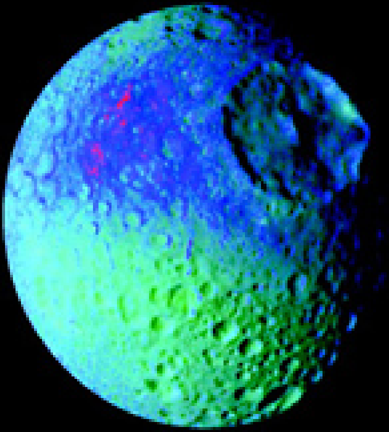
CLASSICAL MECHANICS



R. DOUGLAS GREGORY

CAMBRIDGE

CAMBRIDGE www.cambridge.org/9780521826785

CLASSICAL MECHANICS

Gregory's *Classical Mechanics* is a major new textbook for undergraduates in mathematics and physics. It is a thorough, self-contained and highly readable account of a subject many students find difficult. The author's clear and systematic style promotes a good understanding of the subject: each concept is motivated and illustrated by worked examples, while problem sets provide plenty of practice for understanding and technique. Computer assisted problems, some suitable for projects, are also included. The book is structured to make learning the subject easy; there is a natural progression from core topics to more advanced ones and hard topics are treated with particular care. A theme of the book is the importance of conservation principles. These appear first in vectorial mechanics where they are proved and applied to problem solving. They reappear in analytical mechanics, where they are shown to be related to symmetries of the Lagrangian, culminating in Noether's theorem.

• Suitable for a wide range of undergraduate mechanics courses given in mathematics and physics departments: no prior knowledge of the subject is assumed

• Profusely illustrated and thoroughly class-tested, with a clear direct style that makes the subject easy to understand: all concepts are motivated and illustrated by the many worked examples included

• Good, accurately-set problems, with answers in the book: computer assisted problems and projects are also provided. Model solutions for problems available to teachers from `www.cambridge.org/Gregory`

The author

Douglas Gregory is Professor of Mathematics at the University of Manchester. He is a researcher of international standing in the field of elasticity, and has held visiting positions at New York University, the University of British Columbia, and the University of Washington. He is highly regarded as a teacher of applied mathematics: this, his first book, is the product of many years of teaching experience.

Bloody instructions, which, being taught,
Return to plague th' inventor.

SHAKESPEARE, *Macbeth*, act I, sc. 7

Front Cover The photograph on the front cover shows Mimas, one of the many moons of Saturn; the huge crater was formed by an impact. Mimas takes 22 hours 37 minutes to orbit Saturn, the radius of its orbit being 185,500 kilometers. After reading Chapter 7, you will be able to estimate the mass of Saturn from this data!

CLASSICAL MECHANICS

AN UNDERGRADUATE TEXT

R. DOUGLAS GREGORY

University of Manchester

CAMBRIDGE
UNIVERSITY PRESS

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

Cambridge University Press
The Edinburgh Building, Cambridge CB2 2RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521826785

First published in print format 2006

ISBN-13 978-0-511-16097-4 eBook (EBL)
ISBN-10 0-511-16097-6 eBook (EBL)

ISBN-13 978-0-521-82678-5 hardback
ISBN-10 0-521-82678-0 hardback

ISBN-13 978-0-521-53409-3
ISBN-10 0-521-53409-7

Contents

Preface

Information for readers

What is this book about and who is it for?

This is a book on **classical mechanics** for **university undergraduates**. It aims to cover all the material normally taught in classical mechanics courses from Newton's laws to Hamilton's equations. If you are attending such a course, you will be unlucky not to find the course material in this book.

What prerequisites are needed to read this book?

It is expected that the reader will have attended an elementary **calculus** course and an elementary course on **differential equations** (ODEs). A previous course in mechanics is helpful but not essential. *This book is self-contained in the sense that it starts from the beginning and assumes no prior knowledge of mechanics*. However, in a general text such as this, the early material is presented at a brisker pace than in books that are specifically aimed at the beginner.

What is the style of the book?

The book is written in a crisp, no nonsense style; in short, there is no waffle! The object is to get the reader to the important points as quickly and easily as possible, consistent with good understanding.

Are there plenty of examples with full solutions?

Yes there are. Every new concept and technique is reinforced by **fully worked examples**. The author's advice is that the reader should think how he or she would do each worked example *before* reading the solution; much more will be learned this way!

Are there plenty of problems with answers?

Yes there are. At the end of each chapter there is a large collection of problems. For convenience, these are arranged by topic and trickier problems are marked with a star. **Answers are provided to all of the problems**. A feature of the book is the inclusion of computer assisted problems. These are interesting physical problems that cannot be solved analytically, but can be solved easily with computer assistance.

Where can I find more information?

More information about this book can be found on the book's homepage

`http://www.cambridge.org/Gregory`

All feedback from readers is welcomed. Please e-mail your comments, corrections and good ideas by clicking on the comments button on the book's homepage.

Information for lecturers

Scope of the book and prerequisites

This book aims to cover all the material normally taught in undergraduate mechanics courses from Newton's laws to Hamilton's equations. It assumes that the students have attended an elementary calculus course and an elementary course on ODEs, but no more. The book is self contained and, in principle, it is not essential that the students should have studied mechanics before. However, their lives will be made easier if they have!

Inspection copy and Solutions Manual

Any lecturer who is giving an undergraduate course on classical mechanics can request an **inspection copy** of this book. Simply go to the book's homepage

`http://www.cambridge.org/Gregory`

and follow the links.

Lecturers who adopt this book for their course may receive the **Solutions Manual**. This has a **complete set of detailed solutions** to the problems at the end of the chapters. To obtain the Solutions Manual, just send an e-mail giving your name, affiliation, and details of the course to `solutions@cambridge.org`

Feedback

All feedback from instructors and lecturers is welcomed. Please e-mail your comments via the link on the book's homepage

Acknowledgements

I am very grateful to many friends and colleagues for their helpful comments and suggestions while this book was in preparation. But most of all I thank my wife Win for her unstinting support and encouragement, without which the book could not have been written at all.

Part One

NEWTONIAN MECHANICS OF A SINGLE PARTICLE

CHAPTERS IN PART ONE

Chapter One

The algebra and calculus of vectors

KEY FEATURES

The key features of this chapter are the **rules of vector algebra** and **differentiation of vector functions** of a scalar variable.

This chapter begins with a review of the rules and applications of **vector algebra**. Almost every student taking a mechanics course will already have attended a course on vector algebra, and so, instead of covering the subject in full detail, we present, for easy reference, a summary of vector operations and their important properties, together with a selection of worked examples.

The chapter closes with an account of the **differentiation of vector functions** of a scalar variable. Unlike the vector algebra sections, this is treated in full detail. Applications include the **tangent vector** and **normal vector** to a curve. These will be needed in the next chapter in order to interpret the velocity and acceleration vectors.

1.1 VECTORS AND VECTOR QUANTITIES

Most physical quantities can be classified as being **scalar quantities** or **vector quantities**. The temperature in a room is an example of a scalar quantity. It is so called because its *value* is a scalar, which, in the present context, means a *real number*. Other examples of scalar quantities are the volume of a can, the density of iron, and the pressure of air in a tyre. Vector quantities are defined as follows:

Definition 1.1 ***Vector quantity*** *If a quantity Q has a* ***magnitude*** *and a* ***direction*** *associated with it, then Q is said to be a* ***vector quantity***. [Here, magnitude means a positive real number and direction is specified relative to some underlying reference frame* that we regard as fixed.]

The **displacement** of a particle† is an example of a vector quantity. Suppose the particle starts from the point A and, after moving in a general manner, ends up at the

* See section 2.2 for an explanation of the term 'reference frame'.
† A particle is an idealised body that occupies only a single point of space.

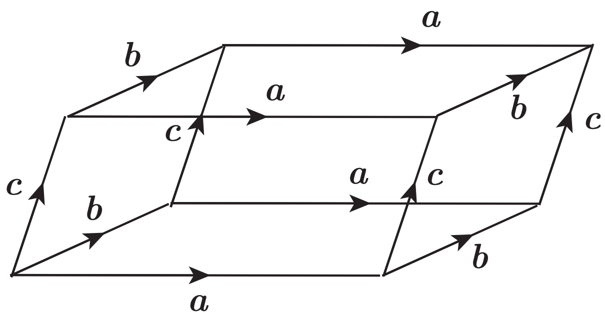


FIGURE 1.1 Four different representations of each of the vectors ***a***, ***b*** ***c*** form the twelve edges of the parallelopiped box.

point B. The *magnitude* of the displacement is the distance AB and the *direction* of the displacement is the direction of the straight line joining A to B (in that order). Another example is the **force** applied to a body by a rope. In this case, the *magnitude* is the strength of the force (a real positive quantity) and the *direction* is the direction of the rope (away from the body). Other examples of vector quantities are the velocity of a body and the value of the electric (or magnetic) field. In order to manipulate all such quantities without regard to their physical origin, we introduce the concept of a vector as an *abstract quantity*.

Definition 1.2 ***Vector*** *A* ***vector*** *is an* ***abstract*** *quantity characterised by the two properties* ***magnitude*** *and* ***direction****. Thus two vectors are equal if they have the same magnitude and the same direction.**

Notation. Vectors are written in bold type, for example ***a***, ***b***, ***r*** or ***F***. The **magnitude** of the vector ***a***, which is a real positive number, is written $|\boldsymbol{a}|$, or sometimes† simply a.

It is convenient to define operations involving abstract vectors by reference to some simple, easily visualised vector quantity. The standard choice is the set of directed **line segments**. Each straight line joining two points (P and Q say, in that order) is a vector quantity, where the magnitude is the distance PQ and the direction is the direction of Q relative to P. We call this the line segment $\overrightarrow{PQ}$ and we say that it *represents* some abstract vector ***a***.‡ Note that each vector ***a*** is represented by infinitely many different line segments, as indicated in Figure 1.1.

* In order that our set of vectors should have a standard algebra, we also include a special vector whose magnitude is zero and whose direction is not defined. This is called the **zero vector** and written **0**. The zero vector is not the same thing as the number zero!

† It is often useful to denote the magnitudes of the vectors ***a***, ***b***, ***c***, … by a, b, c, …, but this does risk confusion. Take care!

‡ The zero vector is represented by line segments whose end point and starting point are coincident.

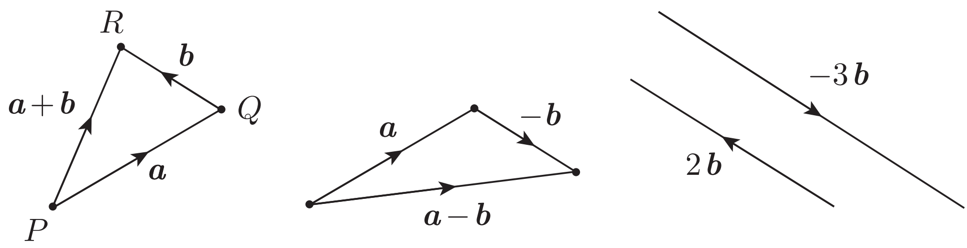


FIGURE 1.2 Addition, subtraction and scalar multiplication of vectors.

1.2 LINEAR OPERATIONS: $\boldsymbol{a}+\boldsymbol{b}$ AND $\lambda\boldsymbol{a}$

Since vectors are abstract quantities, we can define sums and products of vectors in any way we like. However, in order to be of any use, the definitions must create some coherent algebra and represent something of interest when applied to a range of vector quantities. Also, our definitions must be independent of the particular representations used to construct them. The definitions that follow satisfy all these requirements.

The vector sum $\boldsymbol{a}+\boldsymbol{b}$

Definition 1.3 ***Sum of vectors*** *Let* $\boldsymbol{a}$ *and* $\boldsymbol{b}$ *be any two vectors. Take any representation* $\overrightarrow{PQ}$ *of* $\boldsymbol{a}$ *and suppose the line segment* $\overrightarrow{QR}$ *represents* $\boldsymbol{b}$. *Then the* ***sum*** $\boldsymbol{a}+\boldsymbol{b}$ *of* $\boldsymbol{a}$ *and* $\boldsymbol{b}$ *is the* ***vector*** *represented by the line segment* $\overrightarrow{PR}$, *as shown in Figure 1.2 (left).*

Laws of algebra for the vector sum

(i)	$\boldsymbol{b}+\boldsymbol{a}=\boldsymbol{a}+\boldsymbol{b}$	(commutative law)
(ii)	$\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})=(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}$	(associative law)

Definition 1.4 ***Negative of a vector*** *Let* $\boldsymbol{b}$ *be any vector. Then the vector with the same magnitude as* $\boldsymbol{b}$ *and the* ***opposite*** *direction is called the* ***negative*** *of* $\boldsymbol{b}$ *and is written* $-\boldsymbol{b}$. ***Subtraction*** *by* $\boldsymbol{b}$ *is then defined by*

$$\boldsymbol{a}-\boldsymbol{b}=\boldsymbol{a}+(-\boldsymbol{b}).$$

[That is, to subtract $\boldsymbol{b}$ just add $-\boldsymbol{b}$, as shown in Figure 1.2 (centre).]

The scalar multiple $\lambda\boldsymbol{a}$

Definition 1.5 ***Scalar multiple*** *Let* $\boldsymbol{a}$ *be a vector and* λ *be a scalar (a real number). Then the* ***scalar multiple*** $\lambda\boldsymbol{a}$ *is the vector whose magnitude is* $|\lambda||\boldsymbol{a}|$ *and whose direction is*

(i) *the same as* $\boldsymbol{a}$ *if* λ *is positive,*
(ii) *undefined if* λ *is zero (the answer is the zero vector),*
(iii) *the same as* $-\boldsymbol{a}$ *if* λ *is negative.*

It follows that $-(\lambda\boldsymbol{a}) = (-\lambda)\boldsymbol{a}$.

Laws of algebra for the scalar multiple

(i) $\lambda(\mu\boldsymbol{a}) = (\lambda\mu)\boldsymbol{a}$ (associative law)

(ii) $\lambda(\boldsymbol{a}+\boldsymbol{b}) = \lambda\boldsymbol{a}+\lambda\boldsymbol{b}$ and $(\lambda+\mu)\boldsymbol{a} = \lambda\boldsymbol{a}+\mu\boldsymbol{a}$ (distributive laws)

The effect of the above laws is that **linear combinations** of vectors can be manipulated just *as if* the vectors were symbols representing real or complex numbers.

Example 1.1 *Laws for vector sum and scalar multiple*

Simplify the expression $3(2\boldsymbol{a}-4\boldsymbol{b})-2(2\boldsymbol{a}-\boldsymbol{b})$.

Solution

On this one occasion we will do the simplification by strict application of the laws. It is instructive to decide which laws are being used at each step!

$$\begin{aligned}
3(2\boldsymbol{a}-4\boldsymbol{b})-2(2\boldsymbol{a}-\boldsymbol{b}) &= 3\Big(2\boldsymbol{a}+(-4)\boldsymbol{b}\Big)+(-2)\Big(2\boldsymbol{a}+(-1)\boldsymbol{b}\Big)\\
&= \Big(6\boldsymbol{a}+(-12)\boldsymbol{b}\Big)+\Big((-4)\boldsymbol{a}+2\boldsymbol{b}\Big)\\
&= \Big(6\boldsymbol{a}+(-4)\boldsymbol{a}\Big)+\Big((-12)\boldsymbol{b}+2\boldsymbol{b}\Big)\\
&= 2\boldsymbol{a}+(-10)\boldsymbol{b} = 2\boldsymbol{a}-10\boldsymbol{b}. \blacksquare
\end{aligned}$$

Unit vectors

A vector of **unit magnitude** is called a **unit vector**. If any vector $\boldsymbol{a}$ is divided by its own magnitude, the result is a *unit vector* having the *same direction* as $\boldsymbol{a}$. This new vector is denoted by $\widehat{\boldsymbol{a}}$ so that

$$\widehat{\boldsymbol{a}} = \boldsymbol{a}/|\boldsymbol{a}|.$$

Basis sets

Suppose $\boldsymbol{a}$ and $\boldsymbol{b}$ are two non-zero vectors, with the direction of $\boldsymbol{b}$ neither the same nor opposite to that of $\boldsymbol{a}$. Let $\overrightarrow{OA}$, $\overrightarrow{OB}$ be representations of $\boldsymbol{a}$, $\boldsymbol{b}$ and let $\mathcal{P}$ be the plane

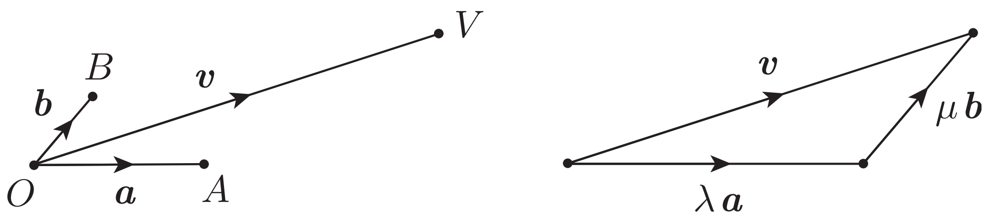


FIGURE 1.3 The set $\{\boldsymbol{a}, \boldsymbol{b}\}$ is a basis for all vectors lying in the plane OAB.

containing the triangle OAB. Then (see Figure 1.3) any vector $\boldsymbol{v}$ whose representation $\overrightarrow{OV}$ lies in the plane $\mathcal{P}$ can be written in the form

$$\boldsymbol{v} = \lambda \boldsymbol{a} + \mu \boldsymbol{b}, \tag{1.1}$$

where the coefficients λ, μ are unique. Vectors that have their directions parallel to the same plane are said to be **coplanar**. Thus we have shown that *any vector coplanar with* $\boldsymbol{a}$ *and* $\boldsymbol{b}$ *can be expanded uniquely in the form* (1.1). It is also apparent that this expansion set cannot be reduced in number (in this case to a single vector). For these reasons the pair of vectors $\{\boldsymbol{a}, \boldsymbol{b}\}$ is said to be a **basis set** for vectors lying* in the plane $\mathcal{P}$.

Suppose now that $\{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\}$ is a set of three non-coplanar vectors. Then any vector $\boldsymbol{v}$, *without restriction*, can be written in the form

$$\boldsymbol{v} = \lambda \boldsymbol{a} + \mu \boldsymbol{b} + \nu \boldsymbol{c}, \tag{1.2}$$

where the coefficients λ, μ, ν are unique. In this case we say that the set $\{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\}$ is a **basis set** for all three-dimensional vectors. Although *any* set of three non-coplanar vectors forms a basis, it is most convenient to take the basis vectors to be *orthogonal unit vectors*. In this case the basis set† is usually denoted by $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ and is said to be an **orthonormal basis**. The representation of a general vector $\boldsymbol{v}$ in the form

$$\boldsymbol{v} = \lambda \boldsymbol{i} + \mu \boldsymbol{j} + \nu \boldsymbol{k}$$

is common in problem solving.

In applications involving the cross product of vectors, the distinction between **right-** and **left-handed** basis sets actually matters. There is no experiment in classical mechanics or eletromagnetism that can distinguish between right- and left-handed sets. The difference can only be exhibited by a model or some familiar object that exhibits 'handedness', such as a corkscrew.‡ Figure 1.4 shows a **right-handed orthonormal basis set** attached to a well known object.

---

* Strictly speaking vectors are abstract quantities that do not *lie* anywhere. This phrase should be taken to mean '*vectors whose directions are parallel to the plane* $\mathcal{P}$'.

† It should be remembered that there are *infinitely* many basis sets made up of orthogonal unit vectors.

‡ Suppose that the non-coplanar vectors $\{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\}$ have representations $\overrightarrow{OA}$, $\overrightarrow{OB}$, $\overrightarrow{OC}$ respectively. Place an ordinary corkscrew with the screw lying along the line through O perpendicular to the plane OAB,

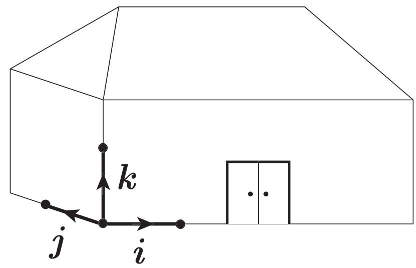


FIGURE 1.4 A **standard basis** set $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ is both **orthonormal** and **right-handed**.

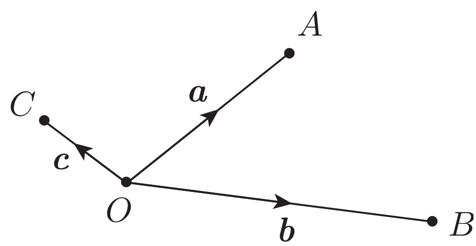


FIGURE 1.5 The points A, B, C have position vectors $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ relative to the origin O.

Definition 1.6 ***Standard basis set*** *If an* ***orthonormal basis*** $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ *is also* ***right-handed*** *(as shown in Figure 1.4), we will call it a* ***standard basis****.*

Position vectors and vector geometry

Suppose that O is a fixed point of space. Then relative to the origin O (and relative to the underlying reference frame), any point of space, such as A, has an associated line segment, $\overrightarrow{OA}$, which represents some vector $\boldsymbol{a}$. Conversely, the vector $\boldsymbol{a}$ is sufficient to specify the position of the point A.

Definition 1.7 ***Position vector*** *The vector* $\boldsymbol{a}$ *is called the* ***position vector*** *of the point A relative to the origin O, [It is standard practice, and very convenient, to denote the position vectors of the points $A, B, C, \dots$ by* $\boldsymbol{a}$*,* $\boldsymbol{b}$*,* $\boldsymbol{c}$*, and so on, as shown in Figure 1.5.]*

Since vectors can be used to specify the positions of points in space, we can now use the laws of vector algebra to prove* results in Euclidean geometry. This is not just an academic exercise. Familiarity with geometrical concepts is an important part of mechanics. We begin with the following useful result:

---

and the handle parallel to OA. Now turn the corkscrew until the handle is parallel to OB and note the direction in which the corkscrew *would* move if it were 'in action'. (The direction of the turn must be such that the angle turned through is at most 180°.) If OC makes an *acute angle* with this direction, the set $\{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\}$ (in that order) is *right-handed*; if OC makes an *obtuse* angle with this direction then the set is *left-handed*.

* Some properties of Euclidean geometry have been used to prove the laws of vector algebra. However, this does not prevent us from giving valid proofs of other results.

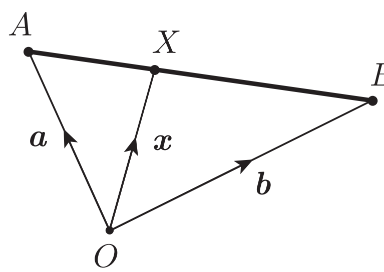


FIGURE 1.6 The point X divides the line AB in the ratio $\lambda : \mu$.

Example 1.2 *Point dividing a line in a given ratio*

The points A and B have position vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ relative to an origin O. Find the position vector $\boldsymbol{x}$ of the point X that divides the line AB in the ratio $\lambda : \mu$ (that is $AX/XB = \lambda/\mu$).

Solution

It follows from Figure 1.6 that $\boldsymbol{x}$ is given by*

$$\begin{aligned} \boldsymbol{x} &= \boldsymbol{a} + \overrightarrow{AX} = \boldsymbol{a} + \left(\frac{\lambda}{\lambda+\mu}\right)\overrightarrow{AB} \\ &= \boldsymbol{a} + \left(\frac{\lambda}{\lambda+\mu}\right)(\boldsymbol{b}-\boldsymbol{a}) = \frac{\mu\boldsymbol{a}+\lambda\boldsymbol{b}}{\lambda+\mu}. \end{aligned}$$

In particular, the **mid-point** of the line AB has position vector $\frac{1}{2}(\boldsymbol{a}+\boldsymbol{b})$. ■

Example 1.3 *Centroid of a triangle*

Show that the three medians of any triangle meet in a point (the centroid) which divides each of them in the ratio 2:1.

Solution

Let the triangle be ABC where the points A, B, C have position vectors $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ relative to some origin O. Then the mid-point P of the side BC has position vector $\boldsymbol{p} = \frac{1}{2}(\boldsymbol{b}+\boldsymbol{c})$. The point X that divides the median AP in the ratio 2:1 therefore has position vector

$$\boldsymbol{x} = \frac{\boldsymbol{a}+2\boldsymbol{p}}{2+1} = \frac{\boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}}{3}.$$

The position vectors of the corresponding points on the other two medians can be found by cyclic permutation of the vectors $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ and clearly give the same value. Hence all three points are coincident and so the three medians meet there. ■

---

* Strictly speaking we should not write expressions like $\boldsymbol{a} + \overrightarrow{AX}$ since the sum we defined was the sum of two *vectors*, not a vector and a line segment. What we really mean is '*the sum of* $\boldsymbol{a}$ *and the vector represented by the line segment* $\overrightarrow{AX}$'. Pure mathematicians would not approve but this notation is so convenient we will use it anyway. It's all part of living dangerously!

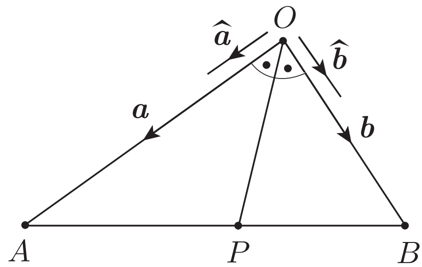


FIGURE 1.7 The bisector theorem: $AP/PB = OA/OB$.

Example 1.4 *The bisector theorem*

In a triangle OAB, the bisector of the angle $A\widehat{O}B$ meets the line AB at the point P. Show that $AP/PB = OA/OB$.

Solution

Let the vertex O be the origin of vectors* and let the position vectors of the vertices A, B relative to O be $\boldsymbol{a}$, $\boldsymbol{b}$ as shown in Figure 1.7. The point with position vector $\boldsymbol{a} + \boldsymbol{b}$ does *not* lie the bisector OP in general since the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ have different magnitudes a and b. However, by symmetry, the point with position vector $\widehat{\boldsymbol{a}} + \widehat{\boldsymbol{b}}$ *does* lie on the bisector and a general point X on the bisector has a position vector $\boldsymbol{x}$ of the form

$$\boldsymbol{x} = \lambda\left(\widehat{\boldsymbol{a}} + \widehat{\boldsymbol{b}}\right) = \lambda\left(\frac{\boldsymbol{a}}{a} + \frac{\boldsymbol{b}}{b}\right) = \lambda\left(\frac{b\,\boldsymbol{a} + a\,\boldsymbol{b}}{ab}\right) = \left(\frac{b\,\boldsymbol{a} + a\,\boldsymbol{b}}{K}\right),$$

where $K = ab/\lambda$ is a new constant. Now X will lie on the line AB if its position vector has the form $(\mu\boldsymbol{a} + \lambda\boldsymbol{b})/(\lambda + \mu)$, that is, if $K = a + b$. Hence the position vector $\boldsymbol{p}$ of P is

$$\boldsymbol{p} = \frac{b\,\boldsymbol{a} + a\,\boldsymbol{b}}{a + b}.$$

Moreover we see that P divides that line AB in the ratio $a : b$, that is, $AP/PB = OA/OB$ as required. ■

1.3 THE SCALAR PRODUCT $\boldsymbol{a} \cdot \boldsymbol{b}$

Definition 1.8 ***Scalar product*** *Suppose the vectors* **a** *and* **b** *have representations* $\overrightarrow{OA}$ *and* $\overrightarrow{OB}$. *Then the* **scalar product a · b** *of* **a** *and* **b** *is defined by*

$$\boldsymbol{a} \cdot \boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}|\cos\theta, \tag{1.3}$$

where θ *is the angle between* OA *and* OB. [Note that $\boldsymbol{a} \cdot \boldsymbol{b}$ is a **scalar** quantity.]

* One can always take a special point of the figure as origin. The penalty is that the symmetry of the labelling is lost.

Laws of algebra for the scalar product

(i) $\boldsymbol{b} \cdot \boldsymbol{a} = \boldsymbol{a} \cdot \boldsymbol{b}$ (commutative law)

(ii) $\boldsymbol{a} \cdot (\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{a} \cdot \boldsymbol{c}$ (distributive law)

(iii) $(\lambda \boldsymbol{a}) \cdot \boldsymbol{b} = \lambda(\boldsymbol{a} \cdot \boldsymbol{b})$ (associative with scalar multiplication)

Properties of the scalar product

(i) $\boldsymbol{a} \cdot \boldsymbol{a} = |\boldsymbol{a}|^2$.

(ii) The scalar product $\boldsymbol{a} \cdot \boldsymbol{b} = 0$ if (and only if) $\boldsymbol{a}$ and $\boldsymbol{b}$ are perpendicular (or one of them is zero).

(iii) If $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ is an orthonormal basis then

$$\boldsymbol{i} \cdot \boldsymbol{i} = \boldsymbol{j} \cdot \boldsymbol{j} = \boldsymbol{k} \cdot \boldsymbol{k} = 1, \qquad \boldsymbol{i} \cdot \boldsymbol{j} = \boldsymbol{j} \cdot \boldsymbol{k} = \boldsymbol{k} \cdot \boldsymbol{i} = 0.$$

(iv) If $\boldsymbol{a}_1 = \lambda_1 \boldsymbol{i} + \mu_1 \boldsymbol{j} + \nu_1 \boldsymbol{k}$ and $\boldsymbol{a}_2 = \lambda_2 \boldsymbol{i} + \mu_2 \boldsymbol{j} + \nu_2 \boldsymbol{k}$ then

$$\boldsymbol{a}_1 \cdot \boldsymbol{a}_2 = \lambda_1 \lambda_2 + \mu_1 \mu_2 + \nu_1 \nu_2.$$

Example 1.5 *Numerical example on the scalar product*

If $\boldsymbol{a} = 2\boldsymbol{i} - \boldsymbol{j} + 2\boldsymbol{k}$ and $\boldsymbol{b} = 4\boldsymbol{i} - 3\boldsymbol{k}$, find the magnitudes of $\boldsymbol{a}$ and $\boldsymbol{b}$ and the angle between them.

Solution

$|\boldsymbol{a}|^2 = \boldsymbol{a} \cdot \boldsymbol{a} = (2\boldsymbol{i} - \boldsymbol{j} + 2\boldsymbol{k}) \cdot (2\boldsymbol{i} - \boldsymbol{j} + 2\boldsymbol{k}) = 2^2 + (-1)^2 + 2^2 = 9$. Hence $|\boldsymbol{a}| = 3$. Similarly $|\boldsymbol{b}|^2 = 4^2 + 0^2 + (-3)^2 = 25$ so that $|\boldsymbol{b}| = 5$. Also $\boldsymbol{a} \cdot \boldsymbol{b} = 8 + 0 + (-6) = 2$. Since $\boldsymbol{a} \cdot \boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}| \cos\theta$, it follows that $2 = 3 \times 5 \times \cos\theta$ so that $\cos\theta = 2/15$.

Hence the magnitudes of $\boldsymbol{a}$ and $\boldsymbol{b}$ are 3 and 5, and the angle between them is $\cos^{-1}(2/15)$. ■

Example 1.6 *Apollonius's theorem*

In the triangle OAB, M is the mid-point of AB. Show that $(OA)^2 + (OB)^2 = 2(OM)^2 + 2(AM)^2$.

Solution

Let the vertex O be the origin of vectors and let the position vectors of A and B be $\boldsymbol{a}$ and $\boldsymbol{b}$. Then the position vector of M is $\frac{1}{2}(\boldsymbol{a} + \boldsymbol{b})$. Then

$$\begin{aligned} 4(OM)^2 &= |\boldsymbol{a} + \boldsymbol{b}|^2 = (\boldsymbol{a} + \boldsymbol{b}) \cdot (\boldsymbol{a} + \boldsymbol{b}) \\ &= \boldsymbol{a} \cdot \boldsymbol{a} + \boldsymbol{b} \cdot \boldsymbol{b} + 2\boldsymbol{a} \cdot \boldsymbol{b} = |\boldsymbol{a}|^2 + |\boldsymbol{b}|^2 + 2\boldsymbol{a} \cdot \boldsymbol{b} \end{aligned}$$

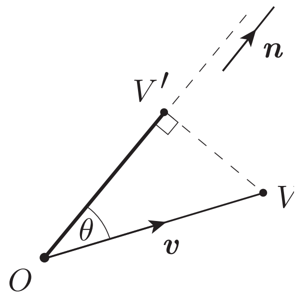


FIGURE 1.8 The component of $\boldsymbol{v}$ in the direction of the unit vector $\boldsymbol{n}$ is equal to OV', the **projection** of OV onto the line through O parallel to $\boldsymbol{n}$.

and

$$\begin{aligned} 4(AM)^2 &= (AB)^2 = |\boldsymbol{a}-\boldsymbol{b}|^2 = (\boldsymbol{a}-\boldsymbol{b})\cdot(\boldsymbol{a}-\boldsymbol{b}) \\ &= \boldsymbol{a}\cdot\boldsymbol{a}+\boldsymbol{b}\cdot\boldsymbol{b}-2\boldsymbol{a}\cdot\boldsymbol{b} = |\boldsymbol{a}|^2+|\boldsymbol{b}|^2-2\boldsymbol{a}\cdot\boldsymbol{b}. \end{aligned}$$

Hence

$$2(OM)^2 + 2(AM)^2 = |\boldsymbol{a}|^2 + |\boldsymbol{b}|^2 = (OA)^2 + (OB)^2$$

as required. ■

Components of a vector

Definition 1.9 ***Components of a vector*** *Let* $\boldsymbol{n}$ *be a unit vector. Then the* ***component*** *of the vector* $\boldsymbol{v}$ *in the direction of* $\boldsymbol{n}$ *is defined to be* $\boldsymbol{v}\cdot\boldsymbol{n}$. *The component of* $\boldsymbol{v}$ *in the direction of a general vector* $\boldsymbol{a}$ *is therefore* $\boldsymbol{v}\cdot\widehat{\boldsymbol{a}}$.

Properties of components

(i) The component $\boldsymbol{v}\cdot\boldsymbol{n}$ has a simple geometrical significance. Let $\overrightarrow{OV}$ be a representation of $\boldsymbol{v}$ as shown in Figure 1.8. Then

$$\boldsymbol{v}\cdot\boldsymbol{n} = |\boldsymbol{v}||\boldsymbol{n}|\cos\theta = OV\cos\theta = OV',$$

where OV' is the **projection** of OV onto the line through O parallel to $\boldsymbol{n}$.

(ii) Suppose that $\boldsymbol{v}$ is a sum of vectors, $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2 + \boldsymbol{v}_3$ say. Then the component of $\boldsymbol{v}$ in the direction of $\boldsymbol{n}$ is

$$\boldsymbol{v}\cdot\boldsymbol{n} = (\boldsymbol{v}_1+\boldsymbol{v}_2+\boldsymbol{v}_3)\cdot\boldsymbol{n} = (\boldsymbol{v}_1\cdot\boldsymbol{n}) + (\boldsymbol{v}_2\cdot\boldsymbol{n}) + (\boldsymbol{v}_3\cdot\boldsymbol{n}),$$

by the distributive law for the scalar product. Thus, the *component of the sum of a number of vectors in a given direction is equal to the sum of the components of the individual vectors* in that direction.

