RICHARD FITZPATRICK

AN INTRODUCTION TO CELESTIAL MECHANICS

CAMBRIDGE

An Introduction to Celestial Mechanics

RICHARD FITZPATRICK
University of Texas at Austin

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town,
Singapore, São Paulo, Delhi, Mexico City

Cambridge University Press
32 Avenue of the Americas, New York, NY 10013-2473, USA

www.cambridge.org
Information on this title: www.cambridge.org/9781107023819

First published 2012

Printed in the United States of America

A catalog record for this publication is available from the British Library.

Library of Congress Cataloging in Publication data

Fitzpatrick, Richard, 1963–
An introduction to celestial mechanics / Richard Fitzpatrick.
p. cm.
Includes bibliographical references and index.
ISBN 978-1-107-02381-9 (hardback)
1. Celestial mechanics. I. Title.
QB351.F565 2012
521–dc23 2012000780

ISBN 978-1-107-02381-9 Hardback

An Introduction to Celestial Mechanics

This accessible text on classical celestial mechanics—the principles governing the motions of bodies in the solar system—provides a clear and concise treatment of virtually all the major features of solar system dynamics. Building on advanced topics in classical mechanics, such as rigid body rotation, Lagrangian mechanics, and orbital perturbation theory, this text has been written for well-prepared undergraduates and beginning graduate students in astronomy, physics, mathematics, and related fields. Specific topics covered include Keplerian orbits; the perihelion precession of the planets; tidal interactions among the Earth, Moon, and Sun; the Roche radius; the stability of Lagrange points in the three-body problem; and lunar motion. More than 100 exercises allow students to gauge their understanding; a solutions manual is available to instructors. Suitable for a first course in celestial mechanics, this text is the ideal bridge to higher-level treatments.

Richard Fitzpatrick is professor of physics at the University of Texas at Austin, where he has been a faculty member since 1994. He earned his MA in physics at the University of Cambridge and his DPhil in astronomy at the University of Sussex. He is a longstanding Fellow of the Royal Astronomical Society and the American Physical Society and author of *Maxwell's Equations and the Principles of Electromagnetism* (2008).

Preface

The aim of this book is to bridge the considerable gap that exists between standard undergraduate mechanics texts, which rarely cover topics in celestial mechanics more advanced than two-body orbit theory, and graduate-level celestial mechanics texts, such as the well-known books by Moulton (1914), Brouwer and Clemence (1961), Danby (1992), Murray and Dermott (1999), and Roy (2005). The material presented here is intended to be intelligible to an advanced undergraduate or beginning graduate student with a firm grasp of multivariate integral and differential calculus, linear algebra, vector algebra, and vector calculus.

The book starts with a discussion of the fundamental concepts of Newtonian mechanics, as these are also the fundamental concepts of celestial mechanics. A number of more advanced topics in Newtonian mechanics that are needed to investigate the motions of celestial bodies (e.g., gravitational potential theory, motion in rotating reference frames, Lagrangian mechanics, Eulerian rigid body rotation theory) are also described in detail in the text. However, any discussion of the application of Hamiltonian mechanics, Hamilton-Jacobi theory, canonical variables, and action-angle variables to problems in celestial mechanics is left to more advanced texts (see, for instance, Goldstein, Poole, and Safko 2001).

Celestial mechanics (a term coined by Laplace in 1799) is the branch of astronomy that is concerned with the motions of celestial objects—in particular, the objects that make up the solar system—under the influence of gravity. The aim of celestial mechanics is to reconcile these motions with the predictions of Newtonian mechanics. Modern analytic celestial mechanics started in 1687 with the publication of the *Principia* by Isaac Newton (1643–1727) and was subsequently developed into a mature science by celebrated scientists such as Euler (1707–1783), Clairaut (1713–1765), D'Alembert (1717–1783), Lagrange (1736–1813), Laplace (1749–1827), and Gauss (1777–1855). This book is largely devoted to the study of the "classical" problems of celestial mechanics that were investigated by these scientists. These problems include the figure of the Earth; tidal interactions among the Earth, Moon, and Sun; the free and forced precession and nutation of the Earth; the three-body problem; the secular evolution of the solar system; the orbit of the Moon; and the axial rotation of the Moon. However, any discussion of the highly complex problems that arise in modern celestial mechanics, such as the mutual gravitational interaction between the various satellites of Jupiter and Saturn, the formation of the Kirkwood gaps, the dynamics of planetary rings, and the ultimate stability of the solar system, is again left to more advanced texts (see, in particular, Murray and Dermott 1999).

