, \quad (4.337)$$

$$(2) \quad \sigma_l^{(1)}(\sigma, t, J_3, J_4)$$

The effects of the J_3 term are

$$a_l^{(1)}(t) = 0, \quad (4.338)$$

$$e_l^{(1)}(t) = \frac{3}{4} \left(\frac{-J_3}{p^3} \right) \sin i \left(2 - \frac{5}{2} \sin^2 i \right) \left(\frac{n}{\omega_1} \right) (1 - e^2) \sin \omega, \quad (4.339)$$

$$i_l^{(1)}(t) = -\frac{3}{4} \left(\frac{-J_3}{p^3} \right) \cos i \left(2 - \frac{5}{2} \sin^2 i \right) \left(\frac{n}{\omega_1} \right) e \sin \omega, \quad (4.340)$$

$$\Omega_l^{(1)}(t) = \frac{3}{4} \left(\frac{-J_3}{p^3} \right) \frac{\cos i}{\sin i} \left(2 - \frac{15}{2} \sin^2 i \right) \left(\frac{n}{\omega_1} \right) e \cos \omega, \quad (4.341)$$

$$\begin{aligned} \omega_l^{(1)}(t) &= \frac{3}{4} \left(\frac{-J_3}{p^3} \right) \frac{1}{\sin i} \\ &\times \left[\sin^2 i \left(2 - \frac{5}{2} \sin^2 i \right) - \frac{1}{2} e^2 (4 - 35 \sin^2 i + 35 \sin^4 i) \right] \left(\frac{n}{\omega_1} \right) \frac{1}{e} \cos \omega, \end{aligned} \quad (4.342)$$

$$M_l^{(1)}(t) = \frac{3}{4} \left(\frac{J_3}{p^3} \right) \sqrt{1-e^2} \sin i \left[\left(2 - \frac{5}{2} \sin^2 i \right) (1 - 4e^2) \right] \left(\frac{n}{\omega_1} \right) \frac{1}{e} \cos \omega. \quad (4.343)$$

The effects of the J_4 term are

$$a_l^{(1)}(t) = 0, \quad (4.344)$$

$$e_l^{(1)}(t) = -\frac{35}{4} \left(\frac{-J_4}{p^4} \right) \sin^2 i \left(\frac{9}{14} - \frac{3}{4} \sin^2 i \right) \left(\frac{n}{2\omega_1} \right) e (1 - e^2) \cos 2\omega, \quad (4.345)$$

$$i_l^{(1)}(t) = \frac{35}{8} \left(\frac{-J_4}{p^4} \right) \cos i \left[\sin i \left(\frac{9}{14} - \frac{3}{4} \sin^2 i \right) \right] \left(\frac{n}{2\omega_1} \right) e^2 \cos 2\omega, \quad (4.346)$$

$$\Omega_l^{(1)}(t) = \frac{35}{8} \left(\frac{-J_4}{p^4} \right) \cos i \left(\frac{9}{14} - \frac{3}{2} \sin^2 i \right) \left(\frac{n}{2\omega_1} \right) e^2 \sin 2\omega, \quad (4.347)$$

$$\omega_l^{(1)}(t) = -\frac{35J_4}{16p^4} \times \left[\sin^2 i \left(\frac{9}{14} - \frac{3}{4} \sin^2 i \right) - e^2 \left(\frac{9}{14} - \frac{15}{4} \sin^2 i + \frac{27}{8} \sin^2 i \right) \right] \left(\frac{n}{2\omega_1} \right) \sin 2\omega \quad (4.348)$$

$$M_l^{(1)}(t) = \frac{35}{16} \left(\frac{J_4}{p^4} \right) \sqrt{1-e^2} \left[\sin^2 i \left(\frac{9}{14} - \frac{3}{4} \sin^2 i \right) \left(1 - \frac{5}{2} e^2 \right) \right] \left(\frac{n}{2\omega_1} \right) \sin 2\omega. \quad (4.349)$$

On the right sides of the above formulas, all elements, a , e , i , and related n and p , are quasi-mean elements, which are

$$\begin{cases} \bar{a} = \bar{a}_0, & \bar{e} = \bar{e}_0, & \bar{i} = \bar{i}_0 \\ \bar{p} = \bar{p}_0 = \bar{a}_0(1 - \bar{e}_0^2), & \bar{n}_0 = \bar{a}_0^{-3/2} \end{cases}$$

and ω_1 is the variation rate of the first-order secular term of ω given in (4.66).

(3) $a_s^{(2)}(\sigma, t, J_3, J_4)$

The part of $a_s^{(2)}(t)$ by J_2 is given in (4.138) or (4.140), and the part by J_3 and J_4 is given by

$$a_s^{(2)}(t, J_3, J_4) = \int^t (f_{2s})_a dt = \int^t \frac{2}{n^2 a} \frac{\partial R_{2s}}{\partial M} dM = \frac{2}{n^2 a} R_{2s}(J_3, J_4), \quad (4.350)$$

where $R_{2s}(J_3, J_4)$ is given in (4.320).

4.5.5.4 The Non-singularity Perturbation Solution by the Main Zonal Harmonic Terms (J_3, J_4)

(1) The non-singularity perturbation solution of the first type

① Expressions in Kepler orbital elements.

By the same method for J_2 perturbation, the forms of secular terms (including the long-period terms) are given by

$$\Delta a(t) = 0, \quad (4.351)$$

$$\Delta i(t) = \Delta i_l(t), \quad (4.352)$$

$$\Delta \Omega(t) = \Delta \Omega_c(t) + \Delta \Omega_l(t), \quad (4.353)$$

$$\Delta \xi(t) = \cos \omega[\Delta e(t)] - \sin \omega[e \Delta \omega(t)], \quad (4.354)$$

$$\Delta \eta(t) = \sin \omega[\Delta e(t)] + \cos \omega[e \Delta \omega(t)], \quad (4.355)$$

$$\Delta \lambda(t) = \bar{n}(t - t_0) + [\Delta M(t) + \Delta \omega(t)], \quad (4.356)$$

where $\bar{n} = \bar{a}^{-3/2} = \bar{a}_0^{-3/2}$, and $\Delta \sigma(t)$ includes the secular term and the long-period term.

The only short-period term of the second order needed is $a_s^{(2)}(t)$ given in (4.350).

② The long-period terms in $\Delta \sigma_l^{(1)}(t) = \sigma_l^{(1)}(t) - \sigma_l^{(1)}(t_0)$ are

$$\Delta a_l^{(1)}(t) = 0 \quad (4.357)$$

$$\begin{aligned} e_l^{(1)}(t) = & \frac{3}{4} \left(\frac{-J_3}{p^3} \right) \sin i \left(2 - \frac{5}{2} \sin^2 i \right) \left(\frac{n}{\omega_1} \right) (1 - e^2) \sin \omega \\ & - \frac{35}{8} \left(\frac{-J_4}{p^4} \right) \sin^2 i \left(\frac{9}{14} - \frac{3}{4} \sin^2 i \right) \left(\frac{n}{2\omega_1} \right) e (1 - e^2) \cos 2\omega \end{aligned} \quad (4.358)$$

$$\begin{aligned} i_l^{(1)}(t) = & -\frac{3}{4} \left(\frac{-J_3}{p^3} \right) \cos i \left(2 - \frac{5}{2} \sin^2 i \right) \left(\frac{n}{\omega_1} \right) e \sin \omega \\ & + \frac{35}{8} \left(\frac{-J_4}{p^4} \right) \cos i \left[\sin i \left(\frac{9}{14} - \frac{3}{4} \sin^2 i \right) \right] \left(\frac{n}{2\omega_1} \right) e^2 \cos 2\omega \end{aligned} \quad (4.359)$$

$$\begin{aligned} \Omega_l^{(1)}(t) = & \frac{3}{4} \left(\frac{-J_3}{p^3} \right) \frac{\cos i}{\sin i} \left(2 - \frac{15}{2} \sin^2 i \right) \left(\frac{n}{\omega_1} \right) e \cos \omega \\ & + \frac{35}{8} \left(\frac{-J_4}{p^4} \right) \left(\frac{9}{14} - \frac{3}{2} \sin^2 i \right) \left(\frac{n}{2\omega_1} \right) e^2 \sin 2\omega, \end{aligned} \quad (4.360)$$

$$\begin{aligned} \omega_l^{(1)}(t) = & \frac{3}{4} \left(\frac{-J_3}{p^3} \right) \frac{1}{\sin i} \left[\sin^2 i \left(2 - \frac{5}{2} \sin^2 i \right) - e^2 \left(2 - \frac{35}{2} \sin^2 i + \frac{35}{2} \sin^4 i \right) \right] \left(\frac{n}{\omega_1} \right) \frac{1}{e} \cos \omega \\ & - \frac{35}{16} \left(\frac{J_4}{p^4} \right) \left[\sin^2 i \left(\frac{9}{14} - \frac{3}{4} \sin^2 i \right) - e^2 \left(\frac{9}{14} - \frac{15}{4} \sin^2 i + \frac{27}{8} \sin^2 i \right) \right] \left(\frac{n}{2\omega_1} \right) \sin 2\omega, \end{aligned} \quad (4.361)$$

$$\begin{aligned} M_l^{(1)}(t) = & \frac{3}{4} \left(\frac{J_3}{p^3} \right) \sqrt{1 - e^2} \sin i \left[\left(2 - \frac{5}{2} \sin^2 i \right) (1 - 4e^2) \right] \left(\frac{n}{\omega_1} \right) \frac{1}{e} \cos \omega \\ & + \frac{35}{16} \left(\frac{J_4}{p^4} \right) \sqrt{1 - e^2} \left[\sin^2 i \left(\frac{9}{14} - \frac{3}{4} \sin^2 i \right) \left(1 - \frac{5}{2} e^2 \right) \right] \left(\frac{n}{2\omega_1} \right) \sin 2\omega. \end{aligned} \quad (4.362)$$

The above non-singularity solution no longer has the factor $(1/e)$, but $\Delta \sigma_l^{(1)}(t)$ in the secular term still has a small divisor due to the critical inclination. In order

to eliminate the small divisor, we apply the same method used for J_2 perturbation. There are two calculation methods given later in (4.369)–(4.370).

③ The long-period terms $\Delta\sigma_l^{(1)}(t) = \sigma_l^{(1)}(t) - \sigma_l^{(1)}(t_0)$

$$\Delta a_l^{(1)}(t) = 0, \quad (4.363)$$

$$\begin{aligned} \Delta e_l^{(1)}(t) = & \frac{1}{2p} \left(\frac{J_3}{J_2} \right) \sin i \left(2 - \frac{5}{2} \sin^2 i \right) (1 - e^2) G_{1c} \\ & + \frac{35}{8p^2} \left(\frac{J_4}{J_2} \right) \sin^2 i \left(\frac{3}{7} - \frac{1}{2} \sin^2 i \right) e (1 - e^2) G_{2s} \end{aligned} \quad (4.364)$$

$$\begin{aligned} \Delta i_l^{(1)}(t) = & -\frac{1}{2p} \left(\frac{J_3}{J_2} \right) \cos i \left(2 - \frac{5}{2} \sin^2 i \right) e G_{1c} \\ & - \frac{35}{8p^2} \left(\frac{J_4}{J_2} \right) \cos i \left[\sin i \left(\frac{3}{7} - \frac{1}{2} \sin^2 i \right) \right] e^2 G_{2s} \end{aligned} \quad (4.365)$$

$$\begin{aligned} \Delta \Omega_l^{(1)}(t) = & -\frac{1}{p} \left(\frac{J_3}{J_2} \right) \frac{\cos i}{\sin i} \left(1 - \frac{15}{4} \sin^2 i \right) e G_{1s} \\ & + \frac{35}{8p^2} \left(\frac{J_4}{J_2} \right) \left(\frac{3}{7} - \sin^2 i \right) e^2 G_{2c} \end{aligned} \quad (4.366)$$

$$\begin{aligned} e \Delta \omega_l^{(1)}(t) = & -\frac{1}{2p} \left(\frac{J_3}{J_2} \right) \frac{1}{\sin i} \left[\sin^2 i \left(2 - \frac{5}{2} \sin^2 i \right) - \frac{1}{2} e^2 (4 - 35 \sin^2 i + 35 \sin^4 i) \right] G_{1s} \\ & + \frac{35}{16p^2} \left(\frac{J_4}{J_2} \right) \left[\sin^2 i \left(\frac{3}{7} - \frac{1}{2} \sin^2 i \right) - e^2 \left(\frac{3}{7} - \frac{5}{2} \sin^2 i + \frac{9}{4} \sin^4 i \right) \right] e G_{2c} \end{aligned} \quad (4.367)$$

$$\begin{aligned} \Delta M_l^{(1)}(t) + \omega_l^{(1)}(t) = & \frac{1}{2p} \left(\frac{J_3}{J_2} \right) \frac{1}{\sin i} \left[\left(2 - \frac{35}{2} \sin^2 i + \frac{35}{2} \sin^4 i \right) - \sin^2 i \left(2 - \frac{5}{2} \sin^2 i \right) (F_1(e) + 4\sqrt{1-e^2}) \right] e G_{1s} \\ & + \frac{35}{16p^2} \left(\frac{J_4}{J_2} \right) \left[\sin^2 i \left(\frac{3}{7} - \frac{1}{2} \sin^2 i \right) (F_1(e) + \frac{5}{2}\sqrt{1-e^2}) - \left(\frac{3}{7} - \frac{5}{2} \sin^2 i + \frac{9}{4} \sin^4 i \right) \right] e^2 G_{2c} \end{aligned} \quad (4.368)$$

where the elements $a, e, i, n, p = a(1-e^2)$ are given by the quasi-mean elements, same as in the perturbation solution due to J_2 , which are $\bar{a}_0, \bar{e}_0, \bar{i}_0, \bar{n}_0 = \sqrt{\mu \bar{a}_0^{(-3/2)}}$, $\bar{p}_0 = \bar{a}_0(1 - \bar{e}_0^2)$. The expression of $F_1(e)$ is given in Sect. 4.2.2 (4.144) that

$$F_1(e) = \frac{1}{(1 + \sqrt{1 - e^2})}$$

and $G_{1s}, G_{1c}, G_{2s}, G_{2c}$ are

$$\begin{cases} G1s = \frac{(\cos \bar{\omega} - \cos \bar{\omega}_0)}{(2 - \frac{5}{2} \sin^2 i)} \rightarrow -\sin \bar{\omega}_0 \left(\frac{3J_2}{2p^2} \right) n(t - t_0) \\ G1c = \frac{-(\sin \bar{\omega} - \sin \bar{\omega}_0)}{(2 - \frac{5}{2} \sin^2 i)} \rightarrow -\cos \bar{\omega}_0 \left(\frac{3J_2}{2p^2} \right) n(t - t_0) \end{cases} \quad (4.369)$$

$$\begin{cases} G2s = \frac{(\cos 2\bar{\omega} - \cos 2\bar{\omega}_0)}{2(2 - \frac{5}{2} \sin^2 i)} \rightarrow -\sin 2\bar{\omega}_0 \left(\frac{3J_2}{2p^2} \right) n(t - t_0) \\ G2c = \frac{-(\sin 2\bar{\omega} - \sin 2\bar{\omega}_0)}{2(2 - \frac{5}{2} \sin^2 i)} \rightarrow -\cos 2\bar{\omega}_0 \left(\frac{3J_2}{2p^2} \right) n(t - t_0) \end{cases} \quad (4.370)$$

where $\bar{\omega}_0$ is the quasi-mean argument of perigee $\bar{\omega}_0(t_0)$ at time t_0 . The above formulas show that there are two methods of calculation, when the inclination i is near the critic inclination $i_c = 63.4^\circ$, use the right formulas, otherwise, use the left ones. By this method, the problem of the commensurable small divisor is avoided.

(2) The non-singularity perturbation solution of the second type

For a high Earth orbit satellite, the singularities caused by both a small e and a small i may occur, but for this kind of satellite the zonal harmonic terms of the second order, J_3 and J_4 are not important and do not need special treatment, therefore we do not provide the related expressions.

4.6 The Perturbation Solution Due to the High-Order Tesseral Harmonic Terms $J_{l,m}$ ($l \geq 3, M = 1, 2, \dots, l$) of Earth's Non-spherical Gravitation

4.6.1 The General Expression of the Perturbation Function of the Tesseral Harmonic Terms $J_{l,m}$ ($l \geq 3, M = 1, 2, \dots, l$)

The general perturbation function of the high-order tesseral harmonic terms is

$$R = \Delta V = \frac{\mu}{r} \sum_{l \geq 3} \sum_{m=1}^l \left(\frac{a_e}{r} \right)^l P_{l,m}(\sin \varphi) [C_{l,m} \cos m\lambda + S_{l,m} \sin m\lambda], \quad (4.371)$$

Similarly adapting the dimensionless form and using the method for $J_{2,2}$ as in Sect. 4.3.1, the harmonic coefficients, $C_{l,m}$ and $S_{l,m}$, are expressed by $J_{l,m}$ and $\lambda_{l,m}$, that

$$\begin{cases} R = \sum_{l \geq 3} \sum_{m=1}^l R_{l,m}, \\ R_{l,m} = \frac{J_{l,m}}{r^{l+1}} P_{l,m}(\mu) \cos m\bar{\lambda}, \end{cases} \quad (4.372)$$

where

$$\begin{cases} J_{l,m} = (C_{l,m}^2 + S_{l,m}^2)^{1/2}, \\ \bar{\lambda} = \lambda - \lambda_{l,m}, \\ m\lambda_{l,m} = \arctan\left(\frac{S_{l,m}}{C_{l,m}}\right). \end{cases} \quad (4.373)$$

The argument $\lambda_{l,m}$ is the geographic longitude of the “symmetric axis” direction of Earth’s equator, which is defined by the tesseral harmonic coefficients $C_{l,m}$ and $S_{l,m}$. Then $C_{l,m}$ and $S_{l,m}$ can be expressed by

$$C_{l,m} = J_{l,m} \cos m\lambda_{l,m}, \quad S_{l,m} = J_{l,m} \sin m\lambda_{l,m}. \quad (4.374)$$

4.6.2 The Perturbation Solution Due to the Tesseral Harmonic Terms $J_{l,m}$ ($l \geq 3, m = 1 - l$)

$$\sigma_s^{(2)}(t) = \sum_{l \geq 2} \sum_{m=1}^l \sum_{p=0}^l \sum_{q=-\infty}^{\infty} \Delta\sigma_{lmpq}, \quad (4.375)$$

$$\Delta\sigma_{lmpq} = \begin{cases} C_{lmpq}^{\sigma} S_{lmpq}, & \text{for } a, e, i, \\ C_{lmpq}^{\sigma} S_{lmpq}^*, & \text{for } \Omega, \omega, M, \end{cases} \quad (4.376)$$

where

$$S_{lmpq} = J_{l,m} [(1 - \delta_{l,m}) \cos \psi_{lmpq}^* + \delta_{l,m} \sin \psi_{lmpq}^*], \quad (4.377)$$

$$S_{lmpq}^* = J_{l,m} [(1 - \delta_{l,m}) \sin \psi_{lmpq}^* - \delta_{l,m} \cos \psi_{lmpq}^*], \quad (4.378)$$

$$\begin{aligned} \psi_{lmpq}^* &= \psi_{lmpq} - m\lambda_{l,m} \\ &= (l - 2p + q)M + (l - 2p)\omega + m\Omega_{l,m}, \end{aligned} \quad (4.379)$$

$$\psi_{lmpq} = (l - 2p + q)M + (l - 2p)\omega + m(\Omega - S_G), \quad (4.380)$$

$$\Omega_{l,m} = \Omega - (S_G + \lambda_{l,m}), \quad (4.381)$$

$$\delta_{l,m} = \frac{1}{2} [1 - (-1)^{l-m}] = \begin{cases} 1, & (l - m) \text{ is odd,} \\ 0, & (l - m) \text{ is even,} \end{cases} \quad (4.382)$$

where S_G is the Greenwich sidereal time, given in Chap. 1 (1.35) or (1.64).

The expressions of C_{lmpq}^{σ} for a, e, i, ω, Ω , and M are

$$C_{lmpq}^a = 2a \left(\frac{a_e}{a} \right)^l (l - 2p + q) F_{lmp}(i) G_{lpq}(e) \left(\frac{\bar{n}}{\dot{\psi}_{lmpq}} \right), \quad (4.383)$$

$$C_{lmpq}^e = \left(\frac{a_e}{a} \right)^l \frac{\sqrt{1-e^2}}{e} \left[(l - 2p + q) \sqrt{1-e^2} - (l - 2p) \right] F_{lmp}(i) G_{lpq}(e) \left(\frac{\bar{n}}{\dot{\psi}_{lmpq}} \right), \quad (4.384)$$

$$C_{lmpq}^i = \left(\frac{a_e}{a} \right)^l \frac{1}{\sqrt{1-e^2} \sin i} [(l - 2p) \cos i - m] F_{lmp}(i) G_{lpq}(e) \left(\frac{\bar{n}}{\dot{\psi}_{lmpq}} \right), \quad (4.385)$$

$$C_{lmpq}^\Omega = \left(\frac{a_e}{a} \right)^l \frac{1}{\sqrt{1-e^2} \sin i} F'_{lmp}(i) G_{lpq}(e) \left(\frac{\bar{n}}{\dot{\psi}_{lmpq}} \right), \quad (4.386)$$

$$C_{lmpq}^\omega = -\cos i C_{lmpq}^\Omega + \left(\frac{a_e}{a} \right)^l \frac{\sqrt{1-e^2}}{e} F_{lmp}(i) G'_{lpq}(e) \left(\frac{\bar{n}}{\dot{\psi}_{lmpq}} \right), \quad (4.387)$$

$$\begin{aligned} C_{lmpq}^M &= -\sqrt{1-e^2} \left(C_{lmpq}^\omega + \cos i C_{lmpq}^\Omega \right) \\ &+ \left(\frac{a_e}{a} \right)^l \left[2(l+1) - 3(l-2p+q) \left(\frac{n}{\dot{\psi}_{lmpq}} \right) \right] F_{lmp}(i) G_{lpq}(e) \left(\frac{\bar{n}}{\dot{\psi}_{lmpq}} \right), \end{aligned} \quad (4.388)$$

where $\dot{\psi}_{lmpq}$ is approximately given by

$$\begin{cases} \dot{\psi}_{lmpq} = (l - 2p + q) \dot{\bar{M}} + (l - 2p) \dot{\bar{\omega}} + m \left(\dot{\bar{\Omega}} - n_e \right) \\ \approx (l - 2p + q) \bar{n} - mn_e = \bar{n} [(l - 2p + q) - n\alpha], \\ \alpha = n_e / \bar{n} \quad \bar{n} = \sqrt{\mu a}^{-3/2}, \end{cases} \quad (4.389)$$

and n_e is Earth's rotational angular speed, i.e., the variation rate of the sidereal time. It can be approximately given by

$$n_e \approx \dot{S}_G = 360^\circ.98564745/\text{d}. \quad (4.390)$$

In the above formulas the inclination function $F_{lmp}(i)$ and Hansen coefficient $G_{lpq}(e)$, and their derivatives are given by

$$F_{lmp}(i) = \frac{(l+m)!}{2^l p!(l-p)!} \sum_{k=k_1}^{k_2} (-1)^{k+(l-m+\delta_{lm})/2} \binom{2l-2p}{k} \binom{2p}{l-m-k} \times \left(\sin \frac{i}{2}\right)^{-(l-m-2p-2k)} \left(\cos \frac{i}{2}\right)^{(3l-m-2p-2k)} \quad (4.391)$$

$$= \frac{(l+m)!}{2^{2l} p!(l-p)!} \sum_{k=k_1}^{k_2} (-1)^{k+(l-m+\delta_{lm})/2} \binom{2l-2p}{k} \binom{2p}{l-m-k} \times \left(\sin \frac{i}{2}\right)^{-(l-m-2p-2k)} (1 + \cos i)^{(2l-m-2p-2k)},$$

$$k_1 = \max(0, l-m-2p), \quad k_2 = \min(l-m, 2l-2p), \quad (4.392)$$

$$F'_{lmp}(i) = \frac{d}{di} F_{lmp}(i) = \frac{(l+m)!}{2^l p!(l-p)!} \left(\frac{1}{\sin i}\right) \sum_{k=k_1}^{k_2} (-1)^{k+(l-m+\delta_{lm})/2} \binom{2l-2p}{k} \binom{2p}{l-m-k} \times \left[-2l \sin^2 \frac{i}{2} - (l-m-2p-2k)\right] \left(\sin \frac{i}{2}\right)^{-(l-m-2p-2k)} \left(\cos \frac{i}{2}\right)^{(3l-m-2p-2k)}, \quad (4.393)$$

$$G_{lpq}(e) = X_{(l-2p)+q}^{-(l+1), (l-2p)}(e) = O(e^{|q|}). \quad (4.394)$$

For the accuracy of $O(e^2)$, we have

$$X_p^{l,p}(e) = 1 + \frac{1}{4}(l^2 + l - 4p^2)e^2, \quad (4.395)$$

$$\begin{cases} X_{p+1}^{l,p}(e) = -\frac{1}{2}(l-2p)e, \\ X_{p-1}^{l,p}(e) = -\frac{1}{2}(l+2p)e, \end{cases} \quad (4.396)$$

$$\begin{cases} X_{p+2}^{l,p}(e) = \frac{1}{8}[l^2 - (4p+3)l + p(4p+5)]e^2 \\ X_{p-2}^{l,p}(e) = \frac{1}{8}[l^2 + (4p-3)l + p(4p-5)]e^2 \end{cases} \quad (4.397)$$

$$\frac{d}{de}(X_p^{l,p}(e)) = \frac{1}{2}(l^2 + l - 4p^2)e \quad (4.398)$$

$$\begin{cases} \frac{d}{de}(X_{p+1}^{l,p}(e)) = -\frac{1}{2}(l-2p) \\ \frac{d}{de}(X_{p-1}^{l,p}(e)) = -\frac{1}{2}(l+2p) \end{cases} \quad (4.399)$$

$$\begin{cases} \frac{d}{de}(X_{p+2}^{l,p}(e)) = \frac{1}{4}[l^2 - (4p+3)l + p(4p+5)]e \\ \frac{d}{de}(X_{p-2}^{l,p}(e)) = \frac{1}{4}[l^2 + (4p-3)l + p(4p-5)]e \end{cases} \quad (4.400)$$

4.6.3 The Non-singularity Perturbation Solution of the First Type Due to the Tesseral Harmonic $J_{l,m}$ ($l \geq 3, m = 1 - l$) Terms

The short-period terms are

$$a_s^{(2)}(t) = a_s^{(2)}(t), \quad (4.401)$$

$$i_s^{(2)}(t) = i_s^{(2)}(t), \quad (4.402)$$

$$\Omega_s^{(2)}(t) = \Omega_s^{(2)}(t), \quad (4.403)$$

$$\xi_s^{(2)}(t) = \cos \omega [e_s^{(2)}(t)] - \sin \omega [e \omega_s^{(2)}(t)], \quad (4.404)$$

$$\eta_s^{(2)}(t) = \sin \omega [e_s^{(2)}(t)] + \cos \omega [e \omega_s^{(2)}(t)], \quad (4.405)$$

$$\lambda_s^{(2)}(t) = [M_s^{(2)}(t)] + [\omega_s^{(2)}(t)] = C_{lmpq}^{M+\omega} S_{lmpq}^*, \quad (4.406)$$

where

$$\begin{aligned} C_{lmpq}^{M+\omega} &= C_{lmpq}^M + C_{lmpq}^\omega \\ &= -\cos i C_{lmpq}^\Omega \frac{e\sqrt{1-e^2}}{1+\sqrt{1-e^2}} \left[\left(\frac{a_e}{a} \right)^l F_{lmp}(i) G'_{lpq}(e) \left(\frac{\bar{n}}{\dot{\psi}_{lmpq}} \right) \right] \\ &\quad + \left(\frac{a_e}{a} \right)^l \left[2(l+1) - 3(l-2p+q) \left(\frac{n}{\dot{\psi}_{lmpq}} \right) \right] F_{lmp}(i) G_{lpq}(e) \left(\frac{\bar{n}}{\dot{\psi}_{lmpq}} \right) \end{aligned} \quad (4.407)$$

The term S_{lmpq}^* is given by (4.378).

4.6.4 The Non-Singularity Perturbation Solution of the Second Type Due to the Tesseral Harmonic $J_{l,m}$ ($l \geq 3, m = 1 - l$) Terms

The short-period terms are

$$a_s^{(2)}(t) = a_s^{(2)}(t), \quad (4.408)$$

$$\xi_s^{(2)}(t) = \cos \tilde{\omega} [e_s^{(2)}(t)] - \sin \tilde{\omega} [e\omega_s^{(2)}(t) + e\Omega_s^{(2)}(t)], \quad (4.409)$$

$$\eta_s^{(2)}(t) = \sin \tilde{\omega} [e_s^{(2)}(t)] + \cos \tilde{\omega} [e\omega_s^{(2)}(t) + e\Omega_s^{(2)}(t)], \quad (4.410)$$

$$h_s^{(2)}(t) = \frac{1}{2} \cos \frac{i}{2} \cos \Omega [i_s^{(2)}(t)] - \sin \Omega \left[\sin \frac{i}{2} \Omega_s^{(2)}(t) \right], \quad (4.411)$$

$$k_s^{(2)}(t) = \frac{1}{2} \cos \frac{i}{2} \sin \Omega [i_s^{(2)}(t)] + \cos \Omega \left[\sin \frac{i}{2} \Omega_s^{(2)}(t) \right], \quad (4.412)$$

$$\lambda_s^{(2)}(t) = [M_s^{(2)}(t)] + [\omega_s^{(2)}(t)] + [\Omega_s^{(2)}(t)] = C_{lmpq}^{M+\omega+\Omega} S_{lmpq}^*, \quad (4.413)$$

where

$$\begin{aligned} C_{lmpq}^{M+\omega} &= C_{lmpq}^M + C_{lmpq}^\omega + C_{lmpq}^\Omega \\ &= \frac{\sin^2 i}{1 + \cos i} C_{lmpq}^\Omega + \frac{e\sqrt{1-e^2}}{1 + \sqrt{1-e^2}} \left[\left(\frac{a_e}{a} \right)^l F_{lmp}(i) G'_{lpq}(e) \left(\frac{\bar{n}}{\dot{\psi}_{lmpq}} \right) \right] \\ &\quad + \left(\frac{a_e}{a} \right)^l \left[2(l+1) - 3(l-2p+q) \left(\frac{n}{\dot{\psi}_{lmpq}} \right) \right] F_{lmp}(i) G_{lpq}(e) \left(\frac{\bar{n}}{\dot{\psi}_{lmpq}} \right) \end{aligned} \quad (4.414)$$

The term S_{lmpq}^* is given by (4.378).

4.6.5 The Perturbation Solution Due to the Tesseral Terms, $J_{3,m}$ ($M = 1, 2, 3$) and $J_{4,m}$ ($M = 1, 2, 3, 4$) in Kepler Elements

As mentioned in Sect. 4.5.5 for a low Earth orbit satellite, the main perturbation in Earth's gravity potential is Earth's dynamical form-factor J_2 term, and non-spherical terms J_3 and J_4 . The characteristics of the influence of the three terms can be further confirmed if we consider the entire Earth's non-spherical potential up to the fourth order. In this section we provide the perturbation solution due to $J_{3,m}$ ($m = 1, 2, 3$) and $J_{4,m}$ ($m = 1, 2, 3, 4$).

Similar to the case of $J_{2,2}$, these terms involve Earth's rotational angular speed n_e , i.e., the variation rate of the sidereal time, which is given in (4.390):

$$n_e \approx \dot{S}_G = 360^\circ.98564745/\text{d}.$$

The variables are all dimensionless in normalized units. For convenience the two tesseral harmonic coefficients, $C_{l,m}$ and $S_{l,m}$ are given by the following forms:

$$C_{l,m} = J_{l,m} \cos m\lambda_{l,m}, \quad S_{l,m} = J_{l,m} \sin m\lambda_{l,m}, \quad (4.415)$$

$$J_{l,m} = (C_{l,m}^2 + S_{l,m}^2)^{1/2}, \quad (4.416)$$

$$\begin{cases} m\lambda_{l,m} = \arctan(S_{l,m}/C_{l,m}), \\ \Omega_{l,m} = \Omega - (\bar{S}_G + \lambda_{l,m}), \end{cases} \quad (4.417)$$

for $l = 3$ and 4 , there are

$$C_{3,1}, S_{3,1}, C_{3,2}, S_{3,2}, C_{3,3}, S_{3,3}, \\ C_{4,1}, S_{4,1}, C_{4,2}, S_{4,2}, C_{4,3}, S_{4,3}, C_{4,4}, S_{4,4},$$

In practice, $C_{l,m}$ and $S_{l,m}$ are usually expressed by $J_{l,m}$ and $m\lambda_{l,m}$, which are

$$J_{3,1}, J_{3,2}, J_{3,3}, \quad J_{4,1}, J_{4,2}, J_{4,3}, J_{4,4}$$

$$\lambda_{3,1}, \lambda_{3,2}, \lambda_{3,3}, \quad \lambda_{4,1}, \lambda_{4,2}, \lambda_{4,3}, \lambda_{4,4}$$

Except for the term $a_s^{(2)}(t, J_{l,m})$, the other perturbation solutions need only to keep the terms related to Earth's rotation, the rotation factor α is

$$\alpha = n_e/\bar{n}, \quad \bar{n} = \sqrt{\mu}a^{-3/2} = \bar{a}^{-3/2}, \quad (4.418)$$

(1) $J_{3,1}$

$$\begin{aligned} a_s^{(2)}(t) = & \frac{2(J_{3,1})}{a^2} \left\{ F_{310}(i) [-e \cos(2M + 3\omega + \Omega_{3,1}) \right. \\ & + \frac{1}{1 - \alpha/3} \cos(3M + 3\omega + \Omega_{3,1}) + 5e \cos(4M + 3\omega + \Omega_{3,1})] \\ & + F_{311}(i) \left[\frac{1}{(1 - \alpha)} \cos(M + \omega + \Omega_{3,1}) + 3e \cos(2M + \omega + \Omega_{3,1}) \right] \\ & + F_{312}(i) \left[3e \cos(2M + \omega - \Omega_{3,1}) + \frac{1}{1 + \alpha} \cos(M + \omega - \Omega_{3,1}) \right] \\ & + F_{313}(i) [-e \cos(2M + 3\omega - \Omega_{3,1}) \\ & + \frac{1}{1 + \alpha/3} \cos(3M + 3\omega - \Omega_{3,1}) + 5e \cos(4M + 3\omega - \Omega_{3,1})] \left. \right\}, \end{aligned} \quad (4.419)$$

$$e_s^{(2)}(t) = \frac{(J_{3,1})}{a^3} \frac{1}{\alpha} \left\{ F_{311}(i) [\cos(\omega + \Omega_{3,1})] + F_{312}(i) [-\cos(\omega - \Omega_{3,1})] \right\}, \quad (4.420)$$

$$i_s^{(2)}(t) = 0, \quad (4.421)$$

$$\Omega_s^{(2)}(t) = 0, \quad (4.422)$$

$$\omega_s^{(2)}(t) = \frac{(J_{3,1})}{a^3} \frac{1}{\alpha e} \{ F_{311}(i) [-\sin(\omega + \Omega_{3,1})] + F_{312}(i) [\sin(\omega - \Omega_{3,1})] \}, \quad (4.423)$$

$$M_s^{(2)}(t) + \omega_s^{(2)}(t) = 0. \quad (4.424)$$

(2) $J_{3,2}$

$$\begin{aligned} a_s^{(2)}(t) = & \frac{2(J_{3,2})}{a^2} \{ F_{320}(i) [-e \sin(2M + 3\omega + 2\Omega_{3,2}) \\ & + \frac{1}{1 - 2\alpha/3} \sin(3M + 3\omega + 2\Omega_{3,2}) + 5e \sin(4M + 3\omega + 2\Omega_{3,2})] \\ & + F_{321}(i) \left[\frac{1}{(1 - 2\alpha)} \sin(M + \omega + 2\Omega_{3,2}) + 3e \sin(2M + \omega + 2\Omega_{3,2}) \right] \\ & + F_{322}(i) \left[-\frac{1}{(1 + 2\alpha)} \sin(M + \omega - 2\Omega_{3,2}) - 3e \sin(2M + \omega - 2\Omega_{3,2}) \right] \\ & + F_{323}(i) [e \sin(2M + 3\omega - 2\Omega_{3,2}) \\ & - \frac{1}{1 + 2\alpha/3} \sin(3M + 3\omega - 2\Omega_{3,2}) - 5e \sin(4M + 3\omega - 2\Omega_{3,2})] \}, \end{aligned} \quad (4.425)$$

$$e_s^{(2)}(t) = \frac{(J_{3,2})}{a^3} \frac{1}{2\alpha} \{ F_{321}(i) [\sin(\omega + 2\Omega_{3,2})] + F_{322}(i) [\sin(\omega - 2\Omega_{3,2})] \}, \quad (4.426)$$

$$i_s^{(2)}(t) = 0, \quad (4.427)$$

$$\Omega_s^{(2)}(t) = 0, \quad (4.428)$$

$$\omega_s^{(2)}(t) = \frac{(J_{3,2})}{a^3} \frac{1}{2\alpha e} \{ F_{321}(i) [\cos(\omega + 2\Omega_{3,2})] + F_{322}(i) [\cos(\omega - 2\Omega_{3,2})] \}, \quad (4.429)$$

$$M_s^{(2)}(t) + \omega_s^{(2)}(t) = 0. \quad (4.430)$$

(3) $J_{3,3}$

$$\begin{aligned}
a_s^{(2)}(t) = & \frac{2(J_{3,3})}{a^2} \left\{ F_{330}(i) [-e \cos(2M + 3\omega + 3\Omega_{3,3}) \right. \\
& + \frac{1}{(1-\alpha)} \cos(3M + 3\omega + 3\Omega_{3,3}) + 5e \cos(4M + 3\omega + 3\Omega_{3,3})] \\
& + F_{331}(i) \left[\frac{1}{(1-3\alpha)} \cos(M + \omega + 3\Omega_{3,3}) + 3e \cos(2M + \omega + 3\Omega_{3,3}) \right] \\
& + F_{332}(i) \left[-\frac{1}{1+3\alpha} \cos(M + \omega - 3\Omega_{3,3}) + 3e \cos(2M + \omega - 3\Omega_{3,3}) \right] \\
& + F_{333}(i) [-e \cos(2M + 3\omega - 3\Omega_{3,3}) \\
& \left. + \frac{1}{(1+\alpha)} \cos(3M + 3\omega - 3\Omega_{3,3}) + 5e \cos(4M + 3\omega - 3\Omega_{3,3}) \right] \Big\}, \tag{4.431}
\end{aligned}$$

$$e_s^{(2)}(t) = \frac{(J_{3,3})}{a^3} \frac{1}{3\alpha} \{ F_{331}(i) [\cos(\omega + 3\Omega_{3,3})] - F_{332}(i) [\cos(\omega - 3\Omega_{3,3})] \}, \tag{4.432}$$

$$i_s^{(2)}(t) = 0, \tag{4.433}$$

$$\Omega_s^{(2)}(t) = 0, \tag{4.434}$$

$$\omega_s^{(2)}(t) = \frac{(J_{3,3})}{a^3} \frac{1}{3\alpha e} \{ F_{331}(i) [-\sin(\omega + 3\Omega_{3,3})] + F_{332}(i) [\sin(\omega - 3\Omega_{3,3})] \}, \tag{4.435}$$

$$M_s^{(2)}(t) + \omega_s^{(2)}(t) = 0. \tag{4.436}$$

(4) $J_{4,1}$

$$\begin{aligned}
a_s^{(2)}(t) = & \frac{(J_{4,1})}{a^3} \left\{ F_{410}(i) [-3e \sin(3M + 4\omega + \Omega_{4,1}) \right. \\
& + \frac{2}{(1-\alpha/4)} \sin(4M + 4\omega + \Omega_{4,1}) + 13e \sin(5M + 4\omega + \Omega_{4,1})] \\
& + F_{411}(i) [e \sin(M + 2\omega + \Omega_{4,1}) \\
& + \frac{2}{1-\alpha/2} \sin(2M + 2\omega + \Omega_{4,1}) + 9e \sin(3M + 2\omega + \Omega_{4,1})] \\
& + F_{412}(i) [-5e \sin(M - \Omega_{4,1}) - 5e \sin(M + \Omega_{4,1})] \\
& \left. + F_{413}(i) [-e \sin(M + 2\omega - \Omega_{4,1})] \right\}
\end{aligned}$$

$$\begin{aligned}
& - \frac{2}{(1 + \alpha/2)} \sin(2M + 2\omega - \Omega_{4,1}) - 9e \sin(3M + 2\omega - \Omega_{4,1}) \Big] \\
& \quad + F_{414}(i) [3e \sin(3M + 4\omega - \Omega_{4,1}) \\
& - \frac{2}{(1 + \alpha/4)} \sin(4M + 4\omega - \Omega_{4,1}) - 13e \sin(5M + 4\omega - \Omega_{4,1}) \Big] \Big\} \\
& \hspace{25em} (4.437)
\end{aligned}$$

$$e_s^{(2)}(t) = 0. \quad (4.438)$$

$$i_s^{(2)}(t) = \frac{(J_{4,1})}{a^4} \frac{1}{\alpha} \{ F_{412}^*(i) [-\sin(\Omega_{4,1})] \}. \quad (4.439)$$

$$\Omega_s^{(2)}(t) = \frac{(J_{4,1})}{a^4} \frac{1}{\alpha \sin i} \{ F_{412}^{**}(i) [\cos(\Omega_{4,1})] \}, \quad (4.440)$$

$$\omega_s^{(2)}(t) = 0, \quad (4.441)$$

$$\begin{aligned}
M_s^{(2)}(t) + \omega_s^{(2)}(t) &= -\cos i \Omega_s^{(2)}(t) \\
&+ \frac{(J_{4,1})}{a^4} \frac{1}{\alpha} \{ F_{412}(i) [10 \cos \Omega_{4,1}] \} \\
& \hspace{25em} (4.442)
\end{aligned}$$

(5) $J_{4,2}$

$$\begin{aligned}
a_s^{(2)}(t) &= \frac{(J_{4,2})}{a^3} \{ F_{420}(i) [-3e \cos(3M + 4\omega + 2\Omega_{4,2}) \\
&+ \frac{2}{(1 - \alpha/2)} \cos(4M + 4\omega + 2\Omega_{4,2}) + 13e \cos(5M + 4\omega + 2\Omega_{4,2})] \\
&+ F_{421}(i) [e \cos(M + 2\omega + 2\Omega_{4,2}) \\
&+ \frac{2}{(1 - \alpha)} \cos(2M + 2\omega + 2\Omega_{4,2}) + 9e \cos(3M + 2\omega + 2\Omega_{4,2})] \\
&+ F_{422}(i) [5e \cos(M - 2\Omega_{4,2}) + 5e \cos(M + 2\Omega_{4,2})] \\
&+ F_{423}(i) [e \cos(M + 2\omega - 2\Omega_{4,2}) \\
&+ \frac{2}{(1 + \alpha)} \cos(2M + 2\omega - 2\Omega_{4,2}) + 9e \cos(3M + 2\omega - 2\Omega_{4,2})] \\
&+ F_{424}(i) [-3e \cos(3M + 4\omega - 2\Omega_{4,2}) \\
&+ \frac{2}{(1 + \alpha/2)} \cos(4M + 4\omega - 2\Omega_{4,2}) + 13e \cos(5M + 4\omega - 2\Omega_{4,2})] \} \\
& \hspace{25em} (4.443)
\end{aligned}$$

$$e_s^{(2)}(t) = 0, \quad (4.444)$$

$$i_s^{(2)}(t) = \frac{(J_{4,2})}{a^4} \frac{1}{2\alpha} \sin i \{F_{422}^*(i)[- \cos(2\Omega_{4,2})]\} \quad (4.445)$$

$$\Omega_s^{(2)}(t) = \frac{(J_{4,2})}{a^4} \frac{1}{2\alpha} \{F_{422}^{**}(i)[- \sin(2\Omega_{4,2})]\}, \quad (4.446)$$

$$\omega_s^{(2)}(t) = 0, \quad (4.447)$$

$$\begin{aligned} M_s^{(2)}(t) + \omega_s^{(2)}(t) = & -\cos i \Omega_s^{(2)}(t) \\ & + \frac{(J_{4,2})}{a^4} \frac{1}{2\alpha} \{F_{422}(i)[-10 \sin(2\Omega_{4,2})]\} \end{aligned} \quad (4.448)$$

(6) $J_{4,3}$

$$\begin{aligned} a_s^{(2)}(t) = & \frac{(J_{4,3})}{a^3} \{F_{430}(i)[-3e \sin(3M + 4\omega + 3\Omega_{4,3}) \\ & + \frac{2}{(1 - 3\alpha/4)} \sin(4M + 4\omega + 3\Omega_{4,3}) + 13e \sin(5M + 4\omega + 3\Omega_{4,3})] \\ & + F_{431}(i)[e \sin(M + 2\omega + 3\Omega_{4,3}) \\ & + \frac{2}{1 - 3\alpha/2} \sin(2M + 2\omega + 3\Omega_{4,3}) + 9e \sin(3M + 2\omega + 3\Omega_{4,3})] \\ & + F_{432}(i)[-5e \sin(M - 3\Omega_{4,3}) + 5e \sin(M + 3\Omega_{4,3})] \\ & + F_{433}(i)[-e \sin(M + 2\omega - 3\Omega_{4,3}) \\ & - \frac{2}{(1 + 3\alpha/2)} \sin(2M + 2\omega - 3\Omega_{4,3}) - 9e \sin(3M + 2\omega - 3\Omega_{4,3})] \\ & + F_{434}(i)[3e \sin(3M + 4\omega - 3\Omega_{4,3}) \\ & - \frac{2}{(1 + 3\alpha/4)} \sin(4M + 4\omega - 3\Omega_{4,3}) - 13e \sin(5M + 4\omega - 3\Omega_{4,3})]\} \end{aligned} \quad (4.449)$$

$$e_s^{(2)}(t) = 0, \quad (4.450)$$

$$i_s^{(2)}(t) = \frac{(J_{4,3})}{a^4} \frac{1}{3\alpha} \{F_{432}^*(i)[- \sin(3\Omega_{4,3})]\}, \quad (4.451)$$

$$\Omega_s^{(2)}(t) = \frac{(J_{4,3})}{a^4} \frac{1}{3\alpha \sin i} \{F_{432}^{**}(i) [\cos(3\Omega_{4,3})]\}, \quad (4.452)$$

$$\omega_s^{(2)}(t) = 0, \quad (4.453)$$

$$\begin{aligned} M_s^{(2)}(t) + \omega_s^{(2)}(t) = & -\cos i \Omega_s^{(2)}(t) \\ & + \frac{(J_{4,3})}{a^4} \frac{1}{3\alpha} \{F_{433}(i) [10 \cos(3\Omega_{4,3})]\} \end{aligned} \quad (4.454)$$

(7) $J_{4,4}$

$$\begin{aligned} a_s^{(2)}(t) = & \frac{(J_{4,4})}{a^3} \{F_{440}(i) [-3e \cos(3M + 4\omega + 4\Omega_{4,4}) \\ & + \frac{2}{(1-\alpha)} \cos(4M + 4\omega + 4\Omega_{4,4}) + 13e \cos(5M + 4\omega + 4\Omega_{4,4})] \\ & + F_{441}(i) \left[e \cos(M + 2\omega + 4\Omega_{4,4}) + \frac{2}{(1-2\alpha)} \cos(2M + 2\omega + 4\Omega_{4,4}) \right. \\ & \quad \left. + 9e \cos(3M + 2\omega + 4\Omega_{4,4}) \right] \\ & + F_{442}(i) [5e \cos(M - 4\Omega_{4,4}) + 5e \cos(M + 4\Omega_{4,4})] \\ & + F_{443}(i) \left[e \cos(M + 2\omega - 4\Omega_{4,4}) + \frac{2}{(1+2\alpha)} \cos(2M + 2\omega - 4\Omega_{4,4}) \right. \\ & \quad \left. + 9e \cos(3M + 2\omega - 4\Omega_{4,4}) \right] \\ & + F_{444}(i) [-3e \cos(3M + 4\omega - 4\Omega_{4,4}) \\ & \quad \left. + \frac{2}{(1+\alpha)} \cos(4M + 4\omega - 4\Omega_{4,4}) + 13e \cos(5M + 4\omega - 4\Omega_{4,4}) \right]\} \end{aligned} \quad (4.455)$$

$$e_s^{(2)}(t) = 0, \quad (4.456)$$

$$i_s^{(2)}(t) = \frac{(J_{4,4})}{a^4} \frac{1}{4\alpha} \sin i \{F_{442}^*(i) [-\cos(4\Omega_{4,4})]\}, \quad (4.457)$$

$$\Omega_s^{(2)}(t) = \frac{(J_{4,4})}{a^4} \frac{1}{4\alpha} \{F_{442}^{**}(i) [-\sin(4\Omega_{4,4})]\}, \quad (4.458)$$

$$\omega_s^{(2)}(t) = 0, \quad (4.459)$$

$$M_s^{(2)}(t) + \omega_s^{(2)}(t) = -\cos i \Omega_s^{(2)}(t)$$

$$+ \frac{(J_{4,4})}{a^4} \frac{1}{4\alpha} \{F_{442}(i)[-10 \sin(4\Omega_{4,4})]\} \quad (4.460)$$

In the above expressions, the terms having the factor $\frac{1}{m\alpha}$ are called the Earth rotation terms. For low Earth orbit satellites, because the value of α is relatively small, the Earth rotation terms are important and should not be simply treated as short-period terms of the second order.

The inclination functions $F_{lmp}(i)$ for all possible values of l, m , and p are given by

$$\begin{cases} F_{310} = -\frac{15}{16} \sin^2 i (1 + \cos i) \\ F_{311} = +\frac{15}{16} \sin^2 i (1 + 3 \cos i) - \frac{3}{4} (1 + \cos i) \\ F_{312} = +\frac{15}{16} \sin^2 i (1 - 3 \cos i) - \frac{3}{4} (1 - \cos i) \\ F_{313} = -\frac{15}{16} \sin^2 i (1 - \cos i) \end{cases} \quad (4.461)$$

$$\begin{cases} F_{310}^* = +\frac{15}{16} (1 + \cos i)(1 - 3 \cos i) \\ F_{311}^* = -\frac{15}{16} (1 - \cos i)(1 + 3 \cos i) + \frac{3}{4} \\ F_{312}^* = -\frac{15}{16} (1 + \cos i)(1 - 3 \cos i) + \frac{3}{4} \\ F_{313}^* = +\frac{15}{16} (1 - \cos i)(1 + 3 \cos i) \end{cases} \quad (4.462)$$

$$\begin{cases} F_{310}^{**} = +\frac{15}{16} (1 + \cos i)(1 - 3 \cos i) \\ F_{311}^{**} = -\frac{3}{16} (11 - 10 \cos i - 45 \cos^2 i) \\ F_{312}^{**} = +\frac{3}{16} (11 + 10 \cos i - 45 \cos^2 i) \\ F_{313}^{**} = -\frac{15}{16} (1 - \cos i)(1 + 3 \cos i) \end{cases} \quad (4.463)$$

$$\begin{cases} F_{320} = +\frac{15}{8} \sin i (1 + \cos i)^2 \\ F_{321} = +\frac{15}{8} \sin i (1 + \cos i)(1 - 3 \cos i) \\ F_{322} = -\frac{15}{8} \sin i (1 - \cos i)(1 + 3 \cos i) \\ F_{323} = -\frac{15}{8} \sin i (1 - \cos i)^2 \end{cases} \quad (4.464)$$

$$\begin{cases} F_{320}^* = -\frac{15}{8} (1 + \cos i)^2 (2 - 3 \cos i) \\ F_{321}^* = -\frac{15}{8} (1 + \cos i)(1 - 3 \cos i)(2 - \cos i) \\ F_{322}^* = +\frac{15}{8} (1 - \cos i)(1 + 3 \cos i)(2 + \cos i) \\ F_{323}^* = +\frac{15}{8} (1 - \cos i)^2 (2 + 3 \cos i) \end{cases} \quad (4.465)$$

$$\begin{cases} F_{320}^{**} = F_{320}' = -\frac{15}{8} (1 + \cos i)^2 (2 - 3 \cos i) \\ F_{321}^{**} = F_{321}' = +\frac{15}{8} (1 + \cos i)(2 + 5 \cos i - 9 \cos^2 i) \\ F_{322}^{**} = F_{322}' = +\frac{15}{8} (1 - \cos i)(2 - 5 \cos i - 9 \cos^2 i) \\ F_{323}^{**} = F_{323}' = -\frac{15}{8} (1 - \cos i)^2 (2 + 3 \cos i) \end{cases} \quad (4.466)$$

$$\begin{cases} F_{330} = \frac{15}{8} (1 + \cos i)^3 \\ F_{331} = \frac{45}{8} \sin^2 i (1 + \cos i) \\ F_{332} = \frac{45}{8} \sin^2 i (1 - \cos i) \\ F_{333} = \frac{15}{8} (1 - \cos i)^3 \end{cases} \quad (4.467)$$

$$\begin{cases} F_{330}^* = -\frac{45}{8}(1 + \cos i)^2 \\ F_{331}^* = -\frac{45}{8}(1 + \cos i)(3 - \cos i) \\ F_{332}^* = -\frac{45}{8}(1 - \cos i)(3 + \cos i) \\ F_{333}^* = -\frac{45}{8}(1 - \cos i)^2 \end{cases} \quad (4.468)$$

$$\begin{cases} F_{330}^{**} = -\frac{45}{8}(1 + \cos i)^2 \\ F_{331}^{**} = -\frac{45}{8}(1 + \cos i)(1 - 3 \cos i) \\ F_{332}^{**} = +\frac{45}{8}(1 - \cos i)(1 + 3 \cos i) \\ F_{333}^{**} = +\frac{45}{8}(1 - \cos i)^2 \end{cases} \quad (4.469)$$

$$\begin{cases} F_{410} = -\frac{35}{32} \sin^3 i (1 + \cos i) \\ F_{411} = +\frac{35}{16} \sin^3 i (1 + 2 \cos i) - \frac{15}{8} \sin i (1 + \cos i) \\ F_{412} = -\frac{105}{16} \sin^3 i \cos i + \frac{15}{4} \sin i \cos i \\ F_{413} = -\frac{35}{16} \sin^3 i (1 - 2 \cos i) + \frac{15}{8} \sin i (1 - \cos i) \\ F_{414} = +\frac{35}{32} \sin^3 i (1 - \cos i) \end{cases} \quad (4.470)$$

$$\begin{cases} F_{410}^* = +\frac{35}{32} \sin^2 i (1 + \cos i)(1 - 4 \cos i) \\ F_{411}^* = \left[-\frac{35}{16} \sin^2 i (1 + 2 \cos i) + \frac{15}{8} (1 + \cos i) \right] (1 - 2 \cos i) \\ F_{412}^* = \left[\frac{105}{16} \sin^2 i - \frac{15}{4} \right] \cos i \\ F_{413}^* = \left[\frac{35}{16} \sin^2 i (1 - 2 \cos i) - \frac{15}{8} (1 - \cos i) \right] (1 + 2 \cos i) \\ F_{414}^* = -\frac{35}{32} \sin^2 i (1 - \cos i)(1 + 4 \cos i) \end{cases} \quad (4.471)$$

$$\begin{cases} F_{410}^{**} = F_{410}' = +\frac{35}{32} \sin^2 i (1 + \cos i)(1 - 4 \cos i) \\ F_{411}^{**} = F_{411}' = -\frac{35}{16} \sin^2 i (2 - 3 \cos i - 8 \cos^2 i) + \frac{15}{8} (1 + \cos i)(1 - 2 \cos i) \\ F_{412}^{**} = F_{412}' = \frac{15}{4} - \frac{535}{16} \sin^2 i + \frac{65}{2} \sin^4 i \\ F_{413}^{**} = F_{413}' = -\frac{35}{16} \sin^2 i (2 + 3 \cos i - 8 \cos^2 i) + \frac{15}{8} (1 - \cos i)(1 + 2 \cos i) \\ F_{414}^{**} = F_{414}' = +\frac{35}{32} \sin^2 i (1 - \cos i)(1 + 4 \cos i) \end{cases} \quad (4.472)$$

$$\begin{cases} F_{420} = -\frac{105}{32} \sin^2 i (1 + \cos i)^2 \\ F_{421} = +\frac{105}{8} \sin^2 i \cos i (1 + \cos i) - \frac{15}{8} (1 + \cos i)^2 \\ F_{422} = +\frac{45}{16} \sin^2 i (1 - 7 \cos^2 i) \\ F_{423} = -\frac{105}{8} \sin^2 i \cos i (1 - \cos i) - \frac{15}{8} (1 - \cos i)^2 \\ F_{424} = -\frac{105}{32} \sin^2 i (1 - \cos i)^2 \end{cases} \quad (4.473)$$

$$\begin{cases} F_{420}^* = +\frac{105}{16} (1 + \cos i)^2 (1 - 2 \cos i) \\ F_{421}^* = -\frac{105}{4} \sin^2 i \cos i + \frac{15}{4} (1 + \cos i) \\ F_{422}^* = -\frac{45}{8} (1 - 7 \cos^2 i) \\ F_{423}^* = +\frac{105}{4} \sin^2 i \cos i + \frac{15}{4} (1 - \cos i) \\ F_{424}^* = +\frac{105}{16} (1 - \cos i)^2 (1 + 2 \cos i) \end{cases} \quad (4.474)$$

$$\left\{ \begin{array}{l} F_{420}^{**} = F'_{420}/\sin i = +\frac{105}{16}(1 + \cos i)^2(1 - 2\cos i) \\ F_{421}^{**} = F'_{421}/\sin i = -\frac{15}{8}(1 + \cos i)(5 + 7\cos i - 28\cos^2 i) \\ F_{422}^{**} = F'_{422}/\sin i = +\frac{45}{4}\cos i(4 - 7\cos^2 i) \\ F_{423}^{**} = F'_{423}/\sin i = +\frac{15}{8}(1 - \cos i)(5 - 7\cos i - 28\cos^2 i) \\ F_{424}^{**} = F'_{424}/\sin i = -\frac{105}{16}(1 - \cos i)^2(1 + 2\cos i) \end{array} \right. \quad (4.475)$$

$$\left\{ \begin{array}{l} F_{430} = +\frac{105}{16}\sin i(1 + \cos i)^3 \\ F_{431} = +\frac{105}{8}\sin i(1 + \cos i)^2(1 - 2\cos i) \\ F_{432} = -\frac{315}{8}\sin^3 i \cos i \\ F_{433} = -\frac{105}{8}\sin i(1 - \cos i)^2(1 + 2\cos i) \\ F_{434} = -\frac{105}{16}\sin i(1 - \cos i)^3 \end{array} \right. \quad (4.476)$$

$$\left\{ \begin{array}{l} F_{430}^* = -\frac{105}{16}(1 + \cos i)^3(3 - 4\cos i) \\ F_{431}^* = -\frac{105}{8}(1 + \cos i)^2(1 - 2\cos i)(3 - 2\cos i) \\ F_{432}^* = +\frac{945}{8}\sin^2 i \cos i \\ F_{433}^* = \frac{105}{8}(1 - \cos i)^2(1 + 2\cos i)(3 + 2\cos i) \\ F_{434}^* = \frac{105}{16}(1 - \cos i)^3(3 + 4\cos i) \end{array} \right. \quad (4.477)$$

$$\left\{ \begin{array}{l} F_{430}^{**} = F'_{430} = -\frac{105}{16}(1 + \cos i)^3(3 - 4\cos i) \\ F_{431}^{**} = F'_{431} = +\frac{105}{8}\cos i(1 + \cos i)^2(7 - 8\cos i) \\ F_{432}^{**} = F'_{432} = -\frac{315}{8}\sin^2 i(3 - 4\sin^2 i) \\ F_{433}^{**} = F'_{433} = -\frac{105}{8}\cos i(1 - \cos i)^2(7 + 8\cos i) \\ F_{434}^{**} = F'_{434} = -\frac{105}{16}(1 - \cos i)^3(3 + 4\cos i) \end{array} \right. \quad (4.478)$$

$$\left\{ \begin{array}{l} F_{440} = \frac{105}{16}(1 + \cos i)^4 \\ F_{441} = \frac{105}{4}\sin^2 i(1 + \cos i)^2 \\ F_{442} = \frac{315}{8}\sin^4 i \\ F_{443} = \frac{105}{4}\sin^2 i(1 - \cos i)^2 \\ F_{444} = \frac{105}{16}(1 - \cos i)^4 \end{array} \right. \quad (4.479)$$

$$\left\{ \begin{array}{l} F_{440}^* = -\frac{105}{4}(1 + \cos i)^3 \\ F_{441}^* = -\frac{105}{2}(1 + \cos i)^2(2 - \cos i) \\ F_{442}^* = -\frac{315}{2}\sin^2 i \\ F_{443}^* = -\frac{105}{2}(1 - \cos i)^2(2 + \cos i) \\ F_{444}^* = -\frac{105}{4}(1 - \cos i)^3 \end{array} \right. \quad (4.480)$$

$$\left\{ \begin{array}{l} F_{440}^{**} = F'_{440}/\sin i = -\frac{105}{4}(1 + \cos i)^3 \\ F_{441}^{**} = F'_{441}/\sin i = -\frac{105}{2}(1 + \cos i)^2(1 - 2\cos i) \\ F_{442}^{**} = F'_{442}/\sin i = +\frac{315}{2}\sin^2 i \cos i \\ F_{443}^{**} = F'_{443}/\sin i = +\frac{105}{2}(1 - \cos i)^2(1 + 2\cos i) \\ F_{444}^{**} = F'_{444}/\sin i = +\frac{105}{4}(1 - \cos i)^3 \end{array} \right. \quad (4.481)$$

4.6.6 *The Non-singularity Perturbation Solution of the First Type Due to the Tesseral Harmonic Terms $J_{3,m}$ ($M = 1, 2, 3$) and $J_{4,m}$ ($M = 1, 2, 3, 4$)*

The perturbation solution by $J_{3,m}$ ($m = 1, 2, 3$) and $J_{4,m}$ ($m = 1, 2, 3, 4$) has only short-period terms which are

$$a_s^{(2)}(t) = a_s^{(2)}(t), \quad (4.482)$$

$$i_s^{(2)}(t) = i_s^{(2)}(t), \quad (4.483)$$

$$\Omega_s^{(2)}(t) = \Omega_s^{(2)}(t), \quad (4.484)$$

$$\xi_s^{(2)}(t) = \cos \omega [e_s^{(2)}(t)] - \sin \omega [e_s^{(2)}(t)], \quad (4.485)$$

$$\eta_s^{(2)}(t) = \sin \omega [e_s^{(2)}(t)] + \cos \omega [e_s^{(2)}(t)], \quad (4.486)$$

$$\lambda_s^{(2)}(t) = [M_s^{(2)}(t)] + [\omega_s^{(2)}(t)]. \quad (4.487)$$

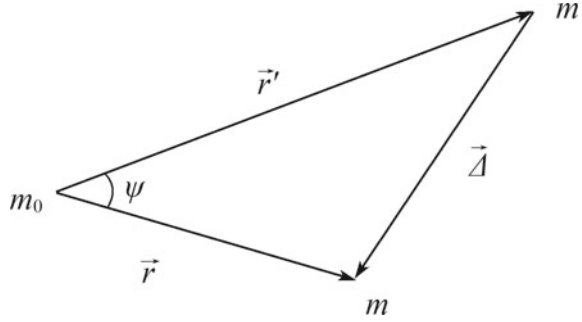
4.6.7 *The Non-singularity Perturbation Solution of the Second Type Due to the Tesseral Harmonic Terms $J_{3,m}$ ($m = 1, 2, 3$) and $J_{4,m}$ ($m = 1, 2, 3, 4$)*

For a high Earth orbit satellite, the situation of having both a small e and a small i can happen, such as GEO satellites. For these satellites the $J_{2,2}$ term is important, but the $J_{3,m}$ ($m = 1, 2, 3$) and $J_{4,m}$ ($m = 1, 2, 3, 4$) terms can be ignored, so the solutions are not provided here.

4.7 The Perturbed Orbit Solution Due to the Gravitational Force of the Sun or the Moon

Usually, this kind of perturbation is called the third-body perturbation. Earth is the central body denoted by m_0 , the moving body is by m , and the third body, the Sun or the Moon, is by m' . The relative positions of the three bodies are shown in Fig. 4.2.

Fig. 4.2 Relative positions of the three bodies



For an Earth's satellite, the property of a third-body (the Sun or the Moon) perturbation is external that the distance between the third body and Earth's center, r' , is much larger than that between the satellite and Earth's center, r , i.e., $(\frac{r}{r'}) < 1$.

4.7.1 The Perturbation Function and Its Decomposition

The perturbing acceleration $\vec{F}_\varepsilon$ due to the gravitational force of the Sun or the Moon is given by the gradient of the perturbing function R , that

$$\vec{F}_\varepsilon = \text{grad}(R) = -m' \left(\frac{\vec{\Delta}}{\Delta^3} + \frac{\vec{r}'}{r'^3} \right), \quad (4.488)$$

where R is given by

$$R = m' \left(\frac{1}{\Delta} - \frac{\vec{r}' \cdot \vec{r}}{r'^3} \right) = m' \left(\frac{1}{\Delta} - \frac{r}{r'^2} \cos \psi \right), \quad (4.489)$$

in which ψ is the phase angle of the satellite and the perturbing body that

$$\cos \psi = \left(\frac{\vec{r}}{r} \right) \cdot \left(\frac{\vec{r}'}{r'} \right), \quad (4.490)$$

and $\Delta = |\vec{\Delta}|$ is the distance between the satellite and the third body, that

$$\vec{\Delta} = \vec{r} - \vec{r}'. \quad (4.491)$$

Because $r < r'$ we have

$$\frac{1}{\Delta} = (r^2 + r'^2 - 2rr' \cos \psi)^{-\frac{1}{2}} = \frac{1}{r'} \sum_{k=0}^{\infty} P_k(\cos \psi) \left(\frac{r}{r'} \right)^k, \quad (4.492)$$

where $P_k(\cos \psi)$ is the Legendre polynomial of $\cos \psi$ that

$$P_0(\cos \psi) = 1, \quad P_1(\cos \psi) = \cos \psi, \quad \frac{3}{2} \cos^2 \psi - \frac{1}{2}, \dots \quad (4.493)$$

Substituting (4.493) into (4.492) then substituting $1/\Delta$ into (4.489) and omitting the terms only related to the position of the third body we obtain

$$\begin{aligned} R &= \frac{m'}{r'} \sum_{k \geq 2} \left(\frac{r}{r'}\right)^k P_k(\cos \psi) \\ &= m' \left[\frac{r^2}{r^3} \left(\frac{3}{2} \cos^2 \psi - \frac{1}{2} \right) + \frac{r^3}{r^4} \left(\frac{5}{2} \cos^3 \psi - \frac{3}{2} \cos \psi \right) \right. \\ &\quad \left. + \frac{r^4}{r^5} \left(\frac{35}{8} \cos^4 \psi - \frac{15}{4} \cos^2 \psi + \frac{3}{8} \right) + \dots \right] \end{aligned} \quad (4.494)$$

For the orbital motion of an Earth's satellite, we only keep the terms up to $P_4(\cos \psi)$ then the resulting R is

$$\begin{aligned} R &= \beta_2 a^2 \left(\frac{r}{a}\right)^2 \left(\frac{3}{2} \cos^2 \psi - \frac{1}{2} \right) + \beta_3 a^3 \left(\frac{r}{a}\right)^3 \left(\frac{5}{2} \cos^3 \psi - \frac{3}{2} \cos \psi \right) \\ &\quad + \beta_4 a^4 \left(\frac{r}{a}\right)^4 \left(\frac{35}{8} \cos^4 \psi - \frac{15}{4} \cos^2 \psi + \frac{3}{8} \right) \end{aligned} \quad (4.495)$$

where

$$\beta_2 = m'/r'^3, \quad \beta_3 = m'/r'^4, \quad \beta_4 = m'/r'^5. \quad (4.496)$$

If the arc of extrapolation is not long, the distance of the third body to the center of Earth can be treated as a constant.

For the external perturbation of $\left(\frac{r}{r'}\right) \ll 1$, the terms of $\cos \psi$ in (4.494) can be expressed as

$$\cos \psi = A \cos f + B \sin f, \quad (4.497)$$

where A and B are given by

$$\begin{aligned} A &= \frac{1}{4} \{ (1 - \cos i) [(1 + \cos i') \cos(\omega - \theta + u') + (1 - \cos i') \cos(\omega - \theta - u')] \\ &\quad + (1 + \cos i) [(1 + \cos i') \cos(\omega + \theta - u') + (1 - \cos i') \cos(\omega + \theta + u')] \\ &\quad + 2 \sin i \sin i' [\cos(\omega - u') - \cos(\omega + u')] \} \end{aligned} \quad (4.498)$$

$$\begin{aligned} B &= -\frac{1}{4} \{ (1 - \cos i) [(1 + \cos i') \sin(\omega - \theta + u') + (1 - \cos i') \sin(\omega - \theta - u')] \\ &\quad + (1 + \cos i) [(1 + \cos i') \sin(\omega + \theta - u') + (1 - \cos i') \sin(\omega + \theta + u')] \\ &\quad + 2 \sin i \sin i' [\sin(\omega - u') - \sin(\omega + u')] \} \end{aligned} \quad (4.499)$$

and

$$\theta = \Omega - \Omega', \quad u' = f' + \omega'. \quad (4.500)$$

All elements marked by an apostrophe, i' , Ω' , ω' , f' , etc., are orbital elements of the perturbing body (the Sun or the Moon) with respect to Earth. In the perturbation function, the only fast changing variable is the true anomaly f of the satellite, whereas A and B are functions of slow changing variables including satellite orbital elements Ω , and ω , and Ω' , ω' , and f' of the third body.

It is easy to decompose the perturbation function R into secular, long-period, and short-period terms. First, we substitute (4.497) into R (4.495) to obtain

$$R = R_2(\beta_2) + R_3(\beta_3) + R_4(\beta_4), \quad (4.501)$$

$$R_2 = \frac{3}{2}\beta_2 a^2 \left(\frac{r}{a}\right)^2 [S_1 + S_2 \cos 2f + S_3 \sin 2f], \quad (4.502)$$

$$R_3 = \frac{3}{2}\beta_3 a^3 \left(\frac{r}{a}\right)^3 [S_4 \cos f + S_5 \cos 3f + S_6 \sin f + S_7 \sin 3f], \quad (4.503)$$

$$R_4 = \frac{15}{8}\beta_4 a^4 \left(\frac{r}{a}\right)^4 [S_8 + S_9 \cos 2f + S_{10} \cos 4f + S_{11} \sin 2f + S_{12} \sin 4f], \quad (4.504)$$

where

$$S_1 = -\frac{1}{3} + \frac{1}{2}(A^2 + B^2), \quad S_2 = \frac{1}{2}(A^2 - B^2), \quad S_3 = AB, \quad (4.505)$$

$$\begin{cases} S_4 = \left[\frac{5}{4}(A^2 + B^2) - 1\right]A, & S_5 = \frac{5}{12}(A^2 - 3B^2)A \\ S_6 = \left[\frac{5}{4}(A^2 + B^2) - 1\right]B, & S_7 = \frac{5}{12}(3A^2 - B^2)B \end{cases} \quad (4.506)$$

$$\begin{cases} S_8 = \frac{1}{5} - (A^2 + B^2) + \frac{7}{8}(A^4 + B^4 + 2A^2 B^2) \\ S_9 = \frac{7}{6}(A^4 - B^4) - (A^2 - B^2), & S_{10} = \frac{7}{24}(A^4 + B^4) - 6A^2 B^2 \\ S_{11} = \frac{7}{3}AB(A^2 + B^2) - 2AB, & S_{12} = \frac{7}{6}AB(A^2 - B^2) \end{cases} \quad (4.507)$$

By the averages of $\left(\frac{r}{a}\right)^2$, $\left(\frac{r}{a}\right)^2 \cos 2f, \dots$, we obtain the secular terms (including the long-period terms), $\bar{R}_2$, $\bar{R}_3$, $\bar{R}_4$, and the short-period terms, R_{2s} , R_{3s} , R_{4s} , as follows.

$$\bar{R}_2 = \frac{3}{2}\beta_2 a^2 \left[S_1 \left(1 + \frac{3}{2}e^2\right) + S_2 \left(\frac{5}{2}e^2\right) \right] \quad (4.508)$$

$$R_{2s} = \left(\frac{3}{2}\beta_2 a^2\right) \left\{ S_1 \left[\left(\frac{r}{a}\right)^2 - \left(1 + \frac{3}{2}e^2\right) \right] + S_2 \left[\left(\frac{r}{a}\right)^2 \cos 2f - \frac{5}{2}e^2 \right] + S_3 \left(\frac{r}{a}\right)^2 \sin 2f \right\} \quad (4.509)$$

$$\bar{R}_3 = -\frac{3}{2}\beta_3 a^3 \left[S_4 \left(\frac{5}{2}e\right) \left(1 + \frac{3}{4}e^2\right) + S_5 \left(\frac{35}{8}e^3\right) \right] \quad (4.510)$$

$$R_{3s} = \frac{3}{2}\beta_3 a^3 \left\{ S_4 \left[\left(\frac{r}{a} \right)^3 \cos f + \frac{5}{2}e \left(1 + \frac{3}{4}e^2 \right) \right] + S_5 \left[\left(\frac{r}{a} \right)^3 \cos 3f + \left(\frac{35}{8}e^3 \right) \right] \right. \\ \left. + S_6 \left(\frac{r}{a} \right)^3 \sin f + S_7 \left(\frac{r}{a} \right)^3 \sin 3f \right\} \quad (4.511)$$

$$\bar{R}_4 = \frac{15}{8}\beta_4 a^4 \left[S_8 \left(1 + 5e^2 + \frac{15}{8}e^4 \right) + S_9 \left(\frac{21}{4}e^2 + \frac{21}{8}e^4 \right) + S_{10} \left(\frac{63}{8}e^4 \right) \right] \quad (4.512)$$

$$R_{4s} = \frac{15}{8}\beta_4 a^4 \left\{ S_8 \left[\left(\frac{r}{a} \right)^4 - \left(1 + 5e^2 + \frac{15}{8}e^4 \right) \right] + S_9 \left[\left(\frac{r}{a} \right)^4 \cos 2f - \left(\frac{21}{4}e^2 + \frac{21}{8}e^4 \right) \right] \right. \\ \left. + S_{10} \left[\left(\frac{r}{a} \right)^4 \cos 4f - \left(\frac{63}{8}e^4 \right) \right] + S_{11} \left(\frac{r}{a} \right)^4 \sin 2f + S_{12} \left(\frac{r}{a} \right)^4 \sin 4f \right\} \quad (4.513)$$

4.7.2 The Perturbation Solution Due to the Gravity of the Sun or the Moon

Since $r \ll (\Delta, r')$ the perturbation magnitude of a third-body can be estimated as

$$\left(\frac{\vec{\Delta}}{\Delta^3} + \frac{\vec{r}'}{r'^3} \right) \approx \frac{1}{r'^3} (\vec{\Delta} + \vec{r}') = \frac{r}{r'^3}.$$

Then the ratio of the perturbation gravity force F_ε to Earth's central gravity force F_0 on the satellite is given by

$$\frac{F_\varepsilon}{F_0} = m' \left(\frac{r}{r'} \right)^3 \quad (4.514)$$

The actual perturbation magnitudes for a low Earth orbits satellite ($r \leq 1.3$) are

$$\frac{F_\varepsilon}{F_0} = \begin{cases} 0.6 \times 10^{-7}, & \text{Sun,} \\ 1.2 \times 10^{-7}, & \text{Moon,} \end{cases} \quad (4.515)$$

and for a high Earth orbit satellite such as a GEO ($r \approx 6.6$)

$$\frac{F_\varepsilon}{F_0} = \begin{cases} 10^{-5}, & \text{Sun} \\ 2 \times 10^{-5}, & \text{Moon} \end{cases} \quad (4.516)$$

The above perturbation magnitudes are estimated for the main part of the perturbation function $R_2(\beta_2)$. According to the value of (r/r') , the influences of $R_3(\beta_3)$ and $R_4(\beta_4)$ are even smaller. Therefore, the characteristics of the third-body perturbation

for both quantitative and qualitative analyses can be obtained by the perturbation function $R_2(\beta_2)$.

4.7.2.1 The Perturbation Solution by the Main Part of the Gravitational Force of the Sun or the Moon

Using a low Earth orbit satellite as an example, and considering both Earth's non-spherical dynamical form-factor J_2 perturbation and the third-body perturbation, we choose $\varepsilon = 10^{-3}$ as the small parameter, then J_2 term has a first-order magnitude, and the main part of the third-body perturbation has a second-order magnitude. The perturbation function can be written as

$$R = R_1(\sigma; J_2) + R_2(\sigma, \sigma', t; \beta_2). \quad (4.517)$$

For the perturbation equation system in terms of the $(\partial R / \partial \sigma)$ and using the same nomenclature we have

$$\dot{\sigma} = f_0(a) + f_1(\sigma; J_2) + f_2(\sigma, \sigma', t; \beta_2), \quad (4.518)$$

where σ is for the six orbital elements $(a, e, i, \Omega, \omega, M)$; the function f_2 is the perturbation function of the third-body, the short-period part of f_2 can be separated, that

$$f_2 = \bar{f}_2 + f_{2s}, \quad (4.519)$$

where $\bar{f}_2$ is given by $\bar{R}_2(\beta_2)$ that

$$\bar{f}_2 = \text{grad}(\bar{R}_2) = f_{2c} + f_{2l}. \quad (4.520)$$

f_{2c} and f_{2l} are the secular part and the long-period part, respectively.

The perturbation solutions in the quasi-mean elements are given as follows. Note that the value β_2 depends on the third body, the Sun or the Moon.

(1) Secular terms and long-period terms

$$\Delta\sigma(t) = \sigma_2(t - t_0) + \Delta\sigma_l^{(1)}(t, \beta_2),$$

$$\Delta a(t, \beta_2) = 0, \quad (4.521)$$

$$\Delta e(t, \beta_2) = -\left(\frac{3}{2}\beta_2 a^3\right) \left[5e\sqrt{1-e^2}S_3\right] n(t - t_0), \quad (4.522)$$

$$\begin{aligned}\Delta i(t, \beta_2) = & -\left(\frac{3}{2}\beta_2 a^3\right) \frac{1}{\sqrt{1-e^2} \sin i} \left[\left(1 + \frac{3}{2}e^2\right) (AA_2 + BB_2) \right. \\ & \left. + 5e^2 \left(\frac{1}{2}(AA_2 - BB_2) - \cos i(AB)\right) \right] n(t - t_0)\end{aligned}\quad (4.523)$$

$$\Delta \Omega(t, \beta_2) = \left(\frac{3}{2}\beta_2 a^3\right) \frac{1}{\sqrt{1-e^2} \sin i} [AA_1(1 + 4e^2) + BB_1(1 - e^2)] n(t - t_0), \quad (4.524)$$

$$\Delta \omega(t, \beta_2) = -\cos i \Delta \Omega(t) + \left(\frac{3}{2}\beta_2 a^3\right) \sqrt{1-e^2} [3S_1 + 5S_2] n(t - t_0), \quad (4.525)$$

$$\begin{aligned}\Delta M(t, \beta_2) = & -\sqrt{1-e^2} [\Delta \omega(t) + \cos i \Delta \Omega(t)] \\ & - \left(\frac{3}{2}\beta_2 a^3\right) \left[4\left(1 + \frac{3}{2}e^2\right) S_1 + 10e^2 S_2 \right] n(t - t_0)\end{aligned}\quad (4.526)$$

The orbital elements a, e, i, Ω , and ω in formulas above and below are for $\bar{a}_0, \bar{e}_0, \bar{i}_0, \bar{\Omega}_0, \bar{\omega}_0$. The auxiliaries are given by

$$\begin{aligned}A_1 = \frac{\partial A}{\partial i} = & \frac{1}{4} \{ \sin i [(1 + \cos i') \cos(\omega - \theta + u') + (1 - \cos i') \cos(\omega - \theta - u')] \\ & - \sin i [(1 + \cos i') \cos(\omega + \theta - u') + (1 - \cos i') \cos(\omega + \theta + u')] \\ & + 2 \cos i \sin i' [\cos(\omega - u') - \cos(\omega + u')] \} \end{aligned} \quad (4.527)$$

$$\begin{aligned}B_1 = \frac{\partial B}{\partial i} = & -\frac{1}{4} \{ \sin i [(1 + \cos i') \sin(\omega - \theta + u') + (1 - \cos i') \sin(\omega - \theta - u')] \\ & - \sin i [(1 + \cos i') \sin(\omega + \theta - u') + (1 - \cos i') \sin(\omega + \theta + u')] \\ & + 2 \cos i \sin i' [\sin(\omega - u') - \sin(\omega + u')] \} \end{aligned} \quad (4.528)$$

$$\begin{aligned}A_2 = \frac{\partial A}{\partial \Omega} = & \frac{1}{4} \{ (1 - \cos i) [(1 + \cos i') \sin(\omega - \theta + u') + (1 - \cos i') \sin(\omega - \theta - u')] \\ & - (1 + \cos i) [(1 + \cos i') \sin(\omega + \theta - u') + (1 - \cos i') \sin(\omega + \theta + u')] \} \end{aligned} \quad (4.529)$$

$$\begin{aligned}B_2 = \frac{\partial B}{\partial \Omega} = & \frac{1}{4} \{ (1 - \cos i) [(1 + \cos i') \cos(\omega - \theta + u') + (1 - \cos i') \cos(\omega - \theta - u')] \\ & - (1 + \cos i) [(1 + \cos i') \cos(\omega + \theta - u') + (1 - \cos i') \cos(\omega + \theta + u')] \} \end{aligned} \quad (4.530)$$

(2) Short-period terms of $\sigma_s^{(2)}(t, \beta_2)$

$$a_s^{(2)}(t, \beta_2) = \frac{2}{n^2 a} R_{2s} = 2a^2 R_{2s} = 2a \left(\frac{3}{2} \beta_2 a^3 \right) G_1 \quad (4.531)$$

$$e_s^{(2)}(t, \beta_2) = \left(\frac{3}{2}\beta_2 a^3\right) G_2 \quad (4.532)$$

$$i_s^{(2)}(t, \beta_2) = \left(\frac{3}{2}\beta_2 a^3\right) \frac{1}{\sqrt{1-e^2} \sin i} (-\cos i G_3 + G_4) \quad (4.533)$$

$$\Omega_s^{(2)}(t, \beta_2) = \left(\frac{3}{2}\beta_2 a^3\right) \frac{1}{\sqrt{1-e^2} \sin i} (-G_5) \quad (4.534)$$

$$\omega_s^{(2)}(t, \beta_2) = \left(\frac{3}{2}\beta_2 a^3\right) \frac{\sqrt{1-e^2}}{e} G_6 - \cos i \Omega_s^{(2)}(t) \quad (4.535)$$

$$M_s^{(2)}(t, \beta_2) = \left(\frac{3}{2}\beta_2 a^3\right) \left(-\frac{1-e^2}{e} G_6 + 7G_7\right) \quad (4.536)$$

where the auxiliaries $G_1, G_2, \dots, G_{10}$ are

$$\begin{aligned} G_1 &= S_1 \left[\left(\frac{r}{a}\right)^2 - \left(1 + \frac{3}{2}e^2\right) \right] + S_2 \left[\left(\frac{r}{a}\right)^2 \cos 2f - \frac{5}{2}e^2 \right] + S_3 \left[\left(\frac{r}{a}\right)^2 \sin 2f \right] \\ &= S_1 \left[-e^2 - 2e \cos E + \frac{1}{2}e^2 \cos 2E \right] + S_2 \left[-e^2 - 2e \cos E + \left(1 - \frac{1}{2}e^2\right) \cos 2E \right] \\ &\quad + S_3 \sqrt{1-e^2} [-2e \sin E + \sin 2E] \end{aligned} \quad (4.537)$$

$$\begin{aligned} G_2 &= S_1 (1-e^2) \left[-e - 2 \cos E + \frac{1}{2}e \cos 2E \right] \\ &\quad + S_2 (1-e^2) \left[-e + 3 \cos E - \frac{3}{2}e \cos 2E + \frac{1}{3} \cos 3E \right] \\ &\quad + S_3 \sqrt{1-e^2} \left[\left(3 - \frac{1}{2}e^2\right) \sin E - \frac{3}{2}e \sin 2E + \left(\frac{1}{3} - \frac{1}{6}e^2\right) \sin 3E \right] \end{aligned} \quad (4.538)$$

$$G_3 = S_2 G_8 + S_3 G_9 = \frac{1}{2} (A^2 - B^2) G_8 + (AB) G_9 \quad (4.539)$$

$$G_4 = -\frac{1}{2} (AB_2 + A_2 B) G_8 + \frac{1}{2} (AA_2 - BB_2) G_9 + (AA_2 + BB_2) G_{10} \quad (4.540)$$

$$G_5 = -\frac{1}{2} (AB_1 + A_1 B) G_8 + \frac{1}{2} (AA_1 - BB_1) G_9 + (AA_1 + BB_1) G_{10} \quad (4.541)$$

$$\begin{aligned} G_6 &= S_1 \left[(-2 + e^2) \sin E + \frac{1}{2}e \sin 2E \right] + S_2 \left[(-3 + 3e^2) \sin E - \frac{1}{2}e \sin 2E + \frac{1}{3} \sin 3E \right] \\ &\quad + \frac{S_3}{\sqrt{1-e^2}} \left[\left(3 - \frac{11}{2}e^2\right) \cos E + \left(\frac{1}{2} + \frac{1}{2}e^2\right) e \cos 2E - \left(\frac{1}{3} - \frac{1}{6}e^2\right) \cos 3E \right] \end{aligned} \quad (4.542)$$

$$G_7 = S_1 G_{10} + \frac{1}{2} S_2 G_9 - \frac{1}{2} S_3 G_8 \quad (4.543)$$

$$G_8 = \sqrt{1 - e^2} \left[5e \cos E - (1 + e^2) \cos 2E + \frac{1}{3} e \cos 3E \right] \quad (4.544)$$

$$G_9 = \left(5 - \frac{5}{2} e^2 \right) e \sin E - \left(1 + \frac{1}{2} e^2 \right) \sin 2E + \left(\frac{1}{3} - \frac{1}{6} e^2 \right) e \sin 3E \quad (4.545)$$

$$G_{10} = \left(2 - \frac{3}{4} e^2 \right) e \sin E - \frac{3}{4} e^2 \sin 2E + \frac{1}{12} e^3 \sin 3E \quad (4.546)$$

In the fields of orbital designs, orbital measurement and control, for providing the necessary information of satellite orbital variation, we may need to know the details of the perturbation solution due to the Sun and the Moon. Then we must separate the solution rigorously into the secular terms, the long-period terms, and the short-period terms. The method is to decompose $\Delta\sigma(t)$ which includes the secular and the long-period terms, into $\sigma_2(t - t_0)$ and $\Delta\sigma_l^{(1)}(t, \beta_2)$. Omitting the procedure, the results of the secular variation rates are

$$\begin{cases} a_2 = 0, & e_2 = 0, & i_2 = 0 \\ \Omega_2 = -\left(\frac{3}{4}\beta_2 a^3\right)\left(1 - \frac{3}{2}\sin^2 i'\right)\left(1 + \frac{3}{2}e^2\right)(1 - e^2)^{-1/2}n \cos i \\ \omega_2 = \left(\frac{3}{4}\beta_2 a^3\right)\left(1 - \frac{3}{2}\sin^2 i'\right)\left[\left(2 - \frac{5}{2}\sin^2 i\right) + \frac{1}{2}e^2\right](1 - e^2)^{-1/2}n \\ M_2 = -\left(\frac{3}{4}\beta_2 a^3\right)\left(1 - \frac{3}{2}\sin^2 i'\right)\left(1 - \frac{3}{2}\sin^2 i\right)\left(\frac{7}{3} + e^2\right)n \end{cases} \quad (4.547)$$

Also in the long-period terms, there are terms with commensurable divisors caused by slow variables, which can be seen from the definitions of A and B in (4.498) and (4.499), respectively. Readers who need more information about the structures of the secular and long-period variations of the Sun and the Moon perturbations may refer to the related works by the author [3–5].

(3) Orbital elements of the Sun and the Moon

If we only need the values of orbital elements of the Sun or the Moon at a given time, we can transfer the coordinates of the position vector and velocity from the Earth's heliocentric ecliptic frame (or the Moon's geocentric ecliptic frame) to the corresponding position vector and velocity in J2000 geocentric celestial frame, then transfer to orbital elements σ' .

For some problems in the process of constructing the perturbation solutions, we may need σ' as a function of t , then we need a different method, which is provided as follows. For the Sun there are

$$i' = \varepsilon, \quad \Omega' = 0, \quad u' = L_\odot, \quad (4.548)$$

where ε is the obliquity, $L_\odot = M_\odot + \omega_\odot + \Omega_\odot$ is the mean ecliptic longitude of the Sun. For the Moon,

$$u' = M' + \omega' = L_m - (\Omega_m - \theta_m), \quad (4.549)$$

$$\begin{cases} \cos i' = \cos \varepsilon \cos J - \sin \varepsilon \sin J \cos \Omega_m \\ \sin i' = (1 - \cos^2 i')^{1/2} \end{cases} \quad (4.550)$$

$$\begin{cases} \sin \Omega' = \frac{\sin J \sin \Omega_m}{\sin i'} \\ \cos \Omega' = \frac{\cos J - \cos \varepsilon \cos i'}{\sin \varepsilon \sin i'} \end{cases} \quad (4.551)$$

$$\sin(\Omega_m - \theta_m) = \frac{1}{\sin i'} \left(\cos \varepsilon \sin \Omega_m \sin J - \sin \varepsilon \sin 2\Omega_m \sin^2 \frac{J}{2} \right) \quad (4.552)$$

where $L_m = M_m + \omega_m + \Omega_m$. L_m and Ω_m are the mean ecliptic longitudes of the Moon and its orbital ascending node, respectively. Because the orbital intersection of the ecliptic and the Moon's orbit is $J = 5^\circ 09'$, the angle $\Omega_m - \theta_m = O(J)$ in (4.552) has a defined quadrant.

The Sun's orbital elements, $\sigma_\odot$, and the Moon's orbital elements, σ_m , can be given by a simple average ephemeris, the calculation formulas are given in Appendix 2. But for the Moon, the variation of its orbit due to perturbation is large, the largest amplitude of the periodic variation can be 2×10^{-2} , the accuracy of its orbital elements given by the average ephemeris is relatively low, and can reduce the accuracy of the analytical perturbation solutions for high Earth orbit satellites. Then if the periodic terms with large amplitudes are included the entire problem becomes complicated.

4.7.2.2 The Non-singularity Solution of the First Type

The above perturbation solution given by the quasi-mean elements method eliminates the commensurable singularity. We can further eliminate the small e singularity by rearranging the solution using the solution of Kepler orbital elements. The results are:

(1) Secular and long-period terms $\Delta\sigma(t) = \sigma_2(t - t_0) + \Delta\sigma_I^{(1)}(t, \beta_2)$

$$\Delta a(t, \beta_2) = 0, \quad (4.553)$$

$$\begin{aligned} \Delta i(t, \beta_2) = & -\left(\frac{3}{2}\beta_2 a^3\right) \frac{1}{\sqrt{1-e^2} \sin i} \left[\left(1 + \frac{3}{2}e^2\right) (AA_2 + BB_2) \right. \\ & \left. + 5e^2 \left(\frac{1}{2}(AA_2 - BB_2) - \cos i(AB)\right) \right] n(t - t_0) \end{aligned} \quad (4.554)$$

$$\Delta\Omega(t, \beta_2) = \left(\frac{3}{2}\beta_2 a^3\right) \frac{1}{\sqrt{1-e^2} \sin i} [AA_1(1+4e^2) + BB_1(1-e^2)]n(t-t_0), \quad (4.555)$$

$$\Delta\xi(t, \beta_2) = \cos \omega[\Delta e(t, \beta_2)] - \sin \omega[e\Delta\omega(t, \beta_2)], \quad (4.556)$$

$$\Delta\eta(t, \beta_2) = \sin \omega[\Delta e(t, \beta_2)] + \cos \omega[e\Delta\omega(t, \beta_2)], \quad (4.557)$$

$$\begin{aligned} \Delta\lambda(t, \beta_2) &= [\Delta M(t, \beta_2) + \Delta\omega(t, \beta_2)] \\ &= -\cos i \Delta\Omega(t) \\ &\quad + \frac{e^2}{1+\sqrt{1-e^2}} \left(\frac{3}{2}\beta_2 a^3\right) \sqrt{1-e^2} [3S_1 + 5S_2]n(t-t_0) \\ &\quad - \left(\frac{3}{2}\beta_2 a^3\right) \left[4\left(1+\frac{3}{2}e^2\right)S_1 + 10e^2 S_2\right]n(t-t_0) \end{aligned} \quad (4.558)$$

where expressions of $\Delta e(t)$ and $\Delta\omega(t)$ are given by (4.522) and (4.525), respectively.

(2) Short-period terms $\sigma_s^{(2)}(t, \beta_2)$

$$a_s^{(2)}(t, \beta_2) = 2a\left(\frac{3}{2}\beta_2 a^3\right)G_1, \quad (4.559)$$

$$i_s^{(2)}(t, \beta_2) = \left(\frac{3}{2}\beta_2 a^3\right) \frac{1}{\sqrt{1-e^2}} (I_{S1}G_8 + I_{S2}G_9 + I_{S3}G_{10}), \quad (4.560)$$

$$\Omega_s^{(2)}(t, \beta_2) = \left(\frac{3}{2}\beta_2 a^3\right) \frac{1}{\sqrt{1-e^2} \sin i} (-G_5), \quad (4.561)$$

$$\xi_s^{(2)}(t, \beta_2) = \cos \omega[e_s^{(2)}(t, \beta_2)] - \sin \omega[e\omega_s^{(2)}(t, \beta_2)], \quad (4.562)$$

$$\eta_s^{(2)}(t, \beta_2) = \sin \omega[e_s^{(2)}(t, \beta_2)] + \cos \omega[e\omega_s^{(2)}(t, \beta_2)], \quad (4.563)$$

$$\begin{aligned} \lambda_s^{(2)}(t, \beta_2) &= M_s^{(2)}(t, \beta_2) + \omega_s^{(2)}(t, \beta_2) \\ &= -\cos i \Omega_s^{(2)}(t, \beta_2) + \left(\frac{3}{2}\beta_2 a^3\right) \left[\sqrt{1-e^2} \left(\frac{e}{1+\sqrt{1-e^2}} \right) (G_6) + 7G_7 \right], \end{aligned} \quad (4.564)$$

where $e_s^{(2)}(t)$ and $\omega_s^{(2)}(t)$ are given by (4.532) and (4.535), respectively.

4.7.2.3 The Non-singularity Solution of the Second Type

The non-singularity solution of the second type, which may be needed for geosynchronous satellites, can also be given by rearranging the solution of Kepler orbital elements to eliminate the small i singularity, the results are as follows.

(1) Secular terms and long-period terms $\Delta\sigma(t) = \sigma_2(t - t_0) + \Delta\sigma_l^{(1)}(t, \beta_2)$

$$\Delta a(t, \beta_2) = 0, \quad (4.565)$$

$$\Delta\xi(t, \beta_2) = \cos\tilde{\omega}[\Delta e(t, \beta_2)] - \sin\tilde{\omega}[e\Delta\tilde{\omega}(t, \beta_2)], \quad (4.566)$$

$$\Delta\eta(t, \beta_2) = \sin\tilde{\omega}[\Delta e(t, \beta_2)] + \cos\tilde{\omega}[e\Delta\tilde{\omega}(t, \beta_2)], \quad (4.567)$$

$$\Delta h(t, \beta_2) = \frac{1}{2} \cos(i/2) \cos\Omega[\Delta i(t, \beta_2)] - \frac{1}{2 \cos(i/2)} \sin\Omega[\sin i \Delta\Omega(t, \beta_2)], \quad (4.568)$$

$$\Delta k(t, \beta_2) = \frac{1}{2} \cos(i/2) \sin\Omega[\Delta i(t, \beta_2)] + \frac{1}{2 \cos(i/2)} \cos\Omega[\sin i \Delta\Omega(t, \beta_2)], \quad (4.569)$$

$$\begin{aligned} \Delta\lambda(t, \beta_2) &= [\Delta M(t, \beta_2) + \Delta\tilde{\omega}(t, \beta_2)] \\ &= \frac{\sin i}{1 + \cos i} [\sin i \Delta\Omega(t)] + \left(\frac{3}{2}\beta_2 a^3\right) \left\{ (1 - \sqrt{1 - e^2}) \sqrt{1 - e^2} [3S_1 + 5S_2] \right. \\ &\quad \left. - [4(1 + 3e^2/2)S_1 + 10e^2S_2] \right\} n(t - t_0) \end{aligned} \quad (4.570)$$

where $\Delta e(t)$ and $\Delta i(t)$ are given by (4.522) and (4.523), respectively; and $\sin i \Delta\Omega(t)$ and $\Delta\tilde{\omega}(t)$ are given by

$$\sin i \Delta\Omega(t, \beta_2) = \left(\frac{3}{2}\beta_2 a^3\right) \frac{1}{\sqrt{1 - e^2}} [AA_1(1 + 4e^2) + BB_1(1 - e^2)] n(t - t_0), \quad (4.571)$$

$$\begin{aligned} \Delta\tilde{\omega}(t, \beta_2) &= \frac{\sin i}{1 + \cos i} [\sin i \Delta\Omega(t, \beta_2)] \\ &\quad + \left(\frac{3}{2}\beta_2 a^3\right) \sqrt{1 - e^2} [3S_1 + 5S_2] n(t - t_0) \end{aligned} \quad (4.572)$$

I_{S2} and I_{S3} in $\Delta i(t)$ are given in the short-period terms.

(2) Short-period terms $\sigma_s^{(2)}(t, \beta_2)$

$$a_s^{(2)}(t, \beta_2) = 2a \left(\frac{3}{2} \beta_2 a^3 \right) G_1, \quad (4.573)$$

$$\xi_s^{(2)}(t, \beta_2) = \cos \tilde{\omega} \left[e_s^{(2)}(t, \beta_2) \right] - \sin \tilde{\omega} \left[e \tilde{\omega}_s^{(2)}(t, \beta_2) \right], \quad (4.574)$$

$$\eta_s^{(2)}(t, \beta_2) = \sin \tilde{\omega} \left[e_s^{(2)}(t, \beta_2) \right] + \cos \tilde{\omega} \left[e \tilde{\omega}_s^{(2)}(t, \beta_2) \right], \quad (4.575)$$

$$h_s^{(2)}(t, \beta_2) = \frac{1}{2} \cos(i/2) \cos \Omega \left[i_s^{(2)}(t, \beta_2) \right] - \frac{1}{2 \cos(i/2)} \sin \Omega \left[\sin i \Omega_s^{(2)}(t, \beta_2) \right], \quad (4.576)$$

$$k_s^{(2)}(t, \beta_2) = \frac{1}{2} \cos(i/2) \sin \Omega \left[i_s^{(2)}(t, \beta_2) \right] + \frac{1}{2 \cos(i/2)} \cos \Omega \left[\sin i \Omega_s^{(2)}(t, \beta_2) \right], \quad (4.577)$$

$$\begin{aligned} \lambda_s^{(2)}(t, \beta_2) &= M_s^{(2)}(t, \beta_2) + \omega_s^{(2)}(t) + \Omega_s^{(2)}(t, \beta_2) \\ &= \left(\frac{3}{2} \beta_2 a^3 \right) \left[\frac{1}{\sqrt{1-e^2}} \left(\frac{\sin i}{1+\cos i} \right) (-G_5) + \sqrt{1-e^2} \left(\frac{e}{1+\sqrt{1-e^2}} \right) (G_6) + 7G_7 \right] \end{aligned} \quad (4.578)$$

where $e_s^{(2)}(t, \beta_2)$ is given by (4.532), and

$$i_s^{(2)}(t, \beta_2) = \left(\frac{3}{2} \beta_2 a^3 \right) \frac{1}{\sqrt{1-e^2}} (I_{S1} G_8 + I_{S2} G_9 + I_{S3} G_{10}), \quad (4.579)$$

$$\sin i \Omega_s^{(2)}(t, \beta_2) = \left(\frac{3}{2} \beta_2 a^3 \right) \frac{1}{\sqrt{1-e^2}} (-G_5), \quad (4.580)$$

$$e \tilde{\omega}_s^{(2)}(t, \beta_2) = \left(\frac{3}{2} \beta_2 a^3 \right) \left[\frac{e}{\sqrt{1-e^2}} \left(\frac{\sin i}{1+\cos i} \right) (-G_5) + \sqrt{1-e^2} (G_6) \right], \quad (4.581)$$

where I_{S1} , I_{S2} , and I_{S3} in $\Delta i(t)$ and $i_s^{(2)}(t, \beta_2)$ are given by

$$I_{S1} = \frac{1}{\sin i} \left\{ -\frac{1}{2} [(AB_2 + A_2B) + \cos i (A^2 - B^2)] \right\}, \quad (4.582)$$

$$I_{S2} = \frac{1}{\sin i} \left\{ \frac{1}{2} (AA_2 - BB_2) - \cos i (AB) \right\}, \quad (4.583)$$

$$I_{S3} = \frac{1}{\sin i} (AA_2 + BB_2), \quad (4.584)$$

These expressions are for eliminating the factor $(1/\sin i)$ in the original expression of $i_s^{(2)}(t, \beta_2)$ (4.533). The factor $(1/\sin i)$ can be eliminated after the term $i_s^{(2)}(t, \beta_2)$ is transferred to (4.579), and with A, B, A_2 , and B_2 re-written as follows.

$$A = A_0 + \Delta A$$

$$\begin{aligned} A_0 &= \frac{1}{4} \sin i \left\{ \frac{\sin i}{(1 + \cos i)} [(1 + \cos i') \cos(\omega - \theta + u') + (1 - \cos i') \cos(\omega - \theta - u')] \right. \\ &\quad \left. + 2 \sin i' [\cos(\omega - u') - \cos(\omega + u')] \right\} \\ \Delta A &= \frac{1}{4} \{ (1 + \cos i) [(1 + \cos i') \cos(\omega + \theta - u') + (1 - \cos i') \cos(\omega + \theta + u')] \} \end{aligned} \quad (4.585)$$

$$B = B_0 + \Delta B$$

$$\begin{aligned} B_0 &= -\frac{1}{4} \sin i \{ \sin i / (1 + \cos i) [(1 + \cos i') \sin(\omega - \theta + u') + (1 - \cos i') \sin(\omega - \theta - u')] \\ &\quad + 2 \sin i' [\sin(\omega - u') - \sin(\omega + u')] \} \\ \Delta B &= -\frac{1}{4} \{ (1 + \cos i) [(1 + \cos i') \sin(\omega + \theta - u') + (1 - \cos i') \sin(\omega + \theta + u')] \} \end{aligned} \quad (4.586)$$

$$A_2 = A_{20} + \Delta A_2$$

$$\begin{aligned} A_{20} &= \frac{1}{4} \sin i \{ \sin i / (1 + \cos i) [(1 + \cos i') \sin(\omega - \theta + u') + (1 - \cos i') \sin(\omega - \theta - u')] \} \\ \Delta A_2 &= -\frac{1}{4} \{ (1 + \cos i) [(1 + \cos i') \sin(\omega + \theta - u') + (1 - \cos i') \sin(\omega + \theta + u')] \} \end{aligned} \quad (4.587)$$

$$B_2 = B_{20} + \Delta B_2$$

$$\begin{aligned} B_{20} &= \frac{1}{4} \sin i \{ \sin i / (1 + \cos i) [(1 + \cos i') \cos(\omega - \theta + u') + (1 - \cos i') \cos(\omega - \theta - u')] \} \\ \Delta B_2 &= -\frac{1}{4} \{ (1 + \cos i) [(1 + \cos i') \cos(\omega + \theta - u') + (1 - \cos i') \cos(\omega + \theta + u')] \} \end{aligned} \quad (4.588)$$

Note that there are the following relationships

$$\Delta A_2 = \Delta B, \quad \Delta B_2 = -\Delta A \quad (4.589)$$

Finally, I_{S1} , I_{S2} , and I_{S3} are given by

$$\begin{aligned} I_{S1} &= -\frac{1}{2 \sin i} \{ [(A_0 B_2 + A_{20} B) + (B_0 \Delta B + B_{20} \Delta A)] + \cos i [A_0 (A + \Delta A) - B_0 (B + \Delta B)] \} \\ &\quad + \frac{\sin i}{2(1 + \cos i)} (\Delta A^2 - \Delta B^2) \end{aligned} \quad (4.590)$$

$$\begin{aligned} I_{S2} &= \frac{1}{\sin i} \left\{ \frac{1}{2} [(A_0 A_2 + A_{20} \Delta A) - (B_0 B_2 + B_{20} \Delta B)] - \cos i [A_0 B + B_0 \Delta A] \right\} \\ &\quad + \frac{\sin i}{(1 + \cos i)} (\Delta A \Delta B) \end{aligned} \quad (4.591)$$

$$I_{S3} = \frac{1}{\sin i} [(A_0 A_2 + B_0 B_2) + (A_{20} \Delta A + B_{20} \Delta B)] \quad (4.592)$$

Because in the expressions of A_0 , B_0 , A_{20} , and B_{20} (4.585)–(4.588), there is the factor $\sin i$, then I_{S1} , I_{S2} , and I_{S3} no longer have the factor $(1/\sin i)$, and the problem of the small i in the solution $i_s^{(1)}(t, \beta_2)$ is solved. This treatment is also used in (4.560). It should be careful when developing computer software for calculating these variables.

4.7.2.4 The Perturbation Solution Due to the Part of the Perturbation Functions of the Sun or the Moon Related to $P_3(\cos \psi)$

In the perturbation function of the Sun or the Moon, the part of R_3 is formed by $P_3(\cos \psi)$. After decomposing R_3 , the results are given by (4.503), (4.510), and (4.511), which are presented here again that

$$R_3 = \frac{3}{2}\beta_3 a^3 \left(\frac{r}{a}\right)^3 [S_4 \cos f + S_5 \cos 3f + S_6 \sin f + S_7 \sin 3f], \quad (4.593)$$

$$\bar{R}_3 = -\frac{3}{2}\beta_3 a^3 \left[S_4 \left(\frac{5}{2}e\right) \left(1 + \frac{3}{4}e^2\right) + S_5 \left(\frac{35}{8}e^3\right) \right], \quad (4.594)$$

$$R_{3s} = \frac{3}{2}\beta_3 a^3 \left\{ S_4 \left[\left(\frac{r}{a}\right)^3 \cos f + \frac{5}{2}e \left(1 + \frac{3}{4}e^2\right) \right] + S_5 \left[\left(\frac{r}{a}\right)^3 \cos 3f + \left(\frac{35}{8}e^3\right) \right] \right. \\ \left. + S_6 \left(\frac{r}{a}\right)^3 \sin f + S_7 \left(\frac{r}{a}\right)^3 \sin 3f \right\} \quad (4.595)$$

$$\begin{cases} S_4 = \left[\frac{5}{4}(A^2 + B^2) - 1 \right] A, & S_5 = \frac{5}{12}(A^2 - 3B^2) A \\ S_6 = \left[\frac{5}{4}(A^2 + B^2) - 1 \right] B, & S_7 = \frac{5}{12}(3A^2 - B^2) B \end{cases} \quad (4.596)$$

(1) The perturbation solution in Kepler elements

① Secular terms and Long-period terms $\Delta\sigma(t) = \Delta\sigma_l(t, \beta_3)$

$$\Delta a_l(t, \beta_3) = 0, \quad (4.597)$$

$$\Delta e_l(t, \beta_3) = \left(\frac{3}{2}\beta_3 a^4\right) \sqrt{1-e^2} \left[\frac{5}{2} \left(1 + \frac{3}{4}e^2\right) S_6 + \frac{105}{8} e^2 S_7 \right] n(t - t_0), \quad (4.598)$$

$$\Delta i_l(t, \beta_3) = \left(\frac{3}{2}\beta_3 a^4\right) \frac{1}{\sqrt{1-e^2} \sin i} \left\{ \left[\frac{5}{2} \left(1 + \frac{3}{4}e^2\right) e G_{11} + \frac{35}{8} e^3 G_{12} \right] \right. \\ \left. - \cos i \left[\frac{5}{2} \left(1 + \frac{3}{4}e^2\right) e S_6 + \frac{35}{8} e^3 (3S_7) \right] \right\} n(t - t_0) \quad (4.599)$$

$$\Delta\Omega_l(t, \beta_3) = -\left(\frac{3}{2}\beta_3 a^4\right) \frac{1}{\sqrt{1-e^2} \sin i} \left[\frac{5}{2} \left(1 + \frac{3}{4}e^2\right) e G_{13} + \frac{35}{8} e^3 G_{14} \right] n(t-t_0), \quad (4.600)$$

$$\begin{aligned} \Delta\omega_l(t, \beta_3) = & -\cos i \cdot \Delta\Omega_l(t, \beta_3) \\ & - \left(\frac{3}{2}\beta_3 a^4\right) \frac{\sqrt{1-e^2}}{e} \left[\frac{5}{2} \left(1 + \frac{9}{4}e^2\right) S_4 + \left(\frac{105}{8}e^2\right) S_5 \right] n(t-t_0), \end{aligned} \quad (4.601)$$

$$\begin{aligned} \Delta M_l(t, \beta_3) = & -\sqrt{1-e^2} [\Delta\omega_l(t, \beta_3) + \cos i \cdot \Delta\Omega_l(t, \beta_3)] \\ & + \left(\frac{3}{2}\beta_3 a^4\right) \left[15e \left(1 + \frac{3}{4}e^2\right) S_4 + \left(\frac{105}{4}e^3\right) S_5 \right] n(t-t_0), \end{aligned} \quad (4.602)$$

$$\begin{cases} G_{11} = \left[\frac{5}{2}(AA_2 + BB_2)\right]A + \left[\frac{5}{4}(A^2 + B^2) - 1\right]A_2 \\ G_{12} = \left[\frac{5}{6}(AA_2 - 3BB_2)\right]A + \left[\frac{5}{12}(A^2 - 3B^2)\right]A_2 \\ G_{13} = \left[\frac{5}{6}(AA_1 + BB_1)\right]A + \left[\frac{5}{4}(A^2 + B^2) - 1\right]A_1 \\ G_{14} = \left[\frac{5}{6}(AA_1 - 3BB_1)\right]A + \left[\frac{5}{12}(A^2 - 3B^2)\right]A_1 \end{cases} \quad (4.603)$$

② Short-period terms $\sigma_s^{(2)}(t, \beta_3)$.

Usually only the $a_s^{(2)}(t, \beta_3)$ is needed that

$$\begin{aligned} a_s^{(2)}(t, \beta_3) = & 3\beta_3 a^5 \left\{ S_4 \left[\left(\frac{r}{a}\right)^3 \cos f + \frac{5}{2}e \left(1 + \frac{3}{4}e^2\right) \right] + S_5 \left[\left(\frac{r}{a}\right)^3 \cos 3f + \left(\frac{35}{8}e^3\right) \right] \right. \\ & \left. + S_6 \left(\frac{r}{a}\right)^3 \sin f + S_7 \left(\frac{r}{a}\right)^3 \sin 3f \right\} \end{aligned} \quad (4.604)$$

The short-period terms for other orbit elements are given in the results of the non-singularity solution of the second type in the text below.

(2) The non-singularity solution of the first type

Results are similar to those given by (4.553)–(4.558) but in which β_2 is replaced by β_3

① Secular terms and long-period terms $\Delta\sigma(t) = \Delta\sigma_l(t, \beta_3)$

.

$$\Delta a(t, \beta_3) = 0, \quad (4.605)$$

$$\Delta i(t, \beta_3) = \left(\frac{3}{2}\beta_3 a^4\right) \frac{1}{\sqrt{1-e^2}} \left[\frac{5}{2}e \left(1 + \frac{3}{4}e^2\right) (I_{S4}) + \frac{35}{8}e^3 (I_{S5}) \right] n(t-t_0), \quad (4.606)$$

$$\Delta\Omega(t, \beta_3) = -\left(\frac{3}{2}\beta_3 a^4\right) \frac{1}{\sqrt{1-e^2} \sin i} \left[\frac{5}{2} \left(1 + \frac{3}{4}e^2\right) e G_{13} + \frac{35}{8} e^3 G_{14} \right] n(t-t_0), \quad (4.607)$$

$$\Delta\xi(t, \beta_3) = \cos \omega[\Delta e(t, \beta_3)] - \sin \omega[e\Delta\omega(t, \beta_3)], \quad (4.608)$$

$$\Delta\eta(t, \beta_3) = \sin \omega[\Delta e(t, \beta_3)] + \cos \omega[e\Delta\omega(t, \beta_3)], \quad (4.609)$$

$$\begin{aligned} \Delta\lambda_l(t, \beta_3) &= \Delta M_l(t, \beta_3) + \Delta\omega_l(t, \beta_3) \\ &= -\cos i \cdot \Delta\Omega_l(t) + \left(\frac{3}{2}\beta_3 a^4\right) \left\{ \left[15e \left(1 + \frac{3}{4}e^2\right) S_4 + \left(\frac{105}{4}e^3\right) S_5 \right] \right. \\ &\quad \left. - \frac{e}{1 + \sqrt{1 - e^2}} \left[\frac{5}{2} \left(1 + \frac{9}{4}e^2\right) S_4 + \left(\frac{105}{8}e^2\right) S_5 \right] \right\} n(t - t_0) \end{aligned} \quad (4.610)$$

where $\Delta e(t, \beta_3)$ and $\Delta\omega(t, \beta_3)$ are given by (4.608) and (4.611), respectively.

② Short-period terms.

The formula for $a_s^{(2)}(t, \beta_3)$ is given by (4.604).