(iii) If a vector $\boldsymbol{v}$ is expanded in terms of a general basis set $\{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\}$ in the form $\boldsymbol{v} = \lambda\boldsymbol{a}+\mu\boldsymbol{b}+\nu\boldsymbol{c}$, the coefficients λ, μ, ν are *not* the components of the vector $\boldsymbol{v}$ in the

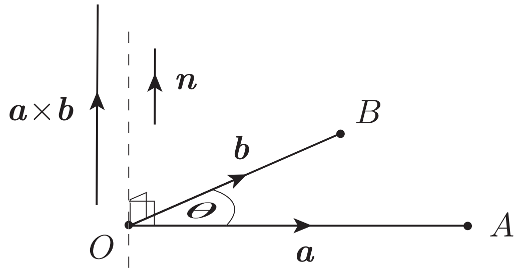


FIGURE 1.9 The vector product $\boldsymbol{a}\times\boldsymbol{b} = (|\boldsymbol{a}||\boldsymbol{b}|\sin\theta)\,\boldsymbol{n}$.

directions of $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$. However if $\boldsymbol{v}$ is expanded in terms of an *orthonormal* basis set $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ in the form $\boldsymbol{v} = \lambda\,\boldsymbol{i} + \mu\,\boldsymbol{j} + \nu\,\boldsymbol{k}$, then the component of $\boldsymbol{v}$ in the $\boldsymbol{i}$-direction is

$$\begin{aligned}\boldsymbol{v}\cdot\boldsymbol{i} &= (\lambda\boldsymbol{i} + \mu\,\boldsymbol{j} + \nu\boldsymbol{k})\cdot\boldsymbol{i} = \lambda(\boldsymbol{i}\cdot\boldsymbol{i}) + \mu(\boldsymbol{j}\cdot\boldsymbol{i}) + \nu(\boldsymbol{k}\cdot\boldsymbol{i})\\ &= \lambda + 0 + 0 = \lambda.\end{aligned}$$

Similarly μ and ν are the components of $\boldsymbol{v}$ in the $\boldsymbol{j}$- and $\boldsymbol{k}$-directions. Hence *when a vector $\boldsymbol{v}$ is expanded in terms of an orthonormal basis set $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ in the form $\boldsymbol{v} = \lambda\,\boldsymbol{i} + \mu\,\boldsymbol{j} + \nu\,\boldsymbol{k}$, the coefficients λ, μ, ν are the components of $\boldsymbol{v}$ in the $\boldsymbol{i}$- $\boldsymbol{j}$- and $\boldsymbol{k}$-directions.*

Example 1.7 ***Numerical example on components***

If $\boldsymbol{v} = 6\boldsymbol{i} - 3\,\boldsymbol{j} + 15\boldsymbol{k}$ and $\boldsymbol{a} = 2\boldsymbol{i} - \boldsymbol{j} - 2\boldsymbol{k}$, find the component of $\boldsymbol{v}$ in the direction of $\boldsymbol{a}$.

Solution

$|\boldsymbol{a}|^2 = \boldsymbol{a}\cdot\boldsymbol{a} = 2^2 + (-1)^2 + (-2)^2 = 9$. Hence $|\boldsymbol{a}| = 3$ and

$$\widehat{\boldsymbol{a}} = \frac{\boldsymbol{a}}{|\boldsymbol{a}|} = \frac{2\boldsymbol{i} - \boldsymbol{j} - 2\boldsymbol{k}}{3}.$$

The required component of $\boldsymbol{v}$ is therefore

$$\boldsymbol{v}\cdot\widehat{\boldsymbol{a}} = (6\boldsymbol{i} - 3\,\boldsymbol{j} + 15\boldsymbol{k})\cdot\left(\frac{2\boldsymbol{i} - \boldsymbol{j} - 2\boldsymbol{k}}{3}\right) = \frac{12 + 3 - 30}{3} = -5. \blacksquare$$

1.4 THE VECTOR PRODUCT $a \times b$

Definition 1.10 ***Vector product*** *Suppose the vectors **a** and **b** have representations $\overrightarrow{OA}$ and $\overrightarrow{OB}$ and let **n** be the unit vector perpendicular to the plane OAB and such that $\{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{n}\}$ is a **right-handed** set. Then the **vector product $\boldsymbol{a}\times\boldsymbol{b}$** of **a** and **b** is defined by*

$$\boldsymbol{a}\times\boldsymbol{b} = (|\boldsymbol{a}||\boldsymbol{b}|\sin\theta)\,\boldsymbol{n}, \tag{1.4}$$

where θ $(0 \le \theta \le 180°)$ *is the angle between* OA *and* OB. [Note that $\boldsymbol{a}\times\boldsymbol{b}$ is a **vector** quantity.]

Laws of algebra for the vector product

(i) $\boldsymbol{b}\times\boldsymbol{a} = -\boldsymbol{a}\times\boldsymbol{b}$ (**anti**-commutative law)

(ii) $\boldsymbol{a}\times(\boldsymbol{b}+\boldsymbol{c}) = \boldsymbol{a}\times\boldsymbol{b}+\boldsymbol{a}\times\boldsymbol{c}$ (distributive law)

(iii) $(\lambda\boldsymbol{a})\times\boldsymbol{b} = \lambda(\boldsymbol{a}\times\boldsymbol{b})$ (associative with scalar multiplication)

Since the vector product is anti-commutative, the *order of the terms in vector products must be preserved*. The vector product is not associative.

Properties of the vector product

(i) $\boldsymbol{a}\times\boldsymbol{a} = \boldsymbol{0}$.

(ii) The vector product $\boldsymbol{a}\times\boldsymbol{b} = \boldsymbol{0}$ if (and only if) $\boldsymbol{a}$ and $\boldsymbol{b}$ are parallel (or one of them is zero).

(iii) If $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ is a standard basis then

$$\boldsymbol{i}\times\boldsymbol{j} = \boldsymbol{k}, \quad \boldsymbol{k}\times\boldsymbol{i} = \boldsymbol{j}, \quad \boldsymbol{j}\times\boldsymbol{k} = \boldsymbol{i}, \qquad \boldsymbol{i}\times\boldsymbol{i} = \boldsymbol{j}\times\boldsymbol{j} = \boldsymbol{k}\times\boldsymbol{k} = \boldsymbol{0}.$$

(iv) If $\boldsymbol{a}_1 = \lambda_1\boldsymbol{i} + \mu_1\boldsymbol{j} + \nu_1\boldsymbol{k}$ and $\boldsymbol{a}_2 = \lambda_2\boldsymbol{i} + \mu_2\boldsymbol{j} + \nu_2\boldsymbol{k}$ then

$$\boldsymbol{a}_1\times\boldsymbol{a}_2 = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \lambda_1 & \mu_1 & \nu_1 \\ \lambda_2 & \mu_2 & \nu_2 \end{vmatrix}$$

where the determinant is to be evaluated by the first row.

Example 1.8 ***Numerical example on vector product***

If $\boldsymbol{a} = 2\boldsymbol{i} - \boldsymbol{j} + 2\boldsymbol{k}$ and $\boldsymbol{b} = -\boldsymbol{i} - 3\boldsymbol{k}$, find a unit vector perpendicular to both $\boldsymbol{a}$ and $\boldsymbol{b}$.

Solution

The vector $\boldsymbol{a}\times\boldsymbol{b}$ is perpendicular to both $\boldsymbol{a}$ and $\boldsymbol{b}$. Now

$$\begin{aligned} \boldsymbol{a}\times\boldsymbol{b} &= \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 2 & -1 & 2 \\ -1 & 0 & -3 \end{vmatrix} \\ &= (3-0)\,\boldsymbol{i} - ((-6)-(-2))\,\boldsymbol{j} + (0-1)\,\boldsymbol{k} \\ &= 3\,\boldsymbol{i} + 4\,\boldsymbol{j} - \boldsymbol{k}. \end{aligned}$$

The magnitude of this vector is $\left(3^2 + 4^2 + (-1)^2\right)^{1/2} = (26)^{1/2}$. Hence the required unit vector can be either of $\pm\,(3\,\boldsymbol{i} + 4\,\boldsymbol{j} - \boldsymbol{k})\,/(26)^{1/2}$. ■

1.5 TRIPLE PRODUCTS

Triple products are not new operations but are simply one product followed by another. There are two kinds of triple product whose values are scalar and vector respectively.

Triple scalar product

An expression of the form $\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c})$ is called a **triple scalar product**; its value is a scalar.

Properties of the triple scalar product

(i)

$$\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}) = \boldsymbol{c} \cdot (\boldsymbol{a} \times \boldsymbol{b}) = \boldsymbol{b} \cdot (\boldsymbol{c} \times \boldsymbol{a}), \tag{1.5}$$

that is, *cyclic permutation of the vectors* $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ *in a triple scalar product leaves its value unchanged*. [Interchanging two vectors reverses the sign.]
This formula can alternatively be written

$$\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c}, \tag{1.6}$$

that is, *interchanging the positions of the 'dot' and the 'cross' in a triple scalar product leaves its value unchanged*.
Because of this symmetry, the triple scalar product can be denoted unambiguously by $[\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}]$.

(ii) The triple scalar product $[\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}] = 0$ if (and only if) $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ are **coplanar** (or one of them is zero). In particular *a triple scalar product is zero if two of its vectors are the same*.

(iii) If $[\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}] > 0$ then the set $\{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\}$ is *right-handed*. If $[\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}] < 0$ then the set $\{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\}$ is *left-handed*.

(iv) If $\boldsymbol{a}_1 = \lambda_1 \boldsymbol{i} + \mu_1 \boldsymbol{j} + \nu_1 \boldsymbol{k}$, $\boldsymbol{a}_2 = \lambda_2 \boldsymbol{i} + \mu_2 \boldsymbol{j} + \nu_2 \boldsymbol{k}$, $\boldsymbol{a}_3 = \lambda_3 \boldsymbol{i} + \mu_3 \boldsymbol{j} + \nu_3 \boldsymbol{k}$, where $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ is a standard basis, then

$$[\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3] = \begin{vmatrix} \lambda_1 & \mu_1 & \nu_1 \\ \lambda_2 & \mu_2 & \nu_2 \\ \lambda_3 & \mu_3 & \nu_3 \end{vmatrix}. \tag{1.7}$$

Triple vector product

An expression of the form $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c})$ is called a **triple vector product**; its value is a vector.

Property of the triple vector product

Since $\boldsymbol{b} \times \boldsymbol{c}$ is perpendicular to both $\boldsymbol{b}$ and $\boldsymbol{c}$, it follows that $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c})$ must lie in the same plane as $\boldsymbol{b}$ and $\boldsymbol{c}$. It can therefore be expanded in the form $\lambda \boldsymbol{a} + \mu \boldsymbol{b}$. The actual formula is

$$\boldsymbol{a}\times(\boldsymbol{b}\times\boldsymbol{c}) = (\boldsymbol{a}\cdot\boldsymbol{c})\,\boldsymbol{b} - (\boldsymbol{a}\cdot\boldsymbol{b})\,\boldsymbol{c}. \tag{1.8}$$

Since the vector product is anti-commutative and non-associative, it is wise to use this formula *exactly* as it stands.

Example 1.9 ***Using triple products***

Expand the expression $(\boldsymbol{a}\times\boldsymbol{b})\cdot(\boldsymbol{c}\times\boldsymbol{d})$ in terms of scalar products.

Solution

Use the triple scalar product formula (1.6) to interchange the first ‘dot’ and ‘cross’, and then expand the resulting triple vector product by the formula (1.8), as follows:

$$\begin{aligned}(\boldsymbol{a}\times\boldsymbol{b})\cdot(\boldsymbol{c}\times\boldsymbol{d}) &= \boldsymbol{a}\cdot[\,\boldsymbol{b}\times(\boldsymbol{c}\times\boldsymbol{d})\,] = \boldsymbol{a}\cdot[\,(\boldsymbol{b}\cdot\boldsymbol{d})\,\boldsymbol{c} - (\boldsymbol{b}\cdot\boldsymbol{c})\,\boldsymbol{d}\,]\\ &= (\boldsymbol{a}\cdot\boldsymbol{c})(\boldsymbol{b}\cdot\boldsymbol{d}) - (\boldsymbol{a}\cdot\boldsymbol{d})(\boldsymbol{b}\cdot\boldsymbol{c})\ \blacksquare\end{aligned}$$

1.6 VECTOR FUNCTIONS OF A SCALAR VARIABLE

In practice, the value of a vector quantity often depends on a scalar variable such as the time t. For example, if A is the label of a particle moving through space, then its position vector $\boldsymbol{a}$ (relative to a fixed origin O) will vary with time, that is, $\boldsymbol{a} = \boldsymbol{a}(t)$. The vector $\boldsymbol{a}$ is therefore a function of the scalar variable t.

The time dependence of a vector need not involve motion. The value of the electric or magnetic field at a *fixed point** of space will generally vary with time so that $\boldsymbol{E} = \boldsymbol{E}(t)$ and $\boldsymbol{B} = \boldsymbol{B}(t)$. More generally, the scalar variable need not be the time. Consider the space curve $\mathcal{C}$ shown in Figure 1.10, whose points are parametrised by the parameter α. Each point of the curve has a unique tangent line whose *direction* can be characterised by the unit vector $\boldsymbol{t}$. This is called the **unit tangent vector** to $\mathcal{C}$ and it depends on α, that is, $\boldsymbol{t} = \boldsymbol{t}(\alpha)$. In this case the independent variable is the scalar α and (just to confuse matters) the dependent variable is the vector $\boldsymbol{t}$.

Differentiation

The most important operation that can be carried out on a vector function of a scalar variable is **differentiation**.

Definition 1.11 ***Differentiation of vectors*** *Suppose that the vector $\boldsymbol{v}$ is a function of the scalar variable α, that is, $\boldsymbol{v} = \boldsymbol{v}(\alpha)$. Then the* ***derivative*** *of the function $\boldsymbol{v}(\alpha)$ with respect to α is defined by the limit*†

$$\frac{d\boldsymbol{v}}{d\alpha} = \lim_{\Delta\alpha\to 0}\left(\frac{\boldsymbol{v}(\alpha+\Delta\alpha) - \boldsymbol{v}(\alpha)}{\Delta\alpha}.\right) \tag{1.9}$$

* We will not be concerned here with vector functions of *position*. These are called vector *fields*.

† *Mathematical note*: The statement $\boldsymbol{u}(\alpha) \to \boldsymbol{U}$ as $\alpha \to A$ means that $|\,\boldsymbol{u}(\alpha) - \boldsymbol{U}\,| \to 0$ as $\alpha \to A$.

This looks identical to the definition of the derivative of an ordinary real function, but there is a difference. When α changes to $\alpha+\Delta\alpha$, the function $\boldsymbol{v}$ changes from $\boldsymbol{v}(\alpha)$ to $\boldsymbol{v}(\alpha+\Delta\alpha)$, a difference of $\boldsymbol{v}(\alpha+\Delta\alpha)-\boldsymbol{v}(\alpha)$. However, this 'difference' now means *vector* subtraction and its value is a vector; it remains a vector after dividing by the scalar increment $\Delta\alpha$. Hence $d\boldsymbol{v}/d\alpha$, the limit of this quotient as $\Delta\alpha \to 0$, is a **vector**. Furthermore, since $d\boldsymbol{v}/d\alpha$ depends on α, it is itself a vector function of the scalar variable α. The rules for differentiating combinations of vector functions are similar to those for ordinary scalar functions.

Differentiation rules for vector functions

Let $\boldsymbol{u}(\alpha)$ and $\boldsymbol{v}(\alpha)$ be vector functions of the scalar variable α, and let $\lambda(\alpha)$ be a scalar function. Then:

(i) $$\frac{d}{d\alpha}(\boldsymbol{u}+\boldsymbol{v}) = \dot{\boldsymbol{u}}+\dot{\boldsymbol{v}}$$

(ii) $$\frac{d}{d\alpha}(\lambda\,\boldsymbol{u}) = \dot{\lambda}\,\boldsymbol{u}+\lambda\dot{\boldsymbol{u}}$$

(iii) $$\frac{d}{d\alpha}(\boldsymbol{u}\cdot\boldsymbol{v}) = \dot{\boldsymbol{u}}\cdot\boldsymbol{v}+\boldsymbol{u}\cdot\dot{\boldsymbol{v}}$$

(iv) $$\frac{d}{d\alpha}(\boldsymbol{u}\times\boldsymbol{v}) = \dot{\boldsymbol{u}}\times\boldsymbol{v}+\boldsymbol{u}\times\dot{\boldsymbol{v}}$$

where $\dot{\boldsymbol{u}}$ means $d\boldsymbol{u}/d\alpha$ and so on. Note that the order of the terms in the vector product formula must be preserved.

Example 1.10 *Differentiating vector functions*

(i) The position vector of a particle P at time t is given by

$$\boldsymbol{r} = (2t^2-5t)\,\boldsymbol{i}+(4t+2)\,\boldsymbol{j}+t^3\,\boldsymbol{k},$$

where $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ is a *constant* basis set. Find $d\boldsymbol{r}/dt$ and $d^2\boldsymbol{r}/dt^2$. (These are the velocity and acceleration vectors of P at time t.)

(iii) If $\boldsymbol{a}=\boldsymbol{a}(t)$ and $\boldsymbol{b}$ is a constant vector, show that

$$\frac{d}{dt}[\,\boldsymbol{a}\cdot(\dot{\boldsymbol{a}}\times\boldsymbol{b})\,] = \boldsymbol{a}\cdot(\ddot{\boldsymbol{a}}\times\boldsymbol{b}).$$

Solution

(i) Since $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ are constant vectors, it follows from the differentiation rules that

$$\frac{d\boldsymbol{r}}{dt} = (4t-5)\,\boldsymbol{i}+4\,\boldsymbol{j}+3t^2\,\boldsymbol{k}, \qquad \frac{d^2\boldsymbol{r}}{dt^2} = 4\,\boldsymbol{i}+6t\,\boldsymbol{k}.$$

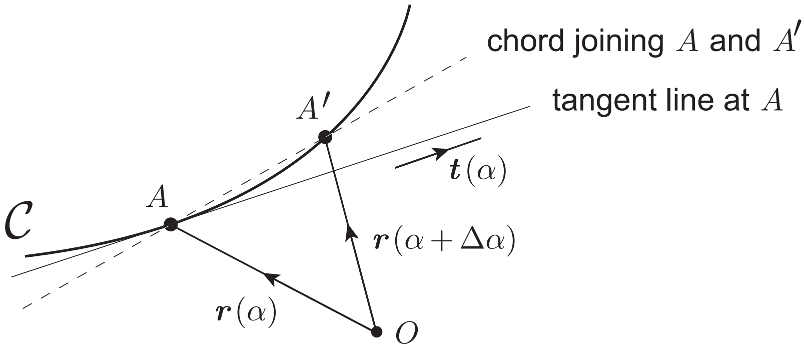


FIGURE 1.10 The unit tangent vector $\boldsymbol{t}(\alpha)$ at a typical point A on the curve $\mathcal{C}$, defined parametrically by $\boldsymbol{r} = \boldsymbol{r}(\alpha)$.

(ii)

$$\frac{d}{dt}[\boldsymbol{a}\cdot(\dot{\boldsymbol{a}}\times\boldsymbol{b})] = \dot{\boldsymbol{a}}\cdot\left(\dot{\boldsymbol{a}}\times\boldsymbol{b}\right) + \boldsymbol{a}\cdot\left(\frac{d}{dt}(\dot{\boldsymbol{a}}\times\boldsymbol{b})\right) = \boldsymbol{0} + \boldsymbol{a}\cdot\left(\ddot{\boldsymbol{a}}\times\boldsymbol{b} + \dot{\boldsymbol{a}}\times\dot{\boldsymbol{b}}\right)$$
$$= \boldsymbol{a}\cdot\left(\ddot{\boldsymbol{a}}\times\boldsymbol{b} + \dot{\boldsymbol{a}}\times\boldsymbol{0}\right) = \boldsymbol{a}\cdot(\ddot{\boldsymbol{a}}\times\boldsymbol{b}),$$

as required. ■

1.7 TANGENT AND NORMAL VECTORS TO A CURVE

In the next chapter we will define the velocity and acceleration of a particle moving in a space of three dimensions. In order to be able to interpret these definitions, we need to know a little about the differential geometry of curves. In particular, it is useful to know what the **unit tangent** and **unit normal** vectors of a curve are.

Unit tangent vector

Consider the curve $\mathcal{C}$ shown in Figure 1.10 which is defined by the parametric equation $\boldsymbol{r} = \boldsymbol{r}(\alpha)$. In general this can be a curve in *three-dimensional* space. Let A be a typical point of $\mathcal{C}$ corresponding to the parameter α and A' a nearby point corresponding to the parameter $\alpha + \Delta\alpha$. The chord $\overrightarrow{AA'}$ represents the vector

$$\Delta\boldsymbol{r} = \boldsymbol{r}(\alpha + \Delta\alpha) - \boldsymbol{r}(\alpha)$$

and so $\Delta\boldsymbol{r}/|\,\Delta\boldsymbol{r}\,|$ is a *unit* vector parallel to the chord $\overrightarrow{AA'}$. The **unit tangent vector** $\boldsymbol{t}(\alpha)$ at the point A is defined to be the *liixit* of this expression as $A' \to A$, that is

$$\boldsymbol{t}(\alpha) = \lim_{\Delta\alpha\to 0}\frac{\Delta\boldsymbol{r}}{|\,\Delta\boldsymbol{r}\,|}.$$

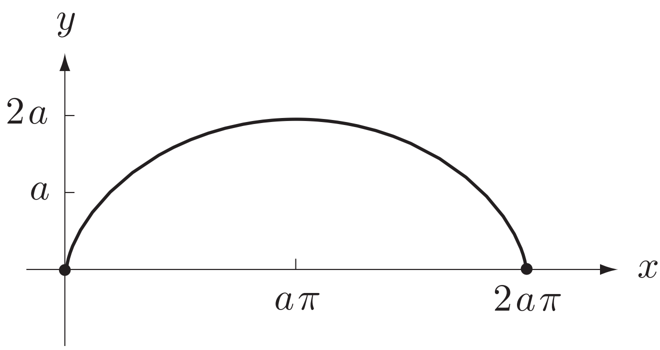


FIGURE 1.11 The cycloid $x = a(\theta - \sin\theta)$, $y = a(1 - \cos\theta)$, $z = 0$, where $0 < \theta < 2\pi$.

The tangent vector $\boldsymbol{t}$ is related to the derivative $d\boldsymbol{r}/d\alpha$ since

$$\frac{d\boldsymbol{r}}{d\alpha} = \lim_{\Delta\alpha\to 0}\frac{\Delta\boldsymbol{r}}{\Delta\alpha} = \lim_{\Delta\alpha\to 0}\frac{\Delta\boldsymbol{r}}{|\Delta\boldsymbol{r}|}\times\lim_{\Delta\alpha\to 0}\frac{|\Delta\boldsymbol{r}|}{\Delta\alpha}$$
$$= \boldsymbol{t}(\alpha)\times\left|\lim_{\Delta\alpha\to 0}\frac{\Delta\boldsymbol{r}}{\Delta\alpha}\right| = \boldsymbol{t}(\alpha)\times\left|\frac{d\boldsymbol{r}}{d\alpha}\right|,$$

that is,

$$\boxed{\frac{d\boldsymbol{r}}{d\alpha} = \left|\frac{d\boldsymbol{r}}{d\alpha}\right|\boldsymbol{t}(\alpha).} \tag{1.10}$$

Example 1.11 *Finding the unit tangent vector*

Figure 1.11 shows the cycloid $x = a(\theta - \sin\theta)$, $y = a(1 - \cos\theta)$, $z = 0$, where $0 < \theta < 2\pi$. Find the unit tangent vector to the cycloid at the point with parameter θ.

Solution

Let $\boldsymbol{i}$, $\boldsymbol{j}$ be unit vectors in the directions Ox, Oy respectively. Then the vector form of the equation for the cycloid is

$$\boldsymbol{r} = a(\theta - \sin\theta)\,\boldsymbol{i} + a(1 - \cos\theta)\,\boldsymbol{j}.$$

Then

$$\frac{d\boldsymbol{r}}{d\theta} = a(1 - \cos\theta)\,\boldsymbol{i} + (a\sin\theta)\,\boldsymbol{j}$$

and

$$\left|\frac{d\boldsymbol{r}}{d\theta}\right| = a\,(2 - 2\cos\theta)^{1/2} = 2a\sin\tfrac{1}{2}\theta.$$

Hence the **unit tangent vector** to the cycloid is

$$\boldsymbol{t}(\theta) = \frac{d\boldsymbol{r}}{d\theta}\bigg/\left|\frac{d\boldsymbol{r}}{d\theta}\right| = (\sin\tfrac{1}{2}\theta)\,\boldsymbol{i} + (\cos\tfrac{1}{2}\theta)\,\boldsymbol{j},$$

after simplification. ■

The formula (1.10) takes its simplest form when the parameter α is taken to be s, the *distance along the curve* measured from some fixed point. In this case,

$$\left|\frac{d\boldsymbol{r}}{ds}\right| = \lim_{\Delta s\to 0}\frac{|\Delta\boldsymbol{r}|}{\Delta s} = 1$$

so that $\boldsymbol{t}$ (pointing in the direction of increasing s) is given by the simple formula

$$\boxed{\boldsymbol{t} = \frac{d\boldsymbol{r}}{ds}.} \tag{1.11}$$

This is the most convenient formula for *theoretical* purposes.

Unit normal vector

Let $\boldsymbol{t}(s)$ be the unit tangent vector to the curve C, where the parameter s represents distance along the curve. Then, since $\boldsymbol{t}$ is a vector function of the scalar variable s, it has a derivative $d\boldsymbol{t}/ds$ which is another vector function of s.

Since $\boldsymbol{t}$ is a unit vector it follows that $\boldsymbol{t}(s)\cdot\boldsymbol{t}(s) = 1$ and if we differentiate this identity with respect to s, we obtain

$$\begin{aligned} 0 &= \frac{d}{ds}(\boldsymbol{t}\cdot\boldsymbol{t}) = \frac{d\boldsymbol{t}}{ds}\cdot\boldsymbol{t} + \boldsymbol{t}\cdot\frac{d\boldsymbol{t}}{ds} \\ &= 2\left(\frac{d\boldsymbol{t}}{ds}\cdot\boldsymbol{t}\right). \end{aligned}$$

It follows that $d\boldsymbol{t}/ds$ is always perpendicular to $\boldsymbol{t}$. It is usual to write $d\boldsymbol{t}/ds$ in the form

$$\boxed{\frac{d\boldsymbol{t}}{ds} = \kappa\,\boldsymbol{n}} \tag{1.12}$$

where $\kappa = |d\boldsymbol{t}/ds|$, a positive scalar called the **curvature**, and $\boldsymbol{n}$ is a *unit* vector called the (principal) **unit normal vector**. At each point of the curve, the unit vectors $\boldsymbol{t}(s)$ and $\boldsymbol{n}(s)$ are mutually perpendicular.

The quantities $\boldsymbol{n}$ and κ have a nice geometrical interpretation. Let A be any point on the curve and suppose that the distance parameter s is measured from A. Then, by Taylor's theorem, the form of the curve $\mathcal{C}$ near A is given approximately by

$$\boldsymbol{r}(s) = \boldsymbol{r}(0) + s\left[\frac{d\boldsymbol{r}}{ds}\right]_{s=0} + \tfrac{1}{2}s^2\left[\frac{d^2\boldsymbol{r}}{ds^2}\right]_{s=0} + O\left(s^3\right),$$

that is,

$$\boldsymbol{r}(s) = \boldsymbol{a} + s\,\boldsymbol{t} + \left(\tfrac{1}{2}\kappa s^2\right)\boldsymbol{n} + O\left(s^3\right), \tag{1.13}$$

where $\boldsymbol{a}$ is the position vector of the point A, and $\boldsymbol{t}$, κ and $\boldsymbol{n}$ are evaluated at the point A. Thus, near A, *the curve $\mathcal{C}$ lies* in the plane through A parallel to the vectors $\boldsymbol{t}$ and $\boldsymbol{n}$*. We can also see from equation (1.13) that, near A, the curve $\mathcal{C}$ is approximately a parabola. To the *same order of approximation*, it is equally true that, near A, the curve $\mathcal{C}$ is given by

$$\boldsymbol{r}(s) = \boldsymbol{a} + \kappa^{-1}(\sin\kappa s)\,\boldsymbol{t} + \kappa^{-1}(1 - \cos\kappa s)\,\boldsymbol{n} + O\left(s^3\right). \tag{1.14}$$

Thus, near A, the curve $\mathcal{C}$ is approximately a **circle** of radius κ^{-1}; the vector $\boldsymbol{t}$ is tangential to this circle and the vector $\boldsymbol{n}$ points towards its centre. The radius κ^{-1} is called the **radius of curvature** of $\mathcal{C}$ at the point A.

Example 1.12 *Finding the unit normal vector and curvature*

Find the unit normal vector and curvature of the cycloid $x = a(\theta - \sin\theta)$, $y = a(1 - \cos\theta)$, $z = 0$, where $0 < \theta < 2\pi$.

Solution

The **tangent vector** to the cycloid has already been found to be

$$\boldsymbol{t}(\theta) = \frac{d\boldsymbol{r}}{d\theta}\bigg/\left|\frac{d\boldsymbol{r}}{d\theta}\right| = (\sin\tfrac{1}{2}\theta)\,\boldsymbol{i} + (\cos\tfrac{1}{2}\theta)\,\boldsymbol{j}.$$

Hence, by the chain rule,

$$\begin{aligned}\frac{d\boldsymbol{t}}{ds} = \frac{d\boldsymbol{t}/d\theta}{ds/d\theta} = \frac{d\boldsymbol{t}/d\theta}{|d\boldsymbol{r}/d\theta|} &= \frac{\tfrac{1}{2}(\cos\tfrac{1}{2}\theta)\,\boldsymbol{i} - \tfrac{1}{2}(\sin\tfrac{1}{2}\theta)\,\boldsymbol{j}}{2a\sin\tfrac{1}{2}\theta}\\ &= \left(4a\sin\tfrac{1}{2}\theta\right)^{-1}\left((\cos\tfrac{1}{2}\theta)\,\boldsymbol{i} - (\sin\tfrac{1}{2}\theta)\,\boldsymbol{j}\right).\end{aligned}$$

Hence the **unit normal vector** and **curvature** of the cycloid are given by

$$\boldsymbol{n}(\theta) = (\cos\tfrac{1}{2}\theta)\,\boldsymbol{i} - (\sin\tfrac{1}{2}\theta)\,\boldsymbol{j}, \qquad \kappa(\theta) = \left(4a\sin\tfrac{1}{2}\theta\right)^{-1}.$$

The **radius of curvature** of the cycloid is therefore $4a\sin\frac{1}{2}\theta$. ■

* More precisely, this plane makes *three point contact* with the curve $\mathcal{C}$ at the point A.

Problems on Chapter 1

Answers and comments are at the end of the book.

Harder problems carry a star (∗).

1.1 In terms of the standard basis set $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$, $\boldsymbol{a} = 2\boldsymbol{i} - \boldsymbol{j} - 2\boldsymbol{k}$, $\boldsymbol{b} = 3\boldsymbol{i} - 4\boldsymbol{k}$ and $\boldsymbol{c} = \boldsymbol{i} - 5\boldsymbol{j} + 3\boldsymbol{k}$.

(i) Find $3\boldsymbol{a} + 2\boldsymbol{b} - 4\boldsymbol{c}$ and $|\boldsymbol{a} - \boldsymbol{b}|^2$.
(ii) Find $|\boldsymbol{a}|$, $|\boldsymbol{b}|$ and $\boldsymbol{a} \cdot \boldsymbol{b}$. Deduce the angle between $\boldsymbol{a}$ and $\boldsymbol{b}$.
(iii) Find the component of $\boldsymbol{c}$ in the direction of $\boldsymbol{a}$ and in the direction of $\boldsymbol{b}$.
(iv) Find $\boldsymbol{a} \times \boldsymbol{b}$, $\boldsymbol{b} \times \boldsymbol{c}$ and $(\boldsymbol{a} \times \boldsymbol{b}) \times (\boldsymbol{b} \times \boldsymbol{c})$.
(v) Find $\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c})$ and $(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c}$ and verify that they are equal. Is the set $\{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\}$ right- or left-handed?
(vi) By evaluating each side, verify the identity $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \boldsymbol{c})\,\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\,\boldsymbol{c}$.

Vector geometry

1.2 Find the angle between any two diagonals of a cube.

1.3 $ABCDEF$ is a regular hexagon with centre O which is also the origin of position vectors. Find the position vectors of the vertices C, D, E, F in terms of the position vectors $\boldsymbol{a}$, $\boldsymbol{b}$ of A and B.

1.4 Let $ABCD$ be a general (skew) quadrilateral and let P, Q, R, S be the mid-points of the sides AB, BC, CD, DA respectively. Show that $PQRS$ is a parallelogram.

1.5 In a general tetrahedron, lines are drawn connecting the mid-point of each side with the mid-point of the side opposite. Show that these three lines meet in a point that bisects each of them.

1.6 Let $ABCD$ be a general tetrahedron and let P, Q, R, S be the median centres of the faces opposite to the vertices A, B, C, D respectively. Show that the lines AP, BQ, CR, DS all meet in a point (called the *centroid* of the tetrahedron), which divides each line in the ratio 3:1.

1.7 A number of particles with masses $m_1, m_2, m_3, \dots$ are situated at the points with position vectors $\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3, \dots$ relative to an origin O. The centre of mass G of the particles is defined to be the point of space with position vector

$$\boldsymbol{R} = \frac{m_1\boldsymbol{r}_1 + m_2\boldsymbol{r}_2 + m_3\boldsymbol{r}_3 + \cdots}{m_1 + m_2 + m_3 + \cdots}$$

Show that if a different origin O' were used, this definition would still place G at the same point of space.

1.8 Prove that the three perpendiculars of a triangle are concurrent.

[Construct the two perpendiculars from A and B and take their intersection point as O, the origin of position vectors. Then prove that OC must be perpendicular to AB.]

Vector algebra

1.9 If $\boldsymbol{a}_1 = \lambda_1 \boldsymbol{i} + \mu_1 \boldsymbol{j} + \nu_1 \boldsymbol{k}$, $\boldsymbol{a}_2 = \lambda_2 \boldsymbol{i} + \mu_2 \boldsymbol{j} + \nu_2 \boldsymbol{k}$, $\boldsymbol{a}_3 = \lambda_3 \boldsymbol{i} + \mu_3 \boldsymbol{j} + \nu_3 \boldsymbol{k}$, where $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ is a standard basis, show that

$$\boldsymbol{a}_1 \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3) = \begin{vmatrix} \lambda_1 & \mu_1 & \nu_1 \\ \lambda_2 & \mu_2 & \nu_2 \\ \lambda_3 & \mu_3 & \nu_3 \end{vmatrix}.$$

Deduce that cyclic rotation of the vectors in a triple scalar product leaves the value of the product unchanged.

1.10 By expressing the vectors $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ in terms of a suitable standard basis, prove the identity $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \boldsymbol{c})\, \boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\, \boldsymbol{c}$.

1.11 Prove the identities

(i) $(\boldsymbol{a} \times \boldsymbol{b}) \cdot (\boldsymbol{c} \times \boldsymbol{d}) = (\boldsymbol{a} \cdot \boldsymbol{c})(\boldsymbol{b} \cdot \boldsymbol{d}) - (\boldsymbol{a} \cdot \boldsymbol{d})(\boldsymbol{b} \cdot \boldsymbol{c})$

(ii) $(\boldsymbol{a} \times \boldsymbol{b}) \times (\boldsymbol{c} \times \boldsymbol{d}) = [\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{d}]\, \boldsymbol{c} - [\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}]\, \boldsymbol{d}$

(iii) $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) + \boldsymbol{c} \times (\boldsymbol{a} \times \boldsymbol{b}) + \boldsymbol{b} \times (\boldsymbol{c} \times \boldsymbol{a}) = \boldsymbol{0}$ (Jacobi's identity)

1.12 ***Reciprocal basis*** Let $\{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\}$ be any basis set. Then the corresponding **reciprocal basis** $\{\boldsymbol{a}^*, \boldsymbol{b}^*, \boldsymbol{c}^*\}$ is defined by

$$\boldsymbol{a}^* = \frac{\boldsymbol{b} \times \boldsymbol{c}}{[\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}]}, \qquad \boldsymbol{b}^* = \frac{\boldsymbol{c} \times \boldsymbol{a}}{[\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}]}, \qquad \boldsymbol{c}^* = \frac{\boldsymbol{a} \times \boldsymbol{b}}{[\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}]}.$$

(i) If $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ is a standard basis, show that $\{\boldsymbol{i}^*, \boldsymbol{j}^*, \boldsymbol{k}^*\} = \{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$.

(ii) Show that $[\boldsymbol{a}^*, \boldsymbol{b}^*, \boldsymbol{c}^*] = 1/[\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}]$. Deduce that if $\{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\}$ is a right handed set then so is $\{\boldsymbol{a}^*, \boldsymbol{b}^*, \boldsymbol{c}^*\}$.