A number of topics, closely related to classical celestial mechanics, are not discussed in this book for the sake of brevity. The first of these is positional astronomy—the branch of astronomy that is concerned with finding the positions of celestial objects in the Earth's sky at a particular instance in time. Interested readers are directed to Smart (1977). The second excluded topic is the development of numerical methods for the solution of problems in celestial mechanics. Interested readers are directed to Danby (1992). The third (mostly) excluded topic is astrodynamics: the application of Newtonian dynamics to the design and analysis of orbits for artificial satellites and space probes. Interested readers are directed to Bate, Mueller, and White (1977). The final excluded topic is the determination of the orbits of celestial objects from observational data. Interested readers are again directed to Danby (1992).

Contents

1 Newtonian mechanics

1.1 Introduction

Newtonian mechanics is a mathematical model whose purpose is to account for the motions of the various objects in the universe. The general principles of this model were first enunciated by Sir Isaac Newton in a work titled *Philosophiae Naturalis Principia Mathematica* (Mathematical Principles of Natural Philosophy). This work, which was published in 1687, is nowadays more commonly referred to as the *Principia*.[1]

Until the beginning of the twentieth century, Newtonian mechanics was thought to constitute a *complete* description of all types of motion occurring in the universe. We now know that this is not the case. The modern view is that Newton's model is only an *approximation* that is valid under certain circumstances. The model breaks down when the velocities of the objects under investigation approach the speed of light in a vacuum, and must be modified in accordance with Einstein's *special theory of relativity*. The model also fails in regions of space that are sufficiently curved that the propositions of Euclidean geometry do not hold to a good approximation, and must be augmented by Einstein's *general theory of relativity*. Finally, the model breaks down on atomic and subatomic length scales, and must be replaced by *quantum mechanics*. In this book, we shall (almost entirely) neglect relativistic and quantum effects. It follows that we must restrict our investigations to the motions of *large* (compared with an atom), *slow* (compared with the speed of light) objects moving in *Euclidean* space. Fortunately, virtually all the motions encountered in conventional celestial mechanics fall into this category.

Newton very deliberately modeled his approach in the *Principia* on that taken in Euclid's *Elements*. Indeed, Newton's theory of motion has much in common with a conventional *axiomatic system*, such as Euclidean geometry. Like all axiomatic systems, Newtonian mechanics starts from a set of terms that are *undefined* within the system. In this case, the fundamental terms are *mass*, *position*, *time*, and *force*. It is taken for granted that we understand what these terms mean, and, furthermore, that they correspond to *measurable* quantities that can be ascribed to, or associated with, objects in the world around us. In particular, it is assumed that the ideas of position in space, distance in space, and position as a function of time in space are correctly described by conventional Euclidean vector algebra and vector calculus. The next component of an axiomatic system is a set of *axioms*. These are a set of *unproven* propositions,

[1] An excellent discussion of the historical development of Newtonian mechanics, as well as the physical and philosophical assumptions that underpin this theory, is given in Barbour 2001.

involving the undefined terms, from which all other propositions in the system can be derived via logic and mathematical analysis. In the present case, the axioms are called *Newton's laws of motion* and can be justified only via experimental observation. Note, incidentally, that Newton's laws, in their primitive form, are applicable only to *point objects* (i.e., objects of negligible spatial extent). However, these laws can be applied to extended objects by treating them as collections of point objects.