(3) The non-singularity solution of the second type

① Secular terms and long-period terms $\Delta\sigma(t) = \Delta\sigma_l(t, \beta_3)$

$$\Delta a(t, \beta_4) = 0, \quad (4.611)$$

$$\Delta\xi(t, \beta_3) = \cos \tilde{\omega}[\Delta e(t, \beta_3)] - \sin \tilde{\omega}[e\Delta\omega(t, \beta_3) + e\Delta\Omega(t, \beta_3)], \quad (4.612)$$

$$\Delta\eta(t, \beta_3) = \sin \tilde{\omega}[\Delta e(t, \beta_3)] + \cos \tilde{\omega}[e\Delta\omega(t, \beta_3) + e\Delta\Omega(t, \beta_3)], \quad (4.613)$$

$$\Delta h(t, \beta_3) = \frac{1}{2} \cos(i/2) \cos \Omega[\Delta i(t, \beta_3)] - \frac{1}{2 \cos(i/2)} \sin \Omega[\sin i \Delta\Omega(t, \beta_3)], \quad (4.614)$$

$$\Delta k(t, \beta_3) = \frac{1}{2} \cos(i/2) \sin \Omega[\Delta i(t, \beta_3)] + \frac{1}{2 \cos(i/2)} \cos \Omega[\sin i \Delta\Omega(t, \beta_3)] \quad (4.615)$$

$$\begin{aligned} \Delta\lambda_l(t, \beta_3) &= \Delta M_l(t, \beta_3) + \Delta\omega_l(t, \beta_3) \\ &= \left(\frac{3}{2}\beta_3 a^4\right) \left\{ \left[15e \left(1 + \frac{3}{4}e^2\right) S_4 + \left(\frac{105}{4}e^3\right) S_5 \right] \right. \\ &\quad \left. - \frac{e}{1 + \sqrt{1 - e^2}} \left[\frac{5}{2} \left(1 + \frac{9}{4}e^2\right) S_4 + \left(\frac{105}{8}e^2\right) S_5 \right] \right. \\ &\quad \left. - \frac{\sin i}{\sqrt{1 - e^2}(1 + \sin i)} \left[\frac{5}{2} \left(1 + \frac{3}{4}e^2\right) eG_{13} + \frac{35}{8}e^3G_{14} \right] \right\} n(t - t_0), \end{aligned} \quad (4.616)$$

② Short-period terms $\sigma_s^{(2)}(t, \beta_3)$

$$a_s^{(2)}(t, \beta_3) = 3\beta_3 a^5 \left\{ S_4 \left[\left(\frac{r}{a} \right)^3 \cos f + \frac{5}{2} e \left(1 + \frac{3}{4} e^2 \right) \right] + S_5 \left[\left(\frac{r}{a} \right)^3 \cos 3f + \left(\frac{35}{8} e^3 \right) \right] \right. \\ \left. + S_6 \left(\frac{r}{a} \right)^3 \sin f + S_7 \left(\frac{r}{a} \right)^3 \sin 3f \right\} \quad (4.617)$$

$$\xi_s^{(2)}(t, \beta_3) = \cos \tilde{\omega} \left[e_s^{(2)}(t, \beta_3) \right] - \sin \tilde{\omega} \left[e \tilde{\omega}_s^{(2)}(t, \beta_3) \right], \quad (4.618)$$

$$\eta_s^{(2)}(t, \beta_3) = \sin \tilde{\omega} \left[e_s^{(2)}(t, \beta_3) \right] + \cos \tilde{\omega} \left[e \tilde{\omega}_s^{(2)}(t, \beta_3) \right], \quad (4.619)$$

$$h_s^{(2)}(t, \beta_3) = \frac{1}{2} \cos(i/2) \cos \Omega \left[i_s^{(2)}(t, \beta_3) \right] - \frac{1}{2 \cos(i/2)} \sin \Omega \left[\sin i \Omega_s^{(2)}(t, \beta_3) \right], \quad (4.620)$$

$$k_s^{(2)}(t, \beta_3) = \frac{1}{2} \cos(i/2) \sin \Omega \left[i_s^{(2)}(t, \beta_3) \right] + \frac{1}{2 \cos(i/2)} \cos \Omega \left[\sin i \Omega_s^{(2)}(t, \beta_3) \right], \quad (4.621)$$

$$\lambda_s^{(2)}(t, \beta_3) = 3 \left(\frac{3}{2} \beta_3 a^4 \right) [S_4(-3 \sin M) + S_5(-\sin 3M) + S_6(3 \cos M) + S_7(\cos 3M)], \quad (4.622)$$

where

$$e_s^{(2)}(t, \beta_3) = \left(\frac{3}{2} \beta_3 a^4 \right) \left[S_4 \left(-\frac{1}{4} \cos 2M \right) + S_5 \left(\frac{9}{4} \cos 2M + \frac{3}{8} \cos 4M \right) \right. \\ \left. + S_6 \left(-\frac{1}{4} \sin 2M \right) + S_7 \left(\frac{9}{4} \sin 2M + \frac{3}{8} \sin 4M \right) \right], \quad (4.623)$$

$$i_s^{(2)}(t, \beta_3) = \left(\frac{3}{2} \beta_3 a^4 \right) \left(-I_{S4} \sin M - \frac{1}{3} I_{S5} \sin 3M + I_{S6} \cos M + \frac{1}{3} I_{S7} \cos 3M \right), \quad (4.624)$$

$$\sin i \Omega_s^{(2)}(t, \beta_3) = \left(\frac{3}{2} \beta_3 a^4 \right) \left[G_{13}(\sin M) + G_{14} \left(\frac{1}{3} \sin 3M \right) \right. \\ \left. + G_{17}(-\cos M) + G_{18} \left(-\frac{1}{3} \cos 3M \right) \right], \quad (4.625)$$

$$e \tilde{\omega}_s^{(2)}(t, \beta_3) = \left(\frac{3}{2} \beta_3 a^4 \right) \left[S_4 \left(-\frac{1}{4} \sin 2M \right) + S_5 \left(-\frac{9}{4} \sin 2M + \frac{3}{8} \sin 4M \right) \right. \\ \left. + S_6 \left(\frac{1}{4} \cos 2M \right) + S_7 \left(\frac{9}{4} \cos 2M - \frac{3}{8} \cos 4M \right) \right], \quad (4.626)$$

The above functions are expanded into series of trigonometric functions of the mean anomaly. The auxiliaries $G_{15}, G_{16}, \dots; I_{S6}, I_{S7}, \dots$ are:

$$\begin{cases} G_{15} = \left[\frac{5}{2} (AA_2 + BB_2) \right] B + \left[\frac{5}{4} (A^2 + B^2) - 1 \right] B_2 \\ G_{16} = \left[\frac{5}{6} (3AA_2 - BB_2) \right] B + \left[\frac{5}{12} (3A^2 - B^2) \right] B_2 \\ G_{17} = \left[\frac{5}{2} (AA_1 + BB_1) \right] B + \left[\frac{5}{4} (A^2 + B^2) - 1 \right] B_1 \\ G_{18} = \left[\frac{5}{6} (3AA_1 - BB_1) \right] B + \left[\frac{5}{12} (3A^2 - B^2) \right] B_1 \end{cases} \quad (4.627)$$

$$\begin{aligned} I_{S6} = \frac{1}{\sin i} & \left\{ \left(\frac{5}{2} \right) [A_0(A_2 \Delta B) + B_0(AA_2 + B_2(B + \Delta B)) + A_{20}(\Delta A \Delta B) + B_{20}(\Delta B^2)] \right. \\ & + [B_{20} + \cos i A_0] \left[\frac{5}{4} (A^2 + B^2) - 1 \right] \Big\} \\ & + \frac{\sin i}{(1 + \cos i)} \left[\left(\frac{5}{4} \right) (A^2 + B^2) - 1 \right] (-\Delta A) \end{aligned} \quad (4.628)$$

$$\begin{aligned} I_{S7} = \left(\frac{5}{12} \right) \frac{1}{\sin i} & \left\{ A_0 [6\Delta B^2 - 3(A + \Delta A)\Delta A] + B_0 [6AA_2 - 2BB_2 + (B + 3\Delta B)\Delta A] \right. \\ & + A_{20} [6A\Delta B] + B_{20} [3A^2 - B^2 - 2B\Delta B] \\ & + 3 \cos i [A_0 (A^2 - 3B^2 + (A + \Delta A)\Delta A) - 3B_0 (B + \Delta B)\Delta A] \Big\} \\ & + \left(\frac{5}{4} \right) \frac{\sin i}{(1 + \cos i)} \left[(3\Delta B^2 - \Delta A^2) \Delta A \right] \end{aligned} \quad (4.629)$$

4.7.2.5 The Perturbation Solution Due to the Part of the Perturbation Functions of the Sun or the Moon Related to $P_4(\cos \psi)$

In the perturbation functions of the Sun and the Moon, the part of R_4 is given by (4.504), (4.512), and (4.513). Here we only give the solution of the short-period term $a_s^{(2)}(t, \beta_4)$ that

$$\begin{aligned} a_s^{(2)}(t, \beta_4) &= 2a^2 R_{4s} \\ &= \frac{15}{4} \beta_4 a^6 \left\{ S_8 \left[\left(\frac{r}{a} \right)^4 - \left(1 + 5e^2 + \frac{15}{8} e^4 \right) \right] + S_9 \left[\left(\frac{r}{a} \right)^4 \cos 2f - \left(\frac{21}{4} e^2 + \frac{21}{8} e^4 \right) \right] \right. \\ &\quad \left. + S_{10} \left[\left(\frac{r}{a} \right)^4 \cos 4f - \left(\frac{63}{8} e^4 \right) \right] + S_{11} \left(\frac{r}{a} \right)^4 \sin 2f + S_{12} \left(\frac{r}{a} \right)^4 \sin 4f \right\} \end{aligned} \quad (4.630)$$

4.8 The Perturbed Orbit Solution Due to Earth's Deformation

When we discuss the perturbation by Earth's non-spherical gravity potential in the above chapters Earth is provided by an averaged model or is assumed to have a rigidity shape, whose total gravity potential is given by

$$V(\vec{R}) = \frac{GE}{R} \left\{ 1 + \sum_{l=1}^{\infty} \sum_{m=0}^l \left(\frac{a_e}{R} \right)^l P_{l,m}(\sin \varphi) [C_{l,m} \cos m\lambda + S_{l,m} \sin m\lambda] \right\}, \quad (4.631)$$

where the constants E and a_e are Earth's total mass and the equatorial radius of the referring ellipsoid, respectively; $C_{l,m}$ and $S_{l,m}$ are coefficients related to Earth's irregular shape (comparing to a sphere) and uneven mass distribution. Using the normalized units Earth's potential has the following form,

$$V(\vec{R}) = \frac{1}{R} \left\{ 1 + \sum_{l=1}^{\infty} \sum_{m=0}^l \left(\frac{1}{R} \right)^l \bar{P}_{l,m}(\sin \varphi) [\bar{C}_{l,m} \cos m\lambda + \bar{S}_{l,m} \sin m\lambda] \right\}. \quad (4.632)$$

In general models of Earth's gravity potential, the coefficients related to Earth's shape and mass distribution, $C_{l,m}$ and $S_{l,m}$ (or $\bar{C}_{l,m}$, $\bar{S}_{l,m}$) are regarded as defined constants.

In reality, all big celestial bodies are elastic bodies. On Earth, the surface covered by land is not rigid, and the other part, about 71% of the total surface, is covered by oceans, therefore deformation is a normal phenomenon. The reasons causing the deformation of a celestial body are complicated, but basically can be attributed to two types, the tidal deformation due to external gravity forces from other celestial bodies, and the rotational deformation due to the uneven rotation of the celestial body itself. In the case of Earth, the tidal deformation includes three types, the solid tide, the ocean tide, and the atmospheric tide, although the atmospheric tide is mainly caused by heat sources. All the deformations may affect the geographic position of an observatory, from the dynamical point of view, they can affect the above-mentioned coefficients $C_{l,m}$ and $S_{l,m}$ in Earth's gravity potential expansion. Therefore, the true Earth's gravity potential is composed of the above-given form in (4.632), which is its "average" (V), and an additional deformation potential (ΔV). For Earth based on the causes of deformations, there are four types of the additional potentials due to the solid tide, the ocean tide, the atmospheric tide, and the rotational deformation, denoted by ΔV_{ST} , ΔV_{OT} , ΔV_{AT} , and ΔV_{RT} , respectively. The influence of the additional deformation potentials is called the deformation perturbation of the central body.

The effect of Earth's deformation (including tidal and rotational) on the Earth's satellite orbital variation is treated as a synthesis tidal perturbation (mainly the solid tide) for general accuracy requirements. The details of ocean tidal modeling and research on ocean tidal models using Earth's satellites (such as ocean satellites)

belong to an independent field and are beyond the scope of this book. If necessary, readers may refer to Chap. 8 of reference [5].

4.8.1 Expression of the Additional Potential of Tidal Deformation

For Earth's tidal deformation the effect on the satellite orbit is given by the uniformed form—the synthesis tide. The corresponding additional deformation potential ΔV is given by [5, 7]

$$\Delta V = \frac{GE}{r} \left(\frac{m'}{E} \right) \sum_{l=2}^{\infty} k_l \left(\frac{a_e}{r} \right)^l \left(\frac{a_e}{r'} \right)^{l+1} P_l(\cos \psi^*), \quad (4.633)$$

where m' is the mass of the Sun or the Moon; r and r' are the radii of a particle, here is a satellite, and the Sun or the Moon from Earth's center, respectively; ψ^* is the angle between the direction of the satellite and the direction of the tidal crest.

The cause of the additional tidal potential is the gravitational forces of the Sun and the Moon. These forces make Earth's elastic deformations both in shape and in mass distribution. The deformations then change Earth's external gravitational field and the change in the potential is given by (4.633) according to the tidal model. The coefficient k_l in (4.633) is called the l th Love number, introduced by English geophysicist A. E. Love in 1909. The Love number is a kind of parameter for adjusting the deformation potential ΔV by the tidal model (i.e., the theory of hydrostatic equilibrium) and the actual Earth's deformation potential. According to early measurements by the Geodetic Earth-Orbiting satellite 1 and 2 (GEOS-1 and GEOS-2) the value of k_2 is about 0.30 for the synthesis tide.

The additional tidal potential can be expressed by the changes of the coefficients $C_{l,m}$ and $S_{l,m}$ in Earth's gravity potential through spherical harmonic expansion. In (4.633), $P_l(\cos \psi^*)$ can be presented as

$$\begin{aligned} P_l(\cos \psi^*) &= \sum_{m=0}^l 2^{\delta_m} \frac{(l-m)!}{(l+m)!} P_{l,m}(\sin \varphi) P_{l,m}(\sin \varphi') \\ &\quad \times [\cos m\lambda \cos m\lambda^* + \sin m\lambda \sin m\lambda^*] \\ &= \sum_{m=0}^l 2^{\delta_m} \frac{(l-m)!}{(l+m)!} P_{l,m}(\sin \varphi) P_{l,m}(\sin \varphi') \cos m(\lambda - \lambda^*) \end{aligned} \quad (4.634)$$

where δ_m is

$$\delta_m = \begin{cases} 0, & m = 0, \\ 1, & m \neq 0, \end{cases} \quad (4.635)$$

and λ^* in (4.634) is defined as

$$\lambda^* = \lambda' + \nu. \quad (4.636)$$

The angle ν is the angle between the direction of the tidal crest and the direction of the Sun or the Moon. The tidal crest is “lagged” by the Sun or the Moon, and ν is called the angle of lag. The position of a satellite in the Earth-fixed coordinate system is given by longitude λ and latitude φ , and the position of the Sun or the Moon is given by longitude λ' and latitude φ' . In (4.634) λ and λ' are appeared in the form $(\lambda - \lambda')$, if we do not consider Earth’s precession, nutation, and polar motion, (in the tidal perturbation problem, it is not necessary to include these factors), then for λ , φ , λ' , and φ' , there is no difference between the geocentric celestial frame and the Earth-fixed frame. But for rigorous formulas, these coordinates are for the Earth-fixed frame. Substituting (4.634) into (4.633) leads

$$\begin{aligned} \Delta V = & \frac{GE}{r} \left(\frac{m'}{E} \right) \sum_{l=2}^{\infty} \sum_{m=0}^l 2^{\delta_m} \frac{(l-m)!}{(l+m)!} \left(\frac{a_e}{r} \right)^l P_{l,m}(\sin \varphi) \\ & \times k_l \left(\frac{a_e}{r'} \right)^{l+1} P_{l,m}(\sin \varphi') [\cos m\lambda^* \cos m\lambda + \sin m\lambda^* \sin m\lambda], \end{aligned} \quad (4.637)$$

and ΔV can be further expressed by the changes of the coefficients of Earth’s gravity potential that

$$\Delta V = \frac{GE}{r} \sum_{l=2}^{\infty} \sum_{m=0}^l \left(\frac{a_e}{r} \right)^{l+1} P_{l,m}(\sin \varphi) [\Delta C_{l,m} \cos m\lambda + \Delta S_{l,m} \sin m\lambda], \quad (4.638)$$

$$\Delta C_{l,m} = k_l \left(\frac{m'}{E} \right) \left(\frac{a_e}{r'} \right)^{l+1} \left[2^{\delta_m} \frac{(l-m)!}{(l+m)!} \right] P_{l,m}(\sin \varphi') \cos m\lambda^*, \quad (4.639)$$

$$\Delta S_{l,m} = k_l \left(\frac{m'}{E} \right) \left(\frac{a_e}{r'} \right)^{l+1} \left[2^{\delta_m} \frac{(l-m)!}{(l+m)!} \right] P_{l,m}(\sin \varphi') \sin m\lambda^*. \quad (4.640)$$

4.8.2 *Effect of the Main Term in the Additional Tidal Deformation Potential (the Second-Order Term of $l = 2$) on a Satellite Orbit*

The potential of $l = 2$ using the normalized units has three parts which are

$$\Delta V_{2,0} = [\beta k_2 P_2(\sin \varphi')] \left(\frac{1}{r} \right)^3 P_2(\sin \varphi), \quad (4.641)$$

$$\Delta V_{2,1} = \left[\frac{1}{3} \beta k_2 P_{2,1}(\sin \varphi') \right] \left(\frac{1}{r} \right)^3 P_{2,1}(\sin \varphi) \cos(\lambda - \lambda'), \quad (4.642)$$

$$\Delta V_{2,2} = \left[\frac{1}{3} \beta k_2 P_{2,2}(\sin \varphi') \right] \left(\frac{1}{r} \right)^3 P_{2,2}(\sin \varphi) \cos 2(\lambda - \lambda'), \quad (4.643)$$

where β is the parameter β_2 in the perturbation function of the Sun or the Moon's gravity potential $P_2(\cos \psi)$ in (4.496), that

$$\beta = \beta_2 = m' / r'^3, \quad (4.644)$$

and m' is the mass of the Sun or the Moon in the unit of Earth's mass, r' is the distance of the Sun or the Moon to Earth's center, for tidal deformation perturbation r' can be treated as a constant.

For a low Earth orbit satellite assuming the magnitude of $\Delta V_{2,m}$ ($m = 0, 1, 2$) perturbation to be the small parameter ε given by the ratio of the perturbing acceleration $F_\varepsilon = |\text{grad} \Delta V_{2,m}|$ to Earth's central gravity acceleration $F_0 = 1/r^2$ that

$$\varepsilon = \frac{F_\varepsilon}{F_0} = O\left(k_2 m' \left(\frac{r}{r'}\right)^3 / r^5\right) = 10^{-8}. \quad (4.645)$$

For a high Earth orbit satellite as r increases the magnitude of the perturbation decreases, and the effect of Earth's deformation decreases too. Based on this estimation the tidal deformation perturbation ($l = 2$) is usually a third-order perturbation. Although for general accuracy requirements this perturbation can be ignored, we still provide the main part of the perturbation solution.

(1) Effect of $\Delta V_{2,0}$ on the perturbation solution of a satellite orbit

Similar to the perturbation of Earth's non-spherical gravity potential $C_{2,0}$ term that from

$$\begin{cases} P_2(\sin \varphi) = \frac{3}{2} \sin^2 \varphi - \frac{1}{2} \\ \sin \varphi = \sin i \sin u, \quad u = \sin(f + \omega) \end{cases}$$

$\Delta V_{2,0}$ in (4.641) can be written as

$$\begin{aligned} \Delta V_{2,0} = \frac{\beta k_2}{4a^3} \left(\frac{a}{r} \right)^3 & \left\{ \left(1 - \frac{3}{2} \sin^2 i' \right) \left[\left(1 - \frac{3}{2} \sin^2 i \right) + \frac{3}{2} \sin^2 i \cos 2u \right] \right. \\ & \left. + \frac{3}{2} \sin^2 i' \left[\left(1 - \frac{3}{2} \sin^2 i \right) + \frac{3}{2} \sin^2 i \cos 2u \right] \cos 2u' \right\} \quad (4.646) \end{aligned}$$

where i' and $u' = f' + \omega'$ are orbit elements of the Sun or the Moon in the epoch geocentric celestial coordinate system. Obviously, u' is a slow variable. By the average values of

$$\overline{\left(\frac{a}{r}\right)^3} = (1 - e^2)^{-3/2}, \quad \overline{\left(\frac{a}{r}\right)^3 \cos 2f} = 0, \quad \overline{\left(\frac{a}{r}\right)^3 \sin 2f} = 0,$$

$\Delta V_{2,0}$ can be decomposed into three parts of secular, long-period, and short-period terms, which are

$$(\Delta V_{2,0})_c = \frac{\beta k_2}{4a^3} \left(1 - \frac{3}{2} \sin^2 i'\right) \left(1 - \frac{3}{2} \sin^2 i\right) (1 - e^2)^{-3/2} \quad (4.647)$$

$$(\Delta V_{2,0})_l = \frac{\beta k_2}{4a^3} \left(\frac{3}{2} \sin^2 i'\right) \left(1 - \frac{3}{2} \sin^2 i\right) (1 - e^2)^{-3/2} \cos 2u' \quad (4.648)$$

$$\begin{aligned} (\Delta V_{2,0})_s = & \frac{\beta k_2}{4a^3} \left(\frac{a}{r}\right)^3 \left[\left(1 - \frac{3}{2} \sin^2 i'\right) + \frac{3}{2} \sin^2 i \cos 2u' \right] \\ & \times \left\{ \left(1 - \frac{3}{2} \sin^2 i\right) \left[\left(\frac{a}{r}\right)^3 - (1 - e^2)^{-3/2} \right] + \frac{3}{2} \sin^2 i \cos 2u \right\} \end{aligned} \quad (4.649)$$

The effect of the perturbation by $(\Delta V_{2,0})_c$ and $(\Delta V_{2,0})_l$ on satellite's orbital motion is secular in the third order (similar to the case of $C_{2,0}$ term in Earth's non-spherical potential). Expressed in the quasi-mean orbital elements the results of the secular and long-period terms are $\Delta\sigma(t) = \sigma_3(t - t_0) + \Delta\sigma_l(t, \beta)$ as

$$\Delta a(t) = 0, \quad (4.650)$$

$$\Delta e(t) = 0, \quad (4.651)$$

$$\Delta i(t) = 0, \quad (4.652)$$

$$\Delta\Omega(t) = -\left(\frac{3}{4p^2} \beta k_2\right) n \cos i W(t), \quad (4.653)$$

$$\Delta\omega(t) = -\left(\frac{3}{4p^2} \beta k_2\right) n \left(2 - \frac{5}{2} \sin^2 i\right) W(t), \quad (4.654)$$

$$\Delta M(t) = -\left(\frac{3}{4p^2} \beta k_2\right) n \left(1 - \frac{3}{2} \sin^2 i\right) \sqrt{1 - e^2} W(t), \quad (4.655)$$

where $p = a(1 - e^2)$ and $W(t)$ is given by

$$W(t) = \left(1 - \frac{3}{2} \sin^2 i'\right) (t - t_0) + \left(\frac{3}{2} \sin^2 i'\right) (\sin 2u' - \sin 2u'_0) / 2n', \quad (4.656)$$

and n' is the variation rate of u' . For a high Earth orbit (such as GEO) satellite, the long-period terms treated as secular terms actually have a period related to the period of the Sun (semi-annual) or the Moon (semi-monthly).

The magnitude of the short-period term $\sigma_s^{(3)}(t)$ is small, we do not provide the formulas here.

The results show that the $\Delta V_{2,0}$ term only has short-period effects on a , e , and i , whereas the secular effects on Ω , ω , and M are not easy to be separated from Earth's non-spherical potential J_2 term. Based on the situation it is unnecessary to give non-singularity solutions of the first type and the second type. If necessary, it is easy to transfer the solution of Kepler orbital elements to the non-singularity solution by the same method as for other perturbations (such as the third-body perturbation), because the perturbation function $\Delta V_{2,0}$ does not need any specific treatment, and the solution does not have any small divisor, see (4.650)–(4.656).

(2) Basic characteristics of $\Delta V_{2,1}$ and $\Delta V_{2,2}$ perturbation on satellite's orbit

The expressions of additional tidal deformation potentials, $\Delta V_{2,1}$ and $\Delta V_{2,2}$ (4.642) and (4.643), show that the effects of these two terms on the satellite's orbit are similar to these by Earth's non-spherical gravity potential, $C_{2,1}$, $S_{2,1}$, and $C_{2,2}$, $S_{2,2}$ terms, and have only periodic effects, including the short-periodic effect (depending on the fast changing variable M) and the long-periodic effect by long-periodic terms such as $\cos(\Omega - \Omega^*)$, $\sin(\Omega - \Omega^*)$, $\cos 2(\Omega - \Omega^*)$, $\sin 2(\Omega - \Omega^*)$, $\dots$, where $\Omega^* = \Omega' + \nu$, ν is the angle of tidal lag, Ω' is the longitude of the orbital ascending node of the Sun or the Moon, for the Sun $\Omega' = 0$. Also, there is another property of tidal perturbation, that the longitude term related to Earth's rotation appears in the form $(\lambda - \lambda')$ (see (4.462) and (4.643)), thus there is no term related to Earth's rotation (it is very different comparing to $J_{2,2}$ perturbation term).

4.9 Post-Newtonian Effect on the Orbital Motion

4.9.1 The Post-Newtonian Effect

In the field of orbital dynamics, the correction of the new theory of gravitation involves the refining of the reference system, measurement treatment, and motion equation of the circling celestial body. In this section, we discuss the effect on the motion equation from the perspective of "method". Because of the very small correction, the influence can be treated as a perturbing source. Usually, the correction term of the motion equation is called the post-Newtonian acceleration, the resulting orbital change is the post-Newtonian effect [3].

For the orbital motion of an Earth's satellite the correction is introduced by Einstein's theory of general relativity to Newton's theory of gravity. For the orbital motion, Earth's central gravity potential is refined by the post-Newtonian acceleration. For the current accuracy requirement, the motion equation of a satellite can be given in the following form

$$\begin{cases} \ddot{\vec{r}} = -\frac{\mu}{r^2} \left(\frac{\vec{r}}{r} \right) + \vec{A}_{PN}, \\ \vec{A}_{PN} = \vec{A}_1 + \vec{A}_2 + \vec{A}_3 + \vec{A}_4. \end{cases} \quad (4.657)$$

The first term on the right side of the equation is Earth's central gravitational acceleration, and the second term $\vec{A}_{PN}$ is the post-Newtonian acceleration which includes $\vec{A}_1$, $\vec{A}_2$, $\vec{A}_3$, and $\vec{A}_4$ due to the Schwarzschild solution, the geodetic precession, the rotation effect, and Earth's dynamical form-factor J_2 effect, given by [3, 9–11]

$$\vec{A}_1 = \frac{\mu}{c^2 r^2} \left[\left(4\frac{\mu}{r} - v^2 \right) \left(\frac{\vec{r}}{r} \right) + 4\dot{r}v \left(\frac{\dot{\vec{r}}}{v} \right) \right], \quad (4.658)$$

$$\vec{A}_2 = 2 \left(\vec{\Omega} \times \vec{v} \right), \quad (4.659)$$

$$\vec{A}_3 = \frac{2\mu}{c^2 r^3} \left[\frac{3}{r^2} \left(\vec{r} \cdot \vec{J} \right) \left(\vec{r} \times \vec{v} \right) + \left(\vec{v} \times \vec{J} \right) \right], \quad (4.660)$$

$$\vec{A}_4 = \frac{1}{c^2} \left[-4\nabla \left(\frac{\mu}{r} R \right) + v^2 \nabla R - 4(\vec{v} \cdot \nabla R) \vec{v} \right], \quad (4.661)$$

The auxiliaries in the four expressions are

$$\vec{\Omega} = \frac{3}{2} \left(\frac{Gm_s}{c^2 r_s^3} \right) \vec{h}_s, \quad \vec{h}_s = \vec{r}_s \times \vec{v}_s, \quad (4.662)$$

$$\vec{J} = J\hat{k}, \quad J = 9.8 \times 10^8 \text{ m}^2/\text{s}, \quad (4.663)$$

$$R = -J_2 \left(\frac{\mu a_e^2}{r^3} \right) \left(\frac{3}{2} \sin^2 \varphi - \frac{1}{2} \right), \quad \sin \varphi = \frac{z}{r}, \quad (4.664)$$

where m_s is the mass of the Sun; r_s and v_s are the radius and velocity of the Sun with respect to Earth's center; $\vec{h}_s$ is a constant vector which is the Sun's area velocity along the fixed elliptic orbit "around Earth"; $\hat{k}$ is the unit vector in the normal direction of Earth's equator plane; R is Earth's non-spherical dynamical form-factor potential; $\mu = GE$, E is Earth's mass; and c is the speed of light. In the calculation, we use the normalized dimensionless units.

For a low Earth orbit satellite, the relative magnitudes of the four perturbations compared to Earth's central gravitational acceleration are about

$$10^{-9}, 10^{-11}, 10^{-12}, 10^{-12}.$$

Based on this estimation, usually only the first term needs to be considered.

4.9.2 Perturbation Solution Due to the Post-Newtonian Effect

Generally, only the Schwarzschild solution is needed, then the post-Newtonian acceleration is given by (4.658) i.e.,

$$\vec{A}_{PN} = \frac{\mu}{c^2 r^2} \left[\left(\frac{4\mu}{r} - v^2 \right) \left(\frac{\vec{r}}{r} \right) + 4\dot{r}v \left(\frac{\dot{\vec{r}}}{v} \right) \right], \quad (4.665)$$

and the small parameter is defined by

$$\varepsilon = \left| \vec{A}_{PN} \right| / \left(\frac{\mu}{r^2} \right) = O\left(\frac{v^2}{c^2} \right), \quad (4.666)$$

In the Solar System, the magnitude of ε for Mercury's motion around the Sun is 10^{-7} , and for a satellite around Earth is 10^{-9} . We can say that the post-Newtonian effect for the Mercury's motion is not too small, but for an Earth's satellite (exactly the Low Earth Orbit satellite), it is only a third-order perturbation compared to Earth's non-spherical dynamical form-factor J_2 term.

The above expression of the post-Newtonian acceleration can be applied to both dynamical systems for a planet moving around the Sun and an artificial satellite moving around Earth. For a planet $\mu = G(m_1 + m_2)$, where m_1 and m_2 are the masses of the Sun and the planet, respectively; for an Earth's satellite $\mu = GE$, where E is Earth's mass. Therefore, these two systems can be treated by the same method as long as we understand that the prime planes of the two frames, i.e., the xy -plane, are different. For a planet of the Sun, the xy -plane is the ecliptic plane and for an Earth's satellite it is Earth's equator plane, and the two orbit elements i and Ω are defined according to each frame.

The coordinates in the radial, transverse, and normal directions of $\vec{r}, \dot{\vec{r}}$ can be given by the orbital elements that

$$\vec{r} = \begin{pmatrix} r \\ 0 \\ 0 \end{pmatrix}, \quad \dot{\vec{r}} = \begin{pmatrix} \dot{r} \\ r\dot{\theta} \\ 0 \end{pmatrix} = \begin{pmatrix} e \sin f \sqrt{\mu/p} \\ (1 + e \cos f) \sqrt{\mu/p} \\ 0 \end{pmatrix}, \quad (4.667)$$

where $p = a(1 - e^2)$. We then derive the three components of $\vec{A}_{PN}$, S , T , and W as

$$\begin{cases} S = \frac{\mu^2}{c^2 a^3} \left[-3 \left(\frac{a}{r} \right)^2 + 10 \left(\frac{a}{r} \right)^3 - 4(1 - e^2) \left(\frac{a}{r} \right)^4 \right], \\ T = \frac{\mu^2}{c^2 a^3} \left[4e \left(\frac{a}{r} \right)^3 \sin f \right], \\ W = 0. \end{cases} \quad (4.668)$$

Substituting (4.668) into the perturbation equations in Kepler elements (3.14) we have

$$\dot{\sigma} = f_0(a) + f_{PN}(\sigma, \varepsilon). \quad (4.669)$$

The post-Newtonian term f_{PN} has only a secular part and a short-period part that

$$f_{PN} = f_c(a, e) + f_s(a, e, M). \quad (4.670)$$

After integrating we obtain the secular terms and the short-period terms of the perturbation solution that

$$\begin{cases} a_c(t - t_0) = 0, \\ e_c(t - t_0) = 0, \\ i_c(t - t_0) = 0, \end{cases} \quad (4.671)$$

$$\Omega_c(t - t_0) = 0, \quad (4.672)$$

$$\omega_c(t - t_0) = \left(\frac{\mu}{c^2}\right)^{\frac{3}{p}} n(t - t_0), \quad (4.673)$$

$$M_c(t - t_0) = -\left(\frac{\mu}{c^2}\right)(1 - e^2)^{-1} \left(3 + \frac{7}{2}e^2 + e^4\right)n(t - t_0), \quad (4.674)$$

$$a_s(t) = -\left(\frac{\mu}{c^2}\right)(1 - e^2)^{-2} [(14 + 6e^2)e \cos f + 5e^2 \cos 2f], \quad (4.675)$$

$$e_s(t) = -\left(\frac{\mu}{c^2}\right)^{\frac{1}{p}} [(3 + 7e^2) \cos f + \frac{5}{2}e \cos 2f], \quad (4.676)$$

$$i_s(t) = 0, \quad (4.677)$$

$$\Omega_s(t) = 0, \quad (4.678)$$

$$\omega_s(t) = \left(\frac{\mu}{c^2}\right)^{\frac{1}{p}} [3(f - M) - \left(\frac{3}{e} - e\right) \sin f - \frac{5}{2} \sin 2f], \quad (4.679)$$

$$M_s(t) = -\left(\frac{\mu}{c^2}\right)^{\frac{1}{p}} \sqrt{1 - e^2} [3e\left(\frac{r}{a}\right) \sin f - \left(\frac{3}{e} + 7e\right) \sin f - \frac{5}{2} \sin 2f]. \quad (4.680)$$

The characteristics of the post-Newtonian perturbation solution are as follows.

- (1) Besides the motion element M (the mean anomaly), only the argument periastris ω has a secular variation. This is exactly one of the evidences of the theory of general relativity, the test of the anomalous perihelion advance of Mercury. This anomalous phenomenon cannot be explained by Newton's theory of gravitation but can be corrected by the post-Newtonian effect.
- (2) The post-Newtonian effect by the Schwarzschild solution has no influence on the orbital plane, because there are no changes in the two elements i and Ω in the perturbation solution.

- (3) There is no long-periodic variation due to the perturbation of the post-Newtonian effect.

4.9.3 Other Post-Newtonian Effects on the Earth's Artificial Satellite Motion

Besides the effect of the Schwarzschild solution, there are the geodetic precession effect, the Earth's rotation effect, and the Earth's flattening effect. These three effects, particularly the Earth's rotation and flattening effects, have a secular influence on the longitude of the right ascension node, Ω , but these effects are too small to be considered.

4.10 Perturbed Orbit Solution Due to the Solar Radiation Pressure

4.10.1 Calculation of Radiation Pressure

The solar radiation pressure acts on all celestial bodies in the Solar System, no matter big or small. The influence of the solar radiation pressure on the motion of natural celestial bodies (planets and natural satellites) did not draw attention until after the launching of Earth's artificial satellite. The reason is that the radiation pressure acts as a surface force, its intensity is related to the efficient area-to-mass ratio of a celestial body that receives the radiation. Because the efficient area-to-mass ratio of an artificial satellite is much larger than that of a natural celestial body, the effect of the radiation pressure on the motion of an artificial satellite becomes significant.

The perturbation due to this kind of surface force has a totally different mechanism and a different form from the above discussed gravitational force. It is necessary to introduce the basic principles and basic calculation formulas of the radiation pressure.

For a moving body, no matter what kind of shape, attitude, and physical properties of surface material it has, we can start with an "infinitesimal" surface element ds as a planar surface to study (Fig. 4.3). The radiative pressure force acting on it is a composition of two forces, that

$$d\vec{F} = d\vec{F}_1 + d\vec{F}_2, \quad (4.681)$$

where $d\vec{F}_1$ and $d\vec{F}_2$ are the radiation pressure incident force and the reflect force, respectively, exerted on ds , and are given by

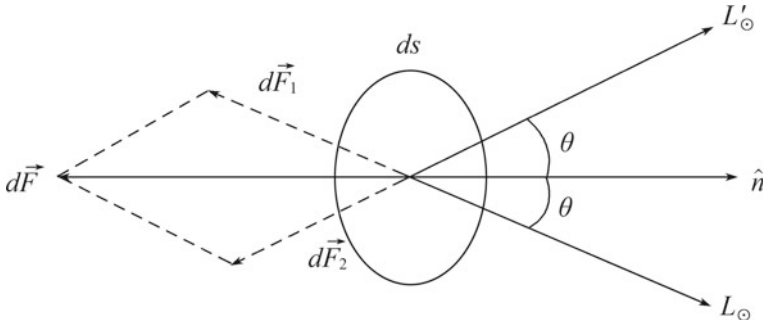


Fig. 4.3 The radiation pressure composition force acts on an infinitesimal surface element ds

$$\begin{cases} d\vec{F}_1 = -\rho_{\odot}(\hat{n} \cdot \hat{L}_{\odot})ds \hat{L}_{\odot}, \\ d\vec{F}_2 = -\eta |d\vec{F}_1| \hat{L}'_{\odot}. \end{cases} \quad (4.682)$$

In (4.682), $\rho_{\odot}$ is the intensity of the radiation pressure on ds ; $\hat{n}$ is the unit vector of the normal of ds ; $\hat{L}_{\odot}$ is the unit vector in the direction of the radiation source, and θ is the angle between $\hat{n}$ and $\hat{L}_{\odot}$; the direction of $d\vec{F}_2$ is given according to the law of reflection (Fig. 4.3), its magnitude depends on η , the coefficient of reflection of ds , that $\eta = 1$ means total reflection whereas $\eta = 0$ total absorption, usually $0 < \eta < 1$. From (4.681) and (4.682) we have the net force of the radiation pressure:

$$d\vec{F} = -\rho_{\odot}ds \cos \theta (\hat{L}_{\odot} + \eta \hat{L}'_{\odot}), \quad (4.683)$$

where $ds \cos \theta$ is the projection of ds in the direction perpendicular to the direction of the source.

In order to calculate $d\vec{F}$, the unit vector $\hat{L}'_{\odot}$ needs to be expressed in the directions of $\hat{L}_{\odot}$ and $\hat{n}$, which are known vectors. For that we introduce $\hat{n}'$ and $\hat{\tau}$ defined by

$$\begin{cases} \hat{n}' = \frac{\hat{L}_{\odot} \times \hat{n}}{\sin \theta}, \\ \hat{\tau} = \hat{n} \times \hat{n}' = \frac{1}{\sin \theta} \hat{n} \times (\hat{L}_{\odot} \times \hat{n}), \end{cases} \quad (4.684)$$

Obviously, the vectors $\hat{n}'$ and $\hat{\tau}$ are on the plane ds and perpendicular to $\hat{n}$. The three vectors, $\hat{n}$, $\hat{n}'$, and $\hat{\tau}$, form a right-handed system. Using the vector cross product formula

$$\vec{a} \times (\vec{b} \times \vec{c}) = (\vec{a} \cdot \vec{c})\vec{b} - (\vec{a} \cdot \vec{b})\vec{c},$$

we obtain

$$\sin \theta \hat{\tau} = (\hat{n} \cdot \hat{n})\hat{L}_{\odot} - (\hat{n} \cdot \hat{L}_{\odot})\hat{n} = \hat{L}_{\odot} - \cos \theta \hat{n}. \quad (4.685)$$

Because $\hat{L}'_{\odot}$, $\hat{n}$, and $\hat{\tau}$ are in the same plane, there is

$$\begin{aligned}\hat{L}'_{\odot} &= (\hat{L}'_{\odot} \cdot \hat{n})\hat{n} + (\hat{L}'_{\odot} \cdot \hat{\tau})\hat{\tau} \\ &= \cos \theta \hat{n} + (\hat{L}'_{\odot} \cdot \hat{L}_{\odot} - \cos \theta \hat{L}'_{\odot} \cdot \hat{n}) \frac{\hat{\tau}}{\sin \theta} \\ &= \cos \theta \hat{n} - \sin \theta \hat{\tau} = 2 \cos \theta \hat{n} - \hat{L}_{\odot}\end{aligned}\quad (4.686)$$

Substituting (4.686) into (4.683) leads

$$d\vec{F} = -\rho_{\odot} ds \cos \theta \left[(1 - \eta) \hat{L}_{\odot} + 2\eta \cos \theta \hat{n} \right]. \quad (4.687)$$

The total radiation force exerted on the moving celestial body then is given by the following areal integration

$$\vec{F}_{\odot} = - \oint_{(\omega)} \rho_{\odot} ds \cos \theta \left[(1 - \eta) \hat{L}_{\odot} + 2\eta \cos \theta \hat{n} \right] ds. \quad (4.688)$$

The integration is on the entire surface (ω) where the radiation force exerts.

If the shape of a satellite is a plane (or an equivalent cross-section) whose area is S and the scale of the plane is much smaller compared to the distance between the satellite and the Sun, then (4.688) can be simplified to

$$\vec{F}_{\odot} = -\rho_{\odot} S \cos \theta \left[(1 - \eta) \hat{L}_{\odot} + 2\eta \cos \theta \hat{n} \right]. \quad (4.689)$$

Furthermore, if $\theta = 0$, which means the normal of the plane always points to the radiation source, then we have

$$\begin{cases} \vec{F}_{\odot} = -\kappa S \rho_{\odot} \hat{L}_{\odot}, \\ \kappa = 1 + \eta. \end{cases} \quad (4.690)$$

Generally, in the orbital perturbation calculation of an Earth's satellite, the accepted radiation pressure model is this simplified one, and the area S in (4.690) is the area of the equivalent cross-section. In reality, the radiation pressure is an extremely complicated problem, which involves the shape, the geometrical information of its attitude in space, the reflection property of the satellite surface material, etc. Only highly precise problems (such as orbital determinations and forecasts of high Earth orbit navigation satellites) need to consider these details. For general orbital determinations and forecasts, it is neither necessary nor possible to involve these factors. In this book, we adopt the simplified model (4.690) for calculating the radiation pressure force to construct the corresponding perturbation solution.

4.10.2 Two States of Radiation Pressure Perturbation

(1) The source of the radiation is the central body.

All celestial bodies and artificial spacecraft in the Solar System receive the radiation pressure from the Sun as the central body. For an Earth's satellite, the radiation force exerted on the satellite also comes from Earth including Earth's reflection radiation pressure (albedo) and the Earth's thermal radiation pressure; then the central body is Earth, and the radiation pressure force given by (4.690) is changed to

$$\vec{F}_{\odot} = \kappa S \rho_{\odot} \left(\frac{r_0^2}{r^2} \right) \left(\frac{\vec{r}}{r} \right), \quad (4.691)$$

where $\rho_{\odot}$ is the radiation pressure intensity at $r = r_0$ denoted by the same symbol $\rho_{\odot}$.

(2) The source of the radiation is a perturbing body.

The motion of an Earth's satellite belongs to this situation. The source of radiation is a perturbing body which is the Sun, then the radiation pressure force is given by

$$\begin{cases} \vec{F}_{\odot} = \kappa S \rho_{\odot} \left(\frac{\Delta_0^2}{\Delta^2} \right) \left(\frac{\vec{\Delta}}{\Delta} \right), \\ \vec{\Delta} = \vec{r} - \vec{r}', \end{cases} \quad (4.692)$$

where $\vec{r}'$ is the position vector of the perturbing body, and $\rho_{\odot}$ the radiation pressure intensity at $\Delta = \Delta_0$.

The above two states of radiation pressure force show that the radiation pressure force actually is a repulsive central force, whose direction is opposite to the gravity force. If we set

$$\mu_{\odot} = \begin{cases} \left(\kappa \frac{S}{m} \right) \rho_{\odot} r_0^2, & \text{source is the central body,} \\ \left(\kappa \frac{S}{m} \right) \rho_{\odot} \Delta_0^2, & \text{source is a perturbing body,} \end{cases} \quad (4.693)$$

then we have the perturbation acceleration $\vec{A}_{\odot}$ corresponding to (4.691) and (4.692) as

$$\vec{A}_{\odot} = \frac{\vec{F}_{\odot}}{m} = \begin{cases} \frac{\mu_{\odot}}{r^3} \vec{r}, & \text{source is the central body,} \\ \frac{\mu_{\odot}}{\Delta^3} \vec{\Delta}, & \text{source is a perturbing body,} \end{cases} \quad (4.694)$$

where, $\mu = G(m_0 + m)$, m_0 and m have the same definitions as before.

For the first state if we only consider the perturbation due to the radiation pressure then the perturbation function is given by

$$\ddot{\vec{r}} = -\frac{\mu}{r^3} \vec{r} + \frac{\mu_{\odot}}{r^3} \vec{r} = -(\mu - \mu_{\odot}) \frac{\vec{r}}{r^3}. \quad (4.695)$$

Comparing to the two-body problem the only difference is that the coefficient μ in the two-body problem now is replaced by $\mu - \mu_\odot$, and the original area integral $n^2 a^3 = \mu$ now becomes

$$n^2 a^3 = \mu - \mu_\odot. \quad (4.696)$$

We do not further discuss this ideal situation of the radiative pressure perturbation.

The radiation pressure force exerted on an Earth's satellite is mainly from the Sun corresponding to the second state given in (4.692). But in reality, things are more complicated, as the radiation pressure from Earth also affects the motion of the satellite. Earth's radiation includes Earth's albedo and thermal radiation which are almost impossible to separate, therefore, it is impossible to build an accurate model. On top of the radiation pressure forces from both the Sun and Earth, the shape and the attitude of a satellite further increase the difficulty.

Based on the situation in practice for an Earth's satellite, the method to deal with the radiation pressure is by a synthetic treatment of the radiation force to build a simplified empirical perturbation model, then to construct the perturbation solution for the radiation pressure.

4.10.3 The Perturbation Solution Due to Radiation Pressure

4.10.3.1 Perturbation Acceleration and Function Due to Radiation Pressure

For the orbital motion of an Earth's satellite the perturbation acceleration due to the radiation pressure $\vec{A}_\odot$ is given by the difference between two terms that

$$\vec{A}_\odot = \frac{\mu_\odot}{\Delta^3} \vec{\Delta} - \left(-\frac{\mu'_\odot}{r'^3} \vec{r}' \right), \quad (4.697)$$

where $\mu'_\odot = \left(\frac{\kappa' S'}{m_0} \right) \rho_\odot \Delta_0^2$, and κ' , S' are parameters for the central body (i.e., the radiation pressure exerted on the Earth). If a satellite is a sphere, whose diameter is 3 m and weight is 1 ton, then the area-to-mass ratio S/m is much larger than that of Earth S'/m_0 , that

$$\left(\frac{S}{m} \right) / \left(\frac{S'}{m_0} \right) \approx 10^9.$$

Therefore, the second term on the right side of (4.697) can be ignored, and the perturbation acceleration by the radiation pressure on the satellite is given by the first term, and the corresponding perturbation function R is given by

$$\begin{cases} R = -\frac{\mu_{\odot}}{\Delta^3}, \\ \mu_{\odot} = \left(\kappa \frac{S}{m}\right) \rho_{\odot} \Delta_0^2, \end{cases} \quad (4.698)$$

where κ is given in (4.690), $\rho_{\odot}$ is the intensity of the radiation pressure at $\Delta = \Delta_0$.

For dealing with the variation of the equivalent cross-section we assume the period of the variation of the equivalent cross-section to be the same as the orbital period of the satellite, then to use an approximate form for the perturbation function R as

$$\begin{cases} R = -\frac{k}{\Delta}, & k = \mu_{\odot}[1 + \alpha \cos(u - u_0)], \\ \mu_{\odot} = \left(\kappa \frac{S}{m}\right) \rho_{\odot} \Delta_0^2, & u = \omega + f, \end{cases} \quad (4.699)$$

where α and u_0 are two empirical parameters reflecting the periodic variation of the radiation pressure, that $\alpha < 1$, and $u_0 = \text{constant}$. In reality, reasonable corrections are needed to obtain an empirical model based on a satellite's basic state and specific requirements of a project. The complicity of the model does not reduce its practical application by modern computer utility, so we do not further discuss the details of treatments for the empirical model. The radiation pressure model given by the expression (4.699) is a reasonable model representing the characteristics of the radiation pressure from both qualitative and quantitative points of view. Therefore, perturbation results by this model are reliable for certain accuracy requirements. The perturbation solution based on this radiation pressure model is derived as follows.

In the epoch geocentric celestial frame (J2000) we have

$$\begin{cases} \frac{1}{\Delta} = (r^2 + r'^2 - 2rr' \cos \psi)^{-1/2} = \frac{1}{r'} \sum_{k=0}^{\infty} \left(\frac{r}{r'}\right)^k P_k(\cos \psi), \\ \cos \psi = \left(\frac{\vec{r}}{r}\right) \cdot \left(\frac{\vec{r}'}{r'}\right), \end{cases}$$

then the perturbation function R is expressed as

$$R = -\frac{\mu_{\odot}}{r'} \left[\frac{r}{r'} \cos \psi + \frac{r^2}{r'^2} \left(\frac{3}{2} \cos^2 \psi - \frac{1}{2} \right) + O\left(\frac{r^3}{r'^3}\right) \right]. \quad (4.700)$$

The first part on the right side corresponds to the perturbing acceleration $\frac{\mu_{\odot}}{r'^3} \vec{r}'$, comparing to (4.697) here $\vec{r}'$ replaces $\vec{\Delta}$, which means it does not include the solar parallax, whereas the second part is related to the solar parallax. The term including $\left(\frac{r}{r'}\right)^3$ is omitted, because its magnitude compared to the first part is $\left(\frac{r}{r'}\right)^2 = O(10^{-8})$ for a low Earth orbit satellite. The perturbation function of the radiation pressure now can be given by

$$R = -\frac{\mu_{\odot}}{r'} \left[\left(\frac{r}{r'}\right) \cos \psi + \left(\frac{r}{r'}\right)^2 \left(\frac{3}{2} \cos^2 \psi - \frac{1}{2} \right) \right]. \quad (4.701)$$

Customarily the three components of the perturbing acceleration $\vec{A}_{\odot}$ S , T , and W are given by

$$\begin{cases} S = \frac{\mu_{\odot}}{\Delta^3} (\vec{\Delta} \cdot \hat{r}) = -\mu_{\odot} \frac{r'}{\Delta^3} (A \cos f + B \sin f) + \mu_{\odot} \frac{r}{\Delta^3}, \\ T = \frac{\mu_{\odot}}{\Delta^3} (\vec{\Delta} \cdot \hat{\theta}) = -\mu_{\odot} \frac{r'}{\Delta^3} (-A \sin f + B \cos f), \\ W = \frac{\mu_{\odot}}{\Delta^3} (\vec{\Delta} \cdot \hat{R}) = -\mu_{\odot} \frac{r'}{\Delta^3} C, \end{cases} \quad (4.702)$$

where $\hat{r}$, $\hat{\theta}$, and $\hat{R}$ are the unit vectors of S , T , and W . To derive (4.702) we also need the relationship

$$\hat{r}' \cdot \hat{r} = \cos \psi = A \cos f + B \sin f, \quad (4.703)$$

where A and B are given by (4.498) and (4.499) in Sect. 4.7.1, respectively, but the third body now is the Sun, therefore in the original expressions of A and B , i' , $\theta = \Omega - \Omega'$, and u' should be

$$i' = \varepsilon, \quad \Omega' = 0, \quad u' = f' + \omega', \quad \theta = \Omega, \quad (4.704)$$

where ε is the obliquity. The expression of C in (4.702) is given by

$$\begin{aligned} C = & \sin i [\sin(\Omega - u') + \sin(\Omega + u')] \\ & + \sin i \cos \varepsilon [\sin(\Omega - u') - \sin(\Omega + u')] \\ & + \cos i \sin \varepsilon [\sin u']. \end{aligned} \quad (4.705)$$

4.10.3.2 Perturbation Motion Equation of the Radiation Pressure

Substituting the expressions of $\cos \psi$ into (4.701) leads the perturbation function R as

$$\begin{aligned} R = & \beta_1 a \left(\frac{r}{a} \right) [A \cos f + B \sin f] \\ & + \beta_2 a^2 \left(\frac{r}{a} \right)^2 \left[-\frac{1}{3} + \frac{1}{2} (A^2 + B^2) + \frac{1}{2} (A^2 - B^2) \cos 2f + AB \sin 2f \right], \end{aligned} \quad (4.706)$$

where

$$\begin{cases} \beta_1 = -\frac{\mu_{\odot}}{r'^2} = -\left(\kappa \frac{S}{m} \right) \rho_{\odot} \left(\frac{\Delta_0^2}{r'^2} \right) \approx -\left(\kappa \frac{S}{m} \right) \rho_s, \\ \beta_2 = -\frac{\mu_{\odot}}{r'^3} \approx -\frac{\left(\kappa \frac{S}{m} \right) \rho_s}{r'}. \end{cases} \quad (4.707)$$

For a low Earth orbit satellite or a medium Earth orbit satellite, the value of ρ_s is given as it is near Earth, that $\rho_s = 4.5606 \times 10^{-6} \text{ (N m}^{-2}\text{)} = 0.3169 \times 10^{-17}$, the second value is given by normalized units.

The second part in (4.706) is almost the same as in (4.502) for the third-body perturbation if m' in the coefficient β_2 given by (4.496) is replaced by $(-\mu_\odot)$. Assuming an Earth's satellite to be a sphere with a diameter of 3 m and weight of 1 ton, then the magnitude of the radiation pressure perturbation compared to the J_2 term is a small value in the second-order, because

$$\varepsilon = \left| \vec{A}_\odot \right| / \left| \vec{F}_0 \right| = \kappa \left(\frac{S}{m} \right) \rho_\odot \frac{r^2}{\mu} = O(10^{-8} - 10^{-7}) = O(J_2^2), \quad (4.708)$$

where $\mu = GE$ and E is Earth's mass. The acceleration discussed here is for the first part of the radiation perturbation function (the part related to β_1 in (4.706)), the second part (the part related to β_2) is even smaller with a magnitude of 10^{-11} . Therefore, for the radiation pressure perturbation, only the first part is needed.

By the averages of $\left(\frac{r}{a}\right) \cos f$ and $\left(\frac{r}{a}\right) \sin f$ the perturbation function R can be decomposed into three parts that

$$R = R_c + R_l + R_s, \quad (4.709)$$

$$R_c = 0, \quad (4.710)$$

$$R_l = \beta_1 a A \left(-\frac{3}{2} e \right), \quad (4.711)$$

$$R_s = \beta_1 a \left[A \left(\frac{r}{a} \cos f + \frac{3}{2} e \right) + B \left(\frac{r}{a} \sin f \right) \right]. \quad (4.712)$$

Substituting R_l and R_s given by (4.711) and (4.712) into the perturbation equation system of the $(\partial R / \partial \sigma)$ -version and including the J_2 term give the perturbation equation system

$$\dot{\sigma} = f_0(a) + f_1(\sigma; J_2) + f_2(\sigma, t; \beta_1). \quad (4.713)$$

The third term is related to the radiation pressure and can be separated into two terms that

$$f_2(\sigma, t; \beta_1) = f_{2l}(\sigma, t; \beta_1) + f_{2s}(\sigma, t; \beta_1). \quad (4.714)$$

The right-side functions contain time t explicitly because they are related to the Sun's coordinates.

4.10.3.3 The Perturbation Solution Due to the Radiation Pressure in Kepler Orbital Elements

By the method of mean orbital elements, the perturbation solution is given by

$$\sigma_l^{(1)}(t) = \int^t \left[f_{2l} + \sum_j \frac{\partial f_{1c}}{\partial \sigma_j} \left(\sigma_l^{(1)} \right)_j \right]_{\bar{\sigma}} dt, \quad (4.715)$$

$$\sigma_s^{(2)}(t) = \int^t (f_{2s})_{\bar{\sigma}} dt \quad (4.716)$$

where f_{1c} is the part related to the J_2 term. For the solution of $M_s^{(2)}(t)$ the integrand also includes $\frac{\partial n}{\partial a} a_s^{(2)}(t)$.

For consistency with previous perturbation solutions by gravitational forces, and for eliminating the commensurable small divisors we use the quasi-mean elements to construct the solution, and in (4.715) the part of $\sum_j \frac{\partial f_{1c}}{\partial \sigma_j} \left(\sigma_l^{(1)} \right)_j$ no longer occurs.

(1) Long-period variations of $\Delta\sigma(t) = \sigma_l^{(1)}(t, \beta_1) - \sigma_l^{(1)}(t_0, \beta_1)$

By integration, the terms of $\sigma_l^{(1)}(t, \beta_1)$ in the long-period variation $\Delta\sigma(t) = \sigma_l^{(1)}(t, \beta_1) - \sigma_l^{(1)}(t_0, \beta_1)$ are as follows.

$$a_l^{(1)}(t, \beta_1) = 0, \quad (4.717)$$

$$e_l^{(1)}(t, \beta_1) = \left(\frac{3}{8} \beta_1 \right) n a^2 \sqrt{1 - e^2} [(1 - \cos i) G_1 + (1 + \cos i) G_2 + 2 \sin i G_3], \quad (4.718)$$

$$i_l^{(1)}(t) = \left(\frac{3}{8} \beta_1 \right) n a^2 \frac{e}{\sqrt{1 - e^2}} [-\sin i G_1 + \sin i G_2 - 2 \cos i G_3], \quad (4.719)$$

$$\Omega_l^{(1)}(t, \beta_1) = - \left(\frac{3}{8} \beta_1 \right) n a^2 \frac{e}{\sqrt{1 - e^2} \sin i} [\sin i G_4 - \sin i G_5 + 2 \cos i G_6], \quad (4.720)$$

$$\begin{aligned} \omega_l^{(1)}(t, \beta_1) = & -\cos i \Omega_l^{(1)}(t, \beta_1) \\ & - \left(\frac{3}{8} \beta_1 \right) n a^2 \frac{\sqrt{1 - e^2}}{e} [(1 - \cos i) G_4 + (1 + \cos i) G_5 + 2 \sin i G_6], \end{aligned} \quad (4.721)$$

$$M_l^{(1)}(t, \beta_1) = \left(\frac{3}{8} \beta_1 \right) n a^2 \frac{1+e^2}{e} [(1 - \cos i) G_4 + (1 + \cos i) G_5 + 2 \sin i G_6], \quad (4.722)$$

where β_1 is given in (4.707), and the six auxiliaries $G_1, G_1, \dots, G_6$ are given by

$$\begin{cases} G_1 = (1 + \cos \varepsilon) \cos(\omega - \Omega + u')/n_1 + (1 - \cos \varepsilon) \cos(\omega - \Omega - u')/n_2 \\ G_2 = (1 + \cos \varepsilon) \cos(\omega + \Omega - u')/n_3 + (1 - \cos \varepsilon) \cos(\omega + \Omega + u')/n_4 \\ G_3 = \sin \varepsilon [\cos(\omega - u')/n_5 - \cos(\omega + u')/n_6] \end{cases} \quad (4.723)$$

$$\begin{cases} G_4 = (1 + \cos \varepsilon) \sin(\omega - \Omega + u')/n_1 + (1 - \cos \varepsilon) \sin(\omega - \Omega - u')/n_2 \\ G_5 = (1 + \cos \varepsilon) \sin(\omega + \Omega - u')/n_3 + (1 - \cos \varepsilon) \sin(\omega + \Omega + u')/n_4 \\ G_6 = \sin \varepsilon [\sin(\omega - u')/n_5 - \sin(\omega + u')/n_6] \end{cases} \quad (4.724)$$

$$\begin{cases} n_1 = \omega_1 - \Omega_1 + n', & n_2 = \omega_1 - \Omega_1 - n' \\ n_3 = \omega_1 + \Omega_1 - n', & n_4 = \omega_1 + \Omega_1 + n' \\ n_5 = \omega_1 - n', & n_6 = \omega_1 + n' \end{cases} \quad (4.725)$$

where Ω_1 and ω_1 are the secular coefficients due to the J_2 perturbation (i.e., the rates of variation) given by (4.65) and (4.66), respectively; n' is the mean angular speed of the Sun, which is the variation rate of the mean ecliptic longitude.

(2) Short-period terms of $\sigma_s^{(2)}(t, \beta_1)$

$$a_s^{(2)}(t, \beta_1) = 2\beta_1 a^3 \left[A \left(\cos E + \frac{1}{2}e \right) + B \sqrt{1 - e^2} \sin E \right], \quad (4.726)$$

$$e_s^{(2)}(t) = \beta_1 a^2 \sqrt{1 - e^2} \left[\sqrt{1 - e^2} A \left(\frac{1}{4} \cos 2E \right) + B \left(-\frac{e}{2} \sin E + \frac{1}{4} \sin 2E \right) \right], \quad (4.727)$$

$$\begin{aligned} i_s^{(2)}(t) = & \frac{1}{4} \beta_1 a^2 / \sqrt{1 - e^2} \left[\sqrt{1 - e^2} \left(\left(\cos E + \frac{e}{2} \right) - \frac{e}{4} \cos 2E \right) H_1 \right. \\ & \left. - \left(\left(1 - \frac{e^2}{2} \right) \sin E - \frac{e}{4} \sin 2E \right) H_2 \right] \end{aligned} \quad (4.728)$$

$$\begin{aligned} \Omega_s^{(2)}(t) = & \frac{1}{4} \beta_1 a^2 / \left(\sqrt{1 - e^2} \sin i \right) \left[\left(\left(1 - \frac{e^2}{2} \right) \sin E - \frac{e}{4} \sin 2E \right) H_1 \right. \\ & \left. + \sqrt{1 - e^2} \left(\left(\cos E + \frac{e}{2} \right) - \frac{e}{4} \cos 2E \right) H_2 \right] \end{aligned} \quad (4.729)$$

$$\begin{aligned} \omega_s^{(2)}(t) = & -\cos i \Omega_s^{(2)}(t) \\ & + \beta_1 a^2 / e \left[\sqrt{1 - e^2} A \left(-\frac{e}{2} \sin E + \frac{1}{4} \sin 2E \right) + B \left(e \left(\cos E + \frac{e}{2} \right) - \frac{1}{4} \cos 2E \right) \right] \end{aligned} \quad (4.730)$$

$$M_s^{(2)}(t) = -\sqrt{1 - e^2} \left[\omega_s^{(2)}(t) + \cos i \Omega_s^{(2)}(t) \right]$$

$$-5\beta_1 a^2 \left\{ A \left[\left(1 - \frac{e^2}{2} \right) \sin E - \frac{e}{4} \sin 2E \right] + B \sqrt{1 - e^2} \left[-\cos E + \frac{e}{4} \cos 2E \right] \right\} \quad (4.731)$$

where E is the eccentric anomaly, and the two auxiliaries H_1 and H_2 are given by

$$\begin{cases} H_1 = \sin i \left[(1 + \cos \varepsilon) \cos(\omega - \Omega + u') + (1 - \cos \varepsilon) \cos(\omega - \Omega - u') \right] \\ \quad - \sin i \left[(1 + \cos \varepsilon) \cos(\omega + \Omega - u') + (1 - \cos \varepsilon) \cos(\omega + \Omega + u') \right] \\ \quad + 2 \cos i \sin \varepsilon [\cos(\omega - u') - \cos(\omega + u')] \\ H_2 = \sin i \left[(1 + \cos \varepsilon) \sin(\omega - \Omega + u') + (1 - \cos \varepsilon) \sin(\omega - \Omega - u') \right] \\ \quad - \sin i \left[(1 + \cos \varepsilon) \sin(\omega + \Omega - u') + (1 - \cos \varepsilon) \sin(\omega + \Omega + u') \right] \\ \quad + 2 \cos i \sin \varepsilon [\sin(\omega - u') - \sin(\omega + u')] \end{cases} \quad (4.732)$$

4.10.3.4 The Problem of Earth's Eclipse

The phenomenon of “eclipse” often happens in the motion of an Earth's satellite, i.e., looking from the satellite the Sun is covered by Earth when the satellite moves into Earth's shadow. The variation of the radiation pressure, which is a continuous process, becomes “discontinuous” due to the existence of Earth's shadow. This phenomenon affects the satellite's orbit, and the above-given solutions need a correction. For general accuracy requirements, the simple way is that first to find the two positions of the satellite when it enters and exists an “eclipse”, denoted by E_1 and E_2 , respectively, in every circling period, then we introduce an “eclipse” factor denoted by ν defined by

$$\nu = 1 - \frac{\Delta E}{2\pi}, \quad (4.733)$$

where $\Delta E = \widehat{E_1 E_2}$ is the arc in the shadow of Earth on the satellite's orbit. The above given long-period term now takes the following form

$$\Delta \sigma(t) = \nu \left[\sigma_l^{(1)}(t) - \sigma_l^{(1)}(t_0) \right]. \quad (4.734)$$

The short-period term $\sigma_s^{(2)}(t)$ requires different treatment, because in this circumstance the characteristic of the short-period term is changed, that it no longer has the same periodic variation, or it becomes a long-period term (the period is very long depending on the relative positions between the Sun, Earth, and the satellite), the calculation formula becomes

$$\Delta \sigma(t) = \frac{1}{2\pi} [\sigma_s^{(2)}(E_1) - \sigma_s^{(2)}(E_2)] n(t - t_0), \quad (4.735)$$

and there are two additional changes in the original $\sigma_s^{(2)}(t, \beta_1)$ that

- (1) Replace $(\cos E + \frac{e}{2})$ by $\cos E$.
- (2) Divide ΔM into two parts, ΔM_1 and ΔM_2 . The first part $\Delta M_1 = [M_s^{(2)}(E_1) - M_s^{(2)}(E_2)]$ corresponds to the integration $\int^t (f_{2s})_{\bar{\sigma}} dt$ that

$$M_s^{(2)}(t) = -\beta_1 \frac{a^2}{e} \left\{ A \left[\frac{1}{2} (3 - e^2) e \sin E + \frac{1}{4} (1 - 3e^2) \sin 2E \right] - \sqrt{1 - e^2} B \left[e \left(\cos E + \frac{e}{2} \right) + \frac{1}{4} (1 - 2e^2) \cos 2E \right] \right\} \quad (4.736)$$

The second part ΔM_2 is given by $\frac{\partial n}{\partial a} a_s^{(2)}(t)$ that

$$\Delta M_2 = \frac{1}{2} \left(-\frac{3}{2a} \Delta a \right) n(t - t_0), \quad (4.737)$$

where Δa is given by (4.735).

To obtain the positions E_1 and E_2 , involves the solution of Earth's eclipse equation, which is given in the next section.

Earth's shadow causes obvious changes in the variation of the semi-major axis of a satellite's orbit, the short-period variation becomes a long-period variation, i.e., in (4.735) the part about a , $[a_s^{(2)}(E_1) - a_s^{(2)}(E_2)]$, takes a long time to change its sign. For some satellites, although their motions are affected by a dissipative force (the atmospheric drag), their semi-major axes may become larger. The radiation pressure is a conservative force, but it is different from other conservative perturbations because the Earth's shadow makes it a discontinuous force.

When there is no Earth's shadow, both parts of $[\Delta M(t)]_s$, the direct part $[\Delta M]_1$ and the indirect part $[\Delta M]_2$ due to Δa from (4.735)–(4.737), are zero. In reality, the effect of Earth's shadow is not complicated as above discussed. It is because the arc in the shadow of a GEO satellite is short, the longest arc in the shadow happens around the equinoxes, during these times the satellite stays in the shadow for no more than 70 min. If the magnitude of the radiation pressure is not large, we do not have to consider the influence of the shadow on the semi-major axis and the transformations of the short-period terms. The only correction that should be considered is the long-period term, which is done by adding the eclipse factor $(1 - \Delta E/2\pi)$ () to related formulas. This treatment does not affect the accuracy of the solutions.

4.10.3.5 Equation of Earth's Shadow and Its Solution

Strictly speaking, there are the umbra and penumbra of Earth's shadow, also the effect of the atmospheric extinction. How to treat this problem depends on the required accuracy. For an Earth's satellite, the above estimated magnitude of the radiation

pressure is $\varepsilon = O(10^{-8} - 10^{-7})$ by (4.708). From Earth to see the Sun the geometric parallax has a magnitude of 10^{-4} , we then can treat Earth's shadow as a cylinder. The error introduced by this simplification to Earth's eclipse model leads to an error in the radiation pressure perturbation, which is only in $O(10^{-12} - 10^{-11})$ and usually can be ignored. We then derive an equation of Earth's shadow based on a cylindrical Earth shadow model, which is

$$\sin(\theta + f) = -\frac{1}{K} \left[1 - \frac{R^2}{p^2} (1 + e \cos f)^2 \right]^{1/2}, \quad (4.738)$$

where

$$\begin{cases} K^2 = A^2 + B^2 & (K > 0), \\ \theta = \arctan(A/B), \end{cases} \quad (4.739)$$

and A and B are given by (4.498) and (4.499). In (4.738) R is the radius of the cross-section of the cylinder, by the normalized unit system, that $R = a_e = 1$.

To solve the equation of Earth's shadow (4.738) means to find two true anomalies, f_1 and f_2 , that f_1 corresponds to the time when the satellite enters the shadow and f_2 the time when the satellite exists in the shadow. The equation can be solved by an iterative process as follows.

(1) First set $e = 0$, then from

$$\sin(\theta + f)^{(0)} = -\frac{1}{K} \left[1 - \left(\frac{R}{p} \right)^2 \right]^{1/2} \quad (4.740)$$

to obtain $f_1^{(0)}$ and $f_2^{(0)}$ as the initial values of f_1 and f_2 , respectively. Obviously one of the angles, say f_2 , is in the fourth quadrant, and the other f_1 in the third quadrant, also according to the properties of Earth's shadow the arc of $\widehat{f_2 f_1} > 180^\circ$, when $(\theta + f)^{(0)}$ is in the fourth quadrant one solution is $f_2^{(0)}$ and the other is $f_1^{(0)}$. If there is no solution, i.e., $|\sin(\theta + f)^{(0)}| > 1$, we take $\sin(\theta + f)^{(0)} = -1$ to give $(\theta + f)^{(0)} = 270^\circ$, then $f_2^{(0)} = f_1^{(0)} = 270^\circ - \theta$.

(2) Substituting $f_1^{(0)}$ and $f_2^{(0)}$ into the original Eq. (4.738) to find the next pair of solution

$$\sin(\theta + f)^{(k)} = -\frac{1}{K} \left[1 - \left(\frac{R}{p} \right)^2 (1 + e \cos f^{(k-1)})^2 \right]^{\frac{1}{2}}, \quad k = 1, 2, \dots \quad (4.741)$$

Repeat the process until f_1 and f_2 satisfy the required accuracy that

$$\left| f_1^{(k)} - f_1^{(k-1)} \right| < \varepsilon^*, \quad \left| f_2^{(k)} - f_2^{(k-1)} \right| < \varepsilon^*$$

there ε^* is a given criterion based on the magnitude of the radiation pressure and the accuracy requirement of a actual project.

Note that during the iterative process $f_1^{(k)}$ and $f_2^{(k)}$ should always be corresponding to $(\theta + f)^{(k)}$ in the fourth quadrant and third quadrant, respectively. If $|\sin(\theta + f)^{(k)}| > 1$ occurs twice in the process, then we assume that there is no shadow, giving $f_2 = 0$ and $f_1 = 2\pi$. From the solutions of f_1 and f_2 , we can obtain E_1 and E_2 which are needed in the perturbation function by

$$\sin E = \frac{\sqrt{1-e^2}}{1+e \cos f} \sin f, \quad \cos E = \frac{1-e^2}{1+e \cos f} \cos f + e. \quad (4.742)$$

4.10.4 The Non-singularity Perturbation Solution of the First Type Due to the Radiation Pressure

Based on the simplified model of the radiation perturbation the non-singularity solution has the form given by (4.734) and (4.735)

$$\Delta \sigma_l(t) = v \left[\sigma_l^{(1)}(t) - \sigma_l^{(1)}(t_0) \right], \quad (4.743)$$

$$\Delta \sigma_s(t) = \frac{1}{2\pi} [\sigma_s^{(2)}(E_1) - \sigma_s^{(2)}(E_2)] n(t - t_0). \quad (4.744)$$

When the satellite is not in an eclipse the corresponding short period term $\Delta \sigma_s(t) = 0$, means that the short-period variation has no influence (or the influence is too small to consider). To know if the satellite is in an eclipse, we use the middle time of $(t - t_0)$ of the extrapolation arc as the standard point.

Note that in the above method $\Delta M_s(t)$ has two parts, ΔM_1 and ΔM_2 , see (4.736) and (4.737).

(1) Long-period terms of $\sigma_l^{(1)}(t, \beta_1)$ in $\Delta \sigma_l(t, \beta_1)$

$$a_l^{(1)}(t, \beta_1) = 0, \quad (4.745)$$

$$i_l^{(1)}(t) = \left(\frac{3}{8} \beta_1 \right) n a^2 \frac{e}{\sqrt{1-e^2}} [-\sin i G_1 + \sin i G_2 - 2 \cos i G_3], \quad (4.746)$$

$$\Omega_l^{(1)}(t, \beta_1) = -\left(\frac{3}{8}\beta_1\right)na^2\frac{e}{\sqrt{1-e^2}\sin i}[\sin i G_4 - \sin i G_5 + 2\cos i G_6], \quad (4.747)$$

$$\xi_l^{(1)}(t, \beta_1) = \cos \omega[e_l^{(1)}(t, \beta_1)] - \sin \omega[e\omega_l^{(1)}(t, \beta_1)], \quad (4.748)$$

$$\eta_l^{(1)}(t, \beta_1) = \sin \omega[e_l^{(1)}(t, \beta_1)] + \cos \omega[e\omega_l^{(1)}(t, \beta_1)], \quad (4.749)$$

$$\begin{aligned} \lambda_l^{(1)}(t, \beta_1) = & -\cos i \Omega_l^{(1)}(t, \beta_1) \\ & + \left(\frac{3}{8}\beta_1\right)na^2\left(\frac{2+\sqrt{1-e^2}}{1+\sqrt{1-e^2}}\right)e[(1-\cos i)G_4 + (1+\cos i)G_5 + 2\sin i G_6] \end{aligned} \quad (4.750)$$

where $e_l^{(1)}(t, \beta_1)$ and $\omega_l^{(1)}(t, \beta_1)$ are given by (4.718) and (4.721), respectively.

(2) Short-period terms of $\sigma_s^{(2)}(t, \beta_1)$ in $\Delta\sigma_s(t, \beta_1)$ (i.e., (4.744))

$$a_s^{(2)}(t, \beta_1) = 2\beta_1 a^3 [A \cos E + B\sqrt{1-e^2} \sin E], \quad (4.751)$$

$$\begin{aligned} i_s^{(2)}(t, \beta_1) = & \frac{1}{4}\beta_1\frac{a^2}{\sqrt{1-e^2}}\left\{\sqrt{1-e^2}\left[(\cos E) - \frac{e}{4}\cos 2E\right]H_1\right. \\ & \left.- \left[\left(1 - \frac{e^2}{2}\right)\sin E - \frac{e}{4}\sin 2E\right]H_2\right\} \end{aligned} \quad (4.752)$$

$$\begin{aligned} \Omega_s^{(2)}(t, \beta_1) = & \frac{1}{4}\beta_1\frac{a^2}{(\sqrt{1-e^2}\sin i)}\left\{\left[\left(1 - \frac{e^2}{2}\right)\sin E - \frac{e}{4}\sin 2E\right]H_1\right. \\ & \left.+ \sqrt{1-e^2}\left[(\cos E + \frac{e}{2}) - \frac{e}{4}\cos 2E\right]H_2\right\} \end{aligned} \quad (4.753)$$

$$\xi_s^{(2)}(t, \beta_1) = \cos \omega[e_s^{(2)}(t, \beta_1)] - \sin \omega[e\omega_s^{(2)}(t, \beta_1)], \quad (4.754)$$

$$\eta_s^{(2)}(t, \beta_1) = \sin \omega[e_s^{(2)}(t, \beta_1)] + \cos \omega[e\omega_s^{(2)}(t, \beta_1)], \quad (4.755)$$

$$\begin{aligned} \lambda_s^{(2)}(t, \beta_1) = & -\cos i \Omega_s^{(2)}(t, \beta_1) \\ & + \beta_1 a^2 \left\{ -A \left[\frac{1}{2} (3 + \sqrt{1-e^2} - e^2) \sin E - \frac{3}{4} e \sin 2E \right] \right. \\ & \left. + B \left[(1 + \sqrt{1-e^2}) \cos E - \frac{1}{2} \sqrt{1-e^2} e \cos 2E \right] \right\} \\ & - \frac{1}{4} \beta_1 a^2 (A + B) \frac{e}{1 + \sqrt{1-e^2}} \sin 2E \end{aligned} \quad (4.756)$$

Note that (4.756) is the first part of $\Delta\lambda_s(t, \beta_1)$, the second part $[\Delta\lambda_s(t, \beta_1)]_2$ is

$$[\Delta\lambda_s(t, \beta_1)]_2 = \frac{1}{2} \left(-\frac{3}{2a} \Delta a_s(t, \beta_1) \right) n(t - t_0), \quad (4.757)$$

The second part comes from the simplified treatment of Earth's shadow as described above in (4.736) and (4.737). Also $e_s^{(2)}(t, \beta_1)$ and $\omega_s^{(2)}(t, \beta_1)$ in (4.754) and (4.755) are given by (4.727) and (4.730), respectively.

4.10.5 The Non-singularity Perturbation Solution of the Second Type Due to the Radiation Pressure

(1) Long-period terms of $\sigma_l^{(1)}(t, \beta_1)$ in $\Delta\sigma_l(t, \beta_1)$

$$a_l^{(1)}(t, \beta_1) = 0, \quad (4.758)$$

$$\xi_l^{(1)}(t, \beta_1) = \cos \tilde{\omega} [e_l^{(1)}(t, \beta_1)] - \sin \tilde{\omega} [e\omega_l^{(1)}(t, \beta_1) + e\Omega_l^{(1)}(t, \beta_1)], \quad (4.759)$$

$$\eta_l^{(1)}(t, \beta_1) = \sin \tilde{\omega} [e_l^{(1)}(t, \beta_1)] + \cos \tilde{\omega} [e\omega_l^{(1)}(t, \beta_1) + e\Omega_l^{(1)}(t, \beta_1)], \quad (4.760)$$

$$h_l(t, \beta_1) = \frac{1}{2} \cos(i/2) \cos \Omega[i_l(t, \beta_1)] - \frac{1}{2 \cos(i/2)} \sin \Omega[\sin i \Omega_l(t, \beta_1)], \quad (4.761)$$

$$k_l(t, \beta_1) = \frac{1}{2} \cos(i/2) \sin \Omega[i_l(t, \beta_1)] + \frac{1}{2 \cos(i/2)} \cos \Omega[\sin i \Omega_l(t, \beta_1)], \quad (4.762)$$

$$\begin{aligned} \lambda_l^{(1)}(t, \beta_1) = & -\left(\frac{3}{8}\beta_1\right)na^2 \left\{ \frac{\sin i}{(1 + \sin i)\sqrt{1 - e^2}} e[\sin i G_4 - \sin i G_5 + 2 \cos i G_6] \right. \\ & \left. - \left(\frac{2 + \sqrt{1 - e^2}}{1 + \sqrt{1 - e^2}} e \right) [(1 - \cos i)G_4 + (1 + \cos i)G_5 + 2 \sin i G_6] \right\} \end{aligned} \quad (4.763)$$

where $e_l^{(1)}(t, \beta_1)$, $i_l^{(1)}(t, \beta_1)$, $\Omega_l^{(1)}(t, \beta_1)$ and $\omega_l^{(1)}(t, \beta_1)$ are given by (4.718)–(4.721).

(2) Short-period terms of $\sigma_s^{(2)}(t, \beta_1)$ in $\Delta\sigma_s(t, \beta_1)$ (i.e., (4.744))

$$a_s^{(2)}(t, \beta_1) = 2\beta_1 a^3 \left[A \cos E + B \sqrt{1 - e^2} \sin E \right], \quad (4.764)$$

$$\xi_s^{(2)}(t, \beta_1) = \cos \tilde{\omega} [e_s^{(2)}(t, \beta_1)] - \sin \tilde{\omega} [e\omega_s^{(2)}(t, \beta_1) + e\Omega_s^{(2)}(t, \beta_1)], \quad (4.765)$$

$$\eta_s^{(2)}(t, \beta_1) = \sin \tilde{\omega} [e_s^{(2)}(t, \beta_1)] + \cos \tilde{\omega} [e\omega_s^{(2)}(t, \beta_1) + e\Omega_s^{(2)}(t, \beta_1)], \quad (4.766)$$

$$h_s(t, \beta_1) = \frac{1}{2} \cos(i/2) \cos \Omega [i_s(t, \beta_1)] - \frac{1}{2 \cos(i/2)} \sin \Omega [\sin i \Omega_s(t, \beta_1)], \quad (4.767)$$

$$k_s(t, \beta_1) = \frac{1}{2} \cos(i/2) \sin \Omega [i_s(t, \beta_1)] + \frac{1}{2 \cos(i/2)} \cos \Omega [\sin i \Omega_s(t, \beta_1)], \quad (4.768)$$

$$\begin{aligned} \lambda_s^{(2)}(t, \beta_1) = & \frac{1}{4} \beta_1 a^2 \left(\frac{\sin i}{\sqrt{1 - e^2} (1 + \cos i)} \right) \left\{ \left[\left(1 - \frac{e^2}{2} \right) \sin E - \frac{e}{4} \sin 2E \right] H_1 \right. \\ & + \sqrt{1 - e^2} \left[\left(\cos E + \frac{e}{2} \right) - \frac{e}{4} \cos 2E \right] H_2 \Big\} \\ & + \beta_1 a^2 \left\{ -A \left[\frac{1}{2} (3 + \sqrt{1 - e^2} - e^2) \sin E - \frac{3}{4} e \sin 2E \right] \right. \\ & + B \left[\left(1 + \sqrt{1 - e^2} \right) \cos E - \frac{1}{2} \sqrt{1 - e^2} e \cos 2E \right] \Big\} \\ & - \frac{1}{4} \beta_1 a^2 (A + B) \frac{e}{1 + \sqrt{1 - e^2}} \sin 2E \end{aligned} \quad (4.769)$$

Same as (4.756), (4.769) is the first part of $\Delta\lambda_s(t, \beta_1)$, the second part $[\Delta\lambda_s(t, \beta_1)]_2$ is

$$[\Delta\lambda_s(t, \beta_1)]_2 = \frac{1}{2} \left(-\frac{3}{2a} \Delta a_s(t, \beta_1) \right) n(t - t_0). \quad (4.770)$$

The expressions of $e_s^{(2)}(t, \beta_1)$, $i_s^{(2)}(t, \beta_1)$, $\Omega_s^{(2)}(t, \beta_1)$ and $\omega_s^{(2)}(t, \beta_1)$ in (4.765)–(4.768) are given by (4.727)–(4.730).

4.11 Perturbed Orbit Solution Due to Atmospheric Drag

4.11.1 Damping Effect: Atmospheric Drag

For an artificial satellite moving in Earth's atmosphere, the dynamic system about the exerting forces on the satellite is an extremely complicated dynamical problem and is of great importance in the field of supersonic fluid dynamics. This problem involves the continuum flow in the lower atmosphere, the free molecular flow in the upper atmosphere, and the transitional flow between them; it also involves neutral atmosphere, ion atmosphere, and mixed atmosphere of many compositions, etc. About the high-speed high-altitude fluid dynamics in the engineering application Earth's atmosphere is generally assumed to be neutral and is divided by the Knudsen number, suggested by Qian, XS of China in 1946. The definition of the Knudsen number, K_n is given by

$$K_n = \frac{\lambda}{L}, \quad (4.771)$$

where λ is the mean free path of gas molecular; L is the representative physical length scale of an aircraft; the value of K_n provides the rarefaction of a flow as follows.

$$\begin{cases} K_n < 0.01, & \text{continuum flow,} \\ K_n = 0.01 - 0.1, & \text{slip flow,} \\ K_n = 0.1 - 10, & \text{transitional flow,} \\ K_n > 10, & \text{free molecular flow.} \end{cases} \quad (4.772)$$

In Earth's atmosphere above the altitude of 100 km the molecular mean free path λ is about 10 cm at 100 km, 1 m at 120 km, and 100 m at 180 km. Then for an aircraft with a representative physical length scale of 10 m, the Knudsen numbers are

$$\begin{cases} K_n = 0.01, & h = 100 \text{ km, continuum flow,} \\ K_n = 0.1, & h = 120 \text{ km, transitional flow,} \\ K_n = 10, & h = 180 \text{ km, free molecular flow,} \end{cases}$$

showing that the altitude of 100 km is the upper boundary of the continuum atmosphere, and 180 km is the lower boundary of the free-molecular atmosphere.

Usually, artificial satellites move in the atmosphere above 200 km, and their representative physical length scale is relatively small. Therefore, they are in the layer of the free molecular atmosphere, and the atmospheric drag exerting on a satellite produces a resistant acceleration. The resistant acceleration can be expressed as

$$\vec{D} = -\frac{1}{2} \left(\frac{C_D S}{m} \right) \rho V \vec{V}, \quad (4.773)$$

where C_D is the drag coefficient, that $C_D = 2.2 \pm 0.2$; S/m is the area-to-mass ratio of a satellite (for a resistant force, S is the equivalent cross-section); ρ is the local atmospheric density where the satellite is; $\vec{V} = \vec{v} - \vec{v}_a$ is the relative velocity of the satellite with respect to the atmosphere, $\vec{v}$ and $\vec{v}_a$ are the velocities of the satellite and the atmosphere with respect to the center of Earth, respectively.

The formula (4.773) is a proper choice, which is proved by satellite data over decades. But there are special satellites (including these need to be retrieved) and some spacecraft to be falling, which fly in the atmosphere below 200 km. In these situations, the satellites are in the transitional and continuum atmosphere layers. Considering the retrieval and falling of a spacecraft it is unavoidable to take special adjustments and certain approximations about the state of the atmosphere at different altitudes.

For the atmosphere between 100 and 200 km, in which the atmosphere changes from the transitional flow to the free-molecular flow, the drag coefficient is approximately given by [12–14]

$$C_D = C_0 + C_1 K_n^{-1} + C_2 K_n^{-2} + C_3 K_n^{-3} \quad (K_n > 3), \quad (4.774)$$

where $C_0 = 2.2$, and C_1 , C_2 , and C_3 are approximately given by

$$C_1 = -0.90, \quad C_2 = 0.16, \quad C_3 = -0.004. \quad (4.775)$$

Considering the changes in atmospheric density models and the current situation, (4.774) can be simplified as

$$C_D = C_0 - C^* L \rho, \quad (4.776)$$

where C^* is a constant to be decided; the units of density ρ and length scale L are kg/m^3 and m, respectively. The value of C_D decreases with decreasing altitude, but it does not mean the drag also decreases, because with decreasing altitude the density ρ may increase much faster than the decreasing of C_D .

It is much more difficult to find an expression for the drag coefficient when the atmosphere changes from the state of slip flow to the state of transitional flow. Below 100 km, it is a continuum flow region, the forces exerting on a spacecraft are complicated. But a spacecraft to be retrieved or to fall moves fast and will stay in this region for a very short time, usually we use a simplified method to deal with it. The buoyant force and the drag are expressed approximately by

$$\begin{cases} \vec{L} = -\frac{1}{2} \left(\frac{C_L S}{m} \right) \rho v^2 \hat{l} \\ \vec{D} = -\frac{1}{2} \left(\frac{C_D S}{m} \right) \rho v^2 \hat{d} \end{cases} \quad (4.777)$$

where C_L is the buoyant coefficient and C_D the drag coefficient, $\hat{l}$ and $\hat{d}$ are unit vectors of directions of buoyant force and drag, respectively. These parameters depend on the flying state of a spacecraft or other available information.

The magnitude of atmospheric drag exerting on a spacecraft depends on the shape and size of its surface, and its attitude, similar to the radiation pressure. About the value of S/m in (4.773) generally the area $S = S(t)$ is given as the equivalent cross-section of the spacecraft.

Earth rotates and it “carries” its atmosphere, but the rule of the atmospheric rotation is complicated. If we denote n_e to the Earth’s rotation angular speed and ω_a to the atmosphere rotation angular speed, then below 200 km, there is $\omega_a = n_e$, and above 200 km the estimation is

$$0.8 \leq \frac{\omega_a}{n_e} \leq 1.4, \quad (4.778)$$

which varies with altitude. This rotation mechanism of the atmosphere is not well understood, usually it is assumed that $\omega_a = n_e$.

According to the above analysis, for precise orbital determination and forecast of Earth’s satellites, the atmospheric drag coefficient can be given by (4.773), but the parameters need complicated adjustments. With the adjustments, it seems that it is not difficult to produce satellite perturbation ephemerides by numerical methods. But to reach high accuracies, the parameters C_D , ω_a , and S/m must be adjusted carefully, more importantly, we need a high precision atmospheric density model $\rho(r, t)$. This is an extremely complicated key factor that constrains high precise orbit determinations and forecasts for Earth’s satellites, especially for the low Earth orbit satellites.

4.11.2 Atmosphere Density Model

(1) The state of the upper atmosphere and its density model

An atmosphere model usually is a mathematical model providing atmospheric state parameters (pressure p , temperature T , density ρ , etc.) and their variations (including all kinds of discrete data and related calculation formulas). Since the launching of Earth’s artificial satellites, many countries have accumulated the ever-increasing data (including information of atmospheric drag exerting on satellites, and direct measurements by instruments on board satellites) over many decades. Based on the new information researchers have obtained important discoveries and results, which have changed our understanding of the structures and dynamics of the upper atmosphere and produced different kinds of atmospheric models. Among the new atmosphere models, the most influential models include the international reference atmosphere models CIRA, the Jacchia models, the DTM models, and the MSIS models. These models, being constantly improved, have many versions. The atmosphere is strongly influenced by the Sun, but the variation of the Sun’s activity is far from understood, therefore, all atmosphere models are still far from ideal and in need of improvement. Although each model has its advantages and shortages, there are two common problems, one is that the errors often exceed 10% when using data provided by any specific model for forecasts; the other is that the calculation

of atmospheric density is too complicated, for high precise orbit determinations it is usually by calculating discrete densities along the satellite orbit (requiring more information of the space environment) to provide satellite ephemerides. We do not further discuss these problems in this book. Our purpose is to provide an approximate atmosphere density model by which to construct the analytic perturbation solution caused by the atmospheric drag.

The above-mentioned upper atmosphere models are different from each other, as they use different information including different physical or chemical activities, and are for different ranges of altitude, etc. But the density distributions from all models have the same basic characteristics given as follows.

- ① The atmospheric density decreases with increasing altitude, the speed of decreasing reduces with increasing altitude, and the isodensity surface has a similar shape as Earth's surface, which is an oblate spheroid.
- ② The distribution of the upper atmosphere density is significantly affected by the solar radiation (ultraviolet and particle radiations), and the influence increases with increasing height. There are diurnal, seasonal, semi-annual, and long-period variations (the 11-year periodic variation related to the solar activity), etc., and the density, temperature, compositions, and motions of the atmosphere change accordingly. From the statistical point of view, the solar radiation effect can be presented by the solar 10.7 cm flux $F_{10.7}$ and the geomagnetic index C_p . These parameters are used approximately as proxies for an approximate model, they do not reflect the fundamental rules.

(2) Approximate formulas of atmosphere density

It is necessary to have a formula for the atmosphere density distribution in order to understand the basic influence of the atmospheric drag perturbation on satellite's orbits. Based on the characteristics of atmosphere density distribution provided by these empirical models, there are some practical approximate models, which are introduced as follows.

Strictly speaking, the atmospheric density ρ influenced by Earth's gravity and the solar radiation is a function of location and time, meaning

$$\rho = \rho(\vec{r}, t), \quad (4.779)$$

The influence of solar radiation depends on time, as mentioned above, and is a problem we do not well understand yet, so can only be given by approximate models.

If we only consider the balance of the gravity force and the buoyant force then according to hydrostatic dynamics, the vertical distribution of the atmosphere density is exponential that

$$\rho = \rho_0 \exp\left(-\frac{r - r_0}{H}\right). \quad (4.780)$$

This provides a spherical atmosphere model, where ρ_0 is the density at the reference sphere surface with a radius $r = r_0$, and H is called the density scale height. This model agrees with the law that atmosphere density decreases with increasing altitude, but according to the available atmosphere models, the speed of density decreasing reduces with increasing altitude, which means the density scale height H is not a constant but increases with altitude slowly. Approximately we assume H to have a linear relationship with height h , that for h between 200 and 600 km, there is

$$H = H(r) = H_0 + \frac{\mu}{2}(r - r_0), \quad (4.781)$$

where $\mu \approx 0.1$, and usually $\mu < 0.2$. Note that here μ is the variation rate of density scale height and do not mix it with Earth's gravitational constant $\mu = GE$. The atmosphere density formula (4.780) now becomes

$$\begin{aligned} \rho &= \rho_0 \exp\left(-\frac{r - r_0}{H}\right) \\ &= \rho_0 \left[1 + \frac{\mu}{2}\left(-\frac{r - r_0}{H_0}\right)^2\right] \exp\left(-\frac{r - r_0}{H_0}\right) \end{aligned} \quad (4.782)$$

The parameters in (4.782) can be obtained from available models. For example, by the CIRA-61 model, at $h = 200$ km, there are

$$\rho_0 = 3.6 \times 10^{-10} \text{ kg/m}^3, \quad H_0 = 37.4 \text{ km}, \quad \mu = 0.1.$$

The next correction is about the isodensity surface, which can be given as an oblate spheroid similar to Earth's surface, then (4.782) takes the following form

$$\begin{aligned} \rho &= \rho_0 \exp\left(-\frac{r - \sigma}{H(r)}\right) \\ &= \rho_0 \left[1 + \frac{\mu}{2}\left(-\frac{r - \sigma}{H_0}\right)^2\right] \exp\left(-\frac{r - \sigma}{H_0}\right) \end{aligned} \quad (4.783)$$

This is the oblate spheroid atmosphere density model, in which σ is the distance from Earth's center to the point on the reference spheroid, at which the atmosphere density is ρ_0 and the density scale height is H_0 (Fig. 4.4). The model (4.783) provides the distribution of atmosphere density due to Earth's gravity force, which is close to the mean atmosphere density by available empirical models. This model was used in the early study of the influence of atmospheric drag on satellite orbits with the initial perigee of a satellite as the reference point p_0 . The density ρ_0 and scale height H_0 at p_0 are ρ_{p_0} and H_{p_0} , respectively.

The influence of solar radiation on atmosphere density distribution causes periodic variations. The most marked is the diurnal variation related to Earth's rotation, that at the same altitude and same latitude the density is much higher during daytime

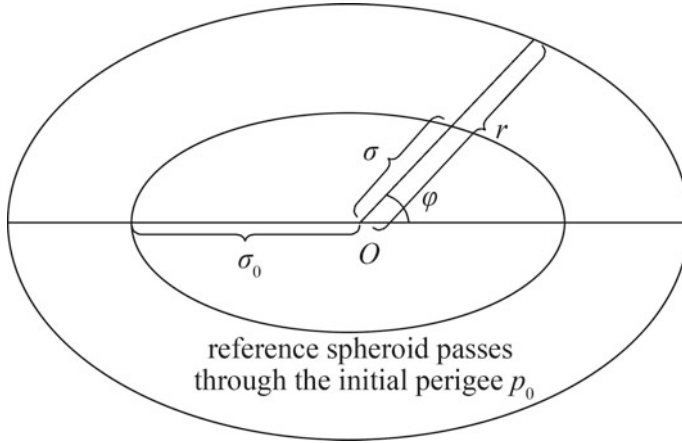


Fig. 4.4 Illustration of isodensity as the surface of an oblate spheroid

than nighttime. Generally, the density reaches its daily maximum at local time 14 o'clock and changes fast around that time, its minimum is around 2–5 o'clock, and changes slowly. Approximately we assume that the center of Earth and the locations of the minimum and maximum densities are collinear, and the isodensity spheroid is symmetric with respect to this line, then the density formula (4.780) is adjusted to

$$\rho = \rho_0(1 + F^* \cos \psi) \exp\left(-\frac{r - \sigma}{H(r)}\right), \quad (4.784)$$

where ψ is the angle of the radius $\vec{r}$ (i.e., the position of a satellite) and the position of the diurnal peak density $\vec{r}_m$, given by

$$\cos \psi = \hat{r} \cdot \hat{r}_m. \quad (4.785)$$

$\hat{r}$ and $\hat{r}_m$ are the unit vectors of $\vec{r}$ and $\vec{r}_m$, respectively. By the assumption of a symmetric diurnal effect, the relationship between the direction of the peak density and the direction of the Sun is given by

$$\begin{cases} \alpha_m = \alpha + \lambda_m \\ \delta_m = \delta \end{cases} \quad (4.786)$$

where α and δ are the Sun's right ascension and declination, respectively; $\lambda_m = 30^\circ$ is given assuming the peak density to appear at 14 o'clock; F^* is the factor of diurnal variation, which is related to the ratio of daily maximum to daily minimum, f^* , that

$$\begin{cases} f^* = \frac{\rho_{\max}}{\rho_{\min}} = \frac{1 + F^*}{1 - F^*} \\ F^* = \frac{f^* - 1}{f^* + 1} \end{cases} \quad (4.787)$$

Considering the diurnal variation, ρ_0 in (4.784) now is the daily mean density on the reference spheroid $r = \sigma$ (corresponding to $\psi = 90^\circ$), which can be calculated by available models and parameters.

There is no question that the mathematical density distribution model can be further adjusted by including other factors, such as the change of diurnal effect with altitude, non-symmetric diurnal effect, seasonal variations, etc., then the formula is certainly much more complicated. The complication in fact does not change the basic form of the density expression, and in the analysis of the influence of the atmospheric drag on satellite orbits, it may not provide fundamental new information. Therefore, it is unnecessary to pursue complicated formulas in high precise orbit determination. In reality, in order to find a density closer to the real situation, the method is directly to calculate density point by point using a selected atmosphere model. In this book, we use a commonly accepted and relatively simple density model based on (4.784) for constructing the perturbation solution due to the atmospheric drag. Results are for general orbit determinations and forecasts. Also, the parameters in density calculation formulas and other parameters appeared in the perturbing acceleration such as C_D and S/m , etc., can be adjusted accordingly.

4.11.3 Atmospheric Rotation and the Expression of Atmospheric Drag

In reality, Earth's atmosphere not only rotates but also in a complicated way. As mentioned above the rotation angular speed of the atmosphere ω_a can be approximately given by the same value as Earth's rotation angular speed n_e . By this assumption we have

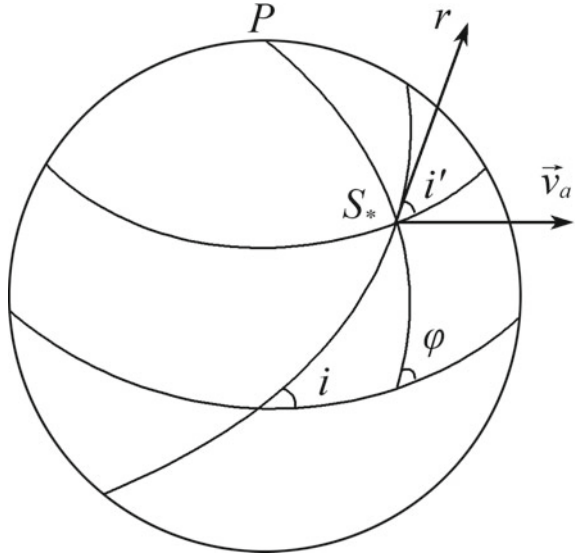
$$v_a = r n_e \cos \varphi, \quad (4.788)$$

where φ is the moving satellite's geographic latitude. The three components of the satellite's velocity $\vec{v}$ in the tangential direction, the primary normal direction, and the second-normal direction (i.e., normal on the orbital plane) are

$$v_U = v, \quad v_N = 0, \quad v_W = 0. \quad (4.789)$$

The three components of the atmospheric rotation velocity $\vec{v}_a$ in the radial direction, the transverse direction, and the normal direction of the orbital plane are (Fig. 4.5)

Fig. 4.5 The auxiliary sphere passing the satellite S_*



$$\begin{cases} (v_a)_S = 0, \\ (v_a)_T = rn_e \cos \varphi \cos i', \\ (v_a)_W = -rn_e \cos \varphi \sin i'. \end{cases} \quad (4.790)$$

Then the three components of the atmospheric velocity $\vec{v}_a$ in the tangential direction, the primary normal direction, and the second-normal direction is (Fig. 4.6)

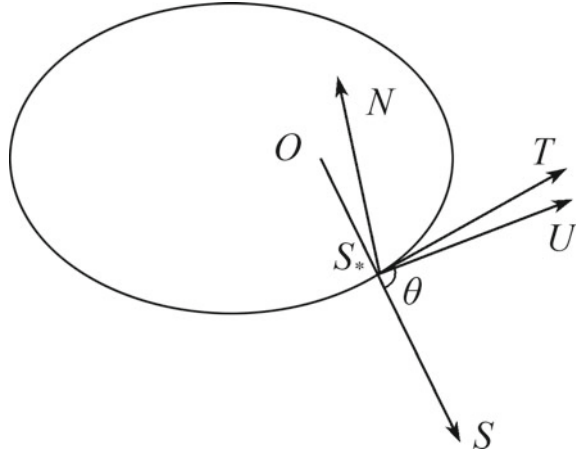
$$\begin{cases} (v_a)_U = (v_a)_T \sin \theta, \\ (v_a)_N = (v_a)_T \cos \theta, \\ (v_a)_W = -rn_e \cos \varphi \sin i'. \end{cases} \quad (4.791)$$

The relationships between (S, T) components and (U, N) components are

$$\begin{cases} \sin \theta = \frac{1+e \cos f}{\sqrt{1+2e \cos f+e^2}} = 1 - \frac{e^2}{2} \sin^2 f + O(e^3) \\ \cos \theta = \frac{e \sin f}{\sqrt{1+2e \cos f+e^2}} = e \sin f - \frac{e^2}{2} \sin 2f + O(e^3) \end{cases} \quad (4.792)$$

Substituting (4.790) and (4.792) into (4.791), and then substituting the result and (4.789) into (4.773), we have the three components of the perturbing acceleration due to a rotating atmospheric drag that

Fig. 4.6 Geometric relationship of vectors in the orbital plane



$$\begin{cases} U = -\frac{\mu}{2} \left(\frac{C_D S}{m} \right) \rho V v \left\{ 1 - \left(\frac{rn_e}{v} \right) \cos \varphi \cos i' \left[1 - \frac{e^2}{2} \sin^2 f + O(e^3) \right] \right\} \\ N = -\frac{\mu}{2} \left(\frac{C_D S}{m} \right) \rho V v \left\{ -\left(\frac{rn_e}{v} \right) \cos \varphi \cos i' \left[e \sin f - \frac{e^2}{2} \sin 2f + O(e^3) \right] \right\} \\ W = -\frac{\mu}{2} \left(\frac{C_D S}{m} \right) \rho V v \left\{ \left(\frac{rn_e}{v} \right) \cos \varphi \sin i' \right\} \end{cases} \quad (4.793)$$

where

$$V = |\vec{v} - \vec{v}_a| = v \left[1 - \left(\frac{rn_e}{v} \right) \cos \varphi \cos i' + \frac{1}{2} \left(\frac{rn_e}{v} \right)^2 \cos^2 \varphi \sin^2 i' + O \left(\left(\frac{rn_e}{v} \right)^2 e^2 \right) \right] \quad (4.794)$$

For a low Earth satellite (such as a 2-h satellite) there is

$$\frac{rn_e}{v} = \frac{1}{12} = 0.8 \times 10^{-1}. \quad (4.795)$$

Usually, a satellite whose motion is affected by the atmospheric drag has a small eccentricity, mostly $e < 0.2$, then we can omit the following terms of small magnitudes:

$$\begin{aligned} \left(\frac{rn_e}{v} \right)^2 &= 7 \times 10^{-3}, \quad \left(\frac{rn_e}{v} \right)^2 e = 10^{-3}, \\ \left(\frac{rn_e}{v} \right) e &= 1.7 \times 10^{-2}, \quad \left(\frac{rn_e}{v} \right) e^2 = 3 \times 10^{-3}. \end{aligned}$$

By the spherical trigonometric formulas (Fig. 4.5)

$$\cos \varphi \cos i' = \cos i, \quad \cos \varphi \sin i' = \cos(f + \omega) \sin i, \quad (4.796)$$

we have the square of the satellite's speed, v^2 as

$$v^2 = \mu \left(\frac{2}{r} - \frac{1}{a} \right) = \mu \frac{(1 + 2e \cos f + e^2)}{a(1 - e^2)}. \quad (4.797)$$

The three components of the perturbing acceleration (4.793) can be simplified to

$$\begin{cases} U = -\frac{\mu}{2} A_1 \frac{(1+2e \cos f + e^2)}{a(1-e^2)} \rho \\ N = 0 \\ W = -\frac{\sqrt{\mu}}{2} A_2 r \cos(f + \omega) \sin i \left[\frac{(1+2e \cos f + e^2)}{a(1-e^2)} \right]^{1/2} \rho \end{cases} \quad (4.798)$$

where the two parameters A_1 and A_2 are defined as

$$A_1 = \left(\frac{C_D S}{m} \right) F^2, \quad A_2 = \left(\frac{C_D S}{m} \right) n_e F, \quad (4.799)$$

$$F = \left(1 - \frac{r n_e}{v} \cos i \right) \approx \left(1 - \frac{r_{p0} n_e}{v_{p0}} \cos i \right). \quad (4.800)$$

In the above formulas, the component W is directly related to the atmospheric rotation. The component N is small, so in the simplified (4.798) it is omitted. Based on the property of the perturbation equations omitting N does not change the characteristics of the perturbation effect by the atmospheric drag.

4.11.4 Structure of the Perturbed Solution Due to the Atmospheric Drag

4.11.4.1 Basic Equations of Atmospheric Drag Perturbation

Substituting the expressions of perturbing accelerations U and W in (4.798) into the perturbation equation system, we obtain the ordinary differential equations of satellite orbital elements due to the dissipative force of atmospheric drag:

$$\frac{da}{dt} = -\frac{A_1 n a^2}{(1 - e^2)^{3/2}} (1 + 2e \cos f + e^2)^{3/2} \rho, \quad (4.801)$$

$$\frac{de}{dt} = -\frac{A_1 n a}{(1 - e^2)^{1/2}} (\cos f + e) (1 + 2e \cos f + e^2)^{1/2} \rho, \quad (4.802)$$

$$\frac{di}{dt} = -\frac{A_2 a}{4(1 - e^2)} \sin i \left(\frac{r}{a} \right)^2 (1 + \cos 2u) (1 + 2e \cos f + e^2)^{1/2} \rho, \quad (4.803)$$

$$\frac{d\Omega}{dt} = -\frac{A_2 a}{4(1-e^2)} \left(\frac{r}{a}\right)^2 \sin 2u (1 + 2e \cos f + e^2)^{1/2} \rho, \quad (4.804)$$

$$\frac{d\omega}{dt} = -\cos i \frac{d\Omega}{dt} - \frac{A_1 n a}{e(1-e^2)^{1/2}} \sin f (1 + 2e \cos f + e^2)^{1/2} \rho, \quad (4.805)$$

$$\frac{dM}{dt} = n + \frac{A_1 n a}{e(1-e^2)} \left(\frac{r}{a}\right) \sin f (1 + e \cos f + e^2)(1 + 2e \cos f + e^2)^{1/2} \rho, \quad (4.806)$$

In (4.803) and (4.804) there is $u = f + \omega$.

The atmosphere density distribution model used in this book is given by (4.784) that

$$\rho = \rho_0 (1 + F^* \cos \psi) \exp\left(-\frac{r - \sigma}{H(r)}\right). \quad (4.807)$$

Under the influence of a dissipative force like the atmospheric drag, the energy of satellite motion reduces, i.e., both the semi-major axis and eccentricity of the orbit become smaller. The basic effect of the atmospheric rotation is to change the position of the orbital plane, as seen in (4.803) that the orbital inclination i decreases. The variation range of the inclination can be estimated by (4.803), (4.801), and the relationship of n and a , that

$$0 \geq \frac{di}{dn} \geq -\frac{n_e a}{3n(n - n_e)}. \quad (4.808)$$

Integrating (4.808) from height h_1 (corresponding to n_1) to a lower height h_0 (corresponding to n_0) leads

$$\Delta i \geq \frac{1}{3} \ln \frac{n_0(n_1 - n_e)}{n_1(n_0 - n_e)}. \quad (4.809)$$

If we choose $h_0 = 100$ km (a satellite descends to 100 km falls quickly), and $h_1 = 1000, 1200$, and 1400 km, the corresponding changes of i would be

$$\Delta i \approx \begin{cases} -0^\circ.27, & h_1 = 1000 \text{ km}, \\ -0^\circ.33, & h_1 = 1200 \text{ km}, \\ -0^\circ.39, & h_1 = 1400 \text{ km}. \end{cases} \quad (4.810)$$

These estimates suggest that during the lifespan of a satellite, its orbital inclination can reduce by about 0.3–0.4 due to the atmospheric rotation. This property provides important information for satellite orbit determinations, forecasts, and control.

4.11.4.2 The Atmospheric Density Model Used for Constructing the Perturbation Solution Due to Atmospheric Drag

We start with the oblate spheroid density model given by (4.783) and assume the density scale height H to be a constant. The initial perigee of a satellite orbit is the reference point p_0 and the spheroid passing point p_0 is the reference spheroid. The radial distance of an arbitrary point on this spheroid and Earth's center denoted by σ is given by

$$\sigma = \sigma_0 [1 - \varepsilon \sin^2 \varphi_{p_0} + O(e^2)], \quad (4.811)$$

where σ_0 is the radius of the “equator” (i.e., the semi-major axis of the ellipse), φ is the geographic latitude, ε is Earth's flattening factor, in the WGS84 system $\varepsilon = 0.00335281$. Since the reference spheroid passes the initial perigee p_0 , then

$$\sigma_{p_0} = \sigma_0 (1 - \varepsilon \sin^2 \varphi_{p_0}) = r_{p_0}.$$

Substituting this formula into (4.811) yields

$$\sigma = r_{p_0} (1 - \varepsilon \sin^2 \varphi) (1 - \varepsilon \sin^2 \varphi_{p_0})^{-1}. \quad (4.812)$$

Then substituting (4.812) into (4.783) yields

$$\rho = \rho_{p_0} \exp \left[-\frac{r - r_{p_0}}{H} + C \cos 2(f + \omega) - C \cos 2\omega_0 + O(Ce) \right], \quad (4.813)$$

$$C = \frac{1}{2} \left(\frac{\varepsilon}{H} r_{p_0} \right) \sin^2 i, \quad (4.814)$$

where C is called the flattening parameter, usually $C \approx 0.1$. For constructing the perturbation solution, it is necessary to express (4.813) as

$$\rho = k \exp \left[\frac{ae}{H} \cos E + C \cos 2(f + \omega) + O(Ce) \right], \quad (4.815)$$

$$k = \rho_{p_0} \exp \left[-\frac{1}{H} (a - a_0 + a_0 e_0) - C \cos 2\omega_0 \right]. \quad (4.816)$$

In the construction of the perturbation solution, the true anomaly f should be expressed by the eccentric anomaly E . Using relationships

$$r = a(1 - e \cos E), \quad r \cos f = a(\cos E - e), \quad r \sin f = a\sqrt{1 - e^2} \sin E, \quad (4.817)$$

we have

$$\begin{aligned}
\rho = k \bigg\{ & \left(1 + \frac{1}{4}C^2\right) + C \cos 2\omega(-e \cos E + \cos 2E + e \cos 3E) \\
& + C \sin 2\omega(e \sin E - \sin 2E - e \sin 3E) \\
& + C^2 \cos 4\omega\left(\frac{1}{4} \cos 4E\right) + C^2 \sin 4\omega\left(-\frac{1}{4} \sin 4E\right) \\
& + O(C\varepsilon) + O(Ce^2) + O(C^2e) + O(C^3) \bigg\} \exp\left(\frac{ae}{H} \cos E\right)
\end{aligned} \tag{4.818}$$

When $e < 0.2$, the terms omitted in (4.818) have magnitudes

$$O(C\varepsilon) = 3 \times 10^{-4}, \quad O(Ce^2) = O(C^2e) = O(C^3) = 10^{-3}.$$

Now we consider the variation of the density scale height H and the diurnal variation of the oblate spheroid atmosphere. As provided above that $\hat{r}$ and $\hat{r}_m$ are the unit vectors of $\vec{r}$ and $\vec{r}_m$, respectively, and there is $\cos \psi = \hat{r} \cdot \hat{r}_m$; the relationship between the direction of the peak density and the direction of the Sun is given by (4.786), which means in the geocentric equatorial frame the Sun's right ascension $\Omega' = 0$, and the direction of the daily peak corresponds to $\Omega' = \lambda_m$. This is similar to the expression (4.703) for the solar radiation pressure perturbation, so we write

$$\cos \psi = \hat{r} \cdot \hat{r}_m = A^* \cos f + B^* \sin f, \tag{4.819}$$

where A^* and B^* are similar to A and B in (4.703) except that θ in (4.703) now is replaced by $\theta = \Omega - \lambda_m$. The expressions of A and B are given by (4.498) and (4.499), respectively, in Sect. 4.7.1. The third body is the Sun so in the original expressions of A and B , the variables i' , $\theta = \Omega - \Omega'$, and u are actually $i' = \varepsilon$, $\Omega' = 0$, $\theta = \Omega$, and $u' = f' + \omega'$.

Based on the assumption about $H(r)$ given by (4.781), and (4.783) the atmospheric density model (4.784) can be further expressed as

$$\rho = \rho_0 \left(1 + F^* A^* \cos f + F^* B^* \sin f\right) \left[1 + \frac{\mu}{2} \left(\frac{r - \sigma}{H_{p_0}}\right)\right] \exp\left(-\frac{r - \sigma}{H_{p_0}}\right). \tag{4.820}$$

It is difficult to determine atmospheric parameters. Since we know $C \approx 0.1$, $\mu \approx 0.1$, and $e < 0.2$, then in the density model we usually only keep the terms with magnitudes of $O(e, C, \mu)$ and greater. This treatment is reasonable and satisfies general accuracy requirements. By the relationship of the satellite's distance r , the true anomaly f , and the eccentric anomaly E (4.817), the density model (4.820) can be simplified to a practically applicable form as

$$\rho = k \left\{1 + C \cos 2\omega \cos 2E - C \sin 2\omega \sin 2E + \Delta(\mu) + \Delta(F^*)\right\} \exp(z \cos E), \tag{4.821}$$

$$k = \rho_{p_0} \exp \left[-\frac{1}{H_{p_0}} (a - a_0 + a_0 e_0) - C \cos 2\omega_0 \right], \quad (4.822)$$

where the parameter z is defined by

$$z = \frac{ae}{H_{p_0}}. \quad (4.823)$$

The key auxiliaries in (4.821) are defined as

$$C = \frac{1}{2} \left(\frac{\varepsilon}{H_{p_0}} r_{p_0} \right) \sin^2 i, \quad (4.824)$$

$$\Delta(\mu) = \mu \left[\frac{3}{4} z_0^2 - z_0^2 \cos E + \frac{1}{4} z_0^2 \cos 2E \right], \quad (4.825)$$

$$\begin{aligned} \Delta(F^*) = & F^* A^* \left[-\left(\frac{e}{2} + \frac{1}{2} \mu z_0^2 \right) + \left(1 + \frac{7}{8} \mu z_0^2 \right) \cos E + \left(\frac{e}{2} - \frac{1}{2} \mu z_0^2 \right) \cos 2E \right. \\ & + \left. \left(\frac{1}{8} \mu z_0^2 \right) \cos 3E + \frac{C}{2} \cos 2\omega (\cos E + \cos 3E) + \frac{C}{2} \sin 2\omega (-\sin E - \sin 3E) \right] \\ & + F^* B^* \left[\left(1 + \frac{5}{8} \mu z_0^2 \right) \sin E + \left(\frac{e}{2} - \frac{1}{2} \mu z_0^2 \right) \sin 2E + \left(\frac{1}{8} \mu z_0^2 \right) \sin 3E \right. \\ & + \left. \frac{C}{2} \cos 2\omega (-\sin E + \cos 3E) + \frac{C}{2} \sin 2\omega (-\cos E + \cos 3E) \right] \quad (4.826) \end{aligned}$$

The model (4.821) includes the main variations of Earth's atmospheric density. It can be used to study the basic characteristics of the perturbation on satellite orbital motion due to the atmospheric drag and the influences of different parameters. In the following sections, we give the perturbation solution due to the atmospheric drag using this density model. In the solution, we only keep the terms with magnitudes of $O(C, \mu, \Delta(F^*))$ (about 10^{-1}). The terms related to these parameters in the order of 10^{-2} can be omitted because the error of the model itself is 5%–10% (in the order of 10^{-1}). The resulting solution shows not only the characteristics of a dissipative perturbation force on satellite orbits but also reaches certain accuracy (with respect to the accuracy of the model) requirements.

4.11.4.3 The Perturbation Solution in Kepler Orbital Elements

The ratio of the perturbing acceleration due to atmospheric drag D to the acceleration F_0 by the central gravity force of Earth's mass is

$$\varepsilon = \frac{D}{F_0} = \frac{1}{2} \left(\frac{C_D S}{m} \right) \rho v^2 / \left(\frac{1}{r^2} \right) \approx \left(\frac{C_D S}{m} \rho \right) \frac{r}{2}. \quad (4.827)$$

For a spherical satellite, whose diameter is 3 m and weight is 1 t, moves at an altitude of 200 km, if the atmospheric density is about 10^{-10} kg/m³, then $\varepsilon = O(10^{-6})$. If the altitude is higher, the atmospheric drag is smaller. By this estimation when we construct the perturbation solution, the perturbation of the atmospheric drag can be treated as a second-order perturbation, $O(J_2^2)$, and the solution has the form as

$$\begin{cases} \sigma_2(t - t_0) = \bar{f}(t - t_0) \\ \sigma_s^{(2)}(t) = \int^t \left(\frac{d\sigma}{dt}\right) dt = \int^E \left(\frac{d\sigma}{dE}\right) dE \end{cases} \quad (4.828)$$

Calculations of both $\bar{f}$ and $\sigma_s^{(2)}(t)$ need the following integrations

$$\begin{cases} \int \sin mE \exp(z \cos E) dE \\ \int \cos mE \exp(z \cos E) dE \end{cases}$$

By the expansions of exponential function and the trigonometric expressions of $\sin^\alpha E$ and $\cos^\alpha E$ (α is a positive integer) derived using double-angle formulas we have

$$\exp(z \cos E) = \sum_{\alpha \geq 0} \sum_{\beta=0}^{\frac{1}{2}(\alpha - \delta_1)} \left(\frac{z^\alpha}{\alpha!}\right) \frac{1}{2^{\alpha - \delta_2}} \binom{\alpha}{\beta} \cos(\alpha - 2\beta)E, \quad (4.829)$$

where

$$\binom{\alpha}{\beta} = \frac{\alpha!}{\beta!(\alpha - \beta)!}, \quad (4.830)$$

$$\delta_1 = \frac{1}{2}[1 - (-1)^\alpha], \quad \delta_2 = \begin{cases} 0, & \alpha - 2\beta = 0 \\ 1, & \alpha - 2\beta \neq 0 \end{cases} \quad (4.831)$$

Then we have

$$\frac{1}{2\pi} \int_0^{2\pi} \sin mE \exp(z \cos E) dE = 0, \quad (4.832)$$

$$\begin{aligned} & \frac{1}{2\pi} \int_0^{2\pi} \cos mE \exp(z \cos E) dE \\ &= \sum_{\alpha(2) \geq m} \frac{1}{[\frac{1}{2}(\alpha - m)]! [\frac{1}{2}(\alpha + m)]!} \left(\frac{z}{2}\right)^\alpha = \sum_{k \geq 0} \frac{1}{k!(m + k)!} \left(\frac{z}{2}\right)^{m+2k} = I_m(z) \end{aligned} \quad (4.833)$$

On the right side of (4.883) $I_m(z)$ is the modified Bessel function of the first kind. In the operation of the sum the sub $\alpha(2) \geq m$ means the value of α is from m and increases by a step-length of 2.