(iii) Show that $\{(\boldsymbol{a}^*)^*, (\boldsymbol{b}^*)^*, (\boldsymbol{c}^*)^*] = \{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\}$.

(iv) If a vector $\boldsymbol{v}$ is expanded in terms of the basis set $\{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\}$ in the form

$$\boldsymbol{v} = \lambda\, \boldsymbol{a} + \mu\, \boldsymbol{b} + \nu\, \boldsymbol{c},$$

show that the coefficients λ, μ, ν are given by $\lambda = \boldsymbol{v} \cdot \boldsymbol{a}^*$, $\mu = \boldsymbol{v} \cdot \boldsymbol{b}^*$, $\nu = \boldsymbol{v} \cdot \boldsymbol{c}^*$.

1.13 *Lamé's equations* The directions in which X-rays are strongly scattered by a crystal are determined from the solutions $\boldsymbol{x}$ of Lamé's equations, namely

$$\boldsymbol{x} \cdot \boldsymbol{a} = L, \qquad \boldsymbol{x} \cdot \boldsymbol{b} = M, \qquad \boldsymbol{x} \cdot \boldsymbol{c} = N,$$

where $\{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\}$ are the basis vectors of the crystal lattice, and L, M, N are *any* integers. Show that the solutions of Lamé's equations are

$$\boldsymbol{x} = L\, \boldsymbol{a}^* + M\, \boldsymbol{b}^* + N\, \boldsymbol{c}^*,$$

where $\{\boldsymbol{a}^*, \boldsymbol{b}^*, \boldsymbol{c}^*\}$ is the reciprocal basis to $\{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\}$.

Differentiation of vectors

1.14 If $\boldsymbol{r}(t) = (3t^2 - 4)\,\boldsymbol{i} + t^3\,\boldsymbol{j} + (t+3)\,\boldsymbol{k}$, where $\{\boldsymbol{i},\ \boldsymbol{j},\ \boldsymbol{k}\}$ is a constant standard basis, find $\dot{\boldsymbol{r}}$ and $\ddot{\boldsymbol{r}}$. Deduce the time derivative of $\boldsymbol{r}\times\dot{\boldsymbol{r}}$.

1.15 The vector $\boldsymbol{v}$ is a function of the time t and $\boldsymbol{k}$ is a constant vector. Find the time derivatives of (i) $|\,\boldsymbol{v}\,|^2$, (ii) $(\boldsymbol{v}\cdot\boldsymbol{k})\,\boldsymbol{v}$, (iii) $\left[\,\boldsymbol{v},\ \dot{\boldsymbol{v}},\ \boldsymbol{k}\,\right]$.

1.16 Find the unit tangent vector, the unit normal vector and the curvature of the circle $x = a\cos\theta$, $y = a\sin\theta$, $z = 0$ at the point with parameter θ.

1.17 Find the unit tangent vector, the unit normal vector and the curvature of the helix $x = a\cos\theta$, $y = a\sin\theta$, $z = b\theta$ at the point with parameter θ.

1.18 Find the unit tangent vector, the unit normal vector and the curvature of the parabola $x = ap^2$, $y = 2ap$, $z = 0$ at the point with parameter p.

Chapter Two

Velocity, acceleration and scalar angular velocity

KEY FEATURES

The key concepts in this chapter are the **velocity** and **acceleration** of a particle and the **angular velocity** of a rigid body in planar motion.

Kinematics is the study of the **motion of material bodies** without regard to the forces that cause their motion. The subject does not seek to answer the question of *why* bodies move as they do; that is the province of dynamics. It merely provides a geometrical description of the possible motions. The basic building block for bodies in mechanics is the **particle**, an idealised body that occupies only a single point of space. The important kinematical quantities in the motion of a particle are its **velocity** and **acceleration**. We begin with the simple case of straight line particle motion, where velocity and acceleration are scalars, and then progress to three-dimensional motion, where velocity and acceleration are vectors.

The other important idealisation that we consider is the **rigid body**, which we regard as a collection of particles linked by a light rigid framework. The important kinematical quantity in the motion of a rigid body is its **angular velocity**. In this chapter, we consider only those rigid body motions that are essentially two-dimensional, so that angular velocity is a scalar quantity. The general three-dimensional case is treated in Chapter 16.

2.1 STRAIGHT LINE MOTION OF A PARTICLE

Consider a particle P moving along the x-axis so that its displacement x from the origin O is a known function of the time t. Then the **mean velocity** of P over the time

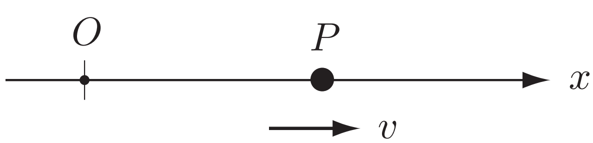


FIGURE 2.1 The particle P moves in a straight line and has displacement x and velocity v at time t.

interval $t_1 \le t \le t_2$ is defined to be the increase in the displacement of P divided by the time taken, that is,

$$\frac{x(t_2) - x(t_1)}{t_2 - t_1}. \tag{2.1}$$

Example 2.1 ***Mean velocity***

Suppose the displacement of P from O at time t is given by $x = t^2 - 6t$, where x is measured in metres and t in seconds. Find the mean velocity of P over the time interval $1 \le t \le 3$.

Solution

In this case, $x(1) = -5$ and $x(3) = -9$ so that the mean velocity of P is $((-9) - (-5))/(3 - 1) = -2\ \text{m s}^{-1}$. ■

The mean velocity of a particle is less important to us than its *instantaneous* velocity, that is, its velocity at a given instant in time. We cannot find the instantaneous velocity of P at time t_1 merely by letting $t_2 = t_1$ in the formula (2.1), since the quotient would then be undefined. However, we can define the instantaneous velocity as the *limit* of the mean velocity as the time interval *tends* to zero, that is, as $t_2 \to t_1$. Thus $v(t_1)$, the instantaneous velocity of P at time t_1 can be defined by

$$v(t_1) = \lim_{t_2 \to t_1} \left(\frac{x(t_2) - x(t_1)}{t_2 - t_1} \right).$$

But this is precisely the definition of dx/dt, the derivative of x with respect to t, evaluated at $t = t_1$. This leads us to the official definition:

Definition 2.1 ***1-D velocity*** *The (instantaneous)* ***velocity*** *v of P, in the positive x-direction, is defined by*

$$v = \frac{dx}{dt}. \tag{2.2}$$

The ***speed*** *of P is defined to be the rate of increase of the total distance travelled and is therefore equal to $|\,v\,|$.*

Similarly, the acceleration of P, the rate of increase of v, is defined as follows:

Definition 2.2 ***1-D acceleration*** *The (instantaneous)* ***acceleration*** *a of P, in the positive x-direction, is defined by*

$$a = \frac{dv}{dt} = \frac{d^2x}{dt^2}. \tag{2.3}$$

Example 2.2 ***Finding rectilinear velocity and acceleration***

Suppose the displacement of P from O at time t is given by $x = t^3 - 6t^2 + 4$, where x is measured in metres and t in seconds. Find the velocity and acceleration of P at time t. Deduce that P comes to rest twice and find the position and acceleration of P at the later of these two times.

Solution

Since $v = dx/dt$ and $a = dv/dt$, we obtain

$$v = 3t^2 - 12t \qquad \text{and} \qquad a = 6t - 12$$

as the velocity and acceleration of P at time t.

P comes to rest when its velocity v is zero, that is, when

$$3t^2 - 12t = 0.$$

This is a quadratic equation for t having the solutions $t = 0, 4$. Thus P is at rest when $t = 0$ s and $t = 4$ s.

When $t = 4$ s, $x = -28$ m and $a = 12\ \mathrm{m\,s^{-2}}$. Note that merely because $v = 0$ at some instant it does not follow that $a = 0$ also. ■

Example 2.3 *Reversing the process*

A particle P moves along the x-axis with its acceleration a at time t given by

$$a = 12t^2 - 6t + 6\ \mathrm{m\,s^{-2}}.$$

Initially P is at the point $x = 4$ m and is moving with speed $8\ \mathrm{m\,s^{-1}}$ in the negative x-direction. Find the velocity and displacement of P at time t.

Solution

Since $a = dv/dt$ we have

$$\frac{dv}{dt} = 12t^2 - 6t + 6,$$

and integrating with respect to t gives

$$v = 4t^3 - 3t^2 + 6t + C,$$

where C is a constant of integration. This constant can be determined by using the given initial condition on v, namely, $v = -8$ when $t = 0$. This gives $C = -8$ so that the velocity of P at time t is

$$v = 4t^3 - 3t^2 + 6t - 8\ \mathrm{m\,s^{-1}}.$$

By writing $v = dx/dt$ and integrating again, we obtain

$$x = t^4 - t^3 + 3t^2 - 8t + D,$$

where D is a second constant of integration. D can now be determined by using the given initial condition on x, namely, $x = 4$ when $t = 0$. This gives $D = 4$ so that the displacement of P at time t is

$$x = t^4 - t^3 + 3t^2 - 8t + 4\ \text{m}. ■$$

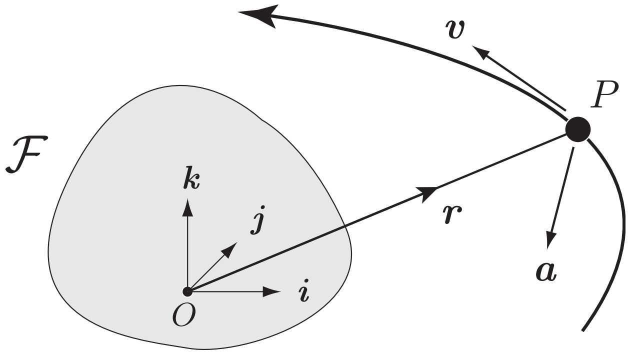


FIGURE 2.2 The particle P moves in three-dimensional space and, relative to the reference frame $\mathcal{F}$ and origin O, has position vector $\boldsymbol{r}$ at time t.

2.2 GENERAL MOTION OF A PARTICLE

When a particle P moves in two or three-dimensional space, its position can be described by its vector displacement $\boldsymbol{r}$ from an origin O that is fixed in a rigid **reference frame** $\mathcal{F}$. Whether $\mathcal{F}$ is moving or not is irrelevant here; the position vector $\boldsymbol{r}$ is simply measured *relative* to $\mathcal{F}$. Figure 2.2 shows a particle P moving in three-dimensional space with position vector $\boldsymbol{r}$ (relative to the reference frame $\mathcal{F}$) at time t.

Question *Reference frames*

What is a reference frame and why do we need one?

Answer

A rigid reference frame $\mathcal{F}$ is essentially a **rigid body** whose particles can be labelled to create reference points. The most familiar such body is the Earth. Relative to a single particle, the only thing that can be specified is distance from that particle. However, relative to a rigid body, one can specify both distance and direction. Thus the value of any vector quantity can be specified relative to $\mathcal{F}$. In particular, if we label some particle O of the body as origin, we can specify the position of any point of space by its position vector relative to the frame $\mathcal{F}$ and the origin O.

The specification of vectors relative to a reference frame is much simplified if we introduce a Cartesian coordinate system. This can be done in infinitely many different ways. Imagine that $\mathcal{F}$ is extended by a set of three mutually orthogonal planes that are *rigidly embedded* in it. The coordinates x, y, z of a point P are then the distances of P from these three planes. Let O be the origin of this coordinate system, and $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ its unit vectors. We can then conveniently refer to the frame $\mathcal{F}$, together with the embedded coordinate system $Oxyz$, by the notation $\mathcal{F}\{O\,;\,\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$. ■

In general motion, the velocity and acceleration of a particle are *vector quantities* and are defined by:

Definition 2.3 ***3-D velocity and acceleration*** *The* ***velocity*** $\boldsymbol{v}$ *and* ***acceleration*** $\boldsymbol{a}$ *of* P *are defined by*

$$\boldsymbol{v} = \frac{d\boldsymbol{r}}{dt} \quad \textit{and} \quad \boldsymbol{a} = \frac{d\boldsymbol{v}}{dt}. \tag{2.4}$$

Connection with the rectilinear case

The scalar velocity and acceleration defined in section 2.1 for the case of straight line motion are simply related to the corresponding vector quantities defined above. It would be possible to use the vector formalism in all cases but, for the case of straight line motion along the x-axis, $\boldsymbol{r}$, $\boldsymbol{v}$, and $\boldsymbol{a}$ would have the form

$$\boldsymbol{r} = x\,\boldsymbol{i}, \qquad \boldsymbol{v} = v\,\boldsymbol{i}, \qquad \boldsymbol{a} = a\,\boldsymbol{i},$$

where $v = dx/dt$ and $a = dv/dt$. It is therefore sufficient to work with the scalar quantities x, v and a; use of the vector formalism would be clumsy and unnecessary.

Example 2.4 *Finding 3-D velocity and acceleration*

Relative to the reference frame $\mathcal{F}\{O\,;\,\boldsymbol{i},\,\boldsymbol{j},\,\boldsymbol{k}\}$, the position vector of a particle P at time t is given by

$$\boldsymbol{r} = (2t^2 - 3)\,\boldsymbol{i} + (4t + 4)\,\boldsymbol{j} + (t^3 + 2t^2)\,\boldsymbol{k}.$$

Find (i) the distance OP when $t = 0$, (ii) the velocity of P when $t = 1$, (iii) the acceleration of P when $t = 2$.

Solution

In this solution we will make use of the rules for differentiation of sums and products involving vector functions of the time. These rules are listed in section 1.6.

(i) When $t = 0$, $\boldsymbol{r} = -3\,\boldsymbol{i} + 4\,\boldsymbol{j}$ so that $OP = |\boldsymbol{r}| = 5$.

(ii) Relative to the reference frame $\mathcal{F}$, the unit vectors $\{\boldsymbol{i},\,\boldsymbol{j},\,\boldsymbol{k}\}$ are *constant* and so their time derivatives are zero. The velocity $\boldsymbol{v}$ of P is therefore

$$\boldsymbol{v} = d\boldsymbol{r}/dt = 4t\,\boldsymbol{i} + 4\,\boldsymbol{j} + (3t^2 + 4t)\,\boldsymbol{k}.$$

When $t = 1$, $\boldsymbol{v} = 4\,\boldsymbol{i} + 4\,\boldsymbol{j} + 7\,\boldsymbol{k}$.

(iii) Relative to the reference frame $\mathcal{F}$, the acceleration $\boldsymbol{a}$ of P is

$$\boldsymbol{a} = d\boldsymbol{v}/dt = 4\,\boldsymbol{i} + (6t + 4)\,\boldsymbol{k}.$$

When $t = 2$, $\boldsymbol{a} = 4\,\boldsymbol{i} + 16\,\boldsymbol{k}$. ■

Interpretation of the vectors $\boldsymbol{v}$ and $\boldsymbol{a}$

The velocity vector $\boldsymbol{v}$ has a simple interpretation. Suppose that s is the arc-length travelled by P, measured from some fixed point of its path, and that s is increasing with time.*

* The arguments that follow assume a familiarity with the unit tangent and normal vectors to a general curve, as described in section 1.7

Then, by the chain rule,

$$\begin{aligned} \boldsymbol{v} &= \frac{d\boldsymbol{r}}{dt} = \frac{d\boldsymbol{r}}{ds} \times \frac{ds}{dt} \\ &= v\,\boldsymbol{t} \end{aligned}$$

where $\boldsymbol{t}$ is the **unit tangent vector** to the path and v $(= ds/dt)$ is the **speed*** of P. Thus, at each instant, *the direction of the velocity vector* $\boldsymbol{v}$ *is along the tangent to its path, and* $|\boldsymbol{v}|$ *is the speed of* P.

The acceleration vector $\boldsymbol{a}$ is harder to picture. This is partly because we are too accustomed to the special case of straight line motion. However, in general,

$$\begin{aligned} \boldsymbol{a} &= \frac{d\boldsymbol{v}}{dt} = \frac{d(v\,\boldsymbol{t})}{dt} = \frac{dv}{dt}\boldsymbol{t} + v\frac{d\boldsymbol{t}}{dt} = \left(\frac{dv}{dt}\right)\boldsymbol{t} + v\left(\frac{d\boldsymbol{t}}{ds} \times \frac{ds}{dt}\right) \\ &= \left(\frac{dv}{dt}\right)\boldsymbol{t} + \left(\frac{v^2}{\rho}\right)\boldsymbol{n}, \end{aligned} \tag{2.5}$$

where $\boldsymbol{n}$ is the unit normal vector to the path of P and ρ $(= \kappa^{-1})$ is its radius of curvature. Hence, *the acceleration vector* $\boldsymbol{a}$ *has a component* dv/dt *tangential to the path and a component* v^2/ρ *normal to the path.*

This formula is surprising. Since each small segment of the path is 'approximately straight' one might be tempted to conclude that only the first term $(dv/dt)\boldsymbol{t}$ should be present. However, what we have shown is that the acceleration vector of P does not generally point along the path but has a component *perpendicular* to the local path direction. The full meaning of formula (2.5) will become clear when we have treated particle motion in polar coordinates.

Uniform circular motion

The simplest example of non-rectilinear motion is motion in a circle. Circular motion is important in practical applications such as rotating machinery. Here we consider the special case of *uniform* circular motion, that is, circular motion with *constant speed*.

Consider a particle P moving with *constant speed* u in the anti-clockwise direction around a circle centre O and radius b, as shown in Figure 2.3. At time $t = 0$, P is at the point $B(b, 0)$. What are its velocity and acceleration vectors at time t?

The first step is to find the position vector of P at time t. Since P moves with constant speed u, the *arc* length BP travelled in time t must be ut. It follows that the angle θ shown in Figure 2.3 is given by $\theta = ut/b$. The position vector of P at time t is therefore

$$\begin{aligned} \boldsymbol{r} &= b\cos\theta\,\boldsymbol{i} + b\sin\theta\,\boldsymbol{j}, \\ &= b\cos(ut/b)\,\boldsymbol{i} + b\sin(ut/b)\,\boldsymbol{j}. \end{aligned}$$

* As in the rectilinear case, *speed* means the rate of increase of the total distance travelled, which, in the present context, is ds/dt, the rate of increase of arc length along the path of P.

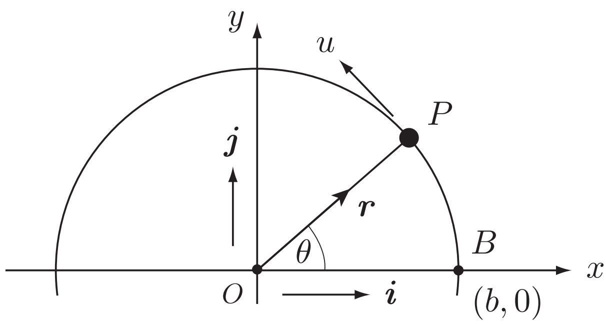


FIGURE 2.3 Particle P moves with constant speed u around a circle of radius b.

It follows that the **velocity** and **acceleration** of P at time t are given by

$$\boldsymbol{v} = \frac{d\boldsymbol{r}}{dt} = -u\sin(ut/b)\,\boldsymbol{i} + u\cos(ut/b)\,\boldsymbol{j},$$
$$\boldsymbol{a} = \frac{d\boldsymbol{v}}{dt} = -\frac{u^2}{b}\cos(ut/b)\,\boldsymbol{i} - \frac{u^2}{b}\sin(ut/b)\,\boldsymbol{j}.$$

We note that the speed of P, calculated from $\boldsymbol{v}$, is

$$|\,\boldsymbol{v}\,| = \left(u^2\cos^2(ut/b) + u^2\sin^2(ut/b)\right)^{1/2} = u,$$

which is what it was specified to be.

The *magnitude* of the acceleration $\boldsymbol{a}$ is given by

$$|\,\boldsymbol{a}\,| = \left(\left(\frac{u^2}{b}\right)^2\cos^2(ut/b) + \left(\frac{u^2}{b}\right)^2\sin^2(ut/b)\right)^{1/2} = \frac{u^2}{b}$$

and, since $\boldsymbol{a} = -(u^2/b^2)\boldsymbol{r}$, the *direction* of $\boldsymbol{a}$ is opposite to that of $\boldsymbol{r}$. This proves the following important result:

Uniform circular motion

When a particle P moves with constant speed u around a fixed circle with centre O and radius b, its acceleration vector is in the direction $\overrightarrow{PO}$ and has constant magnitude u^2/b.

This result is consistent with the general formula (2.5). In this special case, we have $v = u$ and $\rho = b$ so that $dv/dt = 0$ and $\boldsymbol{a} = (u^2/b)\boldsymbol{n}$.

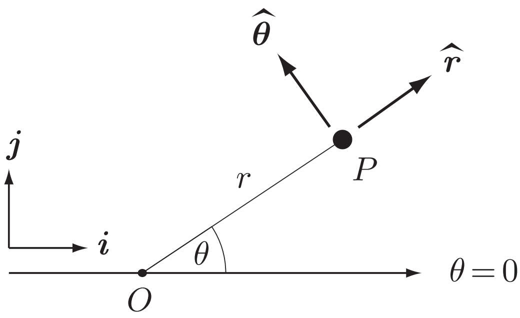


FIGURE 2.4 The plane polar co-ordinates r, θ of the point P and the polar unit vectors $\widehat{\boldsymbol{r}}$ and $\widehat{\boldsymbol{\theta}}$ at P.

Example 2.5 *Uniform circular motion*

A body is being whirled round at $10\ \mathrm{m\,s^{-1}}$ on the end of a rope. If the body moves on a circular path of 2 m radius, find the magnitude and direction of its acceleration.

Solution

The acceleration is directed towards the centre of the circle and its magnitude is $10^2/2 = 50\ \mathrm{m\,s^{-1}}$, five times the acceleration due to Earth's gravity! ■

2.3 PARTICLE MOTION IN POLAR CO-ORDINATES

When a particle is moving in a plane, it is sometimes very convenient to use polar co-ordinates r, θ in the analysis of its motion; the case of circular motion is an obvious example. Less obviously, polar co-ordinates are used in the analysis of the orbits of the planets. This famous problem stimulated Newton to devise his laws of mechanics.

Figure 2.4 shows the polar co-ordinates r, θ of a point P and the **polar unit vectors** $\widehat{\boldsymbol{r}}$, $\widehat{\boldsymbol{\theta}}$ at P. The directions of the vectors $\widehat{\boldsymbol{r}}$ and $\widehat{\boldsymbol{\theta}}$ are called the **radial** and **transverse** directions respectively at the point P. As P moves around, the polar unit vectors do not remain constant. They have constant magnitude (unity) but their directions depend on the θ co-ordinate of P; they are however independent of the r co-ordinate.* In other words, $\widehat{\boldsymbol{r}}$, $\widehat{\boldsymbol{\theta}}$ are *vector functions of the scalar variable* θ.

We will now evaluate the two derivatives $d\widehat{\boldsymbol{r}}/d\theta$, $d\widehat{\boldsymbol{\theta}}/d\theta$. These will be needed when we derive the formulae for the velocity and acceleration of P in polar co-ordinates. First we expand† $\widehat{\boldsymbol{r}}$, $\widehat{\boldsymbol{\theta}}$ in terms of the Cartesian basis vectors $\{\boldsymbol{i}, \boldsymbol{j}\}$. This gives

$$\widehat{\boldsymbol{r}} = \cos\theta\,\boldsymbol{i} + \sin\theta\,\boldsymbol{j}, \tag{2.6}$$

$$\widehat{\boldsymbol{\theta}} = -\sin\theta\,\boldsymbol{i} + \cos\theta\,\boldsymbol{j}. \tag{2.7}$$

Since $\widehat{\boldsymbol{r}}$, $\widehat{\boldsymbol{\theta}}$ are now expressed in terms of the *constant* vectors $\boldsymbol{i}$, $\boldsymbol{j}$, the differentiations with respect to θ are simple and give

* If this is not clear, sketch the directions of the polar unit vectors for P in a few different positions.

† Recall that *any* vector $\boldsymbol{V}$ lying in the plane of $\boldsymbol{i}$, $\boldsymbol{j}$ can be expanded in the form $\boldsymbol{V} = \alpha\,\boldsymbol{i} + \beta\,\boldsymbol{j}$, where the coefficients α, β are the components of $\boldsymbol{V}$ in the $\boldsymbol{i}$- and $\boldsymbol{j}$-directions respectively.

$$\boxed{\frac{d\widehat{\boldsymbol{r}}}{d\theta} = \widehat{\boldsymbol{\theta}} \qquad \frac{d\widehat{\boldsymbol{\theta}}}{d\theta} = -\widehat{\boldsymbol{r}}} \tag{2.8}$$

Suppose now that P is a moving particle with polar co-ordinates r, θ that are functions of the time t. The position vector of P relative to O has magnitude $OP = r$ and direction $\widehat{\boldsymbol{r}}$ and can therefore be written

$$\boldsymbol{r} = r\,\widehat{\boldsymbol{r}}. \tag{2.9}$$

In what follows, one must distinguish carefully between the position vector $\boldsymbol{r}$, which is the *vector* $\overrightarrow{OP}$, the co-ordinate r, which is the *distance* OP, and the polar unit vector $\widehat{\boldsymbol{r}}$.

To obtain the polar formula for the velocity of P, we differentiate formula (2.9) with respect to t. This gives

$$\boldsymbol{v} = \frac{d\boldsymbol{r}}{dt} = \frac{d}{dt}(r\,\widehat{\boldsymbol{r}}) = \left(\frac{dr}{dt}\right)\widehat{\boldsymbol{r}} + r\left(\frac{d\widehat{\boldsymbol{r}}}{dt}\right) \tag{2.10}$$

$$= \dot{r}\,\widehat{\boldsymbol{r}} + r\left(\frac{d\widehat{\boldsymbol{r}}}{dt}\right) \tag{2.11}$$

We will use the **dot notation** for time derivatives throughout this section; $\dot{r}$ means dr/dt, $\dot{\theta}$ means $d\theta/dt$, $\ddot{r}$ means d^2r/dt^2 and $\ddot{\theta}$ means $d^2\theta/dt^2$.

Now $\widehat{\boldsymbol{r}}$ is a function of θ which is, in its turn, a function of t. Hence, by the chain rule and formula (2.8),

$$\frac{d\widehat{\boldsymbol{r}}}{dt} = \frac{d\widehat{\boldsymbol{r}}}{d\theta} \times \frac{d\theta}{dt} = \widehat{\boldsymbol{\theta}} \times \dot{\theta} = \dot{\theta}\,\widehat{\boldsymbol{\theta}}.$$

If we now substitute this formula into equation (2.11) we obtain

$$\boldsymbol{v} = \dot{r}\,\widehat{\boldsymbol{r}} + \left(r\dot{\theta}\right)\widehat{\boldsymbol{\theta}}, \tag{2.12}$$

which is the polar formula for the **velocity** of P.

To obtain the polar formula for acceleration, we differentiate the velocity formula (2.12) with respect to t. This gives*

$$\begin{aligned}
\boldsymbol{a} = \frac{d\boldsymbol{v}}{dt} &= \frac{d}{dt}(\dot{r}\,\widehat{\boldsymbol{r}}) + \frac{d}{dt}\left((r\dot{\theta})\widehat{\boldsymbol{\theta}}\right) \\
&= \ddot{r}\,\widehat{\boldsymbol{r}} + \dot{r}\,\frac{d\widehat{\boldsymbol{r}}}{dt} + \left(\dot{r}\dot{\theta} + r\ddot{\theta}\right)\widehat{\boldsymbol{\theta}} + \left(r\dot{\theta}\right)\frac{d\widehat{\boldsymbol{\theta}}}{dt} \\
&= \ddot{r}\,\widehat{\boldsymbol{r}} + \dot{r}\left(\frac{d\widehat{\boldsymbol{r}}}{d\theta} \times \frac{d\theta}{dt}\right) + \left(\dot{r}\dot{\theta} + r\ddot{\theta}\right)\widehat{\boldsymbol{\theta}} + \left(r\dot{\theta}\right)\left(\frac{d\widehat{\boldsymbol{\theta}}}{d\theta} \times \frac{d\theta}{dt}\right) \\
&= \ddot{r}\,\widehat{\boldsymbol{r}} + \left(\dot{r}\dot{\theta}\right)\widehat{\boldsymbol{\theta}} + \left(\dot{r}\dot{\theta} + r\ddot{\theta}\right)\widehat{\boldsymbol{\theta}} - \left(r\dot{\theta}^2\right)\widehat{\boldsymbol{r}} \\
&= \left(\ddot{r} - r\dot{\theta}^2\right)\widehat{\boldsymbol{r}} + \left(r\ddot{\theta} + 2\dot{r}\dot{\theta}\right)\widehat{\boldsymbol{\theta}},
\end{aligned}$$

* Be a hero. Obtain this formula yourself without looking at the text.

which is the polar formula for the **acceleration** of P. These results are summarised below:

Polar formulae for velocity and acceleration

If a particle is moving in a plane and has polar coordinates r, θ at time t, then its velocity and acceleration vectors are given by

$$\boldsymbol{v} = \dot{r}\,\widehat{\boldsymbol{r}} + \left(r\dot{\theta}\right)\widehat{\boldsymbol{\theta}}, \tag{2.13}$$

$$\boldsymbol{a} = \left(\ddot{r} - r\dot{\theta}^2\right)\widehat{\boldsymbol{r}} + \left(r\ddot{\theta} + 2\dot{r}\dot{\theta}\right)\widehat{\boldsymbol{\theta}}. \tag{2.14}$$

The formula (2.13) shows that the velocity of P is the vector sum of an outward radial velocity $\dot{r}$ and a transverse velocity $r\dot{\theta}$; in other words $\boldsymbol{v}$ is just the sum of the velocities that P *would* have if r and θ varied separately. This is *not* true for the acceleration as it will be observed that adding together the separate accelerations would not yield the term $2\dot{r}\dot{\theta}\,\widehat{\boldsymbol{\theta}}$. This 'Coriolis term' is certainly present however, but is difficult to interpret intuitively.

Example 2.6 *Velocity and acceleration in polar coordinates*

A particle sliding along a radial groove in a rotating turntable has polar coordinates at time t given by

$$r = ct \qquad \theta = \Omega t,$$

where c and Ω are positive constants. Find the velocity and acceleration vectors of the particle at time t and find the speed of the particle at time t.

Deduce that, for $t > 0$, the angle between the velocity and acceleration vectors is always acute.

Solution

From the polar formulae (2.13), (2.14) for velocity and acceleration, we obtain

$$\boldsymbol{v} = c\,\widehat{\boldsymbol{r}} + (ct)\Omega\,\widehat{\boldsymbol{\theta}} = c\left(\widehat{\boldsymbol{r}} + \Omega t\,\widehat{\boldsymbol{\theta}}\right)$$

and

$$\boldsymbol{a} = \left(0 - (ct)\Omega^2\right)\widehat{\boldsymbol{r}} + (0 + 2c\Omega)\,\widehat{\boldsymbol{\theta}} = c\Omega\left(-\Omega t\,\widehat{\boldsymbol{r}} + 2\widehat{\boldsymbol{\theta}}\right).$$

The *speed* of the particle at time t is thus given by $|\boldsymbol{v}| = c\left(1 + \Omega^2 t^2\right)^{1/2}$.

To find the angle between $\boldsymbol{v}$ and $\boldsymbol{a}$, consider

$$\begin{aligned}\boldsymbol{v}\cdot\boldsymbol{a} &= c^2\Omega(-\Omega t + 2\Omega t) = c^2\Omega^2 t\\ &> 0\end{aligned}$$

for $t > 0$. Hence, for $t > 0$, the angle between $\boldsymbol{v}$ and $\boldsymbol{a}$ is acute. ■

General circular motion

An important application of polar coordinates is to circular motion. We have already considered the special case of *uniform* circular motion, but now we suppose that P moves in any manner (not necessarily with constant speed) around a circle with centre O and radius b. If we take O to be the origin of polar coordinates, the condition $r = b$ implies that $\dot{r} = \ddot{r} = 0$ and the formula (2.13) for the **velocity** of P reduces to

$$\boldsymbol{v} = \left(b\dot{\theta}\right)\widehat{\boldsymbol{\theta}}. \tag{2.15}$$

This result is depicted in Figure 2.5. The transverse velocity component $b\dot{\theta}$ (which is not necessarily the speed of P since $\dot{\theta}$ may be negative) is called the **circumferential velocity** of P. Circumferential velocity will be important when we study the motion of a rigid body rotating about a fixed axis; in this case, each particle of the rigid body moves on a circular path.

The corresponding formula for the acceleration of P is

$$\begin{aligned}\boldsymbol{a} &= \left(0 - b\dot{\theta}^2\right)\widehat{\boldsymbol{r}} + + \left((b\ddot{\theta} + 0\right)\widehat{\boldsymbol{\theta}}\\ &= -\left(b\dot{\theta}^2\right)\widehat{\boldsymbol{r}} + + \left((b\ddot{\theta}\right)\widehat{\boldsymbol{\theta}}\\ &= -\left(\frac{v^2}{b}\right)\widehat{\boldsymbol{r}} + +\dot{v}\widehat{\boldsymbol{\theta}}\end{aligned}$$

where v is the circumferential velocity $b\dot{\theta}$. These results are summarised below:

General circular motion

Suppose a particle P moves in any manner around the circle $r = b$, where r, θ are plane polar coordinates. Then the velocity and acceleration vectors of P are given by

$$\boldsymbol{v} = v\widehat{\boldsymbol{\theta}}, \tag{2.16}$$

$$\boldsymbol{a} = -\left(\frac{v^2}{b}\right)\widehat{\boldsymbol{r}} + \dot{v}\widehat{\boldsymbol{\theta}}, \tag{2.17}$$

where v $(= b\dot{\theta})$ is the circumferential velocity of P.

The formula (2.17) shows that, in *general* circular motion, the acceleration of P is the (vector) sum of an inward radial acceleration v^2/b and a transverse acceleration $\dot{v}$. This is consistent with the general formula (2.5). Indeed, what the formula (2.5) says is that, when P moves along a completely general path, its acceleration vector is the same *as if* it were moving on the circle of curvature at each point of its path.

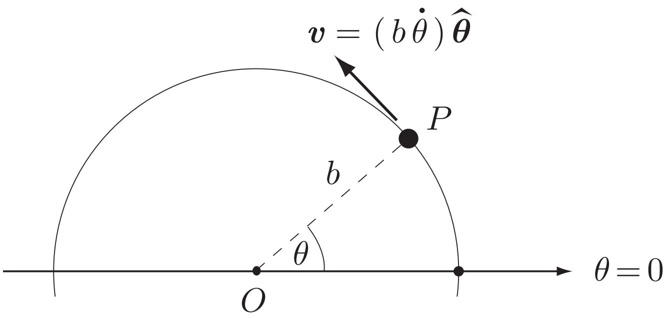


FIGURE 2.5 The particle P moves on the circle with centre O and radius b. At time t its angular displacement is θ and its circumferential velocity is $b\dot{\theta}$.

Example 2.7 *Pendulum motion*

The bob of a certain pendulum moves on a vertical circle of radius b and, when the string makes an angle θ with the downward vertical, the circumferential velocity v of the bob is given by

$$v^2 = 2gb\cos\theta,$$

where g is a positive constant. Find the acceleration of the bob when the string makes angle θ with the downward vertical.