One difference between an axiomatic system and a physical theory is that, in the latter case, even if a given prediction has been shown to follow necessarily from the axioms of the theory, it is still incumbent on us to test the prediction against experimental observations. Lack of agreement might indicate faulty experimental data, faulty application of the theory (for instance, in the case of Newtonian mechanics, there might be forces at work that we have not identified), or, as a last resort, incorrectness of the theory. Fortunately, Newtonian mechanics has been found to give predictions that are in excellent agreement with experimental observations in all situations in which it would be expected to hold.

In the following, it is assumed that we know how to set up a rigid Cartesian frame of reference and how to measure the positions of point objects as functions of time within that frame. It is also taken for granted that we have some basic familiarity with the laws of mechanics.

1.2 Newton's laws of motion

Newton's laws of motion, in the rather obscure language of the *Principia*, take the following form:

1. Every body continues in its state of rest, or uniform motion in a straight line, unless compelled to change that state by forces impressed on it.
2. The change of motion (i.e., momentum) of an object is proportional to the force impressed on it, and is made in the direction of the straight line in which the force is impressed.
3. To every action there is always opposed an equal reaction; or, the mutual actions of two bodies on each other are always equal, and directed to contrary parts.

Let us now examine how these laws can be applied to a system of point objects.

1.3 Newton's first law of motion

Newton's first law of motion essentially states that a point object subject to zero net external force moves in a straight line with a constant speed (i.e., it does not accelerate). However, this is true only in special frames of reference called *inertial frames*. Indeed, we can think of Newton's first law as the definition of an inertial frame: an inertial frame of reference is one in which a point object subject to zero net external force moves in a straight line with constant speed.

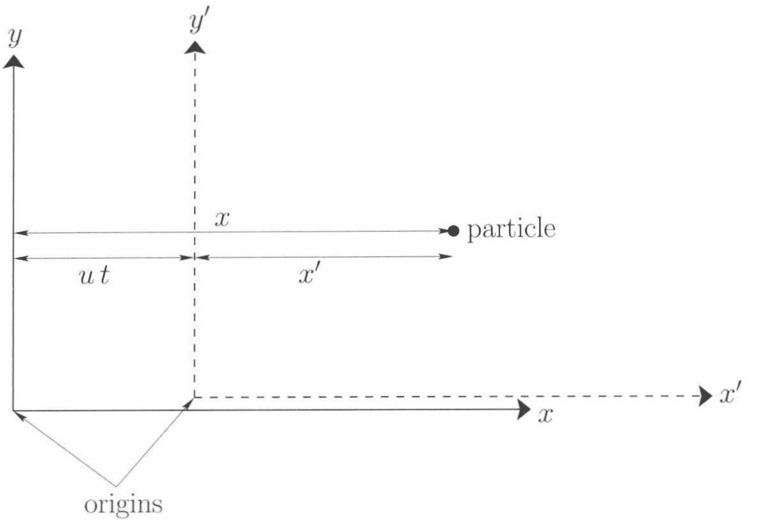

Fig. 1.1 A Galilean coordinate transformation.

Suppose that we have found an inertial frame of reference. Let us set up a Cartesian coordinate system in this frame. The motion of a point object can now be specified by giving its position vector, $\mathbf{r} \equiv (x, y, z)$, with respect to the origin of the coordinate system, as a function of time, t. Consider a second frame of reference moving with some constant velocity $\mathbf{u}$ with respect to the first frame. Without loss of generality, we can suppose that the Cartesian axes in the second frame are parallel to the corresponding axes in the first frame, that $\mathbf{u} \equiv (u, 0, 0)$, and, finally, that the origins of the two frames instantaneously coincide at $t = 0$. (See Figure 1.1.) Suppose that the position vector of our point object is $\mathbf{r}' \equiv (x', y', z')$ in the second frame of reference. It is evident, from Figure 1.1, that at any given time, t, the coordinates of the object in the two reference frames satisfy

$$x' = x - u\,t, \tag{1.1}$$

$$y' = y, \tag{1.2}$$

and

$$z' = z. \tag{1.3}$$

This coordinate transformation was first discovered by Galileo Galilei (1564–1642), and is nowadays known as a *Galilean transformation* in his honor.