The short-period terms involve the following two integrals

$$\begin{aligned} & \int^E \sin m E \exp(z \cos E) dE \\ &= \sum_{\alpha \geq 0} \sum_{\beta=0}^{\frac{1}{2}(\alpha-\delta_1)} \left(\frac{z^\alpha}{\alpha!} \right) \frac{1}{2^{\alpha+1-\delta_2}} \left(\frac{\alpha}{\beta} \right) \left[\begin{aligned} & \delta_3 \frac{\cos(\alpha-2\beta-m)E}{\alpha-2\beta-m} \\ & - \frac{\cos(\alpha-2\beta+m)E}{\alpha-2\beta+m} \end{aligned} \right] \end{aligned} \quad (4.834)$$

$$\begin{aligned} & \int^E \cos m E \exp(z \cos E) dE \\ &= \sum_{\alpha \geq 0} \sum_{\beta=0}^{\frac{1}{2}(\alpha-\delta_1)} \left(\frac{z^\alpha}{\alpha!} \right) \frac{1}{2^{\alpha+1-\delta_2}} \left(\frac{\alpha}{\beta} \right) \left[\delta_3 \frac{\sin(\alpha-2\beta-m)E}{\alpha-2\beta-m} - \frac{\sin(\alpha-2\beta+m)E}{\alpha-2\beta+m} \right] \end{aligned} \quad (4.835)$$

where

$$\delta_3 = \begin{cases} 0, & \alpha - 2\beta - m = 0 \\ 1, & \alpha - 2\beta - m \neq 0 \end{cases} \quad (4.836)$$

Given these integrals, it is easy to construct the perturbation solution of the atmospheric drag. Similar to the method used for other perturbations we use the quasi-mean orbital elements to develop the solution. The long-period terms are treated as secular terms, together with secular terms denoted by $\sigma_2(t - t_0)$. Omitting the process, the results are

$$\begin{aligned} a_2 = -B_1 a^2 n \left\{ (I_0 + 2eI_1) + C(\cos 2\omega I_2) + \mu z_0^2 \left(\frac{3}{4}I_0 - I_1 + \frac{1}{4}I_2 \right) \right. \\ \left. + F^* A^* \left(\frac{e}{2}I_0 + I_1 + \frac{3}{2}eI_2 \right) \right\} \end{aligned} \quad (4.837)$$

$$\begin{aligned} e_2 = -B_1 a n \left\{ \left(\frac{e}{2}I_0 + I_1 + \frac{e}{2}I_2 \right) + \frac{C}{2} \cos 2\omega (I_1 + I_3) \right. \\ \left. + \mu z_0^2 \left(-\frac{1}{2}I_0 + \frac{7}{8}I_1 - \frac{1}{2}I_2 + \frac{1}{8}I_3 \right) + F^* A^* \left(\frac{1}{2}I_0 + \frac{e}{2}I_1 + \frac{1}{2}I_2 + \frac{e}{2}I_3 \right) \right\} \end{aligned} \quad (4.838)$$

$$i_2 = -\frac{1}{4} B_2 a \sin i (I_0 + \cos 2\omega I_2), \quad (4.839)$$

$$\Omega_2 = -\frac{1}{4} B_2 a \sin 2\omega I_2, \quad (4.840)$$

$$\omega_2 = -\cos i \Omega_2$$

$$\begin{aligned} & -B_1 a n \left(\frac{1}{e} \right) \left\{ C \sin 2\omega \left[\frac{e}{4}I_0 - \frac{1}{2}I_1 - eI_2 + \frac{1}{2}I_3 + \frac{3}{4}eI_4 \right] \right. \\ & \left. + F^* B^* \left[\left(\frac{1}{2} + \frac{e^2}{16} \right) I_0 + \frac{e}{2}I_1 - \left(\frac{1}{2} - \frac{e^2}{4} \right) I_2 - \frac{e}{2}I_3 - \frac{5}{16}e^2 I_4 - \frac{e^2}{4}(I_0 - I_2) \right] \right\} \end{aligned} \quad (4.841)$$

$$M_2 = -(\omega_2 + \cos i \Omega_2) + B_1 an \left\{ F^* B^* \left[\frac{e}{4} (I_0 - I_2) \right] \right\} - \frac{1}{2} \left(\frac{3n}{2a} \right) a_2 (t - t_0) \quad (4.842)$$

In the expressions (4.837)–(4.842) I_m is the modified Bessel function of the first kind, given by (4.833). The parameters μ , F^* , and related A^* and B^* are the same as above defined, and the new parameters B_1 , B_2 , C , and F are given by

$$B_1 = \left(\frac{C_D S}{m} \right) \bar{\rho}_{p_0} F^2 \exp \left(-\frac{1}{H_{p_0}} (a - a_0 + a_0 e_0) - C \cos 2\omega_0 \right), \quad (4.843)$$

$$B_2 = \left(\frac{C_D S}{m} \right) \bar{\rho}_{p_0} F n_e \exp \left(-\frac{1}{H_{p_0}} (a - a_0 + a_0 e_0) - C \cos 2\omega_0 \right), \quad (4.844)$$

$$F = 1 - \frac{r_{p_0} n_e}{v_{p_0}} \cos i_0, \quad (4.845)$$

$$C = \frac{1}{2} \left(\frac{\varepsilon}{H_{p_0}} r_{p_0} \right) \sin^2 i_0, \quad (4.846)$$

where $\varepsilon = 1/298.257$ is Earth's geometric flattening factor.

On the right sides of the above formulas, the slow-changing variables Ω and ω , and the Sun's mean ecliptic longitude $u' = L_\odot$ in A^* and B^* take their middle values $\bar{\sigma}_{1/2}$ over the time interval of $(t - t_0)$, that

$$\begin{cases} \Omega = \bar{\Omega}_0 + \frac{1}{2} \Omega_1 (t - t_0) \\ \omega = \bar{\omega}_0 + \frac{1}{2} \omega_1 (t - t_0) \\ L_\odot = L_0 + \frac{1}{2} n' (t - t_0) \end{cases} \quad (4.847)$$

where Ω_1 and ω_1 are given by (4.65) and (4.66); $n' = 0^\circ.9856/\text{d}$ is the Sun's mean angular speed.

Finally, the selections of some parameters require some explanations. There are four atmospheric parameters in the above perturbation solution, ρ_{p_0} , H_{p_0} , μ , f^* , which are needed to be decided. Among them ρ_{p_0} and H_{p_0} are given by the values at the original perigee; the values of μ and f^* , strictly speaking, vary with altitude, but according to the characteristic of atmospheric density variation, the values can be approximately given by the values at the original perigee. The height h_{p_0} at the original perigee p_0 depends on original values of the orbital elements. The original orbital elements may be the instantaneous elements, a_0 , e_0 , i_0 , and ω_0 , or the mean elements, $\bar{a}_0$, $\bar{e}_0$, $\bar{i}_0$, $\bar{\omega}_0$, then with the given Earth's flattening factor the value of h_{p_0} is given by

$$\begin{aligned} h_{p_0} &= a_0(1 - e_0) - (1 - \varepsilon \sin^2 i_0 \sin^2 \omega_0) \\ h_{p_0} &= \bar{a}_0(1 - \bar{e}_0) - (1 - \varepsilon \sin^2 \bar{i}_0 \sin^2 \bar{\omega}_0) \end{aligned} \quad (4.848)$$

Note that the reference values of a_0 , e_0 , i_0 , and ω_0 used in the perturbation formulas should be consistent with the values for calculating h_{p_0} , these reference values affect neither the variations of the orbital elements nor the method of the perturbation calculations.

4.11.5 The Non-singularity Perturbation Solution by the Atmospheric Drag

(1) The non-singularity perturbation solution of the first type

Similar to the method applied to other perturbation forces the non-singularity solution is constructed using the quasi-orbital elements. The long-period terms are treated as secular terms, together are denoted by $\sigma_2(t - t_0)$, the expressions of σ_2 are

$$a_2 = -B_1 a^2 n \left\{ (I_0 + 2eI_1) + C(\cos 2\omega I_2) + \mu z_0^2 \left(\frac{3}{4}I_0 - I_1 + \frac{1}{4}I_2 \right) + F^* A^* \left(\frac{e}{2}I_0 + I_1 + \frac{3}{2}eI_2 \right) \right\} \quad (4.849)$$

$$i_2 = -\frac{1}{4}B_2 a \sin i (I_0 + \cos 2\omega I_2), \quad (4.850)$$

$$\Omega_2 = -\frac{1}{4}B_2 a \sin 2\omega I_2, \quad (4.851)$$

$$\xi_2 = \cos \omega [e_2] - \sin \omega [e\omega_2], \quad (4.852)$$

$$\eta_2 = \sin \omega [e_2] + \cos \omega [e\omega_2], \quad (4.853)$$

$$\lambda_2 = -\cos i \Omega_2 + B_1 a n \left\{ F^* B^* \left[\frac{e}{4}(I_0 - I_2) \right] \right\} - \frac{1}{2} \left(\frac{3n}{2a} \right) a_2 (t - t_0), \quad (4.854)$$

where e_2 and ω_2 are given by (4.838) and (4.841), respectively.

(2) The non-singularity perturbation solution of the second type

$$a_2 = -B_1 a^2 n \left\{ (I_0 + 2eI_1) + C(\cos 2\omega I_2) + \mu z_0^2 \left(\frac{3}{4}I_0 - I_1 + \frac{1}{4}I_2 \right) + F^* A^* \left(\frac{e}{2}I_0 + I_1 + \frac{3}{2}eI_2 \right) \right\} \quad (4.855)$$

$$\xi_2 = \cos \tilde{\omega}[e_2] - \sin \tilde{\omega}[e\omega_2 + e\Omega_2], \quad (4.856)$$

$$\eta_2 = \sin \tilde{\omega}[e_2] + \cos \tilde{\omega}[e\omega_2 + e\Omega_2], \quad (4.857)$$

$$h_2 = \frac{1}{2} \cos \frac{i}{2} \cos \Omega[i_2] - \sin \frac{i}{2} \sin \Omega[\Omega_2], \quad (4.858)$$

$$k_2 = \frac{1}{2} \cos \frac{i}{2} \sin \Omega[i_2] + \sin \frac{i}{2} \cos \Omega[\Omega_2], \quad (4.859)$$

$$\lambda_2 = \frac{\sin i}{1 + \cos i} (\sin i \Omega_2) + B_{1an} \left\{ F^* B^* \left[\frac{e}{4} (I_0 - I_2) \right] \right\} - \frac{1}{2} \left(\frac{3n}{2a} \right) a_2 (t - t_0), \quad (4.860)$$

where e_2 , i_2 , Ω_2 , and ω_2 are given by (4.838)–(4.841).

4.12 Orbital Variations Due to a Small Thruster

In recent years, the development of aerospace technology has made the technique of small thruster reach its maturity. This technique was first used by the Moon prober SMART-1 of the European Space Agency in September 2003. In this book, we do not discuss the details of the technique of small thruster. Our goal is to theoretically analyze how the thrusting process changes the orbital motion of a satellite.

Assuming the weight of a spacecraft to be 1000 kg, during the time it is in its orbit it also receives an onboard thrust. If the thrust is 1 N (1 N = 1 kg m/s²), then the related thrust acceleration is 0.001 m/s². For comparison, the accelerations of Earth's central gravity at altitudes $h = 200$ and 300 km are

$$g(\text{m/s}^2) = \begin{cases} 9.211534, & h = 200.0 \text{ km}, \\ 8.937728, & h = 300.0 \text{ km}. \end{cases}$$

Therefore, the relative magnitude of the thrust is in the order of 10^{-4} . By this estimation, the effect on the satellite's motion caused by the thrust can be regarded as a small perturbation. We discuss two types of the thrust and derive corresponding perturbation models for the orbital motion of a spacecraft. The two types of the thrust are as follows.

- (1) The thrust has three components in the radial, transverse, and normal directions, and the resulting accelerations are denoted to S , T , and W , as the (S, T, W) -type thrust.
- (2) The thrust only acts in the tangential direction, the acceleration of the force is denoted by U , as the U -type thrust.

We analyze the effect only by the thrust perturbation without any other perturbations including the effects of the mixed terms of the J_2 and the thrust perturbation [15].

4.12.1 The Perturbation Solution Due to an (S,T,W)-Type Thrust

The system of perturbation equations of the spacecraft due to the thrust is

$$\left\{ \begin{array}{l} \frac{da}{dt} = \frac{2}{n\sqrt{1-e^2}} [Se \sin f + T(1 + e \cos f)] \\ \frac{de}{dt} = \frac{\sqrt{1-e^2}}{na} [S \sin f + T(\cos f + \cos E)] \\ \frac{di}{dt} = \frac{r \cos u}{na^2 \sqrt{1-e^2}} W \\ \frac{d\Omega}{dt} = \frac{r \sin u}{na^2 \sqrt{1-e^2} \sin i} W \\ \frac{d\omega}{dt} = \frac{\sqrt{1-e^2}}{nae} \left[-S \cos f + T \left(1 + \frac{r}{p} \right) \sin f \right] - \cos i \frac{d\Omega}{dt} \\ \frac{dM}{dt} = n - \frac{1-e^2}{nae} \left[-S \left(\cos f - 2e \frac{r}{p} \right) + T \left(1 + \frac{r}{p} \right) \sin f \right] \end{array} \right. \quad (4.861)$$

where $u = f + \omega$, $p = a(1 - e^2)$, f and E are the true anomaly and the eccentric anomaly, respectively.

The magnitude of the thrust acceleration is small that we can assume

$$S, T, W = O(\varepsilon).$$

Therefore, the perturbation equation system (4.861) can be expressed as a system of equations of a small parameter that

$$\dot{\sigma} = f_0(a) + f_{\text{thrust}}(\sigma, \varepsilon). \quad (4.862)$$

By the following averages

$$\left\{ \begin{array}{l} \overline{\left(\frac{r}{a} \right)} = 1 + \frac{1}{2}e^2 \\ \overline{\sin f} = 0, \quad \overline{\cos f} = -e \\ \overline{\left(\frac{r}{a} \right) \sin f} = 0, \quad \overline{\left(\frac{r}{a} \right) \cos f} = -\frac{3}{2}e \end{array} \right. \quad (4.863)$$

the functions on the right sides of (4.861) can be decomposed into secular, short-period, and long-period terms. We then construct the perturbation solution by the quasi-mean orbital elements. By integration, the solution has the form as

$$\begin{cases} \sigma(t) = \bar{\sigma}_0 + \delta\bar{n}(t - t_0) + \sigma_c(t - t_0) + \Delta\sigma_l(t) + \sigma_s(t), \\ \bar{\sigma}_0 = \sigma_0 - \sigma_s(t_0). \end{cases} \quad (4.864)$$

The secular, long-period, and short-period terms are as follows.

- (1) The secular variation rates of σ_c in the term $\sigma_c(t - t_0)$

$$a_c = 2\sqrt{1 - e^2}(T/n), \quad (4.865)$$

$$e_c = -\frac{3\sqrt{1 - e^2}}{2a}e(T/n), \quad (4.866)$$

$$i_c = 0, \quad (4.867)$$

$$\Omega_c = 0, \quad (4.868)$$

$$\omega_c = \frac{\sqrt{1 - e^2}}{a}(S/n), \quad (4.869)$$

$$M_c = -\left(\frac{3}{a}\right)(S/n) - \left(\frac{3n}{4a}\right)a_c(t - t_0), \quad (4.870)$$

where $n = \sqrt{\mu}a^{-3/2} = \bar{a}^{-3/2}$ is the mean orbital angular speed of the spacecraft.

- (2) Long-period variation rates of σ_l in the term $\Delta\sigma_l(t - t_0) = \sigma_l(t - t_0)$

$$a_l = 0, \quad (4.871)$$

$$e_l = 0, \quad (4.872)$$

$$i_l = -\frac{3}{2a\sqrt{1 - e^2}}e\cos\bar{\omega}_0(W/n), \quad (4.873)$$

$$\Omega_l = -\frac{3}{2a\sqrt{1 - e^2}\sin i}e\sin\bar{\omega}_0(W/n), \quad (4.874)$$

$$\omega_l = -\cos i\Omega_l, \quad (4.875)$$

$$M_l = 0. \quad (4.876)$$

Under the acting of a thrust, the long-period terms do not have any unusual variations, so we just treat the long-period terms as secular terms.

(3) Short-period terms of $\sigma_s(t)$

$$a_s(t) = \frac{2}{n^2} \left[-S e \left(\cos E + \frac{e}{2} \right) + T \sqrt{1-e^2} e \sin E \right], \quad (4.877)$$

$$e_s(t) = \frac{\sqrt{1-e^2}}{n^2 a} \left\{ -S \sqrt{1-e^2} \left(\cos E + \frac{e}{2} \right) + T \left[\left(2 - \frac{3}{2} e^2 \right) \sin E - \frac{e}{4} \sin 2E \right] \right\} \quad (4.878)$$

$$i_s(t) = \frac{W}{n^2 a \sqrt{1-e^2}} \left\{ \left[\left(1 - \frac{e^2}{2} \right) \sin E - \frac{e}{4} \sin 2E \right] \cos \omega + \sqrt{1-e^2} \left[\cos E - \frac{e}{4} \cos 2E \right] \sin \omega \right\}, \quad (4.879)$$

$$\Omega_s(t) = \frac{W}{n^2 a \sqrt{1-e^2} \sin i} \left\{ \left[\left(1 - \frac{e^2}{2} \right) \sin E - \frac{e}{4} \sin 2E \right] \sin \omega - \sqrt{1-e^2} \left[\cos E - \frac{e}{4} \cos 2E \right] \cos \omega \right\}, \quad (4.880)$$

$$\begin{aligned} \omega_s(t) = & -\cos i \Omega_s(t) \\ & - \frac{1}{n^2 a e} \left\{ S (1-e^2)^{3/2} \sin E \right. \\ & \left. + T \left[(2-e^2) \cos E - \frac{e}{4} \cos 2E \right] \right\} \end{aligned} \quad (4.881)$$

$$\begin{aligned} M_s(t) = & \frac{1}{n^2 a e} \left\{ S \left[\left(1 + 3e^2 - \frac{3}{2} e^4 \right) \sin E - \frac{5}{4} e^3 \sin 2E \right] \right. \\ & \left. + T \sqrt{1-e^2} \left[2(1+e^2) \cos E - \frac{e}{4} (1+3e^2) \cos 2E \right] \right\} \end{aligned} \quad (4.882)$$

The orbital elements in the perturbation solution are all quasi-mean elements $\bar{\sigma}(t)$.

4.12.2 The Non-singularity Perturbation Solution Due to an (S,T,W)-Type Thrust

(1) The non-singularity perturbation solution of the first type

The definitions of the non-singularity elements of the first type are given by (4.90) as

$$a, \quad i, \quad \Omega, \quad \xi = e \cos \omega, \quad \eta = e \sin \omega, \quad \lambda = M + \omega.$$

We still use the method of the quasi-mean elements to construct the perturbation solution. The long-period terms are treated the same as secular terms, $\sigma_{c,l}(t - t_0)$, and the variation rates of the secular and long-period terms are $\sigma_{c,l}$ given by

$$a_{c,l} = 2\sqrt{1 - e^2}(T/n), \quad (4.883)$$

$$i_{c,l} = -\frac{3}{2a\sqrt{1 - e^2}}e \cos \bar{\omega}_0(W/n), \quad (4.884)$$

$$\Omega_{c,l} = -\frac{3}{2a\sqrt{1 - e^2} \sin i}e \sin \bar{\omega}_0(W/n), \quad (4.885)$$

$$\xi_{c,l} = \cos \omega[(e_c + e_l)] - e \sin \omega[(\omega_c + \omega_l)], \quad (4.886)$$

$$\eta_{c,l} = \sin \omega[(e_c + e_l)] + e \cos \omega[(\omega_c + \omega_l)], \quad (4.887)$$

$$\begin{aligned} \lambda_{c,l} &= (\omega_c + \omega_l) + (M_c + M_l) \\ &= -\cos i \Omega_l - \left(\frac{1}{a}\right)(3 - \sqrt{1 - e^2})(S/n) - \left(\frac{3n}{4a}\right)a_c(t - t_0) \end{aligned} \quad (4.888)$$

where $(\omega_c + \omega_l)$ is given by (4.869) and (4.875) that

$$\omega_c + \omega_l = \frac{\sqrt{1 - e^2}}{a}(S/n) - \cos i \Omega_l. \quad (4.889)$$

For the eccentricity there is $e_{c,l} = e_c + e_l = e_c$, and e_c is given by (4.866) and $e_l = 0$ by (4.872).

The short-period terms $a_s(t)$, $i_s(t)$, $\Omega_s(t)$ are given by (4.877), (4.878), and (4.880), respectively, and $\sigma_s(t)$ for the other three elements are

$$\xi_s(t) = \cos \omega[e_s(t)] - \sin \omega[e\omega_s(t)], \quad (4.890)$$

$$\eta_s(t) = \sin \omega[e_s(t)] + \cos \omega[e\omega_s(t)], \quad (4.891)$$

$$\lambda_s(t) = M_s(t) + \omega_s(t), \quad (4.892)$$

where $e_s(t)$ is given by (4.878), and

$$e\omega_s(t) = -e \cos i \Omega_s(t)$$

$$-\frac{1}{n^2 a} \left\{ S(1-e^2)^{3/2} \sin + T \left[(2-e^2) \cos E - \frac{e}{4} \cos 2E \right] \right\} E \quad (4.893)$$

$$\begin{aligned} \omega_S(t) + M_S(t) = & -\cos i \Omega_S(t) \\ & + \frac{1}{n^2 a} S \left[\left(\frac{9}{2} - \frac{15}{8} e^2 \right) e \sin E - \frac{5}{4} e^2 \sin 2E \right] \\ & + \frac{1}{n^2 a} T \left[\left(2 - \frac{3}{4} e^2 \right) e \cos E - \frac{1}{4} \left(\frac{5}{2} - \frac{11}{8} e^2 \right) e^2 \cos 2E \right] \end{aligned} \quad (4.894)$$

(2) The non-singularity perturbation solution of the second type

By the definitions of the non-singularity variables of the second type in (4.147) we have

$$\begin{aligned} a, & \quad \xi = e \cos \tilde{\omega}, & \eta = e \sin \tilde{\omega}, \\ h = \sin \frac{i}{2} \cos \Omega, & \quad k = \sin \frac{i}{2} \sin \Omega & \lambda = M + \tilde{\omega}. \end{aligned}$$

Similarly using the method of the quasi-mean elements to construct the perturbation solution, the long-period terms are treated the same as secular terms, as $\sigma_{c,l}(t - t_0)$, and the variation rates of the secular and long-period terms are $\sigma_{c,l}$ given by

$$a_{c,l} = 2\sqrt{1-e^2}(T/n), \quad (4.895)$$

$$\xi_{c,l} = \cos \tilde{\omega}[e_{c,l}] - e \sin \tilde{\omega}[\omega_{c,l} + \Omega_{c,l}], \quad (4.896)$$

$$\eta_{c,l} = \sin \tilde{\omega}[e_{c,l}] + e \cos \tilde{\omega}[\omega_{c,l} + \Omega_{c,l}], \quad (4.897)$$

$$h_{c,l} = \frac{1}{2} \cos \frac{i}{2} \cos \Omega[i_{c,l}] - \sin \Omega \left[\sin \frac{i}{2} \Omega_{c,l} \right], \quad (4.898)$$

$$k_{c,l} = \frac{1}{2} \cos \frac{i}{2} \sin \Omega[i_{c,l}] + \cos \Omega \left[\sin \frac{i}{2} \Omega_{c,l} \right], \quad (4.899)$$

$$\begin{aligned} \lambda_{c,l} = & \Omega_{c,l} + \omega_{c,l} + M_{c,l} \\ = & -\frac{3}{2a\sqrt{1-e^2}} \left(\frac{\sin i}{1+\cos i} \right) e \sin \bar{\omega}_0(W/n) \\ & - \left(\frac{1}{a} \right) \left(3 - \sqrt{1-e^2} \right) (S/n) - \left(\frac{3n}{4a} \right) a_{c,l}(t - t_0), \end{aligned} \quad (4.900)$$

where $e_{c,l}$ is the same as e_c given by (4.866), $\Omega_{c,l} = (\Omega_c + \Omega_l)$ and $\omega_{c,l} = (\omega_c + \omega_l)$ are given by (4.885) and (4.889), respectively, also

$$\begin{aligned}\sin \frac{i}{2} \Omega_{c,l} &= \sin \frac{i}{2} \Omega_l \\ &= -\frac{3}{4a\sqrt{1-e^2} \cos \frac{i}{2}} e \sin \bar{\omega}_0(W/n).\end{aligned}\quad (4.901)$$

The short-period terms are

$$a_s(t) = \frac{2}{n^2} \left[-Se \left(\cos E + \frac{e}{2} \right) + T \sqrt{1-e^2} e \sin E \right], \quad (4.902)$$

$$\xi_s(t) = \cos \tilde{\omega}[e_s(t)] - \sin \tilde{\omega}[e\omega_s(t) + e\Omega_s(t)], \quad (4.903)$$

$$\eta_s(t) = \sin \tilde{\omega}[e_s(t)] + \cos \tilde{\omega}[e\omega_s(t) + e\Omega_s(t)], \quad (4.904)$$

$$h_s(t) = \frac{1}{2} \cos \frac{i}{2} \cos \Omega[i_s(t)] - \sin \Omega \left[\sin \frac{i}{2} \Omega_s(t) \right], \quad (4.905)$$

$$k_s(t) = \frac{1}{2} \cos \frac{i}{2} \sin \Omega[i_s(t)] + \cos \Omega \left[\sin \frac{i}{2} \Omega_s(t) \right], \quad (4.906)$$

$$\lambda_s(t) = M_s(t) + \omega_s(t) + \Omega_s(t), \quad (4.907)$$

where $e_s(t)$ and $i_s(t)$ are given by (4.878) and (4.879), respectively, also

$$\begin{aligned}e\omega_s(t) + e\Omega_s(t) &= e(1 - \cos i)\Omega_s(t) \\ &\quad - \frac{1}{n^2 a} \left\{ S(1 - e^2)^{3/2} \sin E + T \left[(2 - e^2) \cos E - \frac{e}{4} \cos 2E \right] \right\},\end{aligned}\quad (4.908)$$

$$\begin{aligned}\sin \frac{i}{2} \Omega_s(t) &= \frac{W}{2n^2 a \sqrt{1-e^2}} \left(\frac{1}{\cos \frac{i}{2}} \right) \left\{ \left[\left(1 - \frac{e^2}{2} \right) \sin E - \frac{e}{4} \sin 2E \right] \sin \omega \right. \\ &\quad \left. - \sqrt{1-e^2} \left[\cos E - \frac{e}{4} \cos 2E \right] \cos \omega \right\},\end{aligned}\quad (4.909)$$

$$\begin{aligned}M_s(t) + \omega_s(t) + \Omega_s(t) &= \frac{1}{n^2 a} S \left[\left(\frac{9}{2} - \frac{15}{8} e^2 \right) e \sin E - \frac{5}{4} e^2 \sin 2E \right] \\ &\quad + \frac{1}{n^2 a} T \left[\left(2 - \frac{3}{4} e^2 \right) e \cos E - \frac{1}{4} \left(\frac{5}{2} - \frac{11}{8} e^2 \right) e^2 \cos 2E \right] \\ &\quad + \frac{W}{n^2 a \sqrt{1-e^2}} \frac{\sin i}{(1 + \cos i)} \left\{ \left[\left(1 - \frac{e^2}{2} \right) \sin E - \frac{e}{4} \sin 2E \right] \sin \omega \right. \\ &\quad \left. - \sqrt{1-e^2} \left[\cos E - \frac{e}{4} \cos 2E \right] \cos \omega \right\}\end{aligned}\quad (4.910)$$

4.12.3 The Perturbation Solution by a U-type Thrust

The set of the perturbation equation of a spacecraft due to a U -type thrust is

$$\begin{cases} \frac{da}{dt} = \frac{2}{n\sqrt{1-e^2}}(1 + 2e \cos f + e^2)^{1/2}U \\ \frac{de}{dt} = \frac{\sqrt{1-e^2}}{na}(1 + 2e \cos f + e^2)^{-1/2}[2(\cos f + e)U] \\ \frac{di}{dt} = 0, \quad \frac{d\Omega}{dt} = 0 \\ \frac{d\omega}{dt} = \frac{\sqrt{1-e^2}}{nae}(1 + 2e \cos f + e^2)^{-1/2}[(2 \sin f)U] \\ \frac{dM}{dt} = n - \frac{1-e^2}{nae}(1 + 2e \cos f + e^2)^{-1/2}\left[\left(2 \sin f + \frac{2e^2}{\sqrt{1-e^2}} \sin E\right)U\right] \end{cases} \quad (4.911)$$

On the right sides of the equations, there are terms with a factor of $(1 + 2e \cos f + e^2)^{\pm 1/2}$. For solving this kind of equation, we have to deal with the series expanding problem, meaning that the perturbation solution of e cannot be expressed by a closed form. Therefore, we expand the right-side functions with trigonometric functions of the mean anomaly M up to e^2 , resulting

$$\frac{da}{dt} = \frac{2}{n}U \left[\left(1 - \frac{1}{4}e^2\right) + e \cos M + \frac{3}{4}e^2 \cos 2M \right], \quad (4.912)$$

$$\frac{de}{dt} = \frac{2}{na}U \left[-\frac{e}{2} + (1 - e^2) \cos M + \frac{e}{2} \cos 2M + \frac{e^2}{2} \cos 3M \right], \quad (4.913)$$

$$\frac{di}{dt} = 0, \quad (4.914)$$

$$\frac{d\Omega}{dt} = 0, \quad (4.915)$$

$$\frac{d\omega}{dt} = \frac{2}{nae}U \left[\left(1 - \frac{e^2}{2}\right) \sin M + \frac{e}{2} \sin 2M + \frac{e^2}{2} \sin 3M \right], \quad (4.916)$$

$$\frac{dM}{dt} = n - \frac{2}{nae}U \left[\sin M + \frac{e}{2} \sin 2M + \frac{e^2}{2} \sin 3M \right], \quad (4.917)$$

Then we use the quasi-mean elements to construct the perturbation solution, $\sigma(t)$, that

$$\begin{cases} \sigma(t) = \bar{\sigma}_0 + \delta \bar{n}(t - t_0) + \sigma_c(t - t_0) + \sigma_s(t), \\ \bar{\sigma}_0 = \sigma_0 - \sigma_s(t_0). \end{cases} \quad (4.918)$$

The secular and long-period terms are given as follows.

(1) The variation rates of the secular terms of σ_c in $\sigma_c(t - t_0)$ are

$$a_c = 2\left(1 - \frac{1}{4}e^2\right)(U/n), \quad (4.919)$$

$$e_c = -\frac{e}{a}e(U/n), \quad (4.920)$$

$$i_c = 0, \quad (4.921)$$

$$\Omega_c = 0, \quad (4.922)$$

$$\omega_c = 0, \quad (4.923)$$

$$M_c = -\frac{3}{2a}\left(1 - \frac{1}{4}e^2\right)U(t - t_0), \quad (4.924)$$

where $n = \sqrt{\mu}a^{-3/2} = \bar{a}^{-3/2}$ is the mean orbital angular speed of the spacecraft.

(2) Short-period terms of $\sigma_s(t)$

$$a_s(t) = \frac{2}{n^2}\left[e \sin M + \frac{3}{8}e^2 \sin 2M\right]U, \quad (4.925)$$

$$e_s(t) = \frac{2}{n^2a}\left[(1 - e^2) \sin M + \frac{e}{4} \sin 2M + \frac{e^2}{6} \sin 3M\right]U, \quad (4.926)$$

$$i_s(t) = 0, \quad (4.927)$$

$$\Omega_s(t) = 0, \quad (4.928)$$

$$\omega_s(t) = -\frac{2}{n^2ae}\left[\left(1 - \frac{e^2}{2}\right) \cos M + \frac{e}{4} \cos 2M + \frac{e^2}{6} \cos 3M\right]U, \quad (4.929)$$

$$\begin{aligned} M_s(t) = & \frac{2}{n^2ae}\left[\cos M + \frac{e}{4} \cos 2M + \frac{e^2}{6} \cos 3M\right]U \\ & + \left(\frac{3}{2a}\right)\frac{2}{n^2}\left[e \cos M + \frac{3}{16}e^2 \cos 2M\right]U \end{aligned} \quad (4.930)$$

The orbital elements in the perturbation solution are all quasi-mean elements $\bar{\sigma}(t)$.

4.12.4 The Non-singularity Perturbation Solution Due to a U-type Thrust

(1) The non-singularity perturbation solution of the first type

Using the same method as for the (S, T, W) -type thrust, the quasi-mean elements are defined by (4.90) as

$$a, \quad i, \quad \Omega, \quad \xi = e \cos \omega, \quad \eta = e \sin \omega, \quad \lambda = M + \omega.$$

By this set of quasi-mean elements, we construct the perturbation solution and the variation rates of the secular terms of σ_c in $\sigma_c(t - t_0)$ are

$$a_c = 2 \left(1 - \frac{1}{4} e^2 \right) \left(\frac{U}{n} \right), \quad (4.931)$$

$$i_c = 0, \quad (4.932)$$

$$\Omega_c = 0, \quad (4.933)$$

$$\xi_c = \cos \omega(e_c), \quad (4.934)$$

$$\eta_c = \sin \omega(e_c), \quad (4.935)$$

$$\lambda_c = M_c + \omega_c = -\frac{3}{2a} \left(1 - \frac{1}{4} e^2 \right) U(t - t_0), \quad (4.936)$$

where e_c is given by (4.920).

In $\sigma_s(t)$ the short-period terms $a_s(t)$, $i_s(t)$, and $\Omega_s(t)$ are given by (4.925), (4.927) and (4.928), respectively, and $\sigma_s(t)$ for the other three quasi-mean elements are

$$\xi_s(t) = \cos \omega[e_s(t)] - \sin \omega[e\omega_s(t)], \quad (4.937)$$

$$\eta_s(t) = \sin \omega[e_s(t)] + \cos \omega[e\omega_s(t)], \quad (4.938)$$

$$\lambda_s(t) = M_s(t) + \omega_s(t) = \frac{1}{n^2 a} \left[4e \cos M + \frac{9}{16} e^2 \cos 2M \right] U, \quad (4.939)$$

where $e_s(t)$ is given by (4.926), and also

$$e\omega_s(t) = -\frac{2}{n^2 a} \left[\left(1 - \frac{e^2}{2} \right) \cos M + \frac{e}{4} \cos 2M + \frac{e^2}{6} \cos 3M \right] U. \quad (4.940)$$

(2) The non-singularity perturbation solution of the second type

Definitions of the non-singularity variables of the second type given by (4.147) are

$$a, \quad \xi = e \cos \tilde{\omega}, \quad \eta = e \sin \tilde{\omega},$$

$$h = \sin \frac{i}{2} \cos \Omega, \quad k = \sin \frac{i}{2} \sin \Omega, \quad \lambda = M + \tilde{\omega}.$$

The same as before by the set of quasi-mean elements we construct the perturbation solution, and the variation rates of the secular terms of σ_c in $\sigma_c(t - t_0)$ are

$$a_c = 2 \left(1 - \frac{1}{4} e^2 \right) \left(\frac{U}{n} \right), \quad (4.941)$$

$$\xi_c = \cos \tilde{\omega}(e_c), \quad (4.942)$$

$$\eta_c = \sin \tilde{\omega}(e_c), \quad (4.943)$$

$$h_c = 0, \quad (4.944)$$

$$k_c = 0, \quad (4.945)$$

$$\lambda_c = M_c + \omega_c + \Omega_c = -\frac{3}{2a} \left(1 - \frac{1}{4} e^2 \right) U(t - t_0), \quad (4.946)$$

where e_c is given by (4.920). The short-period terms of $\sigma_s(t)$ are

$$a_s(t) = \frac{2}{n^2} \left[e \sin M + \frac{3}{8} e^2 \sin 2M \right] U, \quad (4.947)$$

$$\xi_s(t) = \cos \tilde{\omega}[e_s(t)] - \sin \tilde{\omega}[e\omega_s(t)], \quad (4.948)$$

$$\eta_s(t) = \sin \tilde{\omega}[e_s(t)] + \cos \tilde{\omega}[e\omega_s(t)], \quad (4.949)$$

$$h_s(t) = 0, \quad (4.950)$$

$$k_s(t) = 0, \quad (4.951)$$

$$\lambda_s(t) = M_s(t) + \tilde{\omega}_s(t) = \frac{1}{n^2 a} \left[4e \cos M + \frac{9}{16} e^2 \cos 2M \right] U, \quad (4.952)$$

where $e_s(t)$ and $e\omega_s(t)$ are given by (4.926) and (4.940), respectively.

References

1. Kozai Y (1960) Effect of Precession and Nutation on the Orbital Elements of a Close Earth Satellite. *Astron J* 65(10):621–623
2. Brouwer D, Clemence GM (1961) *Methods of Celestial Mechanics*. Academic Press, New York and London
3. Liu L (1992) *Orbital Dynamics of Artificial Earth Satellite*. Higher Education Press, Beijing
4. Liu L (1998) *Methods of Celestial Mechanics*. Nanjing University Press.
5. Liu L (2000) *Orbital Theory of Spacecraft*. National Defense Industry Press, Beijing
6. Liu L (1974) A Solution of the Motion of an Artificial Satellite in the Vicinity of the Critical Inclination. *ACTA Astronomica Sinica* 15(2): 230–240, and *China Astron. Astrophys.* 1977 1(1): 31–42.
7. Liu L (1975) A Method of Calculating the Perturbation of Artificial Satellites. *ACTA Astronomica Sinica* 16(1): 5–80, and *China Astron. Astrophys.* 1977 1(1): 63–78.
8. Liu L, Tang JS (2008) The additional perturbation of coordinate system and the selection of coordinate system in the Earth's satellite motion. *J. of Space Science, China* 28(2):164–168.
9. Liu L, Tang JS (2015) *Orbital Theory of Satellites and Applications*. Electronic Industry Press, Beijing
10. Liu L, Hou XY (2018) *The Basic of Orbit Theory*. Higher Education Press, Beijing
11. Huang C, Liu L (1992) Analytical solution to the four post-Newtonian effects in a near earth satellite orbit. *Celest Mech* 53(3):172–183
12. Hadjimichalis KS, Brandin CL (1974) Gas Dynamics. *Proceedings of the ninth International Symposium No.2*, D13.1-D13.9.
13. Kienappel KG, Koppenwallner G, Legge H (1972) Rarefied Gas Dynamics. *Proceedings of the Eighth International Symposium* 317–325.
14. Xia CY, Wu YY, Liu L (1982) Joint perturbation on a low Earth satellite orbit due to the variation of the resistant coefficient. *Acta Astronom Sinica* 231(2):175–184
15. Liu L, Hu SJ, Wang X (2006) *Introduction of Aerospace Dynamics*. Nanjing University Press, Nanjing, China

Chapter 5

Satellite Orbit Design and Orbit Lifespan Estimation



Generally speaking, orbital variations of a satellite are mainly due to the central body's non-spherical gravity potential except for low orbit satellites with relatively large area-to-mass ratios (not including the low Moon orbit satellite because the Moon has no atmosphere). For high orbit satellites (such as geosynchronous satellites) the third-body gravitational force is also an important perturbation source. Chapter 4 provides detailed analyses of these two gravitational perturbations as well as the corresponding perturbed orbital solutions (as small parameter power series). Although the perturbation solutions are approximate, they reveal the fundamental characteristics and regularities of satellite orbital variations. In this Chapter, we use some key details in the perturbation solutions to further describe orbital properties for some specific satellites with extensively applicable values, and also provide basic theories and calculation methods for comprehensive works of these kinds of satellite (including orbital design and controlling, etc.).

5.1 Sidereal Period and Nodal Period [1–3]

An orbital period of a satellite is an important parameter in the aspects of application. Perturbations change satellite orbits, as a result for a satellite, there are different orbit periods by definition and application. In this section, we introduce the sidereal period and the nodal period, and provide their rigorous definitions and the transfer relationship between them.

We denote an orbital period of a satellite to T_s that

$$T_s = \frac{2\pi}{n} = \frac{2\pi}{\sqrt{\mu}a^{-3/2}}, \quad (5.1)$$

where n is the mean angular speed of a satellite. Using the normalized dimensionless units described in Chap. 4, the dimensionless form of the orbital period is given by

$$T_s = 2\pi a^{3/2}, \quad (5.2)$$

For a two-body problem, this parameter is a fixed time interval on which a satellite completes one cycle around the central body. Because of perturbations after one cycle the satellite does not come back to the original position, therefore the concept of the period is only instantaneous. The orbital period given by formula (5.1) actually is a function of time, that $T_s = T_s(t)$. This instantaneous period is usually called the sidereal period, which cannot be measured directly. In reality, we often use a few “measurable” time intervals as periods of a satellite, particularly the time interval between successive passages of its sub-satellite through one of its intersections (usually the ascending node) with a reference latitude φ called the nodal period and denoted by T_φ . In the 1960s and 1970s, T_φ was often used in ground-based satellite tracking systems. If the reference latitude is the equator (i.e., $\varphi = 0$), then this period denoted by T_Ω is the nodal period defined in the modern-day aerospace industry. The sidereal and nodal periods are different due to perturbations. In this section, we provide the transformation between the two periods when the main perturbation of J_2 term is involved.

5.1.1 The Transformation Between the Sidereal Period T_s and the Nodal Period T_φ

Combining the area integral (2.7) or (2.15) in Chap. 2 and the definition of the angle θ in (3.12) in Chap. 3, under perturbations there is

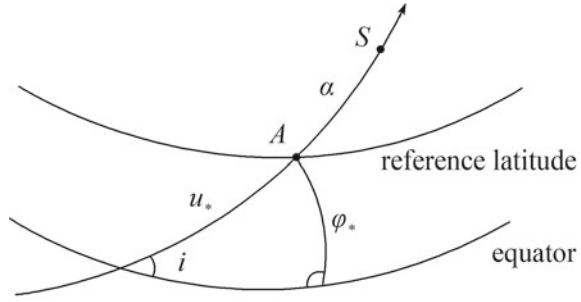
$$r^2(\dot{f} + \dot{\omega} + \cos i \dot{\Omega}) = \sqrt{p} = \sqrt{a(1 - e^2)}. \quad (5.3)$$

Let

$$f + \omega = u_* + \alpha, \quad (5.4)$$

where α is the angle measured from the intersection of the reference latitude (φ_*) and the satellite orbit, A, to the satellite S (Fig. 5.1). When α increases from 0 to 2π , the corresponding sub-satellite completes a cycle with respect to the latitude φ_* , this time interval is called a nodal period T_φ . As mentioned above, if the reference latitude is the equator (i.e., $\varphi_* = 0$), then the nodal period is denoted to T_Ω . But due to perturbations when α increases from 0 to 2π , although the satellite is back to the same latitude the intersection is no longer the previous point A and the related angle u_* is also changed. From the relationship

$$\sin u_* = \frac{\sin \varphi_*}{\sin i}, \quad (5.5)$$

Fig. 5.1 Auxiliary spherical surface

there is

$$\frac{du_*}{dt} = -\tan u_* \cot i \frac{di}{dt}. \quad (5.6)$$

Over a short time interval, such as one period (2π), to satisfy an accuracy of 10^{-5} for transferring one type of period to another, the perturbation terms in need of considering are the first-order secular term $\sigma_1(t - t_0)$ and the short-period term $\sigma_s^{(1)}(t)$ due to the J_2 perturbation. By the definition of the angle α in (5.4) and the area integral (5.3) and (5.6) we have

$$\frac{d\alpha}{dt} = \frac{df}{dt} + \frac{d\omega}{dt} - \frac{du_*}{dt} = \frac{\sqrt{p}}{r^2} - \cos i \frac{d\Omega}{dt} + \tan u_* \cot i \frac{di}{dt}. \quad (5.7)$$

Then the nodal period T_φ can be given by the following integration

$$T_\varphi = \int_0^{2\pi} \frac{dt}{d\alpha} d\alpha = \int_0^{2\pi} \left(\frac{\sqrt{p}}{r^2} - \cos i \frac{d\Omega}{dt} + \tan u_* \cot i \frac{di}{dt} \right)^{-1} d\alpha. \quad (5.8)$$

For the above-given accuracy requirement, the orbital elements in the second and third terms of the integrand can be the non-perturbed elements, but the term $\sqrt{p}/r^2$ should include perturbed variations.

Since we only consider the secular and the short-period terms of the first order due to the J_2 term, T_φ in (5.8) can be expressed as

$$\begin{aligned} T_\varphi &= \int_0^{2\pi} \left[\left(\frac{r^2}{\sqrt{p}} \right)_{\bar{\sigma}_0} + \Delta \left(\frac{r^2}{\sqrt{p}} \right) + \frac{r^4}{p} \left(\cos i \frac{d\Omega}{dt} - \tan u_* \cot i \frac{di}{dt} \right) \right] d\alpha \\ &= \bar{T}_s + \Delta \bar{T}_s, \end{aligned} \quad (5.9)$$

where $\bar{T}_s$ is the mean sidereal period that

$$\bar{T}_s = \int_0^{2\pi} \left(\frac{r^2}{\sqrt{p}} \right)_{\bar{\sigma}_0} d\alpha = 2\pi \bar{a}_0^{\frac{3}{2}}, \quad (5.10)$$

and the perturbed part is

$$\Delta \bar{T}_s = \int_0^{2\pi} \left[\Delta \left(\frac{r^2}{\sqrt{p}} \right) + \frac{r^4}{p} \left(\cos i \frac{d\Omega}{dt} - \tan u_* \cot i \frac{di}{dt} \right) \right] d\alpha. \quad (5.11)$$

To calculate the first part of the integrand of (5.11) we need $\sigma_s^{(1)}(t)$ and $\omega_1(t - t_0)$ given in Sect. 4.2, the result is

$$\begin{aligned} \Delta \left(\frac{r^2}{\sqrt{p}} \right) = & -\frac{3J_2}{2\sqrt{p}} \left[\left(1 - \frac{3}{2} \sin^2 i \right) \left(2 + \frac{3}{2} e^2 \right) + \left(2 - \frac{5}{2} \sin^2 i \right) (e^2) \right. \\ & + \left(\frac{1}{4} e^2 \cos^2 i \right) \cos 2\omega + \tan u_* \left(\frac{1}{4} e^2 \cos^2 i \right) \sin 2\omega \\ & \left. + \left(2 - \frac{5}{2} \sin^2 i \right) (2e \sin f - 3e^2 \sin 2f) \alpha \right] \end{aligned} \quad (5.12)$$

For the second part we need $(f_{1c} + f_{1s})_\Omega$ and $(f_{1s})_i$ in $d\Omega/dt$ and di/dt given in Sect. 4.2. After arranging the terms, we have

$$\begin{aligned} & \frac{r^4}{p} \left(\cos i \frac{d\Omega}{dt} - \tan u_* \cot i \frac{di}{dt} \right) \\ = & \frac{3J_2}{2\sqrt{p}} \cos^2 i \left[-\left(1 + \frac{1}{2} e^2 \right) + \frac{1}{4} e^2 (\cos 2\omega + \tan u_* \sin 2\omega) \right] \end{aligned} \quad (5.13)$$

Substituting (5.12) and (5.13) into (5.11), then integrating, we obtain $\Delta \bar{T}_s$. The nodal period then is given by (5.9) as

$$\begin{aligned} T_\varphi = \bar{T}_s \left\{ 1 + \frac{3J_2}{8a^2} [(-12 - 22e^2) + (16 + 29e^2) \sin^2 i \right. \\ \left. + (16 - 20 \sin^2 i) e \cos(u_* - \omega) - (12 - 15 \sin^2 i) e^2 \cos(2u_* - 2\omega)] \right\} \end{aligned} \quad (5.14)$$

If $\varphi_* = 0$ (i.e., $u_* = 0$), there is

$$\begin{aligned} T_\Omega = \bar{T}_s \left\{ 1 + \frac{3J_2}{8a^2} [(-12 - 22e^2) + (16 + 29e^2) \sin^2 i \right. \\ \left. + (16 - 20 \sin^2 i) e \cos \omega - (12 - 15 \sin^2 i) e^2 \cos 2\omega] \right\} \end{aligned} \quad (5.15)$$

The relationship between T_φ and T_Ω is

$$T_\varphi = T_\Omega \left\{ 1 + \frac{3J_2}{8a^2} \left[(16 - 20 \sin^2 i) e (\cos f_* - \cos \omega) - (12 - 15 \sin^2 i) e^2 (\cos 2f_* - \cos 2\omega) \right] \right\} \quad (5.16)$$

where $f_* = u_* - \omega$. When $e = 0$, there is $T_\varphi = T_\Omega$.

The above result provides the relationship between the nodal period and the mean sidereal period, $\bar{T}_s$. The sidereal period T_s is given by

$$T_s = \int_0^{2\pi} \left(\frac{r^2}{\sqrt{p}} \right)_{\sigma_0} d\alpha = 2\pi a_0^{3/2}.$$

With the relationship of $\bar{a}_0$ and a_0

$$\bar{a}_0 = a_0 - a_s^{(1)}(t_0),$$

we have the sidereal period as

$$T_s = \bar{T}_s \left[1 + \frac{3a_s^{(1)}(t_0)}{2\bar{a}_0} \right]. \quad (5.17)$$

Substituting (5.17) into (5.9) yields.

$$T_\varphi = T_s \left[1 - \frac{3a_s^{(1)}(t_0)}{2\bar{a}_0} \right] + \Delta \bar{T}_s. \quad (5.18)$$

Substituting $a_s^{(1)}(t_0)$ into (5.18) then omitting the terms smaller than $O(J_2 e^2)$ yields

$$\begin{aligned} T_\varphi = T_s - \frac{3J_2}{8a^2} T_s \{ & (12 + 34e^2) - (16 + 47e^2) \sin^2 i + (6 + 27e^2) \sin^2 i \cos 2u_* \\ & - [(4 + \sin^2 i) \sin u_* - 9 \sin^2 i \sin 3u_*] e \sin \omega \\ & - [(4 - 11 \sin^2 i) \cos u_* - 9 \sin^2 i \cos 3u_*] e \cos \omega \\ & + [(18 - 24 \sin^2 i) \cos u_* + 9 \sin^2 i \cos 2u_*] e^2 \cos 2(u_* - \omega) \} \end{aligned} \quad (5.19)$$

and

$$\begin{aligned} T_\Omega = T_s - \frac{3J_2}{8a^2} T_s \{ & (12 + 34e^2) - (10 + 20e^2) \sin^2 i \\ & - (4 - 20 \sin^2 i) e \cos \omega + (18 - 15 \sin^2 i) e^2 \cos 2\omega \} \end{aligned} \quad (5.20)$$

Table 5.1 Transformations of T_s and T_Ω

e	0.001	0.01	0.10	0.20
T_Ω	119.786207	119.784455	119.757554	119.707661
$(T_\Omega)_1$	119.786085	119.784338	119.756505	119.696911
$\Delta T_\Omega/T_\Omega$	1.0×10^{-6}	1.0×10^{-6}	8.7×10^{-6}	8.9×10^{-5}
$(T_\Omega)_2$	119.786550	119.784815	119.757128	119.697762
$\Delta T_\Omega/T_\Omega$	2.9×10^{-6}	3.0×10^{-6}	3.6×10^{-6}	8.2×10^{-5}

Strictly speaking, the values of the orbital elements on the right-sides of the above two expressions should be at the starting time t_0 . For T_φ the starting time is when the satellite is on the reference latitude φ_* (i.e., point A in Fig. 5.1), for T_Ω it is on the ascending node (or descending node). Usually, as long as the required accuracy is satisfied, it is not necessary to separate them.

To test the above transformation of T_s and T_Ω , we use a satellite, whose sidereal period is 2 h, as an example. The initial values are:

$$T_s = 120^m.0, \quad i = 45^\circ, \quad \Omega = 45^\circ, \quad \omega = 0^\circ, \quad M = 0^\circ.$$

Table 5.1 gives the values of calculated T_Ω for different values of eccentricity.

In Table 5.1 the values of T_Ω is calculated by the approximately analytical formula (5.20), whereas $(T_\Omega)_1$ and $(T_\Omega)_2$ are given by numerical calculations using the perturbation equations as the correct values. For $(T_\Omega)_1$ only the J_2 perturbation is included, and for $(T_\Omega)_2$ the terms of J_2, J_3, J_4 , and $J_{2,2}$ are included. The values of ΔT_Ω are the differences between T_Ω and $(T_\Omega)_1$ or between T_Ω and $(T_\Omega)_2$.

Results in Table 5.1 show that the values of T_Ω given by the formula (5.20) which includes only the J_2 perturbation (note that the omitted terms J_3, J_4 , and $J_{2,2}$ have the same magnitude as $(J_2)^2$) can reach an accuracy of 10^{-5} except in the case of $e = 0.20$. Even for the relatively large e , the accuracy agrees with the analysis that the omitted term $J_2 e^3$ after integrating over one cycle has the same magnitude as $\Delta T_\Omega/T_\Omega$.

5.1.2 The Anomalistic Period

The anomalistic period T_ω is the time interval when a satellite successively passes the periapsis (for Earth it is the perigee) twice, therefore we have

$$T_\omega = \int_0^{2\pi} \left(\frac{dt}{dM} \right) dM.$$

If we only consider the J_2 perturbation then

$$\frac{dM}{dt} = n + f_c + f_s,$$

where f_c and f_s are the secular and long-period variation rates of the mean anomaly M , respectively, details are given in Sect. 4.2. By the same method used in Sect. 5.1.1 for T_φ , we can get the relationship between T_ω and the sidereal period T_s to the first-order accuracy. But when the eccentricity is small the position of the periapsis is difficult to decide, and so is the anomalistic period, therefore, in reality, this period is rarely used and we do not provide the transformation of T_ω and T_s .

5.2 Orbital Characteristics of Polar Orbit Satellite [2, 3]

A polar orbit satellite is a satellite with an orbital inclination $i = 90^\circ$ (or $i \approx 90^\circ$). The position of the orbital plane is almost fixed, i.e., $\dot{\Omega} \approx 0$. The variation of Ω and the inclination i have a close relationship. Based on the results in Chap. 4, we further analyze their variations in this section.

5.2.1 Basic Theories

The two orbital elements, i and Ω , decide the orientation of an orbital plane. When there are perturbations the differential equations for i and Ω are

$$\begin{cases} \frac{di}{dt} = \frac{1}{na^2\sqrt{1-e^2}\sin i} \left(\cos i \frac{\partial R}{\partial \omega} - \frac{\partial R}{\partial \Omega} \right), \\ \frac{d\Omega}{dt} = \frac{1}{na^2\sqrt{1-e^2}\sin i} \frac{\partial R}{\partial i}, \end{cases} \quad (5.21)$$

or

$$\begin{cases} \frac{di}{dt} = \frac{r\cos(f+\omega)}{na^2\sqrt{1-e^2}}W, \\ \frac{d\Omega}{dt} = \frac{r\sin(f+\omega)}{na^2\sqrt{1-e^2}\sin i}W, \end{cases} \quad (5.22)$$

where R is the perturbation function, W is the component of the perturbing acceleration in the normal direction to the orbit plane. The results given in Chap. 4 show that:

- (1) The perturbation functions of the zonal harmonic terms, J_l (i.e., $C_{l,0}$, $l = 1, 2, \dots$) due to Earth's non-spherical gravity potential $R(J_l)$ are not related to Ω , but

include $\sin i$ (odd zonal harmonic terms) or $\sin^2 i$ (even zonal harmonic terms). Thus, the corresponding forms of (5.21) are

$$\begin{cases} \frac{di}{dt} = \cos i \Phi_1(J_l; a, e, i, \omega, M), \\ \frac{d\Omega}{dt} = -\cos i \Phi_2(J_l; a, e, i, \omega, M) \end{cases} \quad (5.23)$$

This system of equations has a particular solution given by

$$\begin{cases} a = a(t), e = e(t), \omega = \omega(t), M = M(t), \\ i = i_0 = 90^\circ, \quad \Omega = \Omega_0, \end{cases} \quad (5.24)$$

where Ω_0 is an arbitrary real number on the interval $[0, 2\pi]$.

- (2) For the perturbation due to Earth's non-spherical gravity potential, after eliminating angular variables (the fast variation element M , and the slow variation elements Ω and ω) we obtain the differential equations for the mean elements of $\bar{i}$ and $\bar{\Omega}$ as

$$\frac{d\bar{i}}{dt} = 0, \quad \bar{i} = \bar{i}_0, \quad (5.25)$$

$$\frac{d\bar{\Omega}}{dt} = -\cos i \left[\left(\frac{3J_2}{2p^2} \right) n + O(J_2^2, J_{2l}) \bar{\psi}(\bar{a}_0, \bar{e}_0, \bar{i}_0) \right], \quad (5.26)$$

where $p = a(1 - e^2) = \bar{a}_0(1 - \bar{e}_0^2)$, $n = \bar{n}_0 = \bar{a}_0^{-3/2}$, and $J_{2,l}$ ($l = 1, 2, \dots$) for the even zonal harmonic terms.

The expression of (5.26) shows that the effect of the even terms of zonal harmonic perturbation on the secular variation of the longitude of ascending node has a factor $\cos i$. In fact, in the results of the third-body gravitational perturbation, the variation of the longitude of ascending node also has the same factor $\cos i$, as seen in (4.547). Furthermore, in the results by the gravitational perturbation of the sectoral harmonics terms ($C_{2l,2l}, S_{2l,2l}$), the most obvious terms (i.e., the terms related to the rotation of the central body through the rotation factor $1/\alpha$, see Sect. 4.6) also have the factor $\cos i$. When $i = 90^\circ$, all the effects of these perturbation terms disappear [2]. The orbital variation of a polar orbit satellite can be used for gravitational parameter measurements over a large area, but this characteristic can affect the measurements. To overcome this problem the orbital inclination can be selected slightly away from 90° . It was done in the design for the pair of satellites in the GRACE (Gravity Recovery And Climate Experiment) project in 2002, which was a joint mission of The National Aeronautics and Administration (NASA) USA and the German Aerospace Center (DLR) to measure Earth's gravity field anomalies. The satellite's inclination was 89° .

5.2.2 Preservation of Polar Orbit

Based on the analysis given in the above section that under the influence of gravitational perturbation due to the central body's non-spherical part we can conclude that a polar orbit exists, whose inclination and the longitude of ascending node are

$$i = i_0 = 90^\circ, \Omega = \Omega_0.$$

When we add the secular effect by a third-body perturbation or a hydrostatic atmospheric drag perturbation (i.e., not a rotating atmosphere) we still reach the same conclusion. Therefore, the polar orbit basically exists, the actual polar orbital plane swings due to the periodic effect of perturbations. To understand this phenomenon, we use an Earth's polar satellite as an example to show the amplitude of the oscillation.

Using a 2^h satellite (i.e., the period of the satellite is 2 h) with initial conditions

$$i = 90^\circ, \Omega = 45^\circ,$$

we integrate the perturbation equations over 10^4 cycles (which is a very long arc) to check how the polar orbit maintains its state. Results are presented in Table 5.2.

Table 5.2 gives four combination types of perturbation terms used for the calculation, which are

Type A: effects due to J_2 , J_3 , and J_4 terms.

Type B: effects due to J_2 , J_3 , J_4 , and $J_{2,2}$ terms.

Type C: effects due to J_2 , J_3 , J_4 , $J_{2,2}$, terms, the Sun's gravity, and the Moon's gravity.

Type D: effects due to J_2 , J_3 , J_4 , $J_{2,2}$, terms, the Sun's gravity, the Moon's gravity, the solar radiation pressure, and the atmospheric drag. The magnitudes of the solar radiation pressure and the atmospheric drag are 10^{-7} and 10^{-8} , respectively. The effects of Earth's shadow and the atmospheric rotation are ignored because these effects on the satellite's orbital plane are small as analyzed in Chap. 4.

Results provided in Table 5.2 show that polar orbits can preserve for years, the orbital plane swings but in the vicinity of the initial position even when all kinds of perturbation forces are included. For a high Earth orbit satellite, the results are almost the same as the 2^h satellite.

Table 5.2 Reservation and variation range of a polar orbit

Type	i (deg)	Ω (deg)
A	90.0 (invariant)	45.0 (invariant)
B	89.9993 – 90.0017	44.9999 – 45.0310
C	89.9966 – 90.0376	44.9785 – 45.9459
D	89.9975 – 90.0395	44.9900 – 46.0597

5.3 Existence and Design of Sun-Synchronous Orbit [2–5]

A Sun-synchronous orbit of a satellite means that the satellite's perturbed orbital plane keeps its precession a “constant”, i.e., observing from the central body this precession is the angular speed of the Sun towards the east. If the central body is Earth, then the angular speed of the Sun is $n_s = 0^\circ 0.9856/\text{day}$, accordingly a Sun-synchronous satellite moves around Earth 14 cycles/day on an orbit whose inclination is about 99° , which is close to a polar orbit. But we cannot conclude that a Sun-synchronous orbit is a polar orbit. The fact is that the condition of a Sun-synchronous orbit is not related to the condition of a polar orbit as discussed in the previous section. For example, a Moon's satellite moves around the Moon 12 cycles/day and is Sun-synchronous, but its orbital inclination is about 135° . The condition of a Sun-synchronous orbit depends not only on the main perturbation (the dynamical form-factor of the central body) but also on the revolution speed of the central body.

For a Sun-synchronous satellite the precession speed of its ascending node (whose longitude is Ω) $\dot{\Omega}$ equals the mean speed of the central body moving around the Sun, n_s . In other words, the precession of a satellite's orbital plane is synchronous with the Sun's “eastward motion”. This is a commonly applicable type of satellite, such as the first series of Chinese meteorological satellites, the FY series, and other kinds of geo-observation satellites.

5.3.1 Conditions of Forming a Sun-Synchronous Orbit

Based on the results derived in Chap. 4 due to the main zonal harmonic term J_2 perturbation the secular variation rate of the longitude of ascending node is

$$\dot{\Omega}_1 = -\frac{3J_2}{2p^2}ncos i. \quad (5.27)$$

Therefore, in order to keep the precession of the orbital plane synchronous with the Sun the condition is

$$\dot{\Omega}_1 = -\frac{3J_2}{2p^2}ncos i = n_s. \quad (5.28)$$

When the values of a and e are known, the value of i can be decided by (5.28) to give a Sun-synchronous orbit. This condition is derived only by the first-order secular perturbation (of cause it is the main perturbation). The question is when the effects of other perturbations are considered, is this condition still valid. The answer is given below.

The problem of the Sun-synchronous satellite is about whether the precession speed of the orbital plane can keep constant. Results from Chap. 4 show that under the perturbations of the non-spherical gravity of the central body and a third-body

attraction, after eliminating periodically varying angular variables (they have no secular effects on the precession of Ω), the mean elements, $\bar{i}$ and $\bar{\Omega}$, satisfy the following ordinary differential equations.

$$\frac{d\bar{i}}{dt} = 0, \quad \bar{i} = \bar{i}_0, \quad (5.29)$$

$$\frac{d\bar{\Omega}}{dt} = -\cos i \left[\left(\frac{3J_2}{2p^2} \right) n + O(J_2^2, J_{2l}, m') \bar{\psi}(\bar{a}_0, \bar{e}_0, \bar{i}_0) \right], \quad (5.30)$$

where m' is the mass of a third body. According to the requirement of the Sun-synchronous orbit $\dot{\Omega} = n_s$, then from (5.30) keeping the secular variations only (because they have practical meaning), we have

$$\begin{aligned} n_s = & -\cos i \left(\frac{3J_2}{2p^2} n \right) \\ & \times \left\{ 1 + \left(\frac{3J_2}{2p^2} \right) \left[\left(\frac{3}{2} + \frac{1}{6}e^2 + \sqrt{1-e^2} \right) - \sin^2 i \left(\frac{5}{3} - \frac{5}{24}e^2 + \frac{3}{2}\sqrt{1-e^2} \right) \right. \right. \\ & \left. \left. - \frac{35}{18} \left(\frac{J_4}{J_2^2} \right) \left(\frac{6}{7} + \frac{9}{7}e^2 - \sin^2 i \left(\frac{3}{2} + \frac{9}{4}e^2 \right) \right) + O \left(\frac{J_{2l}}{J_2^2} \right)_{l \geq 3} + O(m') \right] \right\} \end{aligned} \quad (5.31)$$

For given values of a and e , by (5.31) we can obtain the inclination that $i = i(n_s; a, e)$. Strictly speaking, the elements a , e , and i in (5.31) should be $\bar{a}$, $\bar{e}$ and $\bar{i}$, but when we design an orbit, we can decide the mean elements $(\bar{a}, \bar{e}, \bar{i})$ which satisfy $\dot{\Omega} = n_s$ and the corresponding instantaneous elements (a, e, i) simultaneously.

In the design of a Sun-synchronous orbit, there is the question of how to choose the two main orbital elements a and e . We usually choose $e = 0$ for a circular orbit; to determine the value of a , it's necessary to consider the following aspects.

- (1) The value of a depends on the required area coverage of satellite measurements, which decides the nodal period T_φ or T_Ω . By the relationship of the nodal period and the sidereal period given in Sect. 5.1, we can obtain the sidereal period T_s , then from T_s to obtain a , and by (5.31) to obtain the inclination i . But to transfer a nodal period to a sidereal period needs the inclination i , since i can be obtained from (5.31) with a known a , the whole process is actually an iterative process, which is not complicated.
- (2) When we use (5.31) to decide the value of i for given a and e the question is whether we should consider the J_2 perturbation only or to use $J_2, (J_2^2, J_{2l}, \dots)$ and the third-body perturbation together. The fact is that the value of i cannot be rigorously calculated using analytical formulas, even when all perturbations are included, we can only use the secular terms of these perturbations, besides there are other perturbation factors not applicable to (5.31). The answer is that

to include all perturbations is not necessarily better than using J_2 term only. The method to solve this problem is first to consider the J_2 perturbation by which we have

$$\cos i = -n_s / \left(\frac{3J_2}{2p^2} n \right) \quad (5.32)$$

Then we consider all perturbations by numerical methods, and use the above-chosen orbit elements as references to adjust the value of i (i.e., by a simple iteration) until the value of i can provide the required measurement area coverage by which to provide the optimal orbit (corresponding to the optimal value of i).

Now we use a low Earth orbit satellite as an example to demonstrate this method. Assuming that the satellite moves around Earth 14 cycles per day, the initial values of the three main orbital parameters are $e_0 = 0.001$, and a_0 depends on the following two cases

$$\begin{cases} (a) T_\Omega = 1 \text{ side real day}/14 = 102^{\text{m}}.576298, \\ (b) T_\Omega = 1 \text{ mean solar day}/14 = 102^{\text{m}}.857143. \end{cases} \quad (5.33)$$

Results are shown in Tables 5.3 and 5.4. In the tables, Type A is for the inclination i_0 given by (5.31), i.e., by J_2 term only; Type B is for the inclination i_0 given by (5.31), i.e., by J_2 term and the third-body perturbation; Case a and Case b correspond to (a) and (b) of (5.33), respectively. The initial values of i_0 and T_s are given in Table 5.3. The maximum of $(\dot{\Omega} - n_s)$ and the difference between the longitude of ascending node and the longitude of the Sun, $(\Omega - S)$ after integrating the motion equations which include J_2, J_3, J_4 , and $J_{2,2}$ terms over 1120 cycles are given in Table 5.4.

Table 5.3 shows that the results given by numerical calculation are very close for both Type A and Type B, which agree with the analytical conclusion that it is not necessarily better if more perturbation factors are included in the process of deciding

Table 5.3 Initial values of T_s and i_0

Type	T_s (s)		i_0 (deg)	
	a	b	a	b
A	102.648841	102.929720	98.962980	99.020816
B	102.648901	102.929780	98.997442	99.055336

Table 5.4 $\max(\dot{\Omega} - n_s)$ and the value of $(\Omega - S)$

Type	$\max(\dot{\Omega} - n_s)$ (deg/d)		$(\Omega - S)$ (deg)	
	a	b	a	b
A	0.0035	0.0034	0.1547	0.1548
B	0.0073	0.0071	0.4551	0.4547

the value of i . But by this example only, we cannot come to the conclusion that Type A is better than Type B.

The values of a and i are chosen according to different values of T_Ω . The different values of a and i then decide different methods of how a satellite orbit covers the Earth's surface. To a ground station in Case a) when T_Ω equals a sidereal day/14, every day the intersection point of the orbit and the equator moves about 0.98° eastward; in Case b) when T_Ω equals a mean solar day/14, the intersection point of the orbit and the equator does not change (i.e., every day is the same). Because of errors in the initial values and actual effects of perturbations, after 1120 cycles, the longitude difference of the satellite's orbital plane with respect to a fixed point on the ground at the "same time" of every day increases from 0° to 0.15° and 0° to 0.45° corresponding to Type A and Type B, respectively.

5.3.2 Sun-Synchronous Orbits for Different Celestial Bodies

For the major planets and the Moon in the Solar System, the gravity of the J_2 term of an oblate central body is the main perturbation force, it is reasonable to use (5.32) to design a Sun-synchronous satellite orbit for these bodies. The results of a low orbit satellite for Earth, Mars, and the Moon are as follows.

- (1) For a Sun-synchronous orbit of a satellite that moves around Earth about 14 cycles per day (Earth day) there are

$$\begin{aligned}\bar{h} &= 890.0 \text{ km}, \quad T_s = 102^m.8571, \\ a &= 7271.9 \text{ km}, \quad e = 0.0001, \quad i = 98^\circ.9025\end{aligned}\quad (5.34)$$

- (2) For a Sun-synchronous orbit of a satellite that moves around Mars about 12 cycles per day (Mars day) there are

$$\begin{aligned}\bar{h} &= 500.0 \text{ km}, \quad T_s = 123^m.1001 \\ a &= 3897.0 \text{ km}, \quad e = 0.0005, \quad i = 93^\circ.2005.\end{aligned}\quad (5.35)$$

- (3) For a Sun-synchronous orbit of a satellite which moves around the Moon about 12 cycles per day (Earth day) there are

$$\begin{aligned}\bar{h} &= 120.0 \text{ km}, \quad T_s = 120^m.0, \\ a &= 1860.3 \text{ km}, \quad e = 0.0010, \quad i = 134^\circ.9120.\end{aligned}\quad (5.36)$$

It should be mentioned that it is impossible to have a Sun-synchronous spacecraft moving around Venus. Because the rotation of Venus is very slow and Venus is almost a sphere, the magnitude of its J_2 term is small that $J_2 = 4.5 \times 10^{-6}$. By J_2 perturbation the corresponding value of $\dot{\Omega}$ is much too small compared to the Sun's eastward speed n_s , meaning that there is no Sun-synchronous satellite for Venus.

5.4 Existence and Design of Frozen Orbit [2–5]

5.4.1 Basic State of Frozen Orbit

A frozen orbit is an orbit with a stationary apsidal line, i.e., the direction of the periapsis of the orbit is “fixed”. Strictly speaking, the inclination of a frozen orbit can take any possible value (except in the vicinity of the critical inclination, which we will discuss in the next section). This kind of orbit is important in observing terrestrial phenomena (or for other central bodies). For example, the American Navy satellite GEOSAT launched on March 12, 1985, was designed as a frozen orbit for measuring the distance from the satellite to the sea surface.

This special orbit in fact is the mean orbital solution, i.e., the particular solution of the motion equations after eliminating all short-period terms (i.e., the periodic terms composed by the fast-varying element). For a low orbit satellite, the motion equations after eliminating all short-period terms correspond to an average motion, on which the main perturbation is due to the zonal harmonic J_l ($l \geq 2$) terms of the central body's non-spherical gravity. The original equations then are reduced to a four-dimensional dynamical system, which only includes four orbital elements, a , e , i , and ω . The argument of periapsis of a frozen orbit can have one of the following two values [2] that

$$\omega = 90^\circ \text{ or } 270^\circ.$$

For a given semi-major axis and an inclination of an orbit (a , i), the orbital eccentricity can be derived by

$$e = \left| \frac{J_3}{J_2} \right| \frac{1}{2a} \sin i [1 + O(\varepsilon^2)], \quad (5.37)$$

where ε^2 is a small value of higher order with respect to $|J_3/J_2|$. If $(J_3/J_2) > 0$, the related frozen orbit solution has $\omega = 270^\circ$, and if $(J_3/J_2) < 0$, then $\omega = 90^\circ$. Assuming the central body to be Earth we analyze the solution of a frozen orbit based on the results in Chap. 4, and provide the method of designing a frozen orbit as follows.

5.4.2 Basic Equations of a Possible Frozen Orbit

As described in Chap. 4, the motion equations of a satellite in an epoch geocentric celestial coordinate system after eliminating the effects of precession, nutation, and polar motion can be presented as

$$\dot{\sigma} = f(\sigma, t, \beta), \quad (5.38)$$

where σ is a vector of six orbital elements, i.e., $\sigma = (a, e, i, \Omega, \omega, M)^T$. The variable β in the function f is a parameter representing the Earth's non-spherical gravity potential, and the explicitly included time t reflects Earth's rotation (through the tesseral harmonic terms). The equation shows that it is a non-autonomous system, and it does not have a particular solution of $\dot{\omega} = 0$, which means a rigorously defined frozen orbit does not exist.

For a satellite, if it is not a high orbit satellite, and the tesseral harmonic terms do not produce resonance terms, which have a strong perturbing effect, then by eliminating the fast variable terms in the Eq. (5.38) (i.e., the short-period terms related to the mean anomaly M and the Earth's rotation) we have a system of motion equations with reduced dimensions, i.e., a four-dimensional autonomous system as

$$\begin{cases} \dot{X} = f(X; J_l), \\ X = (a, e, i, \omega)^T, \end{cases} \quad (5.39)$$

where X is actually the set of quasi-mean elements $\tilde{\sigma}$ defined in Chap. 4, which includes only the secular and long-period terms. The equations for each element have forms as:

$$\begin{cases} \frac{da}{dt} = f_a(a, e, i, \omega), \\ \frac{de}{dt} = f_e(a, e, i, \omega) = -\left(\frac{1-e^2}{e} \tan i\right) f_i, \\ \frac{di}{dt} = f_i(a, e, i, \omega), \\ \frac{d\omega}{dt} = f_\omega(a, e, i, \omega). \end{cases} \quad (5.40)$$

We do not provide the analytic forms for Eq. (5.40) here, details are given in Chap. 4. The functions in (5.40) include neither time t nor the longitude of ascending node Ω . That is why Ω and M can be separated from the other elements, a, e, i , and ω , and the system of Eq. (5.38) can be reduced to a four-dimensional autonomous system (5.39). The system of Eq. (5.40) is the basic equation system for discussing the existence of frozen orbit, i.e., whether (5.40) has a particular solution corresponding to $\dot{\omega} = 0$.

5.4.3 A Particular Solution of Eq. (5.40): The Frozen Orbit

According to the results due to the central body's non-spherical perturbation, one of the properties of the system of Eq. (5.40) is that for elements a , e , and i , the terms of the right-side functions contain terms of $\sin 2\omega$, $\sin 4\omega$, $\dots$ and $\cos \omega$, $\cos 3\omega$, $\dots$. Therefore, when $\omega = 90^\circ$ or $\omega = 270^\circ$, we have

$$\frac{da}{dt} = 0, \quad \frac{de}{dt} = 0, \quad \frac{di}{dt} = 0. \quad (5.41)$$

Another property of (5.40) is about the right function of element ω , which appears as $\cos 2\omega$, $\cos 4\omega$, $\dots$. When $\omega = 90^\circ$ there are $\cos 2\omega = -1$, $\cos 4\omega = 1$, $\sin \omega = 1$, and $\sin 3\omega = -1$; when $\omega = 270^\circ$ there are $\cos 2\omega = -1$, $\cos 4\omega = 1$, $\sin \omega = -1$, and $\sin 3\omega = 1$.

Based on these two properties, when $\omega = 90^\circ$ or $\omega = 270^\circ$ only the equation of the element ω in the system of Eq. (5.40) is not zero, that

$$\frac{d\omega}{dt} = f_\omega(a, e, i, \omega_0).$$

By setting $f_\omega(a, e, i, \omega_0) = 0$, we obtain a relationship of a , e , and i . Then for given a_0 and i_0 , we can derive a solution that $e_0 = e(a_0, i_0)$, and finally obtain a particular solution of the system of Eq. (5.40) as.

$$a \equiv a_0, \quad e \equiv e_0, \quad i \equiv i_0, \quad \omega \equiv \omega_0 = 90^\circ \text{ or } 270^\circ. \quad (5.42)$$

This set of solutions is the frozen orbit corresponding to the quasi-mean orbital elements. The relationship of the three mean variables of a , e , and i is given by

$$\begin{aligned} e_0 = & \mp \left\{ \frac{3J_3}{4p} \sin i \left[\left(2 - \frac{5}{2} \sin^2 i \right) - \frac{e^2}{\sin^2 i} \left(2 - \frac{35}{2} \sin^2 i + \frac{35}{2} \sin^4 i \right) \right] + O\left(\frac{J_5}{p^3}, \dots\right) \right\} \\ & \times \left\{ \frac{3J_2}{2} \left(2 - \frac{5}{2} \sin^2 i \right) \right. \\ & + \left(\frac{3J_2}{2p} \right)^2 \left[\left(2 - \frac{5}{2} \sin^2 i \right) \left(3 - \frac{79}{24} \sin^2 i \right) - e^2 \left(\frac{5}{12} - \frac{19}{8} \sin^2 i + \frac{75}{32} \sin^4 i \right) \right] \\ & \left. - \left(\frac{3J_2}{2p} \right)^2 \right. \\ & \left[\sin^2 i \left(2 - \frac{5}{2} \sin^2 i \right) \left(\frac{1}{4} + \frac{1-e^2}{3} cf \right) + e^2 \left(\frac{2}{3} - \frac{11}{2} \sin^2 i + \frac{10}{3} \sin^4 i \right) cf \right. \\ & \left. - \sin^2 i \left(\frac{25}{12} - \frac{5}{2} \sin^2 i \right) + e^2 \left(\frac{7}{12} - \frac{79}{24} \sin^2 i + \frac{45}{16} \sin^4 i \right) \right] \\ & \left. - \frac{35(-J_4)}{8p^2} \left[\sin^2 i \left(\frac{9}{14} - \frac{3}{4} \sin^2 i \right) - e^2 \left(\frac{9}{14} - \frac{15}{4} \sin^2 i + \frac{27}{8} \sin^4 i \right) \right] \right\} \end{aligned}$$

$$+ O(J_2^3, J_2 J_4, \dots)_c + O(J_2^3, J_2 J_4, J_2 J_3, \dots)_l \}^{-1}, \quad (5.43)$$

where cf is defined as

$$cf = \frac{1}{e^2} \overline{\cos 2f} = \frac{1 + 2\sqrt{1 - e^2}}{(1 + \sqrt{1 - e^2})^2} = \frac{3}{4} + O(e^2) \quad (5.44)$$

and $n = a^{-3/2}$, $p = a(1 - e^2)$. Note that the calculation units are the same as in Chap. 4, and Earth's gravitational constant $\mu = GE = 1$.

The sign at the beginning on the right side of (5.43), $\mp$, corresponds to $\omega \equiv \omega_0 = 90^\circ/270^\circ$, and the elements a, e, i should be a_0, e_0, i_0 , and $p = a_0(1 - e_0^2)$. The terms of $O(\dots)_c$ and $O(\dots)_l$ are for the high-order secular and long-period terms, respectively. If i_0 is not near the critical inclination $i_c = 63^\circ 26'$, i.e., $|2 - \frac{5}{2}\sin^2 i_0| > 10^{-3}$, then the main part of (5.42) is reduced to

$$e_0 = \mp \left(\frac{J_3}{J_2} \right) \frac{1}{2a_0} \sin i_0 [1 + O(e_0^2)] = O(10^{-3}). \quad (5.45)$$

The value of e_0 by (5.45) shows that an orbit with a stationary apsidal line can only exist if the orbit is near a circle. When $i_0 \rightarrow i_c$, i.e., the inclination of the orbit is in the vicinity of the critical inclination, then we have to deal with the problem of orbital resonance, which is discussed in Sect. 5.4.6.

The expressions (5.43) and (5.45) give the relationships of orbital elements a, e , and i for a frozen orbit. The value of e_0 can be decided by given values of a_0 and i_0 . Because $\omega = 90^\circ$ or 270° , the direction of the periaapsis is fixed. The variation of the element Ω (moves eastward or westward) does not change this relationship. The sign $\mp$ on the right side of (5.43) and (5.45) depends on the sign of (J_3/J_2) of the central body. If $(J_3/J_2) < 0$, then the solution takes “−”, and if $(J_3/J_2) > 0$, the solution takes “+”.

The characteristics of the frozen orbit solution would not be changed if we include perturbation terms of J_{2l-1}, J_{2l} ($l = 2, 3, \dots$) and tesseral terms $J_{l,m}$. This is affirmed by numerical calculations not provided here.

5.4.4 Stability of Frozen Orbit

The above-given orbit solution is a particular solution of an averaged system (after eliminating the short period variation). Whether an orbit can keep a stationary apsidal line in the original complete dynamical system depends on the stability of the particular orbit solution. From a point of view of a linear system, the answer is yes, the particular solution of (5.42) corresponding to a frozen orbit is stable, which can be proved. For a non-linear system, specifically a complete dynamical system including

all sorts of perturbations, the question about the stability of a frozen orbit is not important, what we concern about is how this kind of orbit varies under limited perturbations. For that, we use examples to calculate the variations of Earth's frozen orbits.

- (1) Given the initial epoch 28 October 2014 00:00:00, the orbital period $T_0 = 95^{\text{m}}.7379$, and mean orbital elements

$$a_0 = 6932.3854 \text{ km}, \quad e_0 = 0.00118020, \quad i_0 = 97^\circ.0,$$

$$\Omega_0 = 45^\circ.0, \quad \omega_0 = 90^\circ.0, \quad M_0 = 10^\circ.0$$

by Earth's gravitational model to degree and order 50×50 , for a complete dynamic model, after extrapolating over 200 days, results of the elements of the above given frozen orbit are

$$a = 6932.3455 \text{ km}, \quad e = 0.00136719, \quad i = 96^\circ.878766,$$

$$\Omega = 224^\circ.245937, \quad \omega = 89^\circ.776773, \quad M = 132^\circ.865862.$$

Figure 5.2 shows the variation of element ω . The amplitude of the variation is $\Delta\bar{\omega} \approx \pm 5^\circ$.

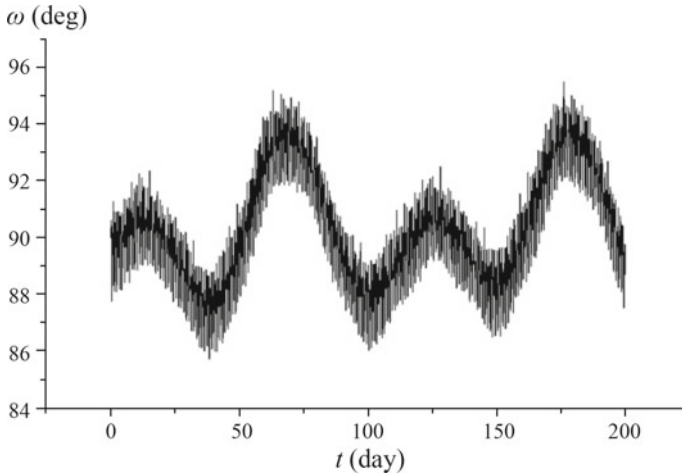


Fig. 5.2 The evolution of the apsidal line direction ($\bar{\omega}$)

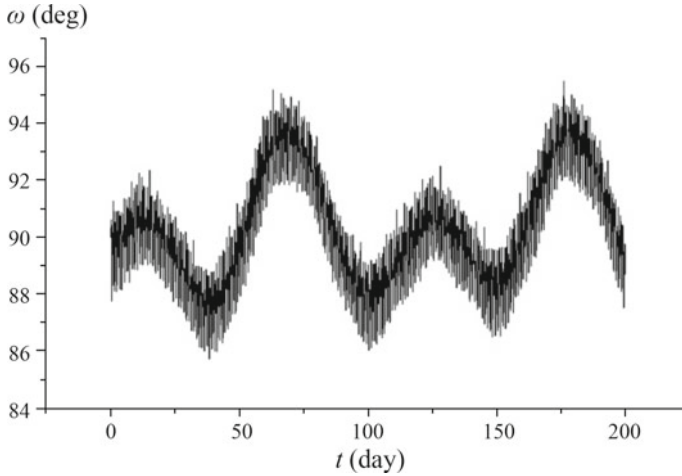


Fig. 5.3 The evolution of the apsidal line direction ($\bar{\omega}$)

- (2) By the same Earth's gravitational model and the perturbations of the Sun and the Moon, and using the same initial conditions, the results of the elements of the frozen orbit after extrapolating over 200 days are

$$a = 6932.3469 \text{ km}, e = 0.00136783, i = 96^\circ.875546,$$

$$\Omega = 224^\circ.237688, \omega = 89^\circ.809594, M = 132^\circ.590597.$$

The evolution of element ω is given in Fig. 5.3, the amplitude is the same that $\Delta\bar{\omega} \approx \pm 5^\circ$.

The above results show that for the averaged system under perturbations the state of the “stationary” apsidal line can keep for a relatively long time. But for the actual dynamical system, which includes the short-period variation, the results are quite different from that of an averaged system, showing that the apsidal line is no longer frozen. That agrees with the analysis given in Sect. 5.4.2 that the non-autonomous system (5.38) does not have a particular solution of $\dot{\omega} = 0$, i.e., the frozen orbit does not exist. But from a practical perspective, what we need is the frozen orbit of an averaged system.

5.4.5 Frozen Orbit for Other Celestial Bodies

If we only consider J_2 and J_3 perturbations, for the argument of periapsis ω the condition of balancing its secular term and the long-period term is

$$\left| \frac{J_3}{e} \right| = O(J_2). \quad (5.46)$$

The relative magnitude of J_2 and J_3 is a very important parameter. As discussed above, this value influences not only the lifespan of a satellite (to be discussed in Sect. 5.6) but also the existence of a frozen orbit, including its property and orientation (i.e., $\omega = 90^\circ$ or 270°). Here we give two important conclusions as follows.

- (1) The value of ω of a frozen orbit depends on the sign of J_2 and J_3 . For Earth and Venus, there is $(J_3/J_2) < 0$, which yields $\omega = 90^\circ$ for a frozen orbit; whereas for Mars and the Moon $(J_3/J_2) > 0$, then $\omega = 270^\circ$.
- (2) According to (5.46) the value of the eccentricity of a frozen orbit depends on the relative value of J_2 and J_3 . For the four celestial bodies (Earth, Venus, Mars, and the Moon), the magnitudes of the ratio (J_3/J_2) are 10^{-3} , 10^{-1} , 10^{-2} , and 10^{-1} , respectively, therefore, an Earth's frozen orbit must have a small eccentricity, i.e., $e = O(10^{-3})$, whereas a frozen orbit of a Moon's satellite can have an eccentricity near 0.1, showing that a frozen orbit is not necessarily an orbit with a small eccentricity.

5.4.6 Characteristics and Applications of Satellite Orbit with a Critical Inclination

The critical inclination is caused by the J_2 perturbation of the central body's non-spherical gravity force. An orbit with the critical inclination also has a "stationary" apsidal line. As given in Chap. 4, formula (4.66) in dimensionless units, the main part of the secular term of the argument of periapsis ω is

$$\dot{\omega} = \frac{3J_2}{2p^2} n \left(2 - \frac{5}{2} \sin^2 i \right). \quad (5.47)$$

When $\dot{\omega} = 0$, from the secular variation point of view, the direction of the periapsis does not change, and the corresponding inclination is called the critical inclination, denoted by i_c , that

$$i_c = 63^\circ 26', 116^\circ 34'. \quad (5.48)$$

This kind of orbit is fundamentally different from the frozen orbit. It involves orbital resonance, and its stability is restricted by the magnitude of the eccentricity

of the orbit, the larger the eccentricity, the more stable the orbit is. One type of communication satellite that belonged to the former Soviet Union, called Molniya, was designed with a large eccentricity and a critical inclination, there were a total of 164 Molniya satellites. For a country mostly located in the high latitude region, using the geosynchronous orbit (above the equator) to transfer information would cost tremendous energy, but using the Molniya type of satellite with a “frozen orbit” as a working orbit for communication could save energy significantly. Molniya satellites kept the apogee over the former Soviet Union, and three Molniya satellites could provide 24-h signal coverage for the entire country.

The “stability” of this kind of orbit actually means that the apsidal line oscillates around a balanced direction, the larger the orbital eccentricity is, the larger the oscillation amplitude can be. This is the special characteristic of the long-periodic variation due to the resonance effect. When $i \approx i_c$, the commensurable small divisor ($4 - 5\sin^2 i \approx 0$) leads to resonance and the apsidal line oscillates around $\omega = 90^\circ$. There is a pair of resonance variables, $G = \sqrt{\mu a(1 - e^2)}$ (angular momentum variable) and ω (angular variable), and two sets of equilibrium solutions, which are

$$\begin{cases} x = \omega = \pi/2, 3\pi/2, \\ y = G = G_c \end{cases} \quad (5.49)$$

and

$$\begin{cases} x = \omega = 0, \pi, \\ y = G = G_c. \end{cases} \quad (5.50)$$

The two sets of solutions correspond to $i_c \approx 63^\circ.4$ and $i_c \approx 116^\circ.6$, respectively. The first equilibrium solution given by (5.49) is at the center (stable) and that by (5.50) is at the saddle (unstable). In the plane of (G, ω) around the center, there are closed orbits.

For proving the above conclusion, we use numerical calculation for an orbit with a critical inclination. The initial conditions are at epoch 6 October 2014 00:00:00 that

$$a_0 = 7873.2753 \text{ km}, \quad e_0 = 0.10063559, \quad i_0 = 63^\circ.447269,$$

$$\Omega_0 = 0^\circ.0, \quad \omega_0 = 45^\circ.0, \quad M_0 = 0^\circ.0.$$

Results of extrapolating over 60,000 days are shown in Figs. 5.4 and 5.5. In the figures $G = \sqrt{\mu a(1 - e^2)}$, $g = \omega$, both are given by quasi-mean elements.

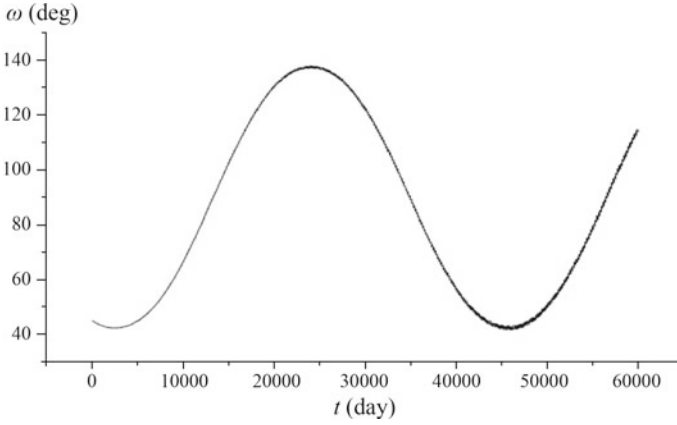


Fig. 5.4 The evolution of the apsidal line direction ($\omega-t$) of the orbit with a critical inclination

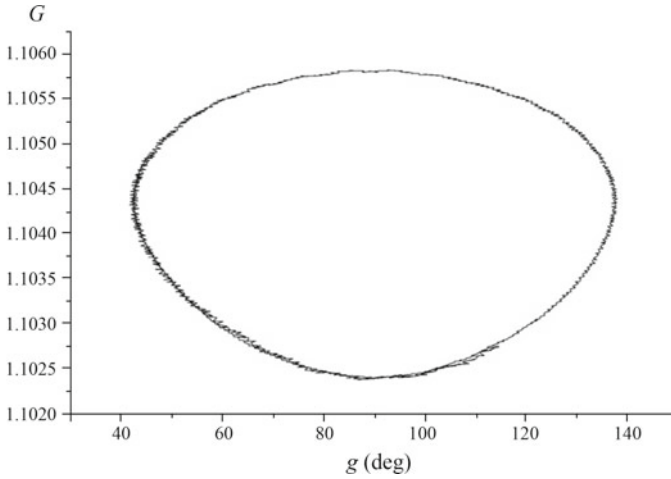


Fig. 5.5 The resonance of ($G-g$) of the orbit with the critical inclination

The two figures show that the direction of the apsidal line of the resonance orbit oscillates around $\bar{\omega} = 90^\circ$ with an amplitude of 50° instead of moving on a cycle of $0^\circ - 360^\circ$, and the variations of the angular momentum $\bar{G}$ and the angle $\bar{g}$ have similar characteristics.

5.5 Existence and Design of Central Body Synchronous Orbit

5.5.1 *Basic State of Central Body Synchronous Satellite Orbit [2, 3, 5]*

The orbital “resonance” caused by the critical inclination as discussed in the previous section, and the “resonance” of the synchronous satellite orbit due to the “third-body” perturbation (as the non-spherical tesseral harmonic part of the central body) discussed in Chap. 4, both have periodic variations with long periods and large amplitudes (usually less than the order of $\varepsilon^{1/2}$). For orbit design, this information is important. In this section, we further discuss the basic characteristics of this kind of orbit and its dynamical mechanism from the application point of view and provide a necessary theoretic foundation for designing synchronous satellite orbits.

5.5.1.1 Possibility of the Existence of Central Body-Synchronous Satellite Orbit

Theoretically to form a central body synchronous satellite orbit is relatively simple, the only requirement is that the orbital period equals the rotational period of the central body. If the synchronous satellite is above the equator of the central body, it becomes a “stationary” satellite. The question is about the possibility of the existence of this kind of satellite orbit in the actual Solar System.

If the central body is Earth, a synchronous satellite has a semi-major axis of about 42,000 km. The satellite is stable whether it is in the Sun-Earth system or in the Earth-Moon system, meaning that the satellite would not be pulled away by the Sun or the Moon. But if the central body is Venus or the Moon, things can be different.

If the central body is Venus, the rotation speed of Venus is very slow, its period is 243^d.0 (Earth day) even longer than the period of Venus’s revolution around the Sun. To form a synchronous orbit, the corresponding semi-major axis of the orbit would be 1.5365×10^6 km, which is beyond the region of Venus’s gravitation (Venus’s Hill sphere has a radius of 1.0112×10^6 km). A satellite in this kind of orbit would soon move away from Venus due to the Sun’s gravity. The conclusion is obvious that the Venus-synchronous satellite cannot exist.

The situation for the Moon is similar to Venus. The rotation speed of the Moon is slow, and the rotation period is 27^d.3217 (Earth day), therefore, the semi-major axis of the corresponding synchronous orbit would be 88.45×10^3 km, which is also greater than the region of its Hill sphere (61.6×10^3 km). If we only consider the perturbation of the Moon’s non-spherical gravity, because the Moon’s equatorial ellipticity $J_{2,2}$ is relatively large, a Moon synchronous satellite in the high altitude along the minor axis (the east-longitude 90°) direction is relatively stable. But in reality, there is Earth, the satellite in that orbit would be attracted by Earth and very

soon leave the Moon. If an initial eccentricity $e = 0.0001$ and an initial inclination $i = 0^\circ.005$ given to a synchronous orbit, then after less than $2^d.8$ due to Earth's gravity force the orbit would be "oblate" and the eccentricity would increase to about 1. The conclusion is that the Moon-synchronous orbit cannot exist.

The above simple analyses show that anyone working on the orbital design for aerospace projects must have the basic dynamical knowledge to avoid unnecessary mistakes.

5.5.1.2 Basic Characteristics of Central Body Synchronous Satellite Orbit

The most important influence on the synchronous orbit's stability comes from the central body's equatorial ellipticity ($J_{2,2}$), which decides the synchronous orbit's stable region and properties. In this section, we take an Earth's satellite as an example and use a different method to discuss the formation of a synchronous orbit and the mechanics of the oscillation of the satellite along the Earth equator's minor axis. Details about the effect and properties of Earth equatorial ellipticity perturbation are given in Sect. 4.3.

(1) Motion equations

Including the main terms of Earth's non-spherical gravity (the dynamical form-factor J_2 and the ellipticity $J_{2,2}$), omitting the influences of precession, nutation, and polar motion, using the dimensionless normalized units introduced in Chap. 4, and in the geocentric equatorial frame $O-r\lambda\varphi$, the corresponding Earth's gravitational potential can be expressed as

$$\begin{aligned} V &= \frac{1}{r} \left[1 - \frac{J_2}{r^2} P_2(\sin \varphi) - \frac{J_{2,2}}{r^2} P_{2,2}(\sin \varphi) \cos 2\bar{\lambda} \right] \\ &= \frac{1}{r} - \frac{J_2}{r^3} \left(\frac{3}{2} \sin^2 \varphi - \frac{1}{2} \right) - \frac{J_{2,2}}{r^3} (3 \cos^2 \varphi) \cos 2\bar{\lambda} \end{aligned} \quad (5.51)$$

where

$$\begin{cases} J_2 = -C_{2,0}, & J_{2,2} = -(C_{2,2}^2 + S_{2,2}^2)^{1/2}, \\ \bar{\lambda} = \lambda - (S_G + \lambda_{2,2}), \\ 2\lambda_{2,2} = \arctan\left(\frac{S_{2,2}}{C_{2,2}}\right). \end{cases} \quad (5.52)$$

All variables in these formulas are given in (4.161), Chap. 4 and Fig. 4.1; S_G is the Greenwich sidereal time; $\bar{\lambda}$ is the longitude of a satellite with respect to the direction of Earth equator's major axis (the longitude of this direction is $\lambda_{2,2}$ in the Earth-fixed frame).

In the geocentric equatorial coordinate system, the position vector $\vec{r}$ and velocity $\dot{\vec{r}}$ of a satellite are given by

$$\vec{r} = \begin{pmatrix} r \\ 0 \\ 0 \end{pmatrix}, \quad (5.53)$$

$$\dot{\vec{r}} = \begin{pmatrix} \dot{r} \\ r \cos \varphi \dot{\lambda} \\ r \dot{\varphi} \end{pmatrix}, \quad (5.54)$$

where

$$\dot{\lambda} = \dot{\bar{\lambda}} + \dot{S}_{2,2} = \dot{\bar{\lambda}} + n_e, \quad (5.55)$$

and n_e is Earth's rotational angular speed. The kinetic energy of the satellite (without the factor of mass) T can be expressed as

$$T = \frac{1}{2}(\dot{r}^2 + r^2 \cos^2 \varphi \dot{\lambda}^2 + r^2 \dot{\varphi}^2). \quad (5.56)$$

Based on the following Lagrange motion equations in dynamics

$$\left\{ \begin{array}{l} \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}} \right) - \frac{\partial T}{\partial q} = \frac{\partial v}{\partial q}, \\ q = \begin{pmatrix} r \\ \lambda \\ \varphi \end{pmatrix}, \quad \dot{q} = \begin{pmatrix} \dot{r} \\ \dot{\lambda} \\ \dot{\varphi} \end{pmatrix}, \end{array} \right. \quad (5.57)$$

we have the basic equations of a satellite's motion as

$$\left\{ \begin{array}{l} \ddot{r} - r \cos^2 \varphi \dot{\lambda}^2 - r \dot{\varphi}^2 = -\frac{1}{r^2} + \frac{3J_2}{r^4} \left(\frac{3}{2} \sin^2 \varphi - \frac{1}{2} \right) + \frac{9J_{2,2}}{r^4} \cos^2 \varphi \cos 2\bar{\lambda}, \\ \frac{d}{dt} (r^2 \cos^2 \varphi \dot{\lambda}) = \frac{6J_{2,2}}{r^3} \cos^2 \varphi \sin 2\bar{\lambda}, \\ \frac{d}{dt} (r^2 \dot{\varphi}) + \frac{1}{2} r^2 \sin 2\varphi \dot{\lambda}^2 = -\frac{3J_2}{2r^3} \sin 2\varphi + \frac{3J_{2,2}}{r^3} \sin 2\varphi \cos 2\bar{\lambda}. \end{array} \right. \quad (5.58)$$

One of the characteristics of geosynchronous satellite motion is that under certain conditions, the satellite oscillates around the direction of the minor axis of Earth's equator. This property is related to the variation of $\bar{\lambda}$.

- (2) Particular solutions of the motion equations and the motion properties of geosynchronous satellite

It is not difficult to prove that the system of Eq. (5.58) has particular solutions as following

$$\begin{cases} r_1 \equiv r_{01}, \bar{\lambda}_1 = 90^\circ, 270^\circ, \varphi_1 \equiv 0, \\ r_2 \equiv r_{02}, \bar{\lambda}_2 = 0^\circ, 180^\circ, \varphi_2 \equiv 0. \end{cases} \quad (5.59)$$

From $J_2 = 1.082637 \times 10^{-3}$ and $J_{2,2} = -1.771156 \times 10^{-6}$, we obtain r_{01} and r_{02} that

$$\begin{cases} r_{01} = 42164.71346 \text{ km (above the major axis),} \\ r_{02} = 42164.72371 \text{ km (above the minor axis).} \end{cases} \quad (5.60)$$

These two solutions are the equilibrium solutions. It can be proved that the solution r_{01} is at the center, which can be proved to be linearly stable, and r_{02} is unstable at the saddle [4]. The question about whether r_{01} is non-linearly stable is not important, what important is that when the other perturbations are considered (perturbations with different characteristics) can this equilibrium state be maintained, or in a complicated dynamical circumstance can a satellite “oscillates” in the vicinity of the equilibrium point (above the equator in the direction of the minor axis). The answer to this question depends on a feature variable, which is $\bar{\lambda}$ in this problem. If $\bar{\lambda}$ is at the stable equilibrium point of 90° (or 270°) with a variation range of $\Delta\bar{\lambda}$ less than $\pm 90^\circ$ then we can say the solution is stable. In the next section, we use numerical calculations to show this property.

5.5.2 Existence and Evolution of a Central Body Synchronous Satellite (Earth, Mars)

- (1) Characteristics of the evolution of geosynchronous satellite orbit

The geosynchronous satellites have become familiar objects to modern society. The orbital period and semi-major axis are about

$$T_s = 1436^{\text{m}}.068176, a = 42164.170 \text{ km.}$$

For a geosynchronous satellite, there are two stable regions above the equator due to the perturbation of Earth's equatorial ellipticity $J_{2,2}$ term (its value is given by the parameters $\bar{C}_{2,2}$, $\bar{S}_{2,2}$). The two stable regions are along the two directions of

the equator's minor axis (longitudes of 75° E and 105° W). A synchronous satellite swings in east–west directions in one of the vicinities of the two stable points (by the orbital resonance mechanism similar to a pendulum phenomenon). The synchronous satellite also swings in north–south directions (actually it is the variation of the orbital inclination) which is caused by a third-body (the Sun or the Moon) gravity force, the variation of the inclination has a long period, about 50 years, and a maximum amplitude of 15° .

For a better understanding, we now use numerical calculations to show the actual states of the east–west-swings and north–south swings of a synchronous satellite orbit (or position).

The initial epoch t_0 is 10 September 2010 00:00:00 (UTC), the initial deviations of period, semi-major axis, and λ (relative to the direction of the minor axis) from the equilibrium solution are

$$\Delta T_s = 6^m 38^s, \Delta a = 10 \text{ km}, \Delta \lambda = 35^\circ.0.$$

The initial deviation of the orbital inclination is small that $\Delta i = 0^\circ.005 \approx O(10^{-4})$, which means that the satellite is basically above the equator. The calculation includes the gravitational perturbations of Earth's non-spherical part, the Sun, and the Moon, and the solar radiation pressure perturbation. After extrapolating 40,000 days (about 110 years) the satellite still maintains in a region above the equator's minor axis (longitude 75° E). The results are

$$\Delta \lambda \in [-43^\circ.13, +41^\circ.32], \Delta \varphi \in [-15^\circ.28, +15^\circ.29].$$

The period of east–west swing (i.e., the variation of $\Delta \lambda$) is about 900 days, over this time interval the variation of the orbital semi-major axis a is $\Delta a \approx \pm 25 \text{ km}$ (Figs. 5.6 and 5.7).

There is another example for showing the resonance of $a - \bar{\lambda}$. The initial epoch t_0 is 26 January 2013 00:00:00 (UTC), the initial orbital elements are

$$a_0 = 42173.8436 \text{ km}, e_0 = 0.00001, i_0 = 0^\circ.0005,$$

$$\Omega_0 = 0^\circ.0, \quad \omega_0 = 45^\circ.0, \quad M_0 = 0^\circ.0.$$

Results (without the solar radiation pressure perturbation) are shown in Fig. 5.8. The swing amplitude of $\bar{\lambda}$ is about $\pm 30^\circ$, and the resonance period is 874 days.

(2) Main properties of Mars-synchronous satellite orbit

Based on the period of Mars's rotation the orbit period and semi-major axis of its synchronous satellite are

$$T_s = 477^m.37772, a = 20427.68425 \text{ km} = 6.013448a_e,$$

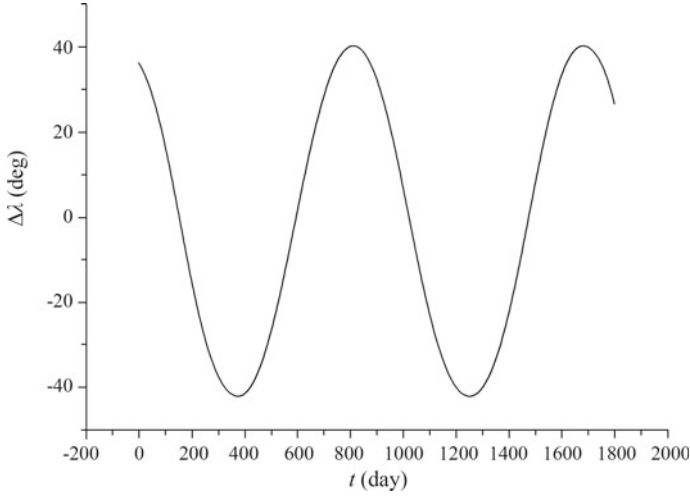


Fig. 5.6 Variations of the longitude from the initial longitude over time, $\Delta\lambda$ versus $(t-t_0)$

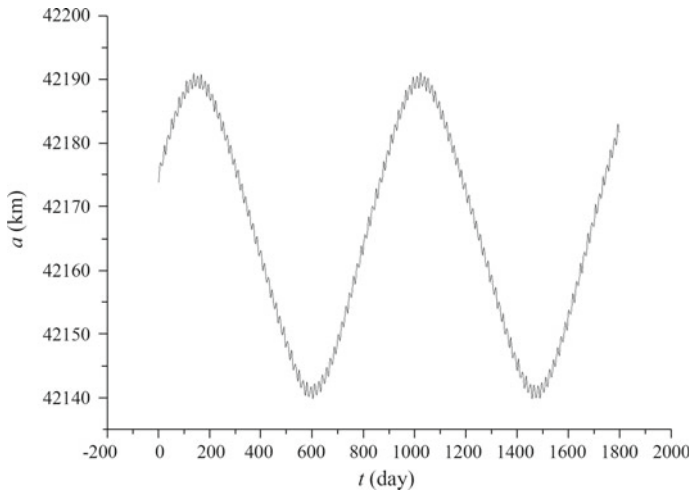


Fig. 5.7 Semi-major axis variations over time, a versus $(t - t_0)$

where a_e is the radius of Mars's equator. If the orbital inclination of a satellite is zero ($i = 0$), then this is a stationary orbit with respect to Mars. The perturbation of Mars's non-spherical gravity force (J_2 and $J_{2,2}$) is similar to that of Earth, but the ratio of $J_{2,2}$ to J_2 is larger, that the magnitude of $(J_{2,2}/J_2)$ is about 10^{-1} . The result is that the effect of Mars-synchronous orbital resonance is stronger than that of Earth, i.e., a spacecraft positioned above Mars's equatorial minor axis (Mars's longitude $164^\circ 0.7$ E) is more stable. Here is an example of numerical results.

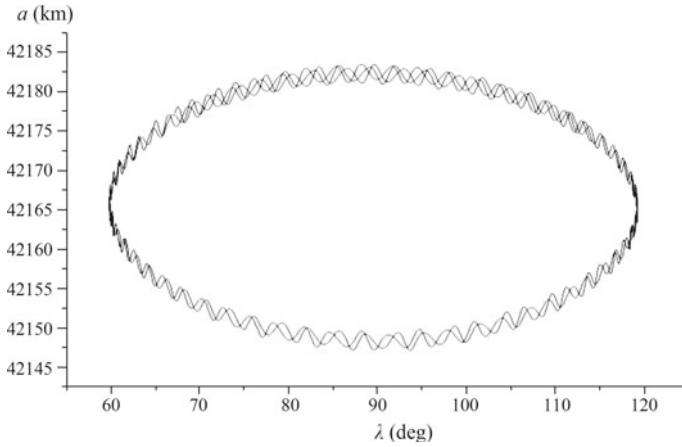


Fig. 5.8 Illustration of the resonance phenomenon of the geosynchronous satellite orbit a versus $\bar{\lambda}$

The initial epoch t_0 is 30 March 2010 4:00:00 (UTC), and the initial orbital elements are

$$a_0 = 20327.684233\text{km}, e_0 = 0.0001, i_0 = 0^\circ.005.$$

The initial deviation of the orbital semi-major axis from that of the synchronous orbit is 100 km, and the deviation of the motion period from the synchronous period is 11 min. Results are given in Figs. 5.9 and 5.10.

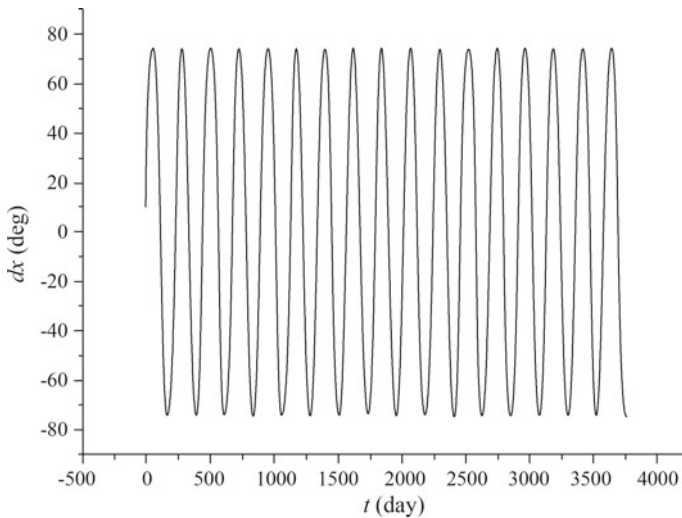


Fig. 5.9 Deviations of the longitude from the initial longitude over time, $\Delta\lambda$ versus $(t - t_0)$

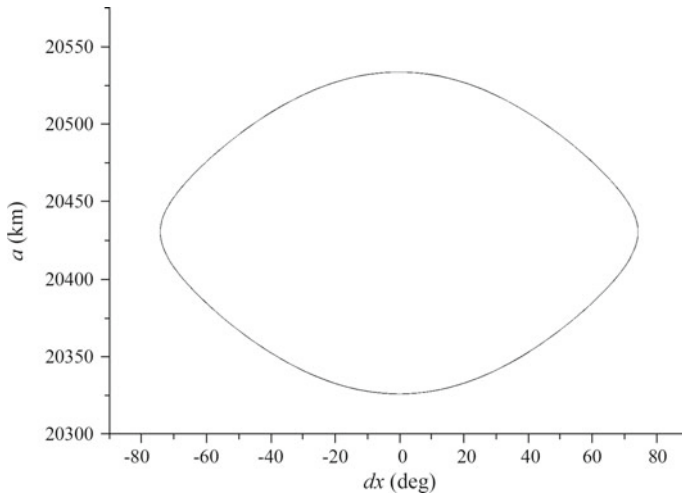


Fig. 5.10 The variation of the semi-major axis and the longitude deviation, a versus $\Delta\lambda$.

In these two figures, $dx = \lambda - \lambda_s$ is the deviation of longitude from $\lambda_s = 164^\circ.7$ which is the longitude of Mars's equatorial minor axis.

The calculation model used to provide the results in Figs. 5.9 and 5.10 includes perturbations of J_2 and $J_{2,2}$. Although the initial deviation is large, the satellite can still move in the vicinity of the space above Mars equator's minor axis. By comparison with the geosynchronous satellite orbit, because the ratio of the equatorial ellipticity $J_{2,2}$ term to the J_2 term of Mars is larger than that of Earth, the orbital resonance effect is stronger, and a spacecraft on the Mars's synchronous orbit is more stable than on the geosynchronous orbit.

5.6 Estimation and Calculation of Satellite's Lifespan Due to the Mechanism of Gravitational Perturbation

About a satellite's lifespan there are three possible reasons to end a satellite. Firstly, a satellite's body is disintegrated; secondly, a satellite's mission is completed (the working lifespan), and thirdly a satellite's orbital motion is stopped (the orbit lifespan). Since this book is about satellite orbits, we only discuss the lifespan of a satellite's orbital motion, i.e., the lifespan of a satellite's orbit.

For a low Earth orbit satellite, the effect of atmospheric dissipation decides its lifespan. The atmospheric drag makes the orbital semi-major axis and eccentricity smaller over time, i.e., the orbit becomes smaller and rounder, and finally the satellite drops into the denser atmosphere layer and falls. The question is what would happen if the central body does not have an atmosphere (like the Moon). In this section, we discuss how the universal gravity mechanism affects a satellite's lifespan. Under the

gravitational perturbations of the central body's non-spherical part and a third-body, the orbital eccentricity e has long-periodic variations, which can reduce the height of the periapsis, eventually, the satellite crashes to the central body, or if the central body has an atmosphere, the satellite falls by the effect of atmospheric dissipation. For a high orbit satellite, the orbital eccentricity increases due to the external perturbation from a third body, eventually when $e > 1$, the orbit becomes a hyperbola and the satellite escapes from the central body. The effect of atmospheric dissipation on the satellite's lifespan is discussed in Sect. 5.7.

5.6.1 Definition and Mechanism of a Low Orbit Satellite Lifespan Due to Gravitational Perturbations [6–10]

As described above the satellite's lifespan largely depends on the variation of the orbital eccentricity. According to the theoretical results given in Chap. 4, the orbital eccentricity e has a long-periodic variation $\Delta e_l(t)$ due to the gravitational perturbations of the central body's non-spherical part and a third body. For a low orbit satellite, the effect from the third body is not important and can be ignored, details are given in Chap. 4. The long-period variation of e due to the central body's non-spherical perturbation is given by (4.284) as

$$e_l^{(1)}(t) = -(1 - e^2) \sum_{l(2) \geq 3} \left(\frac{-J_l}{p_0^l} \right)^{\frac{1}{2}(l-2+\delta_1)} \sum_{p=1}^p \left[\sum_{q=1}^p (-1)^{(l+2q-\delta_1)/2} \left(\frac{1}{2} \right)^{(2l-2p+2q-1)} \right. \\ \left. \times \binom{l}{p-q} \binom{2l-2p+2q}{l} \binom{l-2p+2q}{q} (\sin i)^{(l-2p+2q)} \right] \frac{1}{e} K_3(e) I(\omega), \quad (5.61)$$

where $p_0 = a(1 - e^2)$, and in the sum $l(2)$ means the “step length” is 2, that $l(2) = 3, 5, \dots$. The auxiliaries $K_3(e)$ and $I(\omega)$ that

$$K_3(e) = \sum_{\alpha(2)=l-2p}^{l-2} \binom{l-1}{\alpha} \binom{\alpha}{(\alpha-l+2p)/2} \left(\frac{1}{2} \right)^\alpha e^\alpha, \quad (5.62)$$

$$I(\omega) = \left(\frac{n}{\omega_1} \right) [(1 - \delta_1) \cos(l - 2p)\omega - \delta_1 \sin(l - 2p)\omega]. \quad (5.63)$$

In the sum $\alpha(2)$ means the “step length” is 2, that $\alpha(2) = 1, 3, \dots$. The properties of the variation of the eccentricity can be represented using the J_3 and J_4 perturbation terms. The results of the long period variations of e due to J_3 and J_4 given by (4.339) and (4.345) in Chap. 4, are reprinted here that

$$e_l^{(1)}(t) = \frac{3}{4} \left(\frac{-J_3}{p^3} \right) \sin i \left(2 - \frac{5}{2} \sin^2 i \right) \left(\frac{n}{\omega_1} \right) (1 - e^2) \sin \omega, \quad (5.64)$$

$$e_l^{(1)}(t) = -\frac{35}{4} \left(\frac{-J_4}{p^4} \right) \sin^2 i \left(\frac{9}{14} - \frac{3}{4} \sin^2 i \right) \left(\frac{n}{2\omega_1} \right) e(1 - e^2) \cos 2\omega, \quad (5.65)$$

where

$$\omega_1 = \frac{3J_2}{2p^2} n \left(2 - \frac{5}{2} \sin^2 i \right).$$

The formulas (5.64) and (5.65) represent the effects of odd-terms (J_l , l is an odd number ≥ 3) and even-terms (J_l , l is an even number ≥ 4) of zonal harmonic perturbation terms, respectively, on the long-period variation of e . As we can see that the effect by the odd-terms does not have the factor e , whereas the effect by the even-terms has it. Therefore, the main effect on the satellite's lifespan is the odd terms of the central body's non-spherical zonal harmonic perturbation, especially the J_3 term.