Solution

From the acceleration formula (2.17), we have

$$\boldsymbol{a} = -\left(\frac{v^2}{b}\right)\widehat{\boldsymbol{r}} + \dot{v}\,\widehat{\boldsymbol{\theta}} = -\,(2g\cos\theta)\,\widehat{\boldsymbol{r}} + \dot{v}\,\widehat{\boldsymbol{\theta}}.$$

It remains to express $\dot{v}$ in terms of θ. On differentiating the formula $v^2 = 2gb\cos\theta$ with respect to t, we obtain

$$2v\dot{v} = -\,(2gb\sin\theta)\,\dot{\theta},$$

and, since $b\dot{\theta} = v$, we find that

$$\dot{v} = -g\sin\theta.$$

Hence the **acceleration** of the bob when the string makes angle θ with the downward vertical is

$$\boldsymbol{a} = -\,(2g\cos\theta)\,\widehat{\boldsymbol{r}} - (g\sin\theta)\,\widehat{\boldsymbol{\theta}}. \blacksquare$$

2.4 RIGID BODY ROTATING ABOUT A FIXED AXIS

Some objects that we find in everyday life, such as a brick or a thick steel rod, are so difficult to deform that their shape is virtually unchangeable. We model such an

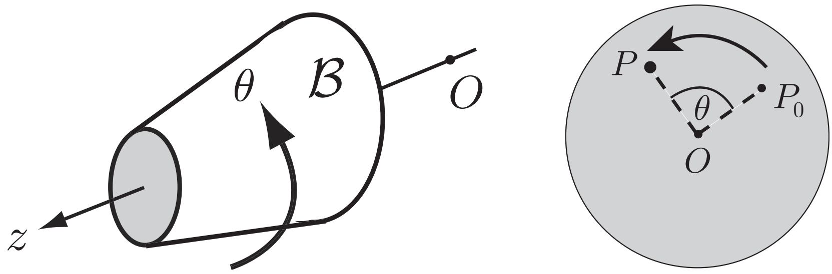


FIGURE 2.6 The rigid body $\mathcal{B}$ rotates about the fixed axis Oz and has angular displacement θ at time t. Each particle P of $\mathcal{S}$ moves on a circular path; the point P_0 is the reference position of P.

object by a **rigid body**, a collection of particles forming a perfectly rigid framework. Any motion of the rigid body must maintain this framework.

An important type of rigid body motion is **rotation about a fixed axis**; a spinning fan, a door opening on its hinges and a playground roundabout are among the many examples of this type of motion. Suppose $\mathcal{B}$ is a rigid body which is constrained to rotate about the fixed axis Oz as shown in Figure 2.6. (This means that the particles of $\mathcal{B}$ that lie on Oz are held fixed. Rotation about Oz is then the only motion of $\mathcal{B}$ consistent with rigidity.) At time t, $\mathcal{B}$ has **angular displacement** θ measured from some reference position. The angular displacement θ is the rotational counterpart of the Cartesian displacement x of a particle in straight line motion. By analogy with the rectilinear case, we make the following definitions:

Definition 2.4 ***Angular velocity*** *The **angular velocity** ω of $\mathcal{B}$ is defined to be $\omega = d\theta/dt$ and the absolute value of ω is called the **angular speed** of $\mathcal{B}$.*

Units. Angular velocity (and angular speed) are measured in radians per second (rad s^{-1}).

Example 2.8 ***Spinning crankshaft 1***

The crankshaft of a motorcycle engine is spinning at 6000 revolutions per minute. What is its angular speed in S.I. units?

Solution

6000 revolutions per minute is 100 revolutions per second which is 200π radians per second. This is the angular speed in S.I. units. ■

Particle velocities in a rotating rigid body

In rotational motion about a fixed axis, each particle P of $\mathcal{B}$ moves on a circle of some radius ρ, where ρ is the (fixed) perpendicular distance of P from the rotation axis. It then follows from (2.16) that the **circumferential velocity** v of P is given by $\rho\dot{\theta}$, that is

$$\boxed{v = \omega\,\rho} \qquad (2.18)$$

Example 2.9 ***Spinning crankshaft 2***

In the crankshaft example above, find the speed of a particle of the crankshaft that has perpendicular distance 5 cm from the rotation axis. Find also the magnitude of its acceleration.

Solution

In this case, $|\omega| = 200\pi$ and $\rho = 1/20$ so that the particle speed (the magnitude of the circumferential velocity v) is $10\pi \approx 31.4 \text{ m s}^{-1}$.

Since the circumferential velocity is constant, $|\boldsymbol{a}| = v^2/\rho = (10\pi)^2/0.05 \approx 2000 \text{ m s}^{-2}$, which is two hundred times the value of the Earth's gravitational acceleration! ■

2.5 RIGID BODY IN PLANAR MOTION

We now consider a more general form of rigid body motion called **planar motion**.

Definition 2.5 ***Planar motion*** *A rigid body $\mathcal{B}$ is said to be in* ***planar motion*** *if each particle of $\mathcal{B}$ moves in a fixed plane and all these planes are parallel to each other.*

Planar motion is quite common. For instance, any flat-bottomed rigid body sliding on a flat table is in planar motion. Another example is a circular cylinder rolling on a rough flat table.

The **particle velocities** in planar motion can be calculated by the following method; the proof is given in Chapter 16. First select some particle C of the body as the reference particle. The velocity of a general particle P of the body is then the vector sum of

(i) a **translational contribution** equal to the velocity of C (as if the body did not rotate) and
(ii) a **rotational contribution** (as if C were fixed and the body were rotating with angular velocity ω about a fixed axis through C).

This result is illustrated in Figure 2.7, where the body is a rectangular plate and the reference particle C is at a corner of the plate. The velocity $\boldsymbol{v}$ of P is given by $\boldsymbol{v} = \boldsymbol{v}^C + \boldsymbol{v}^R$, where the translational contribution $\boldsymbol{v}^C$ is the velocity of C and the rotational contribution $\boldsymbol{v}^R$ is caused by the angular velocity ω about C. Although the reference particle can be any particle of the body, it is usually taken to be the centre of mass or centre of symmetry of the body.

Example 2.10 ***The rolling wheel***

A circular wheel of radius b rolls in a straight line with speed u on a fixed horizontal table. Find the velocities of its particles.

Solution

This is an instance of planar motion and so the particle velocities can be found by the method above. Let the position of the wheel at some instant be that shown in

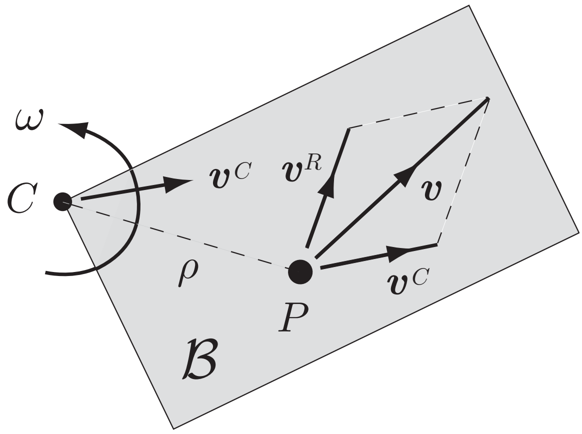


FIGURE 2.7 The velocity of the particle P belonging to the rigid body $\mathcal{B}$ is the sum of the translational contribution $\boldsymbol{v}^C$ and the rotational contribution $\boldsymbol{v}^R$. The reference particle C can be *any* particle of the body.

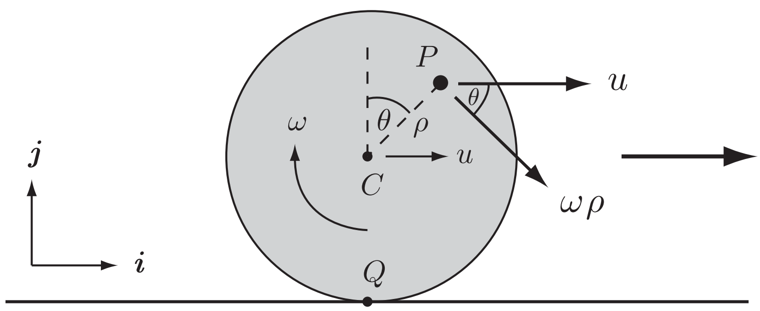


FIGURE 2.8 The circular wheel rolls from left to right on a fixed horizontal table. The reference particle C is taken to be the centre of the wheel and the velocity of a typical particle P is the sum of the two velocities shown.

Figure 2.8. The reference particle C is taken to be the centre of the wheel, and the wheel is supposed to have some angular velocity ω about C. The velocity $\boldsymbol{v}^P$ of a typical particle P is then the sum of the two velocities shown. In terms of the vectors $\{\boldsymbol{i}, \boldsymbol{j}\}$

$$\begin{aligned}\boldsymbol{v}^P &= u\,\boldsymbol{i} + \omega\rho\,(\cos\theta\,\boldsymbol{i} - \sin\theta\,\boldsymbol{j})\\ &= (u + \omega\rho\cos\theta)\,\boldsymbol{i} - (\omega\rho\sin\theta)\,\boldsymbol{j}.\end{aligned} \tag{2.19}$$

In particular, on taking $\rho = b$ and $\theta = \pi$, the velocity $\boldsymbol{v}^Q$ of the contact particle Q is given by

$$\boldsymbol{v}^Q = (u - \omega b)\,\boldsymbol{i}. \tag{2.20}$$

If the wheel is allowed to slip as it moves across the table, there is no restriction on $\boldsymbol{v}^Q$ so that u and ω are unrelated. But rolling, by definition, requires that

$$\boldsymbol{v}^Q = \boldsymbol{0}. \tag{2.21}$$

On applying this **rolling condition** to our formula (2.20) for $\boldsymbol{v}^Q$, we find that ω must be related to u by

$$\omega = \frac{u}{b}, \tag{2.22}$$

and on using this value of ω in (2.19) we find that the **velocity** of the typical particle P is given by

$$\boldsymbol{v}^P = u\left(1 + \frac{\rho}{b}\cos\theta\right)\boldsymbol{i} - u\left(\frac{\rho}{b}\sin\theta\right)\boldsymbol{j}. \tag{2.23}$$

When P lies on the circumference of the wheel, this formula simplifies to

$$\boldsymbol{v}^P = u\,(1 + \cos\theta)\,\boldsymbol{i} - u\sin\theta\,\boldsymbol{j}, \tag{2.24}$$

in which case the *speed* of P is given by

$$|\,\boldsymbol{v}^P\,| = 2u\cos(\theta/2), \qquad (-\pi \le \theta \le \pi).$$

Thus the highest particle of the wheel has the largest speed, $2u$, while the contact particle has speed zero, as we already know. ■

2.6 REFERENCE FRAMES IN RELATIVE MOTION

A reference frame is simply a rigid coordinate system that can be used to specify the positions of points in space. In practice it is convenient to regard a reference frame as being embedded in, or attached to, some rigid body. The most familiar case is that in which the rigid body is the Earth but it could instead be a moving car, or an orbiting space station. In principle, any event, the motion of an aircraft for example, can be observed from any of these reference frames and the motion will appear different to each observer. It is this difference that we now investigate.

Let the motion of a particle P be observed from the reference frames $\mathcal{F}\{\mathcal{O}\,;\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}\}$ and $\mathcal{F}'\{O'\,;\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}\}$ as shown in Figure 2.9. Here we are supposing that the frame $\mathcal{F}'$ *does not rotate* relative to $\mathcal{F}$. This is why, without losing generality, we can suppose that $\mathcal{F}$ and $\mathcal{F}'$ have the same set of unit vectors $\{\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}\}$. For example, P could be an aircraft, $\mathcal{F}$ could be attached to the Earth, and $\mathcal{F}'$ could be attached to a car driving along a *straight* road.

Then, $\boldsymbol{r}$, $\boldsymbol{r}'$, the position vectors of P relative to $\mathcal{F}$, $\mathcal{F}'$ are connected by

$$\boldsymbol{r} = \boldsymbol{r}' + \boldsymbol{D}, \tag{2.25}$$

where $\boldsymbol{D}$ is the position vector of O' relative to $\mathcal{F}$.

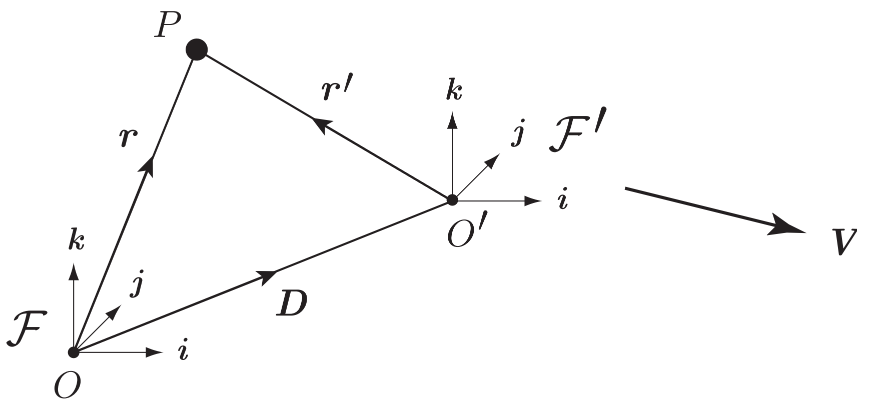


FIGURE 2.9 The particle P is observed from the two reference frames $\mathcal{F}$ and $\mathcal{F}'$.

We now differentiate this equation with respect to t, a step that requires some care. Let us consider the rates of change of the vectors in equation (2.25), as observed from the frame $\mathcal{F}$. Then

$$\boldsymbol{v} = \left(\frac{d\boldsymbol{r}'}{dt}\right)_{\mathcal{F}} + \boldsymbol{V}, \tag{2.26}$$

where $\boldsymbol{v}$ is the velocity of P observed in $\mathcal{F}$ and $\boldsymbol{V}$ is the velocity of $\mathcal{F}'$ relative to $\mathcal{F}$.

Now when two different reference frames are used to observe the same vector, the observed *rates of change* of that vector will generally be different. In particular, it is *not* generally true that

$$\left(\frac{d\boldsymbol{r}'}{dt}\right)_{\mathcal{F}} = \left(\frac{d\boldsymbol{r}'}{dt}\right)_{\mathcal{F}'}.$$

However, as we will show in Chapter 17, these two rates of change *are* equal if the frame $\mathcal{F}$ *does not rotate* relative to $\mathcal{F}$. Hence, in our case, we do have

$$\left(\frac{d\boldsymbol{r}'}{dt}\right)_{\mathcal{F}} = \left(\frac{d\boldsymbol{r}'}{dt}\right)_{\mathcal{F}'} = \boldsymbol{v}',$$

where $\boldsymbol{v}'$ is the velocity of P observed in $\mathcal{F}'$.

Equation (2.26) can then be written

$$\boxed{\boldsymbol{v} = \boldsymbol{v}' + \boldsymbol{V}} \tag{2.27}$$

Thus *the velocity of P observed in $\mathcal{F}$ is the sum of the velocity of P observed in $\mathcal{F}'$ and the velocity of the frame $\mathcal{F}'$ relative to $\mathcal{F}$.* This result applies only when $\mathcal{F}'$ *does not rotate* relative to $\mathcal{F}$.

This is the well known rule for handling 'relative velocities'. In the aircraft example, it means that the true velocity of the aircraft (relative to the ground) is the vector sum of (i) the velocity of the aircraft relative to the car, and (ii) the velocity of the car relative to the road.

Example 2.11 ***Relative velocity***

The Mississippi river is a mile wide and has a uniform flow. A steamboat sailing at full speed takes 12 minutes to cover a mile when sailing upstream, but only 3 minutes when sailing downstream. What is the shortest time in which the steamboat can *cross* the Mississippi to the nearest point on the opposite bank?

Solution

The way to handle this problem is to view the motion of the boat from a reference frame $\mathcal{F}'$ moving with the river. In this reference frame the water is at rest and the boat sails with the same speed *in all directions*. The relative velocity formula (2.27) then gives us the true picture of the motion of the boat relative to the river bank, which is the reference frame $\mathcal{F}$.

Let u^B be the speed of the boat in still water and u^R be the speed of the river, both measured in miles per hour. The upstream and downstream times are just a sneaky way of telling us the values of u^B and u^R. When the boat sails downstream, (2.27) implies that its speed relative to the bank is $u^B + u^R$. But this speed is stated to be 1/3 mile per minute (or 20 miles per hour). Hence

$$u^B + u^R = 20.$$

Similarly the upstream speed is $u^B - u^R$ and is stated to be 1/12 mile per minute (or 5 miles per hour). Hence

$$u^B - u^R = 5.$$

Solving these equations yields

$$u^B = 12.5 \text{ mph}, \qquad u^R = 7.5 \text{ mph}.$$

Now the boat must cross the river. In order to cross by a straight line path to the nearest point on the opposite bank, the boat's velocity (relative to the water) must be directed at some angle α to the required path (as shown in Figure 2.10) so that its *resultant* velocity is perpendicular to the stream. For this to be true, α must satisfy

$$u^B \sin\alpha = u^R,$$

which gives $\sin\alpha = 3/5$. The resultant speed of the boat when crossing the river is therefore $u^B \cos\alpha = 12.5 \times (4/5) = 10$ mph. Since the river is one mile wide, the time taken for the crossing is $1/10$ hour = 6 minutes. ■

The relative velocity formula (2.27) can be differentiated again with repect to t to give a similar connection between accelerations. The result is that

$$\boldsymbol{a} = \boldsymbol{a}' + \boldsymbol{A}, \qquad (2.28)$$

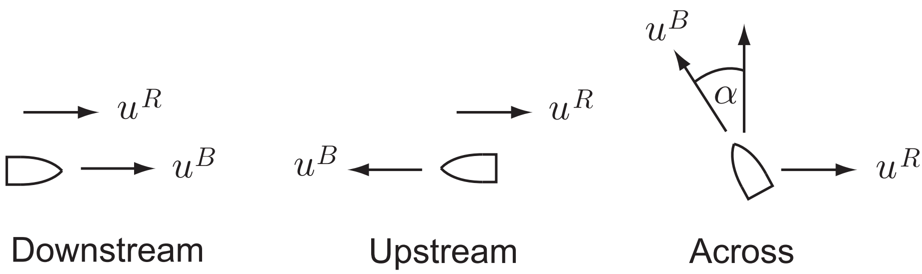


FIGURE 2.10 The river flows from left to right with speed u^R and the boat sails with speed u^B *relative to the river*. In each case the velocity of the boat relative to the bank is the vector sum of the two velocities shown.

where $\boldsymbol{a}$ and $\boldsymbol{a}'$ are the accelerations of P relative to the frames $\mathcal{F}$ and $\mathcal{F}'$ respectively, and $\boldsymbol{A}$ is the acceleration of the frame $\mathcal{F}'$ relative to the frame $\mathcal{F}$. Once again, this result applies only when $\mathcal{F}'$ does not rotate relative to $\mathcal{F}$.

Mutually unaccelerated frames

An important special case of equation (2.28) occurs when the frame $\mathcal{F}'$ is moving with *constant velocity* (and no rotation) relative to $\mathcal{F}$. We will then say that $\mathcal{F}$ and $\mathcal{F}'$ are **mutually unaccelerated** frames. In this case $\boldsymbol{A} = \boldsymbol{0}$ and (2.28) becomes

$$\boldsymbol{a} = \boldsymbol{a}'. \tag{2.29}$$

This means that *when mutually unaccelerated frames are used to observe the motion of a particle P, the observed acceleration of P is the same in each frame*.

This result will be vital in our discussion of *inertial frames* in Chapter 3.

Problems on Chapter 2

Answers and comments are at the end of the book.

Harder problems carry a star ($*$).

Rectilinear particle motion

2.1 A particle P moves along the x-axis with its displacement at time t given by $x = 6t^2 - t^3 + 1$, where x is measured in metres and t in seconds. Find the velocity and acceleration of P at time t. Find the times at which P is at rest and find its position at these times.

2.2 A particle P moves along the x-axis with its acceleration a at time t given by

$$a = 6t - 4 \text{ m s}^{-2}.$$

Initially P is at the point $x = 20$ m and is moving with speed 15 m s^{-1} in the negative x-direction. Find the velocity and displacement of P at time t. Find when P comes to rest and its displacement at this time.

2.3 ***Constant acceleration formulae*** A particle P moves along the x-axis with *constant* acceleration a in the positive x-direction. Initially P is at the origin and is moving with velocity u in the positive x-direction. Show that the velocity v and displacement x of P at time t are given by*

$$v = u + at, \qquad\qquad x = ut + \tfrac{1}{2}at^2,$$

and deduce that

$$v^2 = u^2 + 2ax.$$

In a standing quarter mile test, the Suzuki Bandit 1200 motorcycle covered the quarter mile (from rest) in 11.4 seconds and crossed the finish line doing 116 miles per hour. Are these figures consistent with the assumption of constant acceleration?

General particle motion

2.4 The trajectory of a charged particle moving in a magnetic field is given by

$$\boldsymbol{r} = b\cos\Omega t\,\boldsymbol{i} + b\sin\Omega t\,\boldsymbol{j} + ct\,\boldsymbol{k},$$

where b, Ω and c are positive constants. Show that the particle moves with constant speed and find the magnitude of its acceleration.

2.5 ***Acceleration due to rotation and orbit of the Earth*** A body is at rest at a location on the Earth's equator. Find its acceleration due to the Earth's rotation. [Take the Earth's radius at the equator to be 6400 km.]

Find also the acceleration of the Earth in its orbit around the Sun. [Take the Sun to be fixed and regard the Earth as a particle following a circular path with centre the Sun and radius 15×10^{10} m.

2.6 An insect flies on a spiral trajectory such that its polar coordinates at time t are given by

$$r = be^{\Omega t}, \qquad \theta = \Omega t,$$

where b and Ω are positive constants. Find the velocity and acceleration vectors of the insect at time t, and show that the angle between these vectors is always $\pi/4$.

2.7 A racing car moves on a circular track of radius b. The car starts from rest and its *speed* increases at a constant rate α. Find the angle between its velocity and acceleration vectors at time t.

---

* These are the famous **constant acceleration formulae**. Although they are a mainstay of school mechanics, we will make little use of them since, in most of the problems that we treat, the acceleration is *not* constant. *It is a serious offence to use these formulae in non-constant acceleration problems.*

2.8 A particle P moves on a circle with centre O and radius b. At a certain instant the speed of P is v and its acceleration vector makes an angle α with PO. Find the magnitude of the acceleration vector at this instant.

2.9∗ A bee flies on a trajectory such that its polar coordinates at time t are given by

$$r = \frac{bt}{\tau^2}(2\tau - t) \qquad \theta = \frac{t}{\tau} \qquad (0 \le t \le 2\tau),$$

where b and τ are positive constants. Find the velocity vector of the bee at time t.

Show that the least speed achieved by the bee is b/τ. Find the acceleration of the bee at this instant.

2.10∗ *A pursuit problem: Daniel and the Lion* The luckless Daniel (D) is thrown into a circular arena of radius a containing a lion (L). Initially the lion is at the centre O of the arena while Daniel is at the perimeter. Daniel's strategy is to run with his maximum speed u around the perimeter. The lion responds by running at its maximum speed U in such a way that it remains on the (moving) radius OD. Show that r, the distance of L from O, satisfies the differential equation

$$\dot{r}^2 = \frac{u^2}{a^2}\left(\frac{U^2a^2}{u^2} - r^2\right).$$

Find r as a function of t. If $U \ge u$, show that Daniel will be caught, and find how long this will take.

Show that the path taken by the lion is an arc of a circle. For the special case in which $U = u$, sketch the path taken by the lion and find the point of capture.

2.11 *General motion with constant speed* A particle moves along any path in three-dimensional space with *constant speed*. Show that its velocity and acceleration vectors must always be perpendicular to each other. [*Hint*. Differentiate the formula $\boldsymbol{v} \cdot \boldsymbol{v} = v^2$ with respect to t.]

2.12 A particle P moves so that its position vector $\boldsymbol{r}$ satisfies the differential equation

$$\dot{\boldsymbol{r}} = \boldsymbol{c} \times \boldsymbol{r},$$

where $\boldsymbol{c}$ is a constant vector. Show that P moves with constant speed on a circular path. [*Hint*. Take the dot product of the equation first with $\boldsymbol{c}$ and then with $\boldsymbol{r}$.]

Angular velocity

2.13 A large truck with double rear wheels has a brick jammed between two of its tyres which are 4 ft in diameter. If the truck is travelling at 60 mph, find the maximum speed of the brick and the magnitude of its acceleration. [Express the acceleration as a multiple of $g = 32\ \text{ft s}^{-2}$.]

2.14 A particle is sliding along a smooth radial grove in a circular turntable which is rotating with constant angular speed Ω. The distance of the particle from the rotation axis at time t is

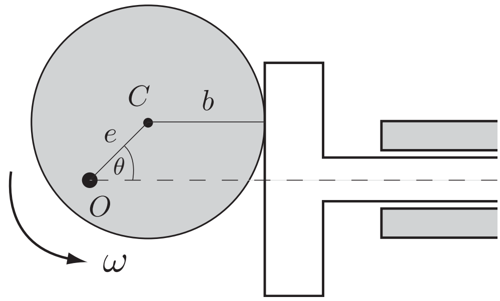


FIGURE 2.11 Cam and valve mechanism

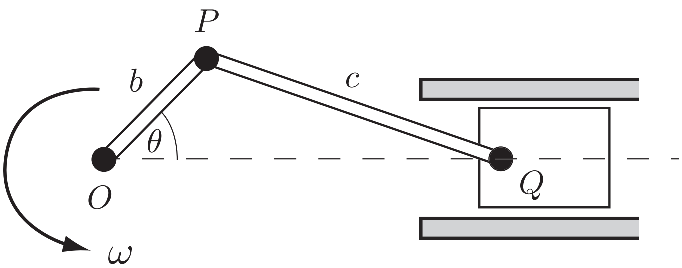


FIGURE 2.12 Crank and piston mechanism

observed to be

$$r = b \cosh \Omega t$$

for $t \geq 0$, where b is a positive constant. Find the speed of the particle (relative to a fixed reference frame) at time t, and find the magnitude and direction of the acceleration.

2.15 Figure 2.11 shows an eccentric circular cam of radius b rotating with constant angular velocity ω about a fixed pivot O which is a distance e from the centre C. The cam drives a valve which slides in a straight guide. Find the maximum speed and maximum acceleration of the valve.

2.16 Figure 2.12 shows a piston driving a crank OP pivoted at the end O. The piston slides in a straight cylinder and the crank is made to rotate with constant angular velocity ω. Find the distance OQ in terms of the lengths b, c and the angle θ. Show that, when b/c is small, OQ is given approximately by

$$OQ = c + b\cos\theta - \frac{b^2}{2c}\sin^2\theta,$$

on neglecting $(b/c)^4$ and higher powers. Using this approximation, find the maximum acceleration of the piston.

2.17 Figure 2.13 shows an epicyclic gear arrangement in which the 'sun' gear $\mathcal{G}_1$ of radius b_1 and the 'ring' gear $\mathcal{G}_2$ of inner radius b_2 rotate with angular velocities ω_1, ω_2 respectively

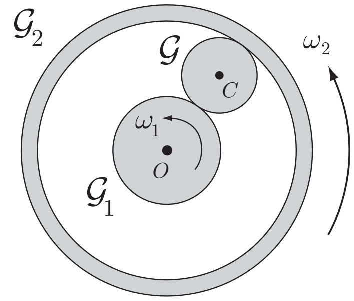


FIGURE 2.13 Epicyclic gear mechanism

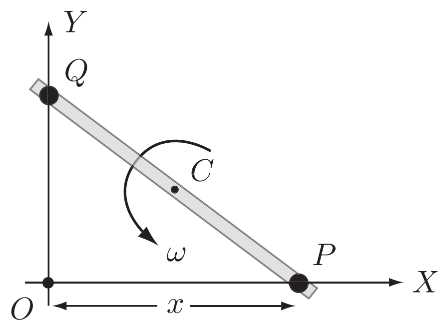


FIGURE 2.14 The pins P and Q at the ends of a rigid link move along the axes OX, OY respectively.

about their fixed common centre O. Between them they grip the 'planet' gear $\mathcal{G}$, whose centre C moves on a circle centre O. Find the circumferential velocity of C and the angular velocity of the planet gear $\mathcal{G}$. If O and C were connected by an arm pivoted at O, what would be the angular velocity of the arm?

2.18 Figure 2.14 shows a straight rigid link of length a whose ends contain pins P, Q that are constrained to move along the axes OX, OY. The displacement x of the pin P at time t is prescribed to be $x = b \sin \Omega t$, where b and Ω are positive constants with $b < a$. Find the angular velocity ω and the speed of the centre C of the link at time t.

Relative velocity

2.19 An aircraft is to fly from a point A to an airfield B 600 km due north of A. If a steady wind of 90 km/h is blowing from the north-west, find the direction the plane should be pointing and the time taken to reach B if the cruising speed of the aircraft in still air is 200 km/h.

2.20 An aircraft takes off from a horizontal runway with constant speed U, climbing at a constant angle α to the horizontal. A car is moving on the runway with constant speed u directly towards the front of the aircraft. The car is distance a from the aircraft at the instant of take-off. Find the distance of closest approach of the car and aircraft. [Don't try this one at home.]

2.21* An aircraft has cruising speed v and a flying range (out and back) of R_0 in still air. Show that, in a north wind of speed u ($u < v$) its range in a direction whose true bearing from

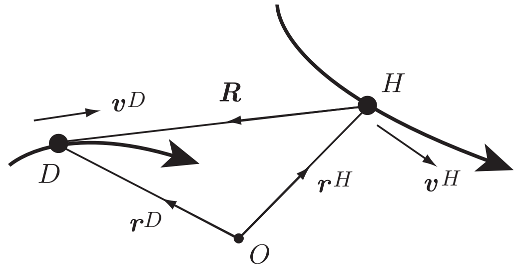


FIGURE 2.15 The dog D chases the hare H by running directly towards the hare's current position.

north is θ is given by

$$\frac{R_0(v^2 - u^2)}{v(v^2 - u^2 \sin^2 \theta)^{1/2}}.$$

What is the maximum value of this range and in what directions is it attained?

Computer assisted problems

2.22 *Dog chasing a hare; another pursuit problem.* Figure 2.15 shows a dog with position vector $\boldsymbol{r}^D$ and velocity $\boldsymbol{v}^D$ chasing a hare with position vector $\boldsymbol{r}^H$ and velocity $\boldsymbol{v}^H$. The dog's strategy is to run directly towards the current position of the hare. Given the motion of the hare and the *speed* of the dog, what path does the dog follow?

Since the dog runs directly towards the hare, its velocity $\boldsymbol{v}^D$ must satisfy

$$\frac{\boldsymbol{v}^D}{v^D} = \frac{\boldsymbol{r}^H - \boldsymbol{r}^D}{|\boldsymbol{r}^H - \boldsymbol{r}^D|}.$$

In terms of the position vector of the dog *relative to the hare*, given by $\boldsymbol{R} = \boldsymbol{r}^D - \boldsymbol{r}^H$, this equation becomes

$$\dot{\boldsymbol{R}} = -\frac{\boldsymbol{R}}{R} v^D - \boldsymbol{v}^H.$$

Given the velocity $\boldsymbol{v}^H$ of the hare and the *speed* v^D of the dog as functions of time, this differential equation determines the trajectory of the dog relative to the hare; capture occurs when $\boldsymbol{R} = \boldsymbol{0}$. The actual trajectory of the dog is given by $\boldsymbol{r}^D = \boldsymbol{R} + \boldsymbol{r}^H$.

If the motion takes place in a plane with $\boldsymbol{R} = X\,\boldsymbol{i} + Y\,\boldsymbol{j}$ then X and Y satisfy the coupled differential equations

$$\dot{X} = -\frac{v^D X}{(X^2 + Y^2)^{1/2}} - v_x^H, \qquad \dot{Y} = -\frac{v^D Y}{(X^2 + Y^2)^{1/2}} - v_y^H,$$

together with initial conditions of the form $X(0) = x_0$ and $Y(0) = y_0$. Such equations cannot usually be solved analytically but are extremely easy to solve with computer assistance. Two interesting cases to consider are as follows. In each case the *speeds* of the dog and the hare are constants.

(i) Initially the hare is at the origin with the dog at some point (x_0, y_0). The hare then runs along the positive x-axis and is chased by the dog. Show that the hare gets caught if $v^D > v^H$, but when $v^D = v^H$ the dog always misses (unless he starts directly in the path of the hare). This remarkable result can be proved analytically.

(ii) The hare runs in a circle (like the lion problem). In this case, with $v^D = v^H$, the dog seems to miss no matter where he starts.

Try some examples of your own and see if you can find interesting paths taken by the dog.

2.23 Consider further the piston problem described in Problem 2.16. Use computer assistance to calculate the exact and approximate accelerations of the piston as functions of θ. Compare the exact and approximate formulae (non-dimensionalised by $\omega^2 b$) by plotting both on the same graph against θ. Show that, when $b/c < 0.5$, the two graphs are close, but when b/c gets close to unity, large errors occur.

Chapter Three

Newton's laws of motion and the law of gravitation

KEY FEATURES

The key features of this chapter are **Newton's laws** of motion, the definitions of **mass** and **force**, the **law of gravitation**, the **principle of equivalence**, and **gravitation by spheres**.

This chapter is concerned with the **foundations of dynamics** and **gravitation**. Kinematics is concerned purely with geometry of motion, but dynamics seeks to answer the question as to what motion will actually occur when specified forces act on a body. The rules that allow one to make this connection are **Newton's laws of motion**. These are laws of physics that are founded upon experimental evidence and stand or fall according to the accuracy of their predictions. In fact, Newton's formulation of mechanics has been astonishingly successful in its accuracy and breadth of application, and has survived, essentially intact, for more than three centuries. The same is true for Newton's universal **law of gravitation** which specifies the forces that all masses exert upon each other.

Taken together, these laws represent virtually the entire foundation of classical mechanics and provide an accurate explanation for a vast range of motions from large molecules to entire galaxies.

3.1 NEWTON'S LAWS OF MOTION

Isaac Newton's* three famous laws of motion were laid down in *Principia*, written in Latin and published in 1687. These laws set out the founding principles of mechanics and have survived, essentially unchanged, to the present day. Even when translated into English, Newton's original words are hard to understand, mainly because the terminology

* Sir Isaac Newton (1643–1727) is arguably the greatest scientific genius of all time. His father was completely uneducated and Isaac himself had no contact with advanced mathematics before the age of twenty. However, by the age of twenty seven, he had been appointed to the Lucasian chair at Cambridge and was one of the foremost scientists in Europe. His greatest achievements were his discovery of the calculus, his laws of motion, and his theory of universal gravitation. On the urging of Halley (the Astronomer Royal), Newton wrote up an account of his new physics and its application to astronomy. *Philosophiae Naturalis Principia Mathematica* was published in 1687 and is generally recognised as the greatest scientific book ever written.

of the seventeenth century is now archaic. Also, the laws are now formulated as applying to particles, a concept never used by Newton. A **particle** is an idealised body that occupies only a **single point of space** and has no internal structure. True particles do not exist† in nature, but it is convenient to regard realistic bodies as being made up of particles. Using modern terminology, Newton's laws may be stated as follows:

Newton's laws of motion

First Law When all external influences on a particle are removed, the particle moves with constant velocity. [This velocity may be zero in which case the particle remains at rest.]

Second Law When a force $\boldsymbol{F}$ acts on a particle of mass m, the particle moves with instantaneous acceleration $\boldsymbol{a}$ given by the formula

$$\boldsymbol{F} = m\boldsymbol{a},$$

where the unit of force is implied by the units of mass and acceleration.

Third Law When two particles exert forces upon each other, these forces are (i) equal in magnitude, (ii) opposite in direction, and (iii) parallel to the straight line joining the two particles.

Units

Any consistent system of units can be used. The standard scientific units are **SI units** in which the unit of mass is the **kilogram**, the unit of length is the **metre**, and the unit of time is the **second**. The unit of force implied by the Second Law is called the **newton**, and written N. An excellent description of the SI system of units can be found on

`http://www.physics.nist.gov/PhysRefData`

the website of the US National Institute of Standards & Technology.

In the Imperial system of units, the unit of mass is the pound, the unit of length is the foot, and the unit of time is the second. The unit of force implied by the Second Law is called the poundal. These units are still used in some industries in the US, a fact which causes frequent confusion.

Interpreting Newton's laws

Newton's laws are clear enough in themselves but they leave some important questions unanswered, namely:

(i) In what **frame of reference** are the laws true?

† The nearest thing to a particle is the electron, which, unlike other elementary particles, does seem to be a *point* mass. The electron does however have an internal structure, having spin and angular momentum.

(ii) What are the definitions of **mass** and **force**?

These questions are answered in the sections that follow. What we do is to set aside Newton's laws for the time being and go back to simple experiments with particles. These are 'thought experiments' in the sense that, although they are perfectly meaningful, they are unlikely to be performed in practice. The supposed 'results' of these experiments are taken to be the *primitive governing laws* of mechanics on which we base our definitions of mass and force. Finally, these laws and definitions are shown to be equivalent to Newton's laws as stated above. This process could be said to provide an *interpretation* of Newton's laws. The interpretation below is quite sophisticated and is probably only suitable for those who have already seen a simpler account, such as that given by French [3].