By definition, the instantaneous velocity of the object in our first reference frame is given by $\mathbf{v} = d\mathbf{r}/dt \equiv (dx/dt, dy/dt, dz/dt)$, with an analogous expression for the velocity, $\mathbf{v}'$, in the second frame. It follows, from differentiation of Equations (1.1)–(1.3) with respect to time, that the velocity components in the two frames satisfy

$$v_x' = v_x - u, \tag{1.4}$$

$$v_y' = v_y, \tag{1.5}$$

and

$$v'_z = v_z. \tag{1.6}$$

These equations can be written more succinctly as

$$\mathbf{v}' = \mathbf{v} - \mathbf{u}. \tag{1.7}$$

Finally, by definition, the instantaneous acceleration of the object in our first reference frame is given by $\mathbf{a} = d\mathbf{v}/dt \equiv (dv_x/dt,\, dv_y/dt,\, dv_z/dt)$, with an analogous expression for the acceleration, $\mathbf{a}'$, in the second frame. It follows, from differentiation of Equations (1.4)–(1.6) with respect to time, that the acceleration components in the two frames satisfy

$$a'_x = a_x, \tag{1.8}$$

$$a'_y = a_y, \tag{1.9}$$

and

$$a'_z = a_z. \tag{1.10}$$

These equations can be written more succinctly as

$$\mathbf{a}' = \mathbf{a}. \tag{1.11}$$

According to Equations (1.7) and (1.11), if an object is moving in a straight line with a constant speed in our original inertial frame (i.e., if $\mathbf{a} = \mathbf{0}$), then it also moves in a (different) straight line with a (different) constant speed in the second frame of reference (i.e., $\mathbf{a}' = \mathbf{0}$). Hence, we conclude that the second frame of reference is also an inertial frame.

A simple extension of the preceding argument allows us to conclude that there is an *infinite* number of different inertial frames moving with constant velocities with respect to one another. Newton thought that one of these inertial frames was special and defined an absolute standard of rest: that is, a static object in this frame was in a state of absolute rest. However, Einstein showed that this is not the case. In fact, there is no absolute standard of rest: in other words, all motion is relative—hence, the name *relativity* for Einstein's theory. Consequently, one inertial frame is just as good as another as far as Newtonian mechanics is concerned.

But what happens if the second frame of reference *accelerates* with respect to the first? In this case, it is not hard to see that Equation (1.11) generalizes to

$$\mathbf{a}' = \mathbf{a} - \frac{d\mathbf{u}}{dt}, \tag{1.12}$$

where $\mathbf{u}(t)$ is the instantaneous velocity of the second frame with respect to the first. According to this formula, if an object is moving in a straight line with a constant speed in the first frame (i.e., if $\mathbf{a} = \mathbf{0}$), then it does not move in a straight line with a constant speed in the second frame (i.e., $\mathbf{a}' \neq \mathbf{0}$). Hence, if the first frame is an inertial frame, then the second is *not*.

A simple extension of the preceding argument allows us to conclude that any frame of reference that accelerates with respect to a given inertial frame is not itself an inertial frame.