According to (5.61) the main effect of the odd terms on the long-periodic variation of e can be presented as

$$\begin{aligned} \Delta e_l(t) &= e_l(t) - e_l(t_0) \\ &= \mu(i) [\sin \omega(t) - \sin \omega(t_0)] \end{aligned} \quad (5.66)$$

where $\mu(i)$ is

$$\mu(i) = \sum_{l(2) \geq 3} (-1)^{(l-1)/2} \frac{l-1}{3p^{l-2}} \left(\frac{J_l}{J_2} \right) F_l^*(i). \quad (5.67)$$

For J_3 term it is

$$\mu(i) = \frac{3}{4} \left(\frac{-J_3}{p^3} \right) \sin i \left(2 - \frac{5}{2} \sin^2 i \right) \left(\frac{n}{\omega_1} \right) (1 - e^2). \quad (5.68)$$

The formula (5.67) shows that there are two key factors, (J_l/J_2) and $F_l^*(i)$, decide the variation of the orbital eccentricity. It is understandable that why the effect of the even-terms has the factor e but the odd-terms have not. The reason is that the magnitude of the odd-terms of the zonal harmonic perturbation, J_{2l-1} ($l \geq 2$), reflects the degree of the asymmetry (including the shape and the mass distribution) of the northern hemisphere and southern hemisphere of the central body. We can conclude that the magnitude of the long-period variation of e depends on the ratio of J_{2l-1} ($l \geq 2$) to J_2 , and the entire effect depends on the variation of $\mu(i)$. In the expression of $\mu(i)$ the factor of $F_l^*(i)$ is a function of $\sin i$, we do not provide the actual expression here, details about $\mu(i)$ are in Sects. 4.5 and 4.6.

For a low Earth orbit satellite there is $O(J_3/J_2) = O(10^{-3})$, so the variation of $e_l(t)$ is small, as a result, the lifespan of the satellite is little affected or none by the non-spherical part of Earth. For a low orbit satellite of Mars or the Moon the magnitude of $|J_3/J_2|$ is relatively large, then the influence of $|J_3/J_2|$ becomes important, particularly for a low Moon orbit satellite. The Moon has no atmosphere to dissipate a satellite's energy, but the strong influence of the odd-term zonal harmonic perturbation J_{2l-1} ($l \geq 2$) on the eccentricity can reduce the height of the orbital perilune and cause a collision. The case of a low Mars orbit satellite is different from that of both Earth and the Moon. The magnitude of $|J_3/J_2|$ of Mars is relatively large which is similar to the case of the Moon, but the non-spherical gravity potential of Mars is different, so the mechanics which cause the reduction of the height of the perimartian are quite different from that of the Moon. Details about this topic are given in the [10].

The expressions of the long-period variations of e and ω indicate the importance of the odd-terms of the zonal harmonic gravitational perturbation. The effect of the J_3 term is obvious no matter how small e is. When e is small the effect of J_3 term on the long-period variation can balance the secular variation of the argument of the periapsis ω due to the J_2 term (see (4.81)). These internal rules are important to deciding the basic characteristics of satellite orbital variation. Definitely the relative magnitude of the two parameters of the central body non-spherical gravity potential, J_3 and J_2 , i.e., $|J_3/J_2|$, is a critical parameter, and its values for Earth, Mars, and the Moon are

$$\left| \frac{J_3}{J_2} \right| \approx O(10^{-3}), O(10^{-2}), O(10^{-2} - 10^{-1}), \quad (5.69)$$

which are used for analyzing related problems in the content below.

5.6.2 Overview of Low Orbit Satellite Lifespan for Earth, the Moon, Mars, and Venus

As shown in (5.66) that the variation of the eccentricity $\Delta e_l(t)$ strongly depends on $\mu(i)$. But for low Earth orbit satellites as discussed above the gravitational perturbation of Earth's non-spherical part has no effect on the satellite lifespan because $O(J_{2l-1}/J_2) = O(10^{-3})$ so $\mu(i)$ is small and has little influence. For low orbit satellites of Venus, Mars, and the Moon, the magnitudes of J_{2l-1}/J_2 are relatively larger, particularly for the Moon, then $\mu(i)$ is considerable.

For the Moon there is no atmosphere therefore the lifespan of a Moon satellite orbit mainly depends on the odd-term zonal harmonic perturbation J_{2l-1} ($l \geq 2$). The magnitude of the corresponding character parameter is

$$\left| \frac{\bar{J}_{2l-1}}{\bar{J}_2} \right| = O(10^{-1} - 10^{-2}), \quad l \geq 2, \quad (5.70)$$

which indicates that the amplitude of long-period variation of e can be as large as 0.05–0.1, which can strongly affect the lifespan of a low Moon orbit prober.

When the inclination of a low Moon orbit satellite $i \neq 0$, the long-period variation of e also depends on the variation of $\sum_{l(2) \geq 3} (J_{2l-1}/J_2) F_l^*(i)$, which has multiple maxima and minima. As a result, the lifespans of low Moon orbit satellites at a similar height but with different inclinations can be totally different. For example, there are two low Moon orbit satellites, each is on a near circular orbit with an initial height of perilune $h_p = 100$ km. One orbit has an inclination of $i_0 = 90^\circ$, the other $i_0 = 40^\circ$, and the corresponding lifespans are only 172 days and 48 days (i.e., when the satellite collides on the Moon), respectively. But for the same kind of orbit with a different inclination, $i_0 = 85^\circ$, after 50 years the lowest height of the perilune is about 60 km, which means there is no collision. These examples indicate that the Moon's non-spherical gravity potential with respect to the equator is asymmetrical, also its mass distribution is uneven, which has mass lumps.

The above theoretical analysis can be verified by numerical calculations. In the Refs. [7] and [8], calculations for low Moon orbit satellites (mean height is 100 km) were performed for the following two cases using the Moon gravitational field model of LP75G by JPL.

- (1) Based on the analytical solution due to the Moon's non-spherical gravity perturbation mechanism the calculations are for the values of $|\Delta e_i|$ for every i from $i = 0^\circ 0.5, 1^\circ 0.5, 2^\circ 0.5, \dots, 179^\circ 0.5$, to show the distribution of maxima and minima of $|\Delta e_i|$.
- (2) Using the complete motion equations which include main perturbations (the Moon's non-spherical gravity force, Earth's gravity force, and the Sun's gravity force) for a low Moon orbit satellite (mean height $\bar{h} = 100$ km, $e_0 = 0.001$) the calculations are for the variations of the height of the perilune related to different i_0 , to show the relationship of satellite lifespan with inclination i .

Results show that there are a few “stable regions” of $|\Delta e_i|$ corresponding to minima of $|\Delta e_i|$. The values of i related to minima $|\Delta e_i|$ are $i = 0^\circ, 27^\circ, 51^\circ, 77^\circ$, and 85° . By the property of $\sin i$, there are also corresponding “stable regions” when $i = 95^\circ, 103^\circ, 129^\circ$, and 153° . If i_0 is selected to be one of the stable values, the low Moon orbit satellite has a long lifespan, otherwise, the lifespan is short. Results of the calculations prove this conclusion. Table 5.5 gives some of the results, in which T_c is the time interval during which the satellite's height changes from 100 km to the minimum value h_p . Details are given in [8, 9].

These results confirm the correctness of the analytical solution of the Moon satellite, which provides the relationship between the lifespan of the low Moon orbit and its inclination caused by the Moon's non-spherical gravity potential.

For Mars, its gravitational field has a character parameter greater than that of Earth, but similar to that of the Moon, as

Table 5.5 Relationships of the variations of low Moon orbit parameters and inclination i

i (deg)	T_c (day)	Max e	Min h_p (km)
1.0	2723.1	0.0362	33.9
27.0	2219.9	0.0419	23.6
51.0	1908.1	0.0337	38.5
77.0	3383.8	0.0381	30.4
85.0	1711.7	0.0220	59.6
95.0	1102.0	0.0172	68.3
10.0	42.5	0.0545	0.0
40.0	47.9	0.0547	0.0
60.0	88.2	0.0548	0.0
82.0	294.7	0.0547	0.0
90.0	172.0	0.0545	0.0
98.0	236.2	0.0546	0.0

$$\left| \frac{\bar{J}_{2l-1}}{\bar{J}_2} \right| \geq O(10^{-2}), \quad l \geq 2. \quad (5.71)$$

The perturbation of the zonal harmonic terms can make the eccentricity of a low Mars orbit have a long-period variation with an amplitude of 0.02–0.03, therefore, can affect the satellite's lifespan. For Mars, the function $\mu(i)$ in (5.67) is also important, which is different from Earth, but the value of $\mu(i)$ mainly depends on $\sin i$, which is different from the Moon. For a Mars's satellite, if it has a polar orbit the amplitude of the long-period variation of its eccentricity, $\Delta e_1(t)$, is much larger than that of an orbit of small inclination, meaning that the height of its perimartian of a polar orbit satellite drops fast. This phenomenon reflects that although the asymmetry of Mars's non-spherical gravity potential with respect to the equator is more obvious than that of Earth, its mass distribution is relatively even unlike that of the Moon. Table 5.6 lists some numerical results (after integrating over an arc of one year), showing that by the gravitational mechanism, only the critical height of the initial perimartian height for a collision between a low Mars orbit satellite and Mars is 80–85 km. The data in Table 5.6 reveal the falling states of the perimartian for different inclinations, which are obviously different from the Moon's satellites.

The content in this section is about the fundamental orbital information, which is the necessary knowledge for working on the overall design of a circling prober around a target celestial body in deep-space exploring. It should never assume that the rules applied to Earth's satellites can also work for satellites of other celestial

Table 5.6 Relationship of characteristics of orbit variations with the inclination i

h (km)	i (deg)	t_c (day)	Max e	Min h_p (km)
80	90	19.2	0.0207	0.0
	45	290.3	0.0157	22.481
	5	316.3	0.0065	59.571

bodies. Each celestial body has its own physical properties, we must follow basic physical principles to deal with each problem.

For a low Venus orbit satellite, the corresponding character parameter, $O(|J_{2l-1}/J_2|) = O(10^{-1})$, is large, which can cause a large long-period variation of e then affect the lifespan of a satellite. But the deviation of the shape of Venus from a sphere is small, as a result, the long-period term of e has a very long period, almost like a secular term, therefore the effect of the long-period term is rather slow. We now use numerical calculations for two cases to check the height variation of periapsis of Venus's satellite assuming that the Venus's atmospheric influence is negligible. The epoch i_0 is 11 January 2011 00:00:00 (UTC), the initial height of periapsis (h_p) and the height of the apoapsis (h_a), and two inclinations are selected as

$$h_p = 700\text{km}, h_a = 800\text{km}, i = 10^\circ.0, 85^\circ.0.$$

Results of the calculated variations of h_p are illustrated in Figs. 5.11 and 5.12.

In the two figures, there are areas where $h_p \leq 0$, which actually means when $h_p = 0.0$ km, the calculation continued in order to know the details of the long-periodic variation. The figures show that the period of the long-period variation of the eccentricity is indeed very long, and after a certain time the height of h_p drops to zero. But the whole process is slow. When $i = 10^\circ.0$, it takes 2,068 days for the satellite crashes to Venus (at $e = 0.0013$); when $i = 85^\circ.0$, it takes 4,545 days. The effect of the inclination on the lifespan of the satellite reflects the uneven mass distribution of Venus.

Results of calculation for perturbations of Venus's non-spherical gravity and the Sun's gravity show little change from results without the Sun's gravity. For example,

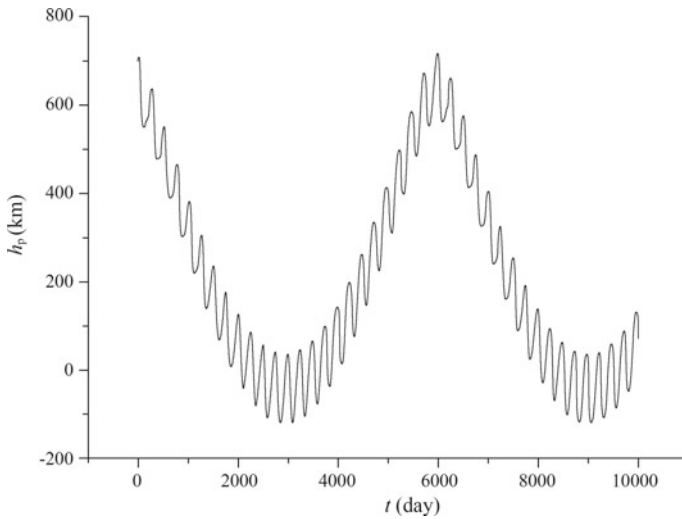


Fig. 5.11 The variation of h_p for initial values: $h_p = 700$ km, $h_a = 800$ km, and $i = 10^\circ.0$

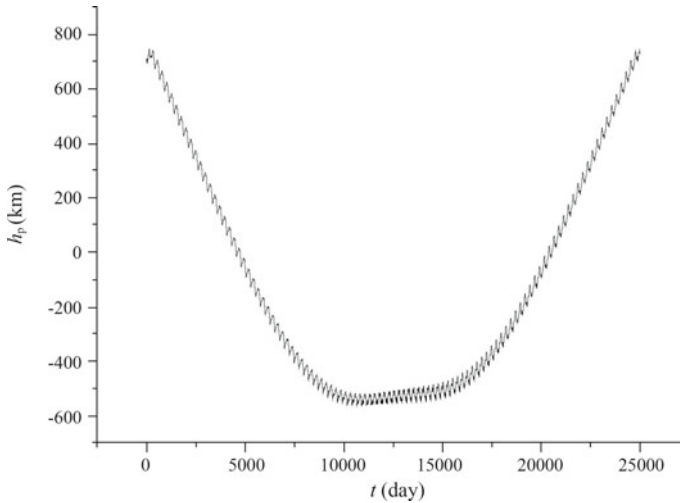


Fig. 5.12 The variation of h_p for initial values: $h_p = 700$ km, $h_a = 800$ km, and $i = 85^\circ.0$

when $i = 10^\circ.0$, considering Venus's non-spherical gravity perturbation, it takes 2,067.9 days for the value of h_p to drop to zero and $e = 0.1103$; when considering both Venus's non-spherical gravity perturbation and the Sun's gravity perturbation it is 2,069.8 days. Obviously, the difference is very small, showing that the main perturbation which affects the satellite's lifespan is Venus's non-spherical gravity force.

5.6.3 Evolution Characteristics and Lifespans of Orbit with a Large Eccentricity [2, 9]

There are spacecraft whose orbits have a large eccentricity but a low height of periapsis, usually $h_p \approx 200$ km. For example, the GTO (Geostationary Transfer Orbit) debris and the twin satellite, TC-1 and TC-2 (Tan-Ce), whose height of periapsis is a little higher than 200 km, but the eccentricity is large, that $e > 0.7$, therefore, over one period of moving around Earth, the spacecraft mostly is far away from Earth. In this situation the effects of both the atmospheric drag and Earth's non-spherical gravity force (except the J_2 term) are weak, the main perturbations are from the J_2 term and a third-body (the Sun or the Moon) attraction. These two perturbation forces are conservative, so have no dissipation effect, then we should check the effect of the orbit eccentricity on the lifespan of this kind of satellite.

(1) Two examples

We now use the numerical method to examine the orbital evolution of a highly eccentric orbit. We put two debris on “reserved” GTOs as GTO-1 and GTO-2, in three hours apart. Numerical calculations show that their lifespans are totally different, the lifespan of GTO-1 is only 160 days, whereas that of GTO-2 is 50 years. The calculations are for epoch 21 December 2006 10:00:00 and 13:00:00 (UTC) for GTO-1 and GTO-2, respectively; and the orbit elements in the Earth-fixed coordinate system are

$$h_p = 200\text{km}, \quad h_a = 36000\text{km}, \quad i = 28^\circ.5, \quad \Omega_G = -10^\circ.0,$$

$$\omega = 180^\circ.0, \quad M = 0^\circ.0$$

and the corresponding semi-major axis $a = 24,478.136$ km, and the eccentricity $e = 0.73126482$. Here “Earth-fixed” means the longitude of ascending node Ω_G is measured from the Greenwich meridian, in the corresponding epoch geocentric coordinate system (J2000.0), the longitudes of ascending node are $\Omega = 229^\circ 0.837399$ and $\Omega = 274^\circ 0.96061$, for the two debris, respectively.

As mentioned above, because of the large eccentricity the two debris move mostly in the space far away from Earth, thus the perturbation effects on the debris are different from those on a low Earth orbit satellite. The dynamical model used for calculations includes the following perturbations:

- ① The Earth’s non-spherical gravity force, $J_{l,m}$ (i.e., $C_{l,m}, S_{l,m}, l = 2, 3, 4, m = 0, 1, \dots, l$) terms.
- ② The Sun and the Moon gravity forces, the orbits of the Sun and the Moon are the average orbits including secular and long-periodic variations.
- ③ The atmospheric drag.

The atmospheric density is from an exponential model including diurnal variation; values of the density and the density scale height at reference points are from the International standard atmosphere model. The product of the equivalent area-to-mass ratio and the coefficient of resistance $C_d S/m$ is 10^9 in the normalized unit system (equivalent to an area of 1 m^2 and a mass of 100 kg). In the calculation, we assume that the orbit ends its life at the time when h_p drops to 100 km .

For a piece of debris with $C_d S/m = 10^9$, the effect of the solar radiation pressure does not exceed the effect of the third-body gravity force, therefore whether to include this perturbation or not has no fundamental influence on the problem discussed here. There are some pieces of debris with a large S/m value, such as $S/m = 1 \text{ m}^2/1 \text{ kg}$, in the normalized units it is about 10^{11} , then the solar radiation pressure cannot be ignored, but we do not deal with this problem in this section.

We assume the calculation arc to be $t - t_0 = 50$ years, use the high precision RKF7(8) integrating algorithm, and accept the even step-length method ([11] and Chap. 8 in this book). Results are listed in Table 5.7.

Table 5.7 Orbital evolution states of the two GTO debris

Model	$J_{l,m}$	Sun, Moon	Atm	Debris	t (day)	a (km)	e	h_p (km)
1	✓	×	✓	GTO-1	11,643.5	6518.2	0.0063	98.9
				GTO-2	11,791.8	6508.3	0.0055	94.6
2	✓	✓	×	GTO-1	144.7	24,411.3	0.7346	100.0
				GTO-2	18,262.5	24,386.3	0.7263	295.4
3	✓	✓	✓	GTO-1	160.7	22,235.9	0.7088	99.9
				GTO-2	17,981.9	6796.5	0.0468	100.0

In Table 5.7, the sign “✓” implies the perturbation force is included, and the sign “×

The data given using Model 1 in Table 5.7 and displayed in Figs. 5.13 and 5.15 show that the effect of the atmospheric dissipation is not the main factor to decide the lifespan of the GTO type orbit. The atmospheric effect is weak, the height of the perigee reduces very slowly (results by Model 1), and the orbital variations of the two GTO debris are similar although their initial conditions are different. The similarities are not only in the orbit’s lifespan but also in the detailed variations of orbit elements, that both a and e become smaller gradually (e becomes smaller slower than a due to a factor of e in the variation rate of a), and the related h_p becomes lower.

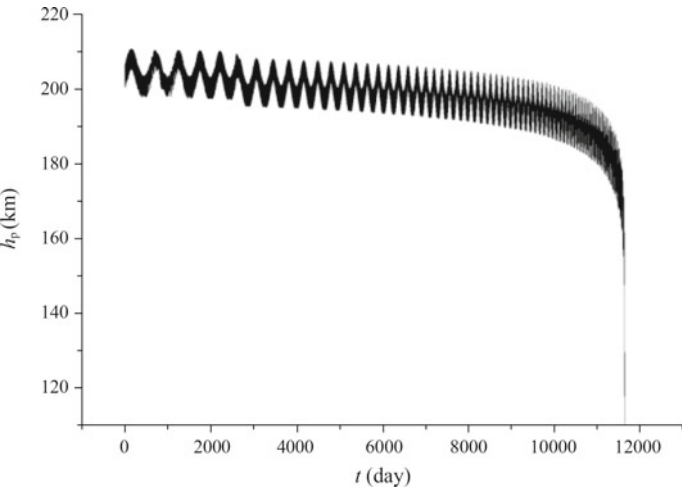


Fig. 5.13 The variation of h_p for GTO-1 using Model 1

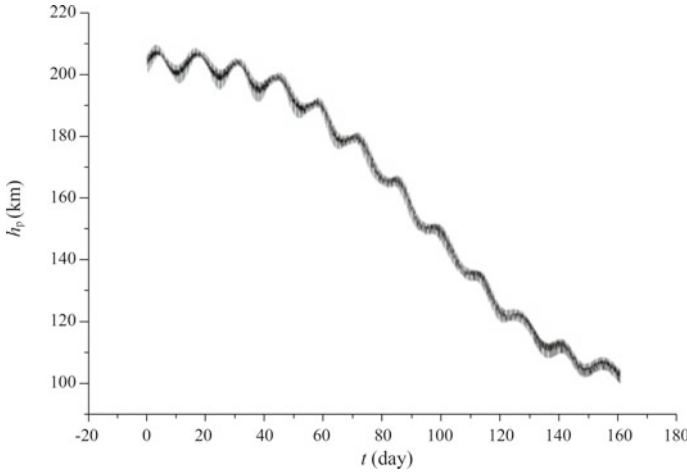


Fig. 5.14 The variation of h_p for GTO-1 using Model 3

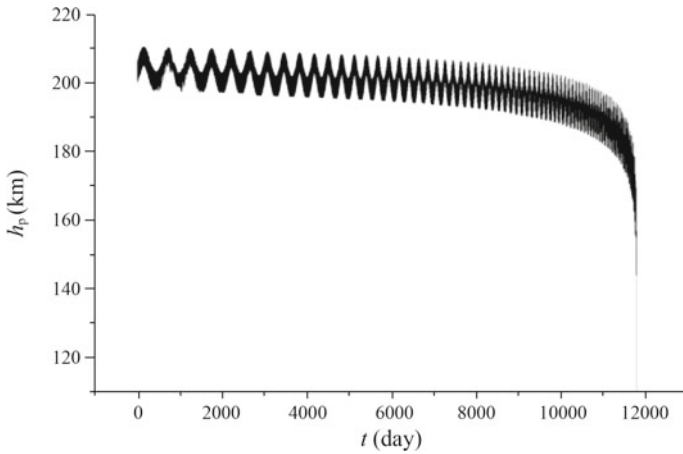


Fig. 5.15 The variation of h_p for GTO-2 using Model 1

Based on the results about the lifespan of GTO type with high eccentricity, we must look deep into the effect of the third-body perturbation using Model 2 and Model 3, especially Model 2, for revealing the importance of the third-body perturbation. The theoretical analysis based on the orbital dynamics principles is given as follows.

(2) Theoretical analysis

Results from the numerical calculations show that the influence of short-period perturbation terms is small. Theoretically, the magnitude of the short-period term is in the order of 10^{-5} – 10^{-4} , no matter it is by Earth's non-spherical gravity force or by the third-body (the Sun or the Moon) gravity force, therefore it is not strong

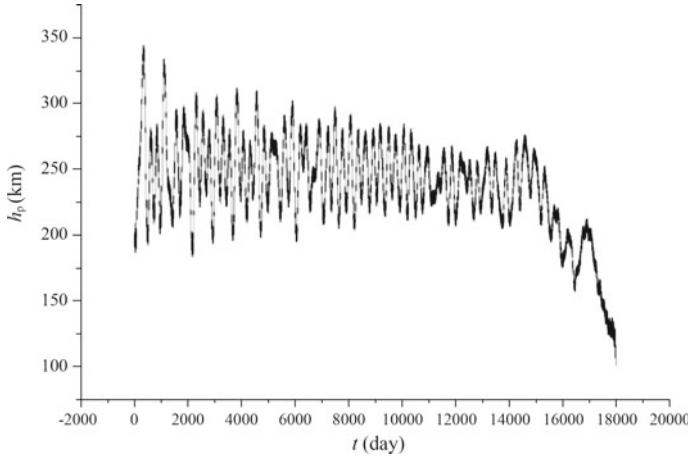


Fig. 5.16 The variation of h_p for GTO-2 using Model 3

enough to change the satellite's lifespan. These two perturbation forces are conservative forces, which have no dissipation effects. Therefore, the question is focused on the effects of long-period terms. In fact, the semi-major axis a only has long-period terms of a higher order because $a_l^{(1)}(t) = 0$, [2] what left to be checked is the long-period variation of e .

The main terms in the long-period variation of e are due to J_2 , J_3 , and J_4 , and the potentials of the Sun and the Moon. Using the normalized units, the long-period variations of e are given as (Sects. 4.5 and 4.7)

$$\begin{aligned}
 e_l(t; J_l) = & \left(\frac{1}{a} \right) \sin i \left\{ \left(\frac{J_2}{4p^2} \right) \sin i \frac{1 + 2\sqrt{1-e^2}}{(1 + \sqrt{1-e^2})^2} e \cos \omega \right. \\
 & + \left(\frac{3}{8p} \right) \frac{\sin i}{(4 - 5 \sin^2 i)} \left[J_2 \left(\frac{7}{3} - \frac{5}{2} \sin^2 i \right) \right. \\
 & \left. \left. + 35 \left(\frac{J_4}{J_2} \right) \left(\frac{1}{7} - \frac{1}{6} \sin^2 i \right) \right] e \cos \omega - \frac{1}{2} \left(\frac{J_3}{J_2} \right) \sin \omega \right\}, \quad (5.72)
 \end{aligned}$$

$$\begin{cases} \Delta e(t, \beta_2) = - \left(\frac{3}{2} \beta_2 a^3 \right) (5e\sqrt{1-e^2} S_3) n(t - t_0), \\ S_3 = AB. \end{cases} \quad (5.73)$$

where $\beta_2 = m'/r'^3$; $S_3 = AB$ is a function of slow variables Ω , ω , and the orbit elements of the Sun and the Moon i' , Ω' , $u' = f' + \omega'$, the actual expressions of these variables are not listed here, details are given in (4.98), (4.99), (4.505), Chap. 4 and [1].

The variation of the orbital periapsis's height of the GTO debris depends on the variations of the semi-major axis a and the eccentricity e , that $h_p = a(1 - e) - a_e$. Under these two perturbations (the gravity force of Earth's non-spherical part and the gravity force of a third body), the semi-major axis a only has a periodic variation in the order of 10^{-4} – 10^{-3} , which is not important, therefore the determining factor is the long-period variation of e . The semi-major axis of the GTO debris a is about 24,500 km, if the variation of e , $\Delta e \geq 0.004$, even if a is unchanged, the height of the periapsis h_p can be reduced to 100 km. Formulas (5.72) and (5.73) show that the influences from the Sun and the Moon are greater than that from the J_2 term. By estimations, we can see that the magnitude of e_l due to the third-body perturbation by (5.73) is about 0.004, which is greater than that of e_l due to the three J_1 terms by (5.72), which is about 10^{-4} – 10^{-3} .

In Table 5.7 numerical results using Model 2 are totally different for the two pieces of debris that GTO-1 fell into the dense atmosphere after 145 days, whereas GTO-2 keeps its h_p above 100 km after 50 years. The difference is caused by the initial condition. The long-period variable $e_l(t)$ is related to slow variables such as Ω and ω , for example, $e_l(t_0)$ depends on ω . In Figs. 5.17 and 5.18 we plot $e_l(t_0)$ as a function of ω for the two debris.

Figure 5.17 shows that at time t_0 GTO-1 is in the state of $e_l(t_0) \approx 0$, and $e_0 = e(t_0)$ which is near the mean value of e according to its long-period variation; whereas from Fig. 5.18 at time t_0 GTO-2 is in the state of a maximum $e_l(t_0)$, corresponding to a maximum e_0 . Therefore, for GTO-1 its e increases over time t , and e_l can be as large as 0.004, soon resulting $h_p = 100$ km as shown in Table 5.7 Model 2. For GTO-2, it is the opposite, at time t_0 its e is near its maximum value and corresponding to $h_p = 200$ km, without atmosphere influence the variation of e does not reduce h_p to

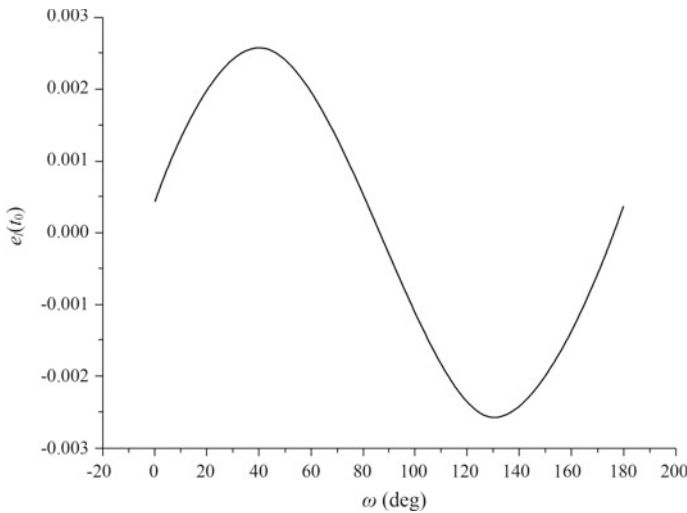


Fig. 5.17 The state curve of $e_l(t_0)$ versus ω for GTO-1

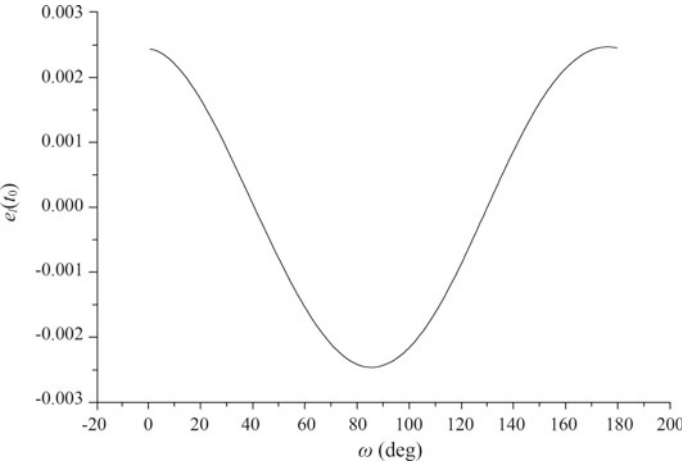


Fig. 5.18 The state curve of $e_I(t_0)$ versus ω for GTO-2

100 km as shown in Table 5.7 Model 2. Based on this mechanism when we add the dissipation effect of the atmosphere for GTO-1, which is the case of Model 3 in Table 5.7, the atmosphere makes e smaller and the overall increase of e is also smaller, and the satellite lifespan increases slightly. For GTO-2, the atmosphere makes the semi-major axis much shorter and h_p lower in the case of Model 3, but compared to Model 1, the speed of h_p decreasing is slower, it takes more than 49 years to reach 100 km. As shown in Table 5.7, the lifespan of GTO-2 is longer by Model 3 than by Model 1 due to the third-body perturbation.

(3) Numerical confirmation of the theoretical analysis

To check the above theoretical analysis, we change the initial conditions for the two GTO debris. For GTO-1, we change ω_0 from 180° to 45° , corresponding to a maximum $e_I(t_0)$ and a maximum e_0 . For GTO-2, we change ω_0 from 180° to 90° , corresponding to a minimum $e_I(t_0)$. According to the theoretical analysis, the lifespans of GTO-1 and GTO-2 would be opposite to the original states, i.e., GTO-2 has a short lifespan, whereas GTO-1 has a long lifespan. Numerical results are shown in Table 5.8.

Table 5.8 Orbital variations of two GTO debris after changing initial conditions

Model	$J_{l,m}$	Sun, Moon	Atm	Debris	t (day)	a (km)	e	h_p (km)
2	✓	✓	×	GTO-1	18,262.5	24,417.2	0.7258	321.1
				GTO-2	281.0	24,499.0	0.7356	99.9
3	✓	✓	✓	GTO-1	18,262.5	19,314.2	0.6570	251.5
				GTO-2	297.1	22,587.7	0.7132	99.9

Also, not given in Table 5.8 is that for GOT-1 by model 2 at $t = 12,977.6^d$ h_p reaches its minimum value of 177.4 km, and by model 3 at $t = 28.1^d$ h_p reaches its minimum value of 191.8 km. It is clear that the numerical results agree with the analytical results.

The results of the theoretical analysis can also be applied for the orbital evolution and lifespan of the twin satellites TC-1 and TC-2. Details are given in [10].

5.6.4 Evolution Characteristics and Lifespans of High Earth Satellite Orbit [6, 10]

For a high Earth satellite orbit, the effect of third-body perturbation increases. For example, for a geosynchronous satellite of $a = 6.6a_e$ (a_e is the radius of Earth's equator), the influence of the Sun and the Moon is in the order of 10^{-5} – 10^{-4} which is close to the influence of Earth's dynamical form-factor J_2 term. The variation of the eccentricity of a GEO is mainly due to the gravity of the Sun and the Moon. Under the influence of the third body, the amplitude of the variation of e contains a factor of $\sin^2 i / \dot{\omega}$, where $\dot{\omega}$ is the rate of the secular variation of ω due to Earth's non-spherical gravity and the third-body gravity, whose minimum magnitude is about 10^{-5} . Therefore, the magnitude of $e_l(t)$ is large, especially for a polar orbit satellite of $i = 90^\circ$.

For a high Earth orbit (HEO) satellite when the semi-major axis $a \geq a_c$ (a_c is the critical value of the semi-major axis for an Earth's satellite), the large amplitude of the long-period variation of the eccentricity can cause the height of perigee $h_p = r_p - a_e$ near zero, and the satellite “crashes” to Earth (or when h_p is below 200 km, the satellite quickly enters into the dense atmosphere and falls). The critical value of the semi-major axis, a_c (Refs. [6] and [10]), is given by

$$a_c = \left[2 \left(\frac{\mu}{\mu'} \right) \left(\frac{a'}{a_e} \right)^3 J_2 \right]^{\frac{1}{5}} a_e, \quad (5.74)$$

$$\mu = GE, \quad \mu' = GM. \quad (5.75)$$

In the formula, GE is Earth's gravitational constant, GM is the gravitational constant of the Sun (or the Moon), a' is the semi-major axis of the orbit of the Sun or the Moon with respect to Earth's center. By (5.74) for Earth's satellites, there is $a_c = 8.2a_e$. To check this criterion, we use numerical calculations for $a = 8.5a_e$, $7.5a_e$, and $6.6a_e$, and $i = 90^\circ$, $28^\circ.5$ and 1° . The initial epoch is 21 December 2006 10:00:00 (UTC), and the initial orbit elements are listed in Table 5.9. The calculations are for an arc of 100 years, when the perigee h_p reduces to 200 km, the calculation ends, and the corresponding time is the approximate orbit lifespan. The dynamical model for the calculations includes Earth's non-spherical gravity terms $J_{l,m}$ ($l = 2$,

Table 5.9 Initial orbit elements of the three high Earth orbit satellites at the epoch t_0

Type	a (km)	e	i (deg)	Ω (deg)	ω (deg)	M
$a = 8.5a_e$	54,214.200	0.001	90, 28.5	230.0	180.0	0.0
$a = 7.5a_e$	47,836.000	0.001	90, 28.5	230.0	180.0	0.0
$a = 6.6a_e$	42,165.000	0.001	90, 1.0	230.0	180.0	0.0

Table 5.10 Characteristics of orbit variation of the three high Earth orbit satellites

Type	i_0 (deg)	t (day)	a (km)	e	h_p (km)
$a = 8.5a_e$	90	19,385.4	54,215.2	0.8787	200.0
	28.5	11,383.6	54,209.6	0.0034	47,649.1
$a = 7.5a_e$	90	24,543.4	47,836.2	0.8265	1921.0
	28.5	28,583.9	47,832.9	0.0020	41,385.2
$a = 6.6a_e$	90	32,735.2	42,139.6	0.6643	7,768.9
	1.0	10,145.4	42,132.4	0.0015	35,692.1

3, 4, $m = 0, 1, \dots, l$) and the Sun and the Moon gravity forces. The results are given in Table 5.10.

The results in Table 5.10 show the following properties of orbital variation.

- (1) When $a > a_c$ ($a = 8.5a_e$) and $i = 90^\circ$, the perturbation of the third body pulls the satellite's orbit to make the orbit oblate (although the semi-major axis changes little, the eccentricity is obviously increased), and the corresponding height of the perigee h_p reduces. With the atmospheric drag in the lower atmosphere, the satellite soon enters the dense layers of Earth's atmosphere and falls. When the initial value of a is much smaller than a_c (such as $a = 6.6a_e$) and $i = 90^\circ$, the variation of the orbital eccentricity is not large enough to make h_p near Earth, and the satellite has a long lifespan.
- (2) It is obvious that for an HEO the variation of e is related to the initial inclination i . When $i = 90^\circ$ the influence of i on the eccentricity is the largest. But only when $a > a_c$, the value of h_p reduces enough to affect the lifespan of the orbit. When $a = 6.6a_e$ to have $h_p \leq 200$ km requires the value of e reaching to 0.84, but after 100 years even for $i = 90^\circ$ the maximum of e is only 0.6643. The results show that the variation amplitude of e is related to the inclination as predicted by the above analysis, it reaches its maximum when $i = 90^\circ$. For a small inclination orbit, the variation of e of an HEO is small, and the satellite's lifespan is long as shown in Table 5.10.

5.6.5 Key Points About Estimating Satellite Orbit Lifespan Due to Gravitational Perturbations

In the previous four sections, the orbital lifespan is estimated for some typical satellite orbits to reveal the regular patterns of orbital variations due to gravitational perturbations. The conclusions provide the necessary theoretical foundation for specific orbital designs of the actual aerospace project. Based on these patterns we then use the complete dynamical model to further perform precision calculations, and to make the design reach the requirement of the project, including tracking and controlling.

The key points of estimating satellite orbit lifespan due to gravitational perturbations are summarized as follows.

(1) Low satellite orbits

First to estimate the relative magnitudes of the odd-terms to J_2 term of the non-spherical zonal harmonic perturbation of the central body, especially the magnitude of $O(J_3/J_2)$. Then to analyze the variation states of each zonal harmonic odd-term to provide the information of the degree of the unevenness of the mass distribution. The unevenness of the mass distribution makes the lifespan of an orbit depend on its inclination.

(2) High eccentric satellite orbits

The mechanism of the long-period variation of the eccentricity e is the main factor that decides the lifespan of a satellite of this kind. This mechanism involves the central body's non-spherical perturbation and a third-body perturbation. The initial conditions of orbital elements, i.e., the initial states of e , Ω , and ω , are also related to the long-period variation of e .

(3) High satellite orbit

The effect of a third-body perturbation increases for this kind of satellite. If a geosynchronous satellite with $a = 6.6a_e$ the magnitude of the Sun and the Moon perturbations can be about $10^{-5} - 10^{-4}$, which is near the magnitude of the J_2 term. The third-body perturbation makes the amplitude of the long-period variation of e increase, and the orbit becomes oblate, the apogee increases to a distance beyond the range of Earth's gravity field, and the satellite can escape from Earth.

5.7 Estimation and Calculation of Satellite Orbit Lifespan in the Perturbed Mechanism of Atmospheric Drag

We use an Earth's satellite as an example. The dissipation effect of Earth's atmosphere is the main factor that decides the lifespan of a low Earth orbit satellite. As mentioned above under the influence of the atmosphere the values of both the semi-major axis and eccentricity of an orbit decrease over time, i.e., the orbit becomes

smaller and rounder, and finally the satellite enters into the dense layers of the atmosphere and falls. Because of the complicity of the atmosphere, there is not any reliable atmospheric dynamical model for this region (altitudes between 100 and 200 km), therefore we cannot provide accurate orbital variations. Usually, we assume that a low Earth orbit satellite ends its life when its perigee, $h_p = a(1 - e) - a_e$, drops to 200 km.

The atmosphere above 200 km is in a state of free molecular flow, in which the motion of an artificial satellite is affected by the atmospheric drag, the acceleration of the resistance is given by

$$\vec{D} = -\frac{1}{2} \left(\frac{C_D S}{m} \right) \rho V \vec{V}. \quad (5.76)$$

Details about the atmospheric drag are given in Sect. 4.11. In (5.77), $C_D = 2.2 \pm 0.2$ is the drag coefficient; S/m is the area-to-mass ratio of a satellite (S is the equivalent area with respect to the drag); ρ is the local atmospheric density at the altitude where the satellite is; $\vec{V} = \vec{v} - \vec{v}_a$ is the velocity of the satellite relative to the atmosphere, $\vec{v}$ and $\vec{v}_a$ are velocities of the satellite and the atmosphere relative to Earth's center, respectively.

By these assumptions, the perturbation solution is given in Chap. 4. The secular variation rates of the semi-major axis and the eccentricity are given by (4.837) and (4.838), respectively, and are rewritten here as

$$\begin{aligned} a_2 = -B_1 a^2 n \left\{ (I_0 + 2eI_1) + C(\cos 2\omega I_2) + \mu z_0^2 \left(\frac{3}{4} I_0 - I_1 + \frac{1}{4} I_2 \right) \right. \\ \left. + F^* A^* \left(\frac{e}{2} I_0 + I_1 + \frac{3}{2} e I_2 \right) \right\} \end{aligned} \quad (5.77)$$

$$\begin{aligned} e_2 = -B_1 a n \left\{ \left(\frac{e}{2} I_0 + I_1 + \frac{e}{2} I_2 \right) + \frac{C}{2} \cos 2\omega (I_1 + I_3) \right. \\ \left. + \mu z_0^2 \left(-\frac{1}{2} I_0 + \frac{7}{8} I_1 - \frac{1}{2} I_2 + \frac{1}{8} I_3 \right) + F^* A^* \left(\frac{1}{2} I_0 + \frac{e}{2} I_1 + \frac{1}{2} I_2 + \frac{e}{2} I_3 \right) \right\} \end{aligned} \quad (5.78)$$

All variables in the above formulas are given in Sect. 4.11.

The above-given formulas are only for estimations. The reasons are of many aspects, including the atmospheric model, the satellite shape and attitude, errors in orbital information by the ground-based orbit controlling system, and non-linear error accumulations during the extrapolation over long arcs, etc. Even by the precision extrapolation method, the error accumulation is unavoidable and serious. Therefore, for the orbital lifespan of a satellite in this atmospheric region, we can only approximately estimate by orbital extrapolation.

About this atmospheric region (altitude $h = 100$ – 200 km), to the present days, the dynamical model for orbital extrapolation is still not available. The state of the

atmosphere can be divided by the Knudsen number, K_n , as mentioned in Sect. 4.11 by (4.771) that

$$K_n = \frac{\lambda}{L}, \quad (5.79)$$

where λ is the mean free path of atmosphere molecular, L is the characteristic scale of a satellite. Further research in the region is needed.

References

1. Liu L (1992) Orbital dynamics of artificial earth satellite. Higher Education Press, Beijing
2. Liu L (2000) Orbital theory of spacecraft. National Defense Industry Press, Beijing
3. Liu L, Tang JS (2015) Orbital theory of satellites and applications. Electronic Industry Press, Beijing
4. Cook GE (1966) Perturbations of near-circular orbits by the earth's gravitational potential. Planet Space Sci 14(3):433–444
5. Liu L, Tang JS (2012) Orbital motions of circling spacecraft around the major planets, the Moon, and asteroids. Space Eng 21(4):4–15
6. Wang X, Liu L (2002) Another mechanism of restricting the lifetime of orbiting satellites. ACTA Astron Sin 43(2):189–196, and China Astron. Astrophys. 2002 26(4):489–496
7. Wang X, Liu L (2002) Another mechanism of restricting the lifetime of orbiting satellites (continued) ACTA Astronomica Sinica 43(4):379–386, and China Astron. Astrophys 2003 27(1):107–113
8. Liu L, Wang HH (2006) Two problems about the motion of low-moon-orbit satellites. ACTA Astronomica Sinica 47(3):275–283, and China Astron. Astrophys 2006 30(4):437–446
9. Liu L, Wang X (2006) Orbital dynamics of spacecraft for the moon's exploration. National Defense Industry Press, Beijing
10. Liu L, Tang JS (2009) Variation characteristics of satellite's orbit eccentricity and its effect on orbital lifespan. Prog Astron 27(1):58–69
11. Fehlberg E (1968) Classical fifth-, sixth-, seventh- and eighth-order runge-kutta formulas with stepsize control. NASA TR R-287

Chapter 6

Orbital Solutions of Satellites of the Moon, Mars, and Venus



6.1 Characteristics of Gravitational Fields of Earth, the Moon, Mars, and Venus

In this section, we use three major planets, Earth, Mars, Venus, and the Moon as examples to show the characteristics of non-spherical gravity potentials of these celestial bodies, which are the background information for building analytical perturbation solutions of spacecraft's orbits around them [1–4]. The data related to Earth's gravity potential are reference parameters for constructing perturbation solutions of satellite orbits for other celestial bodies, in the same time, to show the details of constructing the analytical perturbation solutions and main reasons for the significant differences in characteristics of the solutions.

6.1.1 Basic Characteristics of Earth's Gravity Potential

The size, mass, and density of Earth are similar to these of Venus, and both planets have dense atmospheres. Earth's rotation is faster than Venus but is similar to Mars. The basic parameters of Earth are:

$$M(\text{mass}) = 1/332946.0 (\text{solar mass}),$$

$$P_s (\text{period of revolution}) = 365^{\text{d}}.25636306,$$

$$T_r (\text{period of rotation}) = 23^{\text{h}}56^{\text{m}}04^{\text{s}}.09053.$$

There are more than a dozen models of Earth's gravity potential developed using multi-decades information of ground-based and space-based measurements. The most influential ones are the Smithsonian Earth models (SAO-SE series, especially the SAO-III), the Goddard Earth Models (GEM series), the Joint Earth Gravity Modes (JGM series), and the World Geodetic System (WGS series using EGM-the Earth

Gravitational Models). By comparison, the GEM-T3 and JGM-3 have higher precisions. In recent years there are new Earth’s gravity potential models constructed by scientists in the United States and some European countries using high precision measurements, such as GGM-03C and GGM-03S (the Global Geo-potential Models), etc. For better understanding of Earth’s gravity potential models, Table 6.1 lists the three main parameters, GE (gravitational constant), a_e (radius of the Earth’s equator), and the dynamical form-factor J_2 , given by models of JGM, WGS, and GGM. In the models the flattening factor for Earth is $\varepsilon = 0.00335281$.

Some of the spherical harmonic coefficients of Earth’s gravity potential given by the WGS-83 model and related to the Earth-fixed frame, are listed in Table 6.2.

If we use $\overline{J}_l (l \geq 2)$ to replace $\overline{C}_{l,0}$, then the zonal harmonic coefficients have the following properties,

$$\begin{cases} \overline{J}_l = -\overline{C}_{l,0} \\ \overline{J}_2 = O(10^{-3}), \quad \overline{J}_l (l \geq 3) \leq O(10^{-6}). \end{cases} \tag{6.1}$$

Table 6.1 Main parameters given by three Earth’s gravity potential models

Parameter	JGM-3 (70 × 70)	WGS-84 (180 × 180)	GGM-03S (180 × 180)
$GE \text{ (km}^3/\text{s}^2\text{)}$	398,600.4415	398,600.4418	398,600.44150
$a_e \text{ (km)}$	6378.1363	6378.1370	6378.13630
J_2	$1.08262669 \times 10^{-3}$	$1.082636'022'' \times 10^{-3}$	$1.082635386 \times 10^{-3}$

Table 6.2 Main spherical harmonic coefficients in WGS-84 model

l	$\overline{C}_{l,m}$	m	$\overline{S}_{l,m}$
2	$-0.48416685 \times 10^{-3}$	0	0.0
	$0.24395796 \times 10^{-5}$	1	$-0.55364556 \times 10^{-8}$
	$0.85750651 \times 10^{-6}$	2	$-0.13979548 \times 10^{-5}$
3	$0.95706390 \times 10^{-6}$	0	0.0
	$0.20318729 \times 10^{-5}$	1	$0.25085759 \times 10^{-6}$
	$0.90666113 \times 10^{-6}$	2	$-0.62102428 \times 10^{-6}$
	$0.71770352 \times 10^{-6}$	3	$0.14152388 \times 10^{-5}$
4	$0.53699587 \times 10^{-6}$	0	0.0
	$-0.53548044 \times 10^{-6}$	1	$-0.47420394 \times 10^{-6}$
	$0.34797519 \times 10^{-6}$	2	$0.65579158 \times 10^{-6}$
	$0.99172321 \times 10^{-6}$	3	$-0.19912491 \times 10^{-6}$
	$-0.18686124 \times 10^{-6}$	4	$0.30953114 \times 10^{-6}$

6.1.2 Basic Characteristics of the Moon's Gravity Potential

The Moon is a large natural satellite without an atmosphere and rotates slowly. The basic parameters of the Moon are:

$$M(\text{mass}) = 0.01230002 \text{ (Earth's mass),}$$

$$P_s \text{ (period of revolution around Earth)} = 27^{\text{h}}.32166155 \text{ (Earth day),}$$

$$T_r \text{ (period of rotation)} = 27^{\text{h}}.32166155.$$

Because the Moon rotates slowly and its shape is more round than Earth, its dynamical form-factor J_2 is not only small but also almost in the same magnitudes as the “high-order” harmonic coefficients (including the tesseral harmonic coefficients). More important is that because of the slow rotation the effect of the tesseral harmonic terms on the Moon's satellite orbit is different from that on Earth's satellite orbit. Therefore, the perturbation solution of the low Moon orbit satellite due to the non-spherical gravity force has different expressions than these of Earth's satellite. In Table 6.3 we list the main coefficients of spherical harmonic terms from the American LP75G model. In the model $GM = 4,902.800269 \text{ km}^3/\text{s}^2$, $a_e = 1,738.0 \text{ km}$ (the radius of the Moon's equator).

If we use $\bar{J}_l(l \geq 2)$ to replace $\bar{C}_{l,0}$, then the corresponding zonal harmonic coefficients are

Table 6.3 Main coefficients of spherical harmonic terms given by LP75G

l	$\bar{C}_{l,m}$	m	$\bar{S}_{l,m}$
2	$-0.909759705421 \times 10^{-4}$	0	0.0
	$-0.278742486316 \times 10^{-7}$	1	$0.134650039410 \times 10^{-7}$
	$0.346938467558 \times 10^{-4}$	2	$0.387455709879 \times 10^{-8}$
3	$-0.318354210680 \times 10^{-5}$	0	0.0
	$0.264055021943 \times 10^{-4}$	1	$0.543379807817 \times 10^{-5}$
	$0.142548648260 \times 10^{-4}$	2	$0.487321014419 \times 10^{-5}$
	$0.123166082546 \times 10^{-4}$	3	$-0.176222161812 \times 10^{-5}$
4	$0.317891147722 \times 10^{-5}$	0	0.0
	$-0.595795392676 \times 10^{-5}$	1	$0.158041640821 \times 10^{-5}$
	$-0.712685429394 \times 10^{-5}$	2	$-0.670375016705 \times 10^{-5}$
	$-0.142615915284 \times 10^{-5}$	3	$-0.134130695501 \times 10^{-4}$
	$-0.605805350181 \times 10^{-5}$	4	$0.393236088696 \times 10^{-5}$

$$\begin{cases} \bar{J}_1 = -\bar{C}_{1,0} \\ \bar{J}_2 = O(10^{-4}) \\ \bar{J}_3 = -3.20 \times 10^{-5}, \bar{J}_4 = -3.2 \times 10^{-5}, \bar{J}_5 = 2.6 \times 10^{-7}, \bar{J}_6 = -3.9 \times 10^{-6}, \dots \end{cases} \quad (6.2)$$

The parameter which affects a low Moon satellite's lifespan through changing the eccentricity is

$$\left| \frac{\bar{J}_{2l-1}}{\bar{J}_2} \right| = O(10^{-1} - 10^{-2}), \quad l \geq 2 \quad (6.3)$$

This parameter can affect the magnitude of the long-period eccentricity variation of a Moon's satellite up to 0.05–0.10, therefore, can strongly influence the lifespan of the low Moon orbit satellite. The related details are given in Sect. 5.6.

6.1.3 Basic Characteristics of Mars's Gravity Potential

Mars is similar to Earth, the basic parameters are:

$$\begin{aligned} M(\text{mass}) &= 1/3098710 \text{ (solar mass)}, \\ P_s \text{ (period of revolution)} &= 687^{\text{d}}.0 \text{ (Earth day)}, \\ T_r \text{ (period of rotation)} &= 24^{\text{h}}37^{\text{m}}.3777. \end{aligned}$$

In the Goddard model GGM-1041C for the gravity potential of Mars there are

$$GM = 42828.370245291269 \text{ km}^3/\text{s}^2, \quad a_e = 3397.0 \text{ km}.$$

The main spherical harmonic coefficients of Mars's non-spherical gravity potential are listed in Table 6.4.

Mars's gravity field is somehow similar to Earth, such as the magnitude of the dynamical form-factor J_2 term, but there are also obvious differences, such as the equatorial ellipticity and the north–south asymmetry. If we use $\bar{J}_l (l \geq 2)$ to replace $\bar{C}_{l,0}$, then the corresponding zonal harmonic coefficients are:

$$\begin{cases} \bar{J}_1 = -\bar{C}_{1,0} \\ \bar{J}_2 = 10^{-3}, \quad \bar{J}_l (l \geq 3) = O(10^{-5} - 10^{-6}). \end{cases} \quad (6.4)$$

Similar to the Moon, the parameter which affects a low Mars satellite's lifespan through eccentricity is larger than that of Earth, that

Table 6.4 Main coefficients of spherical harmonic terms in GGM-1041C

l	$\overline{C}_{l,m}$	m	$\overline{S}_{l,m}$
2	$-0.874504613 \times 10^{-3}$	0	0.0
	$3.436153047 \times 10^{-10}$	1	$-2.681273014 \times 10^{-10}$
	$-8.458586426 \times 10^{-5}$	2	$4.890547215 \times 10^{-5}$
3	$-1.188948864 \times 10^{-5}$	0	0.0
	$3.905344232 \times 10^{-6}$	1	$2.513932404 \times 10^{-5}$
	$-1.593364126 \times 10^{-5}$	2	$8.354053132 \times 10^{-6}$
	$3.502347743 \times 10^{-5}$	3	$2.555144464 \times 10^{-5}$
4	$5.122708208 \times 10^{-6}$	0	0.0
	$4.211255659 \times 10^{-6}$	1	$3.758812342 \times 10^{-6}$
	$-9.510694650 \times 10^{-7}$	2	$-8.970113183 \times 10^{-6}$
	$6.448970685 \times 10^{-6}$	3	$-1.926969674 \times 10^{-7}$
	$3.095500461 \times 10^{-7}$	4	$-1.285791357 \times 10^{-5}$

$$\left| \frac{\overline{J}_{2l-1}}{J_2} \right| \geq O(10^{-2}), \quad l \geq 2. \quad (6.5)$$

This parameter can make the magnitude of the long-period eccentricity variation of a Mars's satellite up to 0.02–0.03, therefore, affects the lifespan of the low Mars orbit satellite. The related details are given in Sect. 5.6.

6.1.4 Basic Characteristics of Venus's Gravity Potential

Venus is the nearest planet to Earth, its size, mass, and density are close to these of Earth. It also has a dense atmosphere. But it rotates slowly. The basic parameters about Venus are:

$$M(\text{mass}) = 1/408523.5 \text{ (solar mass)},$$

$$P_s \text{ (period of revolution)} = 224^{\text{d}}.7 \text{ (Earth day)},$$

$$T_r \text{ (period of rotation)} = 243^{\text{d}}.0.$$

Because of Venus's slow rotation, the dynamical form-factor $J_2 = O(10^{-6})$ is much smaller than that of the Moon. About Venus's gravity potential, the parameters are from the new American 70×70 order model MGNP180U, that

$$GM = 324858.592079 \text{ km}^3/\text{s}^2, \quad a_e = 6051.0 \text{ km}.$$

The main spherical harmonic coefficients of Venus's non-spherical gravity potential are listed in Table 6.5.

Table 6.5 Main coefficients of spherical harmonic terms in MGNP180U

l	$\overline{C}_{l,m}$	m	$\overline{S}_{l,m}$
2	$-1.969723357760 \times 10^{-6}$	0	0.0
	$2.680268978050 \times 10^{-8}$	1	$1.324780256340 \times 10^{-8}$
	$8.577798458090 \times 10^{-7}$	2	$-9.553616380010 \times 10^{-8}$
3	$7.968246371910 \times 10^{-7}$	0	0.0
	$2.348303842190 \times 10^{-6}$	1	$5.416288391000 \times 10^{-7}$
	$-8.535261871400 \times 10^{-9}$	2	$8.090612890690 \times 10^{-7}$
	$-1.880193062980 \times 10^{-7}$	3	$2.134850148720 \times 10^{-7}$
4	$7.158087500450 \times 10^{-7}$	0	0.0
	$-4.574232817400 \times 10^{-7}$	1	$4.916077249600 \times 10^{-7}$
	$1.263211981230 \times 10^{-7}$	2	$4.835736762340 \times 10^{-7}$
	$-1.746619213280 \times 10^{-7}$	3	$-1.164975842530 \times 10^{-7}$
	$1.725128369010 \times 10^{-7}$	4	$1.376611668820 \times 10^{-6}$

If we use $\overline{J}_l (l \geq 2)$ to replace $\overline{C}_{l,0}$, then the corresponding zonal harmonic coefficients are

$$\begin{cases} \overline{J}_l = -\overline{C}_{l,0} \\ \overline{J}_2 = O(10^{-6}), \quad \overline{J}_l (l \geq 3) = O(10^{-6} - 10^{-7}). \end{cases} \quad (6.6)$$

Because Venus rotates slowly, like the Moon its dynamical form-factor J_2 is not only small but also almost in the same magnitude as the “high-order” harmonic coefficients (including the tesseral harmonic coefficients). Therefore, the effect of the non-spherical gravity force on a low Venus orbit spacecraft is different from that of Earth, then the method to construct the perturbation solution is also different.

In the following sections, we provide the main concepts for constructing analytical perturbation solutions of circling prober around each of the above-mentioned celestial bodies (except Earth) and the formulas of the perturbation solutions based on the characteristics of gravity potential.

6.2 Perturbed Orbital Solution of the Moon’s Satellite

The analytical high precision perturbation solution of the Moon’s satellite orbit is too complicated to be used in practice. But some works, including orbital designs, forecasts, controlling, etc., still need analytical solutions for understanding orbital variations. Therefore, besides numerical models for precise extrapolation, it is necessary to obtain expressions of analytical solutions with certain precision.

In order to explore the Moon, a Moon’s prober needs a low orbit. Similar to a low Earth orbit satellite, the main perturbation affecting the motion of a low Moon orbit

prober is the Moon's non-spherical gravity force. According to the basic information given in the previous section, for the Moon the magnitudes of the harmonic terms are

$$\begin{aligned} C_{2,0} &= O(10^{-4}), \quad C_{2,2}, S_{2,2} = O(10^{-5}), \\ C_{l,m}, S_{l,m} (l \geq 3) &= O(10^{-6} - 10^{-5}). \end{aligned} \quad (6.7)$$

The Moon's rotation is slow, as a result, the order of long-period term's magnitude reduces. In the process of constructing the power series solution of a small parameter, unlike in the case of Earth, the small parameter is not the Moon's dynamical form-factor of $O(10^{-4})$, but is selected as 10^{-2} that

$$\varepsilon = O(10^{-2}). \quad (6.8)$$

Therefore,

$$C_{2,0} = O(\varepsilon^2), \quad C_{2,2}, S_{2,2} = O(\varepsilon^3), \quad C_{l,m}, S_{l,m} (l \geq 3) = O(\varepsilon^3). \quad (6.9)$$

This decision is for the need for a proper structure of the perturbation solution. In practice deciding which of the terms to keep depends on the precision requirement, it is not necessary to stick to the perturbation orders.

It is convenient for analyzing and calculating in constructing the perturbation solution by using normalized units and a dimensionless system. Therefore, we chose the units for length $[L]$, mass $[M]$, and time $[T]$ as

$$\begin{cases} [L] = a_e = 1738.0 \text{ km (equatorial radius of Moon's reference spheroid),} \\ [M] = M \text{ (Moon's mass),} \\ [T] = \left(\frac{a_e^3}{GM} \right)^{\frac{1}{2}} \approx 17^m \cdot 2465. \end{cases} \quad (6.10)$$

The corresponding Moon's gravitational constant $\mu = GM = 1$, and the Moon's equatorial radius $a_e = 1$.

6.2.1 Selection of Coordinate System

Since the main perturbation acting on a low Moon orbit satellite is the non-spherical gravity force and in order to be consistent with the method used for Earth's satellite orbital solution, we select the epoch (J2000.0) Moon-centric mean equatorial frame. The origin of the frame is at the barycenter of the Moon, which should be the same as the chosen Moon's gravitational model; the fundamental plane (xy-plane) is the Moon's equatorial plane. The Moon's physical libration (the oscillation of the

Moon's equator) affects the non-spherical gravity potential in the equatorial coordinate system, so there is also an additional perturbation of the coordinate system. The relationship and the difference between different coordinate systems are described in Sect. 1.4, which should be considered when constructing the perturbation solution.

6.2.2 Mathematical Model for the Perturbed Motion of the Moon's Satellite

When a satellite moves around the Moon its orbit is a varying ellipse. Similar to Earth's satellite this motion corresponds to an ordinary differential equation with initial values as

$$\begin{cases} \ddot{\vec{r}} = \vec{F}_0(\vec{r}) + \vec{F}_\varepsilon(\vec{r}, \dot{\vec{r}}, t; \varepsilon), \\ \vec{r}(t_0) = \vec{r}_0, \quad \dot{\vec{r}}(t_0) = \dot{\vec{r}}_0, \end{cases} \quad (6.11)$$

or to be described by orbital elements as

$$\begin{cases} \dot{\sigma} = f(\sigma, t; \varepsilon), \\ \sigma(t_0) = \sigma_0. \end{cases} \quad (6.12)$$

In Eq. (6.11) $\vec{r}$, $\dot{\vec{r}}$, and $\ddot{\vec{r}}$ are the position vector, velocity, and acceleration in the Moon-centric equatorial coordinate system, respectively. In Eq. (6.12) σ is the usual set of six orbital elements, a , e , i , Ω , ω , and M . The functions on the right side of (6.11), $\vec{F}_0(\vec{r})$ and $\vec{F}_\varepsilon(\vec{r}, \dot{\vec{r}}, t; \varepsilon)$, are the accelerations due to the Moon's central gravity force and the sum of other perturbations acting on the satellite, respectively, and N is the number of perturbation sources, that

$$\vec{F}_0 = -\frac{\mu}{r^3} \vec{r}, \quad (6.13)$$

$$\vec{F}_\varepsilon = \sum_{j=1}^N \vec{F}_j(\vec{r}, \dot{\vec{r}}, t; \varepsilon^j). \quad (6.14)$$

On the right side of (6.12), $f(\sigma, t; \varepsilon)$ is a set of functions of perturbation accelerations $\vec{F}_\varepsilon$.

We use a low Moon orbit satellite as an example to analyze the forces acting on the satellite, then build a mathematical motion model. The main parameters are:

h (mean altitude) = 100 – 200 km, and $e = 0.001$ (almost a circle). In the Moon-centric mean equatorial frame, generally, there are ten sources of known perturbation producing ten accelerations $\vec{F}_\varepsilon(\vec{r}, \dot{\vec{r}}, t; \varepsilon)$, which are:

- (1) The Moon's non-spherical gravity force (J_2 ; $C_{2,2}$, $S_{2,2}$; $C_{l,m}$, $S_{l,m}$, $l \geq 3$), $\vec{F}_1(C_{l,m}, S_{l,m})$

- (2) The Moon's physical libration (σ, ρ, τ) , $\vec{F}_2(\sigma, \rho, \tau)$
- (3) Earth's gravity force (m'_1) , $\vec{F}_3(m'_1)$
- (4) The Sun's gravity force (m'_2) , $\vec{F}_4(m'_2)$
- (5) The Moon's solid tide $(k_2 m'_1)$, $F_5(k_2 m'_1)$
- (6) Solar radiation pressure, $\vec{F}_6(\rho_s)$
- (7) The indirect perturbation of the Moon's oblateness, $\vec{F}_7(J_2 m_e)$ (here m_e is the mass of the Moon)
- (8) Earth's oblateness, $\vec{F}_8(J'_2 m'_1)$
- (9) Gravity forces of other major planes (Venus, Mars, etc.), $\vec{F}_9(m_p)$ (m_p is the mass of a planet)
- (10) Post-Newtonian effect of Moon's gravity force, $\vec{F}_{10}(GMv^2/c^2)$.

For a low Moon orbit satellite ($\bar{h} = 100 - 200$ km), the corresponding relative perturbation magnitudes are

$$\varepsilon_j = \left| \vec{F}_j \right| / \left| \vec{F}_0 \right|, \quad (j = 1, \dots, 10),$$

then the magnitudes of the ten perturbation forces are

$$\begin{aligned}
 \varepsilon_1(J_2) &= O(10^{-4}), \quad \varepsilon_1(C_{2,2}, S_{2,2}) = O(10^{-5}), \\
 \varepsilon_1(C_{l,m}, S_{l,m}, l \geq 3) &= O(10^{-6} - 10^{-5}), \\
 \varepsilon_2 &= O(10^{-7}), \\
 \varepsilon_3 &= O(10^{-5}), \\
 \varepsilon_4 &= O(10^{-7}), \\
 \varepsilon_5 &= O(10^{-7}), \\
 \varepsilon_6 &= O(10^{-9}), \\
 \varepsilon_7 &= O(10^{-11}), \\
 \varepsilon_8 &= O(10^{-12}), \\
 \varepsilon_9 &= O(10^{-12}), \\
 \varepsilon_{10} &= O(10^{-11}),
 \end{aligned} \tag{6.15}$$

The perturbation magnitude of the solar radiation pressure is given under the assumption of the area-to-mass ratio $(S/M) = 10^8$, which is equivalent to 10^9 of an Earth's satellite, i.e., $0.01 \text{ (m}^2/\text{kg)}$.

By the analysis of the perturbation magnitudes, we conclude that:

- (1) For general orbit analysis, it only needs two perturbation forces, the Moon's non-spherical gravity force and Earth's gravity force, to provide the main characteristics of satellite orbital variation.

- (2) For extrapolation of one or two days (low Moon orbit motion arc $S \sim 10^2$) and a required position accuracy better than 1 km, it also only needs the two main perturbation forces as in (1).
- (3) For the high precision orbit determination (the accuracy of distance $\Delta\rho$ is better than 10 m, the accuracy of angle $\Delta\theta$ is better than $0''.02$ of ground-based sampling measurements), to extrapolate 1–2 days requires at least the first five perturbation sources (some of them are marginal, such as the Moon's solid tide), for making the full use of high accuracy measurements.

The above analysis shows that the motion equation can be solved numerically including all perturbation sources without difficulties. But in reality, only the first six sources have practical meaning. For an analytical solution, taking into account the actual background of constructing a perturbation solution and the actual accuracy requirement (extrapolating 1–2 days with an accuracy of 100 m), only the first five sources are needed, and in most cases, the Moon's solid tide perturbation can be ignored.

Considering the first five perturbations in the Moon-centric mean equatorial frame, the perturbation equation of a satellite is given by

$$\ddot{\vec{r}} = \vec{F}_0(\vec{r}) + \sum_{j=1}^5 \vec{F}_j(\vec{r}, t; \varepsilon_j). \quad (6.16)$$

Adopting the dimensionless system (6.10), the perturbing accelerations due to the five forces are

$$\begin{cases} \vec{F}_1 = \text{grad}(\Delta V(C_{l,m}, S_{l,m})), \\ \Delta V(C_{l,m}, S_{l,m}) = \Delta V_{2,0}(J_2) + \Delta V_{2,2}(C_{2,2}) + \Delta V_{l,m}(C_{l,m}, S_{l,m}), \end{cases} \quad (6.17)$$

$$\Delta V_{2,0}(J_2) = -\frac{J_2}{r^3} P_2(\sin \varphi), \quad (6.18)$$

$$\Delta V_{2,2}(C_{2,2}) = \frac{C_{2,2}}{r^3} P_{2,2}(\sin \varphi) \cos 2\lambda, \quad (6.19)$$

$$\Delta V_{l,m} = \sum_{l \geq 3} \sum_{m=0}^l \frac{1}{r^{l+1}} P_{l,m}(\sin \varphi) [C_{l,m} \cos m\lambda + S_{l,m} \sin m\lambda], \quad (6.20)$$

$$\vec{F}_2 = \text{grad} \Delta V_2(J_2 \theta) \quad (6.21)$$

$$\begin{cases} \vec{F}_3 = \text{grad}(R_e(\mu_1)), \\ R_e(\mu_1) = \left(\frac{\mu_1}{r_1^3}\right) r^2 \left[P_2(\cos \psi_1) + \left(\frac{r}{r_1}\right) P_3(\cos \psi_1) \right], \end{cases} \quad (6.22)$$

$$\begin{cases} \vec{F}_4 = \text{grad} R_s(\mu_2), \\ R_s(\mu_2) = \left(\frac{\mu_2}{r_1^3}\right) r^2 P_2(\cos \psi_2), \end{cases} \quad (6.23)$$

$$\begin{cases} \vec{F}_5 = \text{grad}(\Delta V_2(\mu_1 k_2)) \\ \Delta V_2(\mu_1 k_2) = \left(\frac{\mu_1}{r_1}\right) \frac{k_2}{r_1^3} P_2(\cos \psi_1). \end{cases} \quad (6.24)$$

For the requirement of constructing the analytical perturbation solution, the formulas of the perturbation due to the Moon's non-spherical gravity force (6.18)–(6.20) are not in the normalized forms. In the expression of $\vec{F}_2$ (6.21), θ is the parameter for the Moon's physical libration (Sect. 1.4.3), the related expressions are given late in (6.129) and (6.130). The accelerations $\vec{F}_3$ and $\vec{F}_4$ are the gravity perturbations due to Earth and the Sun, and r_1 and r_2 are the distances from the Moon's center to Earth's center and the Sun's center, respectively, and

$$\mu_1 = \frac{GE}{GM}, \quad \mu_2 = \frac{GS}{GM}, \quad (6.25)$$

$$\cos \psi_1 = \left(\frac{\vec{r}}{r}\right) \cdot \left(\frac{\vec{r}_1}{r_1}\right), \quad \cos \psi_2 = \left(\frac{\vec{r}}{r}\right) \cdot \left(\frac{\vec{r}_2}{r_2}\right). \quad (6.26)$$

The acceleration $\vec{F}_5$ is due to the Moon's solid tide, and the parameter $k_2 = 0.029966$, which is the Love number of the second-order.

The above five perturbations include all perturbations with magnitudes not less than 10^{-7} . The gradient functions (6.17)–(6.24), $\Delta V(C_{lm}, S_{lm})$, $\Delta V_2(J_2\theta)$, $R_e(\mu_1)$, $R_s(\mu_2)$, and $\Delta V_2(\mu_1 k_2)$, are perturbation potentials, usually called perturbation functions.

6.2.3 The Numerical Solution for the High Precise Orbital Extrapolation

If we use X to represent the state of an orbit, then the extrapolation problem corresponds to the following ordinary differential equation with initial values:

$$\begin{cases} \dot{X} = F(t, X; \varepsilon_j), \\ X_0 = X(t_0) \end{cases} \quad (6.27)$$

The function F on the right side includes the accelerations of the Moon's gravity force $\vec{F}_0$ and perturbing accelerations $\vec{F}_j$, which are given in the previous section. In the numerical method, the perturbing acceleration of the Moon's physical libration is given by the transformation of coordinate systems (Sect. 1.4.3). Similar to the calculation of orbital extrapolation for Earth's satellite, this part is assimilated into the calculation of related right-side functions and needs not to be separated from the perturbation sources.

If the state X is given by the position vector and velocity of a satellite, $\vec{r}$ and $\dot{\vec{r}}$, respectively, then in the Moon-centric mean equatorial frame, the position vector $\vec{r}$

satisfies the following ordinary differential equation with initial values:

$$\begin{cases} \ddot{\vec{r}} = \vec{F}_0 + \vec{F}_\varepsilon, \\ \vec{r}(t_0) = \vec{r}_0, \quad \dot{\vec{r}}(t_0) = \dot{\vec{r}}_0, \end{cases} \quad (6.28)$$

where $\vec{F}_0$ is the acceleration by the Moon's central gravity force, $\vec{F}_\varepsilon$ is the sum of perturbing accelerations that

$$\begin{cases} \vec{F}_0 = -\frac{GM}{r^3}\vec{r}, \\ \vec{F}_\varepsilon = \vec{F}_1 + \vec{F}_3 + \vec{F}_4 + \vec{F}_5. \end{cases} \quad (6.29)$$

By the chosen unit system (6.10), there is $GM = 1$. The perturbing accelerations, $\vec{F}_j$ ($j = 1, 3, 4, 5$), correspond to the Moon's non-spherical gravity force, Earth's gravity force, the Sun's gravity force, and the Moon's solid tidal force caused by Earth's gravity force, as described above.

Adopting the same method of solving the Earth's satellite orbit problem, the basic variables are orbital elements, σ , and the related ordinal differential equation system and initial values are

$$\begin{cases} \frac{d\sigma}{dt} = f(t, \sigma, \varepsilon), \\ \sigma(t_0) = \sigma_0. \end{cases} \quad (6.30)$$

The orbital elements usually are Kepler orbital elements or the non-singularity orbital elements of the first type for any eccentricity ($0 \leq e < 1.0$).

6.2.4 The Analytical Perturbation Solution of the Moon's Satellite Orbit

In the Moon-centric mean equatorial frame, including the five perturbing sources the perturbation equation system has the form as

$$\begin{cases} \dot{\sigma} = f_0(a) + f_2(\sigma; \varepsilon^2) + f_3(\sigma, t; \varepsilon^3), \\ \varepsilon = O(10^{-2}). \end{cases} \quad (6.31)$$

The five perturbing forces are all gravitational forces. The selection of a small parameter is entirely for the need of the expressions of the perturbation solution as mentioned above. The method and expression forms for a Moon's satellite have no essential difference from those for an Earth's satellite. But when we use the same method for an Earth's satellite to construct the perturbation solution for a Moon's satellite, we should pay attention to the following three aspects.

- (1) The actual magnitude of the small parameter, $\varepsilon = O(10^{-2})$ or $\varepsilon = O(10^{-3})$, is not important. In the process of constructing the perturbation solution, we may

treat the J_2 term of ε^2 in the perturbation equation system (6.31) as a term of the first order, and the $J_{2,2}$ term of ε^3 as a term of the second order, and the rest terms are all as terms of the second order, as long as we know the actual magnitude of each term in the calculation.

- (2) The Moon's rotation is slow, the variations of the variables related to the Moon's rotation, i.e., the tesseral harmonic terms of the Moon's non-spherical gravity potential ($J_{l,m}$, $m \geq 1$), are all slow variables, and should be treated differently from that for an Earth's satellite.
- (3) The characteristics of the oscillation of the Moon's equator are different from Earth, the corresponding additional coordinate perturbation should also be treated differently from that for Earth.

Except for the above three points, the actual procedure of constructing the perturbation solution of the Moon's satellite does not need to repeat.

6.2.4.1 Expressions of the Perturbation Solution in Kepler Orbit Elements a , e , i , Ω , ω , and M [1, 2]

The above analysis of the perturbation sources for a low Moon orbit satellite shows that the magnitude of the J_2 term is 10^{-4} , the magnitudes of the other perturbations are 10^{-6} – 10^{-7} , and the small parameter is chosen as $\varepsilon = 10^{-2}$. In orbit determination or forecast of 1–2 days ($S \sim 10^2$) with an accuracy of 10^{-5} , which is equivalent to a distance accuracy better than 100 m, the perturbation solution should keep periodic terms of the second order and the secular terms of the third order. This kind of formal treatment is described previously. Therefore, the perturbation equation system due to the five perturbing sources can be expressed in a form of a small parameter equation that

$$\begin{aligned} \frac{d\sigma}{dt} &= f(\sigma, t, \varepsilon) \\ &= f_0(a) + f_2(J_2) + f_2(C_{2,2}) + f_2(\mu_1 P_2(\cos \psi_1)) \\ &\quad + f_3(C_{l,0}; l \geq 3) + f_3(C_{l,m}, S_{l,m}; l \geq 3) \\ &\quad + f_4(\mu_2 P_2(\cos \psi_2)) + f_4(\mu_1 k_2) + f_4(\theta J_2), \end{aligned} \quad (6.32)$$

where $f_0(a) = n = a^{-3/2}$ appears only in the equation of dM/dt . The other auxiliaries appear in Sect. 6.2.2. If the required accuracy is reduced by one order (i.e., 10^{-4}), which is equivalent to a distance accuracy of better than 1 km, then all the terms of f_4 can be omitted.

If we adopt the quasi-mean orbit elements for eliminating small divisors, then the reference solution includes both the secular terms and the long-period terms, and the definitions of the quasi-mean elements are

$$\bar{\sigma}(t) = \bar{\sigma}^{(0)}(t) + (\sigma_{1c} + \sigma_{2c} + \cdots)(t - t_0) + \Delta\sigma_l^{(1)}(t) + \Delta\sigma_l^{(2)}(t) + \cdots, \quad (6.33)$$

$$\bar{\sigma}^{(0)}(t) = \bar{\sigma}(t_0) + \delta \bar{n}_0(t - t_0), \quad (6.34)$$

$$\bar{\sigma}(t_0) = \sigma(t_0) - [\sigma_s^{(1)}(t_0) + \sigma_s^{(2)}(t_0) + \dots], \quad (6.35)$$

$$\delta = (0 \ 0 \ 0 \ 0 \ 0 \ 1)^T, \quad (6.36)$$

$$n_0 = a_0^{-3/2}, \quad (6.37)$$

where $\sigma_{1c}, \sigma_{2c}, \dots$ are secular coefficients of the first-order, the second-order, $\dots$ (i.e., the variation rates); $\Delta\sigma_l^{(1)}(t), \Delta\sigma_l^{(2)}(t), \dots$ are long-period terms of the first-order, the second-order, $\dots$, which are

$$\Delta\sigma_l^{(1)} = \sigma_l^{(1)}(t) - \sigma_l^{(1)}(t_0), \quad \Delta\sigma_l^{(2)} = \sigma_l^{(2)}(t) - \sigma_l^{(2)}(t_0), \quad (6.38)$$

and $\sigma_s^{(1)}(t), \sigma_s^{(2)}(t), \dots$ are short-period terms of the first-order, the second-order, $\dots$.

In the perturbation solution of the motion Eq. (6.32) there are

$$\sigma_1 = 0, \quad \sigma_s^{(1)}(t) = 0. \quad (6.39)$$

The analytical solution then has the form as

$$\sigma(t) = \bar{\sigma}(t) + \sigma_s^{(2)}(t). \quad (6.40)$$

As discussed above for orbital extrapolation of 1–2 days, (i.e., an arc $S \sim 10^2$ of a low Moon orbit) and accuracy of the position better than 100 m, we only need the first five perturbations: the Moon's non-spherical gravity force perturbation including $J_2, J_1(J_1 = -C_{l,0}, l \geq 3)$, and $C_{l,m}, S_{l,m}(m \neq 0)$, the additional coordinate perturbation, Earth and the Sun's gravity force perturbations, and the Moon's solid tide perturbation.

About the Sun's gravity force perturbation, strictly speaking, the Moon's satellite orbit analytical solution should be constructed in the heliocentric ecliptic coordinate system, but the Sun is relatively far from the Earth-Moon system, we can treat the Earth's heliocentric ecliptic coordinate system as the Moon's heliocentric ecliptic coordinate system, which introduces an error of about 2.5×10^{-3} (i.e., the parallax of looking at the Sun from Earth and the Moon), and this error is allowed by the required accuracy. Note that for a low Moon orbit satellite, the magnitude of the Sun's perturbation is only 10^{-7} . Similarly in the Moon-centric equatorial frame, when considering Earth's perturbation, the Earth's orbital analytical solution could be given by transforming the Moon's orbital solution in the geocentric frame. But as mentioned above the Moon's orbit around Earth is perturbed by the Sun, and the largest amplitude of periodic terms can reach to 2.0×10^{-2} . Therefore, from the

analytical point of view, the Earth's perturbation term can be only given to 10^{-7} , beyond which the complicity of the variation of the Moon's orbit can cause a huge problem in analyzing the Moon's satellite orbit. This is an important factor that restricts the accuracy of the pure analytical perturbation solution for a Moon's satellite orbit. We can conclude that for constructing a low Moon satellite orbit solution we only include the five perturbations, also for the Sun and Earth's perturbations the orbits of the Sun and Earth can only use the simplified forms. Under these conditions, the corresponding extrapolation accuracy of a Moon's satellite position has the order of 100 m over 1–2 days.

The perturbation solution has the forms as

$$\begin{cases} \sigma(t) = \bar{\sigma}(t_0) + (\delta\bar{n}_0 + \sigma_{2c} + \sigma_{3c} + \sigma_{4c})(t - t_0) + \Delta\sigma_l^{(1)}(t) + \sigma_s^{(2)}(t), \\ \bar{\sigma}(t_0) = \sigma_0 - \sigma_s^{(2)}(t_0). \end{cases} \quad (6.41)$$

The forms of the perturbation terms are given by

$$\begin{cases} \sigma_{2c} = \sigma_{2c}(J_2), \\ \sigma_{3c} = \sigma_{3c}(J_l, l(2) \geq 4) + \sigma_{3c}(\mu_1), \\ \sigma_{4c} = \sigma_{4c}(J_2^2) + \sigma_{4c}(\mu_2), \end{cases} \quad (6.42)$$

$$\Delta\sigma_l^{(1)}(t) = \Delta\sigma_l(t; J_{l \geq 3}; C_{l,m}, S_{l,m}(l \geq 2); \mu_1; \mu_2; \mu_1 k_2; J_2 \theta), \quad (6.43)$$

$$\begin{cases} a_s^{(2)}(t) = a_s^{(2)}(t; J_2; J_{l \geq 3}; C_{l,m}, S_{l,m}(l \geq 2); \mu_1; \mu_2; \mu_1 k_2; J_2 \theta), \\ \sigma_s^{(2)}(t) = \sigma_s^{(2)}(t; J_2; C_{2,2}, S_{2,2}), \text{ for } e, i, \Omega, \omega, M, \end{cases} \quad (6.44)$$

where the term of $\sigma_{4c}(J_2^2)$ is the secular variation due to J_2^2 , which has a magnitude of 10^{-8} so can be omitted, or simplified and included in the calculation if it is needed. The magnitude of the long-period term $\Delta\sigma_l^{(2)}(t, J_2)$ is equivalent to that of $\sigma_{4c}(J_2^2)$, but the long-period variation due to the zonal harmonic terms has only the terms including the argument of perilune ω which always have a factor of the eccentricity e , therefore can be omitted.

When calculating σ_{2c} , σ_{3c} , σ_{4c} , and $\Delta\sigma_l^{(1)}(t)$, $\Delta\sigma_l^{(2)}(t)$, because $\sigma_{1c} = 0$ the value of $\bar{\sigma}(t)$ can be given by $\bar{\sigma}_0 = \bar{\sigma}(t_0)$. Also if the extrapolating arc is not too lg, the long-period terms of $\Delta\sigma_l^{(1)}(t)$, $\Delta\sigma_l^{(2)}(t)$ can be treated as secular terms $\sigma_{2c}(m_e)$, $\sigma_{3c}(J_l; C_{l,m}, S_{l,m})$, $\sigma_{4c}(\mu_2, J_2^2)$. In order to clearly display the properties of orbital variation, we arrange the variations by the secular terms $\sigma_c(t - t_0)$, the long-period terms $\Delta\sigma(t)$, and the short-period term $\sigma_s(t)$ for each orbital elements as

$$\begin{cases} a(t) = \bar{a}_0 + a_s(t), \\ e(t) = \bar{e}_0 + \Delta e_l(t) + e_s(t), \\ i(t) = \bar{i}_0 + \Delta i_l(t) + i_s(t), \\ \Omega(t) = \bar{\Omega}_0 + \Omega_c(t - t_0) + \Delta \Omega_l(t) + \Omega_s(t), \\ \omega(t) = \bar{\omega}_0 + \omega_c(t - t_0) + \Delta \omega_l(t) + \omega_s(t), \\ M(t) = \bar{M}_0 + (\bar{n}_0 + M_c)(t - t_0) + \Delta M_l(t) + M_s(t). \end{cases} \quad (6.45)$$

The perturbation forces for each term are given in (6.42)–(6.44), and $\bar{n}_0 = \bar{a}_0^{-3/2}$, $\bar{a}_0$ is given by

$$\bar{a}_0 = a_0 - a_s^{(2)}(t_0; J_2; J_{l \geq 3}; C_{l,m}, S_{l,m} (l \geq 2); m_e). \quad (6.46)$$

The expressions of perturbation terms are as follows.

(1) σ_c

For all elements, the terms of the first-order zero, i.e. $\sigma_{1c} = 0$. Also, none of a , e , and i has the secular variation of any order. The only secular terms are for Ω , ω , and M that

$$\Omega_{2c} = -\frac{3J_2}{2p^2} n \cos i, \quad (6.47)$$

$$\omega_{2c} = \frac{3J_2}{2p^2} n \left(2 - \frac{5}{2} \sin^2 i \right), \quad (6.48)$$

$$M_{2c} = \frac{3J_2}{2p^2} n \sqrt{1 - e^2} \left(1 - \frac{3}{2} \sin^2 i \right), \quad (6.49)$$

$$\Omega_{3c}(J_l) = -n \cos i \sum_{l(2) \geq 4} \left(\frac{J_l}{p_0^l} \right) F_2(i) K_1(e), \quad (6.50)$$

$$\begin{aligned} \omega_{3c}(J_l) &= -\cos i \Omega_{3c} \\ &- n \sum_{l(2) \geq 4} \left(\frac{J_l}{p_0^l} \right) F_1(i) [(2l - 1) K_1(e) + (1 - e^2) K_2(e)], \end{aligned} \quad (6.51)$$

$$\begin{aligned} M_{3c}(J_l) &= -\sqrt{1 - e^2} (\omega_{3c} + \cos i \Omega_{3c}) \\ &- n \sqrt{1 - e^2} \sum_{l(2) \geq 4} \left(\frac{J_l}{p_0^l} \right) F_1(i) [2(l + 1) K_1(e)], \end{aligned} \quad (6.52)$$

$$\begin{aligned} \Omega_{4c}(J_2^2) &= -\left(\frac{3J_2}{2p^2} \right)^2 n \cos i \left[\left(\frac{3}{2} + \frac{1}{6} e^2 + \sqrt{1 - e^2} \right) \right. \\ &\quad \left. - \left(\frac{5}{3} - \frac{5}{24} e^2 + \frac{3}{2} \sqrt{1 - e^2} \right) \sin^2 i \right], \end{aligned} \quad (6.53)$$

$$\begin{aligned}\omega_{4c}(J_2^2) = & \left(\frac{3J_2}{2p^2}\right)^2 n \left[\left(4 + \frac{7}{12}e^2 + 2\sqrt{1-e^2}\right) \right. \\ & - \left(\frac{103}{12} + \frac{3}{8}e^2 + \frac{11}{2}\sqrt{1-e^2}\right) \sin^2 i \\ & \left. + \left(\frac{215}{48} - \frac{15}{32}e^2 + \frac{15}{4}\sqrt{1-e^2}\right) \sin^4 i \right],\end{aligned}\quad (6.54)$$

$$\begin{aligned}M_{4c}(J_2^2) = & \left(\frac{3J_2}{2p^2}\right)^2 n \sqrt{1-e^2} \left[\frac{1}{2} \left(1 - \frac{3}{2} \sin^2 i\right)^2 \sqrt{1-e^2} + \left(\frac{5}{2} + \frac{10}{3}e^2\right) \right. \\ & - \left(\frac{19}{3} + \frac{26}{3}e^2\right) \sin^2 i + \left(\frac{233}{48} + \frac{103}{12}e^2\right) \sin^4 i \\ & \left. + \frac{e^4}{1-e^2} \left(\frac{35}{12} - \frac{35}{4} \sin^2 i + \frac{315}{32} \sin^4 i\right) \right],\end{aligned}\quad (6.55)$$

$$\Omega_{3c}(\mu_1) = -n \cos i \left(\frac{3}{4}\beta_1 a^3\right) \left(1 - \frac{3}{2} \sin^2 i'\right) \left(1 + \frac{3}{2}e^2\right) (1-e^2)^{-\frac{1}{2}}, \quad (6.56)$$

$$\omega_{3c}(\mu_1) = n \left(\frac{3}{4}\beta_1 a^3\right) \left(1 - \frac{3}{2} \sin^2 i'\right) \left[\left(2 - \frac{5}{2} \sin^2 i\right) + \frac{1}{2}e^2\right] (1-e^2)^{-\frac{1}{2}}, \quad (6.57)$$

$$M_{3c}(\mu_1) = -n \left(\frac{3}{4}\beta_1 a^3\right) \left(1 - \frac{3}{2} \sin^2 i'\right) \left(1 - \frac{3}{2} \sin^2 i\right) \left(\frac{7}{3} + e^2\right), \quad (6.58)$$

$$\begin{aligned}\Omega_{4c}(\mu_2, \beta_1 k_2) = & -n \cos i \left(\frac{3}{4}\beta_2 a^3\right) \left(1 - \frac{3}{2} \sin^2 i'\right) \left(1 + \frac{3}{2}e^2\right) (1-e^2)^{-\frac{1}{2}} \\ & - n \cos i \left(\frac{3}{4p^2}\beta_1 k_2\right) W(\sigma'),\end{aligned}\quad (6.59)$$

$$\begin{aligned}\omega_{4c}(\mu_2, \beta_1 k_2) = & n \left(\frac{3}{4}\beta_2 a^3\right) \left(1 - \frac{3}{2} \sin^2 i'\right) \left[\left(2 - \frac{5}{2} \sin^2 i\right) + \frac{1}{2}e^2\right] (1-e^2)^{-\frac{1}{2}} \\ & + n \left(\frac{3}{4p^2}\beta_1 k_2\right) \left(2 - \frac{5}{2} \sin^2 i\right) W(\sigma'),\end{aligned}\quad (6.60)$$

$$\begin{aligned}M_{4c}(\mu_2, \beta_1 k_2) = & -n \left(\frac{3}{4}\beta_2 a^3\right) \left(1 - \frac{3}{2} \sin^2 i'\right) \left(1 - \frac{3}{2} \sin^2 i\right) \left(\frac{7}{3} + e^2\right) \\ & + n \left(\frac{3}{4p^2}\beta_1 k_2\right) \left(1 - \frac{3}{2} \sin^2 i\right) \sqrt{1-e^2} W(\sigma').\end{aligned}\quad (6.61)$$

In $\sigma_{3c}(J_l)$, the functions of $F_1(i), \dots, K_1(e), \dots$ are

$$\begin{cases} F_1(i) = \sum_{q=0}^{l/2} (-1)^{\frac{(l+2q)}{2}} \left(\frac{1}{2}\right)^{(l+2q)} C_{l,q} (\sin i)^{2q}, \\ F_2(i) = \sum_{q=1}^{l/2} (-1)^{\frac{(l+2q)}{2}} \left(\frac{1}{2}\right)^{(l+2q)} C_{l,q} \cdot 2q (\sin i)^{(2q-2)}, \\ C_{l,q} = \binom{l}{\frac{l}{2}-q} \binom{l+2q}{l} \binom{2q}{q} = \frac{(l+2q)!}{(q!)^2 (\frac{l}{2}-q)! (\frac{l}{2}+q)!}, \end{cases} \quad (6.62)$$

$$\begin{cases} K_1(e) = \sum_{\alpha(2)=0}^{l-2} C_{l,\alpha} \left(\frac{1}{2}\right)^\alpha e^\alpha, \\ K_2(e) = \sum_{\alpha(2)=2}^{l-2} C_{l,\alpha} \alpha \left(\frac{1}{2}\right)^\alpha e^{\alpha-2}, \\ C_{l,\alpha} = \binom{l-1}{\alpha} \binom{\alpha}{\frac{\alpha}{2}} = \frac{(l-1)!}{(l-1-\alpha)! (\frac{\alpha}{2})!}, \end{cases} \quad (6.63)$$

where the sub $l(2) \geq 4$ means $l = 4, 6, \dots$, and $\alpha(2) \geq 2$ means $\alpha = 2, 4, \dots$.

The definitions of $\beta_1, \beta_2, W(\sigma')$, which are related to the Sun and Earth, are

$$\begin{cases} \beta_1(\text{ or } \beta_2) = \frac{Gm'}{GMr^3} = \left(\frac{m'}{M}\right)/a'^3 = m'/a'^3, \\ W(\sigma') = \left(1 - \frac{3}{2} \sin^2 i'\right) + \frac{3}{2} \sin^2 i' \cos 2u', \end{cases} \quad (6.64)$$

where the value of r' takes the value of $\vec{r}'$, that $\vec{r}' = a'[1 + O(e'^2)]$. In (6.56)–(6.64) the letters with an apostrophe are for the third body (Earth or the Sun); the orbit elements of a third body, $a', e', i', \dots$ are in the Moon-centric mean equatorial frame; GM in (6.64) is the Moon's gravitational constant, M is the Moon's mass, which is the unit of mass in constructing the perturbation solution of the Moon's satellite, do not be confused with the mean anomaly, and we do not explain it again.

The orbit elements on the right sides of the above formulas, a, e, i, n , and $p_0 = a(1 - e^2)$, are all quasi-mean orbit elements, $\bar{a}_0, \bar{e}_0, \bar{i}_0, \bar{n}_0 = \bar{a}_0^{-3/2}$ and $\bar{p}_0 = \bar{a}_0(1 - \bar{e}_0^2)$.

It should be explained that the reason to separate the secular terms due to the perturbations of the Sun in (6.59)–(6.61) and Earth in (6.56)–(6.58) is for analysis, whereas in reality when we calculate extrapolations the actual expressions used are the simple long-period terms given in (6.68)–(6.70).

$$(2) \quad \Delta\sigma_l(t) = \sigma_l(t) - \sigma_l(t_0)$$

For the four perturbation the Moon's non-spherical gravity force, the gravity forces of Earth and the Sun, and the Moon's solid tide, the long-period variations are given by the following expressions (6.65)–(6.70), the part of additional coordinate perturbation ($J_2\theta$) is discussed in the following Sect. 6.2.5.

$$\Delta a_l^{(1)}(t) = 0, \quad (6.65)$$

$$\Delta e_l^{(1)}(t) = (1 - e^2) \sum_{l \geq 3} \left(\frac{J_l}{p_0}\right)^{l-2+\delta/2} \sum_{p=1}^{l-2} (l-2p) F_3(i) \left[\frac{1}{e} K_3(e)\right] I(\omega)n(t - t_0)$$

$$\begin{aligned}
& - (1 - e^2) \sum_{l \geq 2} \sum_{m=1}^l \left(\frac{1}{p_0^l} \right) \sum_{n=1}^{l-1} (l - 2p) F_{lmp}(i) \left(\frac{1}{e} K_3(e) \right) \Phi_{lmp} n(t - t_0) \\
& - \left(\frac{3}{2} \beta a^3 \right) \left[5e \sqrt{1 - e^2} S_3 \right] n(t - t_0), \tag{6.66}
\end{aligned}$$

$$\begin{aligned}
\Delta i_l^{(1)}(t) = & -\cos i \sum_{l \geq 3} \left(\frac{J_l}{p_0^l} \right) \sum_{p=1}^{(l-2+\delta)/2} (l - 2p) \left[\frac{F_3(i)}{\sin i} \right] K_3(e) I(\omega) n(t - t_0) \\
& + \sum_{l \geq 2} \sum_{m=1}^l \left(\frac{1}{p_0^l} \right) \sum_{p=1}^{l-1} [(l - 2p) \cos i - m] \left[\frac{F_{lmp}(i)}{\sin i} \right] K_3(e) \Phi_{lmp} n(t - t_0) \\
& + \left(\frac{3}{2} \beta a^3 \right) \frac{1}{\sqrt{1 - e^2} \sin i} \left[5e^2 \cos i S_3 - A A_2 (1 + 4e^2) - B B_2 (1 - e^2) \right] n(t - t_0), \tag{6.67}
\end{aligned}$$

$$\begin{aligned}
\Delta \Omega_l^{(1)}(t) = & -\cos i \sum_{l \geq 3} \left(\frac{J_l}{p_0^l} \right) \sum_{p=1}^{(l-2+\delta)/2} \left[\frac{F_4(i)}{\sin^2 i} \right] K_3(e) H(\omega) n(t - t_0) \\
& + \sum_{l \geq 2} \sum_{m=1}^l \left(\frac{1}{p_0^l} \right) \sum_{p=1}^{l-1} \left[\frac{F'_{lmp}(i)}{\sin i} \right] K_3(e) \Psi_{lmp} n(t - t_0) \\
& + \left(\frac{3}{2} \beta a^3 \right) \frac{1}{\sqrt{1 - e^2} \sin i} [A A_1 (1 + 4e^2) + B B_1 (1 - e^2)] n(t - t_0), \tag{6.68}
\end{aligned}$$

$$\begin{aligned}
\Delta \omega_l^{(1)}(t) = & -\cos i \Delta \Omega_l^{(1)}(t) \\
& - \sum_{l \geq 3} \left(\frac{J_l}{p_0^l} \right) \sum_{p=1}^{(l-2+\delta)/2} F_3(i) \left[(2l - 1) K_3(e) + (1 - e^2) K_4(e) \right] H(\omega) n(t - t_0) \\
& + \sum_{l \geq 2} \sum_{m=1}^l \left(\frac{1}{p_0^l} \right) \sum_{p=1}^{l-1} F_{lmp}(i) \left[(2l - 1) K_3(e) + (1 - e^2) K_4(e) \right] \Psi_{lmp} n(t - t_0) \\
& + \left(\frac{3}{2} \beta a^3 \right) \sqrt{1 - e^2} (3S_1 + 5S_2) n(t - t_0), \tag{6.69}
\end{aligned}$$

$$\begin{aligned}
\Delta M_l^{(1)}(t) = & -\sqrt{1 - e^2} \left[\Delta \omega_l^{(1)}(t) + \cos i \Delta \Omega_l^{(1)}(t) \right] \\
& - \sqrt{1 - e^2} \sum_{l \geq 3} \left(\frac{J_l}{p_0^l} \right) \sum_{p=1}^{(l-2+\delta)/2} 2(l + 1) F_3(i) K_3(e) H(\omega) n(t - t_0) \\
& + \sqrt{1 - e^2} \sum_{l \geq 2} \sum_{m=1}^l \left(\frac{1}{p_0^l} \right) \sum_{n=1}^{l-1} 2(l + 1) F_{lmp}(i) K_3(e) \Psi_{lmp} n(t - t_0) \\
& - \left(\frac{3}{2} \beta a^3 \right) \left[4 \left(1 + \frac{3}{2} e^2 \right) S_1 + 10e^2 S_2 \right] n(t - t_0). \tag{6.70}
\end{aligned}$$

The long-period variations due to the tidal deformation caused by Earth are given in (6.92)–(6.98).

In the above expressions, p_0 means $a(1 - e^2)$, we replace p by p_0 because the letter p is used in the sums. Also, the orbit elements, a , e , i , Ω , and ω are $\bar{a}_0$, $\bar{e}_0$, $\bar{i}_0$, $\bar{\Omega}_0$, $\bar{\omega}_0$, respectively. The same usages of these symbols are repeated, and we do not explain it again.

The auxiliaries in the solution due to the Moon's non-spherical gravity perturbation, $F_3(i)$, $F_4(i)$, $\dots$, are given by

$$\left\{ \begin{array}{l} F_3(i) = \sum_{q=0}^p (-1)^{(l+2q-\delta_1)/2} \left(\frac{1}{2}\right)^{(2l-2p+2q-\delta_2)} C_{lpq} (\sin i)^{(l-2p+2q)}, \\ F_4(i) = \sum_{q=0}^p (l-2p+2q) (-1)^{(l+2q-\delta_1)/2} \left(\frac{1}{2}\right)^{(2l-2p+2q-\delta_2)} C_{lpq} (\sin i)^{(l-2p+2q)}, \\ C_{lpq} = \frac{\binom{l}{p-q} \binom{2l-2p+2q}{l} \binom{l-2p+2q}{q}}{(2l-2p+2q)!} \\ = \frac{1}{q!(p-q)!(l-p+q)!(l-2p+q)!}, \end{array} \right. \quad (6.71)$$

$$\delta_1 = \frac{1}{2}[1 - (-1)^l], \quad \delta_2 = \begin{cases} 0, & l-2p = 0, \\ 1, & l-2p \neq 0, \end{cases} \quad (6.72)$$

$$\left\{ \begin{array}{l} K_3(e) = \sum_{\alpha(2)=|l-2p|}^{l-2} C_{lp\alpha} \left(\frac{1}{2}\right)^\alpha e^\alpha, \\ K_4(e) = \sum_{\alpha(2)=|l-2p|}^{l-2} C_{lp\alpha} \alpha \left(\frac{1}{2}\right)^\alpha e^{\alpha-2}, \\ C_{lp\alpha} = \frac{\binom{l-1}{\alpha} \binom{\alpha}{\frac{1}{2}(\alpha - |l-2p|)}}{(l-1)!} \\ = \frac{1}{(l-1-\alpha)! \left[\frac{1}{2}(\alpha - |l-2p|)\right]! \left[\frac{1}{2}(\alpha + |l-2p|)\right]!}, \end{array} \right. \quad (6.73)$$

$$\left\{ \begin{array}{l} I(\omega) = -(1 - \delta_1) \sin(l-2p)\omega + \delta_1 \cos(l-2p)\omega, \\ H(\omega) = (1 - \delta_1) \cos(l-2p)\omega + \delta_1 \sin(l-2p)\omega. \end{array} \right. \quad (6.74)$$

The symbols δ_1 and δ_2 in (6.72) are defined for separating the odd terms from the even terms in the non-spherical gravity potential, do not mix them with the symbol δ used in (6.36) which is for distinguishing the mean anomaly from the other five orbit elements.

$$\begin{aligned}
F_{lmp}(i) &= \frac{(l+m)!}{2^l p! (l-p)!} \sum_{k=k_1}^{k_2} (-1)^{k+\frac{(l-m+\delta_{lm})}{2}} \binom{2l-2p}{k} \binom{2p}{l-m-k} \\
&\quad \times \left(\sin \frac{i}{2} \right)^{-(l-m-2p-2k)} \left(\cos \frac{i}{2} \right)^{(3l-m-2p-2k)} \\
&= \frac{(l+m)!}{2^l p! (l-p)!} \sum_{k=k_1}^{k_2} (-1)^{k+\frac{(l-m+\delta_{lm})}{2}} \binom{2l-2p}{k} \binom{2p}{l-m-k} \\
&\quad \times (\sin i)^{(l-m-2p-2k)} (1 + \cos i)^{(3l-m-2p-2k)} \tag{6.75}
\end{aligned}$$

$$\begin{aligned}
F'_{lmp}(i) &= \frac{d}{dt} F_{lmp}(i) \\
&= \frac{(l+m)!}{2^l p! (l-p)!} \left(\frac{1}{\sin i} \right) \sum_{k=k_1}^{k_2} (-1)^{k+(l-m+\delta_{lm})/2} \binom{2l-2p}{k} \binom{2p}{l-m-k} \\
&\quad \times \left[-2l \sin^2 \frac{i}{2} - (l-m-2p-2k) \right] \left(\sin \frac{i}{2} \right)^{-(l-m-2p-2k)} \times \left(\cos \frac{i}{2} \right)^{(3l-m-2p-2k)}, \tag{6.76}
\end{aligned}$$

$$k_1 = \max(0, l-m-2p), \quad k_2 = \min(l-m, 2l-2p), \tag{6.77}$$

$$\left\{ \begin{aligned}
\Phi_{lmp} &= -[(1-\delta_m)C_{l,m} - \delta_m S_{l,m}] \sin((l-2p)\omega + m\Omega_G) \\
&\quad + [(1-\delta_m)S_{l,m} + \delta_m C_{l,m}] \cos((l-2p)\omega + m\Omega_G) \\
\Psi_{lmp} &= [(1-\delta_m)C_{l,m} - \delta_m S_{l,m}] \cos((l-2p)\omega + m\Omega_G) \\
&\quad + [(1-\delta_m)S_{l,m} + \delta_m C_{l,m}] \sin((l-2p)\omega + m\Omega_G) \\
\Omega_G &= \Omega - S_G, \quad S_G = [(f_m + \omega_m + \Omega_m)_0 + 180^\circ] + n_r(t - t_0) \\
\delta_{l,m} &= \frac{1}{2} [1 - (-1)^{l-m}]
\end{aligned} \right. \tag{6.78}$$

In (6.78) S_G is the “sidereal time” on the Moon, i.e., the hour angle of the direction of the X axis in the Moon-fixed frame; f_m , ω_m , and Ω_m are the Moon's orbital elements in the geocentric frame, which are introduced previously.

The definitions of β , Gm' , $\dots$ in the terms of the gravity perturbations of Earth and the Sun, are

$$\beta = \frac{Gm'}{r'^3} = \frac{Gm'}{\bar{r}^3}, \quad Gm' = m_e \text{ or } m_s, \tag{6.79}$$

where m_e and m_s (dimensionless) are the masses of Earth and the Sun, respectively. The formulas (6.66)–(6.70) are for both Earth and the Sun depending on which mass is used, Earth or the Sun. These formulas include the secular terms $\sigma_{3c}(\mu_1)$ and $\sigma_{4c}(\mu_2)$. As explained before, the separation of these secular terms is for analysis, whereas in calculations the terms $\sigma_{3c}(\mu_1)$ and $\sigma_{4c}(\mu_2)$ should not be calculated repeatedly. The definitions of $S_1, S_2, S_3, A, A_1, A_2, B, B_1$, and B_2 in (6.66)–(6.70) are

$$\begin{cases} S_1 = -\frac{1}{3} + \frac{1}{2}(A^2 + B^2), \\ S_2 = \frac{1}{2}(A^2 - B^2), \\ S_3 = AB, \end{cases} \quad (6.80)$$

$$\begin{aligned} A = \frac{1}{4} \{ & (1 - \cos i)[(1 + \cos i') \cos(\omega - \theta + u') + (1 - \cos i') \cos(\omega - \theta - u')] \\ & + (1 + \cos i)[(1 + \cos i') \cos(\omega + \theta - u') + (1 - \cos i') \cos(\omega + \theta + u')] \\ & + 2 \sin i \sin i' [\cos(\omega - u') - \cos(\omega + u')] \}, \end{aligned} \quad (6.81)$$

$$\begin{aligned} B = -\frac{1}{4} \{ & (1 - \cos i)[(1 + \cos i') \sin(\omega - \theta + u') + (1 - \cos i') \sin(\omega - \theta - u')] \\ & + (1 + \cos i)[(1 + \cos i') \sin(\omega + \theta - u') + (1 - \cos i') \sin(\omega + \theta + u')] \\ & + 2 \sin i \sin i' [\sin(\omega - u') - \sin(\omega + u')] \}, \end{aligned} \quad (6.82)$$

$$\begin{aligned} A_1 = \frac{\partial A}{\partial i} = \frac{1}{4} \{ & \sin i [(1 + \cos i') \cos(\omega - \theta + u') + (1 - \cos i') \cos(\omega - \theta - u')] \\ & - \sin i [(1 + \cos i') \cos(\omega + \theta - u') + (1 - \cos i') \cos(\omega + \theta + u')] \\ & + 2 \cos i \sin i' [\cos(\omega - u') - \cos(\omega + u')] \}, \end{aligned} \quad (6.83)$$

$$\begin{aligned} B_1 = \frac{\partial B}{\partial i} = -\frac{1}{4} \{ & \sin i [(1 + \cos i') \sin(\omega - \theta + u') + (1 - \cos i') \sin(\omega - \theta - u')] \\ & - \sin i [(1 + \cos i') \sin(\omega + \theta - u') + (1 - \cos i') \sin(\omega + \theta + u')] \\ & + 2 \cos i \sin i' [\sin(\omega - u') - \sin(\omega + u')] \}, \end{aligned} \quad (6.84)$$

$$\begin{aligned} A_2 = \frac{\partial A}{\partial \Omega} = \frac{1}{4} \{ & (1 - \cos i)[(1 + \cos i') \sin(\omega - \theta + u') + (1 - \cos i') \sin(\omega - \theta - u')] \\ & - (1 + \cos i)[(1 + \cos i') \sin(\omega + \theta - u') + (1 - \cos i') \sin(\omega + \theta + u')] \}, \end{aligned} \quad (6.85)$$

$$\begin{aligned} B_2 = \frac{\partial B}{\partial \Omega} = \frac{1}{4} \{ & (1 - \cos i)[(1 + \cos i') \cos(\omega - \theta + u') + (1 - \cos i') \cos(\omega - \theta - u')] \\ & - (1 + \cos i)[(1 + \cos i') \cos(\omega + \theta - u') + (1 - \cos i') \cos(\omega + \theta + u')] \}, \end{aligned} \quad (6.86)$$

where $\theta = \Omega - \Omega'$. Do not confuse this θ with the Moon's physical libration parameter θ .

The position vectors of Earth and the Sun in the Moon-centric mean equatorial frame (i.e., orbit elements σ') as analyzed at the beginning of this section, can only be treated in simplified forms as follows.

For the Earth's gravity perturbation there are

$$\begin{cases} i' = J + I, & I = 1^\circ 32' 32''.7, \\ \Omega' = \Omega_m, \\ u' = f_m + (\omega_m + 180^\circ), \end{cases} \quad (6.87)$$

where J, f_m, ω_m , and Ω_m are the mean orbital elements of the Moon. The mean orbital elements of the Moon in the J2000.0 geocentric ecliptic frame $\bar{\sigma}'$ are approximately

$$\begin{cases} \bar{a} = 0.0025718814 \text{ AU} = 384747.981 \text{ km}, \\ \bar{e} = 0.054879905, \\ \bar{i} = J = 5^\circ.129835071, \\ \bar{\Omega} = 125^\circ.044556 - 1934^\circ.1361850T, \\ \bar{\omega} = 318^\circ.308686 + 6003^\circ.1498961T, \\ \bar{M} = 134^\circ.963414 + 13^\circ.06499315537d. \end{cases} \quad (6.88)$$

For the Sun's gravity perturbation there are

$$\begin{cases} i' = I, \quad I = 1^\circ 32' 32'' .7, \\ \Omega' = \Omega_m, \\ u' = f_s + \omega_s, \end{cases} \quad (6.89)$$

where f_s and ω_s are the Sun's mean orbital elements in the geocentric frame, the Sun's mean orbital elements, $\bar{\sigma}'$, in the J2000.0 geocentric celestial frame, are

$$\begin{cases} \bar{a} = 1.00000102 \text{ AU}, \quad 1 \text{ AU} = 1.49597870691 \times 10^8 \text{ km}, \\ \bar{e} = 0.016709, \quad \bar{i} = \varepsilon = 23^\circ.4393, \\ \bar{\Omega} = 0^\circ.0, \\ \bar{\omega} = 282^\circ.9373 + 0^\circ.32T, \\ \bar{M} = 357^\circ.5291 + 0^\circ.9856d. \end{cases} \quad (6.90)$$

where d is the Julian day number, T is the Julian century number, which are defined in Sect. 1.1.3. For given mean orbital elements $\bar{\sigma}'$, f_m , and f_s can be calculated using the following simplified format that

$$f = M + 2e \sin M + O(e^2) \quad (6.91)$$

The long-period variations given in (6.66)–(6.70) are simplified. If the high order terms (l, m) in the Moon's non-spherical gravity potential are included, it may cause short-period terms, or if the extrapolating arc is long, treating the long-period terms as secular terms may reduce the accuracy. To solve these problems, we may keep the long-period terms as what they are, and calculate $\sigma_l(t)$ and $\sigma_l(t_0)$ separately. It can be done by the following method. For e and i , in (6.66) and (6.67) replace $I(\omega)n(t - t_0)$ by $H(\omega(t))\alpha_c$, and replace $\Phi_{lmp}n(t - t_0)$ by $\Psi_{lmp}(t)\alpha$; for Ω, ω , and M in (6.68)–(6.70) replace $H(\omega)n(t - t_0)$ by $-I(\omega(t))\alpha_c$ and replace $\Psi_{lmp}n(t - t_0)$ by $-\Phi_{lmp}(t)\alpha$, where $\alpha_c = n/(l - 2p)\omega_{2c}$, $\alpha = n/n_G$, and $n_G = (l - 2p)\omega_{2c} + m(\Omega_{2c} - n_r)$. The definitions of $I(\omega)$ and $H(\omega)$ are given in (6.74), and the definitions of Φ_{lmp} and Ψ_{lmp} are in (6.78).

In the long-period terms, the parts by the Moon's solid tidal perturbation due to Earth's gravity force are as follows.

$$a_l^{(1)}(t) = 0, \quad (6.92)$$

$$e_l^{(1)}(t) = 0, \quad (6.93)$$

$$\begin{aligned} i_l^{(1)}(t) = & -\frac{3}{8p^2}\beta_1k_2 \\ & \left\{ \sin i' \cos i \left[2 \cos i' \left(\frac{n}{\Omega_{2c}} \right) \cos \Omega^* + (1 - \cos i') \left(\frac{n}{\alpha_1} \right) \cos(2u' + \Omega^*) \right. \right. \\ & \quad \left. \left. + (1 + \cos i') \left(\frac{n}{\alpha_2} \right) \cos(2u' - \Omega^*) \right] \right. \\ & \left. + \frac{\sin i}{4} \left[2 \sin^2 i' \left(\frac{n}{\Omega_{2c}} \right) \cos \Omega^* + (1 - \cos i')^2 \left(\frac{n}{\alpha_5} \right) \cos(2u' + 2\Omega^*) \right. \right. \\ & \quad \left. \left. - (1 + \cos i')^2 \left(\frac{n}{\alpha_6} \right) \cos(2u' - 2\Omega^*) \right] \right\}, \end{aligned} \quad (6.94)$$

$$\begin{aligned} \Omega_l^{(2)}(t) = & -\frac{3}{4p^2}\beta_1k_2 \left(\frac{3}{2} \sin^2 i' \right) \cos i \left(\frac{n}{2n'} \right) \sin 2u' \\ & + \frac{3}{8p^2}\beta_1k_2 \left\{ -\sin i' \frac{\cos 2i}{\sin i} \left[2 \cos i' \left(\frac{n}{\Omega_1} \right) \sin \Omega^* \right. \right. \\ & \quad \left. \left. + (1 - \cos i') \left(\frac{n}{\alpha_1} \right) \sin(2u' + \Omega^*) - (1 + \cos i') \left(\frac{n}{\alpha_2} \right) \sin(2u' - \Omega^*) \right] \right. \\ & \quad \left. + \frac{\cos i}{4} \left[2 \sin^2 i' \left(\frac{n}{\Omega_1} \right) \sin 2\Omega^* + (1 - \cos i')^2 \left(\frac{n}{\alpha_5} \right) \sin(2u' + 2\Omega^*) \right. \right. \\ & \quad \left. \left. + (1 + \cos i')^2 \left(\frac{n}{\alpha_6} \right) \sin(2u' - 2\Omega^*) \right] \right\} \\ & + \frac{3}{8p^2}\beta_1k_2 \left\{ \sin i' \cos i' \left[2 \cos i' \left(\frac{n}{\Omega_1} \right) \sin \Omega^* \right. \right. \\ & \quad \left. \left. + (1 - \cos i') \left(\frac{n\Omega_1}{\alpha_1^2} \right) \sin(2u' + \Omega^*) + (1 + \cos i') \left(\frac{n\Omega_1}{\alpha_2^2} \right) \sin(2u' - \Omega^*) \right] \right. \\ & \quad \left. + \frac{1}{8} \left(\frac{\sin^2 i}{\cos i} \right) \left[2 \sin^2 i' \left(\frac{n}{\Omega_1} \right) \sin 2\Omega^* + (1 - \cos i')^2 \left(\frac{n\Omega_1}{\alpha_5^2} \right) \sin(2u' + 2\Omega^*) \right. \right. \\ & \quad \left. \left. - (1 + \cos i')^2 \left(\frac{n\Omega_1}{\alpha_6^2} \right) \sin(2u' - 2\Omega^*) \right] \right\}, \end{aligned} \quad (6.95)$$

$$\begin{aligned} \omega_l^{(2)}(t) = & \frac{3}{4p^2}\beta_1k_2 \left(\frac{3}{2} \sin^2 i' \right) \left(2 - \frac{5}{2} \sin^2 i \right) \left(\frac{n}{2n'} \right) \sin 2u' \\ & - \frac{3}{8p^2}\beta_1k_2 \left\{ \sin i' \cot i (1 - 5 \sin^2 i) \left[2 \cos i' \left(\frac{n}{\Omega_1} \right) \sin \Omega^* \right. \right. \\ & \quad \left. \left. + (1 - \cos i') \left(\frac{n}{\alpha_1} \right) \sin(2u' + \Omega^*) - (1 + \cos i') \left(\frac{n}{\alpha_2} \right) \sin(2u' - \Omega^*) \right] \right\} \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{4} \left(1 - \frac{5}{2} \sin^2 i \right) \left[2 \sin^2 i' \left(\frac{n}{\Omega_1} \right) \sin 2\Omega^* \right. \\
& \quad \left. + (1 - \cos i')^2 \left(\frac{n}{\alpha_5} \right) \sin(2u' + 2\Omega^*) + (1 + \cos i')^2 \left(\frac{n}{\alpha_6} \right) \sin(2u' - 2\Omega^*) \right] \Big\} \\
& - \frac{3}{16p^2} \beta_1 k_2 \left\{ 5 \sin i' \sin 2i \left[2 \cos i' \left(\frac{n}{\Omega_1} \right) \sin \Omega^* \right. \right. \\
& \quad \left. + (1 - \cos i') \left(\frac{n\Omega_1}{\alpha_1^2} \right) \sin(2u' + \Omega^*) + (1 + \cos i') \left(\frac{n\Omega_1}{\alpha_2^2} \right) \sin(2u' - \Omega^*) \right] \\
& \quad + \frac{5}{4} \sin^2 i \left[2 \sin^2 i' \left(\frac{n}{\Omega_1} \right) \sin 2\Omega^* \right. \\
& \quad \left. \left. + (1 - \cos i')^2 \left(\frac{n\Omega_1}{\alpha_5^2} \right) \sin(2u' + 2\Omega^*) - (1 + \cos i')^2 \left(\frac{n\Omega_1}{\alpha_6^2} \right) \sin(2u' - 2\Omega^*) \right] \right\}, \tag{6.96}
\end{aligned}$$

$$\begin{aligned}
M_l^{(2)}(t) = & \frac{3}{4p^2} \beta_1 k_2 \left(\frac{3}{2} \sin^2 i' \right) \left(1 - \frac{3}{2} \sin^2 i \right) (1 - e^2) \left(\frac{n}{2n'} \right) \sin 2u' \\
& + \frac{9}{16p^2} \beta_1 k_2 \sqrt{1 - e^2} \left\{ \sin i' \sin 2i (1 - 5 \sin^2 i) \left[2 \cos i' \left(\frac{n}{\Omega_1} \right) \sin \Omega^* \right. \right. \\
& \quad \left. + (1 - \cos i') \left(\frac{n}{\alpha_1} \right) \sin(2u' + \Omega^*) - (1 + \cos i')^2 \left(\frac{n}{\alpha_2} \right) \sin(2u' - \Omega^*) \right] \\
& \quad + \frac{1}{4} \sqrt{1 - e^2} \sin^2 i \left[2 \sin^2 i' \left(\frac{n}{\Omega_1} \right) \sin 2\Omega^* \right. \\
& \quad \left. + (1 - \cos i')^2 \left(\frac{n}{\alpha_5} \right) \sin(2u' + 2\Omega^*) + (1 + \cos i')^2 \left(\frac{n}{\alpha_6} \right) \sin(2u' - 2\Omega^*) \right] \Big\} \\
& - \frac{9}{16p^2} \beta_1 k_2 \sqrt{1 - e^2} \left\{ \sin i' \sin 2i \left[2 \cos i' \left(\frac{n}{\Omega_1} \right) \sin \Omega^* \right. \right. \\
& \quad \left. + (1 - \cos i') \left(\frac{n\Omega_1}{\alpha_1^2} \right) \sin(2u' + \Omega^*) + (1 + \cos i') \left(\frac{n}{\alpha_2} \right) \sin(2u' - \Omega^*) \right] \\
& \quad + \frac{1}{4} \sqrt{1 - e^2} \sin^2 i \left[2 \sin^2 i' \left(\frac{n}{\Omega_1} \right) \sin 2\Omega^* \right. \\
& \quad \left. \left. + (1 - \cos i')^2 \left(\frac{n\Omega_1}{\alpha_5^2} \right) \sin(2u' + 2\Omega^*) - (1 + \cos i')^2 \left(\frac{n\Omega_1}{\alpha_6^2} \right) \sin(2u' - 2\Omega^*) \right] \right\}. \tag{6.97}
\end{aligned}$$

Here $\Omega^* = \Omega - (\Omega' + \nu)$, and ν is the lagging angle of the Moon's solid tide ($\nu = 30^\circ$), in Sect. 4.8.1, we introduce similar parameters for Earth's solid tide in (4.636); and $\alpha_1, \alpha_2, \alpha_5$, and α_6 , are defined as

$$\begin{cases} \alpha_1 = 2n' + \Omega_{2c}, & \alpha_2 = 2n' - \Omega_{2c}, \\ \alpha_5 = n' + \Omega_{2c}, & \alpha_6 = n' - \Omega_{2c}, \end{cases} \tag{6.98}$$

where Ω_{2c} is the secular variation rate of the second-order of Ω due to the Moon's non-spherical perturbation term (J_2) given in (6.47). The values of Earth's orbital elements i' , Ω' , u' , and mean angular speed n' are described before, so are not repeated.