3.2 INERTIAL FRAMES AND THE LAW OF INERTIA

The first law states that, when a particle is unaffected by external influences, it moves with constant velocity, that is, it moves in a straight line with constant speed. Thus, contrary to Aristotle's view, the particle needs no agency of any kind to maintain its motion.* Since the influence of the Earth's gravity rules out any verification of the First Law by an experiment conducted on Earth, Newton showed remarkable insight in proposing a law he could not possibly verify. In order to verify the First Law, all *external influences* must be removed, which means that we must carry out our thought experiment in a place as remote as possible from any material bodies, such as the almost empty space between the galaxies. In our minds then we go to such a place armed with a selection of test particles† which we release in various ways and observe their motion. According to the First Law, each of these particles should move with constant velocity.

Inertial reference frames

So far we have ignored the awkward question as to what reference frame we should use to observe the motion of our test particles. When confronted with this question for the first time, one's probable response is that the reference frame should be 'fixed'. But fixed to what? The Earth rotates and is in orbital motion around the Sun. Our entire solar system is part of a galaxy that rotates about its centre. The galaxies themselves move relative to each other. The fact is that everything in the universe is moving relative to everything else and nothing can properly be described as fixed. From this it might be concluded that any reference frame is as good as any other, but this is not so, for, if the First Law is true at all, it can only be true in certain special reference frames. Suppose for instance that the First Law has been found to be true in the reference frame $\mathcal{F}$. Then it is also true in any other frame $\mathcal{F}'$ that is mutually unaccelerated relative to $\mathcal{F}$ (see section 2.6). This follows because, if the test particles have constant velocities in $\mathcal{F}$, then they have

* Such a law was proposed prior to Newton by Galileo but, curiously, Galileo did not accept the consequences of his own statement.

† Since true particles do not exist, we will have to make do with uniform rigid spheres of various kinds.

zero accelerations in $\mathcal{F}$. But, since $\mathcal{F}$ and $\mathcal{F}'$ are mutually unaccelerated frames, the test particles must have zero accelerations in $\mathcal{F}'$ and thus have constant velocities in $\mathcal{F}'$. Moreover, the First Law does *not* hold in any other reference frame.

Definition 3.1 ***Inertial frame*** *A reference frame in which the First Law is true is said to be an* ***inertial frame****.*

It follows that, *if* there exists one inertial frame, then there exist infinitely many, with each frame moving with constant velocity (and no rotation) relative to any other.

It may appear that the First Law is without physical content since we are saying that it is true in those reference frames in which it is true. However, this is not so since *inertial frames need not have existed at all, and the fact that they do is the real physical content of the First Law*. *Why* there should exist this special class of reference frames in which the laws of physics take simple forms is a very deep and interesting question that we do not have to answer here!

Our discussion is summarised by the following statement which we take to be a law of physics:

The law of inertia *There exists in nature a unique class of mutually unaccelerated reference frames (the inertial frames) in which the First Law is true.*

Practical inertial frames

The preceding discussion gives no clue as to how to set up an inertial reference frame and, in practice, *exact* inertial frames are not available. Practical reference frames have to be tied to real objects that are actually available. The most common practical reference frame is **the Earth**. Such a frame is sufficiently close to being inertial for the purpose of observing most Earth-bound phenomena. The orbital acceleration of the Earth is insignificant and the effect of the Earth's rotation is normally a small correction. For example, when considering the motion of a football, a pendulum or a spinning top, the Earth may be assumed to be an inertial frame.

However, the Earth is not a suitable reference frame from which to observe the motion of an orbiting satellite. In this case, the **geocentric frame** (which has its origin at the centre of mass of the Earth and has no rotation relative to distant stars) would be appropriate. Similarly, the **heliocentric frame** (which has its origin at the centre of mass of the solar system and has no rotation relative to distant stars) is appropriate when observing the motion of the planets.

Example 3.1 *Inertial frames*

Suppose that a reference frame fixed to the Earth is *exactly* inertial. Which of the following are then inertial frames?

A frame fixed to a motor car which is

(i) moving with constant speed around a flat race track,

(ii) moving with constant speed along a straight undulating road,

(iii) moving with constant speed up a constant gradient,

(iv) freewheeling down a hill.

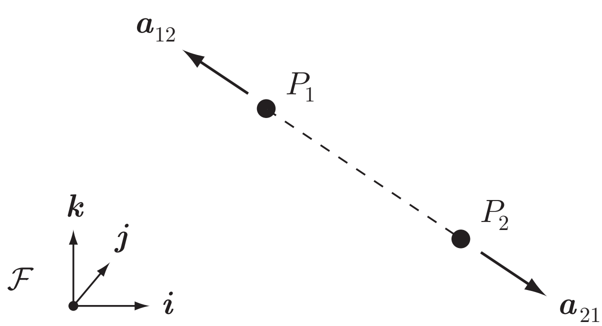


FIGURE 3.1 The particles P_1 and P_2 move under their mutual interaction and, relative to an inertial reference frame $\mathcal{F}$, have accelerations $\boldsymbol{a}_{12}$ and $\boldsymbol{a}_{21}$ repectively. These accelerations are found to satisfy the **law of mutual interaction**.

Solution

Only (iii) is inertial. In the other cases, the frame is accelerating or rotating relative to the Earth.

3.3 THE LAW OF MUTUAL INTERACTION; MASS AND FORCE

We first dispose of the question of what frame of reference should be used to observe the particle motions mentioned in the Second Law. The answer is that any **inertial reference frame** can be used and we will always assume this to be so, unless stated otherwise. As stated earlier, the problem in understanding the Second and Third Laws is that the concepts of mass and force are not defined, which is obviously unsatisfactory.

Our second thought experiment is concerned with the motion of a pair of **mutually interacting particles**. The nature of their mutual interaction can be of any kind* and all other influences are removed. Since each particle is influenced by the other, the First Law does not apply. The particles will, in general, have accelerations, these being independent of the inertial frame in which they are measured. Our second law of physics is concerned with the 'observed' values of these mutually induced accelerations.

The law of mutual interaction *Suppose that two particles P_1 and P_2 interact with each other and that P_2 induces an instantaneous acceleration $\boldsymbol{a}_{12}$ in P_1, while P_1 induces an instantaneous acceleration $\boldsymbol{a}_{21}$ in P_2. Then*

(i) *these accelerations are opposite in direction and parallel to the straight line joining P_1 and P_2,*

---

* The mutual interaction might be, for example, (i) mutual gravitation, (ii) electrostatic interaction, caused by the particles being electrically charged, or (iii) the particles being connected by a fine elastic cord.

(ii) *the ratio of the magnitudes of these accelerations,* $|\boldsymbol{a}_{21}|/|\boldsymbol{a}_{12}|$ *is a constant independent of the nature of the mutual interaction between* P_1 *and* P_2*, and independent of the positions and velocities*† *of* P_1 *and* P_2.

Moreover, suppose that when P_2 *interacts with a third particle* P_3 *the induced accelerations are* $\boldsymbol{a}_{23}$ *and* $\boldsymbol{a}_{32}$*, and when* P_1 *interacts with* P_3 *the induced accelerations are* $\boldsymbol{a}_{13}$ *and* $\boldsymbol{a}_{31}$*. Then the magnitudes of these accelerations satisfy the consistency relation**

$$\frac{|\boldsymbol{a}_{21}|}{|\boldsymbol{a}_{12}|} \times \frac{|\boldsymbol{a}_{32}|}{|\boldsymbol{a}_{23}|} \times \frac{|\boldsymbol{a}_{13}|}{|\boldsymbol{a}_{31}|} = 1. \tag{3.1}$$

Definition of inertial mass

The law of mutual interaction leads us to our definitions of mass and force. *The qualitative definition of the (inertial) mass of a particle is that it is a numerical measure of the reluctance of the particle to being accelerated.* Thus, when particles P_1 and P_2 interact, we attribute the fact that the induced accelerations $\boldsymbol{a}_{12}$ and $\boldsymbol{a}_{21}$ have different magnitudes to the particles having different masses. This point of view is supported by the fact that the ratio $|\boldsymbol{a}_{21}|/|\boldsymbol{a}_{12}|$ depends only upon the particles themselves, and not on the interaction, or where the particles are, or how they are moving. We define the mass ratio m_1/m_2 of the particles P_1, P_2 to be the inverse ratio of the magnitudes of their mutually induced accelerations, as follows:

Definition 3.2 ***Inertial mass*** *The* **mass ratio** m_1/m_2 *of the particles* P_1, P_2 *is defined to be*

$$\frac{m_1}{m_2} = \frac{|\boldsymbol{a}_{21}|}{|\boldsymbol{a}_{12}|}. \tag{3.2}$$

There is however a possible inconsistency in this definition of mass ratio. Suppose that we introduce an third particle P_3. Then, by performing three experiments, we could *independently* determine the three mass ratios m_1/m_2, m_2/m_3 and m_3/m_1 and there is no guarrantee that the product of these three ratios would be unity. However, the **consistency relation** (3.1) assures us that it would be found to be unity, and this means that the above definition defines the mass ratios of particles unambiguously.

In order to have a numerical measure of mass, we simply choose some particle A as the reference mass (having mass one unit), in which case the mass of any other particle can be expressed as a number of 'A-units'. If we were to use a different particle B as the reference mass, we would obtain a second measure of mass in B-units, but this second measure would just be proportional to the first, differing only by a multiplied constant. In SI units, the reference body (having mass one kilogram) is a cylinder of platinum iridium alloy kept under carefully controlled conditions in Paris.

† This is true when relativistic effects are negligible.
* The significance of the consistency relation will be explained shortly.

Example 3.2 *A strange definition of mass*

Suppose the mass ratio m_2/m_1 were defined in some other way, such as

$$\frac{m_1}{m_2} = \left(\frac{|\boldsymbol{a}_{21}|}{|\boldsymbol{a}_{12}|}\right)^{1/2}.$$

Is this just as good as the standard definition?

Solution

For some purposes it would be just as good. It would lead to the non-standard form

$$\boldsymbol{F} = m^2\boldsymbol{a}$$

for the second law, and, for the motion of a single particle, the theory would be essentially unaffected. We will see later however that, if mass were defined in this way, then the mass of a multi-particle system would *not* be equal to the sum of the masses of its constituent particles! This is not contradictory, but it is a very undesirable feature and explains why the standard definition is used. ■

Definition of force

We now turn to the definition of force. *Qualitatively, the presence of a force is the reason we give for the acceleration of a particle*. Thus, when interacting particles cause each other to accelerate, we say it is because they *exert forces upon each other*. How do we know that these forces are present? Because the particles are accelerating! These statements are obviously circular and without real content. Force is therefore a quantity of our own invention, but a very useful one nonetheless and an essential part of the Newtonian formulation of classical mechanics. It should be noted though that the concept of force is not an *essential* part of the Lagrangian or Hamiltonian formulations of classical mechanics.*

In mutual interactions, the forces that the particles exert upon each other are defined as follows:

Definition 3.3 ***Force*** *Suppose that the particles P_1 and P_2 are in mutual interaction and have accelerations $\boldsymbol{a}_{12}$ and $\boldsymbol{a}_{21}$ respectively. Then the force $\boldsymbol{F}_{12}$ that P_2 exerts on P_1, and the force $\boldsymbol{F}_{21}$ that P_1 exerts on P_2 are defined to be*

$$\boldsymbol{F}_{12} = m_1\boldsymbol{a}_{12}, \qquad \boldsymbol{F}_{21} = m_2\boldsymbol{a}_{21}, \tag{3.3}$$

where the unit of force is implied by the units of mass and acceleration.

It follows that, in the case of two-particle interactions, the Second Law is true by the definition of force. Also, since $\boldsymbol{a}_{12}$ and $\boldsymbol{a}_{21}$ are opposite in direction and are parallel

* This fact is important when making connections between classical mechanics and other theories, such as general relativity or quantum mechanics. The concept of force does not appear in either of these theories.

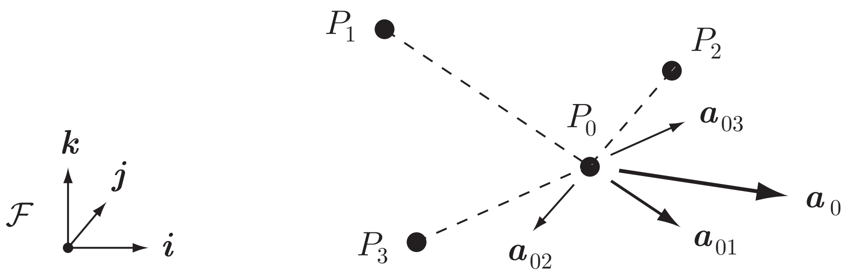


FIGURE 3.2 The **law of multiple interactions**. In the presence of interactions from the particles P_1, P_2, P_3, the acceleration $\boldsymbol{a}_0$ of particle P_0 is given by $\boldsymbol{a}_0 = \boldsymbol{a}_{01} + \boldsymbol{a}_{02} + \boldsymbol{a}_{03}$.

to the line P_1P_2, so then are $\boldsymbol{F}_{12}$ and $\boldsymbol{F}_{21}$; thus parts (ii) and (iii) of the Third Law are automatically true. Furthermore

$$|\boldsymbol{F}_{12}| = m_1|\boldsymbol{a}_{12}| = m_2|\boldsymbol{a}_{21}| = |\boldsymbol{F}_{21}|,$$

on using the definition (3.2) of the mass ratio m_1/m_2. Thus part (i) of the Third Law is also true. Hence the *law of mutual interaction, together with our definitions (3.2), (3.3) of mass and force, implies the truth of the Second and Third Laws.*

3.4 THE LAW OF MULTIPLE INTERACTIONS

Our third and final thought experiment is concerned with what happens when a particle is subject to more than one interaction.

The law of multiple interactions *Suppose the particles $P_0, P_1, \ldots P_n$ are interacting with each other and that all other influences are removed. Then the acceleration $\boldsymbol{a}_0$ induced in P_0 can be expressed as*

$$\boldsymbol{a}_0 = \boldsymbol{a}_{01} + \boldsymbol{a}_{02} + \cdots + \boldsymbol{a}_{0N}, \tag{3.4}$$

where $\boldsymbol{a}_{01}, \boldsymbol{a}_{02} \ldots$ are the accelerations that P_0 would have if the particles $P_1, P_2, \ldots$ were individually interacting with P_0.

This result is sometimes expressed by saying that **interaction forces act independently** of each other. It follows that

$$\begin{aligned} m_0\boldsymbol{a}_0 &= m_0\,(\boldsymbol{a}_{01} + \boldsymbol{a}_{02} + \cdots + \boldsymbol{a}_{0N}) \\ &= \boldsymbol{F}_{01} + \boldsymbol{F}_{02} + \cdots + \boldsymbol{F}_{0N}, \end{aligned}$$

on using the definition (3.3) of mutual interaction forces. Thus the *Second Law remains true for multiple interactions provided that the 'effective force' $\boldsymbol{F}_0$ acting on P_0 is understood to mean the (vector) resultant of the individual interaction forces acting on P_0*, that is

$$\boldsymbol{F}_0 = \boldsymbol{F}_{01} + \boldsymbol{F}_{02} + \cdots + \boldsymbol{F}_{0N}.$$

This result is not always thought of as a law of physics, but it is.* It *could* have been otherwise!

> **Experimental basis of Newton's Laws**
>
> **1.** We accept the **law of inertia**, the **law of mutual interaction** and the **law of multiple interactions** as the 'experimental' basis of mechanics.
>
> **2.** Together with our definitions of mass and force, these experimental laws imply that **Newton's laws are true in any inertial reference frame**.

3.5 CENTRE OF MASS

We can now introduce the notion of the centre of mass of a collection of particles. Suppose we have a system of particles $P_1, P_2, \ldots, P_N$ with masses $m_1, m_2, \ldots, m_N$, and position vectors $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N$ respectively. Then:

Definition 3.4 ***Centre of mass*** *The* ***centre of mass*** *of this system of particles is the* ***point of space*** *whose position vector* $\boldsymbol{R}$ *is defined by*

$$\boldsymbol{R} = \frac{m_1\boldsymbol{r}_1 + m_2\boldsymbol{r}_2 + \cdots + m_N\boldsymbol{r}_N}{m_1 + m_2 + \cdots + m_N} = \frac{\sum_{i=1}^N m_i\boldsymbol{r}_i}{\sum_{i=1}^N m_i} = \frac{\sum_{i=1}^N m_i\boldsymbol{r}_i}{M}, \tag{3.5}$$

where M *is the sum of the separate masses.*

The centre of mass of a system of particles is simply a 'weighted' mean of the position vectors of the particles, where the 'weights' are the particle masses. Centre of mass is an important concept in the mechanics of multi-particle systems. Unfortunately, there is a widespread belief that the centre of mass has a magical ability to describe the behaviour of the system in all circumstances. This is simply not true. For instance, we will show in the next section that it is *not* generally true that the total gravitational force that a system of masses exerts on a test mass is equal to the force that would be exerted by a particle of mass M situated at the centre of mass.

Example 3.3 *Finding centres of mass*

Find the centre of mass of (i) a pair of particles of different masses, (ii) three identical particles.

Solution

(i) For a pair of particles P_1, P_2, the position vector of the centre of mass is given by

$$\boldsymbol{R} = \frac{m_1\boldsymbol{r}_1 + m_2\boldsymbol{r}_2}{m_1 + m_2}.$$

* It certainly does not follow from the observation that 'forces are vector quantities'!

It follows that the centre of mass lies on the line P_1P_2 and divides this line in the ratio $m_2 : m_1$.

(ii) For three identical particles P_1, P_2, P_3, the position vector of the centre of mass is given by

$$\boldsymbol{R} = \frac{m\boldsymbol{r}_1 + m\boldsymbol{r}_2 + m\boldsymbol{r}_3}{m + m + m} = \frac{\boldsymbol{r}_1 + \boldsymbol{r}_2 + \boldsymbol{r}_3}{3}.$$

It follows that the centre of mass lies at the centroid of the triangle $P_1P_2P_3$. ■

The centres of mass of most of the systems we meet in mechanics are easily determined by symmetry considerations. However, when symmetry is lacking, the position of the centre of mass has to be worked out from first principles by using the definition (3.5), or its counterpart for continuous mass distributions. The Appendix at the end of the book contains more details and examples.

3.6 THE LAW OF GRAVITATION

Physicists recognise only four distinct kinds of interaction forces that exist in nature. These are gravitational forces, electromagnetic forces and weak/strong nuclear forces. The nuclear forces are important only within the atomic nucleus and will not concern us at all. The electromagnetic forces include electrostatic attraction and repulsion, but we will encounter them mainly as 'forces of contact' between material bodies. Since such forces are intermolecular, they are ultimately electromagnetic although we will make no use of this fact! The present section however is concerned with **gravitation**.

It is an observed fact that any object with mass attracts any other object with mass with a force called gravitation. When gravitational interaction occurs between particles, the Third Law implies that the interaction forces must be equal in magnitude, opposite in direction and parallel to the straight line joining the particles. The *magnitude* of the gravitational interaction forces is given by:

The law of gravitation

The gravitational forces that two particles exert upon each other each have magnitude

$$\frac{m_1 m_2 G}{R^2}, \tag{3.6}$$

where m_1, m_2 are the particle masses, R is the distance between the particles, and G, the *constant of gravitation*, is a universal constant. Since G is not dimensionless, its numerical value depends on the units of mass, length and force.

This is the famous **inverse square law of gravitation** originally suggested by Robert Hooke,* a scientific contemporary (and adversary) of Newton. In SI units, the constant of gravitation is given approximately by

$$G = 6.67 \times 10^{-11} \quad \mathrm{N\,m^2\,kg^{-2}}, \tag{3.7}$$

this value being determined by observation and experiment. There is presently no theory (general relativity included) that is able to predict the value of G. Indeed, the theory of general relativity does not exclude *repulsion* between masses!

To give some idea of the magnitudes of the forces involved, suppose we have two uniform spheres of lead, each with mass 5000 kg (five metric tons). Their common radius is about 47 cm which means that they can be placed with their centres 1 m apart. What gravitational force do they exert upon each other when they are in this position? We will show later that the gravitational force between uniform spheres of matter is exactly the same *as if* all the mass of each sphere were concentrated at its centre. Given that this result is true, we can find the force that each sphere exerts on the other simply by substituting $m_1 = m_2 = 5000$ and $R = 1$ into equation (3.6). This gives $F = 0.00167$ N approximately, the weight of a few grains of salt! Such forces seem insignificant, but gravitation is the force that keeps the Moon in orbit around the Earth, and the Earth in orbit around the Sun. The reason for this disparity is that the masses involved are so much larger than those of the lead spheres in our example. For instance, the mass of the Sun is about 2×10^{30} kg.

3.7 GRAVITATION BY A DISTRIBUTION OF MASS

It is important to be able to calculate the gravitational force exerted on a particle by a *distribution* of mass, such as a disc or sphere. The Earth, for example, is an approximately spherical mass distribution. We first treat an introductory problem of gravitational attraction by a pair of *particles* and then progress to *continuous distributions* of matter. In all cases, the law of multiple interactions means that the effective force exerted on a particle is the resultant of the individual forces of interaction exerted on that particle.

Example 3.4 *Attraction by a pair of particles*

A particle C, of mass m, and two particles A and B, each of mass M, are placed as shown in Figure 3.3. Find the gravitational force exerted on the particle C.

Solution

By the law of gravitation, each of the particles A and B attracts C with a force of magnitude F' where

$$F' = \frac{mMG}{R^2},$$

* It was Newton however who proved that Kepler's laws of planetary motion follow from the inverse square law.

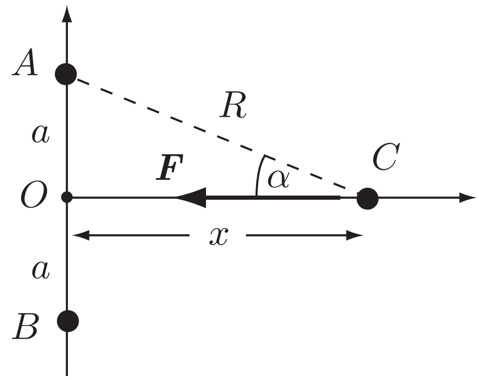


FIGURE 3.3 The particle C, of mass m, is attracted by the particles A and B, each of mass M. The resultant force on C points towards O.

where $R = (a^2 + x^2)^{1/2}$ is the distance AC $(= BC)$. By symmetry, the resultant force $\boldsymbol{F}$ points in the direction CO and so its magnitude F can be found by summing the components of the contributing forces in this direction. Hence

$$F = \frac{2mMG\cos\alpha}{R^2} = 2mMG\left(\frac{R\cos\alpha}{R^3}\right) = 2mMG\left(\frac{x}{(a^2+x^2)^{3/2}}\right)$$

for $x \geq 0$. [The angle α is shown in Figure 3.3.] Thus the resultant force exerted on C looks nothing like the force exerted by a *single* gravitating particle. In particular, it is *not* equal to the force that would be exerted by a mass $2M$ placed at O. However, on writing F in the form

$$F = \frac{2mMG}{x^2}\left(1 + \frac{a^2}{x^2}\right)^{-3/2},$$

we see that

$$F \sim \frac{m(2M)G}{x^2}$$

when x/a is large. Thus, when C is very distant from A and B, the gravitational force exerted on C is *approximately* the same as that of a single particle of mass $2M$ situated at O.

The graph of the *exact* value of F as a function of x is shown in Figure 3.4. Dimensionless variables are used. $F = 0$ when $x = 0$, and rises to a maximum when $x = a/\sqrt{2}$ where $F = 4mMG/3\sqrt{3}a^2$. Thereafter, F decreases, becoming ever closer to its asymptotic form $m(2M)G/x^2$. ∎

General asymptotic form of F as $r \to \infty$

The asymptotic result in the last example is true for attraction by any bounded* distribution of mass. The general result can be stated as follows:

* This excludes mass distributions that extend to infinity, such as an infinite straight wire.

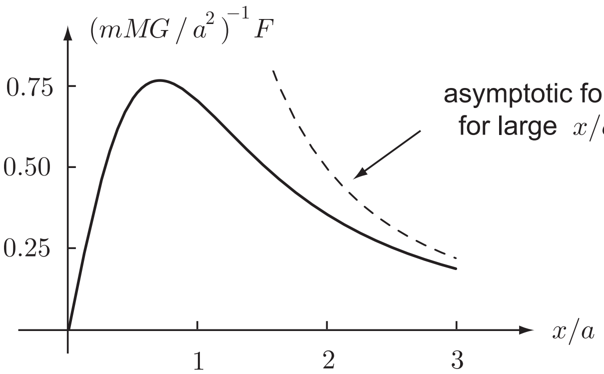


FIGURE 3.4 The dimensionless resultant force $(mMG/a^2)^{-1}F$ plotted against x/a.

Let $\mathcal{S}$ be any bounded system of masses with total mass M. Then the force $\boldsymbol{F}$ exerted by $\mathcal{S}$ on a particle P, of mass m and position vector $\boldsymbol{r}$, has the asymptotic form*

$$\boldsymbol{F} \sim -\frac{mMG}{r^2}\widehat{\boldsymbol{r}},$$

as $r \to \infty$, where $r = |\boldsymbol{r}|$ and $\widehat{\boldsymbol{r}} = \boldsymbol{r}/r$.

In other words, the force exerted by $\mathcal{S}$ on a distant particle is *approximately* the same as that exerted by a particle of mass equal to the total mass of $\mathcal{S}$, situated at the centre of mass of $\mathcal{S}$.

Example 3.5 *Gravitation by a uniform rod*

A particle P, of mass m, and a uniform rod, of length $2a$ and mass M, are placed as shown in Figure 3.5. Find the gravitational force that the rod exerts on the particle.

Solution

Consider the element $[x,\ x + dx]$ of the rod which has mass $M\,dx/2a$ and exerts an attractive force of magnitude

$$\frac{m(M\,dx/2a)G}{R^2}$$

on P, where R is the distance shown in Figure 3.5. By symmetry, the resultant force acts towards the centre O of the rod and can be found by summing the components of the contributing forces in the direction PO. Since the rod is a *continuous* distribution

* The result, as stated, is true for any choice of the origin O of position vectors. However, the asymptotic error is least if O is located at the centre of mass of $\mathcal{S}$. In this case the *relative* error is of order $(a/r)^2$, where a is the maximum 'radius' of the mass distribution about O.

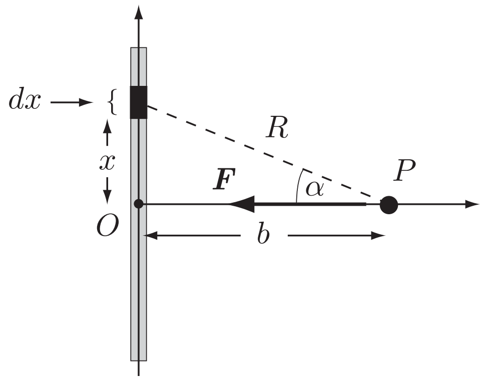


FIGURE 3.5 A particle P, of mass m, is attracted by a uniform rod of length $2a$ and mass M. The resultant force $\boldsymbol{F}$ on P points towards the centre O of the rod.

of mass, this sum becomes an integral. The resultant force exerted by the rod thus has magnitude F given by

$$\begin{aligned} F &= \frac{mMG}{2a}\int_{-a}^{a}\frac{\cos\alpha}{R^2}\,dx = \frac{mMG}{2a}\int_{-a}^{a}\frac{R\cos\alpha}{R^3}\,dx \\ &= \frac{mMG}{2a}\int_{-a}^{a}\frac{b\,dx}{(x^2+b^2)^{3/2}}, \end{aligned}$$

where b is the distance of P from the centre of the rod. This integral can be evaluated by making the substitution $x = b\tan\theta$, the limits on θ being $\theta = \pm\beta$, where $\tan\beta = a/b$. This gives

$$\begin{aligned} F &= \frac{mMG}{2a}\left(\frac{2\sin\beta}{b}\right) = \frac{mMG}{2a}\left(\frac{2a}{b(b^2+a^2)^{1/2}}\right) \\ &= \frac{mMG}{b(b^2+a^2)^{1/2}}. \ \blacksquare \end{aligned}$$

Example 3.6 *Gravitation by a uniform disk*

A particle P, of mass m, is situated on the axis of a uniform disk, of mass M and radius a, as shown in Figure 3.6. Find the gravitational force that the disk exerts on the particle.

Solution

Consider the element of area dA of the disk which has mass $M\,dA/\pi a^2$ and attracts P with a force of magnitude

$$\frac{m(M\,dA/\pi a^2)G}{R^2},$$

where R is the distance shown in Figure 3.6. By symmetry, the resultant force acts towards the centre O of the disk and can be found by summing the components of the

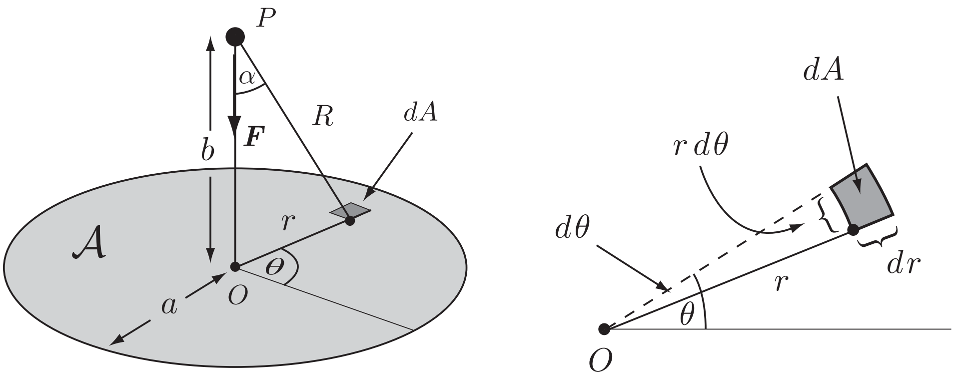


FIGURE 3.6 **Left**: A particle P, of mass m, is attracted by a uniform disk of mass M and radius a. The resultant force $\boldsymbol{F}$ on P points towards the centre O of the disk. **Right**: The element of area dA in polar cordinares r, θ.

contributing forces in the direction PO. The resultant force exerted by the disk thus has magnitude F given by

$$F = \frac{mMG}{\pi a^2} \int_{\mathcal{A}} \frac{\cos\alpha}{R^2}\, dA,$$

where the integral is to be taken over the region $\mathcal{A}$ occupied by the disk. This integral is most easily evaluated using polar coordinates. In this case $dA = (dr)(r\,d\theta) = r\,dr\,d\theta$, and the integrand becomes

$$\frac{\cos\alpha}{R^2} = \frac{R\cos\alpha}{R^3} = \frac{b}{(r^2+b^2)^{3/2}},$$

where b is the distance of P from the centre of the disk. The ranges of integration for r, θ are $0 \le r \le a$ and $0 \le \theta \le 2\pi$. We thus obtain

$$F = \frac{mMG}{\pi a^2} \int_{r=0}^{r=a} \int_{\theta=0}^{\theta=2\pi} \left(\frac{b}{(r^2+b^2)^{3/2}} \right) r\,dr\,d\theta.$$

Since the integrand is independent of θ, the θ-integration is trivial leaving

$$\begin{aligned} F &= \frac{mMG}{\pi a^2} \int_{r=0}^{r=a} \frac{2\pi b r\,dr}{(r^2+b^2)^{3/2}} = \frac{2mMG}{a^2} \left[-b(r^2+b^2)^{-1/2} \right]_{r=0}^{r=a} \\ &= \frac{2mMG}{a^2} \left[1 - \frac{b}{(a^2+b^2)^{1/2}} \right]. \blacksquare \end{aligned}$$

Gravitation by spheres

Because of its applications to astronomy and space travel, and because we live on a nearly spherical body, gravitation by a spherical mass distribution is easily the most important case. We suppose that the mass distribution occupies a spherical volume and is also **spherically symmetric** so that the *mass density depends only on distance from the centre of the*

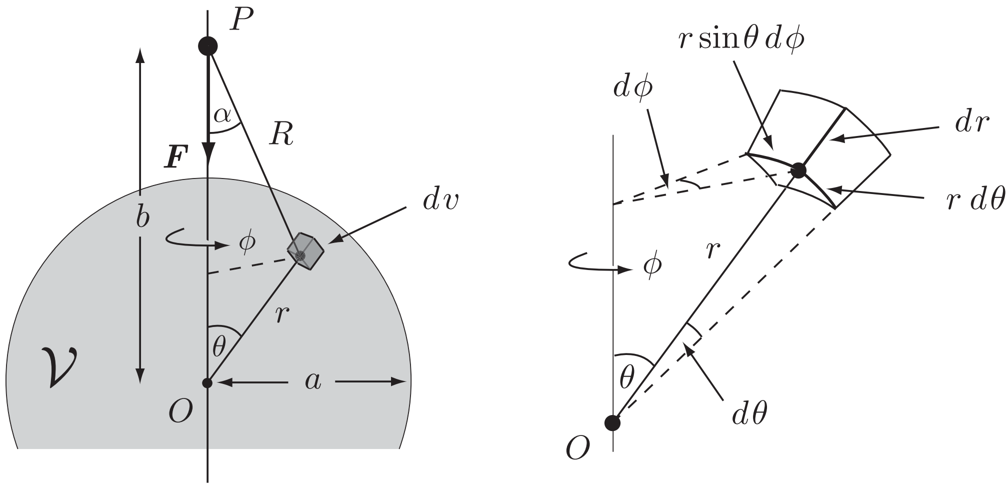


FIGURE 3.7 **Left**: A particle P, of mass m, is attracted by a symmetric sphere of radius a and total mass M. **Right**: The element of volume dV in spherical polar coordinates r, θ, ϕ.

sphere. We call such a body a **symmetric sphere**. The fact that we do not require the density to be uniform is very important in practical applications. The Earth, for instance, has a density of about 3,000 kg m^{-3} near its surface, but its density at the centre is about 16, 500kg m^{-3}. Similar remarks apply to the Sun. Thus, if our results were restricted to spheres of uniform density, they would not apply to the Earth or the Sun, the two most important cases.

The fundamental result concerning gravitation by a symmetric sphere was proved by Newton himself and confirmed his universal theory of gravitation. It is presented here as a theorem.

Theorem 3.1 *The gravitational force exerted by a symmetric sphere of mass M on a particle external to itself is* ***exactly*** *the same as if the sphere were replaced by a particle of mass M located at the centre.*

Proof. Figure 3.7 shows a symmetric sphere with centre O and radius a, and a particle P, of mass m, exterior to the sphere. We wish to calculate the force exerted by the sphere on the particle. The calculation is similar to that in the 'disk' example, but this time the integration must be carried out over the spherical volume occupied by the mass distribution.