For most practical purposes, when studying the motions of objects close to the Earth's surface, a reference frame that is fixed with respect to this surface is approximately inertial. However, if the trajectory of a projectile within such a frame is measured to high precision, then it will be found to deviate slightly from the predictions of Newtonian mechanics. (See Chapter 5.) This deviation is due to the fact that the Earth is rotating, and its surface is therefore *accelerating* toward its axis of rotation. When studying the motions of objects in orbit around the Earth, a reference frame whose origin is the center of the Earth (or, to be more exact, the center of mass of the Earth–Moon system), and whose coordinate axes are fixed with respect to distant stars, is approximately inertial. However, if such orbits are measured to extremely high precision, then they will again be found to deviate very slightly from the predictions of Newtonian mechanics. In this case, the deviation is due to the Earth's orbital motion about the Sun. When studying the orbits of the planets in the solar system, a reference frame whose origin is the center of the Sun (or, to be more exact, the center of mass of the solar system), and whose coordinate axes are fixed with respect to distant stars, is approximately inertial. In this case, any deviations of the orbits from the predictions of Newtonian mechanics due to the orbital motion of the Sun about the galactic center are far too small to be measurable. It should be noted that it is impossible to identify an *absolute* inertial frame—the best approximation to such a frame would be one in which the cosmic microwave background appears to be (approximately) isotropic. However, for a given dynamic problem, it is always possible to identify an *approximate* inertial frame. Furthermore, any deviations of such a frame from a true inertial frame can be incorporated into the framework of Newtonian mechanics via the introduction of so-called fictitious forces. (See Chapter 5.)

1.4 Newton's second law of motion

Newton's second law of motion essentially states that if a point object is subject to an external force, $\mathbf{f}$, then its equation of motion is given by

$$\frac{d\mathbf{p}}{dt} = \mathbf{f}, \tag{1.13}$$

where the momentum, $\mathbf{p}$, is the product of the object's inertial mass, m, and its velocity, $\mathbf{v}$. If m is not a function of time, then Equation (1.13) reduces to the familiar equation

$$m\,\frac{d\mathbf{v}}{dt} = \mathbf{f}. \tag{1.14}$$

This equation is valid only in an *inertial frame*. Clearly, the inertial mass of an object measures its reluctance to deviate from its preferred state of uniform motion in a straight line (in an inertial frame). Of course, the preceding equation of motion can be solved only if we have an independent expression for the force, $\mathbf{f}$ (i.e., a law of force). Let us suppose that this is the case.

An important corollary of Newton's second law is that force is a *vector quantity*. This must be the case, as the law equates force to the product of a scalar (mass) and a vector (acceleration).[2] Note that acceleration is obviously a vector because it is directly related to displacement, which is the prototype of all vectors. One consequence of force being a vector is that two forces, $\mathbf{f}_1$ and $\mathbf{f}_2$, both acting at a given point, have the same effect as a single force, $\mathbf{f} = \mathbf{f}_1 + \mathbf{f}_2$, acting at the same point, where the summation is performed according to the laws of vector addition. Likewise, a single force, $\mathbf{f}$, acting at a given point, has the same effect as two forces, $\mathbf{f}_1$ and $\mathbf{f}_2$, acting at the same point, provided that $\mathbf{f}_1 + \mathbf{f}_2 = \mathbf{f}$. This method of combining and splitting forces is known as the *resolution of forces*; it lies at the heart of many calculations in Newtonian mechanics.

Taking the scalar product of Equation (1.14) with the velocity, $\mathbf{v}$, we obtain

$$m\,\mathbf{v}\cdot\frac{d\mathbf{v}}{dt} = \frac{m}{2}\frac{d(\mathbf{v}\cdot\mathbf{v})}{dt} = \frac{m}{2}\frac{dv^2}{dt} = \mathbf{f}\cdot\mathbf{v}. \tag{1.15}$$

This can be written

$$\frac{dK}{dt} = \mathbf{f}\cdot\mathbf{v}, \tag{1.16}$$

where

$$K = \frac{1}{2}\,m\,v^2. \tag{1.17}$$

The right-hand side of Equation (1.16) represents the rate at which the force does work on the object—that is, the rate at which the force transfers energy to the object. The quantity K represents the energy that the object possesses by virtue of its motion. This type of energy is generally known as *kinetic energy*. Thus, Equation (1.16) states that any work done on a point object by an external force goes to increase the object's kinetic energy.