(3) $\sigma_s^{(2)}(t)$

According to the analysis of the perturbation magnitude, for all $\sigma_s^{(2)}(t)$ except $a_s^{(2)}(t)$, the only perturbation should be included is the Moon's non-spherical gravity force. For $a_s^{(2)}(t)$ when calculating $M(t)$ we need $\bar{n}_0 = \bar{a}_0^{-3/2}$ and $\bar{a}_0 = a_0 - [a_s^{(2)}(t_0) + a_s^{(3)}(t_0)]$, therefore we need to include perturbations of the Moon's non-spherical gravity force, the third-body (Earth and the Sun) gravity forces, the Moon's solid tidal force, and the additional coordinate perturbation, the sum of these perturbations is denoted to $a_s^{(2)}(t)$, that

$$\begin{cases} a_s^{(2)}(t) = a_s^{(2)}(t; J_2; J_{l \geq 3}; C_{l,m}, S_{l,m} (l \geq 2); \mu_1; \mu_2; \mu_1 k_2; J_2 \theta), \\ \sigma_s^{(2)}(t) = \sigma_s^{(2)}(t; J_2; J_{l \geq 3}; C_{2,2}, S_{2,2}) \text{ for } e, i, \Omega, \omega, M. \end{cases} \quad (6.99)$$

$$a_s^{(2)}(t) = \int^t (f_{2s})_a dt = \int^t \frac{2}{na} \frac{\partial R_{2s}}{\partial M} \frac{dM}{n} = \frac{2}{n^2 a} R_{2s} = 2a^2 R_{2s}, \quad (6.100)$$

$$\begin{aligned} R_{2s} &= R_s(J_2) + R_s(J_l)_{l \geq 3} + R_s(C_{l,m}, S_{l,m})_{l \geq 2} \\ &\quad + R_s(\mu_1, \mu_2) + R_s(\mu_1 k_2) + R_s(J_2 \theta). \end{aligned} \quad (6.101)$$

The additional coordinate perturbation ($J_2 \theta$ term) is given in Sect. 6.2.5.

The perturbation function for short-period terms of the other five orbital elements is given by

$$R_{2s} = R_s(J_2) + R_s(J_l)_{l \geq 3} + R_s(C_{2,2}, S_{2,2})_{l \geq 2} + R_s(\mu_1), \quad (6.102)$$

where

$$R_s(J_2) = \frac{3J_2}{2a^3} \left\{ \frac{1}{3} \left(1 - \frac{3}{2} \sin^2 i \right) \left[\left(\frac{a}{r} \right)^3 - (1 - e^2)^{-\frac{3}{2}} \right] + \frac{1}{2} \sin^2 i \left(\frac{a}{r} \right)^3 \cos 2u \right\}, \quad (6.103)$$

$$R_s(J_l) = R(J_l) - R(J_l)_{c,l}, \quad (6.104)$$

$$R(J_l) = \sum_{l \geq 3} \frac{(-J_l)}{a^{l+1}} \sum_{p=0}^{(l-\delta_1)/2} F_3(i) \left[(1 - \delta) \left(\frac{a}{r} \right)^{l+1} \cos(l - 2p)u + \delta_1 \left(\frac{a}{r} \right)^{l+1} \sin(l - 2p)u \right],$$

$$R(J_l)_{c,l} = \frac{\sqrt{1-e^2}}{a} \sum_{l \geq 3} \frac{(-J_l)}{p_0^l} \sum_{p=1}^{(l-\delta_1)/2} F_3(i) K_3(e) H(\omega),$$

$$R_s(C_{l,m}, S_{l,m}) = R(C_{l,m}, S_{l,m}) - R_l(C_{l,m}, S_{l,m}), \quad (6.105)$$

$$\begin{aligned}
R(C_{l,m}, S_{l,m}) &= \sum_{l \geq 2} \sum_{m=1}^l \frac{1}{a^{l+1}} \sum_{p=0}^l F_{lmp}(i) \left(\frac{a}{r}\right)^{l+1} \\
&\quad \times \left\{ [(1 - \delta_m)C_{l,m} - \delta_m S_{l,m}] \cos((l - 2p)u + m\Omega_G) \right. \\
&\quad \left. + [(1 - \delta_m)S_{l,m} + \delta_m C_{l,m}] \sin((l - 2p)u + m\Omega_G) \right\}, \\
R_l(C_{l,m}, S_{l,m}) &= \frac{\sqrt{1 - e^2}}{a} \sum_{l \geq 2} \sum_{m=1}^l \frac{1}{p_0^l} \sum_{p=1}^{l-1} F_{lmp}(i) K_3(e) \Psi_{lmp}, \\
R_s(\mu_1) &= \left(\frac{3}{2}\beta a^2\right) \left\{ S_1 \left[\left(\frac{r}{a}\right)^2 - \left(1 + \frac{3}{2}e^2\right) \right] \right. \\
&\quad \left. + S_2 \left[\left(\frac{r}{a}\right)^2 \cos 2f - \frac{5}{2}e^2 \right] + S_3 \left(\frac{r}{a}\right)^2 \sin 2f \right\}. \quad (6.106)
\end{aligned}$$

In the above expressions, the auxiliaries are already provided.

Based on the above analysis we have $a_s^{(2)}(t) = 2a^2 R_{2s}$, and the short-period terms of the other five orbital elements are

$$\begin{aligned}
e_s^{(2)}(t) &= \frac{3J_2}{2a^2} \left(\frac{1 - e^2}{e}\right) \left\{ \frac{1}{3} \left(1 - \frac{3}{2} \sin^2 i\right) \left[\left(\frac{a}{r}\right)^3 - (1 - e^2)^{-3/2} \right] \right. \\
&\quad \left. + \frac{1}{2} \sin^2 i \left(\frac{a}{r}\right)^3 \cos 2(f + \omega) \right. \\
&\quad \left. - \frac{\sin^2 i}{2(1 - e^2)^2} [e \cos(f + 2\omega) + \cos 2(f + \omega) + \frac{e}{3} \cos(3f + 2\omega)] \right\} \\
&\quad - \frac{3J_2}{2p^2} \sin^2 i \left(\frac{1 - e^2}{6e}\right) \overline{\cos 2f} \cos 2\omega \quad (6.107)
\end{aligned}$$

$$\begin{aligned}
i_s^{(2)}(t) &= \frac{3J_2}{8p^2} \sin 2i \left[e \cos(f + 2\omega) + \cos 2(f + \omega) + \frac{e}{3} \cos(3f + 2\omega) \right] \\
&\quad + \frac{3J_2}{24p^2} \sin 2i \overline{\cos 2f} \cos 2\omega \quad (6.108)
\end{aligned}$$

$$\begin{aligned}
\Omega_s^{(2)}(t) &= -\frac{3J_2}{2p^2} \cos i \left\{ (f - M + e \sin f) \right. \\
&\quad \left. - \frac{1}{2} [e \sin(f + 2\omega) + \sin 2(f + \omega) + \frac{e}{3} \sin(3f + 2\omega)] \right\} \\
&\quad + \frac{3J_2}{12p^2} \cos i \overline{\cos 2f} \sin 2\omega, \quad (6.109)
\end{aligned}$$

$$\begin{aligned}
\omega_s^{(2)} &= \frac{3J_2}{2p^2} \left\{ \left(2 - \frac{5}{2} \sin^2 i\right) (f - M + e \sin f) \right. \\
&\quad \left. + \left(1 - \frac{3}{2} \sin^2 i\right) \left[\left(\frac{1}{e} - \frac{e}{4}\right) \sin f + \frac{1}{2} \sin 2f + \frac{e}{12} \sin 3f \right] \right. \\
&\quad \left. - \left[\frac{1}{4e} \sin^2 i + \left(\frac{1}{2} - \frac{15}{16} \sin^2 i\right) e \right] \sin(f + 2\omega) - \left(\frac{1}{2} - \frac{5}{4} \sin^2 i\right) \sin 2(f + \omega) \right\}
\end{aligned}$$

$$\begin{aligned}
& + \left[\frac{7}{12e} \sin^2 i - \left(\frac{1}{6} - \frac{19}{48} \sin^2 i \right) e \right] \sin(3f + 2\omega) + \frac{3}{8} \sin^2 i \sin(4f + 2\omega) \\
& + \frac{e}{16} \sin^2 i [\sin(5f + 2\omega) + \sin(f - 2\omega)] \Big\} \\
& - \frac{3J_2}{2p^2} \left[\sin^2 i \left(\frac{1}{8} + \frac{1-e^2}{6e^2} \cos 2f \right) + \frac{1}{6} \cos^2 i \cos 2f \right] \sin 2\omega, \quad (6.110) \\
M_s^{(2)}(t) = & \frac{3J_2}{2p^2} \sqrt{1-e^2} \Big\{ - \left(1 - \frac{3}{2} \sin^2 i \right) \left[\left(\frac{1}{e} - \frac{e}{4} \right) \sin f + \frac{1}{2} \sin 2f + \frac{e}{12} \sin 3f \right] \right. \\
& + \sin^2 i \left[\left(\frac{1}{4e} + \frac{5}{16} e \right) \sin(f + 2\omega) \right. \\
& - \left(\frac{7}{12e} - \frac{e}{48} \right) \sin(3f + 2\omega) - \frac{3}{8} \sin(4f + 2\omega) \\
& \left. \left. - \frac{e}{16} \sin(5f + 2\omega) - \frac{e}{16} \sin(f - 2\omega) \right] \right\} \\
& + \frac{3J_2}{2p^2} \sqrt{1-e^2} \sin^2 i \left(\frac{1}{8} + \frac{1+e^2/2}{6e^2} \cos 2f \right) \sin 2\omega \quad (6.111)
\end{aligned}$$

where

$$\left(\frac{a}{r} \right) = \frac{1 + e \cos f}{1 - e^2} = \frac{1}{1 - e \cos E}, \quad (6.112)$$

$$\begin{cases} f = M + 2e \sin M, & f - M = 2e \sin M, \\ E = M + e \sin M, & E - M = e \sin M, \end{cases} \quad (6.113)$$

$$\overline{\cos 2f} = \frac{1 + 2\sqrt{1-e^2}}{(1 + \sqrt{1-e^2})^2} e^2. \quad (6.114)$$

The true anomaly f and the eccentric anomaly E are expressed using simplified approximate formulas, which are accurate enough for short-period terms. If necessary, they can be obtained by solving the Kepler equation.

6.2.4.2 Expressions of the Non-singularity Elements $(a, i, \Omega, \xi, \eta, \lambda)$ Solution

For a Moon's prober, the singularity only appears when e is small. To eliminate it we use the non-singularity elements for $0 \leq e < 1$ which is $\sigma = (a, i, \Omega, \xi, \eta, \lambda)$ that $\xi = e \cos \omega$, $\eta = e \sin \omega$ and $\lambda = M + \omega$.

The coordinate system and calculation units are the same as for Kepler elements. The method of constructing the non-singularity perturbation solution is the same as that used for Earth's satellites.

The secular terms $\Delta\sigma(t)$ (including the long-periodic terms) due to perturbations of gravity forces can be expressed as

$$\Delta a(t) = 0, \quad (6.115)$$

$$\Delta i(t) = \Delta i_l(t), \quad (6.116)$$

$$\Delta \Omega(t) = \Delta \Omega_c(t) + \Delta \Omega_l(t), \quad (6.117)$$

$$\Delta \xi(t) = \cos \omega [\Delta e(t)] - \sin \omega [e \Delta \omega(t)], \quad (6.118)$$

$$\Delta \eta(t) = \sin \omega [\Delta e(t)] + \cos \omega [e \Delta \omega(t)], \quad (6.119)$$

$$\Delta \lambda(t) = \bar{n}(t - t_0) + [\Delta M(t) + \Delta \omega(t)], \quad (6.120)$$

where $\bar{n} = \bar{a}^{-3/2} = \bar{a}_0^{-3/2}$.

The short-period terms are

$$a_s(t) = a_s^{(1)}(t) + a_s^{(2)}(t), \quad (6.121)$$

$$i_s^{(1)}(t) = i_s^{(1)}(t), \quad (6.122)$$

$$\Omega_s^{(1)}(t) = \Omega_s^{(1)}(t), \quad (6.123)$$

$$\xi_s^{(1)}(t) = \cos \omega [e_s^{(1)}(t)] - \sin \omega [e \omega_s^{(1)}(t)], \quad (6.124)$$

$$\eta_s^{(1)}(t) = \sin \omega [e_s^{(1)}(t)] + \cos \omega [e \omega_s^{(1)}(t)], \quad (6.125)$$

$$\lambda_s^{(1)}(t) = [M_s^{(1)}(t)] + [\omega_s^{(1)}(t)]. \quad (6.126)$$

In the above expressions (6.116)–(6.126) the perturbation solutions, $\Delta e(t)$, $\Delta i(t)$, $\Delta \omega(t)$, $\Delta \Omega(t)$, $\Delta M(t)$, and $a_s^{(1)}(t)$, $e_s^{(1)}(t)$, $i_s^{(1)}(t)$, $\omega_s^{(1)}(t)$, $\Omega_s^{(1)}(t)$, $M_s^{(1)}(t)$, are expressed by the results of Kepler elements. By this treatment, the perturbation solution which is recombined by the Kepler element solution no longer has the $(1/e)$ factor. The details in the process of eliminating the small divisor, are not repeated here.

6.2.5 Additional Perturbation of Coordinate System [5, 6]

The oscillation of the Moon's equator is due to the Moon's physical libration. For the general accuracy requirement, only the variation related to the main part of the Moon's non-spherical gravity force, J_2 , is needed. The perturbation potential due to the physical libration can be obtained using the coordinate transformation (1.91) in Sect. 1.4.3. Substituting $\vec{R} = (A)^T \vec{r}$ into the Moon's non-spherical gravity potential $V(J_2)$ yields the additional potential denoted to $\Delta V_2(J_2\theta)$, θ is the libration parameter, that

$$\begin{aligned} \Delta V(J_2\theta) = & \left(\frac{3J_2}{4a^3} \right) \left(\frac{a}{r} \right)^3 \\ & \times \{ 2 \sin i \sin 2u [\sigma_1 \sin(L_m - \omega_m - \Omega) + \sigma_2 \sin(L_m + \omega_m - \Omega) \\ & + \sigma_3 \sin(2L_m - \Omega_m - \Omega) + \sigma_4 \cos(L_m - \Omega)] \\ & - \sin 2i (1 - \cos 2u) [\sigma_1 \cos(L_m - \omega_m - \Omega) + \sigma_2 \cos(L_m + \omega_m - \Omega) \\ & + \sigma_3 \cos(2L_m - \Omega_m - \Omega) - \sigma_4 \sin(L_m - \Omega)] \}. \end{aligned} \quad (6.127)$$

As shown in (6.127), the first-order variation of the zonal harmonic term (as represented by the J_2 term) of the Moon's non-spherical gravity potential due to the physical libration is not related to the longitudinal component of the physical libration. It is understandable because the zonal harmonic terms are rotationally symmetric. Also, the inclination component of the physical libration, ρ , does not explicitly appear in the format, it is because we use the node-component σ instead, which is a reasonable approximation.

(1) The solution of the additional perturbation by J_2 term

The additional gravity potential $\Delta V(J_2\theta)$ is the related additional perturbation function, denoted to $R(J_2\theta)$. For a low Moon orbit satellite, it is a quantity of the second-order with respect to the J_2 term, and by the average method (see Sect. 3.4.2) $R(J_2\theta)$ can be decomposed into two parts as

$$R(J_2\theta) = R_l(J_2\theta) + R_s(J_2\theta), \quad (6.128)$$

where R_l and R_s are the long-period part and the short-period part, respectively. Their actual expressions are

$$\begin{aligned} R_l(J_2\theta) = & - \left(\frac{3J_2}{4a^3} \right) \sin 2i (1 - e^2)^{-\frac{3}{2}} \{ \sigma_1 \cos(L_m - \omega_m - \Omega) + \sigma_2 \cos(L_m + \omega_m - \Omega) \\ & + \sigma_3 \cos(2L_m - \Omega_m - \Omega) - \sigma_4 \sin(L_m - \Omega) \} \\ = & R_l(J_2\theta; e, i, \Omega), \end{aligned} \quad (6.129)$$

$$\begin{aligned}
R_s(J_2\theta) = & \left(\frac{3J_2}{4a^3} \right) \left\{ -\sin 2i \left[\left(\frac{a}{r} \right)^3 - (1 - e^2)^{-\frac{3}{2}} - \left(\frac{a}{r} \right)^3 \cos 2u \right] \right. \\
& \times [\sigma_1 \cos(L_m - \omega_m - \Omega) + \sigma_2 \cos(L_m + \omega_m - \Omega) \\
& + \sigma_3 \cos(2L_m - \Omega_m - \Omega) - \sigma_4 \sin(L_m - \Omega)] \\
& + 2 \sin i \left(\frac{a}{r} \right)^3 \sin 2u [\sigma_1 \sin(L_m - \omega_m - \Omega) + \sigma_2 \sin(L_m + \omega_m - \Omega) \\
& \left. + \sigma_3 \sin(2L_m - \Omega_m - \Omega) + \sigma_4 \cos(L_m - \Omega)] \right\}. \quad (6.130)
\end{aligned}$$

The long-period solution of Kepler elements due to the Moon's physical libration related to the J_2 term can be given by

$$\Delta\sigma_l(t) = \sigma_l(t) - \sigma_l(t_0).$$

The components of $\sigma_l(t)$ are

$$a_l(t) = 0, \quad (6.131)$$

$$e_l(t) = 0, \quad (6.132)$$

$$\begin{aligned}
i_l(t) = & -\left(\frac{3J_2}{2p^2} \right) \cos i \left[\frac{\sigma_1}{\alpha_1} \cos(L_m - \omega_m - \Omega) + \frac{\sigma_2}{\alpha_2} \cos(L_m + \omega_m - \Omega) \right. \\
& \left. + \frac{\sigma_3}{\alpha_3} \cos(2L_m - \Omega_m - \Omega) - \frac{\sigma_4}{\alpha_4} \sin(L_m - \Omega) \right], \quad (6.133)
\end{aligned}$$

$$\begin{aligned}
\Omega_l(t) = & -\left(\frac{3J_2}{2p^2} \right) \frac{\cos 2i}{\sin i} \left[\frac{\sigma_1}{\alpha_1} \sin(L_m - \omega_m - \Omega) + \frac{\sigma_2}{\alpha_2} \sin(L_m + \omega_m - \Omega) \right. \\
& \left. + \frac{\sigma_3}{\alpha_3} \sin(2L_m - \Omega_m - \Omega) + \frac{\sigma_4}{\alpha_4} \cos(L_m - \Omega) \right], \quad (6.134)
\end{aligned}$$

$$\begin{aligned}
\Omega_l(t) = & \left(\frac{3J_2}{2p^2} \right) \frac{\cos i}{\sin i} (1 - 5 \sin^2 i) \left[\frac{\sigma_1}{\alpha_1} \sin(L_m - \omega_m - \Omega) + \frac{\sigma_2}{\alpha_2} \sin(L_m + \omega_m - \Omega) \right. \\
& \left. + \frac{\sigma_3}{\alpha_3} \sin(2L_m - \Omega_m - \Omega) + \frac{\sigma_4}{\alpha_4} \cos(L_m - \Omega) \right], \quad (6.135)
\end{aligned}$$

$$\begin{aligned}
M_l(t) = & -\left(\frac{9J_2}{4p^2} \right) \sqrt{1 - e^2} \sin 2i \left[\frac{\sigma_1}{\alpha_1} \sin(L_m - \omega_m - \Omega) + \frac{\sigma_2}{\alpha_2} \sin(L_m + \omega_m - \Omega) \right. \\
& \left. + \frac{\sigma_3}{\alpha_3} \sin(2L_m - \Omega_m - \Omega) + \frac{\sigma_4}{\alpha_4} \cos(L_m - \Omega) \right], \quad (6.136)
\end{aligned}$$

where $\alpha_1, \alpha_2, \alpha_3$, and α_4 are defined as

$$\begin{cases} \alpha_1 = \frac{1}{n} (\dot{L}_m - \dot{\omega}_m - \dot{\Omega}) = O(10^{-3}), \\ \alpha_2 = \frac{1}{n} (\dot{L}_m + \dot{\omega}_m - \dot{\Omega}) = O(10^{-3}), \\ \alpha_3 = \frac{1}{n} (2\dot{L}_m - \dot{\Omega}_m - \dot{\Omega}) = O(10^{-3}), \\ \alpha_4 = \frac{1}{n} (\dot{L}_m - \dot{\Omega}) = O(10^{-3}), \end{cases} \quad (6.137)$$

and L_m is the Moon's mean ecliptic longitude.

(2) The non-singularity solution due to the additional perturbation of J_2 term

The non-singularity elements are a , i , Ω , $\xi = e \cos \omega$, $\eta = e \sin \omega$, and $\lambda = M + \omega$. The corresponding $a_l(t)$, $i_l(t)$, and $\Omega_l(t)$ have the same expressions given above, and $\xi_l(t)$, $\eta_l(t)$, and $\lambda_l(t)$ are as follows.

$$\xi_l(t) = -\eta \omega_l(t), \quad (6.138)$$

$$\eta_l(t) = \xi \omega_l(t), \quad (6.139)$$

$$\begin{aligned} \lambda_l(t) &= M_l(t) + \omega_l(t) \\ &= \left(\frac{3J_2}{2p^2} \right) \frac{\cos i}{\sin i} \left[(1 - 5 \sin^2 i) - 3\sqrt{1 - e^2} \sin^2 i \right] \\ &\quad \times \left[\frac{\sigma_1}{\alpha_1} \sin(L_m - \omega_m - \Omega) + \frac{\sigma_2}{\alpha_2} \sin(L_m + \omega_m - \Omega) \right. \\ &\quad \left. + \frac{\sigma_3}{\alpha_3} \sin(2L_m - \Omega_m - \Omega) + \frac{\sigma_4}{\alpha_4} \cos(L_m - \Omega) \right]. \end{aligned} \quad (6.140)$$

The physical libration parameters, σ_1 , σ_2 , σ_3 , and σ_4 , are given in (1.95), and approximately there are

$$\begin{aligned} \rho_1 = \sigma_1 &= -99''.1, \quad \rho_2 = \sigma_2 = 24''.6, \quad \rho_3 = \sigma_3 = -10''.1 \\ \rho_4 &= -\sigma_4 = -80''.8 \end{aligned} \quad (6.141)$$

The perturbation potential (6.128) only has a short-period term and a long-period term, R_l and R_s , respectively, which affect periodic variations of the Moon's satellite orbit. For the short-period variation, only the effect on the semi-major axis needs to be considered, which is

$$a_s(J_2\theta) = 2a^2 R_s(J_2\theta) \quad (6.142)$$

6.2.6 Applications of Analytical Orbital Solution in Orbital Design

There are three most concerned types of the Moon's satellite orbit as follows.

- (1) Regression orbit, i.e., the satellite orbital period and the Moon's rotational period form a ratio of simple integers.
- (2) The Sun-synchronous orbit, i.e., the precession of a satellite's orbital plane is in the same direction of the central body's revolution around the Sun. When

looking from the central body, both the orbital plane of the satellite and the Sun “move” eastward with the same speed.

(3) Frozen orbit, i.e., the direction of the major axis (i.e., the line of apsidal) is fixed.

(1) Regression Orbit

Among the three types of orbit, the first type is very simple. For example, for an Earth's satellite, according to the condition of sunshine, the requirement is that the satellite moves around Earth passing the same point over a fixed location on the ground 12 times a day, then the period of the satellite must be 2 h. For a regression orbit of a Moon's satellite the requirement is similar, and there is no other complicated requirement for the realization of this type of orbit. The condition of a regression orbit can be combined with the conditions for the other two types of orbit, therefore, we do not discuss the design of a regression orbit separately.

(2) Design of a Sun-synchronous orbit

The precession of a satellite's orbital plane is mainly caused by the J_2 term of the central body's non-spherical gravity potential. To make the speed of the orbital precession the same as the Sun's speed “around” the central body (actually, it is the central body that moves around the Sun), the condition by the dimensionless system is

$$\dot{\Omega} = n_s, \quad (6.143)$$

where n_s is the Sun's angular speed “around” the Moon. Obviously, it can only happen in an average system, i.e., $\dot{\Omega}$ should be the average rate of the longitude variation of the satellite's ascending node. By means of the “first-order” solution we have

$$\dot{\Omega} = -\frac{3J_2}{2p^2}n \cos i = n_s. \quad (6.144)$$

If we consider the effects of the four main perturbations on $\dot{\Omega}$, then the condition of a Sun-synchronous orbit is

$$\dot{\Omega} = \Omega_{2c}(J_2) + \Omega_{3c}(J_l) + \Omega_{4c}(J_2^2) + \Omega_{3c}(\mu_1) = n_s. \quad (6.145)$$

Each part in (6.145) is given in Sect. 6.2.4. For convenience, they are given here again,

$$\Omega_{3c}(J_l) = -n \cos i \sum_{l(2) \geq 4} \left(\frac{J_l}{p_0^l} \right) F_2(i) K_1(e), \quad (6.146)$$

$$\Omega_{4c}(J_2^2) = -\left(\frac{3J_2}{2p^2} \right)^2 n \cos i \left[\left(\frac{3}{2} + \frac{1}{6}e^2 + \sqrt{1-e^2} \right) \right]$$

$$-\left(\frac{5}{3} - \frac{5}{24}e^2 + \frac{3}{2}\sqrt{1-e^2}\right)\sin^2 i \Big], \quad (6.147)$$

$$\Omega_{3c}(\mu_1) = -n \cos i \left(\frac{3}{4}\beta_1 a^3\right) \left(1 - \frac{3}{2}\sin^2 i'\right) \left(1 + \frac{3}{2}e^2\right) (1-e^2)^{-\frac{1}{2}}, \quad (6.148)$$

where $p_0 = a(1-e^2)$; $F_2(i)$, $K_1(e)$, and β_1 are given in (6.62)–(6.64). In the case of the Moon, the magnitudes of even zonal harmonic terms ($J_{2,l}$, $l \geq 2$), which affect the precession of the orbit plane, are smaller than the J_2 term by 1–2 orders, the Earth's gravity force perturbation is also smaller than the J_2 term, the other perturbations are even smaller, therefore, for a Sun-synchronous orbit, we include only the effect of the J_2 term in (6.144). The results can basically represent the actual motion state of a Sun-synchronous satellite.

We now use a low Moon orbit satellite as an example. Considering the J_2 effect only, the satellite moves around the Moon 12 circles per day (Earth day), the main orbital parameters are

$$\begin{aligned} \bar{h} &= 120.0\text{km}, & T_S &= 120^{\text{m}} \cdot 0, \\ a &= 1860.3\text{km}, & e &= 0.0010, & i &= 148^\circ.4637. \end{aligned} \quad (6.149)$$

If we need the accurate value of the inclination i , it can be obtained by the relationship of a , e , and i that for given values of a and e there is

$$\begin{aligned} \cos i &= -\frac{n_s}{n} \left\{ \frac{3J_2}{2p^2} + \sum_{l(2) \geq 4} \left(\frac{J_l}{p_0^l} \right) F_2(i) K_1(e) \right. \\ &\quad \left. + \left(\frac{3}{4}\beta_1 a^3 \right) \left(1 - \frac{3}{2}\sin^2 i' \right) \left(1 + \frac{3}{2}e^2 \right) (1-e^2)^{-\frac{1}{2}} \right\}^{-1}. \end{aligned} \quad (6.150)$$

Note that the three main orbital elements, a , e , and i are the mean orbit elements, $\bar{a}$, $\bar{e}$, $\bar{i}$, but from the point of view of orbit design, they can be regarded as the instantaneous orbit elements.

Similar to Earth's situation, the Sun-synchronous orbit of the Moon's satellite is also retrograde, i.e., the orbital inclination $i > 90^\circ$. Also because the perturbation of the Moon's dynamics form-factor J_2 is smaller than that of Earth, the inclination of the Sun-synchronous orbit of Moon's satellite is far away from the inclination of a polar orbit, as shown by (6.149).

(3) Design of a frozen orbit

① General information about frozen orbits

Similar to an Earth's satellite, a frozen orbit means an orbit with a fixed line of apsides, i.e., the direction of the perilune is fixed. For keeping a fixed line of apsides there are two kinds of frozen orbit, one is an orbit with the critical inclination, the other is an orbit with any inclination not equal to the critical inclination.

The frozen orbit with the critical inclination is decided by the variation rate of the secular term of ω due to the J_2 term, in the dimensionless system which is

$$\dot{\omega} = \frac{3J_2}{2p^2} n \left(2 - \frac{5}{2} \sin^2 i \right). \quad (6.151)$$

When $\dot{\omega} = 0$, the direction of the perilune is fixed, and the corresponding inclination is called the critical inclination, denoted to i_c , that

$$i_c = 63^\circ 26', \quad 116^\circ 34'. \quad (6.152)$$

The characteristics of this kind of frozen orbit are decided by the non-spherical gravity force of the central body, the same as a frozen orbit of Earth's satellite.

Generally speaking, a frozen orbit is a particular orbit for any inclination. Similar to the Sun-synchronous orbit this kind of orbit is a mean orbit solution of the motion equation. For a low orbit satellite, the mean system depends mainly on the zonal harmonic term J_l ($l \geq 2$) of the non-spherical gravity potential of the central body (which is the Moon here). Similar to the case of Earth, there are two possibilities for the Moon's frozen orbit, that

$$\omega = 90^\circ \text{ or } 270^\circ. \quad (6.153)$$

For a given semi-major axis a and an inclination i , the eccentricity e can be defined by the simplified formula as

$$e = \left| \frac{J_3}{J_2} \right| \frac{1}{2a} \sin i [1 + O(\varepsilon^2)], \quad (6.154)$$

where ε^2 represents all small variables of higher order with respect to $|J_3/J_2|$. If $(J_3/J_2) > 0$, the frozen orbit has $\omega = 270^\circ$; when $(J_3/J_2) < 0$, there is $\omega = 90^\circ$. For the Moon's non-spherical gravity potential, there is $(J_3/J_2) > 0$, therefore, if only the effects of J_2 and J_3 are considered the frozen orbit has $\omega = 270^\circ$, which is different from a frozen orbit of Earth's satellite.

As shown in (6.154) the relative magnitude of the main zonal harmonic odd term J_3 and the J_2 term decides the value of the eccentricity of a frozen orbit. For the Moon, the eccentricity of a frozen orbit can reach 10^{-1} . It is significantly different from a frozen orbit of an Earth's satellite, which is almost a circle because its eccentricity is only 10^{-3} .

② The solution of frozen orbit

Similar to an Earth's satellite, by the action of the Moon's non-spherical gravity force the mean motion system (i.e., the system after eliminating the short-period variations) of a satellite may have a special solution, whose orbital elements are

$$\bar{a}(t) = a_0, \quad \bar{e}(t) = e_0, \quad \bar{l}(t) = i_0, \quad \bar{\omega}(t) = \omega_0 = 90^\circ \text{ or } 270^\circ. \quad (6.155)$$

This solution has a fixed line of apsides, i.e., a frozen orbit. There is no restriction on $\bar{i}_0$, for a given $\bar{i}_0$, there is a corresponding $\bar{e}_0$. Because the Moon's gravitational field is different from that of Earth, the frozen orbit of the Moon is also different from that of Earth. For an Earth's satellite, basically the two terms of J_2 and J_3 can determine a frozen orbit. For Moon, it is different. Using the same mean orbital elements, a , e , i , and ω , we can derive the solution of a frozen orbit (the principle and method are the same as for an Earth's frozen orbit). For a given a and any value of i , the value of e of a frozen orbit satisfies the following condition

$$e = \pm \sin i \sum_{l(2) \geq 3} (-1)^{\frac{l-1}{2}} (l-1) \left(\frac{J_l}{a_0^l} \right) F^*(i) \left(\frac{n}{\omega_c} \right), \quad (6.156)$$

where

$$\begin{cases} F^*(i) = \sum_{q=0}^{(l-1)/2} (-1)^q \left(\frac{1}{2} \right)^{(l+2q+1)} C_{lpq}^* (\sin^2 i)^q, \\ C_{lpq}^* = \binom{l}{(l-1)/2 - q} \binom{l+2q+1}{l} \binom{2q+1}{q}, \end{cases} \quad (6.157)$$

and ω_c is the rate of the secular variation of ω . If we only consider the effect of J_2 term, ω_{2c} , then (6.157) can be reasonably simplified to

$$e = \pm \sin i \left\{ \sum_{l(2) \geq 3} (-1)^{\frac{l-1}{2}} \left(\frac{2}{3a_0^{l-2}} \right) (l-1) \left(\frac{J_l}{J_2} \right) F^*(i) \right\} / \left(2 - \frac{5}{2} \sin^2 i \right). \quad (6.158)$$

On the right side the sign “+” corresponds to a frozen orbit solution with $\omega \equiv \omega_0 = 90^\circ$, and the sign “−” is for $\omega_0 = 270^\circ$. In the operation of the sum, the sub $l(2)$ increases by a step-size of 2, i.e., $l(2) = 3, 5, \dots$.

As shown above for a low Moon orbit satellite the solution of a frozen orbit has two possibilities, i.e., $\omega \equiv \omega_0 = 90^\circ$ or 270° , corresponding to different values of i and a relatively large e . Table 6.6 lists calculated e values for a given i and a period of $T_s \approx 117.849$ min. The calculation is for i changing from 1° to 179° in every degree, only some of the results are listed. An Earth's frozen orbit depends on the J_3 term, the values of e are always small for different values of i , i.e., $e_0 = O(10^{-3})$. The relatively larger values of e of a Moon's frozen orbit reflect the uneven mass distribution compared to Earth such as the existence of mass lumps.

In Table 6.6 there are some orbits with large values of e (such as when $i = 90^\circ$). Under perturbations of external forces, the value of e varies, and it is entirely possible that the height of the perilune reduces to $h_p = 0$, so the satellite crashes on the Moon. This is the kind of dynamical mechanism without a dissipation force that decides the lifespan of a Moon's satellite. By this mechanism, an initially small eccentricity

Table 6.6 The Moon's frozen orbit solutions

i (deg)	ω (deg)	e	i (deg)	ω (deg)	e
1.0	90.0	0.005647	50.0	90.0	0.000870
5.0	90.0	0.025923	55.0	90.0	0.140493
10.0	90.0	0.040836	60.0	270.0	0.091597
20.0	90.0	0.021863	63.0	270.0	0.188417
28.0	270.0	0.002481	75.0	270.0	0.044756
35.0	270.0	0.060784	77.0	270.0	0.009016
40.0	270.0	0.047442	80.0	90.0	0.026043
45.0	270.0	0.046151	85.0	270.0	0.001753
49.5	270.0	0.007062	90.0	270.0	0.043215
50.0	90.0	0.000870	95.0	270.0	0.001728

influenced by the long-period zonal harmonic odd terms of the Moon's non-spherical gravity potential, J_{2l-1} ($l \geq 2$) (related to the slowly varying element ω), can increase and eventually reach a value that leads $h_p = 0$. The lifespan of a Moon's satellite is relatively short, the inclination of a frozen orbit corresponds to an eccentricity which is not too small, therefore, the long-period variation of the eccentricity cannot be restricted by the frozen orbit. As a result, a frozen orbit cannot control the height of h_p , which may reduce to the state of $h_p = 0$ causing a crash. The gravitational mechanism affecting the lifespan of a low Moon orbit satellite was discussed in Sect. 5.6.1.

6.3 Perturbed Orbital Solution of Mars's Satellite

For a low Mars orbit satellite, if the original mean height $\bar{h}$ is between 200 and 400 km, the forces acting on the satellite besides the gravity force of Mars's mass (as a particle) include the usual perturbation forces, which are Mars's non-spherical gravity force, the perturbation due to Mars equatorial oscillation, the Sun's gravity force, the gravity forces of Mars's natural satellites (as the Moon's gravity force on an Earth's artificial satellite), gravity forces of other major planets, the solar radiation pressure perturbation, and Mars's atmospheric drag, etc. There is also the post-Newtonian effect. The question about which perturbations should be considered depends on each specific situation.

6.3.1 Selection of Coordinate System

For a low Mars orbit satellite, the main perturbation is still the Mars's non-spherical gravity potential. Similar to the method of dealing with an Earth's satellite, we choose an epoch J2000.0 Mars-centric mean equatorial frame. As introduced in Sect. 1.5, the origin of the frame is at Mars's barycenter, which should be consistent with the chosen Mars's gravitational field model, the fundamental plane (i.e., the xy -plane) is Mars's mean equatorial plane. Because Mars's equator oscillates, there is also the related additional coordinate perturbation, the details about this perturbation are given in Sect. 1.5, and the effect on the perturbation solution is considered when we construct the solution.

6.3.2 The Mathematical Model of Perturbed Motion for a Mars's Satellite

It is convenient for analyzing and comparing different physical terms to use normalized units and a dimensionless system in constructing the perturbation solution of an orbital motion for a Mars's satellite in the same way dealing with an Earth's satellite. We choose the units for length $[L]$, mass $[M]$, and time $[T]$ as

$$\begin{cases} [L] = a_e \text{ (equatorial radius of Mars's reference spheroid),} \\ [M] = M \text{ (Mars's mass, which value is replaced by } GM), \\ [T] = (a_e^3 / GM)^{\frac{1}{2}}. \end{cases} \quad (6.159)$$

The corresponding Mars's gravitational constant $\mu = GM = 1$. Accepting the Goddard Mars gravity model, GMM-2B, there are

$$\begin{cases} a_e = 3397.0 \text{ km,} \\ GM = 42828.3719 \text{ km}^3/\text{s}^2 \\ [T] = 15.945064755181 \text{ min.} \end{cases} \quad (6.160)$$

As we know the selection of a unit system does not involve the accuracy of the solution. If we use a different Mars's gravity force model, the values of the above units would change a little, and it does not alter the magnitudes of the involved physical terms nor affects the analysis.

For a low Mars orbit satellite, the motion equation in the Mars-centric equatorial frame (as pointed out in Sect. 1.5.1, the nutation of Mars's equator is small, therefore, we do not have to separate a true equator from a mean equator) has the form as

$$\vec{F}_0 = -\frac{\mu}{r^3} \vec{r}, \quad (6.161)$$

$$\vec{F}_\varepsilon = \sum_{j=1}^{10} \vec{F}_j(\vec{r}, \dot{\vec{r}}, t; \varepsilon_j). \quad (6.162)$$

The 10 perturbations included in the motion equation are: Mars's non-spherical gravity force, the perturbation due to Mars equatorial oscillations, the Sun's gravity force, the gravity forces of Mars's natural satellites, the gravity forces of other major planets, the solar radiation pressure perturbation, Mars's atmospheric drag, and the post-Newtonian effect. The magnitudes $\varepsilon_j (j = 1, 2, \dots)$ of the corresponding perturbing accelerations $\vec{F}_j$ can be estimated by

$$\varepsilon_j = \frac{|\vec{F}_j|}{|\vec{F}_0|}. \quad (6.163)$$

For a satellite with a mean height $\bar{h}$ between 200 and 400 km, the estimated magnitudes of each perturbation are as follows.

$$\begin{aligned} \varepsilon_1(J_2) &= O(10^{-3}), \\ \varepsilon_2(J_{2,2}) &= O(10^{-4}), \\ \varepsilon_3(J_3, J_4, \dots, J_{l,m}, \dots) &= O(10^{-5} - 10^{-6}), \\ \varepsilon_4(\text{Precession}) &= O(10^{-7}), \\ \varepsilon_5(\text{Sun}) &= O(10^{-8}), \\ \varepsilon_6(\text{Moon}) &= O(10^{-9}), \\ \varepsilon_7(\text{Planets}) &\leq O(10^{-12}), \\ \varepsilon_8(\text{solar radiation}) &\leq O(10^{-9}), \\ \varepsilon_9(\text{atmosphere}) &= O(10^{-8}), \quad \bar{h} = 300 \text{ km}. \end{aligned} \quad (6.164)$$

The post-Newtonian effect is mainly from the Schwarzschild solution, its magnitude is about

$$\varepsilon_{10} = O(10^{-10}). \quad (6.165)$$

Mars has an atmosphere, but the air is much thinner than Earth's atmosphere, the air pressure on Mars's surface is only 0.007 bar. Although the air density of Mars's atmosphere decreases with increasing height slower compared to Earth's atmosphere (i.e., the scale height of the density is larger and the speed of the density decreasing is slower compared to Earth's air), for a not too low orbit, the effect of the atmosphere on energy dissipation is less obvious compared to an Earth's satellite. Also because the slowly reducing density with increasing altitude (can be treated by an algebraic

model, even a linear model) for the above magnitude analysis we assume the density of Mars's air to be about 1/10 of Earth's air at height $\bar{h} = 300$ km.

The analysis in (6.164) shows that the main perturbation is the Mars's non-spherical gravity force (including the equatorial oscillations). Generally, the perturbation (ε_7) due to other major planes can be omitted, the Sun's perturbation is "marginal", and the non-gravitational forces (radiation pressure and atmospheric drag) depend on the equivalent area-to-mass ratio of a satellite, the estimates given in (6.164) are based on the area-to-mass ratio (S/m) as

$$S/m = 1\text{m}^2/100\text{ kg} = 10^8. \quad (6.166)$$

This ratio is usually not far from the actual value. For some special satellites, it should be treated according to the actual situation.

The above quantitative estimations can be summarized by the following three points.

- (1) Generally only Mars's non-spherical gravity potential, the atmospheric drag (if Mars's air density is as large as 1/10 of Earth's air density, and the area-to-mass ratio of a satellite is large enough), and the additional coordinate perturbation related to Mars's non-spherical force need to be considered.
- (2) The gravity field of Mars's non-spherical part is complicated, the spherical harmonic expansion converges slowly, as a result, more high-order terms are needed.
- (3) Only for high Mars orbit satellites the Sun's gravity perturbation is needed.

6.3.3 *The Analytical Perturbation Solution of Mars's Satellite Orbit [7, 8]*

The physical state of Mars is similar to Earth (mainly the dynamical circumstance of satellite orbital motion, such as the fast rotation, and relatively large non-spherical oblateness), therefore, unlike the Moon's case, the characteristics of a spacecraft orbiting around Mars, the method of constructing the perturbation solution, and actual results are basically similar to these of Earth, only the related parameters need to be adjusted, thus we do not repeat the similar content. But there are three obvious differences, which are the joint perturbation effect of the J_2 term and other zonal harmonic terms of the non-spherical gravity potential; the effect of the additional coordinate perturbation due to the oscillation of Mars's equator; and the perturbation due to the two Mars's natural satellites requiring the orbital information about the two satellites. For providing related results these differences need to be analyzed in detail.

6.3.3.1 The Joint Perturbation Effect of J_2 Term with $J_{2,2}$ and Other Terms

The magnitude of the joint perturbation effect is only $O(10^{-7})$ on the surface, but the actual influence can be 10^{-6} , which is equivalent to $a_s^{(2)}(J_2^2, t)$, therefore, it cannot be ignored. The actual expressions are given as follows.

(1) The short-period term, $a_s^{(2)}(J_2 \cdot J_{2,2}, t)$

By the method of mean orbital elements, the second-order short-period term of a is given by

$$a_s^{(2)}(J_2 \cdot J_{2,2}, t) = \int \left\{ \sum_j \left[\frac{\partial f(J_2)}{\partial \sigma_j} (\sigma_s(J_{2,2}))_j + \frac{\partial f(J_{2,2})}{\partial \sigma_j} (\sigma_s(J_2))_j \right] \right\}_s dt, \quad (6.167)$$

Substituting $f(J_2)$ and $f(J_{2,2})$ and the first-order short-period terms $\sigma_s(J_2)$ and $\sigma_s(J_{2,2})$ into (6.167), then integrating and keeping terms greater than $O(e^0)$, we have

$$\begin{aligned} a_s^{(2)}(J_2 \cdot J_{2,2}, t) = & \frac{9J_2 J_{2,2}}{8a^3} \left\{ \frac{40}{3} \sin^2 i \left(1 - \frac{5}{4} \sin^2 i \right) \cos(2\Omega_{2,2}) \right. \\ & + \left[\sin^2 i (1 + \cos i) \left(\frac{2}{\alpha} + 2 \right) - \sin^4 i \left(\frac{8}{\alpha} + 6 \right) \right. \\ & + \left. \frac{28}{3} \left(1 - \frac{3}{2} \sin^2 i \right) (1 + \cos i)^2 \right] \\ & \times \cos(2M + 2\omega + 2\Omega_{2,2}) \\ & - \left[\sin^2 i (1 - \cos i) \left(\frac{2}{\alpha} + 2 \right) \right. \\ & - \sin^4 i \left(\frac{8}{\alpha} + 6 \right) - \frac{28}{3} \left(1 - \frac{3}{2} \sin^2 i \right) (1 - \cos i)^2 \left. \right] \\ & \times \cos(2M + 2\omega - 2\Omega_{2,2}) \\ & + \frac{4}{3} \sin^2 i \left[(1 + \cos i)^2 \cos(4M + 4\omega + 2\Omega_{2,2}) \right. \\ & + (1 - \cos i)^2 \cos(4M + 4\omega - 2\Omega_{2,2}) \left. \right] \left. \right\}, \quad (6.168) \end{aligned}$$

where

$$\begin{cases} \Omega_{2,2} = \Omega - S_{2,2} - n_m(t - t_0), \\ \alpha = n_m/n. \end{cases} \quad (6.169)$$

In (6.169) Ω is the quasi-mean element at time t , $\overline{\Omega}(t)$; $S_{2,2}$ is the local “sidereal time” of the “symmetry axis” direction of Mars’s equator at epoch t_0 given by the Greenwich sidereal time on Mars, the Greenwich sidereal time is given by (1.109) in Chap. 1, and $S_{2,2}$ is W in (1.109); n_m and n are Mars’s rotation speed and the mean angular speed of the satellite, respectively. The variable α is the rotation factor of Mars, for a low Mars orbit satellite $\alpha \leq 0.1$. In (6.168) because α is a divisor, the terms having factor $1/\alpha$ are larger than other terms by an order; as a result, the magnitude of $a_s^{(2)}(t) = a_s^{(2)}(J_2 \cdot J_{2,2}, t)$ can be larger than 10^{-6} . That is why the terms related to Mars’s rotation are important. The joint part which is related to α in (6.168) is

$$a_s^{(2)}(J_2 \cdot J_{2,2}, t) = \frac{9J_2 J_{2,2}}{4a^3} \left(\frac{1}{\alpha} \right) \left\{ \left[\sin^2 i (1 + \cos i) - 4 \sin^4 i \right] \cos(2M + 2\omega + 2\Omega_{2,2}) - \left[\sin^2 i (1 - \cos i) - 4 \sin^4 i \right] \cos(2M + 2\omega - 2\Omega_{2,2}) \right\}. \quad (6.170)$$

The magnitude of $a_s^{(2)}(J_2 \cdot J_{2,2}, t)$ can be $O(10^{-6})$ – $O(10^{-5})$, which causes an effect of $O(10^{-4})$ or larger along the orbit when extrapolating an arc over 1–2 days.

(2) The short-period term $a_s^{(2)}(J_{2,2} \cdot J_{2,2}, t)$

This term is given by the following integration

$$a_s^{(2)}(J_{2,2} \cdot J_{2,2}, t) = \int_t \left[\sum_j \frac{\partial f_1(J_{2,2})}{\partial \sigma_j} (\sigma_s^{(1)}(J_{2,2}))_j \right]_s dt, \quad (6.171)$$

Substituting $f_1(J_{2,2})$ and the first-order short-period terms of $\sigma_s^{(1)}(J_{2,2})$ into (6.171), then integrating and keeping terms greater than $O(e^0)$, we have the joint terms related to the rotation factor α as

$$a_s^{(2)}(J_{2,2} \cdot J_{2,2}, t) = \frac{9J_{22}^2}{8a^3} \frac{1}{\alpha} \left\{ 24 \sin^2 i \cos i \cos(2M + 2\omega) - 8 \sin^2 i (1 + \cos i)^2 \cos(2M + 2\omega + 4\Omega_{2,2}) + 8 \sin^2 i (1 - \cos i)^2 \cos(2M + 2\omega - 4\Omega_{2,2}) \right\}. \quad (6.172)$$

6.3.3.2 The Additional Coordinate Perturbation Due to Mars’s Precession

The IAU 2000 Mars’s orientation model introduced in Sect. 1.5 (Fig. 1.6) provides the necessary connection and transformation between a Mars’s frame and an Earth’s frame. The model also gives the calculation formulas for the right ascension α and declination δ of the Mars’s mean pole in the Mars’s celestial coordinate system which are

$$\begin{aligned}\alpha &= 317^\circ.68143 - 0^\circ.1061T, \\ \delta &= 52^\circ.88650 - 0^\circ.0609T,\end{aligned}\tag{6.173}$$

where T is the Julian century number of time t starting from J2000.0. The Mars's nutation is small (its amplitude of oscillation is no more than $1''$) and has no “accumulative” effect on an orbit, which generally can be ignored. What needs to be considered is the additional coordinate perturbation due to the precession of Mars, and we do not have to distinguish the true equator and the mean equator.

- (1) The transformation of Mars-fixed coordinate system $O\text{-}XYZ$ and Mars-centric equatorial coordinate system $O\text{-}xyz$

Assume that the position vector of a Mars's satellite is $\vec{r}$ in the Mars-centric equatorial frame and is $\vec{R}$ in the Mars-fixed frame, then without considering the polar motion and nutation, the transformation of the two systems is given in Sect. 1.5 as

$$\vec{R} = (MP)\vec{r},\tag{6.174}$$

where the coordinate transformation matrix (MP) includes only two rotation matrices, that

$$(MP) = (MR)(PR),\tag{6.175}$$

and MR is Mars's rotation matrix, which is given by (1.109), i.e.,

$$(MR) = R_{\text{IAU}}(t) = R_Z(W),\tag{6.176}$$

$$W = 176^\circ.630 + 350^\circ.89198226d.\tag{6.177}$$

In (6.177), d is the Julian day number starting from J2000.0. The other matrix (PR) is Mars's precession matrix given by (1.111), i.e.,

$$(PR) = R_x(90^\circ - \delta)R_z(-(\alpha_0 - \alpha))R_x(\delta_0 - 90^\circ).\tag{6.178}$$

In dealing with Mars's additional coordinate perturbation, we can use the same method as that for Earth, i.e., to simplify the matrix MP .

The IAU 2000 Mars's orientation model provides the variations of the right ascension α and declination δ of Mars's mean pole in the Mars's celestial frame due to the precession, that

$$\begin{cases} \alpha = \alpha_0 - 0^\circ.1061T, & \delta = \delta_0 - 0^\circ.0609T, \\ \alpha_0 = 317^\circ.68143, & \delta_0 = 52^\circ.88650. \end{cases}\tag{6.179}$$

The formulas in (6.179) are the same as in (6.173). The variations of α and δ are like the secular (long-period) variations of Earth's precession. This formula shows

that in 50 years, the variation of Mars's mean pole is small, we may only keep the first-order parts of $\Delta\alpha = (\alpha - \alpha_0)$ and $\Delta\delta = (\delta - \delta_0)$. To do that we rewrite the precession matrix (6.178) as

$$(PR) = R_x(\theta + \Delta\theta)R_z(-\Delta\mu)R_x(-\theta), \quad (6.180)$$

where

$$\theta = 37^\circ.11350, \quad (6.181)$$

$$\Delta\theta = 0^\circ.0609T, \quad \Delta\mu = 0^\circ.1061T. \quad (6.182)$$

The definition of T is given above for (6.173), $\Delta\mu$ is the precession variation, and $\theta = 90^\circ - \delta_0$.

The transformation relationship of $\vec{R}$ and $\vec{r}$ (6.174) can be simplified using matrices (MR) and (PR) to the accuracy of the first order of $\Delta\mu$. The actual forms are

$$\vec{R} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = (MP)\vec{r} = \begin{pmatrix} X \\ Y \\ z - (\Delta\mu \sin \theta)x - (\Delta\theta)y \end{pmatrix}, \quad (6.183)$$

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} (x \cos W + y \sin W) + \cos W [-(\Delta\mu \cos \theta)y + (\Delta\mu \sin \theta)z] \\ \quad + \sin W [(\Delta\mu \cos \theta)y + (\Delta\theta)z] \\ (-x \sin W + y \cos W) - \sin W [-(\Delta\mu \cos \theta)y + (\Delta\mu \sin \theta)z] \\ \quad + \cos W [(\Delta\mu \cos \theta)y + (\Delta\theta)z] \end{pmatrix}, \quad (6.184)$$

In the content below, the analytical solution of the additional coordinate perturbation is given based on this transformation relationship. The zonal harmonic terms J_l ($l \geq 2$) in Mars's non-spherical gravity potential are related to the additional perturbation but not related to Mars's rotation and only in the Z -direction. The Z -component of $\vec{R}$ is given by

$$Z = z - [(\Delta\mu \sin \theta)x + (\Delta\theta)y], \quad (6.185)$$

$$\Delta\mu \sin \theta = 0^\circ.0640T, \quad \Delta\theta = 0^\circ.0609T. \quad (6.186)$$

The relationships given by (6.183)–(6.186) are the basic formulas for constructing the additional perturbation potential due to the oscillation of Mars's equator.

(2) The additional potential of Mars's non-spherical gravity potential

To build the additional potential, again we use the dimensionless system, the related normalized units are given by (6.160).

Because the magnitude of Mars's precession is small, for the additional coordinate perturbation we only need to give the part due to the main zonal harmonic term J_2 . The corresponding non-spherical gravity potential is given by

$$V(J_2) = \left(-\frac{J_2}{a^3}\right) \left(\frac{a}{R}\right)^3 \left[\left(\frac{3}{2}\right) \left(\frac{Z}{R}\right)^2 - \left(\frac{1}{2}\right) \right], \quad (6.187)$$

where $R = r$. In the above transformation of the Mars-fixed frame and the Mars-centric equatorial frame the polar motion and nutation are ignored, and the transformation is not related to Mars's rotation. We then use the transformation (6.183) to build the additional perturbation solution due to the J_2 term.

Substituting the Z -component in (6.183) into the gravity potential of J_2 term (6.187) yields

$$\begin{aligned} V(J_2) = & \left(-\frac{J_2}{a^3}\right) \left(\frac{a}{r}\right)^3 \left[\left(\frac{3}{2}\right) \left(\frac{z}{r}\right)^2 - \left(\frac{1}{2}\right) \right] \\ & + 3 \left(\frac{J_2}{a^3}\right) \left(\frac{a}{r}\right)^3 \left(\frac{z}{r}\right) \left[(\Delta\mu \sin \theta) \left(\frac{x}{r}\right) + (\Delta\theta) \left(\frac{y}{r}\right) \right]. \end{aligned} \quad (6.188)$$

where

$$\begin{cases} \left(\frac{Z}{r}\right) = \sin i \sin u, \\ \left(\frac{x}{r}\right) = \cos \Omega \cos u - \sin \Omega \sin u \cos i, \\ \left(\frac{y}{r}\right) = \sin \Omega \cos u + \cos \Omega \sin u \cos i. \end{cases} \quad (6.189)$$

The variables $u = f + \omega$, a , i , Ω , ω , and f are all usual orbital elements (Kepler elements). From (6.188) we obtain the additional coordinate perturbation potential in the Mars-centric equatorial coordination system as

$$\begin{aligned} \Delta V(J_2) = & \left(\frac{3J_2}{2a^3}\right) \left(\frac{a}{r}\right)^3 \sin i \{ \cos i [(\Delta\theta) \cos \Omega - (\Delta\mu \sin \theta) \sin \Omega] \\ & + [(\Delta\mu \sin \theta) \cos \Omega + (\Delta\theta) \sin \Omega] \sin 2u \\ & + \cos i [(\Delta\mu \sin \theta) \sin \Omega - (\Delta\theta) \cos \Omega] \cos 2u \}. \end{aligned} \quad (6.190)$$

In 10–20 years from the standard epoch T_0 (J2000.0) the precession variation $\Delta\mu$ can be as large as 10^{-4} (6.186), therefore, the magnitude of the additional coordinate perturbation of J_2 term can reach 10^{-7} . The other additional perturbation due to the main tesseral harmonic term $J_{2,2}$ is one order smaller than the one by the J_2 term, usually can be ignored. For a high precision project, this part of the additional potential can be derived using the same method with the transformation relationship of (X, Y) and (x, y) given in (6.184).

- (3) The perturbation solution of the additional coordinate potential by Mars's J_2 term

The potential of additional coordinate perturbation $\Delta V(J_2)$ can be separated into the long-period term ΔV_l and the short-period term ΔV_s that

$$\Delta V(J_2) = \Delta V_l + \Delta V_s. \quad (6.191)$$

The expressions of the two terms are

$$\begin{cases} \Delta V_l = \left(\frac{3J_2}{2a^3}\right)(1-e^2)^{-\frac{3}{2}} \sin i \cos i [(\Delta y \cos \theta_0) \cos \Omega - (\Delta y \sin \theta_0) \sin \Omega], \\ \Delta V_s = \Delta V(J_2) - \Delta V_l. \end{cases} \quad (6.192)$$

It is clear that the additional coordinate perturbation by Mars's J_2 term only has long-periodic and short-periodic effects. But the long-period term ΔV_l actually is a Poisson term of the mixed-type due to the precession variation of $\Delta\mu$. Because $\Delta\mu$ changes slowly, for a satellite orbit at any time t , $\Delta\mu$ can be treated as a constant corresponding to this time.

The short-period term is only needed for the semi-major axis, that

$$a_s(t) = 2a^2 \Delta V_s. \quad (6.193)$$

The long-period term $\Delta\sigma(t)$ is given by

$$\Delta\sigma(t) = \sigma_l(t) - \sigma_l(t_0), \quad (6.194)$$

where the components of $\sigma_l(t)$ are

$$a_l(t) = 0, \quad (6.195)$$

$$e_l(t) = 0, \quad (6.196)$$

$$i_l(t) = I, \quad (6.197)$$

$$\Omega_l(t) = (-\cot i)Q, \quad (6.198)$$

$$\omega_l(t) = (\csc i)Q, \quad (6.199)$$

$$M_l(t) = 0, \quad (6.200)$$

with

$$I = [(\Delta\theta) \cos \Omega - (\Delta\mu \sin \theta) \sin \Omega], \quad (6.201)$$

$$Q = [(\Delta\mu \sin \theta) \cos \Omega + (\Delta\theta) \sin \Omega]. \quad (6.202)$$

The orbital elements in the above formulas are the quasi-mean orbit elements $\tilde{\sigma}(t)$.

6.3.3.3 Effects of Mars's Two Natural Satellites on the Orbit of a Mars's Spacecraft

The perturbation magnitude of Mars's two natural satellites on a low Mars orbit spacecraft given in Sect. 6.3.2 is.

$$\varepsilon_6(\text{Moon}) = O(10^{-9})$$

It is small and usually not important, but for a high Mars orbit spacecraft the effect increases, similar to the Moon's gravity force acting on a high Earth orbit satellite. To analyze the effects of the two Mars's satellites, we need the orbital information about the two satellites.

(1) The states of the two Mars's satellites

In the Mars-centric equatorial frame, the three main orbital elements of the two moons (Phobos and Deimos) are

$$\begin{aligned} a &= 2.7604(a_e), \quad e = 0.0151, \quad i = 1^\circ.082, \\ a &= 6.9070(a_e), \quad e = 0.00033, \quad i = 1^\circ.791. \end{aligned}$$

The masses of Phobos and Deimos relative to Mars's mass are $m_1 = 1.68 \times 10^{-8}$ and $m_2 = 2.80 \times 10^{-9}$, respectively, and a_e is the radius of Mars's equator. From these values, we can simply estimate the influence of the two moons to a particular spacecraft according to the height of an orbit.

(2) The perturbation solution due to Phobos or Deimos

In this book, we mainly consider the low Mars orbit spacecraft in the altitude range of 200–1000 km. The magnitude of the perturbation from Phobos or Deimos can be in the order of $O(10^{-8})$. If we omit terms of $O(10^{-9})$ the perturbation function is given by

$$\begin{cases} R = m' \left[\frac{r^2}{r^3} \left(\frac{3}{2} \cos^2 \psi - \frac{1}{2} \right) + \frac{r^3}{r^4} \left(\frac{5}{2} \cos^2 \psi - \frac{3}{2} \cos \psi \right) \right], \\ \cos \psi = \left(1 - \sin^2 \frac{i}{2} \right) \cos(f + \omega + \Omega - f' - \omega' - \Omega') \\ \quad + \sin^2 \frac{i}{2} \cos(f + \omega - \Omega - f' - \omega' - \Omega'). \end{cases} \quad (6.203)$$

The variables with an apostrophe (including dimensionless variables m' and orbit elements $a', e', i', \dots$) are for the third body (i.e., Phobos or Deimos). Because of the small eccentricities of the two moons, the corresponding value of r' can be replaced by the semi-major axis, the difference is a small value of $O(10^{-10})$. By this treatment, the perturbation function has only a secular part and a short-period part, that

$$R_c = m' \frac{a^2}{a'^3} \left(1 + \frac{3}{2}e^2\right) \left[\frac{3}{4} \left(1 - \sin^2 \frac{i}{2}\right)^2 + \frac{3}{4} \sin^4 \frac{i}{2} - \frac{1}{2} \right], \quad (6.204)$$

$$R_s = R - R_c. \quad (6.205)$$

The corresponding perturbation solution for a low Mars orbit spacecraft is as follows.

① σ_2

$$a_2 = 0, \quad e_2 = 0, \quad i_2 = 0, \quad (6.206)$$

$$\Omega_2 = -\frac{3m'}{4\sqrt{1-e^2}} \frac{a^3}{a'^3} n \left(1 + \frac{3}{2}e^2\right) \cos i, \quad (6.207)$$

$$\begin{aligned} \omega_2 = & \frac{3m'}{4\sqrt{1-e^2}} \frac{a^3}{a'^3} n \left(1 + \frac{3}{2}e^2\right) \cos^2 i \\ & + 3m' \frac{a^3}{a'^3} \sqrt{1-e^2} n \left[\frac{3}{4} \left(1 - \sin^2 \frac{i}{2}\right)^2 + \frac{3}{4} \sin^4 \frac{i}{2} - \frac{1}{2} \right], \end{aligned} \quad (6.208)$$

$$\begin{aligned} \lambda_2 = & M_2 + \omega_2 \\ = & \frac{3m'}{4\sqrt{1-e^2}} \frac{a^3}{a'^3} n \left(1 + \frac{3}{2}e^2\right) \cos^2 i \\ & + 3m' \frac{a^3}{a'^3} \sqrt{1-e^2} \left(1 - \sqrt{1-e^2}\right) n \left[\frac{3}{4} \left(1 - \sin^2 \frac{i}{2}\right)^2 + \frac{3}{4} \sin^4 \frac{i}{2} - \frac{1}{2} \right] \\ & - 4m' \frac{a^3}{a'^3} n \left(1 + \frac{3}{2}e^2\right) \left[\frac{3}{4} \left(1 - \sin^2 \frac{i}{2}\right)^2 + \frac{3}{4} \sin^4 \frac{i}{2} - \frac{1}{2} \right]. \end{aligned} \quad (6.209)$$

② $a_s(t)$

$$a_s(t) = \frac{2}{n^2 a} (R - R_c). \quad (6.210)$$

For a high Mars orbit spacecraft, the problem is more complicated. In some cases, the orbital information of Phobos or Deimos cannot be simplified, like the Moon's perturbation cannot be ignored for a high Earth satellite. In this case, high precision

orbital information of the two moons is needed, details are provided in reference [9], and we do not further discuss this problem.

6.4 Perturbed Orbital Solution of Venus's Satellite

Major planets rotate fast except Venus. The Moon and Venus belong to a type of celestial bodies which rotate slowly, and Venus rotates even slower than the Moon, so its shape is more close to a sphere. In Sect. 6.1.4, Table 6.5 lists the dynamical harmonic coefficients, showing that the dynamical form-factor coefficient J_2 is not only small but also similar to other “high” order harmonic coefficients (including the tesseral harmonic terms). Therefore, the perturbation of Venus's non-spherical gravity on a low Venus orbit satellite is different from either an Earth's satellite or a Moon's satellite. We now use a spacecraft moving around Venus as an example to build the perturbation solution of its orbit, and to provide the results.

Like Earth, Venus's rotation axis also oscillates. Recent theoretical research results show that if Venus's precession and nutation had the same magnitudes as these of Earth, their influences on a spacecraft can be ignored because the non-spherical part of Venus is relatively small. Without precession and nutation, the transformation relationship between the position vector of a spacecraft in the Venus-fixed frame $\vec{R}$ and in the Venus-centric mean equatorial frame $\vec{r}$ only needs to include Venus's rotation, that

$$\vec{R} = (ER)\vec{r} = R_Z(S_G)\vec{r}, \quad (6.211)$$

where S_G is the sidereal time on Venus (the reflection of rotation). From the IAU 2000 Venus's orientation model (Fig. 1.6 or in Appendix 3), the rotation angle of Venus is

$$W(t) = 160^\circ.20 - 1^\circ.4813688d. \quad (6.212)$$

Note that the rotation direction of Venus is opposite to that of Earth. Then we have Venus's sidereal time, which is needed in constructing the perturbed solution of a Venus's prober, given by

$$S_G = S_0 + \dot{S}_G(t - t_0) = 160^\circ.20 - 1^\circ.4813688d, \quad (6.213)$$

where d is the Julian day number starting from J2000.0. The rotating angular speed of Venus is

$$n_e = -1^\circ.4813688/d. \quad (6.214)$$

6.4.1 The Perturbation Function of Venus's Non-Spherical Gravity Potential

In the Venus-fixed frame, Venus's non-spherical gravity potential (i.e., the perturbation function) is given by

$$R = \Delta V = \frac{\mu}{r} \sum_{l \geq 2} \sum_{m=0}^l \left(\frac{a_e}{r}\right)^l P_{l,m}(\sin \varphi) [C_{l,m} \cos m\lambda + S_{l,m} \sin m\lambda], \quad (6.215)$$

which can be further expressed as

$$\begin{aligned} R = & \frac{\mu}{a} \sum_{l=2}^{\infty} \sum_{m=0}^l \left(\frac{a_e}{a}\right)^l \sum_{p=0}^l F_{lmp}(i) \left(\frac{a}{r}\right)^{l+1} \{[(1 - \delta_m)C_{l,m} - \delta_m S_{l,m}] \cos((l - 2p)u + m\bar{\theta}) \\ & + [(1 - \delta_m)S_{l,m} + \delta_m C_{l,m}] \sin((l - 2p)u + m\bar{\theta})\}, \\ \bar{\theta} = & \Omega - S_G, \end{aligned} \quad (6.216)$$

where $\mu = GM$ is Venus's gravitational constant. Similarly, if we use dimensionless units, then $\mu = GM = 1$. The symbol δ_m and inclination function $F_{lmp}(i)$ are introduced in (4.382) and (4.391)–(4.393), that

$$\delta_{l,m} = \frac{1}{2} [1 - (-1)^{l-m}] = \begin{cases} 1, & (l - m) \text{ is odd,} \\ 0, & (l - m) \text{ is even.} \end{cases} \quad (6.217)$$

$$\begin{aligned} F_{lmp}(i) = & \frac{(l+m)!}{2^l p! (l-p)!} \sum_{k=k_1}^{k_2} (-1)^{k+(l-m+\delta_{l,m})/2} \binom{2l-2p}{k} \binom{2p}{l-m-k} \\ & \times \left(\sin \frac{i}{2}\right)^{-(l-m-2p-2k)} \left(\cos \frac{i}{2}\right)^{(3l-m-2p-2k)} \\ = & \frac{(l+m)!}{2^{2l} p! (l-p)!} \sum_{k=k_1}^{k_2} (-1)^{k+(l-m+\delta_{l,m})/2} \binom{2l-2p}{k} \binom{2p}{l-m-k} \\ & \times \left(\sin \frac{i}{2}\right)^{-(l-m-2p-2k)} (1 + \cos i)^{(2l-m-2p-2k)}, \end{aligned} \quad (6.218)$$

$$k_1 = \max(0, l - m - 2p), \quad k_2 = \min(l - m, 2l - 2p), \quad (6.219)$$

$$\begin{aligned} F'_{lmp}(i) = & \frac{d}{di} F_{lmp}(i) \\ = & \frac{(l+m)!}{2^l p! (l-p)!} \left(\frac{1}{\sin i}\right) \sum_{k=k_1}^{k_2} (-1)^{k+(l-m+\delta_{l,m})/2} \binom{2l-2p}{k} \binom{2p}{l-m-k} \end{aligned}$$

$$\times \left[-2l \sin^2 \frac{i}{2} - (l - m - 2p - 2k) \right] \left(\sin \frac{i}{2} \right)^{-(l-m-2p-2k)} \left(\cos \frac{i}{2} \right)^{(3l-m-2p-2k)}. \quad (6.220)$$

Similar to the Moon, because of slow rotation the terms related to time (the tesseral terms) do not need to be expanded into the trigonometric functions of the mean anomaly M to separate the perturbation function R . Using the averages of the following terms

$$\left(\frac{a}{r} \right)^{l+1} \cos(l - 2p)f, \quad \left(\frac{a}{r} \right)^{l+1} \sin(l - 2p)f,$$

then

$$S_p(e) = \overline{(a/r)^{l+1} \sin(|l - 2p|)f} = 0, \quad (6.221)$$

$$\begin{aligned} C_p(e) &= \overline{(a/r)^{l+1} \cos(|l - 2p|)f} \\ &= \delta_p \left\{ (1 - e^2)^{-(l-1/2)} \sum_{k(2)=|l-2p|}^{(l-2)} \binom{l-1}{k} \binom{k}{(k-|l-2p|)/2} \left(\frac{e}{2} \right)^k \right\} \end{aligned} \quad (6.222)$$

$$\delta_p = \begin{cases} 0, & p = 0 \text{ or } l, \\ 1, & p \neq 0 \text{ and } l, \end{cases} \quad (6.223)$$

and the perturbation function R can be decomposed to

$$R = R_c + R_l + R_s. \quad (6.224)$$

In (6.221) the sub of the sum $k(2) = |l - 2p|$ means $k(2) = |l - 2p|, |l - 2p| + 2, \dots$.

The actual expressions of the three parts of R , R_c , R_l , and R_s are

$$R_c = \frac{\mu}{a} \sum_{l(2)=2}^{\infty} \left(\frac{\mu}{a} \right) \left(\frac{a_e}{a} \right)^l C_{l,0} [C_0(e) F_{l0l/2}(i)], \quad l(2) = 2, 4, \dots, \quad (6.225)$$

$$\left\{ \begin{aligned} R_l &= \frac{\mu}{a} \sum_{l=2}^{\infty} \sum_{m=0}^l \left(\frac{\mu}{a} \right) \left(\frac{a_e}{a} \right)^l \sum_{p=1}^{l-1} C_p(e) F_{lmp}(i) \\ &\quad \times [\delta C S \cos((l - 2p)\omega + m\bar{\theta}) + \delta S C \sin((l - 2p)\omega + m\bar{\theta})], \\ &\quad |l - 2p| + m \neq 0, \end{aligned} \right. \quad (6.226)$$

$$R_s = \frac{\mu}{a} \sum_{l=2}^{\infty} \sum_{m=0}^l \left(\frac{\mu}{a} \right) \left(\frac{a_e}{a} \right)^l$$

$$\begin{aligned}
& \times \left\{ F_{lm0}(i) \left(\frac{a}{r} \right)^{l+1} [\delta CS \cos(lu + m\bar{\theta}) + \delta SC \sin(lu + m\bar{\theta})] \right. \\
& + F_{lm1}(i) \left(\frac{a}{r} \right)^{l+1} [\delta CS \cos(-lu + m\bar{\theta}) + \delta SC \sin(-lu + m\bar{\theta})] \\
& + \sum_{p=1}^{l-1} F_{lmp}(i) \left[\delta CS \left(\left(\frac{a}{r} \right)^{l+1} - C_p(e) \right) \cos((l-2p)\omega + m\bar{\theta}) \right. \\
& \left. \left. + \delta SC \left(\left(\frac{a}{r} \right)^{l+1} - C_p(e) \right) \sin((l-2p)\omega + m\bar{\theta}) \right] \right\}, \quad (6.227)
\end{aligned}$$

where δCS and δSC are

$$\begin{cases} \delta CS = [(1 - \delta_m)C_{l,m} - \delta_m S_{l,m}], \\ \delta SC = [(1 - \delta_m)S_{l,m} + \delta_m C_{l,m}]. \end{cases} \quad (6.228)$$

In the expressions of R_c and R_s , $C_0(e)$ is $C_p(e)$ when $|l - 2p| = 0$; $F_{l0l/2}(i)$, $F_{lm0}(i)$, and $F_{lm1}(i)$ are the inclination function $F_{lmp}(i)$ corresponding to $m = 0$, $p = l/2$, and $p = l$, respectively.

6.4.2 The Structure and Results of the Analytical Perturbation Solution

By the characteristics of the perturbation function R and its components (R_c , R_l , and R_s), it is not difficult to build the analytical solution of the first order (i.e., the linear solution) of secular, long-period, and short-period terms [10]. This form of solution is easy for analyzing the properties of orbit variation, because it does not have to be expanded into trigonometric series of the mean anomaly which is restricted by the magnitude of the eccentricity of the spacecraft's orbit. Therefore, this solution is different from the Kaula solution used for geodetic satellites [11]. Kaula solution actually is an analytical solution of the first order constructed by the classical perturbation method, which expands the perturbation function into trigonometric series of the mean anomaly regardless of the particular dynamics.

Substituting R_c , R_l , and R_s into the perturbation Eqs. (3.64) yields

$$\frac{d\sigma}{dt} = f_0(a) + f_c(\sigma; C_{l,0}) + f_l(\sigma, t; C_{l,m}, S_{l,m}) + f_s(\sigma, t; C_{l,m}, S_{l,m}). \quad (6.229)$$

Based on the characteristics of Venus's non-spherical gravity field the harmonic terms in the perturbation function do not have to be separated by their magnitudes as for an Earth's satellite, and are all treated as small terms of the first-order relative to the central gravity force, that

$$f_0(a) = O(1), |f_c, f_l, f_s| = O(\varepsilon). \quad (6.230)$$

With the knowledge given in previous chapters, particularly about the slow rotating Moon, we can use exactly the same method to construct the perturbation solution without dividing the terms by the perturbation magnitude. There should be no difficulty using the perturbation equation system (6.229) to obtain the perturbation solution of the first order. We do not provide the actual results here. Readers can refer to the reference [10], in which we provide the entire expressions of the analytical perturbation solution, also we use high precision integrating software PKF7(8) [12] to exam the orbital extrapolation (for 100 circles) of the analytical perturbation solution for two Venusian probes, PVO (Pioneer Venus Orbiter) and Magellan (their eccentricities were 0.843 and 0.382, respectively), results show that the high precise solutions are not restricted by the magnitude of the orbital eccentricity.

References

1. Liu L, Wang JS (1998) An analytic solution of the orbital variation of Lunar satellites. *ATCA Astronomica Sinica* 39(1):81–102, and *China Astron Astrophys* 1998 22(2):328–351
2. Liu L, Wang X (2006) *Orbital dynamics of spacecraft for the moon's exploration*. National Defense Industry Press, Beijing
3. Liu L, Tang JS (2012) Orbital motions of circling spacecraft around the major planets, the Moon, and asteroids. *Spacecr Eng* 21(4):4–15
4. Liu L, Tang JS (2015) *Orbital theory of satellites and applications*. Electronic Industry Press, Beijing
5. Eckhardt DH (1981) Theory of the libration of the Moon. *Moon Planet* 25:3–49
6. Zhang W, Liu L (2005) Effects of the physical libration of the Moon on Lunar orbiters. *ATCA Astronomica Sinica* 46(2) 196–206, and *China Astron Astrophys* 29(4):438–448
7. Zhou CH, Yu SX, Liu L (2012) On the coupled perturbations of J_2 and tesseral harmonic terms of the Mars orbiters. *ATCA Astron Sinica* 53(3):205–212, and *China Astron Astrophys* 36(4):399–406
8. Liu L, Zhao YH, Zhang W et al (2010) The additional perturbation of coordinate system and the selection of related coordinate system for a Mars's probe. *ATCA Astronomica Sinica* 51(4):412–421
9. Yu SX, Liu L (2011) Construction of analytical ephemeris for Mars's natural satellites. *J Spacecr TT&C Technol* 30(4):60–65
10. Liu L, Shun CK (2000) Analytic perturbation solution to the Venusian orbiter due to the non-spherical gravitational potential. *Sci China (Series A)* 43(5):552–560
11. Kaula WM (1966) *Theory of satellite geodesy: Application of satellites to geodesy* (Chapter 3). Waltham Blaisdell Pub Co
12. Fehlberg E (1968) Classical fifth, sixth, seventh, and eighth-order Runge-Kutta with Stepsize Control. NASA TR R-287

Chapter 7

Orbital Motion and Calculation Method in the Restricted Three-Body Problem



The restricted problem is introduced in the Introduction of this book. This chapter is about the existent solutions and related calculation methods of orbital motion for a small body (i.e., an asteroid or a space explorer) in a **restricted three-body problem (especially the circular restricted three-body problem)** in the Solar System. The two main bodies, denoted by P_1 and P_2 , have masses m_1 and m_2 , respectively, and their motions are defined. The research subject is the motion of a third body P whose mass is denoted by m . This restricted three-body model is a necessarily and reasonably simplified mathematical model of a complicated dynamical system.

In the Solar System, all the major planets, except Mercury, move around the Sun in almost circular orbits, and the Moon's orbit around Earth is also almost a circle (Table 7.1). Because of the small eccentricities, their orbits are often treated as circles except for making high precision ephemerides. Mathematical models of the circular restricted three-body problem mentioned in the Introduction are based on this fact. The influence of the orbital eccentricity of the main body, e' , on the motion of a third body is an important field of theoretical research, but from the perspective of application, it is not necessary to study it in depth, the effect can be quantitatively adjusted.

7.1 Selection of Coordinate System and Motion Equation of a Small Body

In the research of motion of a small body in a restricted three-body problem, the motions of the two main bodies are defined, accordingly, there are three types of coordinate systems to choose based on the requirement of a specific aerospace project and a particular orbit of the small body. The three types of coordinate systems are:

Table 7.1 Orbital eccentricities of major planets and the Moon

Mercury	Venus	Earth	Mars	Jupiter	Saturn	Uranus	Neptune	Moon
0.2056	0.0068	0.0167	0.0934	0.0485	0.0555	0.0463	0.0090	0.0549

- (1) The barycenter coordinate system of one of the main bodies, P_i ($i = 1$ or 2), the origin of the frame is at the barycenter of P_i .
- (2) The barycenter inertial coordinate system, the origin of the frame is at the common barycenter of the two main bodies.
- (3) The barycenter rotational coordinate system, the origin of the frame is at the common barycenter of the two main bodies. This system is also called the synodic coordinate system.

Note that the inertial coordinate system here is only used for the restricted three-body problem and has no universal meaning. The three bodies, P_1 , P_2 , and P are all treated as particles. The fundamental plane (xy -plane) and the primary direction are defined according to different celestial systems and different projects.

The barycenter frame of a main body has two types. The first is the barycenter celestial coordinate system, its fundamental plane (the xy -plane) is this body's mean equatorial plane at the epoch time (presently J2000.0), and the direction of the x -axis towards the mean "March equinox", which is introduced in Chap. 1 and the IAU orientation model is illustrated in Fig. 1.6. The second is the main-body-fixed coordinate system, its fundamental plane is the instantaneous true equatorial plane, and the x -axis is towards the prime meridian, for Earth it is the Greenwich meridian.

The first type of coordinate system is often selected for a small body P moving around and near one of the main bodies (such as natural satellites and circling spacecraft). It is also used for a particular project such as a Moon prober. After launch the Moon prober enters into its first orbit, from this time to the time it changes its orbit to an intermediate orbit, its motion is described in the first type of coordinate system. When a small body moves between the two main bodies (like some asteroids or a spacecraft in the process of transferring from one orbit to another) for orbital analysis and designs its motion is presented by the second type or the third type frame, particularly the third type of synodic coordinate system. But for any aerospace project, eventually, we need to know the actual motion of the spacecraft P with respect to the main body (such as Earth), i.e., to use the first type of coordinate system (usually the geocentric celestial coordinate system).

For the convenience of dynamical analysis and calculation, we usually use the normalized system, as we deal with Earth's artificial satellite, then the physical variables are dimensionless, and their magnitudes are normalized. In this chapter, all physical quantities are dimensionless.

Specifically, when we use the first type coordinate system for a small body moving around one of the main body P_i , the units of mass $[M]$, length $[L]$, and time $[T]$ are defined by

$$\begin{cases} [M] = m_i \ (i = 1 \text{ or } 2), \\ [L] = a_e \ (\text{equatorial radius of } P_i), \\ [T] = \left(\frac{a_e^3}{Gm_i} \right)^{\frac{1}{2}}. \end{cases} \quad (7.1)$$

The gravitational constant of this system is $G = 1$.

The second type of coordinate system is used for a small body moving between the two main bodies, and the motion scale is different from the first type of coordinate system. By convention, the corresponding units are defined by

$$\begin{cases} [M] = m_1 + m_2, \\ [L] = a_{12}, \\ [T] = \left[\frac{a_{12}^3}{G(m_1 + m_2)} \right]^{\frac{1}{2}}. \end{cases} \quad (7.2)$$

Again, the gravitational constant of this system $G = 1$. The unit of length a_{12} is the distance between the two main bodies. By this unit system, the masses of the two main bodies are

$$1 - \mu = \frac{m_1}{m_1 + m_2}, \quad \mu = \frac{m_2}{m_1 + m_2}, \quad (7.3)$$

and the distances between the common barycenter and the two main bodies are

$$r'_1 = \mu, \quad r'_2 = 1 - \mu. \quad (7.4)$$

In a circular restricted three-body problem, the distance between the two main bodies a_{12} is a constant, and the time unit $[T]$ is the reciprocal of the angular speed of the two main bodies, that $[T] = 1/n$.

7.1.1 The Motion Equation of a Small Body in the Barycenter Inertial Coordinate System

Denoting the barycenter inertial coordinate system by C -XYZ, the origin of the frame is at the common barycenter of the two main bodies denoted by C ; the XY -plane is the motion plane of the two main bodies; the X -axis is the line connecting the two main bodies and the direction of the axis towards the smaller main body P_2 at the initial time $t = t_0$ (Fig. 7.1).

In this frame, the coordinate vectors of the small body, and the two main bodies are denoted by $\vec{R}$, $\vec{R}'_1$, and $\vec{R}'_2$, respectively, then the coordinate vectors of the small body relative to the two main bodies are given by

$$\vec{R}'_1 = \vec{R} - \vec{R}'_1, \quad \vec{R}'_2 = \vec{R} - \vec{R}'_2. \quad (7.5)$$

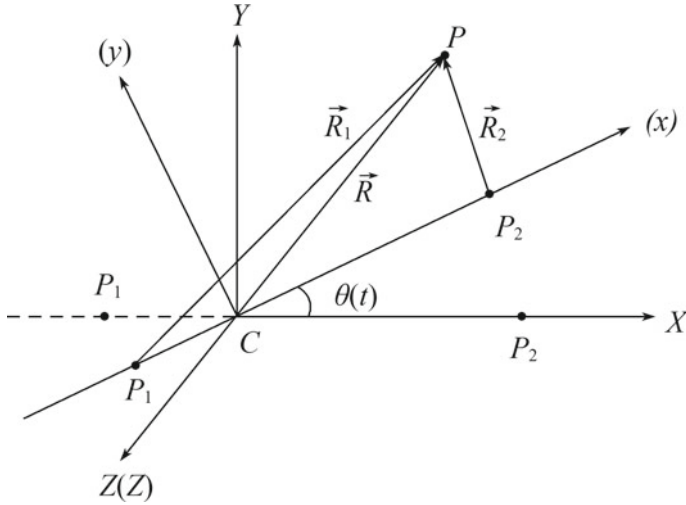


Fig. 7.1 The barycenter inertial frame $C\text{-}XYZ$ and the barycenter rotational frame $C\text{-}xyz$

The geometric relationships of these vectors are illustrated in Fig. 7.1 (for clarity $\vec{R}'_1, \vec{R}'_2$ are not shown in the figure). In the circular restricted three-body problem, the orbits of the two main bodies around the common barycenter are two circles, their coordinate vectors vary over time as

$$\begin{cases} \vec{R}'_1 = \begin{pmatrix} -\mu \cos t \\ -\mu \sin t \\ 0 \end{pmatrix}, \\ \vec{R}'_2 = \begin{pmatrix} (1-\mu) \cos t \\ (1-\mu) \sin t \\ 0 \end{pmatrix}. \end{cases} \quad (7.6)$$

One of the characteristics of the circular restricted problem is that $[T] = 1/n$, therefore

$$\theta(t) = nt^* = nt[T] = n\left(\frac{t}{n}\right) = t, \quad (7.7)$$

where t^* is the original time with demission, i.e., $t^* = t[T]$, and t is dimensionless, which means in the normalized dimensionless system, the angular speed of the two main bodies is $\dot{\theta}(t) = 1$.

In this frame and the selected calculation units, the motion equation of the small body is given by

$$\ddot{\vec{R}} = \left(\frac{\partial U}{\partial \vec{R}} \right)^T = -(1-\mu) \frac{\vec{R}_1}{R_1^3} - \mu \frac{\vec{R}_2}{R_2^3}, \quad (7.8)$$

where U is given by

$$U = U(R_1, R_2) = \frac{1-\mu}{R_1} + \frac{\mu}{R_2}, \quad (7.9)$$

and

$$\begin{cases} R_1 = |\vec{R} - \vec{R}'_1| = [(X + \mu \cos t)^2 + (Y + \mu \sin t)^2 + Z^2]^{\frac{1}{2}}, \\ R_2 = |\vec{R} - \vec{R}'_2| = \{[X - (1 - \mu)\sin t]^2 + [Y - (1 - \mu)\cos t]^2 + Z^2\}^{\frac{1}{2}}. \end{cases} \quad (7.10)$$

7.1.2 The Motion Equation of a Small Body in the Synodic Coordinate System

Denoting the synodic coordinate system (i.e., the barycenter rotational coordinate system) by C -xyz, the angular speed of the frame is the relative angular speed of the two main bodies $\dot{\theta}(t)$, and the two main bodies are fixed on the x -axis (Fig. 7.1). In this frame, the coordinate vectors of a small body and the two main bodies are denoted by $\vec{r}$, $\vec{r}'_1$, and $\vec{r}'_2$, respectively, then the coordinate vectors of the small body relative to the two main bodies are

$$\vec{r}'_1 = \vec{r} - \vec{r}'_1, \quad \vec{r}'_2 = \vec{r} - \vec{r}'_2, \quad (7.11)$$

where

$$\vec{r}'_2 = \begin{pmatrix} -\mu \\ 0 \\ 0 \end{pmatrix}, \quad \vec{r}'_2 = \begin{pmatrix} 1 - \mu \\ 0 \\ 0 \end{pmatrix}. \quad (7.12)$$

We have

$$\begin{cases} r_1 = [(x + \mu)^2 + y^2 + z^2]^{\frac{1}{2}} = R_1 \\ r_2 = [(x - 1 + \mu)^2 + y^2 + z^2]^{\frac{1}{2}} = R_2 \end{cases}. \quad (7.13)$$

The conversion relationships of $\vec{r}$ and $\vec{R}$ are

$$\vec{r} = R_z(t)\vec{R} = \begin{pmatrix} X \cos t + Y \sin t \\ -X \sin t + Y \cos t \\ Z \end{pmatrix}, \quad (7.14)$$

$$\vec{R} = R_z(-t)\vec{r} = \begin{pmatrix} x \cos t - y \sin t \\ x \sin t + y \cos t \\ z \end{pmatrix}, \quad (7.15)$$

$R_z(t)$ and $R_z(-t)$ are rotation matrices. Definitions of matrices $R_x(\theta)$, $R_y(\theta)$, and $R_z(\theta)$ are given in (1.20)–(1.22) of Sect. 1.2, by which we have

$$\begin{aligned} R_z(t) &= \begin{pmatrix} \cos t & \sin t & 0 \\ -\sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\ R_z(-t) &= R_z^T(t) = \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned} \quad (7.16)$$

From (7.15), we derive

$$\begin{cases} \dot{\vec{R}} = \dot{R}_z(-t)\vec{r} + R_z(-t)\dot{\vec{r}}, \\ \ddot{\vec{R}} = \ddot{R}_z(-t)\vec{r} + 2\dot{R}_z(-t)\dot{\vec{r}} + R_z(-t)\ddot{\vec{r}}, \end{cases} \quad (7.17)$$

where

$$\begin{cases} \dot{R}_z(-t) = \begin{pmatrix} -\sin t & -\cos t & 0 \\ \cos t & -\sin t & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\ \ddot{R}_z(-t) = \begin{pmatrix} -\cos t & \sin t & 0 \\ -\sin t & -\cos t & 0 \\ 0 & 0 & 0 \end{pmatrix}. \end{cases} \quad (7.18)$$

By the above conversion relationships and the property of the rotation matrix

$$R_z^{-1}(t) = R_z^T(t) = R_z(-t), \quad (7.19)$$

we can transfer the motion Eq. (7.8) in the barycenter inertial frame to the synodic frame, that

$$\ddot{\vec{r}} + 2 \begin{pmatrix} -\dot{y} \\ \dot{x} \\ 0 \end{pmatrix} = \left(\frac{\partial \Omega}{\partial \vec{r}} \right)^T, \quad (7.20)$$

where

$$\Omega = \frac{1}{2}(x^2 + y^2) + U(r_1, r_2), \quad (7.21)$$

$$U(r_1, r_2) = \frac{1-\mu}{r_1} + \frac{\mu}{r_2}. \quad (7.22)$$

In the restricted problem, it is often necessary to express Ω in the following format [1]

$$\Omega = \frac{1}{2}[(x^2 + y^2) + \mu(1 - \mu)] + \frac{1-\mu}{r_1} + \frac{\mu}{r_2}, \quad (7.23)$$

which can be rearranged in a “symmetric” form as

$$\begin{aligned} \Omega &= \left\{ \frac{1}{2}[(x^2 + y^2 + z^2) + \mu(1 - \mu)] + \frac{1-\mu}{r_1} + \frac{\mu}{r_2} \right\} - \frac{1}{2}z^2 \\ &= \left\{ \frac{1}{2}[(1 - \mu)r_1^2 + \mu r_2^2] + \frac{1-\mu}{r_1} + \frac{\mu}{r_2} \right\} - \frac{1}{2}z^2. \end{aligned} \quad (7.24)$$

In order to be consistent with conventional expression in the text below Ω takes the form (7.23), unless stated otherwise.

7.2 Jacobi Integral and Solution Existence of the Circular Restricted Three-Body Problem

7.2.1 Jacobi Integral in the Circular Restricted Three-Body Problem

There is the main difference between the function Ω in (7.20) and the function U in (7.8), that U is a function of time t but $\Omega = \Omega(x, y, z)$ does not explicitly contain t . Therefore from (7.20) we have

$$\dot{x}\ddot{x} + \dot{y}\ddot{y} + \dot{z}\ddot{z} = \frac{\partial \Omega}{\partial x}\dot{x} + \frac{\partial \Omega}{\partial y}\dot{y} + \frac{\partial \Omega}{\partial z}\dot{z}$$

i.e.,

$$\begin{cases} \frac{1}{2} \frac{d}{dt}(v^2) = \frac{d\Omega}{dt}, \\ v^2 = \dot{x}^2 + \dot{y}^2 + \dot{z}^2, \end{cases}$$

which gives an integral

$$2\Omega - v^2 = C. \quad (7.25)$$

This integral is called the Jacobi integral in the synodic coordinate system, and up to the present day, it is the only integral we know in the circular restricted three-body problem.

In the barycenter inertial coordinates, because U is a function of t explicitly, the integral cannot be given directly. But it is the same circular restricted three-body problem, there must exist an integral. An integral equivalent to the Jacobi integral can be obtained by the transformation of the two coordinate systems. The quantities need to be transformed are

$$r_1, r_2; \quad x^2 + y^2; \quad v^2 = \dot{x}^2 + \dot{y}^2 + \dot{z}^2.$$

Omitting the process of transformation, the resulted Jacobi integral in the barycenter inertial coordinate system has the form as

$$\begin{cases} 2U - [V^2 + 2(\dot{X}Y - X\dot{Y})] = C - \mu(1 - \mu), \\ U = \frac{1 - \mu}{R_1} + \frac{\mu}{R_2}, \quad V^2 = \dot{X}^2 + \dot{Y}^2 + \dot{Z}^2. \end{cases} \quad (7.26)$$

Readers should pay attention to the form of Ω . If using the original form of Ω in (7.21) then on the right side of (7.26) the term $\mu(1 - \mu)$ does not appear.

7.2.2 Existence of Solution of the Circular Restricted Three-Body Problem

When we study the perturbed two-body problem, the reference model is the two-body problem, which has six integrals. Similarly, when we study the perturbed circular restricted three-body problem, the reference model is the circular restricted three-body problem, which has only one integral, the Jacobi integral. Therefore, it is necessary to know whether the circular restricted three-body problem has any solution. The existence of a solution is the premise of actually constructing a motion solution for a small body. This problem is beyond the scope of this book. Here we only provide a simple summary of the results. The complete research work is presented in Chap. 3 of reference [1].

In order to discuss the existence of a solution, we must analyze the motion equation first. The basic Eqs. (7.20)–(7.22) show that there are two singularities, $r_1 = 0$ and $r_2 = 0$. These two singularities are collision singularities, meaning the small body collides on one of the main bodies. Because in the restricted three-body problem (no matter circular or elliptical) the relative motions of the two main bodies are two defined circles (or ellipses), the collision can only be binary. It should be pointed out that here a collision is only an expression from a mathematical point of view. It means that two volume-less particles collide but two real bodies physically collide.

Knowing that there are singularities, also they are singularities of binary collision, we must consider the characteristics of the singularities, whether or not the singularities can be eliminated; if they can be eliminated, what kind of motion state is before

and after a collision. These are the main content related to the existence of the solution in Chap. 3 of reference [1]. The key method to eliminate the binary collision is the transformation of regularization, which includes the independent variable transformation and the related function transformation. The former makes the differential motion equation normalized (i.e., eliminating the two collision singularities in the motion equation), and the latter eliminates the singularities in the solution and makes the speed of the collision limited. These two transformations are related, meaning the selection of the independent variable is connected to the motion quantities (i.e., the state quantities). The final conclusion is that solutions of the circular restricted three-body problem exist in the time range of $(-\infty, \infty)$.

About the transformations of regularization used in the process of proving the above conclusion besides reference [1], there is a monograph written by Stiefel and Scheifele [2]. In this monograph, the independent variable transformation is $t \rightarrow \tau$ by

$$\frac{dt}{d\tau} = g(w), \quad (7.27)$$

which implies changing the time scale. For example, a simple independent variable transformation that can eliminate one collision singularity is given by

$$\frac{dt}{d\tau} = r^p, \quad p \geq 1. \quad (7.28)$$

This transformation can be applied to extrapolation calculation for higher eccentricity satellite orbits by forming the technology of automatically adjusting the integration step size. Details about this method are given in Sect. 8.4.3.

7.3 Calculation and Application of the Libration Point Positions of the Circular Restricted Three-Body Problem

For the circular restricted three-body problem, as mentioned above only one integral, the Jacobi integral, was found, but there is no question that this problem has solutions, and there are ways to find particular solutions which provide the regular patterns of a small body's motion. To the present day, five particular libration solutions are found. The systematic interpretation of these solutions is given in Chap. 4 of reference [1] and references [3, 4]. In the text below we introduce the actual forms of these particular solutions and related calculation methods for their applications.

7.3.1 Conditions of Existence for Libration Solutions

To discuss this problem, it is relatively simple using the synodic coordinate system. The basic motion equation is given by (7.20). The libration solutions are particular solutions satisfying the following conditions

$$x(t) \equiv x_0, \quad y(t) \equiv y_0, \quad z(t) \equiv z_0, \quad (7.29)$$

where x_0 , y_0 , and z_0 are given by initial conditions. From (7.29) we have

$$\dot{x} = 0, \quad \dot{y} = 0, \quad \dot{z} = 0, \quad (7.30)$$

$$\ddot{x} = 0, \quad \ddot{y} = 0, \quad \ddot{z} = 0. \quad (7.31)$$

The points given by (7.29) are the equilibrium points in the synodic coordinate system, which are called libration points, or Lagrange points in the field of Orbital Dynamics. From Eq. (7.20) we know that these libration points must satisfy

$$\Omega_x = 0, \quad \Omega_y = 0, \quad \Omega_z = 0, \quad (7.32)$$

where Ω_x , Ω_y , and Ω_z are partial derivatives of $\Omega(x, y, z)$ with respect to x , y , and z , respectively. The actual expressions of (7.32) are

$$\begin{cases} \Omega_x = x - \frac{(1-\mu)(x+\mu)}{r_1^3} - \frac{\mu(x-1+\mu)}{r_2^3} = 0, \\ \Omega_y = y \left(1 - \frac{1-\mu}{r_1^3} - \frac{\mu}{r_2^3} \right) = 0, \\ \Omega_z = -z \left(\frac{1-\mu}{r_1^3} + \frac{\mu}{r_2^3} \right) = 0. \end{cases} \quad (7.33)$$

Because

$$\frac{1-\mu}{r_1^3} + \frac{\mu}{r_2^3} \neq 0,$$

from Ω_z in (7.33), there is

$$z = z_0 = 0, \quad (7.34)$$

which means the libration points must lie in the xy -plane. With $z = 0$, there are two possible cases satisfy (7.33), that

$$y = 0, \begin{cases} x - \frac{1-\mu}{(x+\mu)^2} + \frac{\mu}{(x-1+\mu)^2} = 0, & \text{for } L_1, \\ x - \frac{1-\mu}{(x+\mu)^2} - \frac{\mu}{(x-1+\mu)^2} = 0, & \text{for } L_2, \\ x + \frac{1-\mu}{(x+\mu)^2} + \frac{\mu}{(x-1+\mu)^2} = 0, & \text{for } L_3, \end{cases} \quad (7.35)$$

$$y \neq 0, \begin{cases} 1 - \frac{1-\mu}{r_1^3} - \frac{\mu}{r_2^3} = 0, \\ x - \frac{(1-\mu)(x+\mu)}{r_1^3} - \frac{\mu(x-1+\mu)}{r_2^3} = 0. \end{cases} \quad (7.36)$$

7.3.2 The Positions of the Three Collinear Libration Points

In the first case of $y = 0$, Eq. (7.35) has three real-number solutions, $x_1(\mu)$, $x_2(\mu)$, and $x_3(\mu)$ corresponding to the three libration points, L_1 , L_2 , and L_3 , which are called collinear libration points (Fig. 7.2).

In Fig. 7.2, $\xi^{(1)}$ and $\xi^{(2)}$ are the distances between the main body P_2 and the libration points, L_1 and L_2 , respectively, and $\xi^{(3)}$ is the distance between the main body P_1 and the libration point L_3 . These distances are given by three power series of μ , which are

$$\xi^{(1)} = \left(\frac{\mu}{3}\right)^{\frac{1}{3}} \left[1 - \frac{1}{3} \left(\frac{\mu}{3}\right)^{\frac{1}{3}} - \frac{1}{9} \left(\frac{\mu}{3}\right)^{\frac{2}{3}} - \dots \right], \quad (7.37)$$

$$\xi^{(2)} = \left(\frac{\mu}{3}\right)^{\frac{1}{3}} \left[1 + \frac{1}{3} \left(\frac{\mu}{3}\right)^{\frac{1}{3}} - \frac{1}{9} \left(\frac{\mu}{3}\right)^{\frac{2}{3}} + \dots \right], \quad (7.38)$$

$$\begin{cases} \xi^{(3)} = 1 - v \left[1 + \frac{23}{84}v^2 + \frac{23}{84}v^3 + \frac{761}{2352}v^4 + \frac{3163}{7056}v^5 + \frac{30703}{49392}v^6 \right] + O(v^8), \\ v = \frac{7}{12}\mu. \end{cases} \quad (7.39)$$

The corresponding solutions of the three collinear libration points $x_i(\mu)$ are

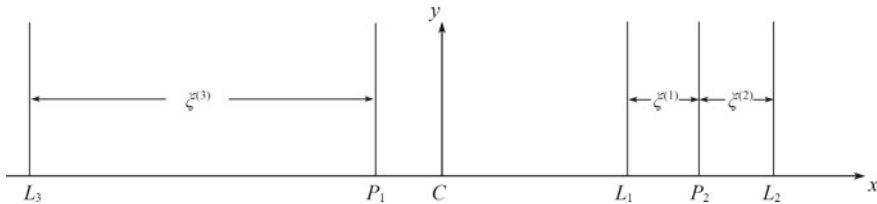


Fig. 7.2 Relative positions of the three collinear libration points and the two main bodies

$$x_1(\mu) = (1 - \mu) - \xi^{(1)}, \quad (7.40)$$

$$x_2(\mu) = (1 - \mu) - \xi^{(2)}, \quad (7.41)$$

$$x_3(\mu) = -(\mu + \xi^{(3)}). \quad (7.42)$$

Note that the order of the arrangement of L_1 and L_2 in Fig. 7.2 is different from the conventional order in reference [1], which is given according to the positions of the libration points on the x -axis. Here the order is given according to the related quantities of energy, which is accepted by the aerospace industry and other fields of application.

When $\mu = 0$, there are

$$\begin{cases} x_1(\mu) = x_2(\mu) = 1, \\ x_3(\mu) = -1, \end{cases} \quad (7.43)$$

and when $\mu = 1/2$, there are

$$\begin{cases} x_1(\mu) = 0, \\ x_2(\mu) = 1.198406, \\ x_3(\mu) = -x_2(\mu). \end{cases} \quad (7.44)$$

The positions of the three collinear libration points, $x_i(\mu)$, and the positions of the two main bodies, $x(P_1)$ and $x(P_2)$ on the x -axis depend on the value of μ . The pattern of the variation is obvious.

The positions of the three collinear libration points, $x_i(\mu)$, are calculated by a simple iteration. The series of expressions of $\xi^{(i)}$ in (7.37)–(7.39) are only used as auxiliaries. Starting with the first-order values of $\xi^{(i)}$ which are $(\mu/3)^{1/3}$, $(\mu/3)^{1/3}$, and $(1 - \nu)$, by (7.40)–(7.42) we get the first approximate of $x_i(\mu)$, then by iterating on (7.35) to get the next $x_i(\mu)$ until the accuracy reaches the requirement.

Table 7.2 lists the positions of the three collinear libration points $x_i(\mu)$ in the restricted three-body problem for the Sun-planet system and the Earth-Moon system; Table 7.3 lists the corresponding Jacobi constants $C_i(\mu)$. Note that in calculating the Jacobi constant the expression of Ω in (7.25) is given by (7.23).

In the above two tables, the values of the basic parameter μ for the major planets and the Moon may be slightly different from the current values, which have no fundamental influence on the results. If needed, the calculations can always be done using new values of parameter μ and by the same method described above.

Table 7.2 Positions of the three collinear libration points $x_i(\mu)$

	μ	x_1	x_2	x_3
Sun-Mercury	0.000000166	0.996193956	1.003815393	-1.000000069
Sun-Venus	0.000002448	0.990682298	1.009371018	-1.000001020
Sun-(Earth+Moon)	0.000003040	0.989985982	1.010075201	-1.000001267
Sun-Mars	0.000000323	0.995251330	1.004763104	-1.000000134
Sun-Jupiter	0.000953875	0.932365587	1.068830521	-1.000397448
Sun-Saturn	0.000285755	0.954747665	1.046070895	-1.000119065
Sun-Uranus	0.000043725	0.975729492	1.024580811	-1.000018219
Sun-Neptune	0.000051773	0.974330318	1.026011304	-1.000021572
Earth-Moon	0.012150568	0.836915214	1.155682096	-1.005062638

Table 7.3 Jacobi constants $C_i(\mu)$ corresponding to the three collinear libration points

	μ	C_1	C_2	C_3
Sun-Mercury	0.000000166	3.000130307	3.000130086	3.000000332
Sun-Venus	0.000002448	3.000780164	3.000776900	3.000004896
Sun-(Earth+Moon)	0.000003040	3.000900982	3.000896928	3.000006081
Sun-Mars	0.000000323	3.000202815	3.000202384	3.000000645
Sun-Jupiter	0.000953875	3.039713802	3.038441715	3.001906822
Sun-Saturn	0.000285755	3.018107577	3.017726518	3.000571427
Sun-Uranus	0.000043725	3.005268402	3.005210099	3.000087449
Sun-Neptune	0.000051773	3.005889908	3.00582087	3.000103544
Earth-Moon	0.012150568	3.200343883	3.184163250	3.024150064

7.3.3 Two Triangle Libration Points

In the second case of $y \neq 0$, there is a solution of Eq. (7.36):

$$r_1 = r_2 = 1. \quad (7.45)$$

It means that the libration point and the two main bodies form an equilateral triangle, therefore the libration point is called an equilateral triangle solution, or just a triangle solution. There are two symmetric triangle libration points L_4 and L_5 , whose positions are given by

$$\begin{cases} x_4 = x_5 = \frac{1}{2} - \mu, \\ y_4 = +\frac{\sqrt{3}}{2}, \quad y_5 = -\frac{\sqrt{3}}{2}. \end{cases} \quad (7.46)$$

7.3.4 Dynamical Characteristics of the Five Libration Points

Although the circular restricted three-body problem is not entirely solved, through the Jacobi integral and the five particular solutions we can extract dynamical properties about this system, therefore, provide theoretical support for orbital motions related to the field of the deep space exploration. In the following content, we provide further analysis connected to some particular requests in aerospace applications.

7.3.4.1 The Jacobi Constant and Its Five Critical Values

From the expression of the Jacobi integral (7.25)

$$2\Omega(x, y, z) - v^2 = C,$$

we can obtain five values of the Jacobi constant corresponding to the five libration points L_i ($i = 1, 2, \dots, 5$), $C_i(\mu)$. At L_i , there is $v^2 = 0$ (note that the speed v is in the synodic coordinate system), then the corresponding values of C_i are

$$C_i(\mu) = 2\Omega(x_i(\mu), 0, 0), \quad i = 1, 2, 3 \quad (7.47)$$

$$C_i(\mu) = 2\Omega(x_i(\mu), y_i, 0), \quad i = 4, 5 \quad (7.48)$$

where $y_4 = \sqrt{3}/2$ and $y_5 = -\sqrt{3}/2$. For L_1 , L_2 , and L_3 ,

$$C_i(\mu) = [x_i^2(\mu) + \mu(1 - \mu)] + 2\left[\frac{1-\mu}{|x_i(\mu)+\mu|} + \frac{\mu}{|x_i(\mu)-(1-\mu)|}\right], \quad (7.49)$$

and for L_4 and L_5 ,

$$C_4(\mu) = C_5(\mu) = 3. \quad (7.50)$$

According to (7.49) and (7.50) and the values of $x_i(\mu)$ and μ , the relationship of the values of the five Jacobi constants at the five libration points is

$$3 = (C_4, C_5) \leq C_3(\mu) \leq C_2(\mu) \leq C_1(\mu) \leq 4.25, \quad (7.51)$$

and the values of $C_1(\mu)$, $C_2(\mu)$, and $C_3(\mu)$ varying with μ are shown in Fig. 7.3.

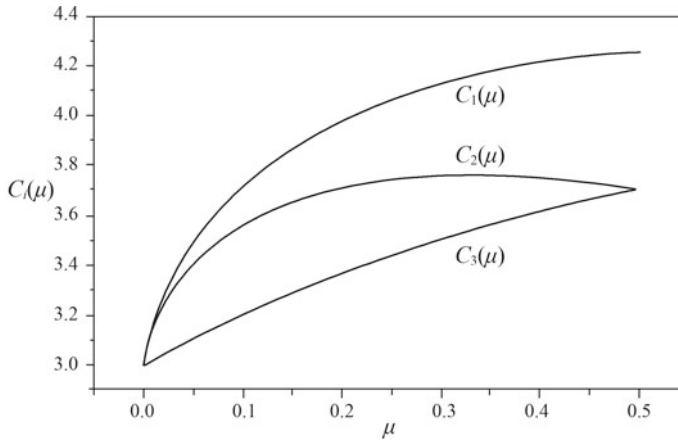


Fig. 7.3 The values of the Jacobi constant $C_i(\mu)$ at the libration points versus μ

7.3.4.2 Surfaces of Zero Velocity and Possible Motion Areas

(1) Surfaces of zero velocity

Since the Jacobi integral (7.25) is an integral of the circular restricted three body-problem then the surface represented by

$$2\Omega(x, y, z) = C \quad (7.52)$$

is a surface of zero velocity, i.e., the motion velocity of a small body on this surface is zero. The value of the integral is given by the initial condition that

$$C = 2\Omega(x_0, y_0, z_0) - v_0^2. \quad (7.53)$$

The geometric structure of a zero-velocity surface depends on the value of the Jacobi constant C . The variations can be illustrated by the cross-sectional curves of the zero-velocity surface in the xy -plane (the zero curves) shown in Figs. 7.4, 7.5, 7.6 and 7.7.

In the four figures, the shadowed areas are for $v > 0$, i.e., the probable motion areas. When the value of C is relatively large (corresponding to a small velocity v), Fig. 7.4 shows that the zero velocity surfaces divide the entire space into four parts. As the value of C decreases, the two zero velocity surfaces surrounding the two main bodies expand, get close, then connect (at L_1), and eventually combine into one (Fig. 7.5). As the value of C continuously decreases the inner zero velocity surface expands then connects to the external zero velocity surface (at L_2), as shown in Fig. 7.6. Finally, the zero velocity surface passes L_3 as seen in Fig. 7.7.

Fig. 7.4 Cross section for $C < C_1$

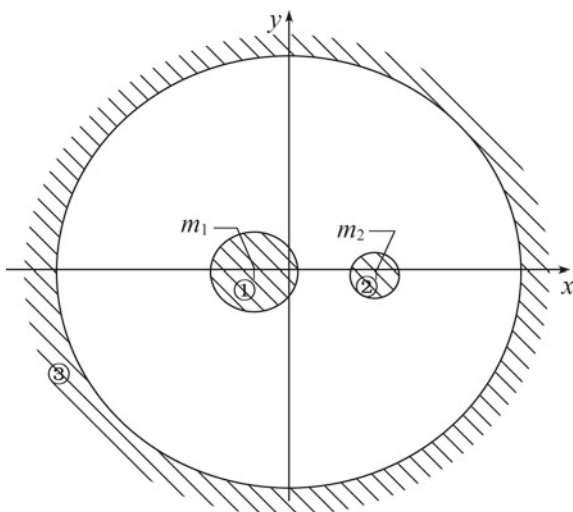
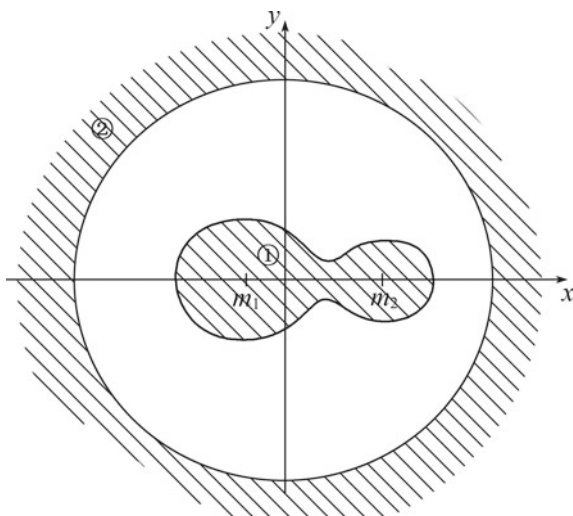


Fig. 7.5 Cross section for $C_2 < C < C_1$



Because the key variations of zero velocity surfaces happen around the libration points ($L_i, i = 1, \dots, 5$), the corresponding five Jacobi constants C_i are often called the critical values.

(2) Probable areas of motion

In the four figures, we can see that the zero velocity surfaces divide the entire space into two types of area, the shadowed areas are for $v > 0$, which are the probable motion areas, and the area outside of the shadowed areas are for $v < 0$, the forbidden

Fig. 7.6 Cross section for $C_3 < C < C_2$

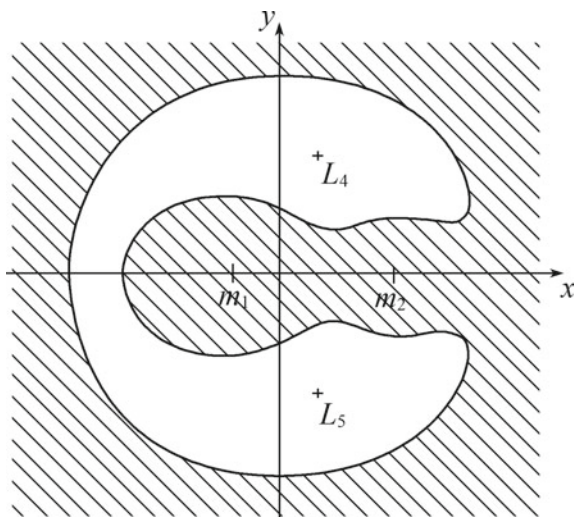
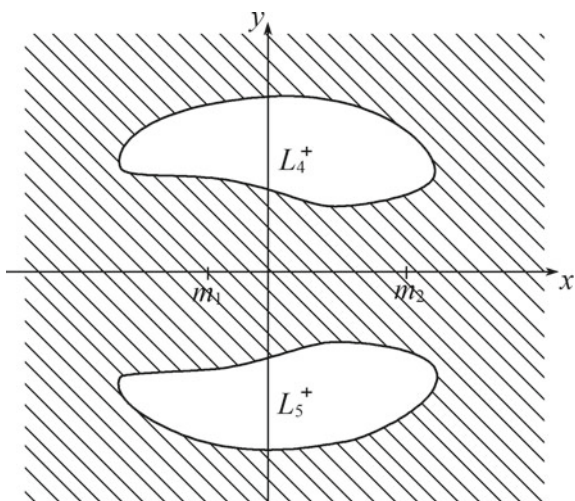


Fig. 7.7 Cross section for $C_4 < C < C_3$



areas of motion. A small body cannot move from a shadowed area passing the zero-velocity surface into a forbidden area (because $v < 0$ is impossible). A small body can only arrive at a zero-surface along a normal line to the surface, then moves back along the same normal line to its original area.