Consider the element of volume dv of the sphere which has mass $\rho\, dv$ and attracts P with a force of magnitude

$$\frac{m(\rho\, dv)G}{R^2},$$

where R is the distance shown in Figure 3.7. By symmetry, the resultant force acts towards the centre O of the sphere and can be found by summing the components of the contributing forces in the direction PO. The resultant force exerted by the sphere thus has magnitude F

given by

$$F = mG \int_{\mathcal{V}} \frac{\rho \cos\alpha}{R^2}\, dv,$$

where the integral is to be taken over the region $\mathcal{V}$ occupied by the sphere. This integral is most easily evaluated using spherical polar coordinates r, θ, ϕ. In this case $dv = (dr)(r\,d\theta)(r\sin\theta\,d\phi) = r^2 \sin\theta\,dr\,d\theta\,d\phi$, and the integrand becomes

$$\frac{\rho\cos\alpha}{R^2} = \frac{\rho R\cos\alpha}{R^3} = \frac{\rho(r)\,(b - r\cos\theta)}{(r^2 + b^2 - 2rb\cos\theta)^{3/2}},$$

on using the cosine rule $R^2 = r^2 + b^2 - 2rb\cos\theta$, where b is the distance of P from the centre of the sphere. The ranges of integration for r, θ, ϕ are $0 \le r \le a$, $0 \le \theta \le \pi$ and $0 \le \phi \le 2\pi$. We thus obtain

$$F = mG \int_{r=0}^{r=a} \int_{\theta=0}^{\theta=\pi} \int_{\phi=0}^{\phi=2\pi} \left(\frac{\rho(r)\,(b - r\cos\theta)}{(r^2 + b^2 - 2rb\cos\theta)^{3/2}} \right) r^2 \sin\theta\,dr\,d\theta\,d\phi.$$

This time the ϕ-integration is trivial leaving

$$\begin{aligned} F &= mG \int_{r=0}^{r=a} \int_{\theta=0}^{\theta=\pi} \frac{2\pi\rho(r)\,(b - r\cos\theta)}{(r^2 + b^2 - 2rb\cos\theta)^{3/2}}\, r^2 \sin\theta\,dr\,d\theta \\ &= 2\pi mG \int_{r=0}^{r=a} r^2 \rho(r) \left\{ \int_{\theta=0}^{\theta=\pi} \frac{(b - r\cos\theta)\sin\theta\,d\theta}{(r^2 + b^2 - 2rb\cos\theta)^{3/2}} \right\} dr, \end{aligned}$$

on taking the θ-integration first and the r-integration second.

The θ-integration is tricky if done directly, but it comes out nicely on making the change of variable from θ to R given by

$$R^2 = r^2 + b^2 - 2rb\cos\theta, \qquad (R > 0).$$

(In this change of variable, r has the status of a constant.) The range of integration for R is $b - r \le R \le b + r$. Then

$$2R\,dR = 2rb\sin\theta\,d\theta,$$

$$b - r\cos\theta = \frac{2b^2 - 2rb\cos\theta}{2b} = \frac{R^2 + (b^2 - r^2)}{2b},$$

and the θ-integral becomes

$$\int_{b-r}^{b+r} \left(\frac{R^2 + (b^2 - r^2)}{2bR^3} \right) \frac{R\,dR}{rb} = \frac{1}{2rb^2} \int_{b-r}^{b+r} \left(1 + \frac{b^2 - r^2}{R^2} \right) dR = \frac{2}{b^2},$$

on performing the now elementary integration.

Hence

$$F = \frac{mG}{b^2} \left(4\pi \int_{r=0}^{r=a} r^2 \rho(r)\,dr \right)$$

and this is as far as we can go without knowing the density function $\rho(r)$. The answer that we are looking for is that $F = mMG/b^2$, where M is the total mass of the sphere. Now M can also be calculated as a volume integral. Since the mass of the volume element dv is $\rho\, dv$, the total mass M is given by

$$M = \int_{\mathcal{V}} \rho\, dv = \int_{r=0}^{r=a} \int_{\theta=0}^{\theta=\pi} \int_{\phi=0}^{\phi=2\pi} r^2 \rho(r) \sin\theta\, dr\, d\theta\, d\phi$$
$$= 4\pi \int_{r=0}^{r=a} r^2 \rho(r)\, dr,$$

on performing the θ- and ϕ-integrations.

Hence, we finally obtain

$$F = \frac{mMG}{b^2},$$

which is the required result.

Since there is no reason why the density $\rho(r)$ should not be zero over part of its range, this result also applies to the case of a particle external to a hollow sphere. The case of a particle *inside* a hollow sphere is different (see Problem 3.5). ■

Spheres attracted by other spheres

Since any element of mass is attracted by a symmetric sphere as if the sphere were a particle, it follows that the force that a symmetric sphere exerts on *any other mass distribution* can be calculated by replacing the sphere by a particle of equal mass located at its centre. In particular then, the force that two symmetric spheres exert upon each other is the same as if *each* sphere were replaced by its equivalent particle. Thus, as far as the forces of gravitational attraction are concerned, *symmetric spheres behave exactly as if they were particles*.

3.8 THE PRINCIPLE OF EQUIVALENCE AND g

Although we have so far not mentioned it, the law of gravitation hides a deep and very surprising fact, namely, that *the force between gravitating particles is proportional to each of their inertial masses*. Now **inertial mass**, as defined by equation (3.2), has no necessary connection with gravitation. It is a measure of the reluctance of that particle to being accelerated and can be determined by non-gravitational means, for instance, by using electrostatic interactions between the particles. It is a matter of extreme surprise then that a quantity that seems to have no necessary connection with gravitation actually *determines* the force of gravitation between particles. What we would have expected was that each particle would have a second property m^*, called **gravitational mass** (not the same as m), which appears in the law of gravitation (3.6) and determines the gravitational force. For example, suppose that we have three uniform spheres of gold, silver and bronze and that the silver and bronze spheres have equal inertial mass. Then the law of gravitation states that, when separated by equal distances, the gold sphere will attract the silver and bronze spheres with *equal forces*, whereas we would have expected these forces to be different.

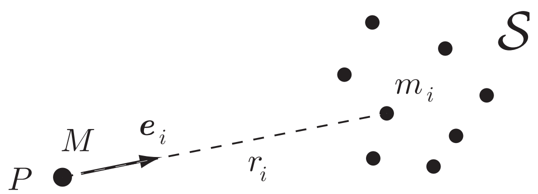


FIGURE 3.8 A particle of mass M is attracted by the gravitation of the system $\mathcal{S}$ which consists of N particles with masses $\{m_i\}$ $(1 \le i \le N)$.

The question arises then as to whether m and m^* are actually equal or just nearly equal so that the difference is difficult to detect. Newton himself did experiments with pendulums made of differing materials, but could not detect any difference in the period. Newton's experiment could have detected a difference of about one part in 10^3. However, the classic experiment of Eötvös (1890) and its later refinements have now shown that any difference between m and m^* is less than one part in 10^{11}. This leads us to believe that m and m^* really are equal and that the law of gravitation means exactly what it says.

The proposition that inertial and gravitational mass are *exactly* equal is called the **principle of equivalence**. Although we accept the principle of equivalence as being true, we still have no explanation why this is so! In this context, it is worth remarking that Einstein made the principle of equivalence into a fundamendal *assumption* of the theory of general relativity.

The gravitational acceleration g

Suppose a particle P of mass M is under the gravitational attraction of the system $\mathcal{S}$, as shown in Figure 3.8. Then, by the law of gravitation, the resultant force $\boldsymbol{F}$ that $\mathcal{S}$ exerts upon P is given by

$$\begin{aligned}\boldsymbol{F} &= \frac{Mm_1G}{r_1^2}\,\boldsymbol{e}_1 + \frac{Mm_2G}{r_2^2}\,\boldsymbol{e}_2 + \cdots + \frac{Mm_NG}{r_N^2}\,\boldsymbol{e}_N = M\left(\sum_{i=1}^{N}\frac{m_iG}{r_i^2}\,\boldsymbol{e}_i\right)\\ &= M\boldsymbol{g},\end{aligned}$$

where the vector $\boldsymbol{g}$, defined by

$$\boldsymbol{g} = \sum_{i=1}^{N}\frac{m_iG}{r_i^2}\,\boldsymbol{e}_i,$$

is *independent of* M. Then, by the Second Law, the induced acceleration $\boldsymbol{a}$ of particle P is determined by the equation

$$M\boldsymbol{g} = M\boldsymbol{a},$$

that is,

$$\boldsymbol{a} = \boldsymbol{g}.$$

Thus *the induced acceleration* $\boldsymbol{g}$ *is the same for any particle* situated at that point. This rather remarkable fact is a direct consequence of the principle of equivalence. Tradition has it that, prior to Newton, Galileo did experiments in which he released different masses from the top of the Tower of Pisa and found that they reached the ground at the same time. Galileo's result is thus a colourful but rather inaccurate verification of the principle of equivalence!

Gravitation by the Earth (rotation neglected)

In the present treatment, the rotation of the Earth is neglected and we regard the Earth as an inertial frame of reference. A more accurate treatment which takes the Earth's rotation into account is given in Chapter 17.

When the system $\mathcal{S}$ is the Earth (or some other celestial body) it is convenient to introduce the notion of the local **vertical** direction. The unit vector $\boldsymbol{k}$, which has the *opposite* direction to $\boldsymbol{g}$, is called the **vertically upwards unit vector** relative to the Earth. In terms of $\boldsymbol{k}$, the force exerted by the Earth on a particle of mass M is given by

$$\boldsymbol{F} = -Mg\boldsymbol{k},$$

where the **gravitational acceleration** g is the *magnitude* of the gravitational acceleration vector $\boldsymbol{g}$. Both g and $\boldsymbol{k}$ are functions of position on the Earth.

Weight

The positive quantity Mg (which is a function of position) is called the **weight** of the particle P. It is the *magnitude of the gravity force* exerted on P by the Earth. Thus the same body will have different weights depending upon where it is situated. However, at a fixed point of space, the weights of bodies are proportional to their masses. This fact, which is a consequence of the principle of equivalence, enables *masses* to be compared simply by comparing their *weights* at the same location (by using a balance, for instance).

The approximation of uniform gravity

It is easy to see that the Earth's gravitational acceleration g and the vertical direction $\boldsymbol{k}$ depend upon position. The Earth is approximately a symmetric sphere which exerts its gravitational force as if all its mass were at its centre. Thus, if the value of g at a point on the Earth's surface is g_1, then the value of g at a height of 6,400 km (the Earth's radius) must be $g_1/4$ approximately. On the other hand, the vertical vector $\boldsymbol{k}$ changes from point to point on the Earth's surface. These changes will be significant for motions whose extent is significant compared to the Earth's radius; this is true for a ballistic missile, for instance. However, most motions taking place on Earth have an extent that is insignificant compared to the Earth's radius and for which the variations of g and $\boldsymbol{k}$ are negligible. Simple examples include the motion of a tennis ball, a javelin or a bullet.

The approximation in which g and $\boldsymbol{k}$ are assumed to be constants is called **uniform gravity**. Uniform gravity is the most common force field in mechanics. Many of the problems solved in this book make this simplifying (and accurate) approximation.

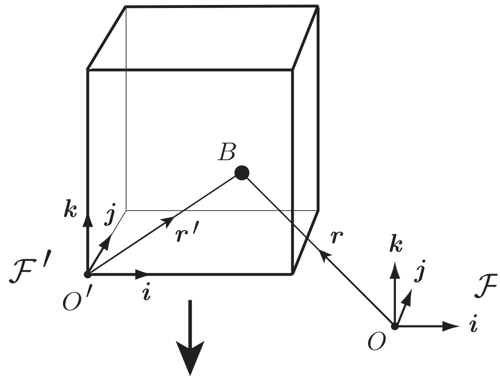


FIGURE 3.9 An elevator contains a ball B and both are freely falling under uniform gravity. $\mathcal{F}$ is an inertial reference frame and $\mathcal{F}'$ is a reference frame attached to the falling elevator.

Numerical values of g

The value of g at any location on the Earth can be measured experimentally (by using a pendulum for instance). The value of g is not quite constant over the Earth's surface since the Earth does not quite have spherical symmetry and different locations have differing altitudes. At sea level on **Earth**, $g = 9.8\ \mathrm{m\,s^{-2}}$ approximately, and a rough value of $10\ \mathrm{m\,s^{-2}}$ is often assumed. The corresponding value for the **Moon** is $1.6\ \mathrm{m\,s^{-2}}$, roughly a sixth of the Earth's value.

Example 3.7 *Particle inside a falling elevator*

An elevator cable has snapped and the elevator and its contents are falling under uniform gravity. One of the passengers takes a ball from his pocket and throws it to another passenger.* What is the motion of the ball *relative to the elevator*?

Solution

Suppose that the ball has mass m and that the local (vector) gravitational acceleration is $-g\boldsymbol{k}$. Then the motion of the ball relative to an *inertial* reference frame $\mathcal{F}$ (fixed to the ground, say) is determined by the Second Law, namely,

$$m\boldsymbol{a} = -mg\boldsymbol{k},$$

where $\boldsymbol{a}$ is the acceleration of the ball measured in $\mathcal{F}$.

Let the frame $\mathcal{F}'$ be attached to the elevator, as shown in Figure 3.9. Then the acceleration $\boldsymbol{a}'$ of the ball measured in $\mathcal{F}'$ is given (see section 2.6) by

$$\boldsymbol{a} = \boldsymbol{a}' + \boldsymbol{A},$$

where $\boldsymbol{A}$ is the acceleration of the frame $\mathcal{F}'$ relative to $\mathcal{F}$. But the elevator, to which the frame $\mathcal{F}'$ is attached, is also moving under uniform gravity and its acceleration $\boldsymbol{A}$ is therefore, by the principle of equivalence, the same as that of the ball, namely,

$$\boldsymbol{A} = -g\boldsymbol{k}.$$

* People do react oddly when put under pressure.

Hence $\boldsymbol{a} = \boldsymbol{a}' - g\boldsymbol{k}$ and so

$$\boldsymbol{a}' = \boldsymbol{0}.$$

Thus, relative to the elevator, the **ball moves with constant velocity**. To observers resident in the frame $\mathcal{F}'$, gravity appears to be absent and $\mathcal{F}'$ appears to be an inertial frame. This provides a practical method for simulating conditions of weightlessness. Fortunately for those wishing to experience weightlessness, there is no need to use an elevator; the same acceleration can be achieved by an aircraft in a vertical dive!

This result is of considerable importance in the theory of general relativity. It shows that, locally at least, a gravitational field can be 'transformed away' by observing the motion of bodies from a freely falling reference frame. ■

Problems on Chapter 3

Answers and comments are at the end of the book.

Harder problems carry a star (∗).

Gravitation

3.1 Four particles, each of mass m, are situated at the vertices of a regular tetrahedron of side a. Find the gravitational force exerted on any one of the particles by the other three.

Three uniform rigid spheres of mass M and radius a are placed on a horizontal table and are pressed together so that their centres are at the vertices of an equilateral triangle. A fourth uniform rigid sphere of mass M and radius a is placed on top of the other three so that all four spheres are in contact with each other. Find the gravitational force exerted on the upper sphere by the three lower ones.

3.2 Eight particles, each of mass m, are situated at the corners of a cube of side a. Find the gravitational force exerted on any one of the particles by the other seven.

Deduce the total gravitational force exerted on the four particles lying on one face of the cube by the four particles lying on the opposite face.

3.3 A uniform rod of mass M and length $2a$ lies along the interval $[-a, a]$ of the x-axis and a particle of mass m is situated at the point $x = x'$. Find the gravitational force exerted by the rod on the particle.

Two uniform rods, each of mass M and length $2a$, lie along the intervals $[-a, a]$ and $[b - a, b + a]$ of the x-axis, so that their centres are a distance b apart ($b > 2a$). Find the gravitational forces that the rods exert upon each other.

3.4 A uniform rigid disk has mass M and radius a, and a uniform rigid rod has mass M' and length b. The rod is placed along the axis of symmetry of the disk with one end in contact with the disk. Find the forces necessary to pull the disk and rod apart. [*Hint*. Make use of the solution in the 'disk' example.]

3.5 Show that the gravitational force exerted on a particle *inside* a hollow symmetric sphere is zero. [*Hint*. The proof is the same as for a particle *outside* a symmetric sphere, except in one detail.]

3.6 A narrow hole is drilled through the centre of a *uniform* sphere of mass M and radius a. Find the gravitational force exerted on a particle of mass m which is inside the hole at a distance r from the centre.

3.7 A symmetric sphere, of radius a and mass M, has its centre a distance b ($b > a$) from an infinite plane containing a uniform distribution of mass σ per unit area. Find the gravitational force exerted on the sphere.

3.8* Two uniform rigid hemispheres, each of mass M and radius a are placed in contact with each other so as to form a complete sphere. Find the forces necessary to pull the hemispheres apart.

Computer assisted problem

3.9 A uniform wire of mass M has the form of a circle of radius a and a particle of mass m lies in the plane of the wire at a distance b ($b < a$) from the centre O. Show that the gravitational force exerted by the wire on the particle (in the direction OP) is given by

$$F = \frac{mMG}{2\pi a^2} \int_0^{2\pi} \frac{(\cos\theta - \xi)d\theta}{\{1 + \xi^2 - 2\xi\cos\theta\}^{3/2}},$$

where the dimensionless distance $\xi = b/a$.

Use computer assistance to plot the graph of (dimensionless) F against ξ for $0 \leq \xi \leq 0.8$ and confirm that F is *positive* for $\xi > 0$. Is the position of equilibrium at the centre of the circle stable? Could the rings of Saturn be solid?

Chapter Four

Problems in particle dynamics

KEY FEATURES

The key features in this chapter are (i) the **vector equation of motion** and its reduction to **scalar equations**, (ii) motion in a **force field**, (iii) geometrical constraints and **forces of constraint**, and (iv) linear and quadratic **resistance forces**.

Particle dynamics is concerned with the problem of calculating the motion of a particle that is acted upon by specified forces. Our starting point is Newton's laws. However, since the First Law merely tells us that we should observe the motion from an inertial frame, and the Third Law will never be used (since there is only one particle), the entirety of particle dynamics is based on the Second Law

$$m\boldsymbol{a} = \boldsymbol{F}_1 + \boldsymbol{F}_2 + \cdots + \boldsymbol{F}_N,$$

where $\boldsymbol{F}_1, \boldsymbol{F}_2, \ldots, \boldsymbol{F}_N$ are the various forces that are acting on the particle. The typical method of solution is to write the Second Law in the form

$$m\frac{d\boldsymbol{v}}{dt} = \boldsymbol{F}_1 + \boldsymbol{F}_2 + \cdots + \boldsymbol{F}_N, \tag{4.1}$$

which is a first order ODE for the unknown velocity function $\boldsymbol{v}(t)$ and is called the **equation of motion** of the particle. If the initial value of $\boldsymbol{v}$ is given, then equation (4.1) can often be solved to yield $\boldsymbol{v}$ as a function of the time t. Once $\boldsymbol{v}$ is determined (and if the initial position of the particle is given), the position vector $\boldsymbol{r}$ of the particle at time t can be found by solving the first order ODE $d\boldsymbol{r}/dt = \boldsymbol{v}$. The sections that follow contain many examples of the implementation of this method. Indeed, it is remarkable how many interesting problems can be solved in this way.

Question ***When can real bodies be modelled as particles?***

Newton's laws apply to particles, but real bodies are not particles. When can real bodies, such as a tennis ball, a spacecraft, or the Earth, be treated as if they were particles?

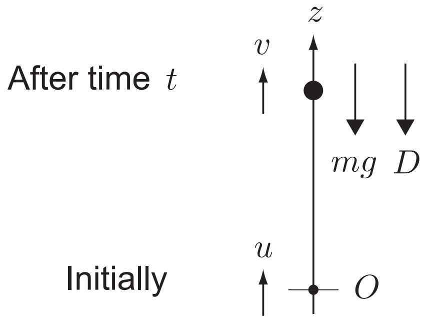


FIGURE 4.1 The particle is initially at the origin and is projected vertically upwards with speed u. The particle moves in a vertical straight line (the axis Oz) under the uniform gravity force mg and possibly a resistance (or drag) force D. At time t the particle has upward velocity v.

Answer

This is quite a tricky question which is not fully discussed until Chapter 10. What we will show is that the **centre of mass** of any body moves *as if* it were a particle of mass equal to the total mass, and all the forces on the body acted upon it. In particular, a rigid body, moving without rotation, can be treated *exactly* as if it were a particle. For example, a block sliding without rotation on a table can be treated exactly as if it were a particle. In other cases we gain only partial information about the motion. If the body is a brick thrown through the air, then particle dynamics can tell us exactly where its centre of mass will go, but not which point of the brick will hit the ground first. ■

4.1 RECTILINEAR MOTION IN A FORCE FIELD

Our first group of problems is concerned with the straight line motion of a particle moving in a force field. A force $\boldsymbol{F}$ is said to be a **field** if it depends only on the **position** of the particle, and not, for instance, on its velocity or the time. For example, the gravitational attraction of any *fixed* mass distribution is a field of force, but resistance forces, which are usually velocity dependent, are not.

If the rectlinear motion takes place along the z-axis, the equation of motion (4.1) reduces to the scalar equation

$$m\frac{dv}{dt} = F(z), \tag{4.2}$$

where v is the (one-dimensional) velocity of the particle and $F(z)$ is the (one-dimensional) force field, both measured in the *positive* z-direction.

First we consider the problem of **vertical motion** of a particle under **uniform gravity** with **no air resistance**. This is fine on the Moon (which has no atmosphere) but, on Earth, the motion of a body is resisted by its passage through the atmosphere and this will introduce errors. The effect of resistance forces is investigated in section 4.3.

Example 4.1 *Vertical motion under uniform gravity*

A particle is projected vertically upwards with speed u and moves in a vertical straight line under uniform gravity with no air resistance. Find the maximum height achieved by the particle and the time taken for it to return to its starting point.

Solution

Let v be the upwards velocity of the particle after time t, as shown in Figure 4.1. Then the scalar equation of motion (4.2) takes the form

$$m\frac{dv}{dt} = -mg,$$

since the drag force D is absent. A simple integration gives

$$v = -gt + C,$$

where C is the integration constant, and, on applying the initial condition $v = u$ when $t = 0$, we obtain $C = u$. Hence the **velocity** v at time t is given by

$$v = u - gt.$$

To find the upward displacement z at time t, write

$$\frac{dz}{dt} = v = u - gt.$$

A second simple integration gives

$$z = ut - \tfrac{1}{2}gt^2 + D,$$

where D a second integration constant, and, on applying the initial condition $z = 0$ when $t = 0$, we obtain $D = 0$. Hence the upward **displacement** of the body at time t is given by

$$z = ut - \tfrac{1}{2}gt^2.$$

The **maximum height** z_{max} is achieved when $dz/dt = 0$, that is, when $v = 0$. Thus z_{max} is achieved when $t = u/g$ and is given by

$$z_{\text{max}} = u\left(\frac{u}{g}\right) - \tfrac{1}{2}g\left(\frac{u}{g}\right)^2 = \frac{u^2}{2g}.$$

The particle returns to O when $z = 0$, that is, when

$$t\left(u - \tfrac{1}{2}gt\right) = 0.$$

Thus the **particle returns** after a time $2u/g$.

For example, if we throw a body vertically upwards with speed 10 m s^{-1}, it will rise to a height of 5 m and return after 2 s. [Here we are neglecting atmospheric resistance and taking $g = 10$ m s^{-2}.] ■

Question ***Saving oneself in a falling elevator***

An elevator cable has snapped and the elevator is heading for the ground. Can the occupants save themselves by leaping into the air just before impact in order to avoid the crash?

Answer

Suppose that the elevator is at rest at a height H when the cable snaps. The elevator will fall and reach the ground with speed $(2gH)^{1/2}$. In order to save themselves, the occupants must leap upwards (relative to the elevator) with this same speed so that their speed relative to the ground is zero. If they were able do this, then they would indeed be saved. However, if they were able to project themselves upwards with this speed, they would also be able to stand outside the building and leap up to the same height H that the elevator fell from! Even athletes cannot jump much more than a metre off the ground, so the answer is that escape is possible in principle but not in practice. ■

Uniform gravity is the simplest force field because it is constant. In the next example we show how to handle a **non-constant force field**.

Example 4.2 ***Rectilinear motion in the inverse square field***

A particle P of mass m moves under the gravitational attraction of a mass M fixed at the origin O. Initially P is at a distance a from O when it is projected with speed u directly away from O. Find the condition that P will 'escape' to infinity.

Solution

By symmetry, the motion of P takes place in a straight line through O. By the law of gravitation, the scalar equation of motion is

$$m\frac{dv}{dt} = -\frac{mMG}{r^2},$$

where r is the distance OP and $v = \dot{r}$. Equations like this can always be integrated once by first eliminating the time. Since

$$\frac{dv}{dt} = \frac{dv}{dr} \times \frac{dr}{dt} = v\frac{dv}{dr},$$

the equation of motion can be written as

$$v\frac{dv}{dr} = -\frac{MG}{r^2},$$

a first order ODE for v as a function of r. This is to be solved with the initial condition $v = u$ when $r = a$. The equation separates to give

$$\int v\, dv = -MG \int \frac{dr}{r^2},$$

and so

$$\tfrac{1}{2}v^2 = \frac{MG}{r} + C,$$

where C is the integration constant. On applying the initial condition $v = u$ when $r = a$, we find that $C = (u^2/2) - (MG/a)$ so that

$$v^2 = \left(u^2 - \frac{2MG}{a}\right) + \frac{2MG}{r}.$$

This determines the outward **velocity** v as a function of r.

Whether the particle escapes to infinity, or not, depends on the *sign* of the bracketed constant term.

(i) Suppose first that this term is *positive* so that

$$u^2 - \frac{2MG}{a} = V^2,$$

where V is a positive constant. Then, since the term $2MG/r$ is positive, it follows that $v > V$ at all times. It further follows that $r > a + Vt$ for all t and so the *particle escapes to infinity*.

(ii) On the other hand, if $u^2 - (2MG/a)$ is *negative*, then v becomes zero when

$$r = \frac{a}{1 - (u^2 a/2MG)},$$

after which the particle falls back towards O and *does not escape*.

(iii) The critical case, in which $u^2 = 2MG/a$ is treated in Problem 4.10; the result is that the *particle escapes*.

Hence the **particle escapes** if (and only if)

$$u^2 \ge \frac{2MG}{a}. \blacksquare$$

Question ***Given u, find r_{max} and the time taken to get there***

For the particular case in which $u^2 = MG/a$, find the maximum distance from O achieved by P and the time taken to reach this position.

Answer

For this value of u, the equation connecting v and r becomes

$$v^2 = MG\left(\frac{2}{r} - \frac{1}{a}\right).$$

Since $r = r_{\text{max}}$ when $v = 0$, it follows that the **maximum distance** from O achieved by P is $2a$.

To find the time taken, we write $v = dr/dt$ and solve the ODE

$$\left(\frac{dr}{dt}\right)^2 = MG\left(\frac{2}{r} - \frac{1}{a}\right)$$

with the initial condition $r = a$ when $t = 0$. After taking the positive square root of each side ($dr/dt \geq 0$ in this motion), the equation separates to give

$$\int_a^{2a} \left(\frac{ar}{2a - r}\right)^{1/2} dr = (MG)^{1/2} \int_0^{\tau} dt.$$

(Here we have introduced the initial and final conditions directly into the limits of integration; τ is the elapsed time.) On simplifying, we obtain

$$\tau = (MG)^{-1/2} \int_a^{2a} \left(\frac{ar}{2a - r}\right)^{1/2} dr.$$

This integral can be evaluated by making the substitution $r = 2a \sin^2 \theta$; the details are unimportant. The result is that the **time taken** for P to progress from $r = a$ to $r = 2a$ is

$$\tau = \left(\frac{a^3}{MG}\right)^{1/2} \left(1 + \tfrac{1}{2}\pi\right). \blacksquare$$

Question ***Speed of escape from the Moon***

A body is projected vertically upwards from the surface of the Moon. What projection speed is neccessary for the body to escape the Moon?

Answer

We regard the Moon as a fixed symmetric sphere of mass M and radius R. In this case, the gravitational force exerted by the Moon is the same as that of a particle of mass M situated at the centre. Thus the preceding theory applies with the distance a replaced by the radius R. The **escape speed** is therefore $(2MG/R)^{1/2}$, which evaluates to about 2.4 km s^{-1}. [For the Moon, $M = 7.35 \times 10^{22}$ kg and $R = 1740$ km.] $\blacksquare$

4.2 CONSTRAINED RECTILINEAR MOTION

Figure 4.2 shows a uniform rigid rectangular block of mass M sliding down the inclined surface of a *fixed* rigid wedge of angle α. The initial conditions are supposed to be such that the block slides, without rotation, down the line of steepest slope of the wedge. The block is subject to uniform gravity, but it is clear that there must be other forces as well. If there were no other forces and the block were released from rest, then the block would move vertically downwards. However, solid bodies cannot pass through each other like ghosts, and interpenetration is prevented by (equal and opposite) forces that they exert upon each other. These are **material contact forces** which come into play only when bodies are in physical contact. They are examples of **forces of constraint**, which are not prescribed beforehand but are sufficient to enforce a specified **geometrical constraint**. Tradition has it that the constraint force that the wedge exerts on the block is

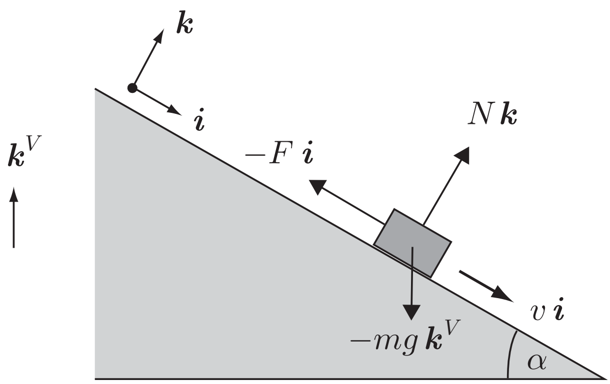


FIGURE 4.2 A rigid rectangluar block slides down the inclined surface of a fixed rigid wedge of angle α. Note that $\boldsymbol{k}^V$ is the vertically upwards unit vector, while $\boldsymbol{i}$ and $\boldsymbol{k}$ are parallel to and perpendicular to the inclined surface of the wedge.

called the total **reaction force $\boldsymbol{R}$**. It is convenient to write this force in the form

$$\boldsymbol{R} = -F\boldsymbol{i} + N\boldsymbol{k},$$

where the unit vectors $\boldsymbol{i}$ and $\boldsymbol{k}$ are parallel to and normal to the slope of the wedge. The scalar N is called the **normal reaction** component and the scalar F is called the **frictional** component.*

The equation of motion of the block is the vector equation (4.1) which becomes

$$M\frac{d(v\boldsymbol{i})}{dt} = -mg\boldsymbol{k}^V - F\boldsymbol{i} + N\boldsymbol{k},$$

where $\boldsymbol{k}^V$ is the vertically upwards unit vector. The easiest way of proceeding is to take components of this vector equation in the $\boldsymbol{i}$- and $\boldsymbol{k}$-directions (the $\boldsymbol{j}$-component gives nothing). On noting that $\boldsymbol{k}^V = -\sin\alpha\,\boldsymbol{i} + \cos\alpha\,\boldsymbol{k}$, this gives

$$M\frac{dv}{dt} = mg\sin\alpha - F \qquad \text{and} \qquad 0 = N - mg\cos\alpha.$$

The second of these equations determines the normal reaction $N = mg\cos\alpha$. However, in the first equation, both v and F are unknown and this prevents any further progress in the solution of this problem.* One can proceed by proposing some empirical 'law of friction', but such laws hold only very roughly. It is not surprising then that, in much of mechanics, frictional forces are neglected. In this case, the total reaction force exerted by the surface is in the normal direction and we describe such surfaces as **smooth**, meaning

* The minus sign is introduced so that F will be positive when the scalar velocity v is positive.
* This reflects the fact that we have said nothing about the roughness of the surface of the wedge!

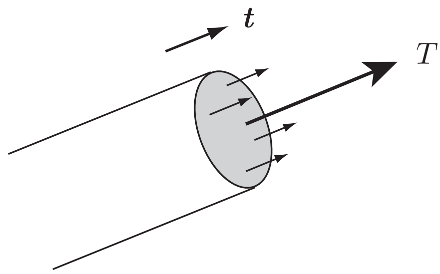


FIGURE 4.3 The idealised string is depicted here as having a small circular cross-section. At each cross-section only tensile stresses exist and their resultant is the tension T in the string at that point.

'perfectly smooth'. Doing away with friction has the advantage of giving us a well-posed problem that we can solve; however the solution will then apply only approximately to real surfaces.

If we now suppose that the inclined surface of the **wedge** is **smooth**, then $F = 0$ and the first equation reduces to

$$\frac{dv}{dt} = g \sin \alpha.$$

Thus, in the absence of friction, the block slides down the plane with the *constant acceleration* $g \sin \alpha$.

Inextensible strings

Another agency that can cause a geometrical constraint is the **inextensible string**. If a particle P of a system is connected to a fixed point O by an inextensible string of length a then, *if the string is taut*, P is constrained to move so that the distance $OP = a$. This geometrical constraint is enforced by the (unknown) constraint force that the string applies to particle P. Our 'string' is an idealisation of real cords and ropes in that it is *infinitely thin*, has *no bending stiffness*, and is *inextensible*. The only force that one part of the string exerts on another is the **tension** T in the string, which acts parallel to the tangent vector $\boldsymbol{t}$ to the string at each point (see Figure 4.3).

It is evident that, in general, T varies from point to point along the string. Suppose for example that a uniform string of mass ρ per unit length is suspended vertically under uniform gravity. Then, since the tension at the lower end is zero, the string will not be in equilibrium unless the tension at a height z above the lowest point is given by $T = \rho g z$; the tension thus rises linearly with height.

The situation is simpler when the mass of the string is negligible; this is the case of the **light inextensible string**.* In this case, it is obvious that the tension is constant when the string is straight. In fact, the tension also remains constant when the string slides over a *smooth* body. This is proved in Chapter 10. The tension in a light string is also constant when the string passes over a *light, smoothly pivoted* pulley wheel.

* In this context, 'light' means 'of zero mass'.

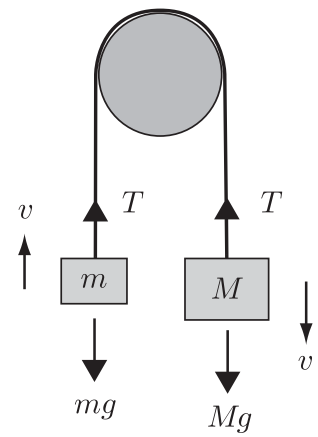


FIGURE 4.4 Attwood's machine: two bodies of masses m and M are connected by a light inextensible string which passes over a smooth rail.

Example 4.3 *Attwood's machine*

Two bodies with masses m, M are connected by a light inextensible string which passes over a smooth horizontal rail. The system moves in a vertical plane with the bodies moving in vertical straight lines. Find the upward acceleration of the mass m and the tension in the string.

Solution

The system is shown in Figure 4.4. Let v be the *upward* velocity of the mass m. Then, since the string is inextensible, v must also be the *downward* velocity of the mass M. Also, since the string is light and the rail is *smooth*, the string has constant tension T. The scalar equations of motion for the two masses are therefore

$$m\frac{dv}{dt} = T - mg, \qquad M\frac{dv}{dt} = Mg - T.$$

It follows that

$$\frac{dv}{dt} = \left(\frac{M-m}{M+m}\right) g \qquad \text{and} \qquad T = \left(\frac{2Mm}{M+m}\right) g. \blacksquare$$

Question *The monkey puzzle*

Suppose that, in the last example, both bodies have the same mass M and one of them is a monkey which begins to climb up the rope. What happens to the other mass?

Answer

Suppose that the monkey climbs with velocity V *relative to the rope*. Then its upward velocity relative to the ground is $V - v$. The equations of upward motion for the mass and the monkey are therefore

$$M\frac{dv}{dt} = T - Mg, \qquad M\frac{d(V-v)}{dt} = T - Mg.$$

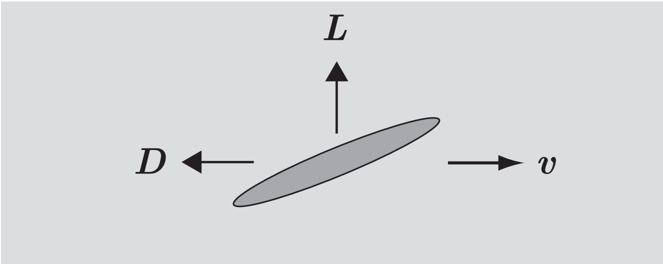


FIGURE 4.5 The drag $\boldsymbol{D}$ and lift $\boldsymbol{L}$ on a body moving through a fluid.