Suppose that under the action of the force, $\mathbf{f}$, our object moves from point P at time t_1 to point Q at time t_2. The net change in the object's kinetic energy is obtained by integrating Equation (1.16):

$$\Delta K = \int_{t_1}^{t_2} \mathbf{f}\cdot\mathbf{v}\,dt = \int_P^Q \mathbf{f}\cdot d\mathbf{r}, \tag{1.18}$$

because $\mathbf{v} = d\mathbf{r}/dt$. Here, $d\mathbf{r}$ is an element of the object's path between points P and Q, and the integral in $\mathbf{r}$ represents the net *work* done by the force as the object moves along the path from P to Q.

As is well known, there are basically two kinds of forces in nature: first, those for which line integrals of the type $\int_P^Q \mathbf{f}\cdot d\mathbf{r}$ depend on the end points but not on the path taken between these points; second, those for which line integrals of the type $\int_P^Q \mathbf{f}\cdot d\mathbf{r}$ depend both on the end points and the path taken between these points. The first kind of force is termed *conservative*, whereas the second kind is termed *non-conservative*. It can be demonstrated that if the line integral $\int_P^Q \mathbf{f}\cdot d\mathbf{r}$ is *path independent*, for all choices of P and Q, then the force $\mathbf{f}$ can be written as the gradient of a scalar field. (See Section A.5.)

[2] A *scalar* is a physical quantity that is invariant under rotation of the coordinate axes. A *vector* is a physical quantity that transforms in an analogous manner to a displacement under rotation of the coordinate axes.

In other words, all conservative forces satisfy

$$\mathbf{f}(\mathbf{r}) = -\nabla U \tag{1.19}$$

for some scalar field $U(\mathbf{r})$. [Incidentally, mathematicians, as opposed to physicists and astronomers, usually write $f(\mathbf{r}) = +\nabla U$.] Note that

$$\int_P^Q \nabla U \cdot d\mathbf{r} \equiv \Delta U = U(Q) - U(P), \tag{1.20}$$

irrespective of the path taken between P and Q. Hence, it follows from Equation (1.18) that

$$\Delta K = -\Delta U \tag{1.21}$$

for conservative forces. Another way of writing this is

$$E = K + U = \text{constant}. \tag{1.22}$$

Of course, we recognize Equation (1.22) as an *energy conservation equation*: E is the object's total energy, which is conserved; K is the energy the object has by virtue of its motion, otherwise known as its *kinetic energy*; and U is the energy the object has by virtue of its position, otherwise known as its *potential energy*. Note, however, that we can write energy conservation equations only for conservative forces. Gravity is an obvious example of such a force. Incidentally, potential energy is undefined to an arbitrary additive constant. In fact, it is only the *difference* in potential energy between different points in space that is well defined.

1.5 Newton's third law of motion

Consider a system of N mutually interacting point objects. Let the ith object, whose mass is m_i, be located at position vector $\mathbf{r}_i$. Suppose that this object exerts a force $\mathbf{f}_{ji}$ on the jth object. Likewise, suppose that the jth object exerts a force $\mathbf{f}_{ij}$ on the ith object. Newton's third law of motion essentially states that these two forces are equal and opposite, irrespective of their nature. In other words,

$$\mathbf{f}_{ij} = -\mathbf{f}_{ji}. \tag{1.23}$$

(See Figure 1.2.) One corollary of Newton's third law is that an object cannot exert a force on itself. Another corollary is that all forces in the universe have corresponding reactions. The only exceptions to this rule are the fictitious forces that arise in non-inertial reference frames (e.g., the centrifugal and Coriolis forces that appear in rotating reference frames—see Chapter 5). Fictitious forces do not generally possess reactions.

Newton's third law implies *action at a distance*. In other words, if the force that object i exerts on object j suddenly changes, then Newton's third law demands that there must be an *immediate* change in the force that object j exerts on object i. Moreover, this must be true irrespective of the distance between the two objects. However, we now know that Einstein's special theory of relativity forbids information from traveling through