It should be reminded that Figs. 7.4, 7.5, 7.6 and 7.7 only show the plane state, when $C < C_5$, in the xy -plane there is no forbidden area, but in the space, there are still forbidden regions.

The Jacobi integral shows that when the value of C becomes smaller then at the same location the speed of a small body increases, i.e., the probable motion area

expands. In the case of $C_2 < C < C_1$, not only the probable motion areas expand but also the property of the probable motion areas changes. If the motion area of a small body at the beginning is restricted in the vicinity of one of the main bodies, when C decreases the small body may be able to move into the vicinity of the other main body. This phenomenon has been used to explain the material exchange between two stars of the close binary and the formation of the Roche lobe [5–7].

One final note is that in the above discussion for a small body moving from the vicinity of one main body to the vicinity of the other main body its initial velocity must satisfy $C_2 < C < C_1$. This is a necessary condition, more discussions and results for related projects are given in reference [8]. Take a Moon's probe as an example, assuming that the Moon moves around Earth in a circle, the probe after being launched into its original parking orbit has to change its orbit to make its speed large enough to satisfy the condition $C_2 < C < C_1$, then it can fly towards the Moon. If its speed is even larger and satisfies $C_3 < C < C_2$, then it not only can go to the Moon but also can fly away from the vicinity of the Moon and leave the Earth-Moon system to become a small artificial planet.

7.3.4.3 The Gravitational Sphere of the Two Main Bodies in the Circular Restricted Three-Body Problem

The motion of a deep-space probe P is often under the combined attraction of two large celestial bodies P_1 and P_2 . During the process of its motion, the probe can be near P_1 or P_2 , therefore its motion cannot be treated as a perturbed two-body problem but a restricted three-body problem. Because the probe must be near the targeted body (assuming it to be P_2), then when P moves into a region dominated by P_2 , the gravity force of P_2 becomes the main force exerted on P . In this region the motion of P can be approximately regarded as a motion related to P_2 in a two-body problem. This method simplifies a complicated problem and is useful for preliminary analysis. The motion region that is dominated by a celestial body has two definitions as follows.

(1) The gravitational sphere

When the magnitudes of the gravitational forces exerted on a small body from two main bodies are equal, then we assume that P is on the border of the gravitational sphere, which is decided by

$$\frac{Gm_2}{r^2} = \frac{Gm_1}{R^2}, \quad (7.54)$$

where m_1 and m_2 are masses of P_1 and P_2 , respectively. When m_2/m_1 is small, according to the geometry of Fig. 7.8, the radius of the gravitational sphere of P_2 can be approximately given by the distance between L and P_2 denoted by r_1 , that

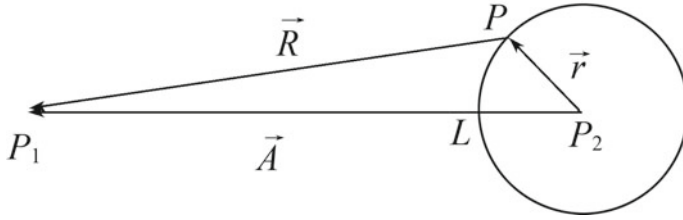


Fig. 7.8 The gravitational sphere of P_2

$$r_1 = \sqrt{\mu}A, \quad \mu = \frac{m_2}{m_1}, \quad (7.55)$$

where $A = |\vec{A}|$ is the distance between P_1 and P_2 .

(2) The gravitational sphere of influence (SOI)

The simplified gravitational sphere given above has a flaw, because the relative motion of P to P_2 (or P_1) is “perturbed” by the other body P_1 (or P_2). To correct it we introduce the sphere of influence. The relative motion equations of P to P_2 and P_1 are given by

$$\begin{cases} \ddot{\vec{r}} = -\frac{Gm_2}{r^2}(\frac{\vec{r}}{r}) + Gm_1\left(\frac{\vec{R}}{R^3} - \frac{\vec{A}}{A^3}\right), \\ \ddot{\vec{R}} = \frac{Gm_1}{R^2}\left(\frac{\vec{R}}{R}\right) - Gm_2\left(\frac{\vec{r}}{r^3} - \frac{\vec{A}}{A^3}\right). \end{cases} \quad (7.56)$$

The sphere of influence is defined by the balance of the two gravity forces of P_1 and P_2 exerted on P that

$$Gm_1\left|\frac{\vec{R}}{R^3} - \frac{\vec{A}}{A^3}\right|\left(\frac{Gm_2}{r^2}\right)^{-1} = Gm_2\left|\frac{\vec{r}}{r^3} - \frac{\vec{A}}{A^3}\right|\left(\frac{Gm_1}{R^2}\right)^{-1}. \quad (7.57)$$

Assuming that the point L in Fig. 7.8 is on the border of the region, and the radius of the sphere is denoted by r_2 , then when m_2/m_1 is small, r_2 can be approximately given by

$$r_2 = \left(\frac{\mu^{2/5}}{2^{1/5}}\right)A, \quad \mu = \frac{m_2}{m_1}. \quad (7.58)$$

The sphere of influence can be used as initial consideration for designing a launching orbit of a deep-space prober. Based on this information a splicing method for the double two-body problem is introduced, which plays a supporting role in the orbit design.

The gravitational sphere (7.55) only reflects a condition of simple “static” balance, which has little meaning for a dynamical problem. In practice we often use the gravitational sphere of influence as the gravitational sphere, so the sphere of influence is also called the gravitational sphere.

(3) Hill sphere

When we discuss the launching condition of a spacecraft, say a Moon prober, we need to know the minimum speed by which the prober can enter the Moon’s gravitational sphere. To decide this speed, we often need another range, which must include the combined effect of gravity forces of P_1 (Earth) and P_2 (the Moon).

As we discussed previously that the original condition of a small body decides the value of the Jacobi constant by (7.53). If the initial position and velocity of P with respect to P_1 are $\vec{r}_0$ and $\vec{v}_0$, respectively, then we have a Jacobi value of C . If $C > C_1$, the motion area of P is given in Fig. 7.4, and if $C_2 < C < C_1$, it is given in Fig. 7.5.

When $C > C_1$, the motion of the prober can only be near Earth. The two areas which are shadowed in Fig. 7.4 expand if the initial speed of P increases (C decreases), and eventually they reach a critical state as shown in Fig. 7.9. The shadowed area in Fig. 7.9 has two parts, one is around P_1 and the other around P_2 . These two parts are called the Hill spheres of the two bodies, the point at which the two Hill spheres meet is the first libration point L_1 . If the distance between L_1 and P_2 is defined as the radius of the Hill sphere of P_2 , denoted by r_3 , then by (7.37) we have

$$r_3 = \left(\frac{\mu}{3}\right)^{\frac{1}{3}} A, \quad \mu = \frac{m_2}{m_1 + m_2}. \quad (7.59)$$

For the main systems in the Solar System (such as the Earth-Moon system, the Sun-(Earth + Moon) system, etc.) the three ranges (r_1 , r_2 , and r_3) for the less massive main body P_2 of the two in these systems are listed in Table 7.4. The mean distance of P_1 and P_2 in the table denoted by A is given by the semi-major axis of the orbit of P_2 except for the Earth-Moon system. The corresponding ranges for other major planets can be calculated by (7.55), (7.58), and (7.59), but are not listed here.

Fig. 7.9 The Hill spheres

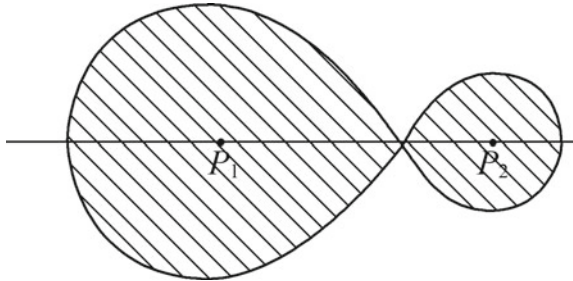


Table 7.4 The ranges of the gravitational spheres, the spheres of influence, and the Hill spheres of planets and the Moon (unit length is 10,000 km)

System	Gravitational sphere	SOI	Hill sphere	A
Sun-Mercury	2.36	9.79	22.07	5790.9083
Sun-Venus	16.9	53.6	101.1	10,820.8600
Sun-(Earth+Moon)	26.1	80.9	150.3	14,959.7870
Sun-Mars	13.0	50.2	108.4	22,793.9184
Sun-Jupiter	2404.9	4196.6	5313.8	77,829.8356
Earth-Moon	4.27	5.78	6.14	38.4401

7.3.4.4 The Second Cosmic Velocity v_2 and the Minimum Velocity of Launching a Deep-Space Prober

The second cosmic velocity v_2 is the minimum velocity to escape Earth's gravity field, i.e., the parabolic velocity of launching a prober from the ground of Earth, which is given by

$$v_2 = \sqrt{\frac{2GE}{a_e}} = \sqrt{2}v_1 = 11.1799 \text{ km/s}, \quad (7.60)$$

where v_1 is the first cosmic velocity for a circling orbit, and v_2 is also called the escape velocity.

To launch a spacecraft, the first thing to be concerned is the escape velocity v_2 . If the spacecraft is launched from a near Earth parking orbit at a height of about 200 km, the corresponding v_2 is 11.0087 km/s, which is not much different from the value of v_2 by (7.60). If the spacecraft is a Moon prober, considering the effect of the Moon's gravity force then the launching velocity can be less.

Suppose that Earth, the Moon, and a prober form a circular restricted three-body problem, we consider "launch" the prober from a parking orbit (a circular orbit with an altitude of 200 km). Assuming that the launch time is 21 March 2011 0 h (UTC), the corresponding modified Julian day number is $\text{MJD} = \text{JD} - 2400000.5 = 55641.0$ (JD is the Julian Day number), then if the initial speed of the prober at $r_p = a_e + 200 \text{ km}$ (a_e is Earth's radius, r_p is the perigee of the packing orbit), and the velocity of the prober when it enters the packing orbit is $v_p = 10.865664 \text{ km/s}$, then the corresponding Jacobi constant $C = 3.200373$, which is near the value of $C_1 = 3.200344$ (Table 7.3), meaning after launching the prober enters the orbit, and can fly to the vicinity of the Moon by the Moon's gravity force (note that $C \leq C_1$ is only the necessary condition for the prober reaching the Moon, see Sect. 7.3.2). The initial orbit elements are

$$\begin{aligned} a &= 127478.137 \text{ km}, \quad e = 0.948398, \quad i = 45.0^\circ, \\ \Omega &= 45.0^\circ, \quad \omega = 10.0^\circ, \quad M = 0.0^\circ. \end{aligned}$$

The corresponding orbit period is $T_s = 5^d \cdot 2426$. As discussed above by the initial condition it is possible to send this prober to the Moon, but the prober has

to fly around Earth several circles before it arrives at the Moon, the actual time is certainly much longer than T_s . Usually, to launch a Moon probe does not choose the minimum speed orbit, because we must consider a different method that costs less energy and takes a shorter time interval. If a probe can pass through the narrow area around the libration point L_1 to reach the Moon, it can certainly save energy.

The best method to launch a Moon probe or a deep-space spacecraft into the target orbit depends on the actual purpose of the space project and the launching conditions. A probe can be launched using the Hohmann transfer to reach the target body directly, or by the sphere of influence using the splicing method of the double two-body problem or a small thrust as transition, references [3, 4] provide related content.

7.3.5 Characteristics and Applications of the Stability of the Five Libration Points

The Jacobi integral and five particular solutions of the circular restricted three-body problem are analyzed and calculated in the previous a few sections from the perspective of aerospace dynamical point (especially about orbits). In this section, we introduce the characteristics of the stability of the five libration solutions and their applications in the aerospace field. The content is selected for the need for deep-space exploration.

7.3.5.1 The Concept of Stability

For a motion, if we denote its state quantity by X , the corresponding motion equation is given by an ordinary differential equation that

$$\frac{dX}{dt} = F(X, t), \quad (7.61)$$

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad F(X, t) = \begin{pmatrix} f_1(x_1, \dots, x_n, t) \\ f_2(x_1, \dots, x_n, t) \\ \vdots \\ f_n(x_1, \dots, x_n, t) \end{pmatrix}. \quad (7.62)$$

If F is a continuous vector function on X and t , also can ensure the existence and the uniqueness of a solution, then there is one and only one solution of the Eq. (7.61) for given initial values of X_0 at initial time t_0 , that the solution

$$X = X(t; t_0, X_0) \quad (7.63)$$

satisfies the initial conditions as

$$X(t_0; t_0, X_0) = X_0. \quad (7.64)$$

Since this solution represents a motion, then the stability of the solution is often a focus of attention. There are various types of stability, here we introduce stabilities related to the motion of a spacecraft.

(1) Stability of solutions: the stability of initial values [9, 10]

The unique solution of Eq. (7.61) which satisfies the initial condition (7.64) is a particular solution that represents an undisturbed motion. If the initial values change a little, then the motion corresponding to the changed initial values is called the disturbed motion. Usually, the motion stability (i.e., the solution stability) implies the stability related to the variations of the difference between the undisturbed and the disturbed motions, which is the stability of the initial values, i.e., the stability according to Lyapunov stability theory or Lyapunov stability.

By this definition, the difference between an unperturbed motion and a perturbed motion is caused only by the difference in the initial values. Therefore, we assume that the initial values of the undisturbed motion and the disturbed motion at $t = t_0$ are

$$x_k^0 \text{ and } x_k^0 + \varepsilon_k \quad (k = 1, 2, \dots, n), \quad (7.65)$$

respectively, and the quantity ε_k is called the initial disturbance. The corresponding disturbed and undisturbed solutions of (7.61) are denoted respectively by

$$x_k(t) \text{ and } \psi_k(t; t_0, x_1^0, \dots, x_n^0).$$

When $t > t_0$, the difference

$$x_k(t) - \psi_k(t; t_0, x_1^0, \dots, x_n^0) \quad (7.66)$$

is called the following disturbance or just the disturbance. Obviously, if all initial disturbances are zero, then all disturbances $x_k(t) - \psi_k(t)$ are zero.

The question about the stability of initial values is that if the components of the initial disturbance ε_k are not all zero, can we determine some small numbers as limits that the values of the following disturbances of $|x_k(t) - \psi_k(t)|$ would never exceed these limits? If the answer is yes, then the solution is stable with respect to the initial values.

The problem with the stability of initial values is related to the dynamical characteristics of the five libration points, and also to the orbital motion of spacecraft. In the orbital forecast, there are always errors in the initial values no matter by analytical methods or by numerical methods. The question is during the process of the motion would the error accumulation cause the forecast results beyond a certain range? Because Kepler motion is unstable with respect to initial values, therefore in

the orbital forecast, it is necessary to decide certain limits for restricting the errors of initial values. In the following content, we introduce the state of stability of initial values at the five libration points and related problems.

(2) The stability of the structure

In dealing with practical problems, we often need to study that when a system shows “disturbance” whether or not there are changes in the topological structure of the system, and under what kind of condition the system can reserve its topological structure. Those are questions about the stability of structures. For example, in the research on non-linear vibration problems, the concept of structural stability in a flat disc system is raised. This problem is presented by a system of equations that

$$\begin{cases} \frac{dx}{dt} = f(x, y), \\ \frac{dy}{dt} = g(x, y), \\ x^2 + y^2 \leq R^2. \end{cases} \quad (7.67)$$

When this system shows “disturbance” there is a source of the disturbance given by $\Delta f(x, y)$ and $\Delta g(x, y)$, then the original system of equations changes to

$$*20c \begin{cases} \frac{dx}{dt} = f(x, y) + \Delta f(x, y), \\ \frac{dy}{dt} = g(x, y) + \Delta g(x, y). \end{cases} \quad (7.68)$$

When $f(x, y)$ and $g(x, y)$ are analytical the necessary and sufficient conditions for the structural stability of this system were provided. In the 1960s, this problem attracted the attention of a group of mathematicians, such as Liao (廖山涛) [11, 12] of China and Smale and his group of USA, [13], etc., but to the present day, there is no substantial progress. The circular restricted three-body problem is an approximation of an actual dynamical model, there is no question about the existence of disturbance. The question of whether or not the structural property of the five libration solutions can maintain is important as it is directly related to the problem of the spacecraft’s position in the space.

7.3.5.2 Overview of the Stability of Libration Solutions

In a circular restricted three-body problem, assuming that a libration solution is given by

$$x_0 = a, \quad y_0 = b, \quad z_0 = 0, \quad (7.69)$$

and an initial disturbance is given by

$$\Delta x = \xi, \quad \Delta y = \eta, \quad \Delta z = \zeta, \quad (7.70)$$

Substituting $x = x_0 + \Delta x = a + \xi$, $y = y_0 + \Delta y = b + \eta$, and $z = \zeta$ into the motion Eq. (7.20) gives

$$\begin{cases} \ddot{\xi} - 2\dot{\eta} = \Omega_{xx}^0 \xi + \Omega_{xy}^0 \eta + \Omega_{xz}^0 \zeta + O(2), \\ \ddot{\eta} + 2\dot{\xi} = \Omega_{yx}^0 \xi + \Omega_{yy}^0 \eta + \Omega_{yz}^0 \zeta + O(2), \\ \ddot{\zeta} = \Omega_{zx}^0 \xi + \Omega_{zy}^0 \eta + \Omega_{zz}^0 \zeta + O(2), \end{cases} \quad (7.71)$$

where $O(2)$ is for small quantities of the second- and higher-order of ξ , η , and ζ , and $\Omega_{xx}^0, \Omega_{xy}^0, \dots, \Omega_{zz}^0$ are the values of the second-order partial derivatives of Ω at a libration point, that

$$\begin{cases} \Omega_x^0 = 0, \quad \Omega_y^0 = 0, \quad \Omega_z^0 = 0 \\ \Omega_{xy}^0 = \Omega_{yx}^0 \\ \Omega_{xz}^0 = \Omega_{zx}^0 = \Omega_{yz}^0 = \Omega_{zy}^0 = 0. \end{cases} \quad (7.72)$$

After omitting the small terms of higher orders in the right-side functions of (7.71), there are

$$\begin{cases} \ddot{\xi} - 2\dot{\eta} = \Omega_{xx}^0 \xi + \Omega_{xy}^0 \eta, \\ \ddot{\eta} + 2\dot{\xi} = \Omega_{yy}^0 \eta + \Omega_{yx}^0 \xi, \\ \ddot{\xi} = \Omega_{zz}^0 \zeta. \end{cases} \quad (7.73)$$

Equation (7.73) is a system of linear homogenous equations with constant coefficients. The component ζ can be separated, and the solution of ζ corresponds to a simple harmonic vibration, which means the small body is not far away from the xy -plane. Now we only need to discuss the motion of the small body in the xy -plane. From the two equations of (7.73) about ξ and η , the characteristic equation is

$$\begin{vmatrix} \lambda^2 - \Omega_{xx}^0 & -2\lambda - \Omega_{xy}^0 \\ 2\lambda - \Omega_{xy}^0 & \lambda^2 - \Omega_{yy}^0 \end{vmatrix} = 0.$$

This is a fourth-degree algebraic equation of λ that

$$\lambda^4 + (4 - \Omega_{xx}^0 - \Omega_{yy}^0)\lambda^2 + [\Omega_{xx}^0 \Omega_{yy}^0 - (\Omega_{xy}^0)^2] = 0. \quad (7.74)$$

(1) The situations of the three collinear libration solutions L_1, L_2 , and L_3

Since there is $0 < \mu < \frac{1}{2}$ (which is true for any restricted three-body problem except $m_1 = m_2$) at the three collinear libration points, there are

$$\begin{cases} \Omega_{xx}^0 = 1 + 2C_0 > 0, & \Omega_{yy}^0 = 1 - C_0 < 0, & \Omega_{zz}^0 = -C_0 < 0 \\ \Omega_{xy}^0 = 0, & \Omega_{xx}^0 > 0, & \Omega_{yy}^0 < 0 \\ \Omega_{xx}^0 \Omega_{yy}^0 - (\Omega_{xy}^0)^2 < 0 \end{cases} \quad (7.75)$$

$$C_0 = \frac{(1-\mu)}{r_1^3} + \frac{\mu}{r_2^3}, \quad (7.76)$$

where r_1 and r_2 in (7.76) are the distances between a libration point and the two main bodies, respectively, i.e., $r_1 = |x_i + \mu|$ and $r_2 = |x_i - 1 + \mu|$ where $i = 1, 2$, and 3 for the three collinear libration points. The four eigenvalues of (7.74) are

$$\begin{cases} \lambda_{1,2} = \pm S_1^{\frac{1}{2}}, \\ \lambda_{3,4} = \pm S_2^{\frac{1}{2}}. \end{cases} \quad (7.77)$$

$$\begin{cases} S_1 = \frac{1}{2} \left\{ -(4 - \Omega_{xx}^0 - \Omega_{yy}^0) + \left[(4 - \Omega_{xx}^0 - \Omega_{yy}^0)^2 - 4\Omega_{xx}^0 \Omega_{yy}^0 \right]^{\frac{1}{2}} \right\}, \\ S_2 = \frac{1}{2} \left\{ -(4 - \Omega_{xx}^0 - \Omega_{yy}^0) - \left[(4 - \Omega_{xx}^0 - \Omega_{yy}^0)^2 - 4\Omega_{xx}^0 \Omega_{yy}^0 \right]^{\frac{1}{2}} \right\}. \end{cases} \quad (7.78)$$

With (7.75) we know $S_1 > 0$ and $S_2 < 0$, therefore there is one positive real root, and according to the definition of solution stability the three collinear libration solutions are unstable. When the high-order terms are included in (7.71) the results are the same.

The four eigenvalues of (7.74) and the two eigenvalues corresponding to the component ζ can be given by the following forms that

$$\begin{cases} \lambda_{1,2} = \pm d_1, \\ \lambda_{3,4} = \pm id_2, \quad \lambda_{5,6} = \pm id_3, \quad i = \sqrt{-1}, \end{cases} \quad (7.79)$$

where $d_3 = \sqrt{C_0} > 0$, $d_1 > 0$, $d_2 > 0$, and

$$\begin{cases} d_1 = \left[\frac{1}{2}(9C_0^2 - 8C_0)^{\frac{1}{2}} - \left(1 - \frac{C_0}{2}\right) \right]^{\frac{1}{2}}, \\ d_2 = \left[\frac{1}{2}(9C_0^2 - 8C_0)^{\frac{1}{2}} + \left(1 - \frac{C_0}{2}\right) \right]^{\frac{1}{2}}. \end{cases} \quad (7.80)$$

In the linear forms, the motion of the small body in the vicinities of the three collinear libration points is given by

$$\begin{cases} \xi = C_1 e^{d_1 t} + C_2 e^{-d_1 t} + C_3 \cos d_2 t + C_4 \sin d_2 t, \\ \eta = \alpha_1 C_1 e^{d_1 t} - \alpha_1 C_2 e^{-d_1 t} - \alpha_2 C_3 \sin d_2 t + \alpha_2 C_4 \cos d_2 t, \\ \zeta = C_5 \cos d_3 t + C_6 \sin d_3 t, \end{cases} \quad (7.81)$$

where

$$\alpha_1 = \frac{1}{2}(d_1 - \Omega_{xx}^0/d_1), \quad \alpha_2 = \frac{1}{2}(d_2 + \Omega_{xx}^0/d_2). \quad (7.82)$$

The six integral constants, $C_1, C_2, \dots$, and C_6 are decided by the initial disturbance conditions that at $t_0 = 0$, the initial disturbances are $\xi_0, \dot{\xi}_0, \eta_0, \dot{\eta}_0, \zeta_0, \dot{\zeta}_0$. Because $d_1 > 0$ by (7.81) we know that although the initial motion state of the small body satisfies the condition of the libration solution, after a small disturbance the small body leaves the libration point, the large the value of d_1 is the fast the small body leaves. For example, in the Sun-(Earth + Moon)-small body system, the values of d_1 at the three collinear libration points are 2.532659, 2.484317, and 0.002825, respectively, which imply that a small body starts near L_1 or L_2 and leaves faster than it starts near L_3 . It is understandable according to their positions.

(2) The situations of the two triangle libration solutions L_4 , and L_5

For the two triangle libration points, it is easy to give that

$$\begin{cases} \Omega_{xx}(L_{4,5}) = 3[x^2 + \mu(1 - \mu)](L_4, L_5) = \frac{3}{4}, \\ \Omega_{yy}(L_{4,5}) = 3y^2(L_4, L_5) = \frac{9}{4}, \\ \Omega_{xy}(L_4) = 3xy(L_4) = +\frac{3\sqrt{3}}{2}\left(\frac{1}{2} - \mu\right), \quad \Omega_{xy}(L_5) = 3xy(L_5) = -\frac{3\sqrt{3}}{2}\left(\frac{1}{2} - \mu\right). \end{cases} \quad (7.83)$$

The corresponding characteristic equation is

$$\lambda^4 + \lambda^2 + \frac{27}{4}\mu(1 - \mu) = 0. \quad (7.84)$$

Let $S = \lambda^2$, then (7.84) becomes a second-degree algebraic equation of S that

$$S^2 + S + \frac{27}{4}\mu(1 - \mu) = 0. \quad (7.85)$$

The solutions are

$$\begin{cases} S_1 = \frac{1}{2} \left\{ -1 + [1 - 27\mu(1 - \mu)]^{\frac{1}{2}} \right\}, \\ S_2 = \frac{1}{2} \left\{ -1 - [1 - 27\mu(1 - \mu)]^{\frac{1}{2}} \right\}. \end{cases} \quad (7.86)$$

The corresponding eigenvalues are

$$\lambda_{1,2} = \pm\sqrt{S_1}, \quad \lambda_{3,4} = \pm\sqrt{S_2}. \quad (7.87)$$

The property of the eigenvalues depends on the value of the discriminant in (7.86), i.e., $d = 1 - 27\mu(1 - \mu)$. When

$$0 < 1 - 27\mu(1 - \mu) < 1,$$

there is

$$\mu(1 - \mu) < \frac{1}{27}, \quad (7.88)$$

so $S_1 < 0$ and $S_2 < 0$, and there are two pairs of conjugate imaginary roots that

$$\begin{cases} \lambda_{1,2} = \pm iS_1, & \lambda_{3,4} = \pm iS_2, \\ S_1 = \sqrt{|S_1|}, & S_2 = \sqrt{|S_2|}. \end{cases} \quad (7.89)$$

In the linear forms based on the definition of stability, the two triangle solutions are stable, and the critical value of μ_0 is given by

$$\mu_0(1 - \mu_0) = \frac{1}{27}.$$

Note that $\mu < \frac{1}{2}$. Now we have the linear stability conditions as

$$\begin{cases} 0 < \mu < \mu_0 \\ \mu_0 = \frac{1}{2} \left(1 - \frac{\sqrt{69}}{9} \right) = 0.038520896504551 \dots \end{cases} \quad (7.90)$$

In the Solar System, there are Sun-Jupiter-asteroid system, Sun-Earth-Moon system, etc., which can be treated as restricted three-body problems, and the corresponding values of μ all satisfy the condition (7.90). The value of μ_0 is usually called the Routh stability criterion.

When $\mu_0 < \mu < \frac{1}{2}$, the triangle libration points are unstable. If $\mu = 0$ or μ_0 they are also unstable. More information about the stability of the triangle libration points is given in Chap. 9 reference [4].

7.3.5.3 Applications of the Stability Characteristics of Libration Solutions

(1) Applications of the collinear libration solutions (L_1, L_2, L_3)

Theoretically, the three collinear libration solutions are unstable, but for linearized motion Eq. (7.73) the motion near any of the libration points given by (7.81) shows that it is possible that certain initial conditions of disturbance can make $C_1 = 0$ and $C_2 = 0$, then the corresponding disturbed solution becomes a periodic or quasi-periodic motion. This property is used for particular aerospace projects. One of the applications is to provide the theoretical foundation for rendezvous, which means a controlled process of assembling satellites together so that they do not collide but actually rendezvous. The method of design rendezvous is given in Sect. 7.4. Another important application is for deep-space exploration to build a special periodic or

quasi-periodic orbit around one of the collinear libration points, which is called Halo orbit or Lissajous orbit. This method is a hot topic in present deep-space exploration, which is systematically presented in reference [14].

(2) Applications of the triangle libration solutions

The two triangle libration solutions are stable with respect to initial values, but are not necessarily structurally stable. So far, the conclusion that the triangle solutions are structurally unstable has not been rigidly proved, but it can be shown simply using the method of disproval given below.

In reference [14] the authors provide the state of deviation of a prober from one of the triangle libration points using three models, which are given as follows.

- ① The model of an ideal circular restricted three-body problem (Earth-Moon-probe).
- ② The model of the perturbed restricted three-body problem (Earth-Moon-probe), in which the perturbing body is the Sun, the orbit of the second main body (the Moon) in the three-body problem is given by the “actual” orbit (which is the secular precession orbit after eliminating short-period variations), and the initial position of the prober is at the triangle libration point without any deviation.
- ③ The model is the same as the second model except for the perturbing force of the Sun is reduced artificially by one order.

By model ①, the initial deviations of the position of the small body from the triangle libration point are 10,000 km in three directions of x , y , z , which are inside of the “stable range”. The deviations are relatively large, but after 130 periods of the Moon moving around Earth the prober is still in a stable state as shown in Fig. 7.10, in which the unit of length RU is the mean distance between Earth and Moon. It is a result of an ideal circular restricted problem.

By model ② (not only the relative motion orbits of the two main bodies are ellipses but also a fourth body’s gravity force is included which affects all the three bodies in the model), although there are no initial deviations of the small body’s position, the small body moves away from the libration point and soon leaves the stable region, as shown in Fig. 7.11.

By model ③ in the Earth-Moon-Sun-prober system, the perturbing force of the Sun is artificially reduced by one order of quantity (equivalent to reducing the mass of the Sun by one order) to make it less than the value of μ (0.012) related to the mass of the Moon. Results show that after the same time interval as it is in model ① (130 periods of the Moon’s evolution around Earth in Fig. 7.10), the small body can maintain in the vicinity of the triangle libration point, and the size of the wandering region in the geocentric coordinates is not too much larger than in model ① (Fig. 7.12).

The above numerical results and analyses indicate that the triangle libration points in the restricted three-body problem are stable even in a non-linear state. In reality, although the value of μ related to the Moon is relatively large the gravitational force of the fourth body (the Sun) is too strong, and there is always an initial deviation, therefore the question about how stable the triangle libration points in the Earth-Moon restricted three-body problem has no practical meaning. If we want to use the

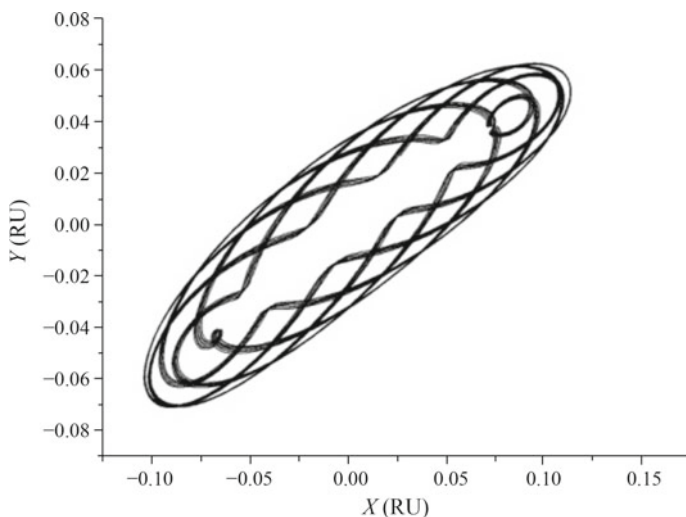


Fig. 7.10 The wandering state of a prober in the restricted three-body problem (Earth-Moon-prober) around a triangle libration point

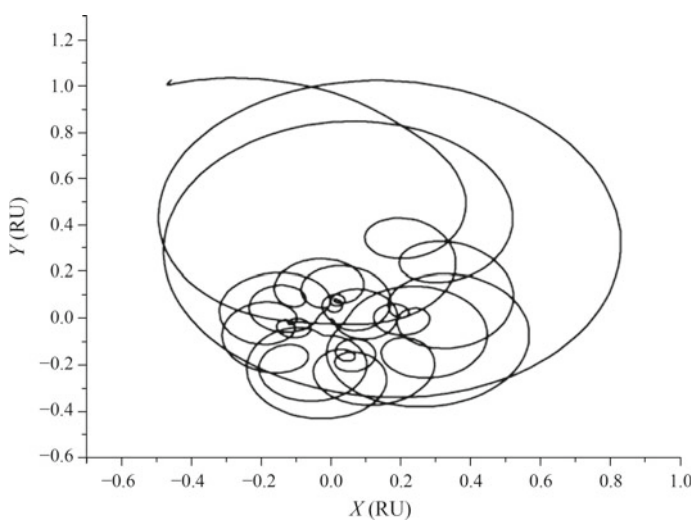


Fig. 7.11 The motion state of a prober in the Earth-centric frame of a complete dynamical model (Earth-Moon-Sun-prober) around a triangle libration point

triangle libration points to build a station in the Earth-Moon system, then we have to use an orbital controlling system that performs according to the deviation pattern, further discussion is given in reference [15].

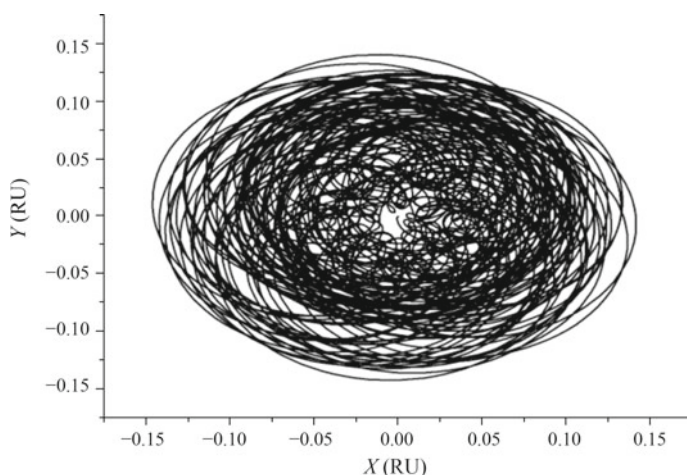


Fig. 7.12 The motion state of a probe in the Earth-centric frame of a dynamical model (Earth-Moon-Sun-probe, but the Sun's mass was artificially reduced by an order) around a triangle libration point

7.3.6 *Calculations and Applications of Libration Points in the Restricted Problem of (2+2) Bodies*

The restricted problem of (2+2) bodies is a kind of expansion of the circular restricted three-body problem. In the Solar System, this kind of structure actually exists. For example, in the main asteroid belt, the distance of two asteroids can be very close, and if the effects of gravitational forces between them cannot be ignored then the Sun, Jupiter, and the two asteroids form a restricted problem of (2+2) bodies. Another case is that assume two or more than two geosynchronous satellites to be launched at a point above Earth's equator, each of them is weighted a few tons, and the distances of them are no more than 1 km, then for high precision problems, the gravity forces between them need to be considered. In this situation, Earth, which as an ellipsoid is equivalent to two bodies (a mass evenly distributed sphere and the non-spherical part), and the two satellites form a restricted problem of (2+2) bodies. In deep-space exploration we may also meet a similar problem, for example, if we want to launch two probes to be stationed around a particular spot, the gravity forces between them cannot be ignored, thus the two satellites and two related big bodies form a restricted problem of (2+2) bodies.

In a circular restricted problem of (2+2) bodies, any one of the small bodies and the two big bodies form a particular perturbed circular restricted three-body problem, the perturbation is from the other small body. In this dynamical model, the question is whether the motion states of the two small bodies are similar to that in a circular restricted three-body problem. This kind of orbital dynamical problem and the related calculation method should be studied.

7.3.6.1 The Mathematical Model of the Circular Restricted Problem of (2+2) Bodies and Its Libration Points

In order to relate this model to the circular restricted three-body problem, the basic motion Eq. (7.20) in the synodic coordinate system can be changed to the following forms [16, 17]

$$\begin{cases} \ddot{x}_\alpha - 2\dot{y}_\alpha = \frac{\partial U}{\partial x_\alpha}, \\ \ddot{y}_\alpha + 2\dot{x}_\alpha = \frac{\partial U}{\partial y_\alpha}, \\ \ddot{z}_\alpha = \frac{\partial U}{\partial z_\alpha}, \end{cases} \quad \alpha = 1, 2 \quad (7.91)$$

where U is the “gravity potential” of the system given by

$$U = \sum_{\alpha=1}^2 \left[\frac{1}{2} (x_\alpha^2 + y_\alpha^2) + \frac{1-\mu}{r_{\alpha 1}} + \frac{\mu}{r_{\alpha 2}} + \frac{\mu_{3-\alpha}}{r} \right] \quad (7.92)$$

with

$$\begin{cases} \vec{r}_{\alpha 1} = \vec{r}_\alpha - \vec{R}_1, \vec{r}_{\alpha 2} = \vec{r}_\alpha - \vec{R}_2, \vec{r} = \vec{r}_1 - \vec{r}_2, \\ \vec{R}_1 = \begin{pmatrix} -\mu \\ 0 \\ 0 \end{pmatrix}, \vec{R}_2 = \begin{pmatrix} 1-\mu \\ 0 \\ 0 \end{pmatrix}, \end{cases} \quad (7.93)$$

$$\mu = \frac{M_2}{M_1+M_2}, \quad \mu_\alpha = \frac{m_\alpha}{M_1+M_2}. \quad (7.94)$$

For the motion of each small body, this system does not have an integral, but for the whole system there is a general Jacobi integral, which is

$$\sum_{\alpha=1}^2 \frac{1}{2} v_\alpha^2 = 2U - C. \quad (7.95)$$

Similar to the circular restricted three-body problem this is the only integral we know. If there are any particular solutions in the synodic frame, they should satisfy the following conditions:

$$\frac{\partial U}{\partial x_\alpha} = 0, \quad \frac{\partial U}{\partial y_\alpha} = 0, \quad \frac{\partial U}{\partial z_\alpha} = 0.$$

The corresponding specific forms are

$$\begin{cases} \left[x_1 - \frac{(1-\mu)(x_1+\mu)}{r_{11}^3} - \frac{\mu(x_1-1+\mu)}{r_{12}^3} \right] - \frac{\mu_2}{r^3}(x_1 - x_2) = 0, \\ y_1 \left(1 - \frac{1-\mu}{r_{11}^3} - \frac{\mu}{r_{12}^3} \right) - \frac{\mu_2}{r^3}(y_1 - y_2) = 0, \\ z_1 \left(\frac{1-\mu}{r_{11}^3} + \frac{\mu}{r_{12}^3} \right) + \frac{\mu_2}{r^3}(z_1 - z_2) = 0, \end{cases} \quad (7.96)$$

$$\begin{cases} \left[x_2 - \frac{(1-\mu)(x_2+\mu)}{r_{21}^3} - \frac{\mu(x_2-1+\mu)}{r_{22}^3} \right] - \frac{\mu_1}{r^3}(x_2 - x_1) = 0, \\ y_2 \left(1 - \frac{1-\mu}{r_{21}^3} - \frac{\mu}{r_{22}^3} \right) - \frac{\mu_1}{r^3}(y_2 - y_1) = 0, \\ z_2 \left(\frac{1-\mu}{r_{21}^3} + \frac{\mu}{r_{22}^3} \right) + \frac{\mu_1}{r^3}(z_2 - z_1) = 0. \end{cases} \quad (7.97)$$

From the two components in the z direction in (7.96) and (7.97), we have

$$z_2 = z_1 \left[1 + \frac{r^3}{\mu_2} \left(\frac{1-\mu}{r_{11}^3} + \frac{\mu}{r_{12}^3} \right) \right], \quad (7.98)$$

$$z_1 \left[\left(\frac{1-\mu}{r_{21}^3} + \frac{\mu}{r_{22}^3} \right) + \frac{r^3}{\mu_2} \left(\frac{1-\mu}{r_{11}^3} + \frac{\mu}{r_{12}^3} \right) \left(\frac{1-\mu}{r_{21}^3} + \frac{\mu}{r_{22}^3} + \frac{\mu_1}{r^3} \right) \right] = 0. \quad (7.99)$$

Obviously, in the synodic frame, the circular restricted problem of (2+2) bodies also has libration solutions that lie in the xy -plane, i.e.,

$$z_1 \equiv z_{10} = 0, \quad z_2 \equiv z_{20} = 0. \quad (7.100)$$

Comparing the x - and y -components in (7.96) and (7.97) with (7.35) of the circular restricted three-body problem, the only difference is a small quantity related to the other small body. Therefore, if the circular restricted problem of (2+2) bodies has any libration points, they should be not far from the libration points in the circular restricted three-body problem.

For each small body, there are collinear libration points and triangle libration points. Furthermore, each of the collinear libration points becomes two, and each of the triangle libration points becomes four forming a symmetric configuration. Based on the original solutions of the circular restricted three-body problem these results are not difficult to obtain, also from the analytical point of view, it is reasonable that these libration points are distributed symmetrically. The method to derive these solutions is similar to the method used in the circular restricted three-body problem, which is by iteration. Using the solutions of the circular restricted three-body problem as initial values, and substituting the initial values into (7.96) and (7.97) by the iteration method we can obtain the positions of the libration points.

7.3.6.2 Motion States of Two Satellites in a Binary System on the Geosynchronous Satellite Orbit

The geosynchronous satellites are communicational instruments, which are stationed above the equator. Because the space above the equator is limited, in order to improve

efficiency, it is possible to place multiple geosynchronous satellites at the “same” location. If there are two satellites, usually the distance between them is small (about 100 m), the gravity forces between them form a binary system and cannot be ignored. As mentioned above, Earth as an ellipsoid can be treated as two bodies, the masses of the two geosynchronous satellites compared to Earth can be ignored, therefore the two satellites and Earth form a circular restricted problem of (2+2) bodies.

The dynamical characteristics of the libration points in the restricted problem of (2+2) bodies can be efficiently used to form a binary system of geosynchronous satellites. Initially, each of the two satellites is placed at its own libration point. Similar to the motion of a single satellite, each of the two synchronous satellites would oscillate in the vicinity of the related libration point (the libration phenomenon). Because there are small disturbances in the initial positions of the two satellites, their drifting motions around the libration points are not completely “synchronous”, therefore they cannot maintain the initial distance between them, but the distance between them can be adjusted by an onboard control system. Details about this topic are given in reference [18].

7.4 Orbit Design for Formation Flying of Satellites and Companion-Flying in the Exploration of Asteroids

7.4.1 The Principle of Satellite Formation Flying

The simplest equation for the formation flying of two spacecraft (or formation flying of two satellites moving around Earth) is the Clohessy-Wiltshire equation (i.e., the C-W equation) [19]. This equation is a linearization of the relative motion of two spacecraft in a two-body problem with a circular reference orbit. This equation can be derived in a different way given here. Considering the dynamical mechanism, this problem can be related to the degenerated state of the circular restricted three-body problem when $\mu = 0$, which implies that the mass of the less massive body of the two main bodies in the original model can be omitted.

We now use the formation of Earth’s satellites as an example to explain this problem. Assume that Earth is the main body P_1 ; the central body is a satellite which is the smaller main body P_2 and moves around Earth in a circular orbit; then a small body P is the companion-flying satellite. The masses of the two satellites are small, when the distance between them is not too close (e.g., the distance of the two satellites is larger than 100 m and each of them is weighted a few tons) the attraction forces between them can be ignored, then this three-body problem becomes a restricted three-body problem with $\mu = 0$. Then the central body without mass can be an imaginary body, the circular restricted three-body problem is reduced to a restricted two-body problem, and the origin of the synodic frame C can be moved to P_1 (the center of Earth). The original positions of the five libration points when $\mu \neq 0$ shown in Fig. 7.13 now are changed according to $\mu = 0$.

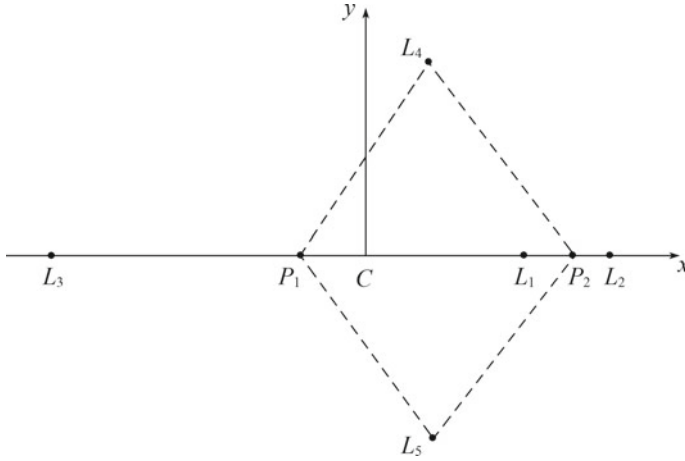


Fig. 7.13 The five libration points in the circular restricted three-body problem

In the geocentric rotation frame (the rotation angular speed is the angular speed of the center satellite moving around Earth) for this kind of restricted two-body problem the motion equation of the companion-flying satellite is given by

$$\begin{cases} \ddot{x} - 2\dot{y} = U_x = \left(x - \frac{x}{r^3}\right), \\ \ddot{y} + 2\dot{x} = U_y = \left(y - \frac{y}{r^3}\right), \\ \ddot{z} = U_z = -\frac{z}{r^3}. \end{cases} \quad (7.101)$$

The original five libration points L_i ($i = 1, 2, 3$) and L_j ($j = 4, 5$) are degenerated that:

L_1 and L_2 together with P_2 are united into one at the point of $x = 1$ and $y = 0$;

L_3 is at the point of $x = -1$ and $y = 0$;

L_4 and L_5 are at $(x = 1/2, y = \pm\sqrt{3}/2)$.

In fact, from (7.101) the case of $U_x = 0$, $U_y = 0$, and $U_z = 0$, corresponds to

$$z = 0, \quad r = 1, \quad (7.102)$$

which means that in the xy -plane, every point on the unit circle centered at the center of Earth (P_1) is a libration point.

Omitting the process of proving, the conclusion is that the five degenerated libration points are all unstable. For L_4 and L_5 because $\mu = 0$, one of the sources of force, P_2 , no longer exists, therefore the balance of gravity is altered.

The two collinear libration points L_1 and L_2 now are combined into one, although they are unstable, similar to the circular restricted three-body problem, in the vicinity

of the point the small body can still maintain a periodic or a quasi-periodic motion state.

Now we give a small disturbance at L_i ($i = 1, 2, 3$) whose coordinate components are ξ , η , and ζ . By substituting ξ , η , and ζ into (7.101), we have

$$\begin{cases} \ddot{\xi} - 2\dot{\eta} = U_{xx}\xi + U_{xy}\eta + U_{xz}\zeta + O(2), \\ \ddot{\eta} + 2\dot{\xi} = U_{yx}\xi + U_{yy}\eta + U_{yz}\zeta + O(2), \\ \ddot{\zeta} = U_{zx}\xi + U_{zy}\eta + U_{zz}\zeta + O(2), \end{cases} \quad (7.103)$$

where $O(2)$ is of higher-order quantity. At L_i ($x = \pm 1, y = 0, z = 0$) there are

$$\begin{cases} U_{xx} = 3, & U_{xy} = 0, & U_{xz} = 0, \\ U_{yx} = 0, & U_{yy} = 0, & U_{yz} = 0, \\ U_{zx} = 0, & U_{zy} = 0, & U_{zz} = -1. \end{cases} \quad (7.104)$$

Substituting (7.104) into (7.103) and omitting higher-order terms, we obtain the linearized equation for the small disturbance around L_i as

$$\begin{cases} \ddot{\xi} - 2\dot{\eta} = 3\xi, \\ \ddot{\eta} + 2\dot{\xi} = 0, \\ \ddot{\zeta} + \xi = 0. \end{cases} \quad (7.105)$$

These are exactly the C-W equations of the companion-flying satellite. As we mentioned above there are no gravity forces between the two satellites, the central satellite can be imaginary, i.e., it does not matter if it exists or not as long the speed of the rotating frame is the angular speed of the imaginary satellite on a circular orbit. If we choose two libration points at $x = 0, y = \pm 1$, then around these two points the system of the linearized motion equations of the small disturbance takes the form as

$$\begin{cases} \ddot{\xi} - 2\dot{\eta} = 0, \\ \ddot{\eta} + 2\dot{\xi} = 3\eta, \\ \ddot{\zeta} + \zeta = 0, \end{cases} \quad (7.106)$$

which corresponds to

$$U_{xx} = 0, \quad U_{yy} = 3, \quad U_{zz} = -1. \quad (7.107)$$

The above discussion shows that although there is no dynamical connection between the two satellites, as long they are close, the conformation of their motions can be explained by the dynamical mechanism around the libration points in a restricted two-body problem. In either (7.105) or (7.106) the ζ component can be separated from the system of equations, and the motion in this direction is a harmonic vibration, i.e., the companion-flying satellite has a small vibration with respect to the xy -plane. For the components of ζ and η , the corresponding characteristic equation is

$$\lambda^4 + \lambda^2 = 0, \quad (7.108)$$

which has a pair of equal roots λ_1 and λ_2 , and a pair of conjugate imaginary roots λ_3 and λ_4 , that

$$\lambda_{1,2} = 0, \quad \lambda_{3,4} = \pm\sqrt{-1}. \quad (7.109)$$

The corresponding motion solutions are

$$\begin{cases} \xi = -C_1 + C_3 \cos t + C_4 \sin t, \\ \eta = \frac{3}{2}C_1 t + C_2 - 2C_3 \sin t + 2C_4 \cos t, \\ \dot{\xi} = -C_3 \sin t + C_4 \cos t, \\ \dot{\eta} = \frac{3}{2}C_1 - 2C_3 \cos t - 2C_4 \sin t. \end{cases} \quad (7.110)$$

or

$$\begin{cases} \xi = \frac{3}{2}C_1 t + C_2 - 2C_3 \sin t + 2C_4 \cos t, \\ \eta = C_1 - C_3 \cos t - C_4 \sin t, \\ \dot{\xi} = \frac{3}{2}C_1 - 2C_3 \cos t - 2C_4 \sin t, \\ \dot{\eta} = C_3 \sin t - C_4 \cos t. \end{cases} \quad (7.111)$$

Choosing proper initial conditions can make $C_1 = 0$ and $C_2 = 0$, then the solution is a conditioned periodic motion for the linearized equation. The chosen initial conditions are

$$t = t_0 : \xi_0, \eta_0, \dot{\xi}_0 = \frac{\eta_0}{2}, \dot{\eta}_0 = -2\xi_0, \zeta_0, \dot{\zeta}_0. \quad (7.112)$$

The disturbed motion solution is now “quasi-periodic” given by

$$\begin{cases} \xi = \xi_0 \cos t + \frac{\eta_0}{2} \sin t, & \dot{\xi} = -\xi_0 \sin t + \frac{\eta_0}{2} \cos t \\ \eta = -2\xi_0 \sin t + \eta_0 \cos t, & \dot{\eta} = -2\xi_0 \cos t - \eta_0 \sin t \\ \zeta = \zeta_0 \cos t + \dot{\zeta}_0 \sin t, & \dot{\zeta} = -\zeta_0 \sin t + \dot{\zeta}_0 \cos t \end{cases} \quad (7.113)$$

In the xy -plane, the configuration of the solution is an ellipse. If the initial disturbance ζ_0 satisfies the following condition

$$\dot{\zeta}_0 = \pm \frac{\eta_0}{2\xi_0} \zeta_0, \quad (7.114)$$

then the configuration in the yz -plane is also an ellipse. The equations of the two ellipses are

$$\frac{\xi^2}{A^2} + \frac{\eta^2}{B^2} = 1, \quad (7.115)$$

$$\frac{\eta^2}{C^2} + \frac{\zeta^2}{D^2} = 1, \quad (7.116)$$

where

$$A^2 = \xi_0^2 + \frac{\eta_0^2}{4}, \quad B^2 = 4A^2, \quad (7.117)$$

$$C^2 = 4A^2 = B^2, \quad D^2 = \left(\frac{4\xi_0^2}{\xi_0^2} \right) A^2. \quad (7.118)$$

The above conditioned periodic motion given by the linearized motion equation is the dynamical mechanics of the formation flying or companion flying of satellites. But the distance between the two satellites cannot be large. If the distance is large the higher-order terms $O(2)$ in (7.103) cannot be omitted, even when the conditions (7.112) and (7.114) are satisfied, the companion flying configuration can still be destroyed. But in the problem of formation flying when $\mu = 0$, the motion of the “center” satellite is unstable, the small body leaves the libration point in a linear way with time t , so it is a weak instability. When $\mu \neq 0$, the corresponding instability is a strong one, because $d_1 \neq 0$, the separation is exponential. Therefore, in the formation flying of small satellites, as long as the distance between the two satellites is not large, the requirement for control conditions can be satisfied. The question about whether or not the high-order terms should be included has certain meanings for pure theoretical study, but has no practical meanings. Because the problem is about formation flying of small satellites (not for formation flying of navigating constellation when the distance of satellites is large), the relative distance is small, and the magnitudes of high-order terms generally are less than the terms of eccentricity and other perturbation terms. Therefore, the influence of the high-order terms does not need to be considered independently, but can be combined with other factors as deviations from (7.112) and managed by an orbital controlling system.

From the above discussion, it is not hard to see that the configuration of satellite formation flying in fact is the result of a restrained unstable motion by condition (7.112) corresponding to a two-body problem. Regarding the characteristics of motion around the libration points, the method of obtaining the condition for formatting the configuration is by transiting a restricted three-body problem to a degenerated restricted two-body problem. This method not only has a simple process but also clearly underlines the dynamical mechanism of this problem rather than a problem of the geometric configuration of a relative motion.

7.4.2 The Problem with the Eccentricity in Orbit Design of Formation Flying of Satellites

In the orbital design of satellite formation flying the problem of whether we need to consider the eccentricity depends on the actual requirement of the aerospace project. Because the eccentricity is relatively small, in most cases it is reasonable to use a circular orbit. If the eccentricity is relatively large, it is possible to use a simplified model of an elliptic restricted three-body problem to obtain the solution for

the formation flying of satellites. In this model the independent variable time t is replaced by the true anomaly f in the relative motion of the two main bodies, then the motion equation of the small body can be given in a variable-scale coordinate system (omitting the derivation) as

$$\begin{cases} x'' - 2y' = \frac{1}{1+e'\cos f}(U_x + x), \\ y'' + 2x' = \frac{1}{1+e'\cos f}(U_y + y), \\ z'' + z = \frac{1}{1+e'\cos f}(U_z + z), \end{cases} \quad (7.119)$$

where

$$U(r_1, r_2) = \frac{1-\mu}{r_1} + \frac{\mu}{r_2}. \quad (7.120)$$

This system of equations also has five libration points, and when $\mu = 0$, there is

$$U(r_1, r_2) = U(r_1) = \frac{1-\mu}{r_1} = \frac{1}{r}. \quad (7.121)$$

As in the circular problem, every point on the unit circle centered at Earth's center (P_1) (the unit of length is the radius r of the circular orbit of the central satellite around Earth) is a libration point. Around a libration point L ($x = 1, y = 0, z = 0$), the linearized motion equations of a disturbance are given by

$$\begin{cases} \xi'' - 2\eta' = 3\xi/(1 + e \cos f), \\ \eta'' + 2\xi' = 0, \\ \zeta'' + \zeta = 0. \end{cases} \quad (7.122)$$

The solution of the disturbance is easy to obtain if only the $O(e)$ term is included, and the conditions for the configuration of the formation flying satellite can also be derived with a relatively simple form [20, 21]. Although this treatment does not provide new information it has a certain value for further discussion. From the practical point of view, the result is much more complicated than the circular orbit problem, particularly the introduction of the variable-scale coordinate system, which is inconvenient for the original physical space. The actual dynamical background is complicated, there are other perturbation factors to be considered. In practice, besides the circular reference orbit, all the influences on the formation of flying conditions can be managed by a comprehensive treatment method. Of course what kind of reference orbit to use depends on each aerospace mission, it is not proper to give a general decisive conclusion.

7.4.3 Extension of the Principle and Related Orbit Design Method of Satellite Formation Flying

The principle of satellite formation flying and the rationally simplified method of orbit design can be extended to build Halo orbit or Lissajous orbit around the collinear libration points in the circular restricted three-body problem based on the solution given by (7.81). Comparing the two linear solutions (7.110) and (7.81) for the two models, we can see that the key technique is to choose proper initial conditions to eliminate the unstable terms in the solutions. The methods are not exactly the same. To build a Halo orbit or a quasi-periodic orbit in the circular restricted three-body problem the high-order terms need to be included in the basic Eq. (7.71) for adjusting the two critic frequencies d_2 and d_3 in (7.81). The details are complicated so are not provided in this book, but are available in references [3, 4].

7.4.4 Orbital Problem of Companion Flying in Asteroid Exploration

To the present day, to explore an asteroid besides ground-based observations is to send a spacecraft passing or circling around the target asteroid. By the method of circling around the target asteroid, a spacecraft becomes a low orbit satellite of the asteroid and the main force exerted on the spacecraft is from the target asteroid, and other forces from major celestial bodies are perturbations (i.e., the third-body perturbation). For this dynamical system, the critical problem is to know the gravitational field of the target asteroid, especially if one has a weird shape. The corresponding orbital problem of the spacecraft has no essential difference from a low Earth orbit satellite. More content on this problem is given in reference [4] of Chap. 10.

In this section, we introduce another type of asteroid exploration in the companion flying form. There are two methods for this type, one is that the asteroid's mass is relatively large and its gravity force on the prober cannot be ignored, the Sun-asteroid-prober form a restricted three-body problem and the prober is stationed at a collinear libration point; the other is for a small asteroid, the prober directly becomes a companion flying with the asteroid.

(1) Companion flying at a collinear libration point

If an asteroid has a circular orbit and a relatively small mass, then this asteroid, the Sun, and a prober form a circular restricted three-body system. Because of the small mass of the asteroid the collinear libration points (L_1 and L_2) are close to the asteroid, then a prober can be sent to the vicinity of one of the libration points to perform preliminary exploration. After obtaining the basic information including the shape of the asteroid, the orbit of the prober can be reduced and the prober becomes a satellite of the asteroid or directly lands on it.

Assuming that the mass of the asteroid is m , then in the Sun-asteroid-prober restricted three-body problem the mass parameter μ is given by

$$\mu = \frac{m}{M_{\odot} + m}. \quad (7.123)$$

If the asteroid is a mass evenly distributed sphere, its mass density and radius are denoted to ρ and R , respectively, then we have $m = 4\pi\rho R^3/3$. To use real numbers as an example, assuming $\rho = 2 \times 10^3 \text{ kg/m}^3$ and $R = 1 \text{ km}$, $m = 8.4 \times 10^{12} \text{ kg}$, then $\mu = 4.2 \times 10^{-18}$, which is very small, therefore

$$\mu \approx \frac{m}{M_{\odot}}. \quad (7.124)$$

If the distance between the Sun and the asteroid is a , from (7.37) and (7.38) the distance between the collinear libration points L_1 or L_2 and the asteroid (omitted the high-order terms, which are very small) is given by

$$R_H \approx a \left(\frac{\mu}{3} \right)^{\frac{1}{3}} \approx a \left(\frac{4\pi\rho R^3}{9M_{\odot}} \right)^{\frac{1}{3}} \approx \left(\frac{4\pi\rho}{9M_{\odot}} \right)^{\frac{1}{3}} a R. \quad (7.125)$$

The position of the libration points should be located outside of the asteroid, which requires

$$\left(\frac{4\pi\rho}{9M_{\odot}} \right)^{\frac{1}{3}} a R > R \Rightarrow \rho > \frac{9M_{\odot}}{4\pi} \frac{1}{a^3}. \quad (7.126)$$

If the asteroid is near Earth, e.g., $a = 1 \text{ AU}$, the request (7.126) becomes

$$\rho > 4.3 \times 10^{-4} \text{ kg/m}^3. \quad (7.127)$$

For an asteroid, these criteria can be satisfied.

(2) Direct companion flying

In this case, the main body is the Sun P_1 , the other main body is the target asteroid whose mass is very small and can be ignored, and the prober is the fly companion. The direct companion flying problem of the prober and the target asteroid actually is the formation flying problem when the gravity of the asteroid “disappears”.

(3) A note about the companion flying

For the formation flying of a prober and an asteroid theoretically, it requires $r_2 \gg (\mu/3)^{1/3}$. When r_2 is relatively small, the gravity force of the asteroid can affect the prober, but the effect is small, by the formation flying method the energy for controlling the target orbit is still small. But the value of r_2 cannot be smaller than the gravitational sphere of the asteroid, because then the motion of the prober would mainly depend on the gravity force of the asteroid.

7.5 Geometric Characteristics of Libration Point Orbit and Orbit Determination

In Sect. 7.3.1, the libration solutions and their geometric properties are given in the synodic coordinate system. By this coordinate system, we obtain simple mathematical expressions of the libration solutions and their important dynamical characteristics. But in an actual aerospace mission, observations, orbit controlling, etc., have to be operated in the real physical space. For example, how to deal with the L_2 libration point orbit in the Earth-Moon system including telemetry and control from a ground-based observatory. Obviously, we have to deal with it in the corresponding geocentric celestial coordinate system, in which the L_2 point orbit is actually a high Earth satellite orbit. For this orbit, we only need to know concepts, such as the gravitational sphere of influence, etc. in the circular restricted three-body problem given in Sect. 7.3.4.

7.5.1 Geometric Characteristics of Libration Point Orbits

We now use two simple numerical examples to clearly show the geometric characteristics of the libration solution in the Earth-Moon restricted three-body problem. The initial epoch is 30 September 2016 00:00:00 (UTC), the corresponding TDT is 57,661.0007891667 (MID). Put one probe on the L_1 orbit and another prober on the L_2 orbit in the Earth-Moon system, respectively, and use the simple coordinate transformation we obtain the initial positions, velocities, and orbital elements of the two orbits in the J2000.0 geocentric celestial coordinate system, which are given in Tables 7.5 and 7.6.

The two initial orbits are presented in Figs. 7.14 and 7.15 for the two probers on L_1 and L_2 , respectively, showing that the two initial orbits actually are high eccentric ellipses around Earth, one prober is at the apogee and the other at the perigee (which is easy to understand), the unit of the frame axes A_e is the equatorial radius of Earth's reference ellipsoid (which is previously denoted by a_e). In the Earth-Moon + prober system, these are initial instantaneous orbits, by the gravity force of the Moon the probers and the Moon all move "synchronously" on circular orbits. Note that these two original orbits are given in the Earth-Moon circular restricted three-body problem model, so they are circles.

7.5.2 Analysis of Forces on a Prober's Motion in a Libration Orbit

Theoretically, a prober on the libration point in the Earth-Moon system moves around Earth has the same orbital "period" as the Moon, but the effect on the prober from

Table 7.5 Initial positions and velocities of the two probers on L_1 and L_2

L_i	x (km)	y (km)	z (km)	$\dot{x}$ (km/s)	$\dot{y}$ (km/s)	$\dot{z}$ (km/s)
$i = 1$	-337,774.810825	-337,774.810825	16,503.725924	-0.08734133	-0.793809190	-0.262191839
$i = 2$	-464,586.522898	27,663.672934	22,699.764448	-0.120132126	-1.091831124	-0.360627231

Table 7.6 Initial orbital elements of the two probers on L_1 and L_2

L_i	a (km)	e	i (deg)	Ω (deg)	ω (deg)	M (deg)
$i = 1$	242,063.297	0.40063038	18.507748	4.970405	353.948868	353.948868
$i = 2$	1,064,951.700	0.56309178	18.507748	4.970405	166.026309	1.189763

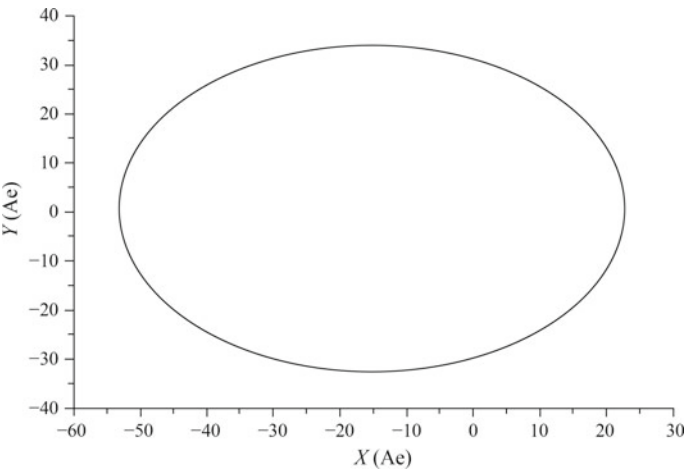


Fig. 7.14 The initial orbit on L_1 in J2000.0 geocentric celestial coordinate system (in the equatorial plane)

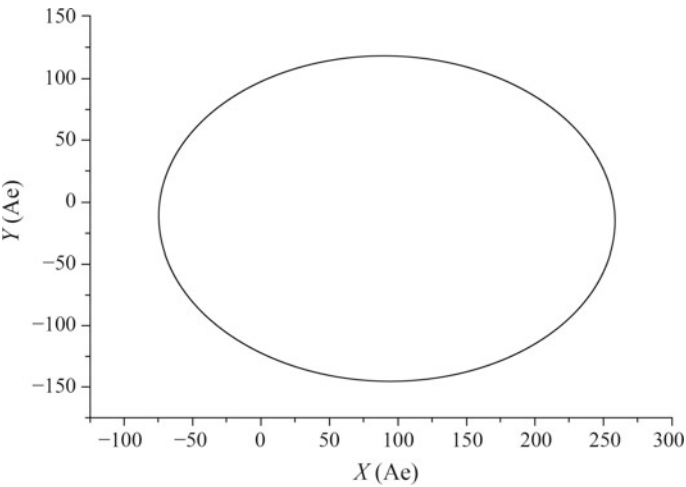


Fig. 7.15 The initial orbit on L_2 in J2000.0 geocentric celestial coordinate system (in the equatorial plane)

the Moon is not a small perturbation but “almost” as important as Earth’s gravity force. If we treat the motion of the prober as a perturbed two-body problem, its state motion equation is given by

$$\ddot{\vec{r}} = -\frac{\mu}{r^3}\vec{r} + \sum_{j=1}^N \vec{F}_j(m_j) + \vec{F}_\varepsilon(\varepsilon), \quad (7.128)$$

where $\mu = G(E + m)$, m is the mass of the prober which can be ignored so $m = 0$, GE is Earth’s gravitational constant; m_j ($j = 1, 2, \dots$) is the mass for a perturbing celestial body, which can be the Moon, the Sun, or other bodies. The Moon and the Sun are the main bodies with masses m_1 and m_2 , respectively; $\vec{F}_\varepsilon(\varepsilon)$ is the perturbation acceleration due to other external factors (including the gravity forces of non-spherical parts of major bodies, solar radiation pressure, etc.).

Similar to low Earth orbit satellites and general high Earth orbit satellites in the motion equation there are different physical quantities with different dimensions and sizes which are not convenient for analyzing. As usual, we use the dimensionless units (or normalized units) to study this problem. The units of length $[L]$, mass $[M]$, and time $[T]$ as in (4.4) are defined by

$$\begin{cases} [L] = a_e (\text{equatorial radius of Earth's reference spheroid}), \\ [M] = E (\text{Earth's mass}), \\ [T] = \left(\frac{a_e^3}{GE}\right)^{\frac{1}{2}} \end{cases}$$

where the time unit $[T]$ is a derived unit. Now in this normalized system, the gravitational constant $G = 1$, and the gravitational constant of the central body is $\mu = GE = 1$.

With the normalized units, the state motion Eq. (7.128) takes the following form

$$\ddot{\vec{r}} = -\frac{1}{r^3}\vec{r} + \sum_{j=1}^N \vec{F}_j(m'_j) + \vec{F}_\varepsilon(\varepsilon) \quad (7.129)$$

where m'_j is the third body’s dimensionless mass (the Moon or the Sun), that

$$m'_1 = \frac{GM}{GE}, \quad m'_2 = \frac{GS}{GE}, \quad \dots, \quad (7.130)$$

GM and GS are gravitational constants of the Moon and the Sun, respectively. Then the acceleration by the third body’s gravity force in the motion Eq. (7.129), $\vec{F}_j(m'_j)$, is given by

$$\vec{F}_j(m'_j) = -m'_j \left(\frac{\vec{\Delta}_j}{\Delta_j^3} + \frac{\vec{r}'_j}{r_j^3} \right), \quad (7.131)$$

where $\vec{\Delta}_j = \vec{r} - \vec{r}'_j$ ($j = 1, 2, \dots$), and $\vec{r}'_j$ is the coordinate vector of the third body, the Moon or the Sun, in the geocentric celestial coordinate system.

(1) Estimations of gravitational perturbation magnitudes of major celestial bodies

Earth's gravitational constant is $GE = 398,600.4418 \text{ km}^3/\text{s}^2$, and the ratios of the gravitational constants of the Moon, the Sun, Mercury, Venus, Mars, Jupiter, and Saturn to Earth are 0.0123000383, 332,946.050895, 0.055273598, 0.814998108, 0.107446732, 317.8942053, and 95.1574041, respectively.

The perturbation magnitude is estimated approximately by

$$\varepsilon_j = m'_j \left(\frac{r}{r'_j} \right)^3, \quad (7.132)$$

where m'_j is the ratio of the mass of a major celestial body to Earth's mass which is the same ratio of the gravitational constants; r and r'_j are the distances of the prober and the perturbing body to the center of Earth, respectively, for estimating the value of r'_j can take its average except Mercury. For Mercury because of the high eccentricity ($e > 0.2$) the value of r' may take one average value around the perihelion and another average value around the aphelion.

Considering a prober placed at the collinear libration points L_1 or L_2 in the Earth-Moon system, the distances from L_1 and L_2 to Earth's center are

$$\begin{cases} r_1 = 0.849065782 \overline{EL} = 51.2a_e, \\ r_2 = 1.167832664 \overline{EL} = 70.4a_e, \end{cases} \quad (7.133)$$

where $\overline{EL}$ is the average distance between Earth and the Moon. From (7.132) the perturbation magnitudes of the above listed celestial bodies (i.e., the Moon, the Sun, Mercury, Venus, Mars, Jupiter, and Saturn) exerted on the prober at L_1 and L_2 are estimated as

$$\varepsilon = m' \left(\frac{r}{r'} \right)^3 = \begin{cases} 0.8 \times 10^{-2}, & 2.0 \times 10^{-2}, \\ 3.5 \times 10^{-3}, & 0.9 \times 10^{-2}, \\ (3.8 - 1.7) \times 10^{-9}, & (1.0 - 0.45) \times 10^{-8}, \\ 4.0 \times 10^{-7}, & 1.0 \times 10^{-6}, \\ 0.8 \times 10^{-8}, & 2.0 \times 10^{-8}, \\ 4.5 \times 10^{-8}, & 1.2 \times 10^{-7}, \\ 1.1 \times 10^{-9}, & 2.9 \times 10^{-9}. \end{cases} \quad (7.134)$$

(2) Perturbation magnitudes of the Earth's non-spherical gravity force

The estimated perturbation magnitudes of the main term J_2 at L_1 and L_2 are

$$\varepsilon = \varepsilon(J_2) = J_2 \left(\frac{3}{r^2} \right) = \begin{cases} 1.2 \times 10^{-6}, & L_1, \\ 0.6 \times 10^{-6}, & L_2. \end{cases} \quad (7.135)$$

(3) Perturbation magnitudes of solar radiation pressure

For a prober located at L_1 or L_2 in the Earth-Moon system with a regular size (including the mass and the equivalent area of radiation pressure) the estimated perturbation magnitudes of the solar radiation pressure are

$$\varepsilon = \varepsilon(\rho_\odot) = \left(\frac{\kappa S}{m} \right) \rho_\odot r^2 = \begin{cases} 1.2 \times 10^{-5}, & L_1, \\ 2.3 \times 10^{-5}, & L_2, \end{cases} \quad (7.136)$$

where $\kappa = 1.44$, $(S/m) = 10^9$, $\rho_\odot = 0.3169 \times 10^{-17}$.

Based on the estimated magnitudes of all possible perturbation sources for a prober located at L_1 or L_2 to build a dynamical model the perturbations greater than 10^{-6} are the gravity forces of the Moon, the Sun, and Mars as particles, the gravity force of Earth's non-spherical part J_2 term, and the solar radiation pressure. The most important forces are the gravity forces of the Moon and the Sun.

7.5.3 Orbit Determination and Forecast Method of Libration Point Orbit

As mentioned above the perturbation from the Moon exerted on a prober at the libration point of the Earth-Moon system in fact is not a small disturbance but an external force almost as important as Earth's gravity force, therefore, this problem cannot be treated as a simple perturbed two-body problem. For ground-based telemetry, tracking, and control, including orbit determination, orbital forecast method, and studying characteristics of Earth-Moon libration point orbits, it is proper to use the J2000.0 geocentric celestial coordinate system and numerical methods. The mathematical model is the perturbed two-body problem, the state motion equation is (7.129) by the dimensionless format that

$$\ddot{\vec{r}} = -\frac{1}{r^3} \vec{r} + \sum_{j=1}^N \vec{F}_j(m'_j) + \vec{F}_\varepsilon(\varepsilon).$$

The perturbation terms are analyzed and provided previously.

The precise orbit determination and position forecast are procedures of numerically solving the state motion Eq. (7.129) to provide a perturbed ephemeris about the orbital position of the prober according to the accuracy requirements. For the libration point orbit, the method is the same as for any other satellite, and the only method is the numerical method. The method of orbit determination is given in Chaps. 9 and 10, and the numerical method is given in Chap. 8.

7.5.4 Orbit Determination of Libration Point Orbit and Precision Examination of Short-Arc Forecast [22]

A rigorous orbital design for a libration point is unachievable. To place a prober at a libration point can only be done approximately. It is because the libration point itself is unstable, that the errors of initial values can spread much faster and more severe than in a general case of a circling orbit of a spacecraft. Keeping a prober near a libration point we must use frequent orbital controlling. Ground-based or space-based measurements and orbital controlling can be only provided through short-arcs.

The libration point orbital determination and forecast using short-arcs are similar to dealing with low Earth orbits or high Earth orbits, there are no specific difficulties and particular problems to solve. Our research group uses the conventional orbit determination method and software, and optical measurements provided by Chinese USB, without any other additional information to determine the orbit of Chang'e 3 spacecraft, and then compare our results of the orbit determination with the results determined by a Chinese aerospace group after the mission. Based on the results we then use simple numerical extrapolation (only including the gravitational forces of Earth, the Moon, and the Sun as particles, and the solar radiation pressure; and the initial six orbital elements used for extrapolation are given by the initial orbit determination of related mission) we provide high precision orbital forecasts without difficulty. The results are listed in Tables 7.7 and 7.8.

The results, A and B, given by the Chinese aerospace group and our group, respectively, show excellent agreements. For a better understanding of the precisions of orbital determination and extrapolations, we give further explanations as follows.

- (1) About the solar radiation pressure model. Since we do not have detailed information on the structure of the spacecraft, we use related information to estimate the equivalent thermal area-to-mass ratio ($\kappa S/m$), and derive an empirical model, which is an equivalent plane model.
- (2) Although the details about the orbit determination and extrapolations given by the two groups are not provided, the results in the two tables show that high precision is achieved. The extrapolations are for three days and seven days, the differences between A and B are basically in a range of less than 500 m. The agreements are provided by two independent working groups, which imply the correctness of the dynamical model for this particular orbit described in the section.

The final conclusion is that although the errors of the initial values of this kind of orbit spread much more notable than other general circling orbits, there are no particular difficulties in high precision orbit determination and short-arc orbit forecast.

Table 7.7 Comparisons of an L_2 Halo orbit by orbit extrapolation over 3 days and afterward precise orbit determination

Method	x (km)	y (km)	z (km)	$\dot{x}$ (km/s)	$\dot{y}$ (km/s)	$\dot{z}$ (km/s)
Precise orbit determination A	-451,897.422	138,113.307	36,463.951	-0.237782	-0.978285	-0.204902
Precise orbit determination B	-451,897.388	138,112.927	36,463.914	-0.237782	-0.978287	-0.204900
Short-arcs extrapolation A	-451,897.458	138,113.145	36,463.888	-0.237783	-0.978286	-0.204902
Short-arcs extrapolation B	-451,897.346	138,112.816	36,463.866	-0.237782	-0.978288	-0.204900

Table 7.8 Comparisons of an L_2 Halo orbit by orbit extrapolation over 7 days and afterward precise orbit determination

Method	x (km)	y (km)	z (km)	$\dot{x}$ (km/s)	$\dot{y}$ (km/s)	$\dot{z}$ (km/s)
Precise orbit determination A	-371,590.823	- 223,419,646	-55,369,560	0.740132	-0.975794	-0.327870
Precise orbit determination B	-371,590.387	-223,420.334	-55,369.450	0.740133	-0.975794	-0.327871
Short-arcs extrapolation A	-371,590.759	-223,420.292	-55,369.879	0.740132	-0.975795	-0.327871
Short-arcs extrapolation B	-371,589,646	-223,420,816	-55,369,813	0.740138	-0.975793	-0.327873

7.5.5 *Orbital Transformation Between the Two Coordinate Systems for a Libration Point Orbit Prober*

For orbit determination and forecast we need the J2000.0 geocentric celestial coordinate system, but for the particular libration point orbit, we need the Earth-Moon rotational synodic coordinate system. The transformation between the two coordinate systems is given in Sect. 7.1. In specific aerospace missions based on different requirements, there are different ways to define the Earth-Moon rotational coordinate system, we do not further discuss this problem.

References

1. Szebehely V (1967) Theory of orbits (chapter 1). New York and London, Academic Press
2. Stiefel EL, Scheifele G (1971) Linear and regular celestial mechanics. New York, Springer-Verlag, Berlin, Heidelberg
3. Liu L, Hou XY (2012) Orbit dynamics of deep spacecraft. Electronic Industry Press, Beijing
4. Liu L, Hou XY (2015) Theory and application of deep spacecraft orbit dynamics. Electronic Industry Press, Beijing
5. Liu L, Huang C (1984) A few problems about close binary dynamics. Res Astron Astrophys 4(4):253–263
6. Liu L (1987) A brief discussion about the Roche model. Res Astron Astrophys 7(3):169–176
7. Liu L (2007) Discussion du modele de Roche, in Jean-Michel Faidit ed: Limites et Lobes de Roche, Societe astronomique de France VUIBERT.
8. Zhao ZY, Liu L (1994) The stability region of a P-type retrograde asteroid in the Solar System. Acta Astronom Sinica 35(4):434–438
9. Ye YQ (1982) Lectures of Ordinary Differential Equation (Chapter 6), Higher Education Press, Beijing
10. Arnold VI (1989) Geometric methods in the theory of ordinary differential equations (chapter 3). Science Press, Beijing
11. Liao ST (1962) Certain ergodic properties of a differential system on a compact differential manifold in an ordinary differential equation system. J Beijing Univ 3:241–265, and (1963) 4:309–324.
12. Liao ST (1979) The structural stability of the ordinary differential equation and related questions. Appl Comput Appl Math 7:52–64
13. Smale S (1967) Differentiable dynamical systems. Bull Amer Math Soc 73:747–817
14. Liu L, Liu HG (2008) The position wandering and controlling of a prober stationed on a triangle libration point in the earth and moon system. J Astron 29(4):1222–1227
15. Liu L, Hou XY (2009) Prospect of the application of libration point in the deep space exploration. Progr Astron 27(2):174–182
16. Whipple AL, Szebehely V (1984) The restricted problem of $n+v$ bodies. Celest Mech 32(2):137–144
17. Whipple AL (1984) Equilibrium solutions of the restricted problem of 2+2 bodies. Celest Mech 33(3):271–294
18. Zhang Q, Liu L (1999) Binary system of geosynchronous satellites. J Nanjing Univ 35(1):7–13
19. Clohessy WH, Wiltshire RS (1960) Terminal guidance system for satellite rendezvous. J Aerospace Sci 27(9):653–674
20. Inalhan G, Tillerson M, How JP (2002) Relative dynamics and control of spacecraft formations in eccentric orbits. J Guidance Control Dyn 25(1):48–59

21. Hou XY, Zhao YH, Liu L (2012) Formation flying in elliptic orbits with the J_2 perturbation. *Res Astron Astrophys* 12(11):1563–1575
22. Liu L, Tang JS, Hou XY (2018) The characteristics and related problems of the orbits around the Earth-Moon libration points. *ACTA Astronomica Sinica* 59(3):29 1–12, and *China Astron Astro Phys* 2019 43:278–291

Chapter 8

Numerical Method for Satellite Orbit Extrapolations



As given in previous chapters that the differential equations, which describe the motions of celestial bodies and spacecraft in the Solar System, are extremely complicated. To the present day, except for the simplest two-body problem, these equations are not solved rigorously. Under certain conditions, based on the mathematical model for a perturbed two-body problem, solutions of power series of small parameters can be obtained. But for high precision requirements, the solutions including high-order terms are extremely long, which are not only difficult to derive but also difficult to use. Furthermore, some perturbations have no analytical expressions, so are impossible to build models for them. In reality, there are problems that do not require analytical solutions, such as afterward orbit determinations and related orbit forecasts, as well as orbit designs under certain restrictions. These problems only need discrete solutions of related motion equations with the required accuracy. The method to provide discrete solutions of ordinary differential motion equations is the numerical method. The numerical method has a very important role in solving problems in orbit dynamics and is particularly prominent in the era of highly developed computing technology. Of course it does not mean that the numerical method can replace the analytical method and the qualitative method. In the field of orbit dynamics, it is profoundly important to know the characteristics and patterns of orbital variations of celestial bodies and spacecraft, which can only be provided by analytical and qualitative methods. From the perspective of research, the numerical method is a necessary auxiliary method to supply valuable information.

8.1 Basic Knowledge of Numerical Method in Solving the Motion Equation

8.1.1 Basic Principles of Numerical Method in Solving Motion Equation

A differential equation with a given initial value has the form as

$$\begin{cases} \frac{dx}{dt} = f(x, t), \\ x(a) = x_0, \quad a \leq t \leq b, \end{cases} \quad (8.1)$$

where x can be a vector (such as a coordinate position $\vec{r}$ or a velocity $\dot{\vec{r}}$). To numerically solve this equation is to calculate the approximations of the solution at a series of discrete points given by t_n ($n = 1, 2, \dots, m$) over a time interval $[a, b]$ that

$$a = t_0 < t_1 < t_2 < \dots < t_m = b. \quad (8.2)$$

Usually, if no specific requirement is given, the values of t_n are distributed evenly, i.e.,

$$t_n = t_0 + nh, \quad n = 1, 2, \dots, m, \quad (8.3)$$

where $t_0 = a$, and h is the step size. From the initial value $x_0 = x(t_0) = x(a)$ we obtain x_1 , then repeat the step to get the next value, eventually, we have all x_n corresponding to t_n ($n = 1, 2, \dots, m$). By this discrete method, the continuous Eq. (8.1) is solved approximately.

One of the direct ways to obtain x_{n+1} from x_n is by the Taylor expansion in the power of h which is

$$x_{n+1} \approx x_n + h \left(\frac{dx}{dt} \right)_{t_n} + \frac{h^2}{2!} \left(\frac{d^2x}{dt^2} \right)_{t_n} + \dots + \frac{h^p}{p!} \left(\frac{d^p x}{dt^p} \right)_{t_n}, \quad (8.4)$$

where p is a positive integer. From (8.1) the derivatives are given by

$$\left(\frac{dx}{dt} \right)_{t_n} = f(x_n, t_n), \quad \left(\frac{d^2x}{dt^2} \right)_{t_n} = \left(\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} \frac{dx}{dt} \right)_{t_n}, \dots \quad (8.5)$$

Obviously, it is a kind of discretized method, but we must calculate high-order derivatives of function $f(x, t)$. The calculation of higher order derivatives can be difficult if $f(x, t)$ is complicated. For the motion equations of celestial bodies (especially various spacecraft) the right-side functions are extremely complicated. Therefore,

generally, it is not practical to directly use the Taylor series. In order to simplify the calculation almost all numerical methods use a few values of the right-side function to replace the high-order derivatives, which is the basic concept of solving the differential equation numerically.

8.1.2 Basic Concepts

In a Taylor series, keeping the first two terms on the right side of (8.4) and discarding all other higher-order terms, we have the Euler method:

$$\begin{cases} x_{n+1} = x_n + hf(x_n, t_n), & n = 0, 1, 2, \dots, \\ x_0 = x(t_0). \end{cases} \quad (8.6)$$

This is a discretized derivative equation. Some basic concepts of the numerical method can be introduced by the Euler method. The Euler method (8.6) shows that for each step at t_n as long as we know the value of x_n , we can forward get the value of x_{n+1} at t_{n+1} . This method is called the one-step method. If for every step forward it needs more than one of the previous values of x , then it is called the multi-step method. The above described single-step method from x_n to x_{n+1} is also called the explicit method.

Another method is the backward Euler method given as

$$\begin{cases} x_{n+1} = x_n + hf(x_{n+1}, t_{n+1}), \\ x_0 = x(t_0). \end{cases} \quad (8.7)$$

The wanted x_{n+1} is included in the right-side function, so this method is also called the implicit method. The method of solving (8.7) is by iteration with the provided initial value of x_{n+1} which is $x_{n+1}^{(0)}$.

Because the value of the numerical solution at t_n , x_n , is an approximation of $x(t_n)$, it has an error. There are different kinds of errors, such as the initial error, the truncation error, and the rounding error. We now use the Euler method (8.6) to explain the truncation error. The accurate solution should satisfy the Taylor series in h with a remainder that

$$x(t_{n+1}) = x(t_n) + h\dot{x}(t_n) + \frac{h^2}{2}\ddot{x}(\xi_n), \quad (8.8)$$

where ξ_n is a point on the time interval (t_n, t_{n+1}) . Since

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

(8.8) can be written as

$$x(t_{n+1}) = x(t_n) + hf(x(t_n), t) + \frac{h^2}{2}\ddot{x}(\xi_n). \quad (8.9)$$

We can see that the Euler method (8.6) is actually the approximation of (8.9) after the term $\frac{h^2}{2}\ddot{x}(\xi_n)$ is truncated. Therefore, the term $\frac{h^2}{2}\ddot{x}(\xi_n)$ is the local truncation error of the Euler method (also called the truncation error), or the local discrete error, which has the same order as h^2 , and is denoted by $O(h^2)$. Obviously, its actual value also depends on the value of $\ddot{x}(t_n)$, which is related to the characteristics of $x(t)$. Because the truncated term is the second-order derivative, the Euler method is called a first-order method. The order of a commonly used single-step method is one order lower than the order of the truncated error, e.g., if the order of the truncated error is $O(h^{p+1})$, then the method is regarded as a p -order method, thus the order of a method can be used as an important sign to indicate the accuracy of the method.

The rounding error is produced by a few factors: a limited length of calculating numbers no matter which computing tools to be used; calculations between numbers with limited lengths (involving computer software); the accuracy of calculating the right-side function $f(x, t)$, etc. Therefore, the rounding error is more complicated than the truncation error, but it is still possible to estimate its range, by treating it as a random quantity, then using a statistical method to estimate its value.

There is a question about the truncation error, that when $h \rightarrow 0$, whether $x_n \rightarrow x(t_n)$ or not. Here $x(t_n)$ and x_n are the actual solution and the approximation of the solution of (8.1) at $t = t_n$, respectively. This question is about the convergence of the numerical method. Another question is about the stability of a numerical method. It does not matter if we use a single-step method or a multi-step method, errors in each step (including initial error and rounding error) can pass on and accumulate (i.e., accumulation of the global error). Only if the accumulation of errors is controlled then the numerical method is stable, otherwise, it is unstable. The problem of stability is related to the order of the method and the step size. The stability of a numerical method decides whether this method is applicable or not, therefore is of vital importance.

There are different definitions of stability. Some are about the method itself (we do not discuss this aspect in this book), and some are about the equation. The differential equation in our discussion is about a motion of a celestial body (including all types of spacecraft), but the corresponding Kepler motion is unstable. Therefore, no matter how stable a numerical method is, it cannot control the increase of the along-track errors (along the direction of the motion) step by step. If a numerical method does not make the increase of the along-track error worse, then this method is regarded relatively stable. In the orbital dynamics for orbital determination by long arcs, it is critical whether or not a numerical method can control the speed of the along-track error increase. This problem is discussed later.

In the theoretical study of the numerical method, convergence, stability, and error estimation are important topics. References [1, 2] provide more systematic materials for further understanding. In this chapter, we only introduce the commonly used numerical methods and related problems.

8.2 Conventional Singer-Step Method: The Runge–Kutta Method

One of the frequently used single-step methods is the well-known Runge–Kutta method (abbreviated as RK method). This method indirectly uses the Taylor series by forming a linear combination of a few values of the right-side function f on the interval $[t_n, t_{n+1}]$ to replace the derivatives of f , then by the Taylor expansion of f to decide the corresponding coefficients. This method avoids calculating high-order derivatives of the function f and assures accuracy.

We now take the commonly used fourth-order RK method as an example. Let

$$x_{n+1} = x_n + \sum_{i=1}^4 c_i k_i, \quad (8.10)$$

where the values of c_i are weight factors to be decided, and variables of k_i satisfy the following equations,

$$\begin{cases} k_i = hf \left(t_n + \alpha_i h, x_n + \sum_{j=1}^{i-1} \beta_{ij} k_j \right), \\ \alpha_1 = 0, \quad \beta_{ij} = 0, \quad i = 1, 2, 3, 4, \end{cases} \quad (8.11)$$

i.e.,

$$\begin{cases} k_1 = hf(t_n, x_n), \\ k_2 = hf(t_n + \alpha_2 h, x_n + \beta_{21} k_1), \\ k_3 = hf(t_n + \alpha_3 h, x_n + \beta_{31} k_1 + \beta_{32} k_2), \\ k_4 = hf(t_n + \alpha_4 h, x_n + \beta_{41} k_1 + \beta_{42} k_2 + \beta_{43} k_3), \end{cases} \quad (8.12)$$

To decide the coefficients α_i and β_{ij} we expand the right-side function f of (8.12) at (t_n, x_n) that

$$k_1 = hf_n, \quad k_2 = hf_n + h^2(\alpha_2 f'_{t_n} + \beta_{21} f'_{x_n} f_n) + \cdots, \cdots$$

Substituting these k_i into (8.10) and comparing the terms with the same order of h, h^2, h^3 , and h^4 in the Taylor expansion of $x(t_n + h)$ at t_n

$$x(t_n + h) = x(t_n) + hf(t_n, x(t_n)) + \frac{h^2}{2}(f'_{t_n}, f'_{x_n} f_n),$$

we obtain the relationships for deciding c_i and β_{ij} , which are

$$c_1 + c_2 = 1, \quad c_2 \alpha_2 = \frac{1}{2}, \quad c_2 \beta_{21} = \frac{1}{2}, \cdots \quad (8.13)$$

The selection of c_i , α_i and β_{ij} is not unique, because there are free parameters, thus different selections give different RK formulas.

The above RK formulas (8.10)–(8.11) are explicit formulas. Rearranging them gives the following forms

$$\begin{cases} x_{n+1} = x_n + \sum_{i=1}^m c_i k_i, \\ k_i = hf \left(t_n + \alpha_i h, x_n + \sum_{j=1}^m \beta_{ij} k_j \right), \end{cases} \quad (8.14)$$

which are the implicit RK formulas. The formula

$$k_i = hf \left(t_n + \alpha_i h, x_n + \sum_{j=1}^i \beta_{ij} k_j \right), \quad i = 1, 2, \dots, m \quad (8.15)$$

is a diagonal implicit RK formula. In orbit dynamics, we usually use the explicit RK formulas.

For the fourth-order RK method, the times of calculating function f in each forward step are consistent with the order of the method. But for a higher-order RK method, the times of calculating f are greater than the order of the method. If $N(m)$ is for the order of RK method given by calculating f for m times, then there are

$$\begin{cases} N(m) = m, & 1 \leq m \leq 4, \\ N(5) = 4, & N(6) = 5, & N(7) = 6, & N(8) = 6, & N(10) = 7, \dots \end{cases} \quad (8.16)$$

8.2.1 The Fourth-Order RK Method (RK4)

(1) Classical formula

$$x_{n+1} = x_n + \frac{1}{6}[k_1 + 2k_2 + 2k_3 + k_4], \quad (8.17)$$

$$\begin{cases} k_1 = hf(t_n, x_n), \\ k_2 = hf(t_n + \frac{1}{2}h, x_n + \frac{1}{2}k_1), \\ k_3 = hf(t_n + \frac{1}{2}h, x_n + \frac{1}{2}k_2), \\ k_4 = hf(t_n + h, x_n + k_3). \end{cases} \quad (8.18)$$

(2) Gill formula (can reduce rounding errors)

$$x_{n+1} = x_n + \frac{1}{6} \left[k_1 + (2 - \sqrt{2})k_2 + (2 + \sqrt{2})k_3 + k_4 \right], \quad (8.19)$$

$$\begin{cases} k_1 = hf(t_n, x_n), \\ k_2 = hf\left(t_n + \frac{1}{2}h, x_n + \frac{1}{2}k_1\right), \\ k_3 = hf\left(t_n + \frac{1}{2}h, x_n + \frac{\sqrt{2}-1}{2}k_1 + \left(1 - \frac{\sqrt{2}}{2}\right)k_2\right), \\ k_4 = hf\left(t_n + h, x_n - \frac{\sqrt{2}}{2}k_2 + \left(1 + \frac{\sqrt{2}}{2}\right)k_3\right). \end{cases} \quad (8.20)$$

The fourth-order RK is a widely used single-step method, which can be applied to problems of orbit dynamics if the required accuracy is not too high and the right-side functions are relatively simple.

8.2.2 The Runge–Kutta–Fehlberg (RKF) Method

One problem of using RK method is the difficulty to estimate the truncation error. To overcome this shortage Fehlberg developed an embedded method. In RK method, the parameters c_i , α_i , and β_{ij} can be chosen differently, by this property, after obtaining the m -order and $(m+1)$ -order RK formulas at the same time, we can get two values of x_{n+1} , and the difference between them can give the local truncation error. Based on the local truncation error the next step size can be decided automatically. This method is called Runge–Kutta–Fehlberg method (RKF method). [3] The embedded method uses the property of free parameter selection of RK method to achieve the embedding of the m -order formula and the $(m+1)$ -order formula. The difference between the m -order and $(m+1)$ -order formulas is small, the RKF method provides the local truncation error by calculating the right-side function a few times more than RK method (shown in (8.21)–(8.26)). The advantage of the RKF method makes it the most accepted single-step method.

Reference [3] of this chapter provides the embedded formulas of 5(6)-order, 6(7)-order, 7(8)-order, and 8(9)-order, denoted to RKF5(6), RKF6(7), RKF7(8), and RKF8(9), respectively. In the following content based on our experience gained from actual projects, we provide two sets of formulas of RKF5(6) and RKF7(8), which readers can use directly.

(1) RKF5(6) formulas

$$\begin{cases} x_{n+1} = x_n + h \sum_{i=0}^5 c_i f_i + O(h^6), \\ \hat{x}_{n+1} = x_n + h \sum_{i=0}^7 \hat{c}_i f_i + O(h^7), \end{cases} \quad (8.21)$$

Table 8.1 The coefficients in PKF5(6)

j	α_i	β_{ij}							c_i	$\hat{c}_i$
i		0	1	2	3	4	5	6		
0	0	0							31/384	7/1408
1	1/6	1/6								0
2	4/15	4/75	16/75							1125/2816
3	2/3	5/6	-8/3	5/2						9/32
4	4/5	-8/5	144/25	-4	16/25					125/768
5	1	361/320	-18/5	407/128	-11/80	55/128			5/66	0
6	0	-11/640	0	11/256	-11/160	11/256	0			5/66
7	1	93/640	-18/5	803/256	-11/160	99/256	0	1		5/66

$$\begin{cases} f_0 = f(t_n, x_n), \\ f_i = f\left(t_n + \alpha_i h, x_n + h \sum_{j=0}^{i-1} \beta_{ij} f_j\right), \quad i = 1, 2, \dots, 7. \end{cases} \quad (8.22)$$

The truncation error of the $n + 1$ step is

$$TE = \frac{5}{66}(f_0 + f_5 - f_6 - f_7)h. \quad (8.23)$$

The coefficients of c_i , α_i , β_{ij} , and $\hat{c}_i$ are listed in Table 8.1.

Formula (8.23) shows that for the Fifth formula only two more calculations of the right-side function, f_6 and f_7 , are needed.

(2) RKF7(8) formulas

$$\begin{cases} x_{n+1} = x_n + h \sum_{i=0}^{10} c_i f_i + O(h^8), \\ \hat{x}_{n+1} = x_n + h \sum_{i=0}^{12} \hat{c}_i f_i + O(h^9), \end{cases} \quad (8.24)$$

$$\begin{cases} f_0 = f(t_n, x_n), \\ f_i = f\left(t_n + \alpha_i h, x_n + h \sum_{j=0}^{i-1} \beta_{ij} f_j\right), \quad i = 1, 2, \dots, 12. \end{cases} \quad (8.25)$$

The truncation error of the $n + 1$ step is

$$TE = \frac{41}{840}(f_0 + f_{10} - f_{11} - f_{12})h. \quad (8.26)$$

The coefficients are listed in Table 8.2.

Table 8.2 The coefficients in PKF7(8)

α_i		β_{ij}											c_i	$\hat{c}_i$	
		0	1	2	3	4	5	6	7	8	9	10			11
j	i														
0	0	0												41/840	0
1	2/27	2/27												0	
2	1/9	1/36	1/12											0	
3	1/6	1/24	0	1/8										0	
4	5/12	5/12	0	-25/108	0	0	125/108	-65/27	125/54					0	
5	1/2	1/20	0	0	1/4	1/5								34/105	
6	5/6	-25/108	0	0	125/108	-65/27	125/54							9/35	
7	1/6	31/300	0	0	0	61/225	-2/9	13/900						9/35	
8	3/2	2	0	0	-53/6	704/45	-107/9	67/90	3					9/280	
9	1/3	-91/108	0	0	23/108	-976/135	311/54	-19/60	17/6	-1/12				9/280	
10	1	2383/4100	0	0	-341/164	4496/1025	-301/82	2133/4100	45/82	45/162	18/41			41/840	0
11	0	3/205	0	0	0	0	-6/41	-3/205	-3/41	3/41	6/41	0			41/840
12	1	-1777/4100	0	0	-341/164	4496/1025	-289/82	2193/4100	51/82	33/164	12/41	0	1		41/840

Formula (8.26) shows that for the Seventh formula also only two more calculations of the right-side function, f_{11} and f_{12} , are needed.

8.3 Linear Multistep Methods: Adams Method and Cowell Method

For a differential equation with a given initial value (8.1), the general linear multistep formula is

$$\begin{aligned} & \alpha_k x_{n+k} + \alpha_{k-1} x_{n+k-1} + \cdots + \alpha_0 x_n \\ &= h(\beta_k f_{n+k} + \beta_{k-1} f_{n+k-1} + \cdots + \beta_0 f_n), \end{aligned} \quad (8.27)$$

where α_i and β_i ($i = 0, 1, 2, \dots, k$) are constants unrelated to n ; $f_i = f(t_i, x_i)$. Generally $\alpha_k = 1, |\alpha_0| + |\beta_0| > 0$ (i.e., α_0 and β_0 do not equal zero at the same time). This formula shows that to calculate the value of x_{n+k} at t_{n+k} we need to know each value of x_i in the previous steps at $t_{n+k-1}, t_{n+k-2}, \dots, t_n$. The method defined by (8.27) is called the k -step method; when $k = 1$, it is a singer-step method, and when $k > 1$ it is a multistep method. Also because the formula x_i and f_i are in linear forms, it is also called the linear multistep method. When $\beta_k = 0$, the formula is explicit, otherwise implicit.

The characteristic polynomial of the k -step method

$$\rho(\xi) = \alpha_k \xi^k + \alpha_{k-1} \xi^{k-1} + \cdots + \alpha_0 \quad (8.28)$$

is very important. If the right-side function of (8.1) $f(t, x)$ is continuous on the intervals of $a \leq t \leq b$ and $-\infty < x < \infty$, and x satisfies the Lipschitz condition, then the necessary and sufficient condition of the k -step formula (8.27) being stable for all $f(t, x)$ is that the k -step formula satisfies the condition of eigenvalues, i.e., the roots of the characteristic polynomial (8.28) all have modulus less than or equal to 1, and the roots of modulus 1 are single roots [1].

In the following content, we introduce a few commonly used calculation formulas of multistep methods.

8.3.1 Adams Methods: Explicit Methods and Implicit Methods

The Adams explicit methods are also called the Adams–Bashforth methods, and the Adams implicit methods are also called the Adams–Moulton methods.

(1) Adams explicit methods

Integrating the differential Eq. (8.1) from t_n to t_{n+1} yields an equivalent integration equation

$$x(t_{n+1}) = x(t_n) + \int_{t_n}^{t_{n+1}} f(t, x(t)) dt. \quad (8.29)$$

Then using an interpolation polynomial to discrete the right-side integrand and to obtain numerical formulas. If we use the Newton backward difference format, which is given by

$$\nabla^m f_n = \sum_{l=0}^m (-1)^l \binom{m}{l} f_{n-l}, \quad (8.30)$$

where the symbol ∇ is for the backward difference calculation that

$$\begin{cases} \nabla f_n = \nabla f(x_n) = f(x_n) - f(x_n - h), \\ \nabla^2 f_n = \nabla f(x_n) - \nabla f(x_n - h), \\ \quad = f(x_n) - 2f(x_n - h) + f(x_n - 2h), \\ \dots \end{cases} \quad (8.31)$$

Then the corresponding backward interpolation polynomial is

$$p(t) = \sum_{m=0}^{k-1} (-1)^m \binom{-s}{m} \nabla^m f_n, \quad (8.32)$$

where k is the number of interpolation points, and s is the assistant variable defined as

$$s = \frac{t-t_0}{h}, \quad (8.33)$$

and there is

$$s+1 = \frac{t-t_0}{h} + 1 = \frac{t-t_{n-1}}{h}, \quad \dots, \quad s+m-1 = \frac{t-t_{n+m-1}}{h}. \quad (8.34)$$

In (8.32) $\binom{-s}{m}$ are the general binomial coefficients defined by

$$\binom{-s}{m} = (-1)^m \binom{s+m-1}{m}. \quad (8.35)$$

Substituting the interpolation polynomial $p(t)$ into (8.29) leads to

$$x_{n+1} = x_n + h \sum_{m=0}^{k-1} \gamma_m \nabla^m f_n, \quad (8.36)$$

where

$$\gamma_m = \int_{t_n}^{t_{n+1}} \frac{1}{h} (-1)^m \binom{-s}{m} dt = \int_0^1 \binom{s+m-1}{m} ds. \quad (8.37)$$

Omitting the process of derivation, the recursive relationship for γ_m is

$$\gamma_m + \frac{1}{2} \gamma_{m-1} + \frac{1}{3} \gamma_{m-2} + \cdots + \frac{1}{m+1} \gamma_0 = 1. \quad (8.38)$$

Substituting the backward difference format (8.30) into (8.36) leads to the calculation formula expressed by the values of the right-side function

$$x_{n+1} = x_n + h \sum_{m=0}^{k-1} \beta_{kl} f_{n-1}, \quad k = 1, 2, \dots, \quad (8.39)$$

$$\beta_{kl} = (-1)^l \sum_{m=l}^{k-1} \binom{m}{l} \gamma_m = (-1)^l \left[\binom{l}{l} \gamma_l + \binom{l+1}{l} \gamma_{l+1} + \cdots + \binom{k-1}{l} \gamma_{k-1} \right]. \quad (8.40)$$

Formulas (8.36) and (8.39) are the Adams explicit formulas of k -step given by the backward difference format and the values of the right-side function, respectively, and to obtain x_{k+1} requires values of $x_{n-k+1}, x_{n-k+2}, \dots, x_n$ at k steps $t_{n-k+1}, t_{n-k+2}, \dots, t_n$.

(2) Adams implicit methods

Integrating the differential Eq. (8.1) from t_{n-1} to t_n yields

$$x(t_n) - x(t_{n-1}) = \int_{t_{n-1}}^{t_n} f(t, x(t)) dt. \quad (8.41)$$

Similarly using the backward difference of f_n for interpolation leads

$$x_{n+1} - x_n = h \sum_{m=0}^{k-1} \gamma_m^* \nabla^m f_n. \quad (8.42)$$

Following the process of the explicit formulas yields the Adams implicit methods as

$$x_{n+1} - x_n = h \sum_{m=0}^{k-1} \beta_{kl}^* f_{n+1-l}, \quad (8.43)$$

$$\beta_{kl}^* = (-1)^l \sum_{m=l}^{k-1} \binom{m}{l} \gamma_m^*, \quad (8.44)$$

where γ_m^* satisfies the following recursive relationship

$$\gamma_m^* + \frac{1}{2}\gamma_{m-1}^* + \frac{1}{3}\gamma_{m-2}^* + \cdots + \frac{1}{m+1}\gamma_0^* = \begin{cases} 1, & m = 0 \\ 0, & m \neq 0 \end{cases} \quad (8.45)$$

The relationship between $\gamma_0^*, \gamma_1^*, \dots, \gamma_m^*$ and γ_m is

$$\sum_{i=0}^m \gamma_i^* = \gamma_m, \quad m = 0, 1, 2, \dots \quad (8.46)$$

The Adams explicit method and implicit method are often combined together, i.e., first by the Adams explicit method to obtain an approximation $x_{n+1}^{(0)}$, which is the predicted evaluation (PE), then by the implicit method to correct and evaluate (CE) to provide the needed x_{n+1} . This combined method is called the PECE technique. For a perturbed two-body problem we can separate the right-side function into two parts according to the characteristics of the motion equation, the main part (i.e., the gravitational acceleration due to the central body) and a second part (i.e., the perturbed acceleration), the former part is the main “content” to be corrected in the PECE process.

The calculations of the coefficients $\gamma_m, \gamma_m^*, \beta_{kl}$ and β_{kl}^* are easy by computer. Here we list the values of γ_m and γ_m^* ($m = 1, 2, \dots, 11$) in Table 8.3.

8.3.2 Cowell's Method and Størmer's Method

Cowell's method was developed by Cowell and Crommelin in 1910 for predicting the return of Halley's Comet. [4] This method is a linear multistep numerical method for solving a second-order differential equation with initial values. The equation and initial conditions have the following form as

$$\begin{cases} \ddot{x} = f(x, t), \\ x(t_0) = x_0, \quad \dot{x}(t_0) = \dot{x}_0. \end{cases} \quad (8.47)$$

One of the properties of this equation is that the right-side function f does not contain the velocity $\dot{x}$. A motion equation for a celestial body is like that if the motion is only connected to gravitational forces. It is why Cowell's method is a commonly used numerical method for solving problems related to Celestial Mechanics. By this method in each step x_n is directly calculated without calculating $\dot{x}_n$, therefore it is simpler than Adams methods as described in the previous section.

Table 8.3 Values of γ_m and γ_m^*

m	0	1	2	3	4	5
γ_m	1	1/2	5/12	3/8	251/720	95/288
γ_m^*	1	-1/2	-1/12	-1/24	-19/720	-3/160
m	6	7	8	9	10	11
γ_m	19087/60480	36799/120960	1070017/3628800	25713/89600	26842253/95800320	4777223/17418240
γ_m^*	-863/60480	-1375/120960	-33953/3628800	-8183/1036800	-3250433/479001600	-4671/788480

The general linear multistep formula for the Eq. (8.47) with initial values is different from (8.27), and is given by

$$\alpha_k x_{n+k} + \alpha_{k-1} x_{n+k-1} + \cdots + \alpha_0 x_n = h^2 (\beta_k f_{n+k} + \beta_{k-1} f_{n+k-1} + \cdots + \beta_0 f_n) \quad (8.48)$$

or written as

$$\sum_{i=0}^k \alpha_i x_{n+i} = h^2 \sum_{i=0}^k \beta_i f_{n+i}. \quad (8.49)$$

Usually $\alpha_k = 1$ and $|\alpha_0| + |\beta_0| > 0$.

(1) Størmer's explicit method

Integrating the differential Eq. (8.47) yields

$$\dot{x}(t) = \dot{x}(t_n) + \int_{t_n}^t f(t, x(t)) dt. \quad (8.50)$$

The next step is to integrate the left side from t_n to t_{n+1} and to integrate the right side from t_n to t_{n+1} , then $\dot{x}(t_n)$ can be eliminated in the results, and we have the following equation of integration:

$$\begin{aligned} x(t_{n+1}) - 2x(t_n) + x(t_{n-1}) \\ = \int_{t_n}^{t_{n+1}} \int_{t_0}^t f(t, x(t)) dt^2 + \int_{t_n}^{t_{n-1}} \int_{t_0}^t f(t, x(t)) dt^2 \end{aligned} \quad (8.51)$$

The integrand can be replaced by an interpolation polynomial, and we obtain a discrete numerical formula. The process of derivation is similar to that of the Adams explicit methods, the result is

$$x(t_{n+1}) - 2x(t_n) + x(t_{n-1}) = h^2 \sum_{m=0}^{k-1} \sigma_m \nabla^m f_n, \quad (8.52)$$

where

$$\begin{cases} \sigma_m = (-1)^m \int_0^1 (1-s) \left[\binom{-s}{m} + \binom{s}{m} \right] ds \\ s = \frac{t - t_0}{h} \end{cases} \quad (8.53)$$

Integrating (8.53) we obtain the recursive relationship of the coefficient σ_m : [1]

$$\begin{cases} \sigma_0 = 1 \\ \sigma_m = 1 - \frac{3}{2}h_2\sigma_{m-1} - \frac{2}{4}h_3\sigma_{m-2} - \cdots - \frac{2}{m+2}h_{m+1}\sigma_0 \\ \quad = 1 - \sum_{i=1}^m \left(\frac{2}{i+2}h_{i+1} \right) \sigma_{m-1}, \quad m = 1, 2, \dots \end{cases} \quad (8.54)$$

where h_i is the sum of the first i terms of the harmonic series that

$$h_i = 1 + \frac{1}{2} + \cdots + \frac{1}{i}. \quad (8.55)$$

Using the backward interpolation polynomial (8.32), we can express (8.52) by function f as

$$x_{n+1} - 2x_n + x_{n-1} = h^2 \sum_{l=0}^{k-1} \alpha_{kl} f_{n-1}, \quad k = 1, 2, \dots, \quad (8.56)$$

where

$$\begin{aligned} \alpha_{kl} &= (-1)^l \sum_{m=l}^{k-1} \sigma_m \binom{m}{l} \\ &= (-1)^l \left[\binom{l}{l} \sigma_l + \binom{l+1}{l} \sigma_{l+1} + \cdots + \binom{k-1}{l} \sigma_{k-1} \right] \end{aligned} \quad (8.57)$$

Formulas (8.52) and (8.56) are Størmer's explicit method.

(2) Cowell's implicit method

By the same procedure used to derive the Adams implicit methods, it is easy to obtain that [1]

$$x_n - 2x_{n-1} + x_{n-2} = h^2 \sum_{m=0}^{k-1} \sigma_m^* \nabla^m f_n, \quad (8.58)$$

$$\sigma_m^* = (-1)^m \int_{-1}^0 (1-s) \left[\binom{-s}{m} + \binom{s+2}{m} \right] ds, \quad (8.59)$$

where s is the same as given in (8.33), and σ_m^* and σ_m have similar recursive relationship as

$$\begin{cases} \sigma_0^* = 1 \\ \sigma_m^* = -\frac{3}{2}h_2\sigma_{m-1}^* - \frac{2}{4}h_3\sigma_{m-2}^* - \cdots - \frac{2}{m+2}h_{m+1}\sigma_0^* \\ = 1 - \sum_{i=1}^m \left(\frac{2}{i+2}h_{i+1} \right) \sigma_{m-1}^*, \quad m = 1, 2, \dots \end{cases} \quad (8.60)$$

also, there is

$$\sigma_m^* = \sigma_m - \sigma_{m-1}. \quad (8.61)$$

The corresponding Cowell implicit method expressed by function f is (use the same subscript as for the explicit methods)

$$x_{n+1} - 2x_n + x_{n-1} = h^2 \sum_{l=0}^{k-1} \alpha_{kl}^* f_{n+1-l}, \quad k = 1, 2, \dots, \quad (8.62)$$

where

$$\alpha_{kl}^* = (-1)^l \sum_{m=l}^{k-1} \binom{m}{l} \sigma_m^* = (-1)^l \left[\binom{l}{l} \sigma_l^* + \binom{l+1}{l} \sigma_{l+1}^* + \cdots + \binom{k-1}{l} \sigma_{k-1}^* \right]. \quad (8.63)$$

A complete Cowell method is a predictor-correct method by combining Størmer's explicit method and Cowell's implicit method, which is also called the first Cowell method. Late Cowell and Crommelin developed a method using the second-order integration, which is called the second Cowell method. Because the second one does not show obvious advancement over the first one, the first one is still frequently used in the community and is called the Cowell method. It should be remembered that the right-side motion function of (8.47) does not contain $\dot{x}$, this method is for solving the motion of a natural celestial body which depends only on the gravitational forces. For the motion of artificial satellites, particularly the low Earth orbit motion, the effect of atmospheric drag cannot be ignored, and in Sect. 8.3.3 we discuss how to apply Cowell's method to this situation.

The coefficients σ_m and σ_m^* ($m = 0, 1, \dots, 11$) in Cowell's method are listed in Table 8.4 which are in need of practical applications.

8.3.3 Adams-Cowell Method

The local truncation error of the k -order (i.e., k -step) Adams method is $O(h^{k+1})$, it is $O(h^{k+2})$ by the Cowell method, therefore by the same order of formulas the Cowell

Table 8.4 The coefficients σ_m and σ_m^*

m	0	1	2	3	4	5
σ_m	1	0	1/12	1/12	19/240	3/40
σ_m^*	1	-1	1/12	0	-1/240	-1/240
m	6	7	8	9	10	11
σ_m	863/12096	825/12096	237671/3628800	229124/3628800	3250433/53222400	60723/1025024
σ_m^*	-221/60480	-19/6048	-9829/3628800	-8547/3628800	-330157/159667200	-24377/13305600

method is more accurate than the Adams method. But with the development of the orbital dynamics, more factors need to be considered, and $\dot{x}$ often appeared in the right-side function; also even for pure gravitational problems, sometimes $\dot{x}$ has to be calculated, such as in calculating ephemerides for perturbations of artificial Earth satellites. Thus, it is necessary to combine the Adams method with the Cowell method to deal with these situations.

The following differential equation with initial values

$$\begin{cases} \ddot{x} = f(x, \dot{x}, t), \\ x(t_0) = x_0, \quad \dot{x}(t_0) = \dot{x}_0 \end{cases} \quad (8.64)$$

can be treated as a first-order differential equation of $\dot{x}$. We can use the Adams method to provide $\dot{x}$, at the same time use the Cowell method to calculate x . This combined method is called the Adams-Cowell method, which is more efficient than the Adams method, and is often used to make precision ephemerides for celestial bodies, natural and artificial. In practice, the explicit method and implicit method are combined to perform the predictor-correct procedure.

Our experience can be summarized as follows.

- (1) The Adams-Cowell method is obviously superior to the Adams method, and the predict-correct method is better than the predictor only.
- (2) In order to keep the higher-order multistep method stable it is necessary to limit the step size.
- (3) For the AC-PECE method (Adams-Cowell predict-evaluate-correct-evaluate) the accuracy does not simply increase when the order of the method increases, it is also affected by the stability (i.e., the problem of error increase). For a perturbation problem, usually the right order is 12–14.

8.4 Key Issues in Applications of the Numerical Method in Orbital Dynamics

8.4.1 Selections of Variables and Corresponding Basic Equations

In solving orbital motion equations the question of whether or not the basic variables are selected properly can affect the efficiency of the entire calculation. For a perturbed two-body problem there are three types of variables to be chosen as follows.

① Position vector $\vec{r}$ and velocity $\dot{\vec{r}}$ of a moving body. These two variables are commonly used not only for a perturbed two-body problem but also for a general N -body problem ($N \geq 3$) and a restricted three-body problem.

② Perturbed position vector $\vec{u}$ and related velocity $\dot{\vec{u}}$ of a moving body defined by

$$\vec{u} = \vec{r} - \vec{r}_c, \quad \dot{\vec{u}} = \dot{\vec{r}} - \dot{\vec{r}}_c, \quad (8.65)$$

where $\vec{r}_c$ and $\dot{\vec{r}}_c$ are the position vector and velocity of the moving body related to a reference orbit, which is usually the non-perturbed orbit, respectively.

③ The orbital elements $\sigma(a, e, i, \dots)$. The six orbital elements can directly replace the perturbed position vector and velocity of the moving body in ②.