On eliminating T, we find that

$$\frac{dv}{dt} = \tfrac{1}{2}\frac{dV}{dt},$$

so that, if the whole system starts from rest,

$$v = \tfrac{1}{2}V, \qquad \text{and} \qquad V - v = \tfrac{1}{2}V.$$

Thus the monkey and the mass rise (relative to the ground) with the same velocity; the monkey cannot avoid hauling up the mass as well as itself! ■

4.3 MOTION THROUGH A RESISTING MEDIUM

The physics of fluid drag

When a body moves through a fluid such as air or water, the fluid exerts forces on the surface of the body. This is because the body must push the fluid out of the way, and to do this the body must exert forces on the fluid. By the Third Law, the fluid must then exert equal and opposite forces on the body. A person wading through water or riding a motorcycle is well aware of the existence of such forces, which fall into the general category of **material contact forces**. We are interested in the *resultant* force that the fluid exerts on the body and it is convenient to write this resultant in the form

$$\boldsymbol{F} = \boldsymbol{D} + \boldsymbol{L},$$

where the vector **drag force** $\boldsymbol{D}$ has the opposite direction to the velocity of the body, and the vector **lift force** $\boldsymbol{L}$ is at right angles to this velocity. The existence of lift makes air travel possible and is obviously very important. However, we will be concerned only with drag since we will restrict our attention to those cases in which the body is a rigid body of revolution which moves (without rotation) in the direction of its axis of symmetry. In this case, the lift is zero, by symmetry. We are then left with the scalar drag D, acting in the opposite direction to the velocity of the body.

The theoretical determination of lift and drag forces is one of the great unsolved problems of hydrodynamics and most of the available data has been obtained by experiment. Even for the case of a rigid sphere moving with constant velocity through an incompressible* fluid, a *general* theoretical solution for the drag is not available. In this problem, the drag depends on the radius a and speed V of the sphere, and the density ρ and viscosity μ of the fluid. Straightforward dimensional analysis shows that D must have the form

$$D = \rho a^2 V^2 F\left(\frac{\rho V a}{\mu}\right),$$

where F is a function of a single variable.

Definition 4.1 ***Reynold's number*** *The dimensionless quantity $R = \rho V a/\mu$ is called the* ***Reynolds number***.† *It is more commonly written $R = V a/\nu$, where the quantity $\nu = \mu/\rho$ is called the* ***kinematic viscosity*** *of the fluid.*

The function $F(R)$ has never been calculated theoretically, and experimental data must be used. It is a surprising fact that, for a wide range of values of R (about $1000 < R < 100,000$), the *function F is found to be roughly constant*. Subject to this approximation, the formula for the drag becomes

$$D = C\rho a^2 V^2,$$

where the dimensionless constant C is called the **drag coefficient**‡ for the sphere; its value is about 0.8.

A similar formula holds (with a different value of C) for any body of revolution moving parallel to its axis. In this case a is the radius of the maximum cross sectional area of the body perpendicular to the direction of motion. For example, the drag coefficient for a circular disk moving at right angles to its own plane is about 1.7. We thus obtain the result that (subject to the conditions mentioned above) the *drag is proportional to the square of the speed of the body through the fluid*. This is the **quadratic law of resistance**.

This result does not hold for low Reynolds numbers. This was shown theoretically by Stokes§ in his analysis of the creeping flow of a fluid past a sphere. Stokes proved that, as $R \to 0$, the function $F(R) \sim 6\pi/R$ so that the drag formula becomes

$$D \sim 6\pi a\mu V.$$

On dimensional grounds, a similar formula (with a different coefficient) should hold for other bodies of revolution. Thus *at low Reynolds numbers*¶ *the drag is proportional to speed of the body through the fluid*. This is the **linear law of resistance**.

Which (if either) of these laws is appropriate in any particular case depends on the Reynolds number. However, it is quickly apparent that the low Reynolds number condition requires quite special physical conditions, as the following example shows.

* In this treatment, the effects of fluid compressibility are neglected. In practice, this means that the speed of the body must be well below the speed of sound in the fluid.

† After the great English hydrodynamicist Osborne Reynolds 1842–1912. At the age of twenty six he was appointed to the University of Manchester's first professorship of engineering.

‡ The drag coefficient C_D used by aerodynamicists is $2C/\pi$.

§ George Gabriel Stokes 1819–1903, a major figure in British applied mathematics.

¶ *Low* means R less than about 0.5.

Table 1 Some fluid properties relevant to drag calculations (Kaye & Laby [14]).

	Density ρ ($\mathrm{kg\,m^{-3}}$)	Kinematic viscosity ν ($\mathrm{m^2 s^{-1}}$)	Sound speed ($\mathrm{m\,s^{-1}}$)
Air (20°C, 1 atm.)	1.20	1.50×10^{-5}	343
Water (20°C)	998	1.00×10^{-6}	1480
Castor oil (20°C)	950	1.04×10^{-3}	1420

Example 4.4 *Which law of resistance?*

A stainless steel ball bearing of radius 1 mm is falling vertically with constant speed in air. Find the speed of the ball bearing. [The density of stainless steel is 7800 $\mathrm{kg\,m^{-3}}$.]

If the medium were castor oil, what then would be the speed of the ball bearing?

Solution

Suppose the ball bearing is falling with constant speed V. (We will later call V the *terminal speed* of the ball bearing.) Then, since its acceleration is zero, the total of the forces acting upon it must also be zero. Thus

$$mg + D = m'g,$$

where m' is the mass of the ball bearing, m is the mass of the displaced fluid, and D is the drag. The term $m'g$ is the gravity force acting downwards and the term mg is the (Archimedes) buoyancy force acting upwards.* (In air, the buoyancy force is negligible.)

Hence, if the **linear** resistance law holds, then

$$6\pi a\rho\nu V = \tfrac{4}{3}\pi a^3 \left(\rho' - \rho\right) g,$$

where a is the radius of the ball bearing, and ρ', ρ are the densities of the ball bearing and air respectively. This gives

$$V = \frac{2a^2 g}{9\nu}\left(\frac{\rho'}{\rho} - 1\right).$$

On using the numerical values given in Table 4.3 we obtain $V = 940\ \mathrm{m\,s^{-1}}$ with the corresponding Reynolds number $R = 63,000$. Quite apart from the fact that the calculated speed is nearly three times the speed of sound, this solution is disqualified on the grounds that the Reynolds number is 100,000 times too large for the low Reynolds number approximation to hold!

On the other hand, if the **quadratic** law of resistance holds then

$$C\rho a^2 V^2 = \tfrac{4}{3}\pi a^3 \left(\rho' - \rho\right) g,$$

* It is not entirely obvious that the total force exerted by the fluid on the sphere is the sum of the drag and buoyancy forces, but it is true for an incompressible fluid.

where C is the drag coefficient for a sphere which we will take to be 0.8. In this case

$$V^2 = \frac{4\pi a g}{3C}\left(\frac{\rho'}{\rho} - 1\right).$$

This gives the value $V = 19\ \mathrm{m\,s^{-1}}$ with the corresponding Reynolds number $R = 1250$. This Reynolds number is nicely within the range in which the quadratic resistance law is applicable, and so provides a consistent solution. Thus the answer is that, **in air, the ball bearing falls with a speed of** $19\ \mathrm{m\,s^{-1}}$.

When the medium is **castor oil**, a similar calculation shows that it is the **linear resistance** law which provides the consistent solution. The answer is that, **in castor oil, the ball bearing falls with a speed of** $1.5\ \mathrm{cm\,s^{-1}}$, the Reynolds number being 0.015.

This example illustrates the conditions needed for low Reynolds number flow: slow motion of a small body through a sticky fluid. Perhaps the most celebrated application of the low Reynolds number drag formula is **Millikan**'s oil drop method of determining the electronic charge (see Problem 4.20 at the end of the chapter). ■

Example 4.5 *Vertical motion under gravity with linear resistance*

A body is projected vertically upwards with speed u in a medium that exerts a drag force $-mK\boldsymbol{v}$, where K is a positive constant and $\boldsymbol{v}$ is the velocity of the body.* Find the maximum height acheived by the body, the time taken to reach that height, and the terminal speed.

Solution

On including the linear resistance force, the scalar equation of motion becomes

$$m\frac{dv}{dt} = -mg - mKv,$$

with the initial condition $v = u$ when $t = 0$ (see Figure 4.1). This first order ODE for v separates in the form

$$\int \frac{dv}{g + Kv} = -\int dt,$$

and, on integration, gives

$$\frac{1}{K}\ln(g + Kv) = -t + C,$$

where C is the integration constant. On applying the initial condition $v = u$ when $t = 0$, we obtain $C = K^{-1}\ln(g + Ku)$ and so

$$t = \frac{1}{K}\ln\left(\frac{g + Ku}{g + Kv}\right).$$

* This is the vector drag force acting on the body; hence the minus sign. The coefficient is taken in the form mK for algebraic convenience.

This expression gives t in terms of v, which is what we need for finding the time taken to reach the maximum height. The maximum height is achieved when $v = 0$ so that τ, **the time taken to reach the maximum height**, is given by

$$\tau = \frac{1}{K} \ln\left(1 + \frac{Ku}{g}\right).$$

The expression for t in terms of v can be inverted to give

$$v = ue^{-Kt} - \frac{g}{K}\left(1 - e^{-Kt}\right)$$

for the upward **velocity** of the body at time t.

The terminal speed of the body is the limit of $|v|$ as $t \to \infty$. In this limit, the exponential terms tend to zero and

$$v \to -\frac{g}{K}.$$

Thus, in contrast to motion with no resistance, the speed of the body does not increase without limit as it falls, but tends to the finite value g/K. Thus the **terminal speed** of the body is g/K.

The terminal speed can also be deduced directly from the equation of motion. If the body is falling with the terminal speed, then $dv/dt = 0$ and the equation of motion implies that $0 = -mg - mKv$. It follows that the (upward) terminal velocity is $-g/K$.

The **maximum height** z_{max} can now be found by integrating the equation $dz/dt = v$ and then putting $t = \tau$. However we can also obtain z_{max} by starting again with a modified equation of motion. For some laws of resistance, this trick is essential. If we write

$$\frac{dv}{dt} = \frac{dv}{dz} \times \frac{dz}{dt} = v\frac{dv}{dz},$$

the equation of motion becomes

$$v\frac{dv}{dz} = -g - Kv,$$

with the initial condition $v = u$ when $z = 0$. This equation also separates to give

$$\begin{aligned} -\int dz &= \int \frac{v\,dv}{g + Kv} = \frac{1}{K}\int \left(1 - \frac{g}{g + Kv}\right) dv \\ &= \frac{v}{K} - \frac{g}{K^2}\ln(g + Kv) + D, \end{aligned}$$

where D is the integration constant. On applying the initial condition $v = u$ when $z = 0$, we obtain

$$\begin{aligned} z &= -\frac{v}{K} + \frac{g}{K^2}\ln(g + Kv) + \frac{u}{K} - \frac{g}{K^2}\ln(g + Ku) \\ &= \frac{1}{K}(u - v) - \frac{g}{K^2}\ln\left(\frac{g + Ku}{g + Kv}\right). \end{aligned}$$

This expression cannot be inverted to give v as a function of z, but it is exactly what we need to find z_{max}. Since z_{max} is achieved when $v = 0$, we find that the **maximum height** acheived by the body is given by

$$z_{max} = \frac{u}{K} - \frac{g}{K^2} \ln\left(1 + \frac{Ku}{g}\right). \blacksquare$$

Question ***Approximate form of z_{max} for small Ku/g***

Find an approximate expression for z_{max} when Ku/g is small.

Answer

When Ku/g is small, the log term can be expanded as a power series. This gives

$$\begin{aligned} z_{max} &= \frac{u}{K} - \frac{g}{K^2}\left[\frac{Ku}{g} - \tfrac{1}{2}\left(\frac{Ku}{g}\right)^2 + \tfrac{1}{3}\left(\frac{Ku}{g}\right)^3 + \cdots\right] \\ &= \frac{u^2}{2g}\left[1 - \tfrac{2}{3}\left(\frac{Ku}{g}\right) + \cdots\right]. \end{aligned}$$

In this expression, the leading term $u^2/2g$ is just the value of z_{max} in the absence of resistance. The first correction term has a negative sign which means that z_{max} is reduced by the presence of resistance, as would be expected. ■

Question ***Ball bearing released in castor oil***

The ball bearing in Example 4.4 is released from rest in castor oil. How long does it take for the ball bearing to achieve 99% of its terminal speed?

Answer

Recall that the linear law of resistance *is* appropriate for this motion. Since the motion is entirely downwards, it is more convenient to measure v downwards in this problem, in which case the solution for v becomes

$$v = \frac{g}{K}\left(1 - e^{-Kt}\right) = V\left(1 - e^{-gt/V}\right),$$

where V is the terminal velocity. When $v = 0.99V$, $e^{-gt/V} = 0.01$ and so the **time required** is

$$t = \ln(100)V/g,$$

which evaluates to about 7 milliseconds on using the value for V calculated in Example 4.4. ■

Note on the sign of resistance forces In the last example we used the same scalar equation of motion whether the body was rising or falling. This is correct in the case of linear resistance since, when the sign of v is reversed, so is the sign of Kv. In the case of quadratic resistance however, when the sign of v is reversed, the sign of Kv^2 remains unchanged and so the correct sign must be inserted manually. Thus, for *quadratic resistance, the scalar equations of motion for ascent and descent are different*. The same is true when the drag is proportional to *any even power* of v.

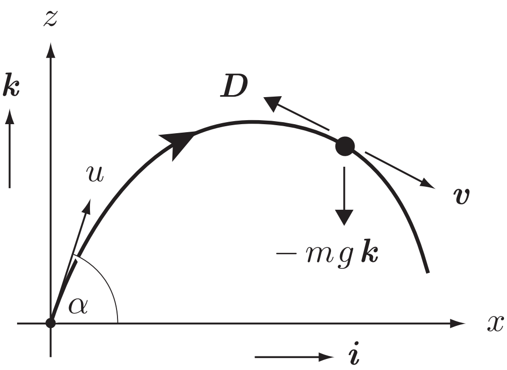


FIGURE 4.6 A particle, initially at the origin, is projected with speed u in a direction making an angle α with the horizontal. The particle moves under the uniform gravity force $-mg\boldsymbol{k}$ and the resistance (drag) force $\boldsymbol{D}$.

4.4 PROJECTILES

A body that moves freely under uniform gravity, and possibly air resistance, is called a **projectile**. Projectile motion is very common. In ball games, the ball is a projectile, and controlling its trajectory is a large part of the skill of the game. On a larger scale, artillery shells are projectiles, but guided missiles, which have rocket propulsion, are not.

The projectile problem differs from the problems considered in section 4.3 in that projectile motion is not restricted to take place in a vertical straight line. However, we will continue to assume that the effect of the air is to exert a drag force opposing the current velocity of the projectile.* It is then evident by symmetry that each **projectile motion takes place in a vertical plane**; this vertical plane contains the initial position of the projectile and is parallel to its initial velocity.

Projectiles without resistance

The first (and easiest) problem is that of a projectile moving without air resistance. This is fine on the Moon, but will be only an approximation to projectile motion on Earth. The effect of air resistance can be very significant, as our later examples will show.

Example 4.6 *Projectile without air resistance*

A particle which is subject solely to uniform gravity is projected with speeed u in a direction making an angle α with the horizontal. Find the subsequent motion.

Solution

Suppose that the motion takes place in the (x, z)-plane as shown in Figure 4.6. In the absence of the drag force, the vector equation of motion becomes

$$m\frac{d\boldsymbol{v}}{dt} = -mg\boldsymbol{k},$$

with the initial condition $\boldsymbol{v} = (u\cos\alpha)\boldsymbol{i} + (u\sin\alpha)\boldsymbol{k}$ when $t = 0$. If we now write $\boldsymbol{v} = v_x\boldsymbol{i} + v_z\boldsymbol{k}$ and take components of this equation (and initial condition) in the

* This will be true if the projectile is a rigid sphere moving without rotation.

$\boldsymbol{i}$- and $\boldsymbol{k}$-directions, we obtain the two scalar equations of motion

$$\frac{dv_x}{dt} = 0, \qquad \frac{dv_z}{dt} = -g,$$

with the respective initial conditions $v_x = u\cos\alpha$ and $v_z = u\sin\alpha$ when $t = 0$. Simple integrations then give the components of the particle **velocity** to be

$$v_x = u\cos\alpha, \qquad v_z = u\sin\alpha - gt.$$

The position of the particle at time t can now be found by integrating the expressions for v_x, v_z and applying the initial conditions $x = 0$ and $z = 0$ when $t = 0$. This gives

$$x = (u\cos\alpha)\,t, \qquad z = (u\sin\alpha)\,t - \tfrac{1}{2}gt^2,$$

the solution for the **trajectory** of the particle. ■

Question ***Form of the path***

Show that the path taken by the particle is an inverted parabola.

Answer

To find the path, eliminate t from the trajectory equations. This gives

$$z = (\tan\alpha)\,x - \left(\frac{g}{2u^2\cos^2\alpha}\right)x^2,$$

which is indeed an inverted parabola. ■

Question ***Time of flight and the range***

Find the time of flight and the range of the projectile on level ground.

Answer

On level ground, the motion will terminate when $z = 0$ again. From the second trajectory equation, this happens when $(u\sin\alpha)\,t - \frac{1}{2}gt^2 = 0$. Hence the time of flight τ is given by $\tau = 2u\sin\alpha/g$. The horizontal range R is then obtained by putting $t = \tau$ in the first trajectory equation, which gives

$$R = \frac{u^2\sin 2\alpha}{g}. \ ■$$

Question ***Maximum range***

Find the value of α that gives the maximum range on level ground when u is fixed.

Answer

R is a maximum when $\sin 2\alpha = 1$, that is when $\alpha = \pi/4$ in which case $R_{\max} = u^2/g$. Thus, if an artillery shell is to be projected over a horizontal range of 4 km, then the gun must have a muzzle speed of at least 200 m s^{-1}. ■

There is a myriad of problems that can be found on the projectile with no air resistance, and some interesting examples are included at the end of the chapter. It should be noted though that all these problems are *dynamically* equivalent to the problem solved above. Any difficulties lie in the geometry!

Projectiles with resistance

We now proceed to include the effect of air resistance. From our earlier discussion of fluid drag, it is evident that in most practical instances of projectile motion through the Earth's atmosphere, it is the **quadratic law** of resistance that is appropriate. On the other hand, only the **linear law** of resistance gives rise to linear equations of motion and simple analytical solutions. This explains why mechanics textbooks contain extensive coverage of the linear case, even though this case is almost never appropriate in practice; the case that *is* appropriate cannot be solved! In the following example, we treat the linear resistance case.

Example 4.7 *Projectile with linear resistance*

A particle is subject to uniform gravity and the linear resistance force $-mK\boldsymbol{v}$, where K is a positive constant and $\boldsymbol{v}$ is the velocity of the particle. Initially the particle is projected with speed u in a direction making an angle α with the horizontal. Find the subsequent motion.

Solution

With the linear resistance term included, the vector equation of motion becomes

$$m\frac{d\boldsymbol{v}}{dt} = -mK\boldsymbol{v} - mg\boldsymbol{k},$$

with the initial condition $\boldsymbol{v} = (u\cos\alpha)\boldsymbol{i} + (u\sin\alpha)\boldsymbol{k}$ when $t = 0$. As in the last example, this equation resolves into the two scalar equations of motion

$$\frac{dv_x}{dt} + Kv_x = 0, \qquad \frac{dv_z}{dt} + Kv_z = -g,$$

with the respective initial conditions $v_x = u\cos\alpha$ and $v_z = u\sin\alpha$ when $t = 0$. These first order ODEs are both separable and linear and can be solved by either method; if they are regarded as linear, the integrating factor is e^{Kt}. The equations integrate to give the components of the particle **velocity** to be

$$v_x = (u\cos\alpha)e^{-Kt}, \qquad v_z = (u\sin\alpha)e^{-Kt} - \frac{g}{K}\left(1 - e^{-Kt}\right).$$

The position of the particle at time t can now be found by integrating the expressions for v_x, v_z and applying the initial conditions $x = 0$ and $z = 0$ when $t = 0$. This gives

$$x = \frac{u\cos\alpha}{K}\left(1 - e^{-Kt}\right), \quad z = \frac{Ku\sin\alpha + g}{K^2}\left(1 - e^{-Kt}\right) - \frac{g}{K}t, \tag{4.3}$$

the solution for the **trajectory of the particle**. Figure 4.7 shows typical paths taken by the particle for the same initial conditions and three different values of the dimensionless resistance parameter λ ($= Ku/g$). (The case $\lambda = 0$ corresponds to zero resistance so that the path is a parabola.) It is apparent that resistance can have a dramatic effect on the motion. ■

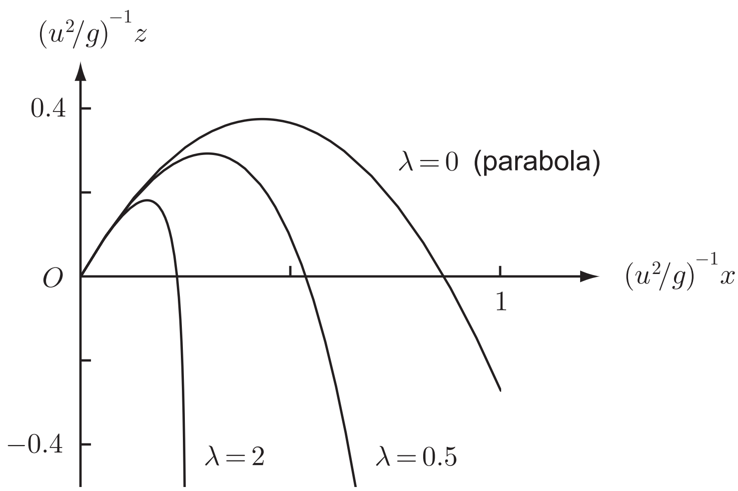


FIGURE 4.7 Projectile motion under uniform gravity and **linear resistance**. The graphs show the paths of the particle for $\alpha = \pi/3$ and three different values of the dimensionless resistance parameter λ ($= Ku/g$). Except when $\lambda = 0$, the paths have vertical asymptotes.

Question *Vertical asymptote of the path*

Show that the path has a vertical asymptote.

Answer

Since e^{-Kt} decreases and tends to zero as $t \to \infty$, it follows from equations (4.3) that the horizontal displacement x *increases* and tends to the value $u \cos\alpha/K$ as $t \to \infty$, while the vertical displacement z tends to negative infinity. Thus the vertical line $x = u\cos\alpha/K$ is an asymptote to the path. In terms of the dimensionless variables used in Figure 4.7, this is the line $(u^2/g)^{-1}x = \cos\alpha/\lambda$. ■

Question *Approximate formula for the range when* λ *is small*

Find an approximate formula for the range on level ground when the resistance parameter λ is small.

Answer

Since the particle returns to Earth again when $z = 0$, it follows from the second of equations (4.3) that the flight time τ satisfies the equation

$$(Ku\sin\alpha + g)\left(1 - e^{-K\tau}\right) - Kg\tau = 0,$$

which can be written in the form

$$(\lambda\sin\alpha + 1)\left(1 - e^{-K\tau}\right) - K\tau = 0, \tag{4.4}$$

where $\lambda (= Ku/g)$ is the dimensionless resistance parameter. Unfortunately, this equation cannot be solved explicity for τ, and hence the need for an approximate solution. We know from the last example that, in the absence of resistance, the flight time τ is given by $\tau = 2u\sin\alpha/g$. It is

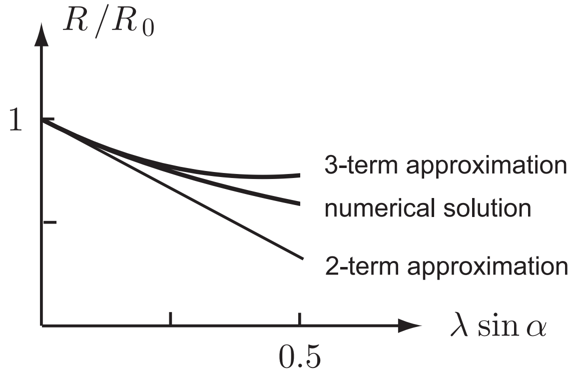


FIGURE 4.8 The ratio R/R_0 plotted against $\lambda \sin\alpha$.

reasonable then, when λ is small, to seek a solution for τ in the form

$$\tau = \frac{2u \sin\alpha}{g}\left[1 + b_1\lambda + b_2\lambda^2 + \cdots\right], \tag{4.5}$$

where the coefficients $b_1, b_2, \ldots$ are to be determined. To find the expansion coefficients we substitute the expansion (4.5) (truncated after the required number of terms) into the left side of equation (4.4), re-expand in powers of λ, and then set the coefficients in this expansion equal to zero. The corresponding formula for the range can then be found by substituting this approximate formula for τ into the first equation of (4.3) and re-expanding in powers of λ. The details are tedious and, in fact, such operations are best done with computer assistance. The completion of this solution is the subject of computer assisted Problem 4.34 at the end of this chapter. The answer (to three terms) is that the range R on level ground is given by

$$\frac{R}{R_0} = 1 - \left(\frac{4\sin\alpha}{3}\right)\lambda + \left(\frac{14\sin^2\alpha}{9}\right)\lambda^2 + O(\lambda^3),$$

where R_0 is the range when resistance is absent. Figure 4.8 compares two different approximations to R with the ‘exact’ value obtained by numerical solution of equation (4.4). As would be expected, the three term approximation is closer to the exact value. ■

4.5 CIRCULAR MOTION

In this section we examine some important problems in which a body moves on a circular path. Our first problem is concerned with a body executing a circular orbit under the gravitational attraction of a fixed mass. This is a fairly accurate model of the motion of the planets* around the Sun.

Example 4.8 *Circular orbit in the inverse square field*

A particle of mass m moves under the gravitational attraction of a fixed mass M situated at the origin. Show that circular orbits with centre O and any radius are

* The orbits of Mercury, Mars and Pluto are the most elliptical with eccentricities of 0.206, 0.093 and 0.249 respectively. The eccentricity of Earth’s orbit is 0.017.

possible, and find the speed of the particle in such an orbit. Deduce the period of the orbit.

Solution

Note that we are not required to find the *general* orbit; we may assume from the start that the orbit is a circle. Suppose then that the particle is executing a circular orbit with centre O and radius R. We need to confirm that the vector equation of motion can be satisfied. Take polar coordinates r, θ with centre at O. Then the acceleration $\boldsymbol{a}$ of the particle is given in terms of the usual polar unit vectors by the formula (2.14), that is,

$$\begin{aligned}\boldsymbol{a} &= \left(\ddot{r} - r\dot{\theta}^2\right)\widehat{\boldsymbol{r}} + \left(r\ddot{\theta} + 2\dot{r}\dot{\theta}\right)\widehat{\boldsymbol{\theta}}\\ &= -\frac{v^2}{R}\widehat{\boldsymbol{r}} + \dot{v}\,\widehat{\boldsymbol{\theta}}\end{aligned}$$

for motion on the circle $r = R$, where the circumferential velocity $v = R\dot{\theta}$. The equation of motion for the particle is therefore

$$m\left[-\frac{v^2}{R}\widehat{\boldsymbol{r}} + \dot{v}\widehat{\boldsymbol{\theta}}\right] = -\frac{mMG}{R^2}\widehat{\boldsymbol{r}},$$

which, on taking components in the radial and transverse directions, gives

$$\frac{v^2}{R} = \frac{MG}{R^2} \qquad \text{and} \qquad \dot{v} = 0.$$

Hence the equation of motion is satisfied if v is a constant given by

$$v^2 = \frac{MG}{R}.$$

Thus a circular orbit of radius R is possible provided that the particle has **constant speed** $(MG/R)^{1/2}$.

The **period** τ of the orbit is the time taken for one circuit and is given by

$$\tau = \frac{2\pi R}{v} = \left(\frac{4\pi^2 R^3}{MG}\right)^{1/2}.$$

Thus the square of the period of a circular orbit is proportional to the cube of its radius. This is a special case of Kepler's third law of planetary motion (see Chapter 7). ■

A particle may move on a circular path because it is *constrained* to do so. The simplest and most important example of this is the simple pendulum, a mass suspended from a fixed point by a string.

Example 4.9 *The simple pendulum*

A particle P is suspended from a fixed point O by a light inextensible string of length b. P is subject to uniform gravity and moves in a vertical plane through O with the string taut. Find the equation of motion.

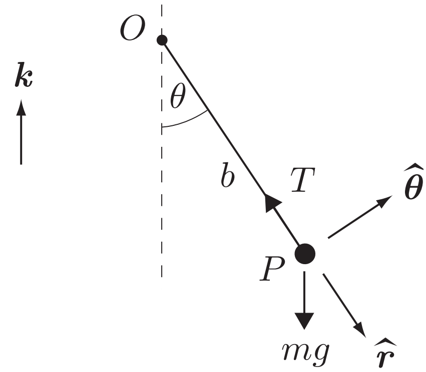


FIGURE 4.9 The simple pendulum

Solution

The system is shown in Figure 4.9. Since the string is of fixed length b, the position of P is determined by the angle θ shown. The acceleration of P can be expressed in the polar form

$$\boldsymbol{a} = -\left(b\dot{\theta}^2\right)\widehat{\boldsymbol{r}} + \left(b\ddot{\theta}\right)\widehat{\boldsymbol{\theta}},$$

where $\widehat{\boldsymbol{r}}$ and $\widehat{\boldsymbol{\theta}}$ are the polar unit vectors shown in Figure 4.9.

P moves under the uniform gravity force $-mg\boldsymbol{k}$ and the tension T in the string which acts in the direction $-\widehat{\boldsymbol{r}}$. It should be noted that the tension T is a force of constraint and not known beforehand. The equation of motion is therefore

$$m\left[-\left(b\dot{\theta}^2\right)\widehat{\boldsymbol{r}} + \left(b\ddot{\theta}\right)\widehat{\boldsymbol{\theta}}\right] = -mg\boldsymbol{k} - T\widehat{\boldsymbol{r}}.$$

If we now take components of this equation in the radial and transverse directions we obtain

$$-mb\dot{\theta}^2 = mg\cos\theta - T, \qquad mb\ddot{\theta} = -mg\sin\theta.$$

The second of these equations is the effective equation of motion in terms of the 'coordinate' θ, namely,

$$\ddot{\theta} + \left(\frac{g}{b}\right)\sin\theta = 0, \tag{4.6}$$

while the first equation determines the unknown tension T once $\theta(t)$ is known.

Equation (4.6) is the **exact equation of motion** for the simple pendulum. Because of the presence of the term in $\sin\theta$, this second order ODE is **non-linear** and cannot be solved by using the standard technique for linear ODEs with constant coefficients. ■

Question *The linear theory for small amplitude oscillations*

Find an approximate linear equation for the case in which the pendulum undergoes oscillations of small amplitude.

Answer

If θ is always small then $\sin\theta$ can be approximated by θ in which case the equation of motion becomes

$$\ddot{\theta} + \left(\frac{g}{b}\right)\theta = 0. \tag{4.7}$$

This is the **linearised equation** for the simple pendulum, which holds approximately for oscillations of small amplitude. Although we do not cover linear oscillations until Chapter 5, many readers will recognise equation (4.7) as the simple harmonic motion equation and will know that the period τ of the oscillations is given by $\tau = 2\pi(b/g)^{1/2}$, independent of the (small) amplitude. ■

Question *Period of large oscillations*

Find the period of the pendulum when the (angular) amplitude of its oscillations is α, where α may not be small.

Answer

This requires that we integrate the exact equation of motion (4.6). We start with a familiar trick. If we write $\Omega = \dot{\theta}$, then

$$\ddot{\theta} = \frac{d\Omega}{dt} = \frac{d\Omega}{d\theta} \times \frac{d\theta}{dt} = \Omega\frac{d\Omega}{d\theta}$$

and the equation of motion becomes

$$\Omega\frac{d\Omega}{d\theta} = -\left(\frac{g}{b}\right)\sin\theta.$$

This is a separable first order ODE for Ω which integrates to give

$$\tfrac{1}{2}\Omega^2 = \left(\frac{g}{b}\right)\cos\theta + C,$$

where C is the constant of integration. On applying the initial condition $\Omega = 0$ when $\theta = \alpha$, we find that $C = -(g/b)^{1/2}\cos\alpha$ and the integrated equation can be written

$$\left(\frac{d\theta}{dt}\right)^2 = \left(\frac{2g}{b}\right)(\cos\theta - \cos\alpha), \tag{4.8}$$

where we have now replaced Ω by $d\theta/dt$.

Since the pendulum comes to rest only when $d\theta/dt = 0$ (that is, when $\theta = \pm\alpha$) it follows that θ must oscillate in the range $-\alpha \le \theta \le \alpha$. The period τ is the time taken for one complete oscillation but, by the symmetry of equation (4.8) under the transformation $\theta \to -\theta$, it follows that the time taken for the pendulum to swing from $\theta = 0$ to $\theta = +\alpha$ is $\tau/4$. To evaluate this time we take the positive square root of each side of equation (4.8) and integrate over the time interval $0 \le t \le \tau/4$. This gives

$$\int_0^\alpha \frac{d\theta}{(\cos\theta - \cos\alpha)^{1/2}} = \left(\frac{2g}{b}\right)^{1/2}\int_0^{\tau/4} dt,$$

so that

$$\tau = \left(\frac{8b}{g}\right)^{1/2}\int_0^\alpha \frac{d\theta}{(\cos\theta - \cos\alpha)^{1/2}}. \tag{4.9}$$

This is the **exact period** of the pendulum when the amplitude of its oscillations is α. It is not possible to perform this integration in terms of standard functions* and so the integral must either be evaluated numerically or be approximated.

Numerical evaluation shows that the *exact period is longer than that predicted by the linearised theory*. When $\alpha = \pi/6$, the period is 1.7% longer, and when $\alpha = \pi/3$ it is 7.3% longer. The period can also be approximated by expanding the integral in equation (4.9) as a power series in α. This is the subject of Problem 4.35 at the end of the chapter. The answer is that, expanded to two terms,

$$\tau = 2\pi \left(\frac{b}{g}\right)^{1/2} \left[1 + \frac{\alpha^2}{16} + O(\alpha^2)\right]. \tag{4.10}$$

This two term approximation predicts an increase in the period of 1.7% when $\alpha = \pi/6$. Note that there is no term in this expansion proportional to α and that the term in α^2 has the small coefficient $1/16$. This explains why the prediction of the linearised theory is rather accurate even when α is not so small! ■

In our final example, we solve the important problem of an **electrically charged particle moving in a uniform magnetic field**. It turns out that plane motions are circular, but the most general motion is helical. The solution in this case differs from the previous examples in that we use Cartesian coordinates instead of polars. This is because we do not know beforehand where the centre of the circle (or the axis of the helix) is, which means that we do not know on which point (or axis) to centre the polar coordinates.

Example 4.10 *Charged particle in a magnetic field*

A particle of mass m and charge e moves in a uniform magnetic field of strength B_0. Show that the most general motion is helical with the axis of the helix parallel to the direction of the magnetic field.