Type 1: using $\vec{r}$ and $\dot{\vec{r}}$ as the basic variables

The differential equation with initial values is given by

$$\begin{cases} \ddot{\vec{r}} = \vec{F}(\vec{r}, \dot{\vec{r}}, t), \\ \vec{r}(t_0) = \vec{r}_0, \quad \dot{\vec{r}}(t_0) = \dot{\vec{r}}_0. \end{cases} \quad (8.66)$$

For a general motion problem, the variables, $\vec{r}, \dot{\vec{r}}, \ddot{\vec{r}}$, are the position, velocity, and acceleration of the celestial body. For a perturbed two-body problem, the right-side function $\vec{F}$ can be separated into two parts that

$$\vec{F} = \vec{F}_0(r) + \vec{F}_1(\vec{r}, \dot{\vec{r}}, t; \varepsilon), \quad \varepsilon \ll 1, \quad (8.67)$$

$$\vec{F}_0 = -G(m_0 + m) \frac{\vec{r}}{r^3}, \quad (8.68)$$

where $\vec{F}_0$ is the acceleration due to the central gravity force; m_0 and m are the masses of the central body and the moving body, respectively; $\vec{F}_1$ is the perturbed acceleration.

The reasons for selecting $\vec{r}$ and $\dot{\vec{r}}$ as the basic variables are that firstly they are not restricted by the type of motion; secondly the form of the right-side function $\vec{F}$ is simple, and the efficiency of calculation often depends on the form of the right-side function. The shortage of this selection for a perturbed two-body problem is that the right-side function includes the unperturbed part, which changes fast over time, for certain accuracy the step size of integration must be restricted to a small value.

Type 2: using $\vec{u}$ and $\dot{\vec{u}}$ as the basic variables

The differential equation with initial values is now given by

$$\begin{cases} \ddot{\vec{u}} = \ddot{\vec{r}} - \ddot{\vec{r}}_c = -\left(\mu \frac{\vec{r}}{r^3} + \ddot{\vec{r}}_c\right) + \vec{F}_1(\vec{u}, \dot{\vec{u}}, \vec{r}_c, \dot{\vec{r}}_c, t; \varepsilon), \\ \vec{u}(t_0) = \vec{u}_0, \quad \dot{\vec{u}}(t_0) = \dot{\vec{u}}_0, \end{cases} \quad (8.69)$$

where $\mu = G(m_0 + m)$, and the initial conditions of $\vec{u}_0$ and $\dot{\vec{u}}_0$ are

$$\vec{u}_0 = \vec{r}_0 - \vec{r}_c(t_0), \quad \dot{\vec{u}}_0 = \dot{\vec{r}}_0 - \dot{\vec{r}}_c(t_0). \quad (8.70)$$

The reference orbit is a fixed ellipse, which satisfies the motion equation $\ddot{\vec{r}}_c = -\mu \vec{r}_c / r_c^3$. This motion equation has strict solutions of $\vec{r}_c(t)$ and $\dot{\vec{r}}_c(t)$ with initial values $\vec{u}_0 = 0$ and $\dot{\vec{u}}_0 = 0$. The perturbed variables $\vec{u}$ and $\dot{\vec{u}}$ can be given by numerical

calculation, then by definitions to give $\vec{r}(t)$ and $\dot{\vec{r}}(t)$. This method is called Encke's method [5].

For the same accuracy, the integrating step size can be larger by selecting this set of variables than by selecting the first type of variables. But for each step, the calculation of the ephemeris of the reference orbit is required. Furthermore, when the perturbation varies fast, it requires corrections of the reference orbit section by section, i.e., repeating the initialization constantly. In reality for numerical methods, the value of the right-side function is important but the speed of its variation cannot be ignored. Although by Encke's method the direct result is the perturbed position $\vec{u}$ (may call it by Encke's vector), but because of perturbations, the deviation of the actual orbit from the reference orbit becomes larger and larger, eventually the magnitude of $|\vec{u}|$ can reach $O(r)$, then the efficiency of calculation $\vec{u}$ and $\dot{\vec{u}}$ is not different from directly calculating $\vec{r}$ and $\dot{\vec{r}}$. Therefore, this method cannot solve the long-arc calculation problem by providing high precision results. To overcome this problem, there are a few different methods for improving the reference orbit when calculating $\vec{u}$ and $\dot{\vec{u}}$, mainly to make the reference orbit include the main perturbation effects (similar to an intermediate orbit) and at the same time not to make the calculation too complicated. [5–8] For example, for the orbit of an Earth's satellite, the reference orbit can include the main part of the Earth's non-spherical gravity perturbation, ideally the first-order secular terms due to the J_2 term (as given in Sect. 4.2, the formulas are not complicated), so the largest term in $\vec{F}_1$ is removed, and the calculation does not become too difficult.

The above-described method of adjusting the reference orbit is called the modified perturbation coordinate method or modified Encke's method. Obviously, when the calculating arc is not too long, the results are better than the original method, but when the arc is long, the influence of the perturbation coordinate variations due to higher-order perturbation terms can be shown in the right-side function, and the calculating "zero" point has to be changed section by section (i.e., correcting the reference orbit). To avoid repeatedly calculating the reference orbit we can directly use the orbit elements as the basic variables.

Type 3: using Kepler orbital elements as the basic variables

The differential equation system with initial values is given by

$$\begin{cases} \dot{\sigma} = f(\sigma, t, \varepsilon), & \varepsilon \ll 1, \\ \sigma(t_0) = \sigma_0. \end{cases} \quad (8.71)$$

The basic variables in the equation are the six commonly used Kepler elements a , e , i , Ω , ω , and M , and the perturbation equation is in a general form. If the right-side function f is presented by the orbit elements in the form of $f(\sigma, t, \varepsilon)$, the expression is complicated, especially if there are multiple perturbations. In order to apply the numerical method, the forms of the right-side functions for this system of basic variables are given by

$$\left\{ \begin{aligned} \frac{da}{dt} &= \frac{2}{n\sqrt{1-e^2}} [Se \sin f + T(1 + e \cos f)] \\ \frac{de}{dt} &= \frac{\sqrt{1-e^2}}{na} [S \sin f + T(\cos f + \cos E)] \\ \frac{di}{dt} &= \frac{r \cos u}{na^2 \sqrt{1-e^2}} W \\ \frac{d\Omega}{dt} &= \frac{r \sin u}{na^2 \sqrt{1-e^2} \sin i} W \\ \frac{d\omega}{dt} &= \frac{\sqrt{1-e^2}}{nae} \left[-S \cos f + T \left(1 + \frac{r}{p} \right) \sin f \right] - \cos i \frac{d\Omega}{dt} \\ \frac{dM}{dt} &= n - \frac{1-e^2}{nae} \left[-S \left(\cos f - 2e \frac{r}{p} \right) + T \left(1 + \frac{r}{p} \right) \sin f \right] \end{aligned} \right. \quad (8.72)$$

where $u = f + \omega$, $p = a(1 - e^2)$; f and E are the true anomaly and the eccentric anomaly, respectively; and S , T , and W are given in the following content.

Type 4: using the eccentric anomaly as the time element

If we use M as the time element then to calculate the right-side function, we have to solve the Kepler equation, which is complicated. This problem can be avoided if the time element M is replaced by E , the eccentric anomaly, and the six orbit elements are a , e , i , Ω , ω , and E . The corresponding right-side function $f(\sigma, t, \varepsilon)$ is given by $\vec{F}_1(\vec{r}, \dot{\vec{r}}, t; \varepsilon)$ in (8.67), then the actual forms are

$$\left\{ \begin{aligned} \frac{dE}{dt} &= \frac{a}{r} \left[n - \sqrt{1-e^2} \left(\frac{d\omega}{dt} \right)_1 + \sin E \frac{de}{dt} \right] - \frac{2}{na} S, \\ \left(\frac{d\omega}{dt} \right)_1 &= \frac{1}{nae} \left[\sqrt{1-e^2} (-S \cos f + T \sin f) + T \sin E \right], \end{aligned} \right. \quad (8.73)$$

where $n = \sqrt{\mu}a^{-3/2}$, $\mu = G(m_0 + m)$, $\sin f$ and $\cos f$ can be given by $\sin E$ and $\cos E$ that

$$\begin{cases} r \sin f = a\sqrt{1-e^2} \sin E, & r \cos f = a(\cos E - e), \\ r = a(1 - e \cos E). \end{cases} \quad (8.74)$$

The last problem is to express S , T , and W by $\vec{F}_1$, and the relationships are

$$S = \vec{F}_1 \cdot \hat{r}, \quad T = \vec{F}_1 \cdot \hat{t}, \quad W = \vec{F}_1 \cdot \hat{w}, \quad (8.75)$$

where $\hat{r}$, $\hat{t}$, $\hat{w}$ are unit vectors in the directions of radial, transverse, and normal, respectively, that

$$\hat{r} = \cos u \hat{P}^* + \sin u \hat{Q}^*, \quad \hat{t} = -\sin u \hat{P}^* + \cos u \hat{Q}^*, \quad \hat{w} = \hat{r} \times \hat{t}, \quad (8.76)$$

where $u = f + \omega$. For the convenience of calculation, the unit vectors $\hat{P}^*$ and $\hat{Q}^*$ are defined as

$$\hat{P}^* = \begin{pmatrix} \cos \Omega \\ \sin \Omega \\ 0 \end{pmatrix}, \quad \hat{Q}^* = \begin{pmatrix} -\sin \Omega \cos i \\ \cos \Omega \cos i \\ \sin i \end{pmatrix}, \quad (8.77)$$

$$\hat{r} \times \hat{t} = \hat{P}^* \times \hat{Q}^* = \begin{pmatrix} \sin \Omega \sin i \\ -\cos \Omega \cos i \\ \cos i \end{pmatrix}. \quad (8.78)$$

Now $\vec{r}$ and $\dot{\vec{r}}$ are given by

$$\vec{r} = r\hat{r}, \quad \dot{\vec{r}} = \frac{\sqrt{\mu a(1-e^2)}}{r} \left[\left(\frac{e}{\sqrt{1-e^2}} \sin E \right) \hat{r} + \hat{t} \right]. \quad (8.79)$$

The above expressions show that when we use the orbit elements as the basic variables the calculations of the right-side function are not complicated. The complicity is in the perturbed acceleration $\vec{F}_1$ when there are several perturbations. Comparing the calculations between using Type 1 and Type 3 variables, the difference is small, the only extra calculation is the transformation from $\vec{F}_1$ to S , T , and W , which is a small portion of the entire calculation. The advantage of using the orbit elements as the basic variables is the step size of integration can be relatively large then the calculation can be reduced. Using Type 3 variables if choose the same step size as for Type 1 variables the local truncation error is obviously less and the accuracy increases. The conclusion is that when the perturbed acceleration $\vec{F}_1$ is complicated, choosing the orbit elements as the basic variables has practical values.

8.4.2 Singularity Problem

In the case of $e \approx 0$, or both $e \approx 0$ and $i \approx 0$, the basic variables of Kepler elements can be replaced by the non-singularity elements of the first-type or the second-type, as defined in Sect. 3.2.4.

8.4.3 Homogenization of Step-Size

In the N -body problem ($N \geq 3$), there is an essential singularity, i.e., a collision singularity. For a single collision, it can be eliminated by normalization, [9–12] but this collision means a collision of two particles, which only exists in theory. In the real situation, the problem is that two celestial bodies become very close or almost crash. For example, in a restricted three-body problem, when the small body is very near one of the two main bodies; also in a perturbed two-body problem when the

moving body is near the perigee of a high eccentricity orbit. In both cases, because the moving body is close to the main body its acceleration increases suddenly, although a collision does not actually happen, from the point of view of numerical calculation the integrating step size has to be changed fast. The method to solve this problem is by a transformation of the independent variable t , as a result, the integrating step size can be adjusted automatically. For the perturbed two-body problem the transformation is given by

$$\frac{dt}{ds} = r^p, \quad p \geq 1, \quad (8.80)$$

where s is the new independent variable, p is an adjustable parameter. By (8.80) when the step-size for s is a constant, the original independent variable time t has a varying step size depending on r^p . This independent variable transformation method (8.80) changes the time scale to make the step-size homogenized in a normalization method. This method attracts beneficial discussions, especially about how to choose the parameter p . The situation in a single-step method is relatively easy to deal with, but is complicated in a multistep method. More information about this topic is given in the related references.

8.4.4 Control of the Along-Track Errors

Another problem in numerically solving the motion equation for a celestial body (and a spacecraft of any type), besides the selection of basic variables and the transformation of step size for a perturbed two-body problem with high eccentricity, is the error accumulation. For a short-arc calculation, the accuracy mainly depends on the truncation errors, no matter whether it is a single-step method or a multistep method, as long the order of the method is relatively high and the integrating step-size is relatively short, the accuracy can be achieved. But for “tracking” calculations of orbital evolution over a long time period or calculations of a long-arc orbital determination, the error accumulation is a serious problem. Particularly in almost all traditional numerical methods, there is the artificial dissipation problem, which makes the energy error (or the semi-major axis a) increase “linearly”, therefore, distorting the motion characteristics and failing the calculation, or increasing the along-track errors fast and eventually the calculation cannot achieve the required accuracy.

About the calculation accuracy from the quantitative point of view, the author has developed a method with practical value for a conservative system (or a system with very small dissipation), which increases the calculation accuracy over a limited arc. The main idea of the method is that by using the energy integral (or the energy relationship) through adjusting the semi-major axis a , the energy dissipation can be compensated, therefore, the increase of the along-track error can be controlled. By

this method, the increase of the along-track error, $\Delta(M + \omega)$, can be controlled to be linear with $(t - t_0)$ [13, 14].

8.5 Numerical Calculation of the Right-Side Function

By numerical methods to solve the motion equation, we need to calculate the right-side functions in the form of the position vector and velocity. Take an Earth's satellite as an example, the right-side functions are the accelerations related to different forces. Except for the accelerations due to Earth's non-spherical gravity potential and the related deformation, the other accelerations in the forms of position vector and velocity are provided in Chap. 4. In this section, we provide the acceleration due to Earth's non-spherical gravity potential, which mainly deals with Legendre polynomials and their derivatives.

In the Earth-fixed coordinate system, the gravity potential of Earth's non-spherical perturbation ΔV (i.e., the perturbation function originally denoted to R) is given by

$$\begin{aligned}\Delta V &= \sum_{l,m} \Delta V_{l,m} \\ &= \frac{GE}{R} \sum_{l \geq 2} \sum_{m=0}^l \left(\frac{a_e}{R}\right)^l P_{l,m}(\sin \varphi) [C_{l,m} \cos m\lambda_G + S_{l,m} \sin m\lambda_G],\end{aligned}\quad (8.81)$$

where GE is Earth's gravitational constant; a_e is the equatorial radius of Earth's reference ellipsoid; $C_{l,m}$ and $S_{l,m}$ are the supporting harmonic coefficients of the ellipse, which represent the unevenness of Earth's gravity potential as discussed in Chap. 4. The variables, R , φ , and λ_G , are the spherical coordinates of a satellite in the Earth-fixed frame, which are the distance to the center of Earth, the geocentric latitude, and the longitude measured from the Greenwich meridian, respectively.

Applying the dimensionless units to this system, the associated Legendre polynomials and the harmonic coefficients $C_{l,m}$ and $S_{l,m}$ are in the normalized forms, and the normalized form of ΔV is given by

$$\Delta V = \sum_{l \geq 2} \sum_{m=0}^l \left(\frac{1}{R}\right)^{l+1} \bar{P}_{l,m}(\sin \varphi) [\bar{C}_{l,m} \cos m\lambda_G + \bar{S}_{l,m} \sin m\lambda_G], \quad (8.82)$$

where

$$\begin{cases} \bar{P}_{l,m}(\mu) = \frac{P_{l,m}(\mu)}{N_{l,m}}, \\ \bar{C}_{l,m} = C_{l,m} \cdot N_{l,m}, \bar{S}_{l,m} = S_{l,m} \cdot N_{l,m}, \end{cases} \quad (8.83)$$

$$\begin{cases} N_{l,m} = \left[\left(\frac{1}{1+\delta} \right) \frac{(l+m)!}{(2l+1)(l-m)!} \right]^{\frac{1}{2}}, \\ \delta = \begin{cases} 0, & m = 0, \\ 1, & m \neq 0. \end{cases} \end{cases} \quad (8.84)$$

For the convenience of calculation and analysis, we separate ΔV into two parts, the zonal harmonic part and the tesseral harmonic part that

$$\Delta V = \Delta V_l + \Delta V_{l,m}, \quad (8.85)$$

$$\Delta V_l = \sum_{l \geq 2} \bar{C}_{l,0} \left(\frac{1}{R} \right)^{l+1} \bar{P}_l(\sin \varphi), \quad (8.86)$$

$$\Delta V_{l,m} = \sum_{l \geq 2} \sum_{m=0}^l \left(\frac{1}{R} \right)^{l+1} \bar{P}_{l,m}(\sin \varphi) [\bar{C}_{l,m} \cos m\lambda_G + \bar{S}_{l,m} \sin m\lambda_G]. \quad (8.87)$$

The values of $\cos m\lambda_G$ and $\sin m\lambda_G$ can be given by a simple recursive method, that first calculates $\cos \lambda_G$ and $\sin \lambda_G$ then directly by the following recursive formulas to calculate $\cos m\lambda_G$ and $\sin m\lambda_G$ ($m \geq 2$):

$$\begin{cases} \cos m\lambda_G = 2 \cos \lambda_G \cos(m-1)\lambda_G - \cos(m-2)\lambda_G, \\ \sin m\lambda_G = 2 \cos \lambda_G \sin(m-1)\lambda_G - \sin(m-2)\lambda_G, \end{cases} \quad m \geq 2 \quad (8.88)$$

The values of $\sin \varphi$, $\cos \lambda_G$, and $\sin \lambda_G$ are given in the Earth-fixed coordinate system, whereas the motion equation and solution of the satellite are presented in the epoch J2000.0 geocentric celestial coordinate system, therefore we need to deal with the transformation of coordinate systems. In the transformation, we need the expressions of $\sin \varphi$, $\cos \lambda_G$, and $\sin \lambda_G$, and related partial derivatives, which are

$$\sin \varphi = \frac{Z}{R}, \quad (8.89)$$

$$\left[\frac{\partial(\sin \varphi)}{\partial \vec{R}} \right]^T = -\frac{Z}{R^3} \vec{R} + \frac{1}{R} \vec{k}, \quad \vec{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad (8.90)$$

$$\cos \lambda_G = \frac{X}{R'}, \quad \sin \lambda_G = \frac{Y}{R'}, \quad (8.91)$$

$$\left(\frac{\partial \lambda_G}{\partial \vec{R}} \right)^T = \begin{pmatrix} -\frac{Y}{R'^2} \\ \frac{X}{R'^2} \\ 0 \end{pmatrix} = \frac{1}{R'} \begin{pmatrix} -\sin \lambda_G \\ \cos \lambda_G \\ 0 \end{pmatrix}, \quad (8.92)$$

where $R' = (X^2 + Y^2)^{1/2}$. From (8.91) and (8.92), we obtain

$$\left[\frac{\partial(\cos m\lambda_G)}{\partial \vec{R}} \right]^T = \left(-\frac{m}{R'} \sin m\lambda_G \right) \begin{pmatrix} -\sin \lambda_G \\ \cos \lambda_G \\ 0 \end{pmatrix}, \quad (8.93)$$

$$\left[\frac{\partial(\sin m\lambda_G)}{\partial \vec{R}} \right]^T = \left(\frac{m}{R'} \cos m\lambda_G \right) \begin{pmatrix} -\sin \lambda_G \\ \cos \lambda_G \\ 0 \end{pmatrix}. \quad (8.94)$$

8.5.1 The Perturbation Acceleration of the Zonal Harmonic Term $\vec{F}_1(J_l)$

The transformation between the satellite's position vectors in the Earth-fixed coordinate system $\vec{R}$ and in the geocentric celestial coordinate system $\vec{r}$ is given by

$$\vec{R} = (HG)\vec{r}, \quad (8.95)$$

and correspondingly there is

$$\left(\frac{\partial \vec{R}}{\partial \vec{r}} \right) = (HG)^T. \quad (8.96)$$

From these transformations, we can obtain the acceleration of the zonal harmonic term $\vec{F}(J_l)$, which is

$$\vec{F}(J_l) = \left[\frac{\partial(\Delta V_l)}{\partial \vec{r}} \right]^T = \left(\frac{\partial \vec{R}}{\partial \vec{r}} \right)^T \left[\frac{\partial(\Delta V_l)}{\partial \vec{R}} \right]^T. \quad (8.97)$$

The matrix (HG) is given by

$$(HG) = (EP)(ER)(NR)(PR), \quad (8.98)$$

where (EP) , (ER) , (NR) , and (PR) are matrices of the polar motion, rotation, nutation, and precession, respectively. The formulas are given in Sect. 1.3. There is another important matrix,

$$\left[\frac{\partial(\Delta V_l)}{\partial \vec{R}} \right]^T,$$

which is expressed as

$$\left[\frac{\partial(\Delta V_l)}{\partial \vec{R}} \right]^T = - \sum_{l \geq 2} \bar{C}_{l,0} \left(\frac{1}{R} \right)^{l+3} \left\{ [(l+1)\bar{P}_l(\sin \varphi) + \sin \varphi \bar{P}'_l(\sin \varphi)](\vec{R}) - R \bar{P}'_l(\sin \varphi)(\vec{k}) \right\}, \quad (8.99)$$

where

$$\vec{R} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}, \quad \vec{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (8.100)$$

Note that here $R = r$. In the calculation of $\vec{F}(J_l)$ we need the transformation (8.95) to give $\vec{R}$. The main part of the calculation is about Earth's non-spherical gravity acceleration (the zonal harmonic part) which involves the Legendre polynomial $\bar{P}_l(\mu)$ and its derivative $\bar{P}'_\mu(\mu) = \partial \bar{P}_l(\mu) / \partial \mu$. The calculation method is given in Sect. 8.5.3.

8.5.2 The Perturbation Acceleration of the Tesseral Harmonic Term $\vec{F}(J_{l,m})$

The calculation is similar to that for the zonal harmonic acceleration, and the acceleration due to the tesseral harmonic terms $\vec{F}(J_{l,m})$ is

$$\vec{F}(J_{l,m}) = \left[\frac{\partial(\Delta V_{l,m})}{\partial \vec{r}} \right]^T = \left(\frac{\partial \vec{R}}{\partial \vec{r}} \right)^T \left[\frac{\partial(\Delta V_{l,m})}{\partial \vec{R}} \right]^T, \quad (8.101)$$

where $\left(\frac{\partial \vec{R}}{\partial \vec{r}} \right)^T$ is given by (8.96), the other matrix is given by

$$\begin{aligned} \left[\frac{\partial(\Delta V_{l,m})}{\partial \vec{R}} \right]^T = & - \sum_{l \geq 2} \sum_{m=1}^l \left\{ \left(\frac{1}{R} \right)^{l+3} \left[((l+1)\bar{P}_{l,m}(\sin \varphi) + \sin \varphi \bar{P}'_{l,m}(\sin \varphi))(\vec{R}) \right. \right. \\ & \left. \left. - R \bar{P}'_{l,m}(\sin \varphi)(\vec{k}) [\bar{C}_{lm} \cos m\lambda_G + \bar{S}_{lm} \sin m\lambda_G] \right. \right. \\ & \left. \left. + \left(\frac{m}{R'} \right) \left(\frac{1}{R} \right)^{l+1} \bar{P}_{l,m}(\sin \varphi) [\bar{C}_{l,m} \sin m\lambda_G - \bar{S}_{l,m} \cos m\lambda_G](\vec{G}) \right\}, \end{aligned} \quad (8.102)$$

$$(\vec{G}) = (-\sin \lambda_G \cos \lambda_G \ 0)^T, \quad (8.103)$$

where $R' = (X^2 + Y^2)^{1/2}$ is also in (8.92).

8.5.3 The Recursive Formulas of Legendre Polynomials, $P_l(\mu)$ and the Associated Legendre Polynomials $P_{l,m}(\mu)$, and Their Derivatives [15, 16]

The Legendre polynomials $P_l(\mu)$ and the associated Legendre polynomials $P_{l,m}(\mu)$, and their derivatives, $P'_l(\mu)$ and $P'_{l,m}(\mu)$, where $\mu = \sin \varphi \in [-1, 1]$, by principle can be calculated according to their definitions, but when the orders are high (m and l) the calculation becomes tremendous, and the effective digital number can be damaged severely (due to the properties of the functions and the limited lengths of quantities in computing), and eventually, the calculation fails. To deal with this issue we must apply a proper recursive process to control the propagating speed of the calculation error, therefore, to control the accuracy. The following recursive method can achieve this goal.

(1) The recursive formulas of $P_l(\mu)$ and $P_{l,m}(\mu)$

$$\begin{cases} \bar{P}_0(\mu) = 1 & \bar{P}_1(\mu) = \sqrt{3}\mu \\ \bar{P}_l(\mu) = \left(\frac{2l+1}{2l-1}\right)^{1/2} \left[\left(2 - \frac{1}{l}\right)\mu\bar{P}_{l-1}(\mu) - \left(\frac{2l-1}{2l-3}\right)^{1/2} \left(1 - \frac{1}{l}\right)\mu\bar{P}_{l-2}(\mu) \right], & l \geq 2 \end{cases} \quad (8.104)$$

$$\begin{cases} \bar{P}_{1,1}(\mu) = \sqrt{3}(1 - \mu^2)^{1/2} \\ \bar{P}_{l,l}(\mu) = \left(\frac{2l+1}{2l}\right)^{1/2} (1 - \mu^2)^{1/2} \bar{P}_{l-1,l-1}(\mu), & l \geq 2 \\ \bar{P}_{l,m}(\mu) = \left[\frac{(2l+1)(2l-1)}{(l+m)(l-m)} \right]^{1/2} \bar{P}_{l-1,m}(\mu) \\ \quad - \left[\frac{(2l+1)(l-1+m)(l-1-m)}{(2l-3)(l+m)(l-m)} \right]^{1/2} \bar{P}_{l-2,m}(\mu), & l \geq 2, m = 1, \dots, l-1 \end{cases} \quad (8.105)$$

Note that in the above formulas there is

$$\bar{P}_{ij}(\mu) = 0, \quad i < j. \quad (8.106)$$

(2) The recursive formulas of $\bar{P}'_l(\mu)$ and $\bar{P}'_{l,m}(\mu)$.

The following calculation is performed after $P_l(\mu)$ and $P_{l,m}(\mu)$ are calculated,

$$\begin{cases} \bar{P}'_1(\mu) = \sqrt{3}, \\ \bar{P}'_l(\mu) = l(1 - \mu^2)^{-1} \left[\left(\frac{2l+1}{2l-1}\right)^{1/2} \bar{P}_{l-1}(\mu) - \mu\bar{P}_l(\mu) \right], & l > 1, \end{cases} \quad (8.107)$$

$$\bar{P}'_{l,m}(\mu) = (1 - \mu^2)^{-1} \left\{ [(l+m+1)(l-m)]^{1/2} \bar{P}_{l,m+1}(\mu) - m\mu(1 - \mu^2)^{-1/2} \bar{P}_{lm}(\mu) \right\}. \quad (8.108)$$

8.5.4 The Perturbation Acceleration of the Tidal Deformation $\vec{F}(k_2, J_{2,m})$

The perturbation of the tidal deformation on Earth's satellite orbit is discussed in Chap. 4. Usually, it has a second-order effect on a low Earth orbit satellite, and can be treated as a perturbation quantity of the third-order. The corresponding calculation method in a numerical method is given below.

In order to keep consistency in calculations, we also use the dimensionless units. The perturbation potential of the tidal deformation for $l = 2$ is given by

$$\Delta V_{2,0} = [\beta k_2 P_2(\sin \varphi')] \left(\frac{1}{r}\right)^3 P_2(\sin \varphi), \quad (8.109)$$

$$\Delta V_{2,1} = \left[\frac{1}{3}\beta k_2 P_{2,1}(\sin \varphi')\right] \left(\frac{1}{r}\right)^3 P_{2,1}(\sin \varphi) \cos(\lambda - \lambda'), \quad (8.110)$$

$$\Delta V_{2,2} = \left[\frac{1}{3}\beta k_2 P_{2,2}(\sin \varphi')\right] \left(\frac{1}{r}\right)^3 P_{2,2}(\sin \varphi) \cos 2(\lambda - \lambda'), \quad (8.111)$$

where the coefficient k_2 is the Love number of the second-order, β is the parameter β_2 in the perturbation function $P_2(\cos \psi)$ of the Sun or the Moon, see (4.496), which is

$$\beta = \beta_2 = m'/r'^3, \quad (8.112)$$

where m' is the relative mass of the Sun or the Moon to Earth's mass, r' is the distance from a third body to the center of Earth, for the tidal deformation perturbation it can be treated as a constant.

With these considerations, the perturbation acceleration of the tidal deformation is given by

$$\vec{F}(k_2 J_{2,m}) = \left(\frac{\partial \sum \Delta V_{2,m}}{\partial \vec{r}} \right)^T. \quad (8.113)$$

The acceleration can be separated into two parts, the zonal harmonic part ($m = 0$) and the tesseral part ($m = 1, 2$) that

$$\begin{aligned} \left(\frac{\partial \Delta V_{2,0}}{\partial \vec{r}} \right)^T &= -k_2 \beta P_2(\sin \varphi') \left(\frac{1}{r}\right)^5 \left\{ [3P_2(\sin \varphi) + \sin \varphi P_2'(\sin \varphi)] \vec{r} - r P_2'(\sin \varphi) \vec{k} \right\}, \\ \left(\frac{\partial \Delta V_{2,m}}{\partial \vec{r}} \right)^T &= -k_2 \beta \sum_{m=1}^2 \frac{2(2-m)!}{(2+m)!} P_{2,m}(\sin \varphi') \\ &\quad \times \left(\frac{1}{r}\right)^5 \left\{ [(3P_{2,m}(\sin \varphi) + \sin \varphi P_{2,m}'(\sin \varphi)) \vec{r} - r P_{2,m}'(\sin \varphi) \vec{k}] \cos m(\lambda - \lambda^*) \right\} \end{aligned} \quad (8.114)$$

$$+ \frac{m}{\sqrt{x^2 + y^2}} \left(\frac{1}{r} \right)^3 [P_{2,m}(\sin \varphi) \sin m(\lambda - \lambda^*)] \vec{G} \} \quad (8.115)$$

where

$$\vec{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{G} = \begin{pmatrix} -\sin \lambda \\ \cos \lambda \\ 0 \end{pmatrix}. \quad (8.116)$$

$P_2(\sin \varphi)$ and $P_{2,m}(\sin \varphi)$ and their derivatives are

$$P_2(\sin \varphi) = \frac{3}{2} \sin^2 \varphi - \frac{1}{2}, \quad P'_2(\sin \varphi) = 3 \sin \varphi, \quad (8.117)$$

$$\begin{cases} P_{21}(\sin \varphi) = \frac{3}{2} \sin 2\varphi, & P'_{21}(\sin \varphi) = 3 \cos 2\varphi / \cos \varphi \\ P_{22}(\sin \varphi) = 3 \cos^2 \varphi, & P'_{22}(\sin \varphi) = -6 \sin \varphi \end{cases} \quad (8.118)$$

$$\begin{cases} \cos \lambda = x / \sqrt{x^2 + y^2} & \sin \lambda = y / \sqrt{x^2 + y^2} \\ \cos 2\lambda = \cos^2 \lambda - \sin^2 \lambda & \sin 2\lambda = 2 \sin \lambda \cos \lambda \end{cases} \quad (8.119)$$

In (8.114)–(8.115), φ' and $\lambda^* = \lambda' + \nu$ are the spherical coordinates of the Sun or the Moon in the geocentric coordinate system, where ν is the angle of lag, the relationships of λ' and φ' with the geocentric Cartesian coordinates are

$$\begin{cases} \sin \varphi' = \frac{z'}{r'}, \\ \cos \lambda' = \frac{x'}{\sqrt{x'^2 + y'^2}}, \quad \sin \lambda' = \frac{y'}{\sqrt{x'^2 + y'^2}}. \end{cases} \quad (8.120)$$

8.6 The Role of the Hamiltonian Method in the Orbital Evolution

In orbit dynamics, the evolution of a dynamical system, particularly the evolution of the Solar System (the orbital evolution of big and small celestial bodies) is an important research field. When using numerical methods to study this topic, a critical problem is how to avoid the effect of the artificial dissipation caused by the numerical method itself. The commonly used method is the Hamiltonian calculation method which is used to maintain the entire geometric structure of the system and then to provide actual characteristics of the evolution of the dynamical system. The theoretical analysis and related calculation methods are beyond the scope of this book. Readers who need to know more may refer to the references [17–24].

References

1. The Computer science division, Mathematics Department, Nanjing University (1979) Numerical Method of Solving Differential equations. Science Press, Beijing, China
2. Henrici P (1962) Discrete Variable Methods in Ordinary Differential Equations. Wiley, New York
3. Fehlberg E (1968) Classical Fifth-, Sixth-, Seventh-, and Eighth-Order Runge-Kutta with Step-size Control. NASA. TR R-287.
4. Cowell PH, Crommelin ACD (1910) Appendix to Greenwich Observations for 1909, Edinburgh, 84.
5. Brouwer D, Clemence GM (1961) Methods of Celestial Mechanics. Academic Press, New York and London
6. Kyner WT, Bennett MM (1966) Modified Encke special perturbation method. *Astron J* 71:579–582
7. Escobal PR (1965) Methods of Orbit Determination. John Wiley & Sons, New York
8. Liu L, Hu SJ (1996) Selection and correction of the reference orbit in the improved Encke method. *Acta Astronom Sinica* 37(3):285–293
9. Szebehely V (1967) Theory of Orbits. Academic Press, New York
10. Stiefel EL, Scheifele G (1970) Linear and Regular Celestial Mechanics. Springer-Verlag, Berlin, Heidelberg, New York
11. Huang TY, Ding H (1981) Stability and transformation of the independent variable. *ACTA Astronomica Sinica* 22(4):328–335
12. Liu L, Liao XH (1987) A few problems in the numerical calculation of Celestial Dynamics. *Acta Astronom Sinica* 28(3):215–225
13. Liu L, Liao XH (1994) Numerical calculations in the orbital determination of an artificial satellite for a long arc. *Celest Mech* 59(3):221–235
14. Liu L, Liao XH (1997) A few problems about numerical methods in solving motion equations of celestial bodies. *Acta Astronom Sinica* 38(1):75–85
15. Liu L (2000) Orbital Theory of Spacecraft. National Defense Industry Press, Beijing
16. Montenbruck O, Gill E (2000) Satellite Orbits: Models, Methods, and Applications. Springer-Verlag, Berlin, Heidelberg
17. Feng K (ed) (1984) Proc. Science Press, Beijing Symposium on Differential Geometry and Differential Equations, p 42
18. Feng K (1986) Difference Schemes for Hamiltonian Formalism and Symplectic Geometry. *J. Comp. Math.* 4:279
19. Forest E, Ruth RD (1990) Fourth-order Symplectic Integration. *Physica D* 43:105
20. Yoshida H (1990) Construction of Higher Order Symplectic Integrators. *Phys Lett A* 150:262–268
21. Zhao ZY, Liao XH, Liu L (1992) Applications of Symplectic integration in Astronomy Dynamics (I). *Acta Astronom Sinica* 33(1):36–47
22. Liu L, Liao XH, Zhao ZY, Wang CB (1994) Applications of Symplectic integration in Astronomy Dynamics (III). *Acta Astronom Sinica* 35(1):51–56
23. Liao XH, Liu L (1994) An improved Symplectic method. *Chinese J of Computational Physics* 11(2):212–218
24. Liu L (1998) Methods of Celestial Mechanics. Nanjing University Press.

Chapter 9

Formulation and Calculation of Initial Orbit Determination



The orbit determination can be directly defined as the process of determining the motion state of an orbit of a celestial body (including various spacecraft) at a given time t_0 from a certain number of tracking measurements using proper mathematical methods. The motion state of a body means its position vector and velocity $\vec{r}_0, \dot{\vec{r}}_0$ at a given time in a chosen space coordinate system. If a body (it can be one of the major planets or an asteroid in the Solar System or an artificial satellite) circles around a main body then the motion state of the body at a given time can be given by the six orbital elements, which are usually denoted to $\sigma_0 = (a_0, e_0, i_0, \Omega_0, \omega_0, M_0)^T$.

9.1 Formulation of Orbit Determination

In the field of celestial mechanics and orbital dynamics, orbit determination usually has two concepts, the orbit determination of an initial orbit by short-arc measurements and the precise orbit determination by long-arc measurements (or a short-arc sometimes). The initial orbit determination usually uses a non-perturbed two-body problem as the reference model. It is a necessary job no matter it is for a special aerospace project or for discovering a small celestial body in the Solar System (asteroid, natural satellite, or comet). The initial orbit can be used in solving some problems or for providing the initial information for the precise orbit determination. The orbital model of a precise orbit determination is a perturbed two-body problem or a perturbed $(N + 1)$ -body problem of a “complete” dynamical system. The precise orbit determination is called orbit improvement by tradition. Because in the process of its development we also gain other parameters (geometric and physical), so the whole process actually includes the precise orbit determination and parameter estimations; also by the principle and the method of the precise orbit determination it is necessary to have the initial orbit information of $\vec{r}^*, \dot{\vec{r}}^*$ or σ^* , which are initial values for iteration, but the corresponding time t^* is not necessarily the same as the

epoch time t_0 , from these points of view it is more proper to call it the precise orbit determination than orbit improvement.

The observational variables for the orbit determination can be distance ρ , velocity $\dot{\rho}$, angles α and δ (right ascension and declination), angles A and h (azimuth and elevation), and position vector $\vec{r}(x, y, z)$, etc. Usually, we denote the measurement variable by Y , the epoch time of the orbit determination by t_0 , and the state variable at t_0 by X_0 (which includes orbital quantities and some geometric and physical parameters to be decided), the relationships between them are

$$\begin{cases} Y = Y(X, t), \\ X(t) = X(t; t_0, X_0). \end{cases} \quad (9.1)$$

The state variable X is an n -dimensional variable, the series of measurement variable Y_j ($j = 1, 2, \dots, k$) at t_j has $(m \times k)$ dimensions, m ($m \geq 1$) is the dimension of the observed quantity (the sampling of angle quantity has 2 dimensions). When $(m \times k) \geq n$, theoretically the system of Eq. (9.1) can be solved, i.e., from a series of measurements Y_j ($j = 1, 2, \dots, k$) at t_j to give the state variable X_0 at the epoch time t_0 . Actually, it is a problem of solving an implicit equation $\Phi(t, Y; t_0, X_0) = 0$, which involves the certainty of a solution and how to solve it. The following two points should be discussed about these two problems.

- (1) The problem of the certainty of a solution (i.e., the condition of orbit determination) is about whether the only state variable X_0 at t_0 can be determined by a series of measurements Y_j ($j = 1, 2, \dots, k$) at t_j . This problem is also called the observability problem. For example, if the measurements are distances and velocities from a single station over a short arc, it is difficult to determine an orbit. The related mathematical problem is discussed in the following sections.
- (2) When the problem of certainty is solved, then we can think of how to derive a solution. The equation system (9.1) consists of multi-variable non-linear equations, and transcendental functions, therefore, it is impossible to be directly solved. Usually, the orbit is determined by a multi-variable iterative procedure. In the case of a short arc if the measurements are for one variable, distance or speed, then the solution can only be obtained by the general multi-variable iteration; but if the measurements are for angles, it is possible to solve the problem by a special and simple iterative method even if the dynamical model includes the complete dynamic system. This is the major difference between the method of orbit determination using data over a short arc and the method of precise orbit determination based on multi-variable iteration.

9.2 A Review of Initial Orbit Calculation in the Sense of the Two-Body Problem

In the history of celestial mechanics over more than 300 years, there are many different methods to determine the orbits of celestial bodies using angle measurements. Basically, there are two types of methods, Laplace's method and Gauss's method [1–11]. For the modern computing capability Laplace's method is more concise and efficient. In this section, we discuss the principles of the initial orbit calculation by Laplace's method in the sense of the two-body problem combined with the author's decades-long experience.

9.2.1 Basic Conditions for Initial Orbit Determination

First, we choose a frame for a motion. For an Earth's artificial satellite, the space coordinate system $O\text{-}xyz$ is the epoch J2000.0 geocentric celestial coordinate system corresponding to the J2000.0 Earth's mean equator and mean March equinox. According to the formulation of orbit determination, the initial orbit determination involves two types of equations, which are the observational equation of geometric relationships and the differential equation of dynamical motion state.

(1) The observational equation: the geometrical relationship

The geometrical relationship for a moving body in the chosen coordinate system is given by

$$\vec{r} = \vec{\rho} + \vec{R}, \quad (9.2)$$

where $\vec{r}$ is the position vector of the moving body; $\vec{\rho}$ is the observed vector of the moving body from the observatory as $\vec{\rho}(\rho, \alpha, \delta)$ (α is the right ascension and δ is the declination) or $\vec{\rho}(\rho, A, h)$ (A is the azimuth and h is the elevation); and $\vec{R}$ is the coordinate vector of the observatory. In the frame of $O\text{-}xyz$, $\vec{\rho}$ can be expressed as

$$\vec{\rho} = \rho \hat{L}, \quad \hat{L} = (\lambda, \mu, \nu)^T. \quad (9.3)$$

The superscript "T" means the transposition matrix; the rectangular coordinate components (λ, μ, ν) can be given by angles (α, δ) or (A, h) via a simple transformation of the coordinate systems; R is known and is expressed as

$$\vec{R} = (X_e, Y_e, Z_e)^T. \quad (9.4)$$

Here $\vec{R}$ and the position vector of the moving body should be in the same coordinate system.

In fact, the geometrical relationship (9.2) corresponding to the first equation of (9.1), $Y = Y(X, t)$, is the relationship between the measurement variable Y and the state variable X . For example, the measurement variable (length) is ρ , the state variable is $\vec{r}$, and their relationship is $\rho = |\vec{r} - \vec{R}|$, i.e., ρ and $\vec{r}$ are Y and X , respectively.

(2) The motion equation: the dynamical relationship and its solution

In the frame of $O\text{-}xyz$, the differential equation of the relative motion of the observed body with respect to the center of the central body is

$$\begin{cases} \ddot{\vec{r}} = -\frac{GM_0}{r^3}\vec{r}, \\ \vec{r}(t_0) = \vec{r}_0, \quad \dot{\vec{r}}(t_0) = \dot{\vec{r}}_0, \end{cases} \quad (9.5)$$

where GM_0 is the gravitational constant of the central body.

Equation (9.5) has the following solution which satisfies the initial conditions

$$\begin{cases} \vec{r} = \vec{r}(t; t_0, \vec{r}_0, \dot{\vec{r}}_0), \\ \dot{\vec{r}} = \dot{\vec{r}}(t; t_0, \vec{r}_0, \dot{\vec{r}}_0). \end{cases} \quad (9.6)$$

The solution (9.6) corresponds to the second equation of (9.1), $X(t) = X(t; t_0, X_0)$, i.e., the state transition equation, and X is the state variable that

$$X = \begin{pmatrix} \vec{r} \\ \dot{\vec{r}} \end{pmatrix}.$$

9.2.2 Construction of the Basic Equation for an Initial Orbit

The procedure of both the initial orbit determination and the precise orbit determination from the point of view of pure mathematics is to connect the observable geometrical relationship with the related dynamical condition (the dynamical law which the moving body must obey, i.e., the state transition equation), then to form a functional relation of the observed variable Y and the state variable X as

$$\begin{aligned} \Phi(t_j, Y_j, X_j(t_j; t_0, X_0)) &= \Phi(t_j, Y_j; t_0, X_0) = 0, \\ j &= 1, 2, \dots, k. \end{aligned} \quad (9.7)$$

From this relationship, the state variable X_0 at t_0 can be derived that

$$X_0 = \psi(t_0; t_1, t_2, \dots, t_k; Y_1, Y_2, \dots, Y_k). \quad (9.8)$$

Actually, the state variable X_0 is implicitly included in $\Phi(t_j, Y_j; t_0, X_0)$, and the initial orbit determination is to solve the equation system $\Phi(t_j, Y_j; t_0, X_0) = 0$.

For the initial orbit determination, even if we use a non-perturbed two-body motion problem as the model, which is the simplest, the corresponding functions in (9.7) become very complicated and include transcendental functions through the Kepler equation. Over 300 years, many methods have been developed to avoid the “formal way” of directly solving the complicated equation system $\Phi(t_j, Y_j; t_0, X_0) = 0$, but to use some special relationships of the two-body problem and some related treatments. Among the various methods, there are mainly Laplace’s method, Gauss’s method, and methods combining these two.

Laplace’s method of initial orbit calculation uses the dynamical condition by which the Eq. (9.5) can be expanded into power series of the time interval $\Delta t = t - t_0$ as

$$\vec{r}(t) = \vec{r}_0 + \vec{r}_0^{(1)} \Delta t + \frac{1}{2!} \vec{r}_0^{(2)} \Delta t^2 + \cdots + \frac{1}{k!} \vec{r}_0^{(k)} \Delta t^k + \cdots, \quad (9.9)$$

where $\vec{r}_0^{(k)}$ is the value of the k th-order of the derivative of $\vec{r}(t)$ at $t = t_0$ that

$$\vec{r}_0^{(k)} = \left(\frac{d^k \vec{r}}{dt^k} \right)_{t=t_0}.$$

The power series solution of (9.9) can be derived by calculating the value of each order of derivatives of $\vec{r}^{(k)}$ at $t = t_0$. With $\vec{r}_0^{(1)} = \dot{\vec{r}}_0$, the value of the k th-order derivative $\vec{r}_0^{(k)}$ ($k \geq 2$) can be formed by $\vec{r}_0$ and $\dot{\vec{r}}_0$ based on (9.5), that

$$\vec{r}_0^{(k)} = \vec{r}_0^{(k)}(t_0, \vec{r}_0, \dot{\vec{r}}_0), \quad k \geq 2. \quad (9.10)$$

Then the corresponding series of (9.9) can be arranged in the following form:

$$\vec{r}(t) = F(\vec{r}_0, \dot{\vec{r}}_0, \Delta t) \vec{r}_0 + G(\vec{r}_0, \dot{\vec{r}}_0, \Delta t) \dot{\vec{r}}_0. \quad (9.11)$$

For the three components of $\vec{r}(t)$, $x(t)$, $y(t)$, and $z(t)$, F and G are quantity functions with the same form. In the initial orbit determination, F and G are quantity functions of vital importance, and have a common property, which is

$$F = 1 + O(s^2), \quad G = \Delta t[1 + O(s^2)], \quad (9.12)$$

where $\Delta t = t - t_0$; s is the arc of motion corresponding to $\Delta t = t - t_0$, that $s = n\Delta t$, and n is the mean angular speed of the moving body.

When using data over a short arc to determine the initial orbit, there is $|s| < 1$. Therefore, even though we know nothing about the orbit of the target body, we can choose

$$F^{(0)} = 1, \quad G^{(0)} = \Delta t, \quad (9.13)$$

as the initial values for iteration. These two values in fact are excellent initial information. We can say that the solution of the motion Eq. (9.5), which is given as power series of time interval Δt by (9.11) for a short arc, is a “global” dynamical condition for determining the basic equation to construct the initial orbit. The basic equation for orbit determination is a linear equation of $t_0; \vec{r}_0, \dot{\vec{r}}_0$. The equation can be solved by iteration, which is simple with modern computing power. Details about the method are provided in the following content.

9.3 Initial Orbit Determination for Perturbed Motion

Following the improvement of measurements of Earth’s artificial satellite (such as the precision of angle data can be in seconds, and for navigation the precision of length in meters and even better) and requirements of deep-space exploration projects, the initial orbit determination problem now faces varieties of target and different precision demands. It is necessary to improve the traditional method, particularly to promote the reference model of the initial orbit determination from a two-body problem to a general perturbed two-body problem by including the perturbations due to the non-spherical part of the central body and a third-body’s attraction. Also, the orbit determination method should not be restricted to ellipse, i.e., the hyperbola orbit should be included. In this section, we provide the expanded Laplace’s method for initial orbit calculation, and a special iterate procedure for short-arc orbit determinations without prior information. As mentioned above this iterative method is fundamentally different from the multi-variable iteration. The measurements are the conventional tracking data, and the expansion of Laplace’s method is relatively simple, we may call the expanded method as the general Laplace’s method [12–15]. In this section, we first introduce the basic principles of the general Laplace’s method, then apply the actual method to a few different types of measurements and give examples of result tests.

9.3.1 *Construction of the Basic Equation for Initial Orbit Determination*

After comparing different methods, the author has developed an improved Laplace’s method based on today’s observation technique and computing ability. The method emphasizes the advantage of high accuracy precision measurements, and focuses on the physical background of the uncooperative target (i.e., without prior information of the initial target orbit) to build an initial orbit. The main property of the sampling data is that the data are from a short-arc, i.e., an arc $s = n\Delta t < 1$, that is the reason why the method is called short-arc orbit determination. For an Earth’s artificial satellite, the arc is about 1/3 of a circle, which is equivalent to 30 min of a low orbit satellite moving

on its orbit (its period is about 90 min). If the epoch time of the to-be-determined orbit t_0 is at the middle of the arc, then the longest arc of the data coverage satisfies $s_N = n |t_N - t_0| < 1$. Under this promise, the accepted dynamical condition is neither a numerical solution of the perturbed differential equations nor an analytical solution of a power series of a small parameter, but a power series of the time interval $\Delta t = t - t_0$, which is similar to the initial orbit solution of the state differential equation given based on the two-body problem model. This solution in the sense of short-arc also includes all information about the orbit.

The perturbed motion equation with initial conditions is

$$\begin{cases} \ddot{\vec{r}} = -\frac{Gm_0}{r^3}\vec{r} + \vec{F}_\varepsilon(\vec{r}, \dot{\vec{r}}, t; \varepsilon), \\ \vec{r}(t_0) = \vec{r}_0, \quad \dot{\vec{r}}(t_0) = \dot{\vec{r}}_0, \end{cases} \quad (9.14)$$

where Gm_0 is the gravitational constant of the central body as a sphere. The perturbed acceleration $\vec{F}_\varepsilon$ includes all dynamical forces (conservative and dissipative) except the central gravity force. The magnitude of $\vec{F}_\varepsilon$ is irrelevant as long the related forces have mathematical models. For this dynamical model, the initial orbit to be determined is an instantaneous orbit at the epoch time t_0 , which can be an ellipse or a hyperbola.

If the right-side function of the motion Eq. (9.14) (no restriction on the orbit type) satisfies certain conditions, its solution which satisfies the initial conditions can be expanded to a power series of the time interval $\Delta t = t - t_0$ as same as (9.9), i.e.,

$$\vec{r}(t) = \vec{r}_0 + \vec{r}_0^{(1)} \Delta t + \frac{1}{2!} \vec{r}_0^{(2)} \Delta t^2 + \cdots + \frac{1}{k!} \vec{r}_0^{(k)} \Delta t^k + \cdots, \quad (9.15)$$

where $\vec{r}_0^{(k)}$ is the value of the k th-order derivative of $\vec{r}(t)$ at $t = t_0$ that

$$\vec{r}_0^{(k)} = \left(\frac{d^k \vec{r}}{dt^k} \right)_{t=t_0}.$$

The power series solution can be obtained by first calculating the derivatives of $\vec{r}^{(k)}$ for each order according to the perturbed motion Eq. (9.14), then calculating their values at $t = t_0$, $\vec{r}_0^{(k)}$. Similar to the unperturbed situation, the solution (9.15) can be arranged to be the following form,

$$\vec{r}(t) = F^*(\vec{r}_0, \dot{\vec{r}}_0, \Delta t) \vec{r}_0 + G^*(\vec{r}_0, \dot{\vec{r}}_0, \Delta t) \dot{\vec{r}}_0. \quad (9.16)$$

The functions F^* and G^* usually have different forms for the three components of $\vec{r}(t)$, $x(t)$, $y(t)$, and $z(t)$. If the perturbation is only due to the gravity of the non-spherical part of the central body then the x - and y -components of F^* , F_x , and F_y , are the same, also the x - and y -components of G^* , G_x , and G_y are the same. The two functions F^* and G^* are extremely important (their actual forms are discussed later) and have similar properties, which are

$$F_x, F_y, F_z = 1 + O(s^2), \quad G_x, G_y, G_z = \Delta t[1 + O(s^2)], \quad (9.17)$$

where $\Delta t = t - t_0$; s is the arc of motion corresponding to $\Delta t = t - t_0$. For a short-arc orbit determination, there is $|s| < 1$. Therefore, without knowing anything about the target orbit we can use the following values as initial values for iteration:

$$F^{(0)} = 1, \quad G^{(0)} = \tau, \quad F_z^{(0)} = F^{(0)}, \quad G_z^{(0)} = G^{(0)}. \quad (9.18)$$

These values provide excellent information of the initial state. The power series solution of the perturbed Eq. (9.14) given by (9.16) is a linear equation of t_0 ; $\vec{r}_0, \dot{\vec{r}}_0$. The equation can be solved by iteration, which is simple using today's computers.

The above-given power series solution of time interval Δt has no restriction on the type of the target trajectory, which can be an ellipse or a hyperbola. It is also applicable to both perturbed and unperturbed motions. Although the iteration involves complicated functions it is not a multi-variable iteration but a simple one. Also, it does not demand particular initial values for iteration, even though the to-be determined is an uncooperative target. By comparison, the precise orbit determination (to be discussed in Chap. 10) involves multi-variable iteration and requires certain initial values for iteration.

9.3.2 Initial Orbit Determination Using Angle Data Over a Short-Arc

The conventional space measurements are sampling data of angle variables, the right ascension and declination (α, δ) or the azimuth and elevation (A, h). The method of orbit determination uses the above-discussed power series solution of Δt as the dynamical condition to build an equation of orbit determination, then solve this equation directly for the position vector and velocity $(\vec{r}_0, \dot{\vec{r}}_0)$ at a chosen epoch t_0 . We call this method the general Laplace's method. This basic method can be promoted to orbit determinations using different types of data and different measurement models.

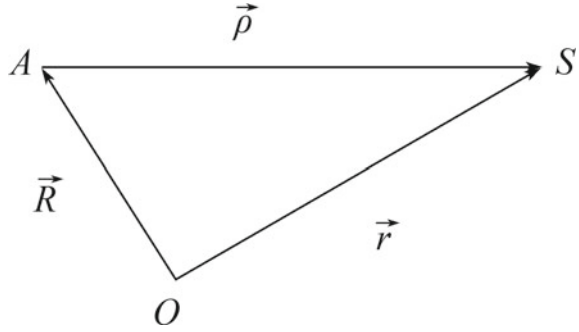
Now we use an Earth's satellite as an example. The coordinate system is the epoch J2000.0 geocentric celestial coordinate system O -xyz. The motion of an Earth's satellite corresponds to a perturbed two-body problem, but the method which we introduce here is not restricted to this dynamical model, readers are suggested to pay attention to the details of the method.

(1) Geometrical condition: the geometrical relationship of measurements

In the chosen frame O -xyz, the tracking station A (strictly speaking, it is a sampling center with measurement equipment) and the target satellite S have the following geometrical relationship:

$$\vec{r} = \vec{\rho} + \vec{R}, \quad (9.19)$$

Fig. 9.1 The geometrical relationship of a tracking station A , a satellite S , and the center of Earth O



where $\vec{r}$ is the position vector of the satellite; $\vec{\rho}$ is the observed vector of the satellite from the tracking station; and $\vec{R}$ is the coordinate vector of the tracking station (Fig. 9.1) that

$$\vec{R} = (X_e, Y_e, Z_e)^T. \quad (9.20)$$

The coordinate vector $\vec{R}$ should be given in the chosen geocentric celestial coordinate system. Because the position of the tracking station is usually given in the Earth-fixed coordinate system, it involves the transformation of the Earth-fixed frame and the geocentric celestial frame, which is provided in (1.20)–(1.30), or (1.75) with the related content (such as the transformation of the precession matrix and the nutation matrix).

The observed vector $\vec{\rho}$ for angle data in the frame O -xyz can be presented as

$$\vec{\rho} = \rho \hat{L}, \quad \hat{L} = (\lambda, \mu, \nu)^T. \quad (9.21)$$

The direct measurements can only provide the orientation, i.e., the unit vector $\hat{L} = (\lambda, \mu, \nu)^T$. The rectangular coordinates (λ, μ, ν) can be given by measured angles (α, δ) or (A, h) through the transformation of coordinate systems. For measured angles (α, δ) usually the sampling data are consistent with the chosen coordinate system O -xyz that

$$\hat{L} = \begin{pmatrix} \lambda \\ \mu \\ \nu \end{pmatrix} = \begin{pmatrix} \cos \delta \cos \alpha \\ \cos \delta \sin \alpha \\ \sin \delta \end{pmatrix}. \quad (9.22)$$

The measured (A, h) are the coordinates in the instantaneous true horizontal coordinate system, the transformation relationship is

$$\hat{L} = \begin{pmatrix} \lambda \\ \mu \\ \nu \end{pmatrix} = (GR)^T(ZR) \begin{pmatrix} \cos h \cos A \\ -\cos h \sin A \\ \sin h \end{pmatrix}, \quad (9.23)$$

where (GR) is the precession-nutation matrix, in the true equatorial coordinate system it is a unit matrix; (ZR) is the transformation matrix for the horizontal coordinate system and the equatorial coordinate system as provided in (1.16), that

$$(ZR) = R_z(180^\circ - S)R_y(90^\circ - \varphi), \quad (9.24)$$

where S and φ are the local sidereal time and geodetic latitude of the tracking station, respectively.

(2) Dynamical condition: the motion equation of the satellite

The orbit determination cannot be achieved by the geometrical relationship only, we must know the corresponding dynamical background of $\vec{r}$ in the geometrical relationship (9.19), which is the dynamical condition for orbit determination. In the frame $O\text{-}xyz$ the motion equation of the satellite related to the center of Earth is given by

$$\begin{cases} \ddot{\vec{r}} = -\frac{\mu}{r^3}\vec{r} + \vec{F}_\varepsilon(\vec{r}, \dot{\vec{r}}, t; \varepsilon), \\ \vec{r}(t_0) = \vec{r}_0, \quad \dot{\vec{r}}(t_0) = \dot{\vec{r}}_0, \end{cases} \quad (9.25)$$

where $\mu = GE$ is Earth's gravitational constant, in the normalized dimensionless system $\mu = 1$. The perturbed acceleration $\vec{F}_\varepsilon$ is for any kind of dynamical force, conservative or dissipative, as long as the force can be expressed by a mathematical model. The method of orbit determination uses a power-series solution of Δt , which satisfies the initial condition, as the dynamical condition to build an equation of orbit determination, then solving this equation directly to obtain the position vector and velocity $(\vec{r}_0, \dot{\vec{r}}_0)$ at the epoch time t_0 .

(3) The basic equation of orbit determination

The form of the power series solution is provided in the last section for a perturbed system, i.e., (9.16) reprinted here:

$$\vec{r}(t) = F^*(\vec{r}_0, \dot{\vec{r}}_0, \Delta t)\vec{r}_0 + G^*(\vec{r}_0, \dot{\vec{r}}_0, \Delta t)\dot{\vec{r}}_0,$$

where F^* and G^* are power series of Δt . Substituting this solution into the geometrical relationship (9.19) yields

$$\hat{L} \times (F^*\vec{r}_0 + G^*\dot{\vec{r}}_0) = \hat{L} \times \vec{R}. \quad (9.26)$$

This is the basic equation system of orbit determination. For one set of angle sampling data, only two of the three equations in (9.26) are independent, which means we need at least three sets of samplings to decide $\vec{r}_0(x_0, y_0, z_0)$ and $\dot{\vec{r}}_0(\dot{x}_0, \dot{y}_0, \dot{z}_0)$ at a given epoch time t_0 . The basic Eq. (9.26) are linear of $\vec{r}_0$ and $\dot{\vec{r}}_0$ “in form” (because the coefficients F^* and G^* also depend on $\vec{r}_0$ and $\dot{\vec{r}}_0$). If $\vec{r}_0$ and $\dot{\vec{r}}_0$ can be decided, then by simple transformations, we can derive the instantaneous trajectory at t_0 —an ellipse or a hyperbola.

The key issue in the process of determining an initial orbit is to obtain the coefficients of F^* and G^* which satisfy the demanding precision. If this problem can be solved, then we can avoid a multi-variable iteration and determine the orbit without prior information of the target. The following content gives how to construct the equations of orbit determination (9.26).

If the right-side function of the motion Eq. (9.25) meets certain conditions as mentioned in Sect. 9.2 (no restriction on the type of the trajectory), then there is a solution of the equation that satisfies the initial conditions, and the solution can be expressed as a power series of the time interval $\Delta t = t - t_0$, i.e., (9.15)

$$\vec{r}(t) = \vec{r}_0 + \vec{r}_0^{(1)} \Delta t + \frac{1}{2!} \vec{r}_0^{(2)} \Delta t^2 + \cdots + \frac{1}{k!} \vec{r}_0^{(k)} \Delta t^k + \cdots,$$

where $\vec{r}_0^{(k)}$ can be given by $\vec{r}_0$ and $\dot{\vec{r}}_0$ in the form of (9.10):

$$\vec{r}_0^{(k)} = \vec{r}_0^{(k)}(t_0, \vec{r}_0, \dot{\vec{r}}_0), \quad k \geq 2.$$

We now give the specific expression for the perturbations due to the main term of the central body's non-spherical part (the J_2 term) and a third body. The combined acceleration of the two perturbations is

$$\vec{F}_\varepsilon = \left(\frac{3J_2}{2}\right) \left[\left(5\frac{z^2}{r^7} - \frac{1}{r^5}\right) \vec{r} - \left(\frac{2z}{r^5}\right) \hat{k} \right] - \mu' \left(\frac{\vec{\Delta}}{\Delta^3} + \frac{\vec{r}'}{r'^3} \right). \quad (9.27)$$

For the convenience of analyzing magnitudes and expressing formats, we use the dimensionless units. The corresponding gravitational constant of Earth is $\mu = GE = 1$, and $\mu' = GM'/GE$, where M' is the mass of the third body (the Sun or the Moon). In (9.27) J_2 is the coefficient of Earth's dynamical form-factor, and

$$\hat{k} = (0 \ 0 \ 1)^T, \quad \vec{\Delta} = \vec{r} - \vec{r}', \quad (9.28)$$

$\vec{r}'$ is the position vector of the third body.

Omitting the procession of derivation, the three components of the solution in the power series of Δt are given in the same format of (9.16) that

$$\begin{cases} x = F(\vec{r}_0, \dot{\vec{r}}_0, \Delta t) x_0 + G(\vec{r}_0, \dot{\vec{r}}_0, \Delta t) \dot{x}_0, \\ y = F(\vec{r}_0, \dot{\vec{r}}_0, \Delta t) y_0 + G(\vec{r}_0, \dot{\vec{r}}_0, \Delta t) \dot{y}_0, \\ z = F_z(\vec{r}_0, \dot{\vec{r}}_0, \Delta t) z_0 + G_z(\vec{r}_0, \dot{\vec{r}}_0, \Delta t) \dot{z}_0. \end{cases} \quad (9.29)$$

Note as mentioned above that the x -component and y -component of F^* are the same, so both are donated by F , also that the x -component and y -component of G^* are the same, so both are donated by G . The z -components of F^* and G^* are denoted to F_z and G_z , respectively. Then (9.26) can be expressed in the three components as

$$\begin{cases} (Fv)x_0 - (F_z\lambda)z_0 + (Gv)\dot{x}_0 - (G_z\lambda)\dot{z}_0 = (vX_e - \lambda Z_e), \\ (Fv)y_0 - (F_z\mu)z_0 + (Gv)\dot{y}_0 - (G_z\mu)\dot{z}_0 = (vY_e - \mu Z_e), \\ (F\mu)x_0 - (F\lambda)y_0 + (G\mu)\dot{x}_0 - (G\lambda)\dot{y}_0 = (\mu X_e - \lambda Y_e). \end{cases} \quad (9.30)$$

Also as mentioned above, one sampling forms the expressions of the three components in (9.30), but there are only two independent, i.e., the third one can be given by any two of the three. Considering the random nature of sampling, it is inappropriate to artificially remove one of the components. Besides to keep all the three components would not affect the process of solving the equations but can provide useful statistical information and improve the precision.

It should be noticed that the geometric state of $\hat{L} \times \vec{R} \approx 0$ must be avoided, otherwise, the system of Eq. (9.30) would have a zero solution. Usually, this situation does not appear, it is properly managed by the ground measurement and control system.

Denoting τ to Δt , the power series of the coefficients F , G , F_z , and G_z have the following forms:

$$\begin{aligned} F = & 1 + \frac{\tau^2}{2} \left[-u^3 + \left(\frac{3J_2}{2} \right) (5u_7z_0^2 - u_5) - \mu'\mu'_3 \right] \\ & + \frac{\tau^3}{6} \left\{ 3u_5\sigma + \left(\frac{3J_2}{2} \right) [5(u_7 - 7u_9z_0^2)\sigma + 10u_7z_0\dot{z}_0] \right\} \\ & + \frac{\tau^4}{24} \left\{ u_5(3v_0^2 - 2u_1 - 15u_2\sigma^2) + \left(\frac{3J_2}{2} \right) [6u_8(4u_2z_0^2 - 1) - 5u_7(7u_2z_0^2 - 1)v_0^2 \right. \\ & \quad \left. + 10u_7\dot{z}_0^2 + 35u_9(9u_2z_0^2 - 1)\sigma^2 - 140u_9\sigma z_0\dot{z}_0] + u_3(\mu'\mu'_3) \right\} \\ & + \frac{\tau^5}{120} u_7 [15\sigma(-3v_0^2 + 2u_1 + 7u_2\sigma^2)] + O(\tau^6), \end{aligned} \quad (9.31)$$

$$\begin{aligned} G = & \tau + \frac{\tau^3}{6} \left[-u_3 + \left(\frac{3J_2}{2} \right) (5u_7z_0^2 - u_5) - \mu'u'_3 \right] \\ & + \frac{\tau^4}{24} \left\{ 6u_5\sigma + \left(\frac{3J_2}{2} \right) [20u_7z_0\dot{z}_0 - 10u_7(7u_2z_0^2 - 1)\sigma] \right\} \\ & + \frac{\tau^5}{120} u_5 [9v_0^2 - 8u_1 - 45u_2\sigma^2] + O(\tau^6), \end{aligned} \quad (9.32)$$

$$F_z = F + \left(\frac{3J_2}{2} \right) \left[\frac{\tau^2}{2} (-2u_5) + \frac{\tau^3}{6} (10u_7\sigma) + \frac{\tau^4}{24} u_7 (10v_0^2 - 6u_1 - 70u_2\sigma^2) \right], \quad (9.33)$$

$$G_z = G + \left(\frac{3J_2}{2} \right) \left[\frac{\tau^3}{2} (-2u_5) + \frac{\tau^4}{24} (20u_7\sigma) \right], \quad (9.34)$$

where

$$\begin{cases} u_n = \frac{1}{r_0^n}, & \sigma = \vec{r}_0 \cdot \dot{\vec{r}}_0, v_0^2 = \dot{\vec{r}}_0 \cdot \dot{\vec{r}}_0, \\ u'_n = \frac{1}{r_0^{n+1}}, r'_0 = |\dot{\vec{r}}'_0|. \end{cases} \quad (9.35)$$

In the process we have simplified the part of the third-body perturbation, because an Earth's satellite is always far from a third body, the perturbation acceleration due to a third body in (9.27) can be replaced by $-\mu'(\vec{r}/r_0'^3)$, where $\vec{r}'_0$ is the position vector of the third-body at the epoch time t_0 , i.e., $\vec{r}'_0 = \vec{r}'(t_0)$. For a deep-space prober near its target celestial body, this simplification may not be proper to apply, but we can obtain the expressions for F , G , F_z , and G_z , which are introduced in Sect. 9.3.6. Note that in (9.31)–(9.35) the symbol $\sigma = O(e)$ is a small quantity in the same magnitude of the eccentricity e , do not be mixed up with the σ of orbit elements.

For an unperturbed motion model (which is the basic model for orbit determination in the early stage), the three components of F^* or G^* are the same, therefore the equations of orbit determination (9.26) are simpler than (9.30) as

$$\begin{cases} (Fv)x_0 - (F\lambda)z_0 + (Gv)\dot{x}_0 - (G\lambda)\dot{z}_0 = (vX_e - \lambda Z_e), \\ (Fv)y_0 - (F\mu)z_0 + (Gv)\dot{y}_0 - (G\mu)\dot{z}_0 = (vY_e - \mu Z_e), \\ (F\mu)x_0 - (F\lambda)y_0 + (G\mu)\dot{x}_0 - (G\lambda)\dot{y}_0 = (\mu X_e - \lambda Y_e). \end{cases} \quad (9.36)$$

For both perturbed and unperturbed models, the actual magnitudes of F^* and G^* using the dimensionless units are given by

$$F^* = 1 + O(s^2), \quad G^* = \tau[1 + O(s^2)], \quad (9.37)$$

$$s = n\tau, \quad n = a^{-\frac{3}{2}}, \quad (9.38)$$

where n is the mean angular speed of the satellite. For a short-arc problem, there is $|s| < 1$ and the power series solution is convergent. When $|s| < 1$ the corresponding arc is less than 1/6 of the orbit. If the epoch time is chosen as the time corresponding to the “middle point” of the whole sampling period, then the arc is less than 1/3 of the orbit, which fits the method of short-arc orbit determination.

For any other type of motion as long it can be provided by a mathematical model we can derive F^* and G^* . It is because the solution of the motion equation is given by derivatives of higher order, and the right-side functions of the equation only depend on $\vec{r}$, $\dot{\vec{r}}$, then the derivatives of higher-order, $\vec{r}_0^{(k)}$ ($k \geq 2$), can always be given by $\vec{r}_0$ and $\dot{\vec{r}}_0$.

(4) Calculation process of initial orbit determination

For a series of given angle measurements: $t_j; (\alpha_j, \delta_j)$ or $(A_j, h_j), j = 1, 2, \dots, k, k \geq 3$, the orbit at epoch t_0 can be determined by $\vec{r}_0$ and $\dot{\vec{r}}_0$, which are solutions of the basic Eqs. (9.30) or (9.36). The Eqs. (9.30) or (9.36) are linear algebra equations of $\vec{r}_0$ and $\dot{\vec{r}}_0$ in form, which can be solved when F , G , F_z , and G_z are known. Since F , G , F_z , and G_z are functions of $\vec{r}_0$ and $\dot{\vec{r}}_0$, to derive them we must use an iterative method. This

iteration is a simple and special process different from the multi-variable iteration used in the precise orbit determination. The magnitudes of F , G , F_z , and G_z are estimated as

$$\begin{cases} F = 1 + O(s^2), & G = \tau[1 + O(s^2)], \\ F_z = 1 + O(s^2), & G_z = \tau[1 + O(s^2)], \end{cases} \quad (9.39)$$

where $|s| < 1$. Without prior information of the target orbit, we can use the following values as the starting values of the iteration

$$F^{(0)} = 1, \quad G^{(0)} = \tau, \quad F_z^{(0)} = F^{(0)}, \quad G_z^{(0)} = G^{(0)}, \quad (9.40)$$

and use (9.31)–(9.34) to calculate F , G , F_z , G_z by iteration. The iteration stops when the results of $\vec{r}_0$ and $\dot{\vec{r}}_0$ meet the required accuracy. Then by the transformation relationship, the orbit elements at epoch t_0 can be derived from $\vec{r}_0$ and $\dot{\vec{r}}_0$. The transformation relationship is given in Sects. 2.4 and 2.6. The calculated orbit can be an ellipse or a hyperbolic, but the process of deriving $\vec{r}_0$ and $\dot{\vec{r}}_0$ does not involve the orbit type. The difference in the orbit types only appears in the transformation.

There are two points needed to be explained:

① About how to solve the linear Eqs. (9.30) or (9.36). Although an orbit can be determined by three sets of angle data, in practice an orbit is determined by as many samplings as possible. If there are k sets of sampling data, then there are $k \times m$ (here $m = 2$) conditions corresponding to $k \times 3$ basic equations, i.e., $k \times 3$ condition states. Usually $k \times 3 \gg n$, $n = 6$ is the dimension of an orbit. To fully use the available information, we need to choose an optimal estimation method to solve the equation of orbit determination (9.30) or (9.36). Based on the characteristics of initial orbit determination, including short-arc data collection and a realistic accuracy requirement, it is impossible and unnecessary to use a complicated estimation method. The commonly used method is the evenly weighted least squares (actually unweighted) estimator, sometimes with a procedure of removing wild data.

② About the criteria of controlling an iteration, i.e., when to stop an iteration. It is about how to decide the controlling parameters and the corresponding accuracy control. The decision is given based on the accuracy of sampling data and the state of the arc. It can be made according to the differences of parameters F , G , F_z , and G_z before and after an iteration, ΔF , ΔG , ΔF_z , and ΔG_z ; or the differences in orbital quantities $\vec{r}_0$ and $\dot{\vec{r}}_0$ before and after an iteration, $|\Delta \vec{r}_0|$, $|\Delta \dot{\vec{r}}_0|$; or the difference of an essential orbit element in σ (such as the orbital semi-major axis a), $\Delta \sigma$, before and after an iteration, etc.

From the above-described principles of orbit determination, although the iteration is for complicated functions, it is still a simple iteration, not a multi-variable iteration. The method does not demand initial information, an orbit can be determined for an uncooperative target. If at the same time we have measurements of distance, the process of the orbit determination method can be even simpler.

9.3.3 Initial Orbit Determination Using (ρ, A, h) Data or Navigation Information

(1) Initial orbit determination using (ρ, A, h) data

Some tracking methods can provide data of distance and position simultaneously, i.e., measurements of (ρ, A, h) . The accuracy of the distant variable ρ is often higher than the position variables (A, h) . Directly using ρ cannot simply decide an orbit, but can improve the accuracy of the result determined by (A, h) .

The position vector of a satellite $\vec{r} = \vec{r}(x, y, z)$ can be given by simultaneous measurements, ρ and (A, h) , that

$$\vec{r} = \vec{\rho} + \vec{R}, \quad (9.41)$$

where $\vec{\rho}$ is the position vector of the satellite from the tracking station. The relationship of $\vec{\rho}$ and the measured angles (A, h) in the O -xyz frame are given by

$$\vec{\rho} = \rho \hat{L}, \quad (9.42)$$

$$\hat{L} = \begin{pmatrix} \lambda \\ \mu \\ \nu \end{pmatrix} = (GR)^T(ZR) \begin{pmatrix} \cos h \cos A \\ -\cos h \sin A \\ \sin h \end{pmatrix}. \quad (9.43)$$

The transformation matrices (GR) and (ZR) and the coordinate vector of the tracking station $\vec{R}$ are provided in Sect. 9.3.2.

With the measurements of (ρ, A, h) , the corresponding basic equations of orbit determination are quite simple, given by

$$\begin{cases} Fx_0 + G\dot{x}_0 = x, \\ Fy_0 + G\dot{y}_0 = y, \\ Fz_0 + G_z\dot{z}_0 = z, \end{cases} \quad (9.44)$$

where F , G , F_z , and G_z are provided by (9.31)–(9.34). This method, therefore, is simpler than the previous one by (A, h) data only. For initial orbit determination, the data of (A, h) ensure the certainty of a solution, and the data of ρ improve the accuracy of the solution. This advantage can be confirmed by the actual results of orbit determination given in Sect. 9.3.4.

(2) Initial orbit determination using navigation information

There are two types of navigation information, one is the satellite navigation system (such as the American GPS and the Chinese BeiDou navigation system), which provides the space position of a target, $\vec{r} = \vec{r}(x, y, z)$; the other is the Astronomical navigation which provides the right ascension and declination of a satellite. We do not further discuss the second type.

For the satellite navigation, the orbit determination method is basically the same method for (ρ, A, h) data, the equations of orbit determination are the same as (9.44) except for the method of providing the data.

9.3.4 Examination of Orbit Determination Method Using Actual Measurements

The target of the examination is a low Earth orbit satellite. Because the arc is relatively short, the effect of the atmospheric drag on the orbit determination can be ignored. For the initial orbit determination, we only need to consider the gravitational force of Earth as a sphere and the main part of Earth's non-spherical gravity potential, the J_2 term. This dynamical model is close enough to the actual situation, and can ensure the reliability of the examination.

The measured photoelectric data (ρ, A, h) are from a ground-based tracking station, over the same arc there are corresponding GPS position measurements used as the test criterion, the accuracy of the GPS measurements is 15 m in the three spatial coordinate directions.

The epoch t_0 is 12 October 2005 17:21:16 (UTC), and the time interval of the arc is 336 s. The results of the orbit determination given in Table 9.1 are by three types of orbit determination method, which are Type 1 by (A, h) data, Type 2 by (ρ, A, h) data, and Type 3 by GPS results. For revealing the reliability of actual results by short-arc orbit determination we also provide the results of Type 4 given by precise orbit determination using GPS long arc data (1 day).

Comparing the results of the first two types, the improvements in the accuracy by the high precise distance data of ρ are clearly displayed. When we use data (ρ, A, h) from a Chinese tracking station (laser distance data of ρ , optical angle data of A and h) for the initial orbit determination results (not provided here) also show the same improvements. The errors of the orbit determination using angles-only data are mainly in the size (the semi-major axis a) and the shape (the eccentricity e) of the orbit. This is a well-known problem when using angles-only data for short-arc orbit determination. The reason can be found in the basic equations used in the process, readers may know the answer. Another question is why in Table 9.1 there are also data of λ ($\lambda = M + \omega$). The answer is that it is for reminding readers to pay attention to the property of a small eccentricity orbit.

Table 9.1 Examination of initial orbit determination results for a low Earth orbit satellite

Type	a (km)	e	i (deg)	Ω (deg)	ω (deg)	M (deg)	λ (deg)
1	6780.0203	0.0097680	42.41975	23.60852	34.55100	353.53114	28.08214
2	6716.1031	0.0012097	42.42771	23.55681	92.56757	295.55009	28.11766
3	6716.3742	0.0012695	42.43379	23.56349	89.09879	299.01709	28.11588
4	6716.3653	0.0012653	42.43365	23.56329	89.00358	299.11240	28.11598

Based on the results of short-arc orbit determination using data (ρ, A, h) and (A, h) we can conclude that the method of short-arc orbit determination introduced in this Chapter is concise and efficient. This method does not require any prior information of a target, with data of certain accuracy over a proper arc length the results are reliable, particularly when there are high accuracy distance data accompanied with the angle data, the results are obviously improved. Readers can figure out the reason. If there are high accuracy position data (such as GPS data) an initial orbit can be determined with high accuracy even by data over a short arc when the main perturbation force is included. The results of Type 3 in Table 9.1 are derived using GPS data, compared to that given by the precise orbit determination using data over a long arc, the error of the semi-major axis a is only in the order of 10 m.

9.3.5 Initial Orbit Determination When a Deep-Space Prober is on a Transfer Orbit

The motion of a prober from launching to entering an orbit around a target celestial body can be divided into three sections, which correspond to three different orbits. The three orbits are: a near Earth parking orbit, a transfer orbit of the transition, and a final orbit to become a satellite of the target body. Both the first orbit around Earth and the last orbit around the target celestial body are typical satellite orbits, about which the orbit determination problem is discussed in the previous content. In this section, we introduce the initial orbit determination for a satellite on the transfer orbit between two main bodies.

The motion of a prober on a transfer orbit is affected by both main bodies (for a Moon's prober, there are Earth and the Moon). When the prober is near Earth, the perturbation of Earth's dynamical form-factor (J_2) should be considered; during the transition period the Moon's gravity force must be included, and the motion of the prober is affected by Earth as an ellipsoid and by the Moon as a particle simultaneously. To deal with this problem we can construct a power series solution over a time interval $\Delta t = t - t_0$ (is referred to as time power series), which is appropriate for short-arc orbit determination. Because during this period the prober moves towards the target body (the Moon), the simplified method previously described for a third body in Sect. 9.3.2 is not proper anymore, we must build a different time power series solution. Specifically, for determining a reliable transfer orbit for ground-based measurement and control system the method to derive (9.29) and (9.31)–(9.34) must be modified [14, 15].

The above consideration has a practical background. When a Moon's prober is on its transfer orbit by today's observation technique the accuracy of angle data can reach $1''$, to fully use the high accuracy data we need a corresponding method of orbit determination.

(1) The time power series solution of $\vec{r}$ and $\dot{\vec{r}}$ with its particular expression

We use a Moon's prober as an example. The coordinate system is the same epoch J2000.0 geocentric celestial coordinate system, and the normalized dimensionless units are defined as

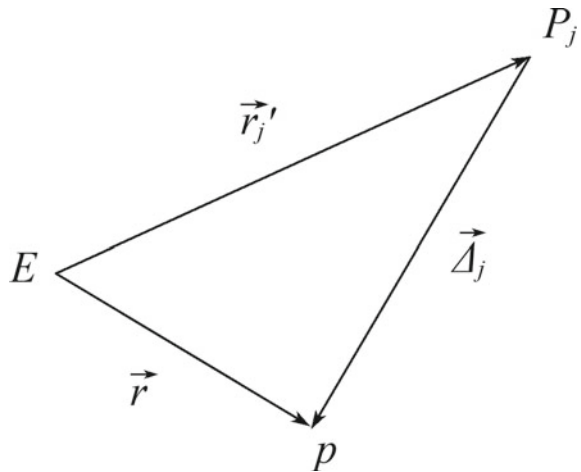
$$\begin{cases} [M] = E \text{ (Earth's mass),} \\ [L] = D \text{ (mean distance to the Moon),} \\ [T] = \left(\frac{D^3}{GE} \right)^{\frac{1}{2}} \approx 4^d.3484. \end{cases} \quad (9.45)$$

In this system, the ordinary differential motion equation of the prober provided by a dynamical model which includes Earth's central gravity force, the gravity force of Earth's dynamical form-factor J_2 term, and the gravity forces of two third-bodies (the Sun and the Moon) is given by

$$\begin{cases} \ddot{\vec{r}} = -\frac{\vec{r}}{r^3} + A_2 \left(\frac{5z^2}{r^7} - \frac{1}{r^5} \right) \vec{r} - A_2 \frac{2z}{r^5} \hat{k} - \sum_{j=1}^2 \mu'_j \left(\frac{\vec{\Delta}_j}{\Delta_j^3} + \frac{\vec{r}'_j}{r_j'^3} \right), \\ \vec{r}(t_0) = \vec{r}_0, \quad \dot{\vec{r}}(t_0) = \dot{\vec{r}}_0, \end{cases} \quad (9.46)$$

where $A_2 = 3J_2a_e^2/2$; $\hat{k} = (0, 0, 1)^T$ is the unit vector in the z -direction; a_e is the dimensionless equatorial radius of Earth's reference ellipsoid; $\vec{\Delta}_j = \vec{r}_j - \vec{r}'_j$; $\mu'_j = (m'_j/M_e)$; m'_j ($j = 1, 2$) is the mass of the Sun or the Moon, and $\vec{r}'_j$ is the position vector of the Sun or the Moon. The geometrical relationship of the center of Earth (E), a prober (p), and a third-body (P , the Sun or the Moon) is illustrated in Fig. 9.2.

Fig. 9.2 Geometrical relationship of Earth E , a main body P_j , and a prober p



In building the power series solution the part related to the Moon can be properly simplified as long as the probe is not near the Moon. Omitting the details and the process of derivation the time power series solution of the motion Eq. (9.46) is given as follows.

$$\begin{cases} x = Fx_0 + G\dot{x}_0 + F'x'_0 + G'\dot{x}'_0, \\ y = Fy_0 + G\dot{y}_0 + F'y'_0 + G'\dot{y}'_0, \\ z = F_z z_0 + G_z \dot{z}_0 + F'_z z'_0 + G'_z \dot{z}'_0, \\ \dot{x} = F_d x_0 + G_d \dot{x}_0 + F'_d x'_0 + G'_d \dot{x}'_0, \\ \dot{y} = F_d y_0 + G_d \dot{y}_0 + F'_d y'_0 + G'_d \dot{y}'_0, \\ \dot{z} = F_{dz} z_0 + G_{dz} \dot{z}_0 + F'_{dz} z'_0 + G'_{dz} \dot{z}'_0, \end{cases} \quad (9.47)$$

where

$$\begin{cases} F = 1 + \left(\frac{\tau^2}{2}\right)f_2 + \left(\frac{\tau^3}{6}\right)f_3 + \left(\frac{\tau^4}{24}\right)f_4 + \left(\frac{\tau^5}{120}\right)f_5 + \left(\frac{\tau^6}{720}\right)f_6 + O(\tau^7), \\ F_z = F + A_2 \left[\left(\frac{\tau^2}{2}\right)(-2u_5) + \left(\frac{\tau^3}{6}\right)(10u_7\sigma) + \left(\frac{\tau^4}{24}\right)(10u_7V_0^2 - 6u_8 - 70u_9\sigma^2) \right], \\ G = \tau + \left(\frac{\tau^3}{6}\right)g_3 + \left(\frac{\tau^4}{24}\right)g_4 + \left(\frac{\tau^5}{120}\right)g_5 + \left(\frac{\tau^6}{720}\right)g_6 + O(\tau^7), \\ G_z = G + A_2 \left[\left(\frac{\tau^3}{6}\right)(-2u_5) + \left(\frac{\tau^4}{24}\right)(20u_7\sigma) \right], \\ F' = \mu' \left[\left(\frac{\tau^2}{2}\right)f'_2 + \left(\frac{\tau^3}{6}\right)f'_3 + O(\tau^4) \right], \\ G' = \mu' \left[\left(\frac{\tau^3}{6}\right)g'_3 + O(\tau^4) \right], \end{cases} \quad (9.48)$$

$$\begin{cases} F_d = \tau f_2 + \left(\frac{\tau^2}{2}\right)f_3 + \left(\frac{\tau^3}{6}\right)f_4 + \left(\frac{\tau^4}{24}\right)f_5 + \left(\frac{\tau^5}{120}\right)f_6 + O(\tau^6), \\ F_{dz} = F_d + A_2 \left[\tau(-2u_5) + \left(\frac{\tau^2}{2}\right)(10u_7\sigma) + \left(\frac{\tau^3}{6}\right)(10u_7V_0^2 - 6u_8 - 70u_9\sigma^2) \right], \\ G_d = 1 + \left(\frac{\tau^2}{2}\right)g_3 + \left(\frac{\tau^3}{6}\right)g_4 + \left(\frac{\tau^4}{24}\right)g_5 + \left(\frac{\tau^5}{120}\right)g_6 + O(\tau^6), \\ G_{dz} = G_d + A_2 \left[\left(\frac{\tau^2}{2}\right)(-2u_5) + \left(\frac{\tau^3}{6}\right)(20u_7\sigma) \right], \\ F'_d = \mu' \left[\tau f'_2 + \left(\frac{\tau^2}{2}\right)f'_3 + O(\tau^4) \right], \\ G'_d = \mu' \left[\left(\frac{\tau^2}{2}\right)f'_3 + O(\tau^4) \right]. \end{cases} \quad (9.49)$$

The related quantities are given by

$$\begin{cases} f_2 = -u_3 + A_2 u_5 (5z_0^2 u_2 - 1) - \mu' u_3', \\ f_3 = 3\sigma u_5 + A_2 u_7 [5\sigma(1 - 7z_0^2 u_2) + 10z_0 \dot{z}_0] + \mu'(3u_5' \sigma'), \\ f_4 = u_5 (3V_0^2 - 2u_1 - 15\sigma^2 u_2) + A_2 [6u_8 (4z_0^2 u_2 - 1) - 5u_7 V_0^2 (7z_0^2 u_2 - 1) \\ \quad + 35u_9 \sigma^2 (9z_0^2 u_2 - 1) + 10u_7 \dot{z}^2 - 140u_9 \sigma z_0 \dot{z}_0], \\ f_5 = 15u_7 \sigma [2u_1 - 3V_0^2 + 7u_2 \sigma^2], \\ f_6 = u_7 [u_2 \sigma^2 (630V_0^2 - 420u_1 - 945u_2 \sigma^2) - (22u_2 - 66u_1 V_0^2 + 45V_0^4)], \end{cases} \quad (9.50)$$

$$\begin{cases} g_3 = f_2, \\ g_4 = 2f_3, \\ g_5 = u_5 [9V_0^2 - 8u_1 - 45u_2 \sigma^2], \\ g_6 = 30u_7 \sigma [5u_1 - 6V_0^2 + 14u_2 \sigma^2], \end{cases} \quad (9.51)$$

$$\begin{cases} f_2' = u_3' - \tilde{u}_3, \\ f_3' = 3(\tilde{u}_5 \sigma_1 - u_5' \sigma') \end{cases}, \quad (9.52)$$

$$g_3' = f_2', \quad (9.53)$$

$$\begin{cases} u_n = 1/r_0^n, & \sigma = \vec{r}_0 \cdot \dot{\vec{r}}_0, & V_0^2 = \dot{\vec{r}}_0^2, \\ u_n' = 1/\Delta_0^n, & \sigma' = \vec{\Delta}_0 \cdot \dot{\vec{\Delta}}_0, & V_0'^2 = \dot{\vec{\Delta}}_0^2, & \alpha = \vec{r}_0 \cdot \vec{\Delta}_0, \\ \tilde{u}_n = 1/r_0'^n, & \sigma_1 = \vec{r}_0' \cdot \dot{\vec{r}}_0', & V_1^2 = \dot{\vec{r}}_0'^2, & \alpha_1 = \vec{r}_0 \cdot \vec{r}_0', \\ \vec{\Delta}_0 = \vec{r}_0 - \vec{r}_0', & \dot{\vec{\Delta}}_0 = \dot{\vec{r}}_0 - \dot{\vec{r}}_0', \end{cases} \quad (9.54)$$

where $\vec{r}_0'$ and $\dot{\vec{r}}_0'$ are the values of $\vec{r}_j'$, $\dot{\vec{r}}_j'$ of a third-body at time t_0 .

(2) Method of orbit determination

When the probe is on its transfer orbit the method of orbit determination is similar to that for an Earth's satellite using short-arc data. The basic equation is constructed by the time power series solution and the measurements can be angles-only data or angle and position data.

① Angle data (α, δ)

The measurements (α, δ) are provided by the Deep-space network of measurement and control, we do not discuss the method and accuracy of data here. The three components of the unit vector $\hat{L}$ are (λ, μ, ν) , the relationship between (α, δ) and (λ, μ, ν) is given in (9.22), and the three components of the coordinate vector of the tracking station $\vec{R}$ is (X, Y, Z) . The corresponding basic equations of orbit determination are

$$\begin{cases} (F\nu)x_0 - (F_z\lambda)z_0 + (G\nu)\dot{x}_0 - (G_z\lambda)\dot{z}_0 = (\nu X - \lambda Z) - \Delta_1(\vec{r}_0', \dot{\vec{r}}_0'), \\ (F\nu)y_0 - (F_z\mu)z_0 + (G\nu)\dot{y}_0 - (G_z\mu)\dot{z}_0 = (\nu Y - \mu Z) - \Delta_2(\vec{r}_0', \dot{\vec{r}}_0'), \\ (F\mu)x_0 - (F\lambda)y_0 + (G\mu)\dot{x}_0 - (G\lambda)\dot{y}_0 = (\mu X - \lambda Y) - \Delta_3(\vec{r}_0', \dot{\vec{r}}_0'), \end{cases} \quad (9.55)$$

$$\begin{cases} \Delta_1(\vec{r}'_0, \dot{\vec{r}}'_0) = (F'\nu)x'_0 - (F'\lambda)z'_0 + (G'\nu)\dot{x}'_0 - (G'\lambda)\dot{z}'_0, \\ \Delta_2(\vec{r}'_0, \dot{\vec{r}}'_0) = (F'\nu)y'_0 - (F'\mu)z'_0 + (G'\nu)\dot{y}'_0 - (G'\mu)\dot{z}'_0, \\ \Delta_3(\vec{r}'_0, \dot{\vec{r}}'_0) = (F'\mu)x'_0 - (F'\lambda)y'_0 + (G'\mu)\dot{x}'_0 - (G'\lambda)\dot{y}'_0. \end{cases} \quad (9.56)$$

② Point position data $\vec{r}(x, y, z)$

The basic equations of orbit determination for point position data can be given directly by power series solutions as

$$\begin{cases} Fx_0 + G\dot{x}_0 = x - (F'x'_0 + G'\dot{x}'_0) \\ Fy_0 + G\dot{y}_0 = y - (F'y'_0 + G'\dot{y}'_0) \\ Fz_0 + G\dot{z}_0 = z - (F'z'_0 + G'\dot{z}'_0) \end{cases} \quad (9.57)$$

③ Procedure of orbit determination and numerical examination

The procedure is the same as previously described for an Earth's satellite, which is by iterating the Eqs. (9.55) or (9.57) without prior orbital information. The initial values for the iterating are

$$\begin{cases} F^{(0)} = 1, & F_z^{(0)} = 1, & F'^{(0)} = 0, \\ G^{(0)} = \tau, & G_z^{(0)} = 1, & G'^{(0)} = 0. \end{cases} \quad (9.58)$$

The results are $\vec{r}_0$ and $\dot{\vec{r}}_0$ at epoch t_0 , which can be transferred to orbital elements of an ellipse or a hyperbola.

Now we use a set of simulated data to examine this method. Choosing 10 January 2008 00:00:00 (TUC) as the initial time T_0 , and the initial orbit elements of a Moon's probe are

$$\begin{aligned} a &= 160450.0 \text{ km}, & e &= 0.95886569, & i &= 9^\circ.0, \\ \Omega &= 328^\circ.0, & \omega &= 212^\circ.0, & M &= 0^\circ.0. \end{aligned}$$

The calculation uses the normalized dimensionless units defined by (9.45). The available data are simulated angle data (there are 50 sets of samplings taken by every 1 min for 50 min continuously), and the coordinates of the assumed ground-based tracking station on Earth in the geocentric frame are (6372.412 km, 118°0.820916, 31°0.893611), and the results are listed in Table 9.2.

The determined orbit is for epoch t_0 measured from the initial time T_0 , and the 50 samplings are on the time interval of ($t_0 - 25 \text{ min.}$, $t_0 + 25 \text{ min.}$). Results are given by four models. For Model A, results are given by extrapolation using the complete dynamical model from T_0 to epoch t_0 , the resulting orbit is regarded as a standard orbit. For Model B, results are derived by the above-described orbit determination method for a two-body problem and the simulated data without adding errors. For

Table 9.2 Examination of initial orbit determination results for a Moon's prober

	a (km)	e	i (deg)	Ω (deg)	ω (deg)	$M + \omega + \Omega$ (deg)
<i>A</i>	156,772.885	0.9579274	8.97521	327.88653	212.16823	204.52762
<i>B</i>	156,995.190	0.9582140	8.95750	327.85395	212.17780	204.38883
<i>C</i>	156,772.805	0.9579272	8.97522	327.88655	212.16823	204.52770
<i>D</i>	156,774.451	0.9578618	8.97996	327.89556	212.16858	204.55371

Model *C*, results are derived for a complete dynamic model, and small errors of 10^{-8} are added to the simulated data. For Model *D*, small errors of 10^{-6} are added to the simulated data. In order to be close to a real situation, the errors added to the simulated data are not purely random. The added errors of 10^{-8} and 10^{-6} are in radian (with respect to the center of Earth).

The numerical results show that the time power series (9.47) itself has a comparable high accuracy, it can be used for short-arc orbit determination when a prober is on a transfer orbit. When a prober is far from Earth, as in this example the distance between the prober and Earth was 20 times the Earth's equatorial radius, and the perturbation effect of the Moon is greater than 10^{-4} . If we still use a two-body problem as a dynamic model, the errors can be large as shown in the results by Model *B*, indicating that the Moon's perturbation must be included. The method provided in this section is applicable to angles-only data, position and angle data, and point position data, but the arc should not be too short (depending on the state of measurements), otherwise the method fails due to the characteristics of the short-arc orbit determination, especially the value of the semi-major axis can be unrealistic. This method of initial orbit determination with a short arc for a prober in the transfer process can be applied by a ground-based measurement and control system.

9.3.6 Initial Orbit Determination Using Space-Based Angle Measurements (α , δ)

In the epoch J2000.0 geocentric celestial coordinate system O -xyz, the measurements satisfy the following geometric relationship (Fig. 9.1)

$$\vec{\rho} = \vec{r} - \vec{R}, \quad (9.59)$$

where $\vec{r}$ is the position vector of a satellite, $\vec{\rho}$ is the observed vector, $\vec{R}$ is the position vector of a space-based platform as a tracking station (strictly speaking, it is a "datum point" from where to observe the target satellite). The vectors of $\vec{\rho}$ and $\vec{r}$ satisfy the measurement equations and the related dynamical rules that

$$\vec{\rho} = \rho \hat{L}, \quad \hat{L} = \begin{pmatrix} \lambda \\ \mu \\ \nu \end{pmatrix} = \begin{pmatrix} \cos \delta \cos \alpha \\ \cos \delta \sin \alpha \\ \sin \delta \end{pmatrix}. \quad (9.60)$$

$$\vec{r} = \vec{r}(\sigma). \quad (9.61)$$

In (9.61) σ is the orbital element of the target orbit. In the same coordinate system, the position vector of the tracking station, now is the datum point, $\vec{R}$, which is given by the orbit of the space-based platform and can be written as

$$\vec{R} = (X_e, Y_e, Z_e)^T. \quad (9.62)$$

Whether or not an orbit can be determined using the angles-only data (α, δ) is a question about the observability. The angles (α, δ) are observable quantities, as long there is not any geometric flaw, theoretically, an orbit can be determined by three sets of samplings. Since a space-based platform can provide the needed samplings, an orbit determination can be achieved.

(1) Basic equations of initial orbit determination

The dynamical condition of the orbit determination is the same as the power series of time interval Δt ($\Delta t = t - t_0$) given by (9.29), i.e.,

$$\vec{r}(t) = F^*(\vec{r}_0, \dot{\vec{r}}_0, \Delta t) \vec{r}_0 + G^*(\vec{r}_0, \dot{\vec{r}}_0, \Delta t) \dot{\vec{r}}_0, \quad (9.63)$$

where F^* and G^* are given before. The basic equations of orbit determination are the same as (9.30) that

$$\begin{cases} (F\nu)x_0 - (F_z\lambda)z_0 + (G\nu)\dot{x}_0 - (G_z\lambda)\dot{z}_0 = (\nu X_e - \lambda Z_e), \\ (F\nu)y_0 - (F_z\mu)z_0 + (G\nu)\dot{y}_0 - (G_z\mu)\dot{z}_0 = (\nu Y_e - \mu Z_e), \\ (F\mu)x_0 - (F\lambda)y_0 + (G\mu)\dot{x}_0 - (G\lambda)\dot{y}_0 = (\mu X_e - \lambda Y_e), \end{cases} \quad (9.64)$$

where the coordinates of the tracking station (X, Y, Z) now are given by the ephemeris of the space-based platform.

(2) Calculation process of initial orbit determination using angle samplings (α, δ)

From a series of samplings $t_j: (\alpha_j, \delta_j), j = 1, 2, \dots, k$, if $k \geq 3$, then $\vec{r}_0$ and $\dot{\vec{r}}_0$ at epoch $t_0, t_0 \in (t_1, t_k)$, can be derived using (9.64), by which an orbit can be determined. The process of iteration is similar to that of a ground-based tracking station.

(3) A numerical example of initial orbit determination by angles-only data (α, δ)

Assuming a space-based platform to be a Sun-synchronous satellite on a near circle orbit, a target to be a satellite on a geosynchronous orbit (GEO), the generated simulation data are from a “true” orbit (epoch UTC = 2,454,101.572916666 or 2007/01/01 01:45:0.0) with orbital elements:

$$\begin{aligned}
 a &= 42241.2122, \quad e = 0.00000210, \quad i = 0.000018, \\
 \Omega &= 350.635953, \quad \omega = 213.580436, \quad M = 13.033569, \\
 M + \omega + \Omega &= 217.249957.
 \end{aligned}$$

The time interval containing the generated samplings is 3.5 h, corresponding to a short arc, about 1/7 of a circle. We then add random errors of 5'' to the simulated samplings (α_j, δ_j) . Results are

$$\begin{aligned}
 a &= 42250.6323, \quad e = 0.00017521, \quad i = 0.000173, \\
 \Omega &= 348.960379, \quad \omega = 230.252322, \quad M = 358.037628, \\
 M + \omega + \Omega &= 217.250329.
 \end{aligned}$$

The accuracy of the results is 10^{-4} and the random difference is 2.5×10^{-5} , which agrees with the property of short-arc (1/7 circle) orbit determination, and the errors are mostly in the semi-major axis a and the eccentricity e .

It should be mentioned that when the orbit of a target body is near the orbit of the space-based platform, in the case of a short-arc, there are $F \approx 1$ and $G \approx \tau$, the iteration may converge to the orbit of the platform. This phenomenon can be explained by the basic equations of orbit determination (9.64). Of course, this phenomenon cannot happen to a ground-based tracking station. When it happens some additional conditions are needed, which are not discussed in this book.

9.3.7 A Brief Summary of Initial Orbit Determination

The method of initial orbit determination provided in this section is not restricted to the non-perturbed motion model (i.e., the model of the two-body problem), it is also applicable for orbit determination by high precise measurements, and for different types of orbits including hyperbolas, even missiles with orbital semi-major axis less than the equatorial radius of Earth, as long as there is a dynamical model.

About the basic principles of orbit determination, no matter whether it is an initial orbit determination or a precise orbit determination (introduced in the next Chapter), by measurements over a short-arc or a long-arc, there is no fundamental difference. But there are major differences in the method and procedure, because there are differences in the demands and prerequisites. It should be underlined that the initial orbit determination has particular limitations, in most cases, there is no initial information (or the prior information). Under this condition to determine an orbit of a space target the iteration method is different from that for a precise orbit determination, which is an iteration of multi-variables requiring initial information.

References

1. Plummer HC (1918) An introductory treatise on dynamical astronomy. Cambridge the University Press
2. Smart WM (1953) Celestial mechanics. University of Glasgow
3. Taff LG (1984) On initial orbit determination. *Astro. J.* 89(12):1426–1428
4. Taff LG (1985) Celestial mechanics: a computational guide for the practitioner. A Wiley-Interscience publication, New York
5. Brouwer D, Clemence GM (1961) Methods of celestial mechanics. Academic Press, New York and London
6. Morton BG, Taff LG (1986) A new method of initial orbit determination. *Celest. Mech.* Springer 39(2):181–190
7. Danby JMA (1992) Fundamentals of celestial mechanics. Willmann-Bell, Richmond, Virginia
8. Battin RH (1999) An Introduction to the mathematics and methods of astrodynamics. AIAA Education Series, American Institute of Aeronautics and Astronautics Inc, Reston, Virginia
9. Beutler G (2005) Methods of celestial mechanics. Springer-Verlag, Berlin, Heidelberg
10. Liu L (1992) Orbital dynamics of artificial earth satellite. Higher Education Press, Beijing
11. Liu L (2000) Orbital theory of spacecraft. National Defense Industry Press, Beijing
12. Liu L, Wang X (2003) A method of orbit computation taking into account the Earth's oblateness. *China Astron. Astrophys.* 27(3):335–339
13. Liu L, Wang JF (2004) Initial orbit calculation. *J of Spacecraft TT&C Technology* 23(3):41–50
14. Liu L, Zhang W (2007) A method of short-arc orbit determination for the transfer orbit of Lunar probes. *ATCA Astronomica Sinica* 48(2):220–227, and *China Astron. Astrophys.* (2007) 31(3):228–295
15. Liu L, Zhang W (2009) Initial orbit determination methods for different types of measurements. *J of Spacecraft TT&C Technology* 28(3):70–76

Chapter 10

Precise Orbit Determination



10.1 Precise Orbit Determination: Orbit Determination and Parameter Estimation

In Sect. 9.1, there is a simple definition of precise orbit determination compared to the initial orbit determination. The precise orbit determination is to decide an orbit of a celestial body or a spacecraft in motion based on a large number of observations, and to provide some geometric and physical parameters related to orbital motion at the same time. By this definition, the precise orbit determination includes the orbit determination and the parameter estimation, and is an expansion of the traditionally simple orbit improvement [1–3].

To be consistent with the initial orbit determination, in the precise orbit determination we also denote X and Y to the state variable and the measurement variable, respectively. The state variable X includes satellite orbital variables, $\vec{r}$, $\dot{\vec{r}}$ (or the set of orbital elements σ), and a to-be-estimated parameter β , i.e.,

$$X = \begin{pmatrix} \vec{r} \\ \dot{\vec{r}} \\ \beta \end{pmatrix} \quad \text{or} \quad X = \begin{pmatrix} \sigma \\ \beta \end{pmatrix}, \quad (10.1)$$

where $\vec{r}$ and $\dot{\vec{r}}$ are the position vector and velocity of a moving body in the chosen frame; σ is for orbital elements (note that the symbol σ does not mean the same as in (9.35)), usually the Kepler elements or the non-singularity elements. The variables $(\vec{r}, \dot{\vec{r}})$ and σ are equivalent, which set of variables to use depends on specific problems. The to-be-estimated β is a parameter, which involves related physical forces, geometric factors of the tracking station, and other parameters related to the moving body. The state variable X is an n -dimensional vector with $n \geq 6$, if the problem is only for improving an existing orbit, then $n = 6$. The types of measurement variable Y include optical angle variables (the equatorial coordinates α and δ , or the horizontal coordinates A and h), radar and laser distance variable ρ , Doppler speed variable $\dot{\rho}$,

radar distance and angle variables (ρ, A, h) , satellite height variable H , and satellite navigation position variable $\vec{r}$, etc. Any quantities related to the target orbit can be used as measurements. The observatory can be Earth-based or space-based, and the two types of observatory have different properties of motion. The measurement variable Y is an m -dimensional vector with $m \geq 1$, such as $m = 2$ for angles-only variables.

The state variable X satisfies the following ordinary differential equation with initial values

$$\begin{cases} \dot{X} = F(X, t), \\ X(t_0) = X_0. \end{cases} \quad (10.2)$$

The solution of (10.2) is called the state equation that

$$X(t) = X(t_0, X_0; t) \quad (10.3)$$

This solution can be analytical formulas obtained by analytically solving the Eq. (10.2), or in a discrete form obtained by numerically solving (10.2). Both solutions are complicated and nonlinear, so it is impossible to provide deduce the state transition matrix $\Phi(t_0, t)$ in a direct and simple way. This matrix is defined as

$$\Phi(t_0, t) = \left(\frac{\partial X}{\partial X_0} \right), \quad (10.4)$$

which is an $(n \times n)$ square matrix.

The measurement variable Y and the state variable X satisfy the following relationship

$$Y = H(X, t). \quad (10.5)$$

The corresponding measurement equation is

$$Y_O = H(X, t) + V, \quad (10.6)$$

where Y_O is for the measured values of Y ; H is for the theoretical values of Y ; V is the error of measurement. In the precise orbit determination, by principle, V is supposed to include only the random errors, and by theory, it is supposed to be white noise. The systematic error of measurement can be given by a systematic error model V' and be added to the right side of Eq. (10.6). According to the above-mentioned measurement types, $H(X, t)$ is a nonlinear function.

To obtain $X(t)$ we must know the to-be-estimated state variable X_0 at epoch t_0 . Because the actual values of X_0 or the ideal values of X_0 with certain accuracy at t_0 usually are unknown in advance, therefore we cannot obtain $X(t)$ directly. If we define an approximation of the to-be-estimated state variable X_0 , denoted by X_0^* , as

the reference state variable, and the corresponding $X(t)$ is denoted by X^* , then we have

$$X^*(t) = X(t_0, X_0^*; t). \quad (10.7)$$

We then expand the Eq. (10.6) at X_0^* and discard the high-order terms (i.e., linearization) to give the basic equation of precise orbit determination—the condition equation, in the form:

$$y = \tilde{B}x_0 + V, \quad (10.8)$$

where

$$y = Y_O - Y_C, \quad Y_C = H(X^*, t), \quad (10.9)$$

$$x_0 = X_0 - X_0^*, \quad (10.10)$$

$$\tilde{B} = \left(\left(\frac{\partial H}{\partial X} \right) \left(\frac{\partial X}{\partial X_0} \right) \right)_{X^*}, \quad (10.11)$$

where Y_C is a calculated “approximation” of Y ; y is called the residual; x_0 is the correction of the to-be-estimated state variable X_0 ; $\left(\frac{\partial H}{\partial X} \right)$ in $\tilde{B}$ is the matrix of measurement, which has a strict formula for any type of data; $\left(\frac{\partial X}{\partial X_0} \right)$ is the above-mentioned state transition matrix Φ in (10.4). The condition Eq. (10.8) is a linear equation given by omitting higher order terms of $O(x_0^2)$, the omitted value is compensated by an iteration process.

The precise orbit determination of a moving body now can be achieved by using a large amount of observed data $t_j, Y_j (j = 1, \dots, k)$ to solve the condition Eq. (10.8), and the result is a correction $\hat{x}_{0/k}$ for the to-be-estimated epoch state variable X_0 that $X_0 = \hat{X}_{0/k}$ and

$$\hat{X}_{0/k} = X_0^* + \hat{x}_{0/k}, \quad (10.12)$$

where $\hat{x}_{0/k}$ and $\hat{X}_{0/k}$ are the solutions of the condition Eq. (10.8), which is solved using a kind of estimator of optimal method; the subscript k is for the number of measurements.

The precise orbit determination in fact is a process of iteration. The first value of $\hat{X}_{0/k}$ by nature does not have the accuracy as demanded, but can be used as an approximation for the next iteration, until the required accuracy is reached. This iteration method is often used for reducing the truncated errors caused by the linearization of the measurement equation. During the process of iteration, the final correction x_0

is relatively small, therefore, in the condition Eq. (10.8) the treatment of matrix $\tilde{B}$ is less strict than that of $H(X^*, t)$. In the calculation of $H(X^*, t)$ we must do our best to match the accuracy of the observed variable Y_0 , otherwise, the high accuracy data would lose their meaning. In the calculation of $\tilde{B}$ the complicity is from the calculation of the transition matrix Φ , because the demand on $\tilde{B}$ is not high, it is reasonable to properly simplify the calculation.

From the process of the precise orbit determination, we know that the accuracy of the result depends on the accuracy of the measured data Y_0 , which involves related data treatment techniques, we do not discuss the details of the treatment in this book. From the point of view of orbit determination, we assume that the measured data Y_0 has certain high accuracy. In this section, we provide the three parts of the calculation of the precise orbit determination as follows.

- (1) Calculation of the theoretical measurements. This is to calculate $Y_C = H(X^*, t)$, which involves solving the state differential Eq. (10.2), analytically or numerically.
- (2) Calculation of matrix $\tilde{B}$.
- (3) Solving the condition Eq. (10.8).

For the first part, in the calculation of the theoretical values of $Y_C = H(X^*, t)$ the key issue is to calculate the state variable X (i.e., the orbital variables), which is to derive $X^*(t)$ from the given initial X_0^* , and $X^*(t)$ is the state variable corresponding to the sampling time t . Because the precise orbit determination uses the information over a long arc, the method of power series solution of time interval $\Delta t = t - t_0$ in the form of (9.15) is not applicable. The power series solution of (9.15) is given by

$$\begin{cases} \vec{r}(t) = \vec{r}_0 + \vec{r}_0^{(1)} \Delta t + \frac{1}{2!} \vec{r}_0^{(2)} \Delta t^2 + \cdots + \frac{1}{k!} \vec{r}_0^{(k)} \Delta t^k + \cdots, \\ \vec{r}_0^{(k)} = \left(\frac{d^k \vec{r}}{dt^k} \right)_{t=t_0}. \end{cases} \quad (10.13)$$

For the precise orbit determination, we must rigorously strictly solve the ordinary differential equation with the initial values condition (10.2) which is

$$\begin{cases} \dot{X} = F(X, t), \\ X(t_0) = X_0. \end{cases}$$

The accuracy of the solution must be agreed with the accuracy of the observed data, otherwise, the available high accuracy data are no longer valuable. This is the main difference between the long-arc precise orbit determination and the short-arc initial orbit determination, and to solve (10.2) requires multi-variable iteration.

There are different ways to solve Eq. (10.2) for obtaining to obtain the state variable $X^*(t)$ at time t . If we use the analytical method provided in Chap. 4 to solve Eq. (10.2), then this method is the analytical orbit determination method. If we use the numerical method provided in Chap. 8 to solve the equation, then the method is called the numerical orbit determination method. The method of transferring the

orbital variables to the corresponding observed geometric variables (α, δ) , (A, h) , ρ , and $\dot{\rho}$, etc., are discussed in Sect. 10.2.

For the second part, the calculation of the matrix $\tilde{B}$ includes calculations of two matrices, the measurement matrix $(\frac{\partial H}{\partial X})$ and the transition matrix of the state variables, $\Phi = (\frac{\partial X}{\partial X_0})$. The first matrix calculation can be completed by strict formulas. The calculation of the second matrix needs to know each element Φ_{ij} . If we use the analytical method, it is relatively easy to give Φ_{ij} by analytical expressions. If we use the numerical method, we have to solve a high-dimension differential equation for Φ , which is

$$\begin{cases} \dot{\Phi} = A\Phi, \\ A = \left(\frac{\partial F}{\partial X} \right)_{X^*}, \end{cases} \quad (10.14)$$

where $F = F(X, t)$ is the right-side function of the state differential Eq. (10.2). Because each element in the matrix $A = (a_{ij})$ is calculated using the reference state variable X^* , therefore Eq. (10.14) is a linear ordinary differential equation.

The third part is about solving the condition Eq. (10.8), which is the most important and critical procedure in the precise orbit determination. Under the condition of available high accurate measured data with precise geometric properties, this procedure decides the optimal result of the precise orbit determination. The goal can be achieved by sufficiently using the statistical information provided by a large amount of measurement and many mathematical methods related to the optimal estimation theory. About the optimal estimation theory, we mainly introduce the commonly used linear least squares estimator; the theory and related problems are given in Sect. 10.5.

10.2 Theoretical Calculation of Measurement Variables

The theoretical calculation of the measurement variable $Y_C = H(X^*, t)$ involves the transformation from the orbital state variables $(\vec{r}, \dot{\vec{r}}$ or orbital elements) to the geometric variables such as the observed variables (α, δ) , (A, h) , ρ , and $\dot{\rho}$. The process of the transformation is not complicated. In this section, we provide the calculation formulas for theoretical values of the measurement variables in the epoch J2000.0 geocentric celestial coordinate system.

The geometric relationship of the measurement vector variable $\vec{\rho}$, the coordinate vector of the tracking station $\vec{r}_e$, and the coordinate vector variable of a moving body $\vec{r}$ in the same frame is

$$\vec{\rho} = \vec{r} - \vec{r}_e. \quad (10.15)$$

The distance ρ can be given by the position vector $\vec{r}(x, y, z)$. and the coordinate vector of the tracking station $\vec{r}_e(x_e, y_e, z_e)$ that

$$\rho = |\vec{r} - \vec{r}_e| = [(x - x_e)^2 + (y - y_e)^2 + (z - z_e)^2]^{\frac{1}{2}}. \quad (10.16)$$

The coordinate vector of the tracking station $\vec{r}_e(x_e, y_e, z_e)$, if it is Earth-based, is usually given in the Earth-fixed coordinate system (which should be related to the reference ellipsoid), denoted by $\vec{R}_e(X_e, Y_e, Z_e)$. The relationship of $\vec{r}_e$ and $\vec{R}_e(X_e, Y_e, Z_e)$ is

$$\vec{r}_e = (HG)^T \vec{R}_e, \quad (10.17)$$

where the transformation matrix (HG) is formed by the precession matrix, the nutation matrix, the rotation matrix, and the polar motion matrix, which are provided in (1.30).

If we write

$$\vec{\rho} = \rho \hat{\rho}, \quad \rho = |\vec{r} - \vec{r}_e|,$$

where ρ and $\hat{\rho}$ are the distance from the tracking station to the moving body and the unit vector in this direction, respectively. We can derive the relative speed

$$\dot{\rho} = \frac{1}{\rho} (\vec{r} - \vec{r}_e) \cdot (\dot{\vec{r}} - \dot{\vec{r}}_e), \quad (10.18)$$

where

$$\dot{\vec{r}}_e = \frac{d}{dt} (HG)^T \vec{R}_e. \quad (10.19)$$

The variations of the precession and nutation are about 3×10^{-9} , and the polar is about 10^{-10} . Both are very small and can be omitted. According to the property of the matrix (HG) , only Earth's rotation variation rate needs to be considered, then

$$\frac{d}{dt} (HG)^T = (HG)_S^T \dot{S}_G, \quad (10.20)$$

where the matrix $(HG)_S$ is the part of matrix (HG) which is $R_z(S_G)$, S_G is the Greenwich sidereal time, given by (1.35) or (1.65) and R_z is given by (1.22). The derivative at t is changed to the derivative at S_G , the sidereal time, that

$$\frac{d}{dS} R_z(S_G) = \begin{pmatrix} -\sin S_G & \cos S_G & 0 \\ -\cos S_G & -\sin S_G & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (10.21)$$

and the value of $\dot{\bar{S}}_G$ is given by

$$\dot{\bar{S}}_G = 360^\circ.985647365/\text{d}. \quad (10.22)$$

From the definition of the tracking coordinates $\vec{r}_e$ and its variation rate $\dot{\vec{r}}_e$ we know that the theoretical calculation of the four types of measurement variables, ρ , $\dot{\rho}, (\alpha, \delta)$, and (A, h) is also related to the calculation of $\vec{r}, \dot{\vec{r}}$ and unit vectors $\hat{\rho}(\alpha, \delta)$ or $\hat{\rho}(A, h)$. We now provide the formulas.

(1) The distance variable ρ

$$\rho = |\vec{r} - \vec{r}_e| = [(x - x_e)^2 + (y - y_e)^2 + (z - z_e)^2]^{\frac{1}{2}}. \quad (10.23)$$

The coordinates of $\vec{r}$ in the epoch geocentric celestial coordinate system are given by

$$\vec{r} = \vec{r}(\sigma) = r \cos u \hat{P} + r \sin u \hat{Q}, \quad (10.24)$$

$$r = |\vec{r}|, \quad u = f + \omega, \quad (10.25)$$

$$\hat{P} = \begin{pmatrix} \cos \Omega \\ \sin \Omega \\ 0 \end{pmatrix}, \quad \hat{Q} = \begin{pmatrix} -\sin \Omega \cos i \\ \cos \Omega \cos i \\ \sin i \end{pmatrix}. \quad (10.26)$$

The formulas for calculating $\vec{r}$ can be used for Kepler orbital elements as well as the non-singularity orbital elements of the first type (when e is small). The calculation method of angle $u = f + \omega$ is as follows.

For Kepler orbital elements,

$$E - M = e \sin E \text{ (Kepler equation)}, \quad (10.27)$$

$$\left(\frac{a}{r}\right) = (1 - e \cos E)^{-1}, \quad (10.28)$$

$$\sin(f - E) = \left(\frac{a}{r}\right) e \sin E \left(1 - \frac{1}{1 + \sqrt{1 - e^2}} e \cos E\right). \quad (10.29)$$

For the non-singularity orbital elements of the first type, it requires to solve the general Kepler equation that

$$\tilde{u} - \lambda = \xi \sin \tilde{u} - \eta \cos \tilde{u}, \quad (10.30)$$

$$\frac{a}{r} = [1 - (\xi \cos \tilde{u} + \eta \sin \tilde{u})]^{-1}, \quad (10.31)$$

$$\sin(u - \tilde{u}) = \left(\frac{a}{r}\right)(\tilde{u} - \lambda) \left[1 - \frac{1}{1 + \sqrt{1 - e^2}} (\xi \cos \tilde{u} + \eta \sin \tilde{u}) \right]. \quad (10.32)$$

For the non-singularity orbital elements of the second type (both e and i are small), the calculations are similar, which are provided in Chap. 3 (3.49)–(3.51).

(2) The speed variable $\dot{\rho}$

$$\dot{\rho} = \frac{1}{\rho} (\vec{r} - \vec{r}_e) \cdot \left(\dot{\vec{r}} - \dot{\vec{r}}_e \right), \quad (10.33)$$

where the formulas for calculating ρ and $\vec{r}$ are already provided, and the formula for $\dot{\vec{r}}$ is similar to that of $\vec{r}$ which is

$$\dot{\vec{r}} = \sqrt{\frac{\mu}{p}} \left[(-\sin u - e \sin \omega) \hat{P} + (\cos u + e \cos \omega) \hat{Q} \right]. \quad (10.34)$$

This formula is also applicable to both Kepler orbital elements and the non-singularity orbital elements of the first type. The calculation of u is the same as given above. When both e and i are small, the calculation formula is (3.60).

(3) The right ascension and declination variables (α , δ)

The right ascension and declination (α , δ) are decided through positions of background stars no matter by photo observation or CCD technique, in the epoch geocentric celestial coordinate system. If we denote (α_0, δ_0) to the theoretical values of (α , δ), then

$$\hat{\rho}_0 = \begin{pmatrix} \cos \delta_0 \cos \alpha_0 \\ \cos \delta_0 \sin \alpha_0 \\ \sin \delta_0 \end{pmatrix} = \frac{(\vec{r} - \vec{r}_e)}{\rho}, \quad (10.35)$$

where $\rho = |\vec{r} - \vec{r}_e|$, $\vec{r}$ and $\vec{r}_e$ are calculated using corresponding orbital elements and the coordinate vector of the tracking station.

Let

$$\vec{\rho} = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = (\vec{r} - \vec{r}_e), \quad (10.36)$$

then the theoretical values of (α_c, δ_c) are calculated by

$$\begin{cases} \alpha_c = \arctan\left(\frac{y'}{x'}\right), \\ \delta_c = \arcsin\left(\frac{z'}{\rho}\right). \end{cases} \quad (10.37)$$

(4) The azimuth and height variables (A, h)

The measurement angles (A, h) correspond to the instantaneous true horizontal coordinates, the related position vector is given by

$$\vec{\rho} = \rho \hat{A}_h, \quad \hat{A}_h = \begin{pmatrix} \cos h \cos A \\ -\cos h \sin A \\ \sin h \end{pmatrix}. \quad (10.38)$$

The azimuth A (also called the horizontal longitude) is measured from the north point in the horizontal plane to the east, the height is measured from the horizontal longitude along the azimuth circle to the zenith direction.

In the epoch geocentric celestial coordinate system, the theoretical values in (10.38), denoted to $(\hat{A}_h)_c$, are given by the calculated values of $(\vec{r}(\sigma) - \vec{r}_e)/\rho$ and coordinate transformations that

$$(\hat{A}_h)_c = \begin{pmatrix} \cos h_c \cos A_c \\ -\cos h_c \sin A_c \\ \sin h_c \end{pmatrix} = (ZR)^T (GR) \left[\frac{(\vec{r} - \vec{r}_e)}{\rho} \right], \quad (10.39)$$

where $(GR) = (NR)(PR)$, and (PR) and (NR) are the precession matrix and the nutation matrix, respectively; (ZR) is the transformation matrix of the instantaneous true equatorial frame and the horizontal frame, that

$$\begin{cases} (ZR) = R_z(\pi - S) R_y\left(\frac{\pi}{2} - \varphi\right), \\ S = S_G + \lambda_G. \end{cases} \quad (10.40)$$

Here S_G is the Greenwich sidereal time; λ_G and φ are the longitude and latitude of the tracking station (astronomical latitude); and the matrices R_y and R_z are given in (1.21) and (1.22), respectively.

From (10.39) we can calculate (A_c, h_c) . Let

$$\vec{\rho} = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = (ZR)^T (GR) (\vec{r} - \vec{r}_e), \quad (10.41)$$

we then have

$$\begin{cases} A_c = \arctan\left(-\frac{y'}{x'}\right), \\ h_c = \arcsin\left(\frac{z'}{\rho}\right). \end{cases} \quad (10.42)$$

10.3 Calculation of Transformation Matrixes

In Sect. 10.1, the relationship among the measurement variable Y , the state variable X , and the matrix $\tilde{B}$ in the condition Eq. (10.8), are presented as

$$Y = H(X, t), \quad (10.43)$$

$$\tilde{B} = \left(\frac{\partial Y}{\partial X} \right) \left(\frac{\partial X}{\partial X_0} \right). \quad (10.44)$$

The matrix $\tilde{B}$ involves two sets of partial derivatives. These are the basic formulas for precise orbit determination. If the orbital variables are the position vector and the velocity of the moving body $\vec{r}, \dot{\vec{r}}$, then the partial derivatives involve

$$\frac{\partial Y}{\partial(\vec{r}, \dot{\vec{r}})}, \quad \frac{\partial(\vec{r}, \dot{\vec{r}})}{\partial(\vec{r}_0, \dot{\vec{r}}_0)}.$$

If the orbital variables are the orbital elements of the motion, then the corresponding partial derivatives involve

$$\frac{\partial Y}{\partial(\vec{r}, \dot{\vec{r}})}, \quad \frac{\partial(\vec{r}, \dot{\vec{r}})}{\partial\sigma}, \quad \frac{\partial\sigma}{\partial\sigma_0}.$$

The partial derivatives $\frac{\partial Y}{\partial(\vec{r}, \dot{\vec{r}})}$ and $\frac{\partial(\vec{r}, \dot{\vec{r}})}{\partial\sigma}$ corresponding to the measurement matrix $\left(\frac{\partial Y}{\partial X} \right)$ can be provided rigorously; whereas $\frac{\partial(\vec{r}, \dot{\vec{r}})}{\partial(\vec{r}_0, \dot{\vec{r}}_0)}$ and $\frac{\partial\sigma}{\partial\sigma_0}$ corresponding to the state transition matrix $\left(\frac{\partial X}{\partial X_0} \right)$ can be provided in simple expressions for certain accuracy.

In this section, we present the partial derivatives with respect to four types of measurement variable Y , which are (ρ) , $(\dot{\rho})$, (α, δ) , and (A, h) at each orbit element in σ . There are three sets of orbital elements σ , which are Kepler orbital elements $(a, e, i, \Omega, \omega, M)$, the non-singularity orbital elements of the first type $(a, i, \Omega, \xi = e \cos \omega, \eta = e \sin \omega, \lambda = M + \omega)$, and the non-singularity of the second type

$$a, \quad \xi = e \cos \tilde{\omega}, \quad \eta = e \sin \tilde{\omega},$$

$$h = \sin \frac{i}{2} \cos \Omega, \quad k = \sin \frac{i}{2} \sin \Omega, \quad \lambda = M + \tilde{\omega}.$$

10.3.1 Matrix $\frac{\partial Y}{\partial(\vec{r}, \dot{\vec{r}})}$

(1) Distance measurement variable ρ

$$\left(\frac{\partial \rho}{\partial(\dot{\vec{r}})} \right) = 0, \text{ and}$$

$$\left(\frac{\partial \rho}{\partial \vec{r}} \right) = \frac{1}{\rho} (\vec{r} - \vec{r}_e)^T = \frac{1}{\rho} ((x - x_e), (y - y_e), (z - z_e))^T, \quad (10.45)$$

$$\rho = |\vec{r} - \vec{r}_e|. \quad (10.46)$$

(2) Speed measurement variable ρ

$$\left(\frac{\partial \dot{\rho}}{\partial \vec{r}} \right) = \frac{1}{\rho} \left((\dot{\vec{r}} - \dot{\vec{r}}_e) - \frac{\dot{\rho}}{\rho} (\vec{r} - \vec{r}_e) \right)^T, \quad (10.47)$$

$$\left(\frac{\partial \dot{\rho}}{\partial \dot{\vec{r}}} \right) = \frac{1}{\rho} (\vec{r} - \vec{r}_e)^T, \quad (10.48)$$

where

$$\dot{\rho} = \frac{1}{\rho} (\vec{r} - \vec{r}_e) \cdot (\dot{\vec{r}} - \dot{\vec{r}}_e). \quad (10.49)$$

(3) Angle measurement variables (α, δ)

$\left(\frac{\partial(\alpha, \delta)}{\partial \dot{\vec{r}}} \right) = 0$, so we only need $\left(\frac{\partial(\alpha, \delta)}{\partial \vec{r}} \right)$. Since the angle α in the related residual y appears in the form of $\cos \delta(\alpha - \alpha_c)$ then the corresponding matrix $\left(\frac{\partial(\alpha, \delta)}{\partial \vec{r}} \right)$ becomes

$$\begin{pmatrix} \cos \delta \frac{\partial \alpha}{\partial \vec{r}} \\ \frac{\partial \delta}{\partial \vec{r}} \end{pmatrix} = \frac{1}{\rho} \begin{pmatrix} -\sin \alpha & \cos \alpha & 0 \\ -\sin \delta \cos \alpha & -\sin \delta \sin \alpha & \cos \delta \end{pmatrix}. \quad (10.50)$$

(4) Angle measurement variables (A, h)

In the epoch geocentric coordinate system, similarly, the angle A in the related residual y appears in the form of $\cos h(A - A_c)$, then the corresponding matrix is

$$\begin{pmatrix} \cos h \frac{\partial A}{\partial \vec{r}} \\ \frac{\partial h}{\partial \vec{r}} \end{pmatrix} = \frac{1}{\rho} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ h_{11} & h_{12} & h_{13} \end{pmatrix} (GR), \quad (10.51)$$

where the precession-nutation matrix (GR) is given previously. The elements a_{ij} and h_{ij} , are

$$\begin{cases} a_{11} = -\sin S \cos A + \cos S \sin A \sin \varphi, \\ a_{12} = \cos S \cos A + \sin S \sin A \sin \varphi, \\ a_{13} = -\cos \varphi \sin A, \end{cases} \quad (10.52)$$

$$\begin{cases} h_{11} = \cos S \cos \varphi \cos h + \sin S \sin h \sin A + \cos S \sin \varphi \sin h \cos A, \\ h_{12} = \sin S \cos \varphi \cos h - \cos S \sin h \sin A + \sin S \sin \varphi \sin h \cos A, \\ h_{13} = \sin \varphi \cos h - \cos \varphi \sin h \cos A. \end{cases} \quad (10.53)$$

10.3.2 Matrix $\left(\frac{\partial(\vec{r}, \dot{\vec{r}})}{\partial \sigma} \right)$

(1) Kepler orbital elements σ ($a, e, i, \Omega, \omega, M$)

$$\begin{cases} \left(\frac{\partial \vec{r}}{\partial \sigma} \right) = \left(\frac{\partial \vec{r}}{\partial a}, \frac{\partial \vec{r}}{\partial e}, \frac{\partial \vec{r}}{\partial i}, \frac{\partial \vec{r}}{\partial \Omega}, \frac{\partial \vec{r}}{\partial \omega}, \frac{\partial \vec{r}}{\partial M} \right), \\ \left(\frac{\partial \dot{\vec{r}}}{\partial \sigma} \right) = \left(\frac{\partial \dot{\vec{r}}}{\partial a}, \frac{\partial \dot{\vec{r}}}{\partial e}, \frac{\partial \dot{\vec{r}}}{\partial i}, \frac{\partial \dot{\vec{r}}}{\partial \Omega}, \frac{\partial \dot{\vec{r}}}{\partial \omega}, \frac{\partial \dot{\vec{r}}}{\partial M} \right), \end{cases} \quad (10.54)$$

and the 12 elements are

$$\begin{cases} \frac{\partial \vec{r}}{\partial a} = \frac{1}{a} \vec{r}, & \frac{\partial \vec{r}}{\partial \Omega} = \vec{\Omega}, \\ \frac{\partial \vec{r}}{\partial e} = H \vec{r} + K \dot{\vec{r}}, & \frac{\partial \vec{r}}{\partial \omega} = \vec{\omega}, \\ \frac{\partial \vec{r}}{\partial i} = \vec{I}, & \frac{\partial \vec{r}}{\partial M} = \frac{1}{n} \dot{\vec{r}}, \end{cases} \quad (10.55)$$

$$\begin{cases} \frac{\partial \dot{\vec{r}}}{\partial a} = -\frac{1}{2a} \dot{\vec{r}}, & \frac{\partial \dot{\vec{r}}}{\partial \Omega} = \vec{\Omega}', \\ \frac{\partial \dot{\vec{r}}}{\partial e} = H' \vec{r} + K' \dot{\vec{r}}, & \frac{\partial \dot{\vec{r}}}{\partial \omega} = \vec{\omega}', \\ \frac{\partial \dot{\vec{r}}}{\partial i} = \vec{I}', & \frac{\partial \dot{\vec{r}}}{\partial M} = -\frac{\mu}{n} \left(\frac{\vec{r}}{r^3} \right), \end{cases} \quad (10.56)$$

where

$$\begin{cases} H = -\frac{a}{p}(\cos E + e), & K = \frac{\sin E}{n}\left(1 + \frac{r}{p}\right), \\ p = a(1 - e^2), & n = \sqrt{\mu}a^{-\frac{3}{2}}, \quad E - e \sin E = M, \end{cases} \quad (10.57)$$

$$\vec{I} = \begin{pmatrix} z \sin \Omega \\ -z \cos \Omega \\ -x \sin \Omega + y \cos \Omega \end{pmatrix}, \quad (10.58)$$

$$\vec{\Omega} = \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}, \quad (10.59)$$

$$\vec{\omega} = \hat{R} \times \vec{r} = \begin{pmatrix} zR_y - yR_z \\ xR_z - zR_x \\ yR_x - xR_y \end{pmatrix}, \quad (10.60)$$

$$\begin{cases} H' = \frac{\sqrt{\mu a}}{rp} \sin E \left[1 - \frac{a}{r}\left(1 + \frac{p}{r}\right)\right], \\ K' = \frac{a}{p} \cos E, \end{cases} \quad (10.61)$$

$$\vec{I}' = \begin{pmatrix} \dot{z} \sin \Omega \\ -\dot{z} \cos \Omega \\ -\dot{x} \sin \Omega + \dot{y} \cos \Omega \end{pmatrix}, \quad (10.62)$$

$$\vec{\Omega}' = \begin{pmatrix} -\dot{y} \\ \dot{x} \\ 0 \end{pmatrix}, \quad (10.63)$$

$$\vec{\omega}' = \hat{R} \times \dot{\vec{r}} = \begin{pmatrix} \dot{z}R_y - \dot{y}R_z \\ \dot{x}R_z - \dot{z}R_x \\ \dot{y}R_x - \dot{x}R_y \end{pmatrix}. \quad (10.64)$$

In (10.60) and (10.64), $\hat{R}$ is the unit vector in the normal direction of the orbit plane, given by

$$\hat{R} = \begin{pmatrix} \sin i \sin \Omega \\ -\sin i \cos \Omega \\ \cos i \end{pmatrix}. \quad (10.65)$$

The calculation formulas of $\vec{r}$ and $\dot{\vec{r}}$ have provided previously.

- (2) The non-singularity orbital elements of the first type σ ($a, i, \Omega, \xi, \eta, \lambda$) where $\xi = e \cos \omega$, $\eta = e \sin \omega$, and $\lambda = M + \omega$

$$\begin{cases} \left(\frac{\partial \vec{r}}{\partial \sigma} \right) = \left(\frac{\partial \vec{r}}{\partial a}, \frac{\partial \vec{r}}{\partial i}, \frac{\partial \vec{r}}{\partial \Omega}, \frac{\partial \vec{r}}{\partial \xi}, \frac{\partial \vec{r}}{\partial \eta}, \frac{\partial \vec{r}}{\partial \lambda} \right), \\ \left(\frac{\partial \dot{\vec{r}}}{\partial \sigma} \right) = \left(\frac{\partial \dot{\vec{r}}}{\partial a}, \frac{\partial \dot{\vec{r}}}{\partial i}, \frac{\partial \dot{\vec{r}}}{\partial \Omega}, \frac{\partial \dot{\vec{r}}}{\partial \xi}, \frac{\partial \dot{\vec{r}}}{\partial \eta}, \frac{\partial \dot{\vec{r}}}{\partial \lambda} \right), \end{cases} \quad (10.66)$$

and the 12 elements are

$$\begin{cases} \frac{\partial \vec{r}}{\partial a} = \frac{1}{a} \vec{r}, & \frac{\partial \vec{r}}{\partial \xi} = A \vec{r} + B \dot{\vec{r}}, \\ \frac{\partial \vec{r}}{\partial i} = \vec{I}, & \frac{\partial \vec{r}}{\partial \eta} = C \vec{r} + D \dot{\vec{r}}, \\ \frac{\partial \vec{r}}{\partial \Omega} = \vec{\Omega}, & \frac{\partial \vec{r}}{\partial \lambda} = \frac{1}{n} \dot{\vec{r}}, \end{cases} \quad (10.67)$$

$$\begin{cases} \frac{\partial \dot{\vec{r}}}{\partial a} = -\frac{1}{2a} \dot{\vec{r}}, & \frac{\partial \dot{\vec{r}}}{\partial \xi} = A' \vec{r} + B' \dot{\vec{r}}, \\ \frac{\partial \dot{\vec{r}}}{\partial i} = \vec{I}', & \frac{\partial \dot{\vec{r}}}{\partial \eta} = C' \vec{r} + D' \dot{\vec{r}}, \\ \frac{\partial \dot{\vec{r}}}{\partial \Omega} = \vec{\Omega}', & \frac{\partial \dot{\vec{r}}}{\partial \lambda} = -\frac{\mu}{n} \left(\frac{\vec{r}}{r^3} \right), \end{cases} \quad (10.68)$$

where the variables $\vec{I}$, $\vec{\Omega}$ and $\vec{I}'$, $\vec{\Omega}'$ are given previously, and the eight variables $A, B, \dots, A', B', \dots$ are given as follows (in the formulas Earth's central gravitational constant $\mu = 1$, all quantities are given in dimensionless units):

$$\begin{cases} A = \frac{a}{p} \left[-(\cos u + \xi) - \left(\frac{r}{p} \right) (\sin u + \eta) (\xi \sin u - \eta \cos u) \right], \\ B = \frac{ar}{\sqrt{p}} \left[\sin u + \left(\frac{a}{r} \right) \frac{\sqrt{1-e^2}}{1+\sqrt{1-e^2}} \eta + \left(\frac{r}{p} \right) (\sin u + \eta) \right], \\ C = -\frac{a}{p} \left[(\sin u + \eta) - \left(\frac{r}{p} \right) (\cos u + \xi) (\xi \sin u - \eta \cos u) \right], \\ D = -\frac{ar}{\sqrt{p}} \left[\cos u + \left(\frac{a}{r} \right) \frac{\sqrt{1-e^2}}{1+\sqrt{1-e^2}} \xi + \left(\frac{r}{p} \right) (\cos u + \xi) \right], \end{cases} \quad (10.69)$$

$$\begin{cases} A' = \sqrt{\frac{1}{p}} \left\{ \frac{1}{p} (\sin u + \eta) - \frac{1}{r} \left[\left(\frac{a}{p} \right) (\sin u + \eta) + \left(\frac{a}{r} \right) \sin u + \left(\frac{a}{r} \right)^2 \frac{\sqrt{1-e^2}}{1+\sqrt{1-e^2}} \eta \right] \right\}, \\ B' = \left(\frac{r}{p} \right) \cos u + \left(\frac{a}{p} \right) \xi, \\ C' = -\sqrt{\frac{1}{p}} \left\{ \frac{1}{p} (\cos u + \xi) - \frac{1}{r} \left[\left(\frac{a}{p} \right) (\cos u + \xi) + \left(\frac{a}{r} \right) \cos u + \left(\frac{a}{r} \right)^2 \frac{\sqrt{1-e^2}}{1+\sqrt{1-e^2}} \xi \right] \right\}, \\ D' = \left(\frac{r}{p} \right) \sin u + \left(\frac{a}{p} \right) \eta, \end{cases} \quad (10.70)$$

where $e^2 = \xi^2 + \eta^2$, $u = f + \omega$, and the formulas of r , $\sin u$, and $\cos u$ are given by (10.30)–(10.32).

(3) The non-singularity orbital elements of the second type $\sigma(a, \xi, \eta, h, k, \lambda)$ which are

$$\begin{aligned} a, \quad \xi &= e \cos \tilde{\omega}, \quad \eta = e \sin \tilde{\omega}, \\ h &= \sin \frac{i}{2} \cos \Omega, \quad k = \sin \frac{i}{2} \sin \Omega, \quad \lambda = M + \tilde{\omega}, \\ \begin{cases} \left(\frac{\partial \vec{r}}{\partial \sigma} \right) &= \left(\frac{\partial \vec{r}}{\partial a}, \frac{\partial \vec{r}}{\partial \xi}, \frac{\partial \vec{r}}{\partial \eta}, \frac{\partial \vec{r}}{\partial h}, \frac{\partial \vec{r}}{\partial k}, \frac{\partial \vec{r}}{\partial \lambda} \right), \\ \left(\frac{\partial \dot{\vec{r}}}{\partial \sigma} \right) &= \left(\frac{\partial \dot{\vec{r}}}{\partial a}, \frac{\partial \dot{\vec{r}}}{\partial \xi}, \frac{\partial \dot{\vec{r}}}{\partial \eta}, \frac{\partial \dot{\vec{r}}}{\partial h}, \frac{\partial \dot{\vec{r}}}{\partial k}, \frac{\partial \dot{\vec{r}}}{\partial \lambda} \right), \end{cases} \end{aligned} \quad (10.71)$$

The 12 elements in (10.71) can be grouped into three types, which are

① The expressions of $\frac{\partial \vec{r}}{\partial a}, \frac{\partial \vec{r}}{\partial \xi}, \frac{\partial \vec{r}}{\partial \lambda}, \frac{\partial \dot{\vec{r}}}{\partial \lambda}$ are exactly the same as provided for the non-singularity orbital elements of the first type in (10.67) and (10.68).

② The expressions of $\frac{\partial \vec{r}}{\partial \xi}, \frac{\partial \vec{r}}{\partial \eta}, \frac{\partial \dot{\vec{r}}}{\partial \xi}, \frac{\partial \dot{\vec{r}}}{\partial \eta}$ have the same forms as those for the non-singularity orbital elements of the first type, except in A, B, C, D of (10.69) and A', B', C', D' of (10.70), where ξ, η , and u should be replaced by

$$\xi = e \cos \tilde{\omega}, \quad \eta = e \sin \tilde{\omega}, \quad u = f + \tilde{\omega}, \quad \tilde{\omega} = \omega + \Omega.$$

③ The expressions of $\frac{\partial \vec{r}}{\partial h}, \frac{\partial \vec{r}}{\partial k}, \frac{\partial \dot{\vec{r}}}{\partial h}, \frac{\partial \dot{\vec{r}}}{\partial k}$, which do not appear before, are (similarly, there is $\mu = 1$):

$$\begin{cases} \frac{\partial \vec{r}}{\partial h} = \vec{H} = r \cos u \hat{P}_h + r \sin u \hat{Q}_h, \\ \frac{\partial \vec{r}}{\partial k} = \vec{K} = r \cos u \hat{P}_k + r \sin u \hat{Q}_k, \end{cases} \quad (10.72)$$

$$\begin{cases} \frac{\partial \dot{\vec{r}}}{\partial h} = \vec{H}' = \frac{1}{\sqrt{p}} \left[-(\sin u + \eta) \hat{P}_h + (\cos u + \xi) \hat{Q}_h \right], \\ \frac{\partial \dot{\vec{r}}}{\partial k} = \vec{K}' = \frac{1}{\sqrt{p}} \left[-(\sin u + \eta) \hat{P}_k + (\cos u + \xi) \hat{Q}_k \right], \end{cases} \quad (10.73)$$

where

$$\hat{P}_h = \begin{pmatrix} 0 \\ 2k \\ 2hk / \cos \frac{i}{2} \end{pmatrix}, \quad \hat{Q}_h = \begin{pmatrix} 2k \\ -4h \\ 2(1 - 2h^2 - k^2) / \cos \frac{i}{2} \end{pmatrix}, \quad (10.74)$$

$$\hat{P}_k = \begin{pmatrix} -4k \\ 2h \\ -2(1 - h^2 - 2k^2) / \cos \frac{i}{2} \end{pmatrix}, \quad \hat{Q}_k = \begin{pmatrix} 2h \\ 0 \\ -2hk / \cos \frac{i}{2} \end{pmatrix}, \quad (10.75)$$

$$\cos \frac{i}{2} = [(1 - (h^2 + k^2))]^{\frac{1}{2}}. \quad (10.76)$$

10.3.3 State Transition Matrix Φ

The process of precise orbit determination has the function to derive estimations for some parameters. We now use an atmospheric parameter as an example to show this function. The atmospheric parameter appears in the precise orbit determination for a low Earth orbit satellite and is denoted by β . We assume

$$\begin{cases} \dot{\beta} = 0, \\ \beta = \beta_0 = \left(\frac{C_D S}{m} \right) \bar{\rho}_{p_0}. \end{cases} \quad (10.77)$$

The parameter β is actually a comprehensive parameter, which involves the coefficient of the atmospheric drag, the equivalent area-to-mass ratio of the satellite's endurance to the drag, and the mean atmospheric density at a chosen location. The three factors are generally impossible to be decided and to be separated in the process of orbit determination, therefore, can be represented by one comprehensive parameter as defined in (10.77). The parameter to-be-estimated β is assumed to be a constant, thus the state variable X is actually a vector of seven dimensions. Although there is only one parameter to be estimated in the example, by providing an estimate of this parameter we can show the capability of the precise orbit determination, therefore, improving the accuracy.

(1) Numerical method to calculate the state transition matrix $\Phi = \left(\frac{\partial X}{\partial X_0} \right)$

The actual form of the numerical solution depends on a selected numerical method, and different methods have different formats. If the right-side function of the state differential Eq. (10.2) were not complicated and the state transition matrix could be given by

$$\Phi(t_n, t_{n+1}) = \left(\frac{\partial X_{n+1}}{\partial X_n} \right), \quad (10.78)$$

the form would still vary with the method, and would not be convenient to use. The fact is that the right-side function of (10.2) is complicated, so it is impractical to use numerical method to directly calculate the right-side function Φ .

As we know that the condition Eq. (10.8) is a result of linearization, by the same principle, the matrix Φ can also be linearized. Let

$$X(t) = \Phi(t, t_0)X_0. \quad (10.79)$$

Correspondingly there are

$$\begin{cases} X^* = \Phi(t, t_0)X_0^*, \\ x = X - X^* = \Phi(t, t_0)x_0. \end{cases} \quad (10.80)$$

Linearizing the state differential equation

$$\begin{cases} \dot{X} = F(X, t), \\ X(t_0) = X_0, \end{cases} \quad (10.81)$$

we obtain

$$\begin{cases} \dot{\Phi} = A\Phi, \\ A = \left(\frac{\partial F}{\partial X} \right)_{X^*}, \quad \Phi(t_0, t_0) = I, \end{cases} \quad (10.82)$$

where $A = A(t)$. Now the problem of calculating $\Phi(t, t_0)$ becomes to solve this matrix differential equation. This is an $(n \times n)$ dimensional linear ordinary differential equation with varying coefficients, and should be integrated with the state differential Eq. (10.2) at the same time. Obviously, it increases the workload of calculation.

If the orbit variables are $(\vec{r}, \dot{\vec{r}})$, then the corresponding state differential Eq. (10.81) has the form as

$$\begin{cases} \frac{d\vec{r}}{dt} = \dot{\vec{r}}, & \frac{d\dot{\vec{r}}}{dt} = \ddot{\vec{r}} = F(\vec{r}, \dot{\vec{r}}, \beta_0), \\ \dot{\beta} = 0, \end{cases} \quad (10.83)$$

and the state transition matrix Φ has the form as

$$\Phi = \left(\frac{\partial(\vec{r}, \dot{\vec{r}}, \beta)}{\partial(\vec{r}_0, \dot{\vec{r}}_0, \beta_0)} \right)_{(n \times n)} = \begin{pmatrix} \left(\frac{\partial \vec{r}}{\partial X_0} \right)_{3 \times n} \\ \left(\frac{\partial \dot{\vec{r}}}{\partial X_0} \right)_{3 \times n} \\ \left(\frac{\partial \beta}{\partial X_0} \right)_{1 \times n} \end{pmatrix}, \quad n = 7, \quad (10.84)$$

The coefficient matrix A in Eq. (10.82) has the form:

$$A = \begin{pmatrix} \left(\frac{\partial \dot{\vec{r}}}{\partial \vec{r}} \right) & \left(\frac{\partial \dot{\vec{r}}}{\partial \dot{\vec{r}}} \right) & \left(\frac{\partial \dot{\vec{r}}}{\partial \beta} \right) \\ \left(\frac{\partial \ddot{\vec{r}}}{\partial \vec{r}} \right) & \left(\frac{\partial \ddot{\vec{r}}}{\partial \dot{\vec{r}}} \right) & \left(\frac{\partial \ddot{\vec{r}}}{\partial \beta} \right) \\ \left(\frac{\partial 0}{\partial \vec{r}} \right) & \left(\frac{\partial 0}{\partial \dot{\vec{r}}} \right) & \left(\frac{\partial 0}{\partial \beta} \right) \end{pmatrix}_{X^*} = \begin{pmatrix} (0) & (1) & (0) \\ \left(\frac{\partial \ddot{\vec{r}}}{\partial \vec{r}} \right) & \left(\frac{\partial \ddot{\vec{r}}}{\partial \dot{\vec{r}}} \right) & \left(\frac{\partial \ddot{\vec{r}}}{\partial \beta} \right) \\ (0) & (0) & (0) \end{pmatrix}_{X^*}. \quad (10.85)$$

If the orbit variables are orbital elements, then the corresponding state differential equation has the following form

$$\begin{cases} \dot{\sigma} = f(\sigma, t; \beta_0), \\ \dot{\beta} = 0. \end{cases} \quad (10.86)$$

The matrix Φ becomes

$$\Phi = \left(\frac{\partial(\sigma, \beta)}{\partial(\sigma_0, \beta_0)} \right)_{(n \times n)} = \begin{pmatrix} \left(\frac{\partial \sigma}{\partial X_0} \right)_{6 \times n} \\ \left(\frac{\partial \beta}{\partial X_0} \right)_{1 \times n} \end{pmatrix}, \quad n = 7, \quad (10.87)$$

and the coefficient matrix A is

$$A = \begin{pmatrix} \left(\frac{\partial f}{\partial \sigma} \right) & \left(\frac{\partial f}{\partial \beta} \right) \\ \left(\frac{\partial 0}{\partial \sigma} \right) & \left(\frac{\partial 0}{\partial \beta} \right) \end{pmatrix}_{X^*} = \begin{pmatrix} \left(\frac{\partial f}{\partial \sigma} \right) & \left(\frac{\partial f}{\partial \beta} \right) \\ (0) & (0) \end{pmatrix}_{X^*}. \quad (10.88)$$

(2) Analytical expression of the state transition matrix $\Phi = \left(\frac{\partial X}{\partial X_0} \right)$

For orbit determination, if there is an analytical solution for the orbit, then the state transition matrix Φ can be provided in analytical form. As mentioned above the process of orbital determination is connected to the linearization, and the whole process is achieved by iteration. Before the convergent of the iteration the theoretical correction, x_0 , of the to-be-estimated state variable X_0 is small, therefore, in the condition equation the treatment of matrix $\tilde{B}$ is less strict than that of $H(X^*, t)$, and can be simplified. Usually, the arc used for the orbit determination is not very long, thus, to give the analytical expression of Φ , we only include the main variation of the orbital variables (such as the first-order secular variations). Actually, for both analytical and numerical methods of determining an orbit, the state transition matrix Φ can be provided by the perturbation solutions of the orbital variables.

If the orbital variables are $(\vec{r}, \dot{\vec{r}})$, the to-be-estimated state variable X_0 includes $\vec{r}_0, \dot{\vec{r}}_0$ and β_0 , then Φ has the following form

$$\Phi = \left(\frac{\partial(\vec{r}, \dot{\vec{r}}, \beta)}{\partial(\vec{r}_0, \dot{\vec{r}}_0, \beta_0)} \right) = \Phi^{(0)} + \Phi^{(1)}, \quad (10.89)$$

where $\Phi^{(0)}$ and $\Phi^{(1)}$ are the non-perturbed and perturbed parts, respectively (only need the main perturbation term, i.e., the first-order secular term), the expression is not given in this book but can be found in reference [4].

If the orbit variables are orbital elements σ , and the to-be-estimated state variables are σ_0 and β_0 , then σ_0 are the mean orbital elements or quasi-mean elements in the analytical orbit determination, while in the numerical orbit determination, σ_0 are the instantaneous orbital elements. The matrix Φ has the following form:

$$\Phi = \left(\frac{\partial(\sigma, \beta)}{\partial(\sigma_0, \beta_0)} \right) = \Phi^{(0)} + \Phi^{(1)}, \quad (10.90)$$

The formulas of $\Phi^{(0)}$ and $\Phi^{(1)}$ for different sets of orbital elements are provided as follows.

① σ are for the Kepler orbital elements $(a, e, i, \Omega, \omega, M)$

$$\Phi^{(0)} = (a_{ij}) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ a_{61} & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad a_{61} = -\frac{3n}{2a}(t - t_0), \quad (10.91)$$

$$\Phi^{(1)} = (b_{ij}) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & b_{17} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ b_{41} & b_{42} & b_{43} & 0 & 0 & 0 & 0 \\ b_{51} & b_{52} & b_{53} & 0 & 0 & 0 & 0 \\ b_{61} & b_{62} & b_{63} & 0 & 0 & 0 & b_{67} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (10.92)$$

where

$$\begin{cases} b_{17} = \frac{\partial a}{\partial \beta_0} = -\left[a^2 I_0(z) \exp\left(-\frac{a-r_{p0}}{H_{p0}}\right) \right] n(t - t_0), \\ b_{67} = \frac{\partial M}{\partial \beta_0} = \left[\frac{3}{2} a I_0(z) \exp\left(-\frac{a-r_{p0}}{H_{p0}}\right) \right]^{\frac{1}{2}} [n(t - t_0)]^2. \end{cases} \quad (10.93)$$

$$b_{41} = -\frac{7}{2a}\Omega_1(t-t_0), \quad b_{42} = \frac{4e}{1-e^2}\Omega_1(t-t_0), \quad b_{43} = -\tan i\Omega_1(t-t_0) \quad (10.94)$$

$$b_{51} = -\frac{7}{2a}\omega_1(t-t_0), \quad b_{52} = \frac{4e}{1-e^2}\omega_1(t-t_0), \quad b_{53} = -\frac{5\sin 2i}{4-5\sin^2 i}\omega_1(t-t_0), \quad (10.95)$$

$$b_{61} = -\frac{7}{2a}M_1(t-t_0), \quad b_{62} = \frac{3e}{1-e^2}M_1(t-t_0), \quad b_{63} = -\frac{3\sin 2i}{2-3\sin^2 i}M_1(t-t_0), \quad (10.96)$$

$$\begin{cases} \Omega_1 = -\frac{3J_2}{2p^2}n \cos i, & \omega_1 = \frac{3J_2}{2p^2}n(2 - \frac{5}{2}\sin^2 i), \\ M_1 = \frac{3J_2}{2p^2}n(1 - \frac{3}{2}\sin^2 i)\sqrt{1-e^2}. \end{cases} \quad (10.97)$$

In the analytical orbit determination, the variables $a, e, i, p = a(1-e^2)$, and $n = a^{-3/2}$ are given by the to-be-estimated mean orbit elements $\bar{a}_0, \bar{e}_0, \bar{i}_0, \dots$ at epoch t_0 , whereas in the numerical orbit determination they are the to-be-estimated instantaneous orbit elements $a_0, e_0, i_0, \dots$. The variables in b_{17} and b_{67} are related to the atmospheric drag, given in Sect. 4.11.

② σ are for the non-singularity orbital elements of the first type ($a, i, \Omega, \xi = e \cos \omega, \eta = e \sin \omega, \lambda = M + \omega$).

$\Phi^{(0)}$ is the same as in (10.91) and $\Phi^{(1)}$ is given by

$$\Phi^{(1)} = (b_{ij}) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & b_{17} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ b_{31} & b_{32} & 0 & b_{34} & b_{35} & 0 & 0 \\ b_{41} & b_{42} & 0 & b_{44} & b_{45} & 0 & 0 \\ b_{51} & b_{52} & 0 & b_{54} & b_{55} & 0 & 0 \\ b_{61} & b_{62} & 0 & b_{64} & b_{65} & 0 & b_{67} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (10.98)$$

where b_{17} and b_{67} are given by (10.93), the other elements are given by

$$\begin{cases} b_{31} = -\frac{7}{2a}\Omega_1(t-t_0), & b_{32} = -\tan i\Omega_1(t-t_0), \\ b_{34} = \frac{4\xi_0}{1-e^2}\Omega_1(t-t_0), & b_{35} = \frac{4\eta_0}{1-e^2}\Omega_1(t-t_0), \end{cases} \quad (10.99)$$

$$\begin{cases} b_{41} = -\eta_0\left[-\frac{7}{2a}\omega_1(t-t_0)\right], \\ b_{42} = -\eta_0\left[-\frac{5\sin 2i}{4-5\sin^2 i}\omega_1(t-t_0)\right], \\ b_{44} = -\eta_0\left[\frac{4\xi_0}{1-e^2}\omega_1(t-t_0)\right], \\ b_{45} = -\left[1 + \frac{4}{1-e^2}\eta_0^2\right]\omega_1(t-t_0), \end{cases} \quad (10.100)$$

$$\begin{cases} b_{51} = \xi_0 \left[-\frac{7}{2a} \omega_1 (t - t_0) \right], \\ b_{52} = \xi_0 \left[-\frac{5 \sin 2i}{4 - 5 \sin^2 i} \omega_1 (t - t_0) \right], \\ b_{54} = \left[1 + \frac{4}{1 - e^2} \xi_0^2 \right] \omega_1 (t - t_0), \\ b_{55} = \xi_0 \left[\frac{4\eta_0}{1 - e^2} \omega_1 (t - t_0) \right], \end{cases} \quad (10.101)$$

$$\begin{cases} b_{61} = -\frac{7}{2a} \lambda_1 (t - t_0), \\ b_{62} = -\left(\frac{3}{2 - 3 \sin^2 i} M_1 + \frac{5}{4 - 5 \sin^2 i} \omega_1 \right) \sin 2i (t - t_0), \\ b_{64} = \frac{\xi_0}{1 - e^2} (3M_1 + 4\omega_1) (t - t_0), \\ b_{65} = \frac{\eta_0}{1 - e^2} (3M_1 + 4\omega_1) (t - t_0). \end{cases} \quad (10.102)$$

The variables a , e , i , and other orbital elements are the to-be-estimated mean orbital elements $\bar{a}_0$, $\bar{e}_0^2 = \bar{\xi}_0^2 + \bar{\eta}_0^2$, $\bar{i}_0$, $\dots$, or the to-be-estimated instantaneous orbit elements a_0 , $e_0^2 = \xi_0^2 + \eta_0^2$, i_0 , $\dots$ at t_0 . The other variables are the same as before, also there is $\lambda_1 = M_1 + \omega_1$.

③ σ are for the non-singularity orbital elements of the second type

$$a, \quad \xi = e \cos \tilde{\omega}, \quad \eta = e \sin \tilde{\omega},$$

$$h = \sin \frac{i}{2} \cos \Omega, \quad k = \sin \frac{i}{2} \sin \Omega, \quad \lambda = M + \tilde{\omega}.$$

$\Phi^{(0)}$ is the same as in (10.91) and $\Phi^{(1)}$ is given by

$$\Phi^{(1)} = (b_{ij}) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & b_{17} \\ b_{21} & b_{22} & b_{23} & b_{24} & b_{25} & 0 & 0 \\ b_{31} & b_{32} & b_{33} & b_{34} & b_{35} & 0 & 0 \\ b_{41} & b_{42} & b_{43} & b_{44} & b_{45} & 0 & 0 \\ b_{51} & b_{52} & b_{53} & b_{54} & b_{55} & 0 & 0 \\ b_{61} & b_{62} & b_{63} & b_{64} & b_{65} & 0 & b_{67} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (10.103)$$

where b_{17} and b_{67} are given by (10.93), the other elements are

$$\begin{cases} b_{21} = -\eta_0 \left[-\frac{7}{2a} \tilde{\omega}_1 (t - t_0) \right], \\ b_{22} = -\eta_0 \left[-\frac{4\xi_0}{1 - e^2} \tilde{\omega}_1 (t - t_0) \right], \\ b_{23} = -\left[1 + \frac{4}{1 - e^2} \eta_0^2 \right] \tilde{\omega}_1 (t - t_0), \\ b_{24} = -\eta_0 \left[\left(-\frac{5}{2} \right) 8h_0 (1 - 2(h_0^2 + k_0^2)) + 4h_0 \right] \left(\frac{3J_2}{2p^2} \right) n(t - t_0), \\ b_{25} = -\eta_0 \left[\left(-\frac{5}{2} \right) 8k_0 (1 - 2(h_0^2 + k_0^2)) + 4k_0 \right] \left(\frac{3J_2}{2p^2} \right) n(t - t_0), \end{cases} \quad (10.104)$$

$$\begin{cases} b_{31} = \xi_0 \left[-\frac{7}{2a} \tilde{\omega}_1(t - t_0) \right], \\ b_{32} = \left[1 + \frac{4}{1-e^2} \xi_0^2 \right] \tilde{\omega}_1(t - t_0), \\ b_{33} = \xi_0 \left[-\frac{4\eta_0}{1-e^2} \tilde{\omega}_1(t - t_0) \right], \\ b_{34} = \xi_0 \left[\left(-\frac{5}{2} \right) 8h_0(1 - 2(h_0^2 + k_0^2)) + 4h_0 \right] \left(\frac{3J_2}{2p^2} \right) n(t - t_0), \\ b_{35} = \xi_0 \left[\left(-\frac{5}{2} \right) 8k_0(1 - 2(h_0^2 + k_0^2)) + 4k_0 \right] \left(\frac{3J_2}{2p^2} \right) n(t - t_0), \end{cases} \quad (10.105)$$

$$\begin{cases} b_{41} = -k_0 \left[-\frac{7}{2a} \Omega_1(t - t_0) \right], \\ b_{42} = -k_0 \left[\frac{4\xi_0}{1-e^2} \Omega_1(t - t_0) \right], \\ b_{43} = -k_0 \left[\frac{4\eta_0}{1-e^2} \Omega_1(t - t_0) \right], \\ b_{44} = -k_0 \left[4h_0 \left(\frac{3J_2}{2p^2} \right) n(t - t_0) \right], \\ b_{45} = - \left[\Omega_1 + 4k_0^2 \left(\frac{3J_2}{2p^2} \right) n \right] (t - t_0), \end{cases} \quad (10.106)$$

$$\begin{cases} b_{51} = h_0 \left[-\frac{7}{2a} \Omega_1(t - t_0) \right], \\ b_{52} = h_0 \left[\frac{4\xi_0}{1-e^2} \Omega_1(t - t_0) \right], \\ b_{53} = h_0 \left[\frac{4\eta_0}{1-e^2} \Omega_1(t - t_0) \right], \\ b_{54} = \left[\Omega_1 + 4h_0^2 \left(\frac{3J_2}{2p^2} \right) n \right] (t - t_0), \\ b_{55} = h_0 \left[4k_0 \left(\frac{3J_2}{2p^2} \right) n(t - t_0) \right], \end{cases} \quad (10.107)$$

$$\begin{cases} b_{61} = -\frac{7}{2a} \lambda_1(t - t_0), \\ b_{62} = \frac{\xi_0}{1-e^2} (3M_1 + 4\tilde{\omega}_1)(t - t_0), \\ b_{63} = \frac{\eta_0}{1-e^2} (3M_1 + 4\tilde{\omega}_1)(t - t_0), \\ b_{64} = \left[\left(-\frac{3}{2} \sqrt{1-e^2} - \frac{5}{2} \right) 8h_0(1 - 2(h_0^2 + k_0^2)) + 4h_0 \right] \left(\frac{3J_2}{2p^2} \right) n(t - t_0), \\ b_{65} = \left[\left(-\frac{3}{2} \sqrt{1-e^2} - \frac{5}{2} \right) 8k_0(1 - 2(h_0^2 + k_0^2)) + 4k_0 \right] \left(\frac{3J_2}{2p^2} \right) n(t - t_0). \end{cases} \quad (10.108)$$

All variables have the same meanings as before, except that $\lambda_1 = M_1 + \tilde{\omega}_1$, $\tilde{\omega}_1 = \omega_1 + \Omega_1$; and e_0 in M_1 , ω_1 , and Ω_1 , is given by $e_0^2 = \xi_0^2 + \eta_0^2$; i_0 is given by $\sin^2(i_0/2) = h_0^2 + k_0^2$; and σ_0 are the mean orbit elements $\bar{\sigma}_0$.

Since $\beta = \beta_0 = \left(\frac{C_{DS}}{m} \right) \bar{\rho}_{p_0}$, for all three types of orbital elements, we have

$$\left(\frac{\partial \beta}{\partial \sigma_0} \right) = (0), \quad \left(\frac{\partial \beta}{\partial \beta_0} \right) = (1) = I, \quad (10.109)$$

where (0) and $(1) = I$ represent a zero-element matrix and a unit matrix, respectively.

In practice, if there are other forms for the atmosphere factors or other perturbation factors, the above transformation matrices need to be changed accordingly, we do not discuss this topic further in this book.

10.4 Estimation of the State Variable: Calculation of Precise Orbit Determination

The basic equation of orbit determination (10.8), i.e., the condition equation, can be formed by the residual matrix y and the matrix $\tilde{B}$ provided in the previous sections, that

$$y = \tilde{B}x_0 + V.$$

In the equation x_0 is the correction of the to-be-estimated state variable X_0 which is n -dimensional; the correction matrix y of the measurements (i.e., the residuals) $t_j, y_j (j = 1, 2, \dots, k)$ is $(m \times k)$ -dimensional, where $m \geq 1$ and m is the dimension of the measurement (the angles-only data are 2-dimensional). When $(m \times k) \geq n$, by principle, the system of Eq. (10.8) can be solved, which means that from a sequence of measurements $t_j, Y_j (j = 1, 2, \dots, k)$, the state variable X_0 at epoch t_0 can be obtained. The usual calculation method is the least squares estimator. This method is discussed in Sect. 10.5, and the details about this method are given in related references [5–7]. In this section, we discuss the certainty of the solution and how to obtain a solution.

10.4.1 Certainty of Solution in the Orbit Determination

The problem about the certainty of solution (i.e., the condition of orbit determination) is about whether we can resolve the only state variable X_0 at the corresponding epoch t_0 from a sequence of measurements $t_j, Y_j (j = 1, 2, \dots, k)$. This problem is also called the problem of observability. Not all kinds of measurements have observability. For example, if measurements are distances and speeds on a short arc from a single tracking station, then it is difficult to determine an orbit.

The observability decides whether the iteration process of solving Eq. (10.8) converges or not. In Sect. 9.1, we discussed this problem. The conclusion is that the observability depends on the characteristics of the matrix $\tilde{B}$, whether it is a positive-defined matrix or whether the corresponding determinant $|\tilde{B}^T \tilde{B}| = 0$. This condition decides the success of the orbit determination. In fact, it involves many aspects, such as the type of measurements, the distribution of an arc used for the project, and the selection of the state variables (e.g., to use Kepler orbital elements a, e, i, Ω, ω , and M would fail to define an orbit with a small eccentricity), etc. These problems have attracted the attention of seriousness in the orbit determination projects for all sorts of spacecraft.

10.4.2 Process of Calculating Solution in the Orbit Determination

(1) The iterative process of orbit determination calculation

A large amount of measurement allows the orbit determination to make use of the statistical advantage of measurement. We now introduce the process of orbit determination by the batch processing method as an example.

For a sequence of given measurements $Y_1, Y_2, \dots, Y_k$, and the initial value of the to-be-estimated state variable X_0 at $t = t_0$, $\hat{X}_{0/0}$, generally $t_1 \leq t_0 \leq t_k$, then the task of the orbit determination is to use the least squares estimator to derive the optimal estimate of X_0 , denoted to $\hat{X}_{0/k}$.

The batch process in fact is an iterative process, achieved by linearizing a non-linear measurement equation. When there is no prior estimate, the so-called reference state variable X_0^* is actually the initial value for iteration, denoted to $\hat{X}_{0/k}^{(j-1)}$, $j = 1, 2, \dots$ where j is the number of iterations. When $j = 1$, there is

$$\hat{X}_{0/k}^{(j-1)} = \hat{X}_{0/0} = X_0^*.$$

In the process every time the Eq. (10.8) is solved, it provides the optimal estimate of the to-be-estimated state variable as

$$\hat{X}_{0/k}^{(j)} = \hat{X}_{0/k}^{(j-1)} + \hat{x}_{0/k}^{(j)}, \quad j = 1, 2, \dots \quad (10.110)$$

Let U as the sum of squared residuals and compare the two adjacent values of U before and after an iteration, if the difference reaches the following criterion

$$|U^{(j)} - U^{(j-1)}| \leq \mu, \quad (10.111)$$

then the iteration is completed, and $\hat{X}_{0/k}^{(j)}$ is the estimate of the state variable by the least squares estimator, otherwise, the iteration continues. The value of μ depends on the requirement of the actual project and the available measurements, we just assume $\mu > 0$. The sum of the squared residuals (or the weighted squared residuals) U , is usually defined by

$$\sum_{l=1}^k U = (y_l^T W_l y_l), \quad (10.112)$$

where W_l is the weight matrix of measurement; y_l^T is the transpose of y_l . In actual orbit determination, it is difficult to know the optimal weight. Usually, the weight is given according to the residual, or all data are given by equal weight.

(2) Selection of the criterion of iteration convergence

The optimal criterion of the least squares estimator is that the sum of squared residuals has its minimum. By principle, it is reasonable that the iteration ends according to (10.111). In reality, the sum of the squared residuals U is often replaced by the root mean square deviation (RMS), denoted by σ^* , which is defined by

$$\sigma^* = \sqrt{\frac{U}{k \times m}}, \quad (10.113)$$

where m is the dimension of measurement variable Y , k is the number of samplings, and thus $k \times m$ is the total amount of measurements. By the definition σ^* has a clear sense of accuracy.

The controlling condition of the iterating convergence is usually chosen from the following two conditions:

$$\begin{cases} |\sigma^{*(j)} - \sigma^{*(j-1)}| < \mu^*, \\ \left| \frac{\sigma^{*(j)} - \sigma^{*(j-1)}}{\sigma^{*(j)}} \right| < \mu^*, \end{cases} \quad (10.114)$$

where μ^* is usually defined according to the accuracy of the measurements and the demand of the orbit determination.

(3) Wild data elimination in orbit determination

Theoretically, there is a “conclusion” that the chance of a residual $y > 3\sigma^*$ is very rare, if it happens, it is assumed to be a “mistake”, and the corresponding measurements should be discarded as a wild value, i.e., not to be used again. In reality, it is more complicated and how to define a wild measurement depends on an actual situation.

10.5 The Least Squares Estimator and Its Application in Precise Orbit Determination

The condition Eq. (10.8) is a system of linear algebra equations. If there are k samplings, then the system of equations is $(m \times k)$ dimensional. Usually, there is a large number of samplings, so $(m \times k) \gg n$, n is the dimension of the to-be-estimated state variable. Therefore, how to sufficiently use the statistic properties to obtain the optimal estimate of the to-be-estimated state variable X_0 is an important procedure in the precise orbit determination. The basic principles of the optimal estimation theory are given in references [5–7]. In this book, we provide a simple introduction.

10.5.1 Estimation Theory and a Few Commonly Used Optimal Estimation Methods

The estimation principle in orbit determination is to estimate the state or parameters of the system based on the state equation and the measurement equation according to the theories of probability statistics and optimization. The state equation providing the variation of the state variable $X(t)$ usually includes the systematic errors, whereas the measurement equation describing the relationship of the state variable $X(t)$ and the measurement variable $Y(t)$ is unavoidably affected by the measurement noise. This system can be described in mathematical forms that:

$$\begin{cases} \dot{X} = F(X, t) + Q, \\ Y = H(X, t) + V, \end{cases} \quad (10.115)$$

where Q and V represent the systematic error in a dynamic model and the noise in measurement, respectively. Generally, both $F(X, t)$ and $H(X, t)$ are complicated non-linear functions. If the dynamic system is the motion of an artificial satellite (or other spacecraft) under the influence of all related forces, then the process of finding a solution is the precise orbit determination. The solution gives the state estimate of the state variables of the satellite, such as $(\vec{r}, \dot{\vec{r}}; \beta)$ or (σ, β) , where β can be a physical parameter in the dynamic model or a parameter related to the satellite itself, etc.

An optimal estimate is obtained according to a certain estimation criterion, if the criterion is changed, then the optimal estimate may not be optimal. In practice based on experience and different requirements, there is a variety of estimation criteria, such as the minimum variance criterion, the maximum likelihood criterion, the maximum posteriori criterion, the linear minimum variance criterion, the least squares criterion, etc. Corresponding to each criterion there is an estimation method (i.e., estimator), such as the minimum variance estimator, the maximum likelihood estimator, the maximum posteriori estimator, the linear minimum variance estimator, the least squares estimator, etc.

The satellite measurement variable Y in fact consists of random data of dynamical measurement. The precise orbit determination of a moving body can be summarized as follows: from a sequence of random measurements (a vector), $Y_1, Y_2, \dots, Y_k$, to estimate a non-random or random vector X (satellite orbital variables and related parameter). Usually, the estimated X is denoted by $\hat{X}$, obviously $\hat{X}$ is a function of the random sequence of $Y_1, Y_2, \dots, Y_k$. Because Y_j ($1 \leq j \leq k$) is random, $\hat{X}$ is also random, generally expressed as

$$\hat{X}_k = g(Y_1, Y_2, \dots, Y_k). \quad (10.116)$$

The subscript k in $\hat{X}_k$ indicates that the estimate is given by k measurements. A different function g gives a different estimate, the question is how to judge the

quality of an estimate. We introduce and discuss a few popular estimation methods as follows.

(1) Consistent estimator

Definition 1 If a sequence of $\{\hat{X}_k\}$ randomly converges to X , then $\hat{X}_k$ is the consistent estimator of X that for any given $\varepsilon > 0$ there is

$$\lim_{k \rightarrow \infty} P\left(\left|\hat{X}_k - X\right| > \varepsilon\right) = 0, \quad (10.117)$$

where P is a probability function. By this definition, if $\hat{X}_k$ is the consistent estimator of X , then for the amount of samplings Y_k , when k is sufficiently large, the estimator $\hat{X}_k$ must be in the vicinity of the true value of X .

(2) Unbiased estimator

If the random sequence of $Y_1, Y_2, \dots, Y_k$ has N repeated independent samplings, the n th sampling is denoted by $Y_1^{(n)}, Y_2^{(n)}, \dots, Y_k^{(n)}$, and an estimate of X derived from the n th sampling is denoted by $\hat{X}_k^{(n)}$, then we expect

$$\frac{1}{N} \sum_{n=1}^N \hat{X}_k^{(n)}$$

to become the mean of X where the number N is sufficiently large. An estimator possessing this property is called an unbiased estimator.

Definition 2 If there is a relationship that

$$E(\hat{X}_k) = E(X), \quad (10.118)$$

then $\hat{X}_k$ is the unbiased estimator of X , where $E(\hat{X}_k)$ is the mathematic expectation of the random vector $\hat{X}_k$, i.e., its mean value. If the relationship is

$$\lim_{k \rightarrow \infty} E(\hat{X}_k) = E(X), \quad (10.119)$$

then $\hat{X}_k$ is the asymptotic unbiased estimator of X .

The characteristic of an unbiased estimator assures that $\hat{X}$ oscillates around the true value X , but it does not provide the amplitude of the oscillation. The following content introduces the minimum-variance estimator which can ensure the mean amplitude of this oscillation to be the minimum.

(3) The minimum-variance estimator

Definition 3 If the estimation variance of $(\hat{X} - X)$ reaches the minimum, which means

$$\text{Var}(\hat{X} - X) = E\left((\hat{X} - X)(\hat{X} - X)^T\right) = \min, \quad (10.120)$$

then $\hat{X}$ is the minimum-variance estimator of X . Here “T” means transposition, and the min is for all possible estimators of $\hat{X}$.

Usually, the minimum-variance unbiased estimator is regarded as the optimal estimator in the sense of mean squares, which assures that the estimator $\hat{X}$ oscillates in the vicinity of the true value X and the mean amplitude of the oscillation reaches the minimum. Here the optimal estimator has the minimum variance among all linear and non-linear unbiased estimators.

(4) Linear minimum-variance unbiased estimator

As mentioned above that the minimum-variance unbiased estimator is not necessarily linear, thus the linear minimum-variance unbiased estimator is not the optimal estimator in the sense of mean squares.

The linear estimation means that the estimator $\hat{X}$ of X is a linear function of measurement Y , and generally has the form

$$\hat{X} = C + AY. \quad (10.121)$$

We assume that X is an n -dimensional random vector, Y is an m -dimensional random vector, C is an n -dimensional non-random vector, and A is an $(n \times m)$ -dimensional matrix. The reason to introduce the linear estimation is for practical purposes, for example, it is applied in the precise orbit determination.

10.5.2 The Least Squares Estimator

The least squares estimator is a classical estimation method. More than 200 years ago this method was used to determine the orbits of a celestial body (i.e., orbit improvement by Gauss, 1809). Following the development of the optimal estimation theory, the ancient method gains new content. The classical least squares estimation method obtains the estimator when the sum of squared residuals reaches the minimum. From the optimal estimation point of view, it is a linear unbiased estimator, but generally, it is not a linear minimum-variance unbiased estimator, only if the optimal weight is applied then it is a linear minimum-variance unbiased estimator. The least squares estimator with optimal weight treatment is called the Markov estimator.

(1) The classical least squares estimator

In a linear system:

$$\begin{cases} X_l = \Phi_{l,0}X_0, \\ Y_l = H_l X_l + V_l, \end{cases} \quad l = 1, 2, \dots, k \quad (10.122)$$

where the sequence of measurements $Y_1, Y_2, \dots, Y_k$ corresponds to a time sequence of $t_1, t_2, \dots, t_k$; X_0 corresponds to $t_0, t_0 \in [t_0, t_k]$. The dynamic equation of the state variable X and the measurement equation of measurement Y are both linear. The state variable X is a non-random vector (n -dimensional), and $\Phi_{l,0}$ is an $(n \times n)$ dimensional state transition matrix. The measurement variable Y is an m -dimensional random vector, and H_l is an $(m \times n)$ dimensional measurement matrix. V_l is the measurement noise; the mean of V_l is assumed to be zero, and the variance matrix is

$$\begin{cases} E(V_l V_l^T) = R_l \quad (R_l > 0), \\ E(V_l V_s^T) = 0 \quad (l \neq s). \end{cases} \quad (10.123)$$

From the measurement sequence of $Y_1, Y_2, \dots, Y_k$ using the least squares method we can obtain the estimators of $X_0, X_{0/k}$, which should make the sum of residual squares minimum:

$$\sum_{l=1}^k U = (Y_l - H_l \Phi_{l,0} \hat{X}_{0/k})^T (Y_l - H_l \Phi_{l,0} \hat{X}_{0/k}) = \min. \quad (10.124)$$

It requires

$$\frac{\partial U}{\partial \hat{X}_{0/k}} = -2 \sum_{l=1}^k \Phi_{l,0}^T H_l^T (Y_l - H_l \Phi_{l,0} \hat{X}_{0/k}) = 0,$$

which leads

$$\left(\sum_{l=1}^k \Phi_{l,0}^T H_l^T H_l \Phi_{l,0} \right) \hat{X}_{0/k} = \sum_{l=1}^k (\Phi_{l,0}^T H_l^T) Y_l. \quad (10.125)$$

This equation is usually called a normal equation, and the matrix

$$\sum_{l=1}^k (H_l \Phi_{l,0})^T (H_l \Phi_{l,0}) \quad (10.126)$$

is called a normal matrix.

The equations in (10.125) belong to a system of linear equations of the n -dimensional vector $\hat{X}_{0/k}$. It has a defined and unique solution only if the $(n \times n)$ dimensional normal matrix (10.126) is positive definite, i.e., the matrix $(H_l \Phi_{l,0})$ is a

full rank matrix. This condition is so called the observability in orbit determination, meaning that the measurement sequence of $Y_1, Y_2, \dots, Y_k$ can provide the only state variable X_0 at t_0 which satisfies (10.122). The solution of (10.125) is

$$\hat{X}_{0/k} = \left(\sum_{l=1}^k \Phi_{l,0}^T H_l^T H_l \Phi_{l,0} \right)^{-1} \left(\sum_{l=1}^k \Phi_{l,0}^T H_l^T Y_l \right). \quad (10.127)$$

It can be proved that the classical least squares estimator $\hat{X}_{0/k}$ is an unbiased estimator of X_0 that

$$E(\hat{X}_{0/k}) = \left(\sum_{l=1}^k \Phi_{l,0}^T H_l^T H_l \Phi_{l,0} \right)^{-1} \left(\sum_{l=1}^k \Phi_{l,0}^T H_l^T E(Y_l) \right).$$

Substituting

$$E(Y_l) = E(H_l X_l + V_l) = H_l X_l + E(V_l) = H_l \Phi_{l,0} X_0$$

into $E(\hat{X}_{0/k})$, we have

$$E(\hat{X}_{0/k}) = \left(\sum_{l=1}^k \Phi_{l,0}^T H_l^T H_l \Phi_{l,0} \right)^{-1} \left(\sum_{l=1}^k \Phi_{l,0}^T H_l^T H_l \Phi_{l,0} \right) X_0 = X_0. \quad (10.128)$$

Since the classical least squares estimator is a linear unbiased estimator and the sum of the squared residuals U is the minimum, then the estimator $\hat{X}_0$ is guaranteed to oscillate in the vicinity of the true value X_0 .

Obviously, the classical least squares estimator does not sufficiently use the statistic property of large amount random samplings due to the limitation of history. The development and improvement of the measurement techniques and the optimal estimation theory bring new content to this method such as the optimal weight treatment which we introduce below.

(2) The weighted least squares estimator

We now consider a situation in which the random samplings have different accuracies. Assuming that there is a measurement sequence of $Y_1, Y_2, \dots, Y_k$ and a corresponding weight sequence $W_1, W_2, \dots, W_k$, both are positive definite matrices, then the sum of the weighted residual squares is given by

$$U = \sum_{l=1}^k \left(Y_l - H_l \Phi_{l,0} \hat{X}_{0/k} \right)^T W_l \left(Y_l - H_l \Phi_{l,0} \hat{X}_{0/k} \right).$$

From

$$\frac{\partial U}{\partial \hat{X}_{0/k}} = 0$$

we have

$$\left(\sum_{l=1}^k \Phi_{l,0}^T H_l^T W_l H_l \Phi_{l,0} \right) \hat{X}_{0/k} = \sum_{l=1}^k (\Phi_{l,0}^T H_l^T W_l Y_l). \quad (10.129)$$

This is the weighted normal equation, which yields the formula for calculating the weighted least squares estimator $\hat{X}_{0/k}$ as

$$\hat{X}_{0/k} = \left(\sum_{l=1}^k \Phi_{l,0}^T H_l^T W_l H_l \Phi_{l,0} \right)^{-1} \left(\sum_{l=1}^k \Phi_{l,0}^T H_l^T W_l Y_l \right). \quad (10.130)$$

It can be proved that the weighted least squares estimator is also an unbiased estimator, i.e.,

$$E(\hat{X}_{0/k}) = X_0.$$

The estimation error is given by

$$\begin{aligned} \hat{X}_{0/k} - X_0 &= \hat{X}_{0/k} - E(\hat{X}_{0/k}) \\ &= \left(\sum_{l=1}^k \Phi_{l,0}^T H_l^T W_l H_l \Phi_{l,0} \right)^{-1} \left(\sum_{l=1}^k \Phi_{l,0}^T H_l^T W_l V_l \right), \end{aligned}$$

and the corresponding variance matrix is

$$\begin{aligned} \text{Var}(\hat{X}_{0/k} - X_0) &= \left(\sum_{l=1}^k \Phi_{l,0}^T H_l^T W_l H_l \Phi_{l,0} \right)^{-1} \\ &\times \left(\sum_{l=1}^k \Phi_{l,0}^T H_l^T W_l R_l W_l H_l \Phi_{l,0} \right) \left(\sum_{l=1}^k \Phi_{l,0}^T H_l^T W_l H_l \Phi_{l,0} \right)^{-1}. \end{aligned} \quad (10.131)$$

Deriving (10.131) uses the property of V_l (10.123) and $W_l = W_l^T$. It can be proved that when $W_l = R_l^{-1}$, this variance has its minimum. Details of proving are given in reference [5] Chap. 1 or reference [7] Chap. 14.

When

$$W_l = R_l^{-1}, \quad (10.132)$$

the estimator is given by

$$\hat{X}_{0/k} = \left(\sum_{l=1}^k \Phi_{l,0}^T H_l^T R_l^{-1} H_l \Phi_{l,0} \right)^{-1} \left(\sum_{l=1}^k \Phi_{l,0}^T H_l^T R_l^{-1} Y_l \right). \quad (10.133)$$

The corresponding variance matrix $\text{Var}(\hat{X}_{0/k} - X_0)$ denoted to $P_{0/k}$ is

$$P_{0/k} = \left(\sum_{l=1}^k \Phi_{l,0}^T H_l^T R_l^{-1} H_l \Phi_{l,0} \right)^{-1}. \quad (10.134)$$

This estimator is the optimal estimator when the linear unbiased variance is minimum, which is often applied in today's precise orbit determination. But it is difficult for the measurement sequence of $\{Y_k\}$ to match the demand of (10.123), therefore the selection of the weight matrix is restricted, in the precise orbit determination, it is often to choose a different weight matrix.

10.5.3 Two Processes of the Least Squares Estimator

The least squares estimator usually has two approaches, one is the batch processing and the other is the sequential processing.

The batch processing is the above given classical least squares estimator. By this method all the samplings in the measurement sequence of $Y_1, Y_2, \dots, Y_k$ are used in the estimation processing at the same time for obtaining the estimator of the state variable X_0 at t_0 , $\hat{X}_{0/k}$. There are historical reasons for accepting this process. As we know in the past the measurements were limited, and the results were not needed in a timely demand. In present days, there is a huge amount of data and the results need to be obtained as fast as possible. Based on this background, the sequence processing has been developed to speed up the calculation and avoid stockpiling huge data.

The sequence processing of the least squares estimator can be outlined as follows: firstly divide the measurement sequence of $Y_1, Y_2, \dots, Y_k, Y_{k+1}, Y_{k+1}, \dots, Y_{k+s}$ into two groups (or more than two), from $Y_1, Y_2, \dots, Y_k$ to obtain a least squares estimator of X_k , $\hat{X}_{k/k}$, then discard the used data and use $\hat{X}_{k/k}$ and the next group of data to obtain the least squares estimator of the state variable X_{k+s} , $\hat{X}_{k+s/k+s}$. This method uses the $k+s$ samplings as one group to obtain $\hat{X}_{0/k+s}$ so is different from the batch processing (here the time t_0 relative to the estimator of X_0 is t_{k+s} in the sequence processing).

According to the formulas of the least squares estimator (10.133) and (10.134), there are

$$\hat{X}_{k+s/k+s} = \left(\sum_{l=1}^{k+s} \Phi_{l,k+s}^T H_l^T R_l^{-1} H_l \Phi_{l,k+s} \right)^{-1} \left(\sum_{l=1}^{k+s} \Phi_{l,k+s}^T H_l^T R_l^{-1} Y_l \right), \quad (10.135)$$

$$P_{k+s/k+s} = \left(\sum_{l=1}^{k+s} \Phi_{l,k+s}^T H_l^T R_l^{-1} H_l \Phi_{l,k+s} \right)^{-1}. \quad (10.136)$$

Let

$$\hat{X}_{k+s/k} = \Phi_{k+s,k} \hat{X}_{k/k}, \quad (10.137)$$

$$P_{k+s/k} = \Phi_{k+s,k} P_{k/k} \Phi_{k+s,k}^T, \quad (10.138)$$

with

$$\Phi_{k+s,k}^{-1} = \Phi_{k,k+s},$$

we have

$$\begin{aligned} \hat{X}_{k+s/k+s} &= \left(P_{k+s/k}^{-1} + \sum_{l=k+1}^{k+s} \Phi_{l,k+s}^T H_l^T R_l^{-1} H_l \Phi_{l,k+s} \right)^{-1} \\ &\times \left(P_{k+1/k}^{-1} \hat{X}_{k+s/k} + \sum_{l=k+1}^{k+s} \Phi_{l,k+s}^T H_l^T R_l^{-1} Y_l \right), \end{aligned} \quad (10.139)$$

$$P_{k+s/k+s} = \left(P_{k+s/k}^{-1} + \sum_{l=k+1}^{k+s} \Phi_{l,k+s}^T H_l^T R_l^{-1} H_l \Phi_{l,k+s} \right)^{-1}. \quad (10.140)$$

Formulas (10.139) and (10.140) are the recursive formulas of the least squares estimator, the value of $\hat{X}_{k+s/k}$ can be regarded as a priori state estimator, and $P_{k+s/k}$ as the priori variance matrix estimator. Clearly that the measurements can be divided into more than two groups. For a linear system, the least squares estimator $\hat{X}_{k+s/k+s}$ derived by (10.139), (10.140) is equivalent to the result by the batch processing.

The amount of samplings in the first group is k , which is arbitrarily chosen, when $k = 1$, then the whole process is a step-by-step recurring process, which is like using the Kalman filter.

10.5.4 Least Squares Estimator with a Priori State Value

Generally, the least squares estimator is applied without a priori value of the state variable X_0 (or a priori state estimator), which is an important difference between the least squares method and the Kalman filter. When we use the above-described sequential processing, we can also derive a least squares estimation formula with a priori estimator.

If we have a measurement sequence of $Y_1, Y_2, \dots, Y_k$, a priori estimator of X_0 at t_0 , $\hat{X}_{0/0}$, and a corresponding variance matrix estimator $P_{0/0}$ ($P_{0/0}^{-1}$ is also called the priori weight matrix), then from (10.137), (10.138), we can derive a least squares estimator with a priori state value. The calculation formulas are

$$\hat{X}_{k/0} = \Phi_{k,0} \hat{X}_{0/0}, \quad (10.141)$$

$$P_{k/0} = \Phi_{k,0} P_{0/0} \Phi_{k,0}^T, \quad (10.142)$$

$$\hat{X}_{k/k} = \left(P_{k/0}^{-1} + \sum_{l=1}^k \Phi_{l,k}^T H_l^T R_l^{-1} H_l \Phi_{l,k} \right)^{-1} \left(P_{k/0}^{-1} \hat{X}_{k/0} + \sum_{l=1}^k \Phi_{l,k}^T H_l^T R_l^{-1} Y_l \right), \quad (10.143)$$

$$P_{k/k} = \left(P_{k/0}^{-1} + \sum_{l=1}^k \Phi_{l,k}^T H_l^T R_l^{-1} H_l \Phi_{l,k} \right)^{-1}, \quad (10.144)$$

or

$$\hat{X}_{0/k} = \left(P_{0/0}^{-1} + \sum_{l=1}^k \Phi_{l,0}^T H_l^T R_l^{-1} H_l \Phi_{l,0} \right)^{-1} \left(\begin{matrix} -1 \\ 0 \end{matrix} P \hat{X}_{0/0} + \sum_{l=1}^k \Phi_{l,0}^T H_l^T R_l^{-1} Y_l \right), \quad (10.145)$$

$$P_{0/k} = \left(P_{0/0}^{-1} + \sum_{l=1}^k \Phi_{l,0}^T H_l^T R_l^{-1} H_l \Phi_{l,0} \right)^{-1}, \quad (10.146)$$

where $\hat{X}_{k/0}$ can be regarded as the initial priori state estimator, and $P_{k/0}$ as the priori estimated variance matrix.

The least squares estimation method described in this section is widely used in the precise orbit determination for artificial satellites and all kinds of spacecraft. In posteriori treatments, the batch processing is also a popular method. The sequential processing using recurring calculation is often used in some real-time orbit determinations.

10.6 Orbit Determination by Ground-Based and Space-Based Joint Network and Autonomous Orbit Determination by Star-To-Star Measurements

For centuries scientifically the motions of natural celestial bodies are observed and measured by ground-based observatories and stations. This was the only method to study the motion of stars before the launching of artificial satellites. Nowadays the

whole system of measurement and control has been profoundly changed. There are all kinds of spacecraft with all kinds of equipment and technique in the space. In the field of orbit determination, there are space-based observatories and tracking stations, which can partially replace ground-based stations and also solve some problems which ground-based stations cannot achieve.

Presently there are two types of space-based systems. One is the Tracking and Data Relay Satellite System, which is a comprehensive aerospace measurement and control system mainly dealing with data relay; the other is the Navigation and Positioning System including global and regional navigations, which provide precise information of position, speed, and time for space or ground targets.

The relay satellite system is an important component of a space-based system. The tracking and data relay satellites are geosynchronous communication satellites, which transfer controlling signals and data from ground to low and medium Earth orbit satellites and spacecraft. It is like elevating ground-based tracking and control stations up to the altitude of the geosynchronous satellite. Three tracking and data relay satellites can form a space-based network of global measurement and control. The Navigation and Positioning System is also an important component of a space-based system, such as the American Global Position System (GPS), Russian GLONSS system, European Galileo system, and Chinese BeiDou system. The High Earth orbit navigation and position system can provide services for high precise positioning (position, speed, and time) of targets between ground to 36,000 km altitude over 24/7 real time.

Another orbit determination system is using a high orbit satellite-constellation to track other spacecraft and then determine the orbit. Every satellite in the constellation is like a space-tracking station, the whole constellation is like a space-based measurement and control network, and the related orbit determination system is the space-based orbit determination network.

Following the fast-increasing demands for satellite service, more and more space projects need satellite constellations to achieve, particularly the kind of constellation where the distances between satellites are close. This kind of constellation is different from the GPS constellation. The ground-based measurement and control system can track one or a few satellites in the constellation (called measurement-stars like space-tracking stations) and determine their orbits; the orbits of other satellites (called user-stars) can be determined using star-star measurements (measurement-star and user-star), or using ground-based measurements and star-star measurements to determine the orbits of the measurement-star and the user-star at the same time. This kind of orbit determination is the orbit determination by ground-based and space-based joint networks.

10.6.1 Outline of Space-Based Network of Orbit Tracking and Determination

The orbit determination includes tracking, measuring, and determining orbits. The space-based network of orbit tracking and determination usually provides measurements of distance and speed.

(1) Orbit determination for user-stars by the tracking and data relay satellite system

The tracking and data relay satellite system provides two-way tracking data of distance and speed for a user-star (or a user-spacecraft) to decide its orbit, or the user-star sends signals and data to a ground-based terminal through the space-based tracking and data relay satellite system to be analyzed. The orbit of the relay satellite also needs to be determined, which corresponds to a high Earth orbit determination.

(2) Orbit tracking and determination for user-stars by the navigation and positioning system

There are two types of orbit determination for a user-star using a navigation system, one uses the pseudocode to measure distances, and the other uses the carrier wave phase data. The data can be converted to the distance between a user-star and a navigation satellite, the ephemeris of the navigation satellite can be obtained from the received navigation information, then the user-star's orbit can be determined autonomously. If the demand for real time is not critical, the navigation satellite can send the measurements to a ground-based station, then the ground-station can obtain the precise ephemeris of the navigation satellite and determine the high precise orbit of the user-star.

(3) Relative measurements between satellites in a constellation

In a constellation, a link can be built between the satellites to track each other. Because they are outside of the atmosphere the influence of the ionosphere on measurements is reduced or even disappeared. The relative measurements can be used for autonomous orbit determination for the whole constellation network.

10.6.2 Basic Principles of the Orbit Determination of Ground-Based and Space-Based Joint Network

The space-based tracking and the ground-based tracking belong to different measurement and control systems, therefore, are related to different orbit determination methods. Currently, for medium and low Earth orbit user-stars in a navigation constellation or in a data relay satellite system, the method is that the orbit of a measurement-star (a high Earth orbit satellite) is given by the ground-based system, the user-star's orbit then is determined by the measurement-star. The measurement-star is like a station moving around Earth, the error in the position of the measurement station

is equivalent to the error in the position of a ground-based station but changing over time, which is an additional error to the orbit determination for the user-star. To overcome this problem, the orbit determination can be achieved by the ground-based and space-ground joint network with simultaneously measured data of the user-star from the ground and the space. Another case is in a constellation of low or medium Earth orbit satellites the orbit determination is achieved by using the data measured between the stars and ground measurements of one or more than one star in the constellation, the principle is the same as in the previous case. In the following context, we use distance data to describe the basic principles of the orbit determination of ground-based and space-based joint networks.

The tracking data include ground tracking data of a measurements-star and space tracking data of a user-star by the measurement-star. Assuming the ground tracking distance to be ρ_1 and the space tracking distance ρ_2 , and the state variable X as

$$X = \begin{pmatrix} \sigma_1 \\ \sigma_2 \end{pmatrix}, \quad (10.147)$$

where σ_1 and σ_2 are the orbital variables of the measurement-star and the user-star, respectively, which can be coordinates, velocities, or orbital elements.

The state differential equation is the same as for a single star that

$$\begin{cases} \dot{X} = F(X, t), \\ X(t_0) = X_0. \end{cases} \quad (10.148)$$

The measurement variable is Y , and the measurement equation is

$$Y = H(X, t) + V. \quad (10.149)$$

Specifically, the measured ρ_1 and ρ_2 are given by

$$\rho_1 = H_1(t, X) + V_1 = |\vec{r}_1 - \vec{r}_e| + V_1, \quad (10.150)$$

$$\rho_2 = H_2(t, X) + V_2 = |\vec{r}_2 - \vec{r}_1| + V_2, \quad (10.151)$$

where $\vec{r}_e$ is the position vector of the ground-based station. The corresponding condition equation has the same form as (10.8) that

$$y = \tilde{B}x + V, \quad (10.152)$$

$$y = Y_O - Y_C = \rho_O - \rho_C, \quad (10.153)$$

where y is the residual, Y_O and Y_C are the measured value and the calculated value of Y , respectively. The matrix $\tilde{B}$ has the form as

$$\tilde{B} = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} \left(\left(\frac{\partial \rho_1}{\partial(\vec{r}_1, \vec{r}_e)} \right) \left(\frac{\partial(\vec{r}_1, \vec{r}_e)}{\partial X} \right) \left(\frac{\partial X}{\partial X_0} \right) \right) \\ \left(\left(\frac{\partial \rho_2}{\partial(\vec{r}_2, \vec{r}_1)} \right) \left(\frac{\partial(\vec{r}_2, \vec{r}_1)}{\partial X} \right) \left(\frac{\partial X}{\partial X_0} \right) \right) \end{pmatrix}, \quad (10.154)$$

If the ground-based coordinates are not involved in the orbit determination, then (10.154) becomes

$$\tilde{B} = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} \left(\left(\frac{\partial \rho_1}{\partial \vec{r}_1} \right) \left(\frac{\partial \vec{r}_1}{\partial X} \right) \left(\frac{\partial X}{\partial X_0} \right) \right) \\ \left(\left(\frac{\partial \rho_2}{\partial(\vec{r}_2, \vec{r}_1)} \right) \left(\frac{\partial(\vec{r}_2, \vec{r}_1)}{\partial X} \right) \left(\frac{\partial X}{\partial X_0} \right) \right) \end{pmatrix}. \quad (10.155)$$

The actual formulas are not essentially different from those for a single star so are not provided here again.

From the above introduction, we can see that the process of orbit determination of ground-based and space-based joint networks is similar to that for a single star. It should be mentioned that without the ground-based support for the measurement-star, depending only on the relative distance data between the two stars the orbit of the user-star cannot be determined. This can be shown in the process of iteration, which would not converge. It is easy to understand intuitively. By the geometry of the orbit, the orbits of the two stars move together, and they are not connected to the ground, therefore, the orbit plane cannot be defined. It is also shown mathematically that the matrix $\tilde{B}$ does not have a full rank, which is discussed in the next section.

10.6.3 The Rank Deficiency in the Autonomous Orbit Determination by Start-To-Star Measurements

We now discuss the problem that whether an orbit can be determined using only the relative star-star distance measurements. The measurement equation is (10.151); the corresponding matrix $\tilde{B}$ is B_2 in (10.155). In order to better understand the problem, we use the state estimator X for Kepler orbital elements of two stars, that $\sigma = (a, e, i, \Omega, \omega, M)^T$, and there is

$$B_2 = \begin{pmatrix} \left(\frac{\partial \rho_2}{\partial(\vec{r}_2, \vec{r}_1)} \right) \left(\frac{\partial(\vec{r}_2, \vec{r}_1)}{\partial(\sigma_2, \sigma_1)} \right) \left(\frac{\partial(\sigma_2, \sigma_1)}{\partial(\sigma_{20}, \sigma_{10})} \right) \end{pmatrix}, \quad (10.156)$$

where

$$\left(\frac{\partial \rho_2}{\partial(\vec{r}_2, \vec{r}_1)} \right) = \frac{1}{\rho_2} ((x_2 - x_1), (y_2 - y_1), (z_2 - z_1), -(x_2 - x_1), -(y_2 - y_1), (z_2 - z_1)), \quad (10.157)$$

$$\begin{pmatrix} \frac{\partial(\vec{r}_2, \vec{r}_1)}{\partial(\sigma_2, \sigma_1)} \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} \frac{\partial \vec{r}_2}{\partial \sigma_2} \end{pmatrix} (0)_{3 \times 6} \\ (0)_{3 \times 6} \begin{pmatrix} \frac{\partial \vec{r}_1}{\partial \sigma_1} \end{pmatrix} \end{pmatrix}, \quad (10.158)$$

$$\begin{cases} \begin{pmatrix} \frac{\partial \vec{r}_2}{\partial \sigma_2} \end{pmatrix} = \begin{pmatrix} \frac{\vec{r}_2}{a_2} & H_2 \vec{r}_2 + K_2 \dot{\vec{r}}_2 & \frac{z_2}{\sin i_2} \hat{R}_2 & \vec{\Omega}_2 & \hat{R}_2 \times \vec{r}_2 & \frac{1}{n_2} \dot{\vec{r}}_2 \end{pmatrix} \\ \begin{pmatrix} \frac{\partial \vec{r}_1}{\partial \sigma_1} \end{pmatrix} = \begin{pmatrix} \frac{\vec{r}_1}{a_1} & H_1 \vec{r}_1 + K_1 \dot{\vec{r}}_1 & \frac{z_1}{\sin i_1} \hat{R}_1 & \vec{\Omega}_1 & \hat{R}_1 \times \vec{r}_1 & \frac{1}{n_1} \dot{\vec{r}}_1 \end{pmatrix} \end{cases} \quad (10.159)$$

For the state transition matrix $\begin{pmatrix} \frac{\partial(\sigma_2, \sigma_1)}{\partial(\sigma_{20}, \sigma_{10})} \end{pmatrix}$ only the unperturbed part is needed, which is

$$\begin{pmatrix} \frac{\partial(\sigma_2, \sigma_1)}{\partial(\sigma_{20}, \sigma_{10})} \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} \frac{\partial \sigma_2}{\partial \sigma_{20}} \end{pmatrix} (0)_{6 \times 6} \\ (0)_{6 \times 6} \begin{pmatrix} \frac{\partial \sigma_1}{\partial \sigma_{10}} \end{pmatrix} \end{pmatrix} + O(\varepsilon), \quad (10.160)$$

where

$$\begin{pmatrix} \frac{\partial \sigma_2}{\partial \sigma_{20}} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ -\frac{3n_2}{2a_2} \Delta t & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad (10.161)$$

$$\begin{pmatrix} \frac{\partial \sigma_1}{\partial \sigma_{10}} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ -\frac{3n_1}{2a_1} \Delta t & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \quad (10.162)$$

In (10.159), the related variables (omitted the subscripts 1 or 2) are

$$\begin{cases} H = -\frac{a}{p}(\cos E + e), & K = \frac{\sin E}{n} \left(1 + \frac{r}{p}\right), \\ p = a(1 - e^2), & n = \sqrt{\mu} a^{-\frac{3}{2}}, \quad \mu = GE, \end{cases} \quad (10.163)$$

$$\vec{\Omega} = \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}, \quad \hat{R} = \begin{pmatrix} \sin i \sin \Omega \\ -\sin i \cos \Omega \\ \cos i \end{pmatrix}. \quad (10.164)$$

In (10.163) E is the eccentric anomaly, $\mu = GE$ is the geocentric gravitational constant, and by the dimensionless normalized system, there is $\mu = 1$.

On the right side of (10.160), the magnitude of $O(\varepsilon)$ is for the perturbed part, which is a small quantity for the motion of an Earth's satellite and has no influence on the problem of rank deficiency. Through simple calculation using one star-to-star distance sampling, we derive the 12 elements in line 1 of the matrix B_2 , that

$$B_2 = \frac{1}{\rho_2} (B_{1,1} \ B_{1,2} \ \cdots \ B_{1,6} \ B_{1,7} \ \cdots \ B_{1,12}), \quad (10.165)$$

where

$$\begin{cases} B_{1,1} = (\vec{r}_2 - \vec{r}_1) \begin{pmatrix} \vec{r}_2 \\ a_2 \end{pmatrix} + (\vec{r}_2 - \vec{r}_1) \begin{pmatrix} \dot{\vec{r}}_2 \\ n_2 \end{pmatrix} \left(-\frac{3n_2}{2a_2} \Delta t \right), \\ B_{1,2} = (\vec{r}_2 - \vec{r}_1) \left(H_2 \vec{r}_2 + K_2 \dot{\vec{r}}_2 \right), \\ B_{1,3} = (\vec{r}_2 - \vec{r}_1) \begin{pmatrix} \vec{z}_2 \\ \sin i_2 \end{pmatrix} \hat{R}_2, \\ B_{1,4} = (\vec{r}_2 - \vec{r}_1) \vec{\Omega}_2, \\ B_{1,5} = (\vec{r}_2 - \vec{r}_1) \left(\hat{R}_2 \times \vec{r}_2 \right), \\ B_{1,6} = (\vec{r}_2 - \vec{r}_1) \begin{pmatrix} \dot{\vec{r}}_2 \\ n_2 \end{pmatrix}, \end{cases} \quad (10.166)$$

$$\begin{cases} B_{1,7} = -(\vec{r}_2 - \vec{r}_1) \begin{pmatrix} \vec{r}_1 \\ a_1 \end{pmatrix} - (\vec{r}_2 - \vec{r}_1) \begin{pmatrix} \dot{\vec{r}}_1 \\ n_1 \end{pmatrix} \left(-\frac{3n_1}{2a_1} \Delta t \right), \\ B_{1,8} = -(\vec{r}_2 - \vec{r}_1) \left(H_1 \vec{r}_1 + K_1 \dot{\vec{r}}_1 \right), \\ B_{1,9} = -(\vec{r}_2 - \vec{r}_1) \begin{pmatrix} \vec{z}_1 \\ \sin i_1 \end{pmatrix} \hat{R}_1, \\ B_{1,10} = -(\vec{r}_2 - \vec{r}_1) \vec{\Omega}_1, \\ B_{1,11} = -(\vec{r}_2 - \vec{r}_1) \left(\hat{R}_1 \times \vec{r}_1 \right), \\ B_{1,12} = -(\vec{r}_2 - \vec{r}_1) \begin{pmatrix} \dot{\vec{r}}_1 \\ n_1 \end{pmatrix}. \end{cases} \quad (10.167)$$

Based on the formulas of $\vec{\Omega}_2$ and $\vec{\Omega}_1$ in (10.164), $B_{1,4}$ and $B_{1,10}$ have the following forms:

$$\begin{cases} B_{1,4} = x_1 y_2 - y_1 x_2, \\ B_{1,10} = -(x_1 y_2 - y_1 x_2) = -B_{1,4}. \end{cases} \quad (10.168)$$

From (10.168) we can see that the absolute values of the elements in the matrix $\tilde{B}$ about Ω_1 and Ω_2 are strictly equal, therefore, this matrix is rank-defective, which means that the (12×12) determinant of the matrix $\tilde{B}^T \tilde{B}$ is zero, i.e., $|\tilde{B}^T \tilde{B}| = 0$, and the orbit cannot be determined. This is the problem of rank deficiency. The corresponding geometric property is that the orbits of the two stars drift together. More discussion about this problem is given in reference [8].

Another point about the orbit determination that should be mentioned is for other types of spacecraft, for example, a Moon's prober. When the prober is near the Moon and before becoming a Moon's satellite, it is still a high Earth orbit satellite, but the

perturbation from the Moon is as large as Earth's gravity force, then in the above discussion, $O(\varepsilon)$ is no longer a small quantity, and the above conclusion is not valid.

References

1. Liu L (2000) Orbital theory of spacecraft. National Defense Industry Press, Beijing
2. Schutz BE, Tapley BD, Born GH (2004) Statistical orbit determination. Academic Press, Amsterdam, Boston, Heidelberg, London, New York, Oxford, Paris, San Francisco, Singapore, Sydney, Tokyo
3. Liu L, Hu SJ (2015) Theory and application of spacecraft orbit determination. Electronic Industry Press, Beijing
4. Liu L, Zhang Q, Liao XH (1998) Problem of algorithm in precision orbit determination. Sci China (Ser A) 28(9):848–856; 1999 42(5):552–560
5. Jia PZ, Zhu ZT (1984) The optimal estimator and its application. Science Press, China
6. Jia PZ (1992) Error analysis and numerical method. National Defense Industry Press, Beijing
7. Liu L (1992) Orbital dynamics of artificial earth satellite. Higher Education Press, Beijing
8. Liu L, Liu YC (2000) The rank deficiency in autonomous orbit determination using star-to-star measurements. J of Spacecraft TT&C Technology 19(3):13–16

Appendix A

Astronomical Constants

The International System of Units (SI).

Length: meter (m), Mass: kilogram (kg), Time: second (s).

Defining constants

Speed of light	$c = 2.99792458 \times 10^8 \text{ ms}^{-1}$
Gaussian gravitational constant	$k = 0.01720209895 = 3548''.1876069651$
$1 - d(\text{TT})/d(\text{TCG})$	$L_G = 6.968290134 \times 10^{-10}$
$1 - d(\text{TDB})/d(\text{TCB})$	$L_B = 1.550519768 \times 10^{-8}$
$\text{TDB} - \text{TCB}(T_0 = 2443144.5003752)$	$\text{TDB}_0 = -6.55 \times 10^{-5} \text{ s}$
Earth rotation angle (J2000.0 UT1)	$\theta_0 = 2\pi \times 0.7790572732640 \text{ rad}$
Variation rate of Earth rotation Angle	$d\theta/d\text{UT1} = 2\pi \times 1.00273781191135548 \text{ rad/d(UT1)}$

Natural measurable constant

Constant of gravitation	$G = 6.67428 \times 10^{-11} \text{ m}^3/(\text{kg} \cdot \text{s}^2)$

Constants for celestial bodies

Heliocentric gravitational constant	$GS = 1.32712440041 \times 10^{20} \text{ m}^3 \text{ s}^{-2} (\text{TDB})$
Equatorial radius of Earth	$a_e = 6378136.6 \text{ m}$
Dynamical form-factor of Earth	$J_2 = 0.0010826359$
Flattening factor of Earth	$f = 0.0033528197 = 1/298.25642$
Geocentric gravitational constant	$GE = 3.986004356 \times 10^{14} \text{ m}^3 \text{ s}^{-2} (\text{TDB})$
Mean angular speed of Earth rotation	$\omega = 7.292115 \times 10^{-5} \text{ rad/s}$
Ratio of the Moon's mass to Earth's mass	$\mu = 0.0123000371$

Primary constants

An Astronomical Unit	$1 \text{ AU} = 1.49597870700 \times 10^{11} \text{ m}$
$1 - d(\text{TCG})/d(\text{TCB})$ (mean)	$L_C = 1.48082686741 \times 10^{-8}$
Light-time for unit distance	$\tau_A = 499.00478384 \text{ s}$

General precession in longitude per Julian century at standard epoch 2000	$p = 5028''.796195/\text{Juliancentury}$
Obliquity of the ecliptic at standard epoch 2000	$\varepsilon = 23^{\circ}26'21''.406$
Constant of nutation at standard epoch 2000	$N = 9''.2052331$
Solar parallax	$\pi = 8''.794143$
Constant of aberration at standard epoch 2000)	$\kappa = 20''.49551$
Mass of the Sun	$S = GS/G = 1.9884 \times 10^{30}\text{kg}$

Ratio of the Sun’s mass to a planet’s mass

Mercury	6,023,600	Mars	3,098,703.59
Venus	408,523.71	Jupiter	1047.348644
Earth	332,946.0487	Saturn	3497.9018
Earth + Moon	328,900.5596	Uranus	22,902.98
Neptune	19,412.26	Pluto*	136,566,000

*In 2006 IAU 26th General Assembly, Pluto was defined as a dwarf planet

Appendix B

Formulas of Mean Ephemeris of Major Celestial Bodies in the Solar System

There are two ways to obtain the ephemerides of major celestial bodies in the Solar System (which are the eight major planets and the Moon): one is from high precision Ephemerides, since 2005, the Ephemerides for major planets is DE405, which can be downloaded from the internet; the other is by calculation using analytical formulas of orbit elements derived by solving perturbed motion equations. We list the formulas for calculating mean orbit elements provided by reference [1]. Analytical formulas with higher accuracy are given in reference [2].

Mean Orbit Elements of Major Planets

The formulas listed here are for the values in the heliocentric ecliptic coordinate system at epoch J2000.0.

Mean orbit elements of Mercury:

$$\left\{ \begin{array}{l} a = 0.38709831 \text{ AU}, \\ e = 0.20563175 + 0.000020406T - 0.000000028T^2, \\ i = 7^\circ.004986 + 0^\circ.0018215T - 0^\circ.0000181T^2, \\ \Omega = 48^\circ.330893 + 1^\circ.1861882T + 0^\circ.0001759T^2, \\ \tilde{\omega} = 77^\circ.456119 + 1^\circ.5564775T + 0^\circ.0002959T^2, \\ L = 174^\circ.794787 + 4^\circ.092334444960d + 0^\circ.0000081T^2, \\ M = 174^\circ.794787 + 4^\circ.092334444960d + 0^\circ.0000081T^2, \\ n = 4^\circ.092339/d. \end{array} \right. \quad (\text{B.1})$$

Mean orbit elements of Venus:

$$\left\{ \begin{array}{l} a = 0.72332982 \text{ AU}, \\ e = 0.00677188 - 0.000047765T + 0.000000097T^2, \\ i = 3^\circ.394662 + 0^\circ.0010037T - 0^\circ.0000009T^2, \\ \Omega = 76^\circ.679920 + 0^\circ.9011204T + 0^\circ.0004066T^2, \\ \tilde{\omega} = 131^\circ.563707 + 1^\circ.4022289T - 0^\circ.0010729T^2, \\ L = 181^\circ.979801 + 1^\circ.60216873457d + 0^\circ.0003106T^2, \\ M = 50^\circ.416094 + 1^\circ.60213034364d + 0^\circ.0013835T^2, \\ n = 1^\circ.602130/d. \end{array} \right. \quad (\text{B.2})$$

Mean orbit elements of Earth:

$$\left\{ \begin{array}{l} a = 1.00000102 \text{ AU}, \\ e = 0.01670862 - 0.000042040T - 0.000001240T^2, \\ i = 0^\circ.0, \\ \Omega = 0^\circ.0, \\ \tilde{\omega} = 102^\circ.937347 + 0^\circ.3225621T - 0^\circ.0001576T^2, \\ M = 357^\circ.529100 + 0^\circ.98560028169d - 0^\circ.0001561T^2. \end{array} \right. \quad (\text{B.3})$$

Mean orbit elements of Mars:

$$\left\{ \begin{array}{l} a = 1.52367934 \text{ AU}, \\ e = 0.09340062 + 0.000090484T - 0.000000081T^2, \\ i = 1^\circ.849726 - 0^\circ.0006011T - 0^\circ.0000128T^2, \\ \Omega = 49^\circ.558093 + 0^\circ.7720956T + 0^\circ.0000161T^2, \\ \tilde{\omega} = 336^\circ.060234 + 1^\circ.8410446T + 0^\circ.0001351T^2, \\ L = 355^\circ.433275 + 0^\circ.52407108760d + 0^\circ.0003110T^2, \\ M = 19^\circ.373041 + 0^\circ.52402068219d + 0^\circ.0001759T^2, \\ n = 0^\circ.504033/d. \end{array} \right. \quad (\text{B.4})$$

Mean orbit elements of Jupiter:

$$\left\{ \begin{array}{l} a = 5.20260319 + 0.0000001913T (\text{AU}), \\ e = 0.04849485 + 0.000163244T - 0.000000472T^2, \\ i = 1^\circ.303270 - 0^\circ.0054966T + 0^\circ.0000046T^2, \\ \Omega = 100^\circ.464441 + 1^\circ.0209542T + 0^\circ.0004011T^2, \\ \tilde{\omega} = 14^\circ.331309 + 1^\circ.6126383T + 0^\circ.0010314T^2, \\ L = 34^\circ.351484 + 0^\circ.08312943981d + 0^\circ.0002237T^2, \\ M = 20^\circ.020175 + 0^\circ.08308528818d - 0^\circ.0008077T^2, \\ n = 4^\circ.0830912/d. \end{array} \right. \quad (\text{B.5})$$

Mean orbit elements of Saturn:

$$\left\{ \begin{array}{l} a = 9.5549096 - 0.000002139T (\text{AU}), \\ e = 0.05550862 - 0.000346818T - 0.000000646T^2, \\ i = 2^\circ.488878 - 0^\circ.0037362T - 0^\circ.0000152T^2, \\ \Omega = 113^\circ.665524 + 0^\circ.8770949T - 0^\circ.0001208T^2, \\ \tilde{\omega} = 93^\circ.056787 + 1^\circ.9637685T + 0^\circ.0008375T^2, \\ L = 50^\circ.077471 + 0^\circ.03349790593d + 0^\circ.0005195T^2, \\ M = 317^\circ.020684 + 0^\circ.03344414088d - 0^\circ.0003180T^2, \\ n = 0^\circ.0334597/d. \end{array} \right. \quad (\text{B.6})$$

Mean orbit elements of Uranus:

$$\left\{ \begin{array}{l} a = 19.2184461 - 0.00000037T (\text{AU}), \\ e = 0.04629590 - 0.000027337T - 0.000000079T^2, \\ i = 0^\circ.773196 + 0^\circ.0007744T - 0^\circ.0000375T^2, \\ \Omega = 74^\circ.005947 + 0^\circ.5211258T + 0^\circ.0013399T^2, \\ \tilde{\omega} = 173^\circ.005159 + 1^\circ.4863784T + 0^\circ.0002145T^2, \\ L = 314^\circ.055005 + 0^\circ.01176903644d + 0^\circ.0003043T^2, \\ M = 141^\circ.049846 + 0^\circ.01172834162d + 0^\circ.0000898T^2, \\ n = 0^\circ.0117308/d. \end{array} \right. \quad (\text{B.7})$$

Mean orbit elements of Neptune:

$$\left\{ \begin{array}{l} a = 30.1103869 - 0.000000166T (\text{AU}), \\ e = 0.00898809 + 0.000006408T - 0.000000001T^2, \\ i = 1^\circ.769952 + 0^\circ.0093082T - 0^\circ.0000071T^2, \\ \Omega = 131^\circ.784057 + 1^\circ.1022035T + 0^\circ.0002600T^2, \\ \tilde{\omega} = 48^\circ.123691 + 1^\circ.4262678T + 0^\circ.0003792T^2, \\ L = 304^\circ.348665 + 0^\circ.00602007691d + 0^\circ.0003093T^2, \\ M = 256^\circ.224974 + 0^\circ.00598102783d - 0^\circ.00006991T^2, \\ n = 0^\circ.0059818/d. \end{array} \right. \quad (\text{B.8})$$

The above orbit elements (a , e , i , Ω , $\tilde{\omega}$, L) are semi-major axis, eccentricity, inclination, longitude of the ascending node, longitude of the perihelion, and mean longitude; AU is the Astronomical unit of distance that $1\text{AU} = 1.49597870 \times 10^{11}$ m, and $\tilde{\omega}$, L , n are defined as follows.

$$\left\{ \begin{array}{l} \tilde{\omega} = \omega + \Omega, \\ L = \tilde{\omega} + M = \omega + \Omega + M, \\ n = \sqrt{\mu}a^{-\frac{3}{2}}, \mu = G(S + m), \end{array} \right. \quad (\text{B.9})$$

where n is the mean angular speed of a planet moving around the Sun, μ and m are the gravitational constant and mass of the planet, respectively, S is the mass of the

Sun. The time units, d and T , are Julian day and Julian century started at J2000.0, respectively, that

$$T = [\text{JD}(t) - \text{JD}(\text{J2000.0})]/36525.0 \quad (\text{B.10})$$

Orbit Elements of the Moon

Formulas of approximate calculation for the Moon's orbit elements at epoch J2000.0 in the geocentric ecliptic coordinate system:

$$\left\{ \begin{array}{l} a = 0.0025718814\text{AU} = 384747.981\text{km}, \\ e = 0.054879905, \\ i = J = 5^\circ.129835071, \\ \Omega = 125^\circ.044556 - 1934^\circ.1361850T + 0^\circ.0020767T^2, \\ \omega = 318^\circ.308686 + 6003^\circ.1498961T - 0^\circ.0124003T^2, \\ M = 134^\circ.963414 + 13^\circ.06499315537d + 0^\circ.0089939T^2. \end{array} \right. \quad (\text{B.11})$$

The variation of the Moon's orbit is relatively large, the largest amplitude of the periodic terms in the elements due to perturbations can be 2×10^{-2} , therefore the accuracy of the above formulas is low.

Another Calculation Method of Orbit Elements of Major Planets

For comparison, we list another set of formulas for calculating orbit elements of major planets provided in reference [3]. The values are for the heliocentric ecliptic coordinate system at epoch J2000.0.

$$a = a_0 + \dot{a}T, \quad (\text{B.12})$$

$$e = e_0 + \dot{e}T, \quad (\text{B.13})$$

$$i = i_0 + (\dot{i}/3600)T, \quad (\text{B.14})$$

$$\tilde{\omega} = \tilde{\omega}_0 + (\dot{\tilde{\omega}}/3600)T, \quad (\text{B.15})$$

$$\Omega = \Omega_0 + (\dot{\Omega}/3600)T, \quad (\text{B.16})$$

$$\lambda = \lambda_0 + (\dot{\lambda}/3600 + 360N_r)T. \tag{B.17}$$

The definitions of T and orbit elements ($a, e, i, \tilde{\omega}, \Omega, \lambda$) are the same as previous, and $\lambda = L$. The values of the orbit elements at epoch J2000.0 ($a_0, e_0, i_0, \tilde{\omega}_0, \Omega_0, \lambda_0$) and their variabilities $(\dot{a}_0, \dot{e}_0, \dot{i}_0, \dot{\tilde{\omega}}_0, \dot{\Omega}_0, \dot{\lambda}_0, N_r)$ are listed in Tables B.1 and B.2, respectively.

In Table B.2, the variabilities are for every Julian century; the values of $\dot{a}$ and $\dot{e}$ are 10^8 times the actual values, i.e., $\dot{a} \times 10^8$ and $\dot{e} \times 10^8$, the unit of $\dot{i}, \dot{\tilde{\omega}}, \dot{\Omega}$, and $\dot{\lambda}$ is arc-second ($1^\circ = 3600''$). All the values in the rows under Earth in both tables are actually for the Earth-Moon system.

These formulas provided by reference [3] are for the arcs between the year 1800 and the year 2050, the maximum error is $600''$, i.e., 3×10^{-3} , which are suitable for general analytical study.

Table B.1 Orbit elements of major planets at Epoch J2000 (JD2451545.0)

Planet	a_0 (AU)	e_0	i_0 (deg)	$\tilde{\omega}_0$ (deg)	Ω_0 (deg)	λ_0 (deg)
Mercury	0.38709893	0.20563069	7.00487	77.45645	48.33167	252.26084
Venus	0.72333199	0.00677323	3.39471	131.53298	76.68069	181.97973
Earth	1.00000011	0.01671022	0.00005	102.94719	348.73936	100.46435
Mars	1.52366231	0.09341233	1.85061	336.04084	49.57854	355.45332
Jupiter	5.20336301	0.04839266	1.30530	14.75385	100.55615	34.40438
Saturn	9.53707032	0.05415060	2.48446	92.43194	113.71504	49.94432
Uranus	19.19126393	0.04616771	0.76986	170.96424	74.22988	313.23218
Neptune	30.06896348	0.00858587	1.76917	44.97135	131.72169	304.88003

Table B.2 Variabilities of orbit elements of major planets at epoch J2000 (JD2451545.0)

Planet	$\dot{a}_0$	$\dot{e}_0$	$\dot{i}_0$	$\dot{\tilde{\omega}}_0$	$\dot{\Omega}_0$	$\dot{\lambda}_0$	N_r
Mercury	66	2527	−23.51	573.57	−446.30	261,628.29	415
Venus	92	−4938	−2.86	−108.80	−996.89	712,136.06	162
Earth	−5	−3804	−46.94	1198.28	−18,228.25	1,293,740.63	99
Mars	−7221	11,902	−25.47	1560.78	−1020.19	217,103.78	53
Jupiter	60,737	−12,880	−4.15	839.93	1217.17	557,078.35	8
Saturn	−301,530	−36,762	6.11	−1948.89	−1591.05	513,052.95	3
Uranus	152,025	−19,150	−2.09	1312.56	1681.40	246,547.79	1
Neptune	−125,196	2514	−3.64	−844.43	−151.25	786,449.21	0

Appendix C

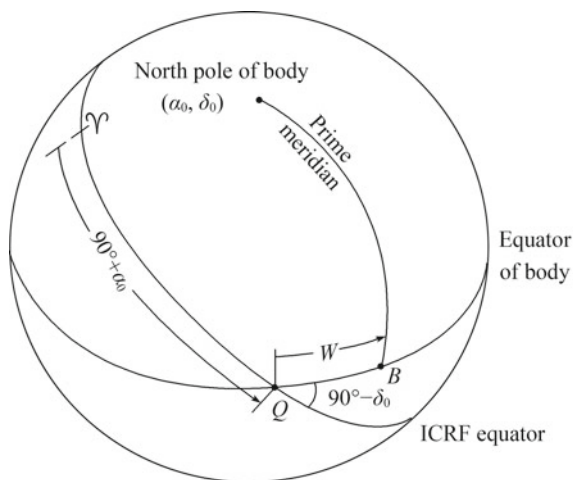
Orientation Models of Major Celestial bodies in the Solar System

Orientation models for major celestial bodies in the Solar System are illustrated in Fig. C.1. In the figure the right ascension and declination of the mean north pole of a celestial body are denoted to α_0 and δ_0 , respectively; Q is the ascending node at the equator, the angle between Q and the March equinox is $90^\circ + \alpha_0$, W is the angle between Q and the prime meridian B . This model directly connects the celestial coordinate system and the body-fixed coordinate system of a celestial body with the Earth's celestial coordinate system and Earth-fixed coordinate system. Details about the model are given in reference [4].

Based on this model, the transformation of $\vec{r}$ in the J2000.0 celestial coordinate system which is the International Celestial Reference Frame (ICRF) to $\vec{R}$ in the body-fixed coordinate system is given by

$$\vec{R} = R_z(W)R_x(90 - \delta_0)R_z(90 + \alpha_0)\vec{r}. \quad (\text{C.1})$$

Fig. C.1 The orientation Model for a celestial body in the Solar System



The definition of the rotation matrix $R_{x,y,z}(\theta)$ is given by (1.20)–(1.22) in Chap. 1.

For the requirements of deep-space exploration, we list the equatorial coordinates of the North Pole and the prime meridian, and other parameters for the major planets and the Moon in the Solar System as follows.

Mercury:

$$\alpha_0 = 281^\circ.0097 - 0^\circ.0328T,$$

$$\delta_0 = 61^\circ.4143 - 0^\circ.0049T,$$

$$\begin{aligned} W = 329^\circ.5469 + 6^\circ.1385025d + 0^\circ.00993822 \sin(M1) - 0^\circ.00104581 \sin(M2) \\ - 0^\circ.00010280 \sin(M3) - 0^\circ.00002364 \sin(M4) \\ - 0^\circ.00000532 \sin(M5), \end{aligned}$$

where,

$$M_1 = 174^\circ.791086 + 4^\circ.092335d,$$

$$M_2 = 349^\circ.582171 + 8^\circ.184670d,$$

$$M_3 = 164^\circ.373257 + 12^\circ.277005d,$$

$$M_4 = 339^\circ.164343 + 16^\circ.369340d,$$

$$M_5 = 153^\circ.955429 + 20^\circ.461675d.$$

Venus:

$$\alpha_0 = 272^\circ.76,$$

$$\delta_0 = 67^\circ.16,$$

$$W = 160^\circ.20 - 1^\circ.4813688d.$$

Earth:

$$\alpha_0 = 0^\circ.00 - 0^\circ.641T,$$

$$\delta_0 = 90^\circ.00 - 0^\circ.557T,$$

$$W = 190^\circ.147 + 360^\circ.9856235d.$$

Mars:

$$\alpha_0 = 317^\circ.68143 - 0^\circ.1061T,$$

$$\delta_0 = 52^\circ.88650 - 0^\circ.0609T,$$

$$W = 176^\circ.630 + 350^\circ.89198226d.$$

Jupiter:

$$\begin{aligned}
 \alpha_0 &= 268^\circ.056595 - 0^\circ.006499T + 0^\circ.000117\sin(J_a) + 0^\circ.000938\sin(J_b) \\
 &\quad + 0^\circ.001432\sin(J_c) + 0^\circ.000030\sin(J_d) + 0^\circ.002150\sin(J_e), \\
 \delta_0 &= 64^\circ.495303 + 0^\circ.002413T + 0^\circ.000050\cos(J_a) \\
 &\quad + 0^\circ.000404\cos(J_b) + 0^\circ.000617\cos(J_c) - 0^\circ.000013\cos(J_d) \\
 &\quad + 0^\circ.000926\cos(J_e), \\
 W &= 284^\circ.95 + 870^\circ.5360000d,
 \end{aligned}$$

where

$$\begin{aligned}
 J_a &= 99^\circ.360714 + 4850^\circ.4046T, \quad J_b = 175^\circ.895369 + 1191^\circ.9605T, \\
 J_c &= 300^\circ.323162 + 262^\circ.5475T, \quad J_d = 114^\circ.012305 + 6070^\circ.2476T, \\
 J_e &= 49^\circ.511251 + 64^\circ.3000T.
 \end{aligned}$$

Saturn:

$$\begin{aligned}
 \alpha_0 &= 40^\circ.589 - 0^\circ.036T, \\
 \delta_0 &= 83^\circ.537 - 0^\circ.004T, \\
 W &= 38^\circ.90 + 810^\circ.7939024d.
 \end{aligned}$$

Uranus:

$$\begin{aligned}
 \alpha_0 &= 257^\circ.311, \\
 \delta_0 &= -15^\circ.175, \\
 W &= 203^\circ.81 - 501^\circ.1600928d.
 \end{aligned}$$

Neptune:

$$\begin{aligned}
 \alpha_0 &= 299^\circ.36 + 0^\circ.70\sin(N), \\
 \delta_0 &= 43^\circ.46 - 0^\circ.51\cos(N), \\
 W &= 253^\circ.18 + 536^\circ.3128492d - 0^\circ.48\sin(N), \\
 N &= 357^\circ.85 + 52^\circ.316T.
 \end{aligned}$$

Moon:

$$\begin{aligned}
 \alpha_0 &= 269^\circ.9949 + 0^\circ.0031T - 3^\circ.8787\sin(E_1) - 0^\circ.1204\sin(E_2) \\
 &\quad + 0^\circ.0700\sin(E_3) - 0^\circ.0172\sin(E_4) + 0^\circ.0072\sin(E_6) \\
 &\quad - 0^\circ.0052\sin(E_{10}) + 0^\circ.0043\sin(E_{13}),
 \end{aligned}$$

$$\begin{aligned}\delta_0 = & 66^\circ.5392 + 0^\circ.0130T + 1^\circ.5419 \cos(E_1) + 0^\circ.0239 \cos(E_2) \\ & - 0^\circ.0278 \cos(E_3) + 0^\circ.0068 \cos(E_4) - 0^\circ.0029 \cos(E_6) \\ & + 0^\circ.0009 \cos(E_7) + 0^\circ.0008 \cos(E_{10}) - 0^\circ.0009 \cos(E_{13}),\end{aligned}$$

$$\begin{aligned}W = & 38^\circ.3213 + 13^\circ.17635815d - 1^\circ.4 \times 10^{-12}d^2 \\ & + 3^\circ.5610 \sin(E_1) + 0^\circ.1208 \sin(E_2) - 0^\circ.0642 \sin(E_3) \\ & + 0^\circ.0158 \sin(E_4) + 0^\circ.0252 \sin(E_5) - 0^\circ.0066 \sin(E_6) \\ & - 0^\circ.0047 \sin(E_7) - 0^\circ.0046 \sin(E_8) + 0^\circ.0028 \sin(E_9) \\ & + 0^\circ.0052 \sin(E_{10}) + 0^\circ.0040 \sin(E_{11}) \\ & + 0^\circ.0019 \sin(E_{12}) - 0^\circ.0044 \sin(E_{13}),\end{aligned}$$

where

$$\begin{aligned}E_1 &= 125^\circ.045 - 0^\circ.0529921d, & E_2 &= 250^\circ.089 - 0^\circ.1059842d, \\ E_3 &= 260^\circ.008 + 13^\circ.0120009d, & E_4 &= 176^\circ.625 + 13^\circ.3407154d, \\ E_5 &= 357^\circ.529 + 0^\circ.9856003d, & E_6 &= 311^\circ.589 + 26^\circ.4057084d, \\ E_7 &= 134^\circ.963 + 13^\circ.0649930d, & E_8 &= 276^\circ.617 + 0^\circ.3287146d, \\ E_9 &= 34^\circ.226 + 1^\circ.7484877d, & E_{10} &= 15^\circ.134 - 0^\circ.1589763d, \\ E_{11} &= 119^\circ.743 + 0^\circ.0036096d, & E_{12} &= 239^\circ.961 + 0^\circ.1643573d, \\ E_{13} &= 25^\circ.053 + 12^\circ.9590088d.\end{aligned}$$

In the above formulas, the time units, d and T , are Julian day and Julian century started at J2000.0, respectively, the standard epoch J2000.0 means 1 January 12:00 2000 TDB.

For Earth and the Moon, there are high precision coordinate transformation formulas that are introduced in this book. The IAU definitions are as additional programs, which are not necessarily the best, especially for Earth, because the nutation of Earth is relatively large, the IAU program is not suitable.

References

1. Bretagnon P (1982) Planetary Ephemerides. VSOP82.
2. Simon J, Bretagnon P, Chapron J et al (1994) Numerical expressions for precision formulae and mean elements for the Moon and the planets. *A & Ap* 282:663–683
3. Murray CD, Dermott SF (1999) *Solar System Dynamics* (Appendix A). Cambridge University Press
4. Archinal BA, A’Hearn MF, Bowell E, Conrad A, Consolmagno GJ, Courtin R, Fukushima T, Hestroffer D, Hilton JL, Neumann KGA, G, Oberst J, Seidelmann P K, Stooke P, Tholen D J, Thomas P C, Williams IP (2011) Report of the IAU working group on cartographic coordinates and rotational elements: 2009. *Celest Mech Dyn Astron* 109:101–135