Solution

The total force $\boldsymbol{F}$ that an electric field $\boldsymbol{E}$ and magnetic field $\boldsymbol{B}$ exert on a charge e is given by the **Lorentz force formula**†

$$\boldsymbol{F} = e\boldsymbol{E} + e\boldsymbol{v} \times \boldsymbol{B},$$

where $\boldsymbol{v}$ is the velocity of the charge. In our problem, there is no electric field and the magnetic field is uniform. If the direction of $\boldsymbol{B}$ is the z-direction of Cartesian coordinates, then $\boldsymbol{B} = B_0\boldsymbol{k}$. The equation of motion of the particle is then

$$m\frac{d\boldsymbol{v}}{dt} = eB_0\boldsymbol{v} \times \boldsymbol{k}.$$

If we now write $\boldsymbol{v}$ in the component form $\boldsymbol{v} = v_x\boldsymbol{i} + v_y\boldsymbol{j} + v_z\boldsymbol{k}$, the vector equation of motion resolves into the three scalar equations

$$\frac{dv_x}{dt} = \Omega v_y, \qquad \frac{dv_y}{dt} = -\Omega v_x, \qquad \frac{dv_z}{dt} = 0, \tag{4.11}$$

* The integral is related to a special function called the *complete elliptic integral of the first kind*.
† This form is correct in SI units.

where

$$\Omega = eB_0/m. \tag{4.12}$$

The last of these equations shows that $v_z = V$, a constant, so that the component of $\boldsymbol{v}$ parallel to the magnetic field is a constant. The first two equations are **first order coupled ODEs** but they are easy to uncouple. If we differentiate the first equation with respect to t and use the second equation, we find that v_x satisfies the equation

$$\frac{d^2 v_x}{dt^2} + \Omega^2 v_x = 0.$$

This equation is a second order ODE with constant coefficients and can be solved in the standard way. However, many readers will recognise this as the SHM equation whose general solution can be written in the form

$$v_x = A \sin(\Omega t + \alpha),$$

where A and α are arbitrary constants. It is more convenient if we introduce a new arbitrary constant R defined by $A = -\Omega R$, so that

$$v_x = -\Omega R \sin(\Omega t + \alpha).$$

If we now substitute this formula for v_x into the first equation of (4.11), we obtain

$$v_y = -\Omega R \cos(\Omega t + \alpha).$$

Having obtained the solution for the three components of $\boldsymbol{v}$, we can now find the trajectory simply by integrating with respect to t. This gives

$$x = R\cos(\Omega t + \alpha) + a, \qquad y = -R\sin(\Omega t + \alpha) + b, \qquad z = Vt + c,$$

where a, b, and c are constants of integration. These constants may be removed by a shift of the origin of coordinates to the point (a, b, c), and the constant α may be removed by a shift in the origin of t. Also, the constant R may be assumed positive; if it is not, make a shift in the origin of t by π/Ω. With these simplifications, the final form for the trajectory is

$$x = R\cos\Omega t, \qquad y = -R\sin\Omega t, \qquad z = Vt, \tag{4.13}$$

where R is a positive constant and $\Omega = eB_0/m$. This is the most **general trajectory** for a charged particle moving in a uniform magnetic field.

To identify this trajectory as a helix, suppose first that $V = 0$ so that the motion takes place in the (x, y)-plane. Then the first two equations of (4.13) imply that the path is a circle of radius R traversed with constant speed $R|\Omega|$ and with period $2\pi/|\Omega|$. When $V \neq 0$, this circular motion is supplemented by a uniform velocity V in the z-direction. The result is a helical path of radius R, with its axis parallel to the magnetic field, which is traversed with constant speed $(V^2 + R^2\Omega^2)^{1/2}$. ■

The above problem has important applications to the **cyclotron** particle accelerator and the **mass spectrograph**. The cyclotron depends for its operation on thefact that the constant

Ω, known as the **cyclotron frequency**, is independent of the velocity of the charged particles. The mass spectrograph is the subject of Problem 4.32 at the end of the chapter.

Problems on Chapter 4

Answers and comments are at the end of the book.
Harder problems carry a star ($*$).

Introductory problems

4.1 Two identical blocks each of mass M are connected by a light inextensible string and can move on the surface of a *rough* horizontal table. The blocks are being towed at constant speed in a straight line by a rope attached to one of them. The tension in the tow rope is T_0. What is the tension in the connecting string? The tension in the tow rope is suddenly increased to $4T_0$. What is the instantaneous acceleration of the blocks and what is the instantaneous tension in the connecting string?

4.2 A body of mass M is suspended from a fixed point O by an inextensible uniform rope of mass m and length b. Find the tension in the rope at a distance z below O. The point of support now begins to rise with acceleration $2g$. What now is the tension in the rope?

4.3 Two uniform lead spheres each have mass 5000 kg and radius 47 cm. They are released from rest with their centres 1 m apart and move under their mutual gravitation. Show that they will collide in *less* than 425 s. [$G = 6.67 \times 10^{-11}\ \mathrm{N\,m^2\,kg^{-2}}$.]

4.4 The block in Figure 4.2 is sliding down the inclined surface of a fixed wedge. This time the frictional force F exerted on the block is given by $F = \mu N$, where N is the normal reaction and μ is a positive constant. Find the acceleration of the block. How do the cases $\mu < \tan\alpha$ and $\mu > \tan\alpha$ differ?

4.5 A stuntwoman is to be fired from a cannon and projected a distance of 40 m over level ground. What is the least projection speed that can be used? If the barrel of the cannon is 5 m long, show that she will experience an acceleration of at least $4g$ in the barrel. [Take $g = 10\ \mathrm{m\,s^{-2}}$.]

4.6 In an air show, a pilot is to execute a circular loop at the speed of sound ($340\ \mathrm{m\,s^{-1}}$). The pilot may black out if his acceleration exceeds $8g$. Find the radius of the smallest circle he can use. [Take $g = 10\ \mathrm{m\,s^{-2}}$.]

4.7 A body has terminal speed V when falling in still air. What is its terminal velocity (relative to the ground) when falling in a steady horizontal wind with speed U?

4.8 ***Cathode ray tube*** A particle of mass m and charge e is moving along the x-axis with speed u when it passes between two charged parallel plates. The plates generate a uniform electric field $E_0\boldsymbol{j}$ in the region $0 \le x \le b$ and no field elsewhere.* Find the angle through

* This is only approximately true.

which the particle is deflected by its passage between the plates. [The cathode ray tube uses this arrangement to deflect the electron beam.]

Straight line motion in a force field

4.9 An object is dropped from the top of a building and is in view for time τ while passing a window of height h some distance lower down. How high is the top of the building above the top of the window?

4.10 A particle P of mass m moves under the gravitational attraction of a mass M fixed at the origin O. Initially P is at a distance a from O when it is projected with the *critical* escape speed $(2MG/a)^{1/2}$ directly away from O. Find the distance of P from O at time t, and confirm that P escapes to infinity.

4.11 A particle P of mass m is attracted towards a fixed origin O by a force of magnitude $m\gamma/r^3$, where r is the distance of P from O and γ is a positive constant. [It's gravity Jim, but not as we know it.] Initially, P is at a distance a from O, and is projected with speed u directly away from O. Show that P will escape to infinity if $u^2 > \gamma/a^2$.

For the case in which $u^2 = \gamma/(2a^2)$, show that the maximum distance from O achieved by P in the subsequent motion is $\sqrt{2}a$, and find the time taken to reach this distance.

4.12 If the Earth were suddenly stopped in its orbit, how long would it take for it to collide with the Sun? [Regard the Sun as a *fixed* point mass. You may make use of the formula for the period of the Earth's orbit.]

Constrained motion

4.13 A particle P of mass m slides on a smooth horizontal table. P is connected to a second particle Q of mass M by a light inextensible string which passes through a small smooth hole O in the table, so that Q hangs below the table while P moves on top. Investigate motions of this system in which Q remains at rest vertically below O, while P describes a circle with centre O and radius b. Show that this is possible provided that P moves with constant speed u, where $u^2 = Mgb/m$.

4.14 A light pulley can rotate freely about its axis of symmetry which is fixed in a horizontal position. A light inextensible string passes over the pulley. At one end the string carries a mass $4m$, while the other end supports a second light pulley. A second string passes over this pulley and carries masses m and $4m$ at its ends. The whole system undergoes planar motion with the masses moving vertically. Find the acceleration of each of the masses.

4.15 A particle P of mass m can slide along a *smooth* rigid straight wire. The wire has one of its points fixed at the origin O, and is made to rotate in the (x, y)-plane with angular speed Ω. By using the vector equation of motion of P in polar co-ordinates, show that r, the distance of P from O, satisfies the equation

$$\ddot{r} - \Omega^2 r = 0,$$

and find a second equation involving N, where $N\widehat{\boldsymbol{\theta}}$ is the force the wire exerts on P. [Ignore gravity in this question.]

Initially, P is at rest (relative to the wire) at a distance a from O. Find r as a function of t in the subsequent motion, and deduce the corresponding formula for N.

Resisted motion

4.16 A body of mass m is projected with speed u in a medium that exerts a resistance force of magnitude (i) $mk|\boldsymbol{v}|$, or (ii) $mK|\boldsymbol{v}|^2$, where k and K are positive constants and $\boldsymbol{v}$ is the velocity of the body. Gravity can be ignored. Determine the subsequent motion in each case. Verify that the motion is bounded in case (i), but not in case (ii).

4.17 A body is projected vertically upwards with speed u and moves under uniform gravity in a medium that exerts a resistance force proportional to the square of its speed and in which the body's terminal speed is V. Find the maximum height above the starting point attained by the body and the time taken to reach that height.

Show also that the speed of the body when it returns to its starting point is $uV/(V^2 + u^2)^{1/2}$. [*Hint*. The equations of motion for ascent and descent are different. See the note at the end of section 4.3.]

4.18* A body is released from rest and moves under uniform gravity in a medium that exerts a resistance force proportional to the square of its speed and in which the body's terminal speed is V. Show that the time taken for the body to fall a distance h is

$$\frac{V}{g}\cosh^{-1}\left(e^{gh/V^2}\right).$$

In his famous (but probably apocryphal) experiment, Galileo dropped different objects from the top of the tower of Pisa and timed how long they took to reach the ground. If Galileo had dropped two iron balls, of 5 mm and 5 cm radius respectively, from a height of 25 m, what would the descent times have been? Is it likely that this difference could have been detected? [Use the quadratic law of resistance with $C = 0.8$. The density of iron is 7500 kg m^{-3}.]

4.19 A body is projected vertically upwards with speed u and moves under uniform gravity in a medium that exerts a resistance force proportional to the fourth power its speed and in which the body's terminal speed is V. Find the maximum height above the starting point attained by the body.

Deduce that, however large u may be, this maximum height is always less than $\pi V^2/4g$.

4.20 ***Millikan's experiment*** A microscopic spherical oil droplet, of density ρ and unknown radius, carries an unknown electric charge. The droplet is observed to have terminal speed v_1 when falling vertically in air of viscosity μ. When a uniform electric field E_0 is applied in the vertically upwards direction, the same droplet was observed to move *upwards* with terminal speed v_2. Find the charge on the droplet. [Use the low Reynolds number approximation for the drag.]

Projectiles

4.21 A mortar gun, with a maximum range of 40 m on level ground, is placed on the edge of a vertical cliff of height 20 m overlooking a horizontal plain. Show that the horizontal range R of the mortar gun is given by

$$R = 40\left\{\sin\alpha + \left(1+\sin^2\alpha\right)^{\frac{1}{2}}\right\}\cos\alpha,$$

where α is the angle of elevation of the mortar above the horizontal. [Take $g = 10\ \mathrm{m\,s^{-2}}$.]

Evaluate R (to the nearest metre) when $\alpha = 45°$ and $35°$ and confirm that $\alpha = 45°$ does not yield the maximum range. [Do not try to find the optimum projection angle this way. See Problem 4.22 below.]

4.22 It is required to project a body from a point on level ground in such a way as to clear a thin vertical barrier of height h placed at distance a from the point of projection. Show that the body will just skim the top of the barrier if

$$\left(\frac{ga^2}{2u^2}\right)\tan^2\alpha - a\tan\alpha + \left(\frac{ga^2}{2u^2}+h\right) = 0,$$

where u is the speed of projection and α is the angle of projection above the horizontal.

Deduce that, if the above trajectory is to exist for some α, then u must satisfy

$$u^4 - 2ghu^2 - g^2a^2 \geq 0.$$

Find the least value of u that satisfies this inequality.

For the special case in which $a = \sqrt{3}h$, show that the minimum projection speed necessary to clear the barrier is $(3gh)^{\frac{1}{2}}$, and find the projection angle that must be used.

4.23 A particle is projected from the origin with speed u in a direction making an angle α with the horizontal. The motion takes place in the (x, z)-plane, where Oz points vertically upwards. If the projection speed u is fixed, show that the particle can be made to pass through the point (a, b) for some choice of α if (a, b) lies *below* the parabola

$$z = \frac{u^2}{2g}\left(1 - \frac{g^2x^2}{u^4}\right).$$

This is called the **parabola of safety**. Points *above* the parabola are 'safe' from the projectile.

An artillery shell explodes on the ground throwing shrapnel in all directions with speeds of up to $30\ \mathrm{m\,s^{-1}}$. A man is standing at an open window 20 m above the ground in a building 60 m from the blast. Is he safe? [Take $g = 10\ \mathrm{m\,s^{-2}}$.]

4.24 A projectile is fired from the top of a conical mound of height h and base radius a. What is the least projection speed that will allow the projectile to clear the mound? [*Hint*. Make use of the parabola of safety.]

A mortar gun is placed on the summit of a conical hill of height 60 m and base diameter 160 m. If the gun has a muzzle speed of $25\ \mathrm{m\,s^{-1}}$, can it shell anywhere on the hill? [Take $g = 10\ \mathrm{m\,s^{-2}}$.]

4.25 An artillery gun is located on a plane surface inclined at an angle β to the horizontal. The gun is aligned with the line of steepest slope of the plane. The gun fires a shell with speed u in the direction making an angle α with the (upward) line of steepest slope. Find where the shell lands.

Deduce the maximum ranges R^U, R^D, up and down the plane, and show that

$$\frac{R^U}{R^D} = \frac{1 - \sin\beta}{1 + \sin\beta}.$$

4.26 Show that, when a particle is projected from the origin in a medium that exerts *linear* resistance, its position vector at time t has the general form

$$\boldsymbol{r} = -\alpha(t)\boldsymbol{k} + \beta(t)\boldsymbol{u},$$

where $\boldsymbol{k}$ is the vertically upwards unit vector and $\boldsymbol{u}$ is the *velocity* of projection. Deduce the following results:

(i) A number of particles are projected simultaneously from the same point, with the same speed, but in *different directions*. Show that, at each later time, the particles all lie on the surface of a sphere.
(ii) A number of particles are projected simultaneously from the same point, in the same direction, but with *different speeds*. Show that, at each later time, the particles all lie on a straight line.
(iii) Three particles are projected simultaneously in a completely general manner. Show that the plane containing the three particles remains parallel to some fixed plane.

4.27 A body is projected in a steady horizontal wind and moves under uniform gravity and *linear* air resistance. Show that the influence of the wind is the same as if the magnitude and direction of gravity were altered. Deduce that it is possible for the body to return to its starting point. What is the shape of the path in this case?

Circular motion and charged particles

4.28 The radius of the Moon's approximately circular orbit is 384,000 km and its period is 27.3 days. Estimate the mass of the Earth. [$G = 6.67 \times 10^{-11}$ N m^2 kg^{-2}.] The actual mass is 5.97×10^{24} kg. What is the main reason for the error in your estimate?

An artificial satellite is to be placed in a circular orbit around the Earth so as to be 'geo-stationary'. What must the radius of its orbit be? [The period of the Earth's rotation is 23 h 56 m, *not* 24 h. Why?]

4.29 ***Conical pendulum*** A particle is suspended from a fixed point by a light inextensible string of length a. Investigate 'conical motions' of this pendulum in which the string maintains a constant angle α with the downward vertical. Show that, for any acute angle α, a conical motion exists and that the particle speed u is given by $u^2 = ag\sin\alpha\tan\alpha$.

4.30 A particle of mass m is attached to the highest point of a *smooth* rigid sphere of radius a by a light inextensible string of length $\pi a/4$. The particle moves in contact with the outer surface of the sphere, with the string taut, and describes a horizontal circle with constant speed u. Find the reaction of the sphere on the particle and the tension in the string. Deduce the maximum value of u for which such a motion could take place. What will happen if u exceeds this value?

4.31 A particle of mass m can move on a *rough* horizontal table and is attached to a fixed point on the table by a light inextensible string of length b. The resistance force exerted on the particle is $-mK\boldsymbol{v}$, where $\boldsymbol{v}$ is the velocity of the particle. Initially the string is taut and the particle is projected horizontally, at right angles to the string, with speed u. Find the angle turned through by the string before the particle comes to rest. Find also the tension in the string at time t.

4.32 ***Mass spectrograph*** A stream of particles of various masses, all carrying the same charge e, is moving along the x-axis in the positive x-direction. When the particles reach the origin they encounter an electronic 'gate' which allows only those particles with a specified speed V to pass. These particles then move in a uniform magnetic field B_0 acting in the z-direction. Show that each particle will execute a semicircle before meeting the y-axis at a point which depends upon its mass. [This provides a method for determining the masses of the particles.]

4.33 ***The magnetron*** An electron of mass m and charge $-e$ is moving under the combined influence of a uniform electric field $E_0\boldsymbol{j}$ and a uniform magnetic field $B_0\boldsymbol{k}$. Initially the electron is at the origin and is moving with velocity $u\,\boldsymbol{i}$. Show that the trajectory of the electron is given by

$$x = a(\Omega t) + b\sin\Omega t, \qquad y = b(1-\cos\Omega t), \qquad z = 0,$$

where $\Omega = eB_0/m$, $a = E_0/\Omega B_0$ and $b = (uB_0 - E_0)/\Omega B_0$. Use computer assistance to plot typical paths of the electron for the cases $a < b$, $a = b$ and $a > b$. [The general path is called a *trochoid*, which becomes a *cycloid* in the special case $a = b$. Cycloidal motion of electrons is used in the **magnetron** vacuum tube, which generates the microwaves in a microwave oven.]

Computer assisted problems

4.34 Complete Example 4.7 on the projectile with linear resistance by obtaining the quoted asymptotic formula for the range of the projectile.

4.35 Find a series approximation for the period of the simple pendulum, in powers of the angular amplitude α. Proceed as follows:

The exact period τ of the pendulum was found in Example 4.9 and is given by the integral (4.9). This integral is not suitable for expansion as it stands. However, if we write $\cos\theta - \cos\alpha = 2(\sin^2(\alpha/2) - \sin^2(\theta/2))$ and make the sneaky substitution $\sin(\theta/2) = \sin(\alpha/2)\sin\phi$, the formula for τ becomes

$$\tau = 4\left(\frac{b}{g}\right)^{1/2}\int_0^{\pi/2}\left(1-\epsilon^2\sin^2\phi\right)^{-1/2}d\phi$$

where $\epsilon = \sin(\alpha/2)$. This new integrand is easy to expand as a power series in the variable ϵ and the limits of integration are now constants. Use computer assistance to expand the integrand to the required number of terms and then integrate term by term over the interval $[0, \pi/2]$. Finally re-expand as a power series in the variable α. The answer to two terms is given by equation (4.10), but it is just as easy to obtain any number of terms.

4.36 ***Baseball trajectory*** A baseball is struck with an initial speed of 45 m s^{-1} (just over 100 mph) at an elevation angle of 40°. Find its path and compare this with the corresponding path when air resistance is neglected. [A baseball has mass 0.30 kg and radius 3.5 cm. Assume the quadratic law of resistance.]

Show that the equation of motion can be written in the form

$$\frac{d\boldsymbol{v}}{dt} = -g\left(\boldsymbol{k} + \frac{\boldsymbol{v}|\boldsymbol{v}|}{V^2}\right),$$

where V is the terminal speed. Resolve this vector equation into two (coupled) scalar equations for v_x and v_z and perform a numerical solution. In this example, air resistance reduces the range by about 35%. It really is easier to hit a home run in Mile High stadium!

Chapter Five

Linear oscillations

and normal modes

KEY FEATURES

The key features of this chapter are the properties of **free undamped** oscillations, **free damped** oscillations, **driven** oscillations, and **coupled** oscillations.

Oscillations are a particularly important part of mechanics and indeed of physics as a whole. This is because of their widespread occurrence and the practical importance of oscillation problems. In this chapter we study the classical **linear theory** of oscillations, which is important for two reasons: (i) the linear theory usually gives a good approximation to the motion when the amplitude of the oscillations is small, and (ii) in the linear theory, most problems can be solved explicitly in closed form. The importance of this last fact should not be underestimated! We develop the theory in the context of the oscillations of a body attached to a spring, but the same equations apply to many different problems in mechanics and throughout physics.

In the course of this chapter we will need to solve linear second order ODEs with constant coefficients. For a description of the standard method of solution see Boyce & DiPrima [8].

5.1 BODY ON A SPRING

Suppose a body of mass m is attached to one end of a light spring. The other end of the spring is attached to a fixed point A on a smooth horizontal table, and the body slides on

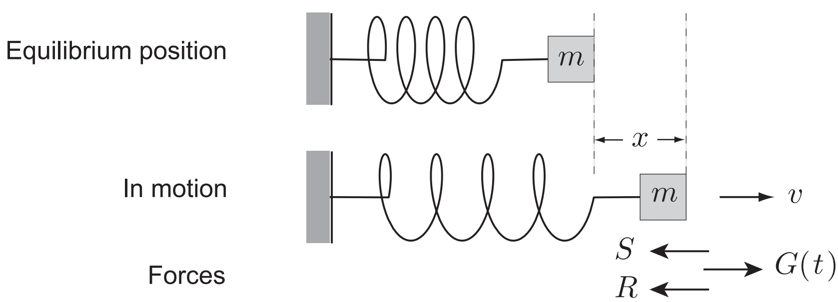


FIGURE 5.1 The body m is attached to one end of a light spring and moves in a straight line.

the table in a straight line through A. Let x be the displacement and v the velocity of the body at time t, as shown in Figure 5.1; note that x is measured from the *equilibrium position* of the body.

Consider now the forces acting on the body. When the spring is extended, it exerts a **restoring force** S in the opposite direction to the extension. Also, the body may encounter a **resistance force** R acting in the opposite direction to its velocity. Finally, there may be an external **driving force** $G(t)$ that is a specified function of the time. The equation of motion for the body is then

$$m\frac{dv}{dt} = -S - R + G(t). \tag{5.1}$$

The **restoring force** S is determined by the design of the spring and the extension x. For sufficiently *small strains*,* the relationship between S and x is approximately *linear*, that is,

$$S = \alpha x, \tag{5.2}$$

where α is a positive constant called the **spring constant** (or **strength**) of the spring. A powerful spring, such as those used in automobile suspensions, has a large value of α; the spring behind a doorbell has a small value of α. The formula (5.2) is called **Hooke's law**† and a spring that obeys Hooke's law exactly is called a **linear** spring.

The **resistance force** R depends on the physical process that is causing the resistance. For fluid resistance, the linear or quadratic resistance laws considered in Chapter 4 may be appropriate. However, neither of these laws represents the frictional force exerted by a rough table. In this chapter we assume the law of **linear resistance**

$$R = \beta v, \tag{5.3}$$

where β is a positive constant called the **resistance constant**; it is a measure of the strength of the resistance. There is no point in disguising the fact that our major reason for assuming linear resistance is that (together with Hooke's law) it leads to a linear equation of motion that can be solved explicitly. However, it does give insight into the general effect of all resistances, and actually is appropriate when the resistance arises from slow viscous flow (automobile shock absorbers, for instance); it is also appropriate in the electric circuit analogue, where it is equivalent to Ohm's law.

With **Hooke's law** and **linear resistance**, the equation of motion (5.1) for the body becomes

$$m\frac{d^2x}{dt^2} + \beta\frac{dx}{dt} + \alpha x = G(t),$$

* The strain is the extension of the spring divided by its natural length. If the strain is large, then the linear approximation will break down and a non-linear approximation, such as $S = ax + bx^3$ must be used instead.

† After Robert Hooke (1635–1703). Hooke was an excellent scientist, full of ideas and a first class experimenter, but he lacked the mathematical skills to develop his ideas. When other scientists (Newton in particular) did so, he accused them of stealing his work and this led to a succession of bitter disputes. So that his rivals could not immediately make use of his discovery, Hooke first published the law that bears his name as an anagram on the Latin phrase '*ut tensio, sic vis*' (as the extension, so the force).

where α is the spring constant, β is the resistance constant and $G(t)$ is the prescribed driving force. *This is a second order, linear ODE with constant coefficients for the unknown displacement* $x(t)$. We could go ahead with the solution of this equation as it stands, but the algebra is made much easier by introducing two new constants Ω and K (instead of α and β) defined by the relations

$$\alpha = m\Omega \qquad\qquad \beta = 2mK.$$

The equation of motion for the body then becomes

$$\boxed{\frac{d^2x}{dt^2} + 2K\,\frac{dx}{dt} + \Omega^2\,x = F(t)} \tag{5.4}$$

where $F(t) = G(t)/m$, the *driving force per unit mass*. This is the standard form of the **equation of motion** for the body. Any system that leads to an equation of this form is called a **damped*** **linear oscillator**. When the force $F(t)$ is absent, the oscillations are said to be **free**; when it is present, the oscillations are said to be **driven**.

5.2 CLASSICAL SIMPLE HARMONIC MOTION

A linear oscillator that is both **undamped** and **undriven** is called a **classical linear oscillator**. This is the simplest case, but arguably the most important system in physics! The equation (5.4) reduces to

$$\frac{d^2x}{dt^2} + \Omega^2 x = 0, \tag{5.5}$$

which, because of the solutions we are about to obtain, is called the **SHM equation**.

Solution procedure

Seek solutions of the form $x = e^{\lambda t}$. Then λ must satisfy the equation

$$\lambda^2 + \Omega^2 = 0,$$

which gives $\lambda = \pm i\Omega$. We have thus found the pair of complex solutions

$$x = e^{\pm i\Omega t},$$

which form a basis for the space of complex solutions. The real and imaginary parts of the first complex solution are

$$x = \begin{cases} \cos \Omega t \\ \sin \Omega t \end{cases}$$

* Damping is another term for resistance. Indeed, automobile shock absorbers are sometimes called dampers.

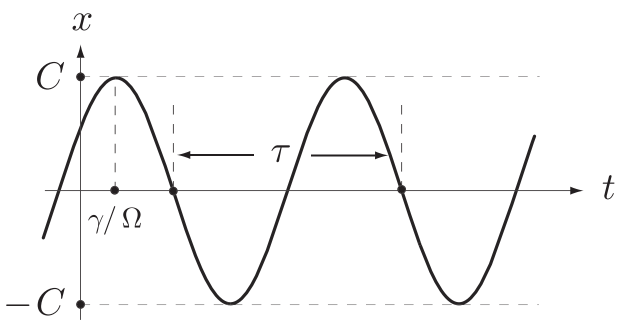


FIGURE 5.2 Classical simple harmonic motion
$x = C\cos(\Omega t - \gamma)$.

and these functions form a basis for the space of real solutions. The **general real solution** of the SHM equation is therefore

$$x = A\cos\Omega t + B\sin\Omega t, \tag{5.6}$$

where A and B are real arbitrary constants. This general solution can be written in the alternative form*

$$x = C\cos(\Omega t - \gamma), \tag{5.7}$$

where C and γ are real arbitrary constants with $C > 0$.

General form of the motion

The general form of the motion is most easily deduced from the form (5.7) and is shown in Figure 5.2. This is called **simple harmonic motion (SHM)**. The body makes infinitely many oscillations of constant **amplitude** C; the constant γ is simply a 'phase factor' which shifts the whole graph by γ/Ω in the t-direction. Since the cosine function repeats itself when the argument Ωt increases by 2π, it follows that the **period** of the oscillations is given by

$$\tau = \frac{2\pi}{\Omega}. \tag{5.8}$$

The quantity Ω, which is related to the frequency ν by $\Omega = 2\pi\nu$, is called the **angular frequency** of the oscillations.

Example 5.1 *An initial value problem for classical SHM*

A body of mass m is suspended from a fixed point by a light spring and can move under uniform gravity. In equilibrium, the spring is found to be extended by a distance b. Find the period of vertical oscillations of the body about this equilibrium position. [Assume small strains.]

* This transformation is based on the result from trigonometry that $a\cos\theta + b\sin\theta$ can always be written in the form $c\cos(\theta - \gamma)$, where $c = (a^2 + b^2)^{1/2}$ and $\tan\gamma = b/a$.

The body is hanging in its equilibrium position when it receives a sudden blow which projects it upwards with speed u. Find the subsequent motion.

Solution

When the spring is subjected to a constant force of magnitude mg, the extension is b. Hence α, the strength of the spring, is given by $\alpha = mg/b$.

Let z be the *downwards* displacement of the body from its equilibrium position. Then the extension of the spring is $b + z$ and the restoring force is $\alpha(b + z) = g(b + z)/b$. The equation of motion for the body is therefore

$$m\frac{d^2z}{dt^2} = mg - \frac{mg(b+z)}{b}$$

that is

$$\frac{d^2z}{dt^2} + \left(\frac{g}{b}\right) z = 0.$$

This is the **SHM equation** with $\Omega^2 = g/b$. It follows that the **period** τ of vertical oscillations about the equilibrium position is given by

$$\tau = \frac{2\pi}{\Omega} = 2\pi \left(\frac{b}{g}\right)^{1/2}.$$

In the **initial value problem**, the subsequent motion must have the form

$$x = A\cos\Omega t + B\sin\Omega t,$$

where $\Omega = (g/b)^{1/2}$. The initial condition $x = 0$ when $t = 0$ shows that $A = 0$ and the initial condition $\dot{x} = -u$ when $t = 0$ then gives $\Omega B = -u$, that is, $B = -u/\Omega$. The subsequent motion is therefore

$$x = -\frac{u}{\Omega}\sin\Omega t,$$

where $\Omega = (g/b)^{1/2}$. ■

5.3 DAMPED SIMPLE HARMONIC MOTION

When **damping** is present but there is no external force, the general equation (5.4) reduces to

$$\frac{d^2x}{dt^2} + 2K\frac{dx}{dt} + \Omega^2 x = 0, \tag{5.9}$$

the **damped SHM equation**.

The solution procedure is the same as in the last section. Seek solutions of the form $x = e^{\lambda t}$. Then λ must satisfy the equation

$$\lambda^2 + 2K\lambda + \Omega^2 = 0,$$

that is

$$(\lambda + K)^2 = K^2 - \Omega^2.$$

We see that *different cases arise depending on whether* $K < \Omega$, $K = \Omega$ *or* $K > \Omega$. These cases give rise to different kinds of solution and must be treated separately.

Under-damping (sub-critical damping): $K < \Omega$

In this case, we write the equation for λ in the form

$$(\lambda + K)^2 = -\Omega_D^2,$$

where $\Omega_D = (\Omega^2 - K^2)^{1/2}$, a *positive real number*. The λ values are then $\lambda = -K \pm i\Omega_D$. We have thus found the pair of complex solutions

$$x = e^{-Kt} e^{\pm i\Omega_D t},$$

which form a basis for the space of complex solutions. The real and imaginary parts of the first complex solution are

$$x = \begin{cases} e^{-Kt} \cos \Omega_D t \\ e^{-Kt} \sin \Omega_D t \end{cases}$$

and these functions form a basis for the space of real solutions. The **general real solution** of the damped SHM equation in this case is therefore

$$x = e^{-Kt} \left(A \cos \Omega_D t + B \sin \Omega_D t\right), \qquad (5.10)$$

where A and B are real arbitrary constants. This general solution can be written in the alternative form

$$x = Ce^{-Kt} \cos(\Omega_D t - \gamma), \qquad (5.11)$$

where C and γ are real arbitrary constants with $C > 0$.

General form of the motion

The general form of the motion is most easily deduced from the form (5.11) and is shown in Figure 5.3. This is called **under-damped SHM**. The body still executes infinitely many oscillations, but now they have *exponentially decaying amplitude* Ce^{-Kt}. Suppose the **period** τ of the oscillations is defined as shown in Figure 5.3.* The introduction of damping decreases the angular frequency of the oscillations from Ω to Ω_D, which *increases* the period of the oscillations from $2\pi/\Omega$ to

$$\tau = \frac{2\pi}{\Omega_D} = \frac{2\pi}{(\Omega^2 - K^2)^{1/2}}. \qquad (5.12)$$

* The period might also be defined as the time interval between successive maxima of the function $x(t)$. Since these maxima do *not* occur at the points at which $x(t)$ touches the bounding curves, it is not obvious that this time interval is even a constant. However, it *is* a constant and has the same value as (5.12) (see Problem 5.5)

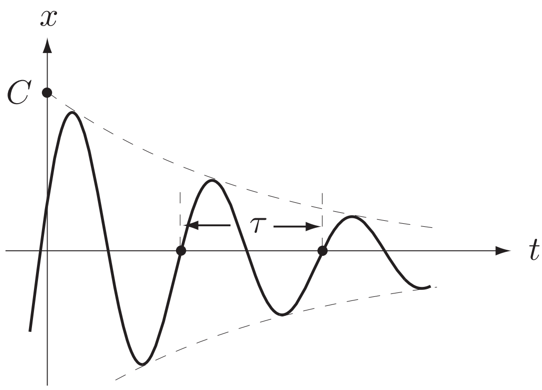


FIGURE 5.3 Under-damped simple harmonic motion $x = Ce^{-Kt}\cos(\Omega_D t - \gamma)$.

Over-damping (super-critical damping): $K > \Omega$

In this case, we write the equation for λ in the form

$$(\lambda + K)^2 = \delta^2,$$

where $\delta = (K^2 - \Omega^2)^{1/2}$, a *positive real number*. The λ values are then $\lambda = -k \pm \delta$, which are now real. We have thus found the pair of real solutions

$$x = e^{-Kt}e^{\pm\delta t},$$

which form a basis for the space of real solutions. The **general real solution** of the damped SHM equation in this case is therefore

$$x = e^{-Kt}\left(Ae^{\delta t} + Be^{-\delta t}\right), \tag{5.13}$$

where A and B are real arbitrary constants.

General form of the motion

Three typical forms for the motion are shown in Figure 5.4. This is called **over-damped SHM**. Somewhat surprisingly, *the body does not oscillate at all*. For example, if the body is released from rest, then it simply drifts back towards the equilibrium position. On the other hand, if the body is projected towards the equilibrium position with sufficient speed, then it passes the equilibrium position once and then drifts back towards it from the other side.

Critical damping: $K = \Omega$

The case of critical damping is solved in Problem 5.6. Qualitatively, the motions look like those in Figure 5.4.

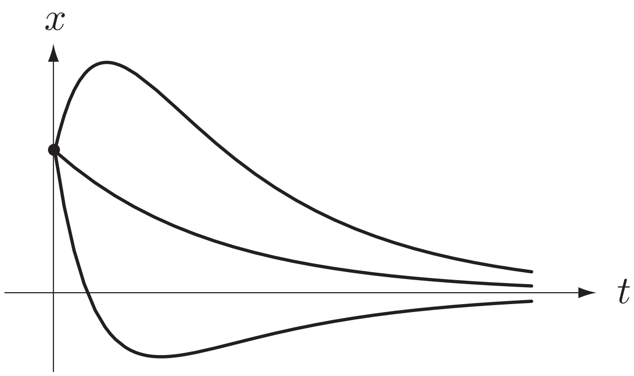


FIGURE 5.4 Three typical cases of over-damped simple harmonic motion.

5.4 DRIVEN (FORCED) MOTION

We now include the effect of an external **driving force** $G(t)$ which we suppose to be a *given* function of the time. In the case of a body suspended by a spring, we could apply such a force directly, but, in practice, the external 'force' often arises indirectly by virtue of the suspension point being made to oscillate in some prescribed way. The seismograph described in the next section is an instance of this.

Whatever the origin of the driving force, the **governing equation** for driven motion is (5.4), namely

$$\frac{d^2x}{dt^2} + 2K\frac{dx}{dt} + \Omega^2 x = F(t), \tag{5.14}$$

where $2mK$ is the damping constant, $m\Omega^2$ is the spring constant and $mF(t)$ is the driving force. Since this equation is linear and inhomogeneous, its general solution is the sum of (i) the general solution of the corresponding homogeneous equation (5.9) (the complementary function) and (ii) *any* particular solution of the inhomogeneous equation (5.14) (the particular integral). The complementary function has already been found in the last section, and it remains to find the particular integral for interesting choices of $F(t)$. Actually there is a (rather complicated) formula for a particular integral of this equation for *any choice* of the driving force $mF(t)$. However, the most important case by far is that of **time harmonic** forcing and, in this case, it is easier to find a particular integral directly. Time harmonic forcing is the case in which

$$F(t) = F_0 \cos pt, \tag{5.15}$$

where F_0 and p are positive constants; mF_0 is the amplitude of the applied force and p is its angular frequency.

Solution procedure

We first replace the forcing term $F_0 \cos pt$ by its complex counterpart $F_0 e^{ipt}$. This gives the complex equation

$$\frac{d^2x}{dt^2} + 2K\frac{dx}{dt} + \Omega^2 x = F_0 e^{ipt}. \tag{5.16}$$

We then seek a particular integral of this complex equation in the form

$$x = ce^{ipt}, \tag{5.17}$$

where c is a complex constant called the **complex amplitude**. On substituting (5.17) into equation (5.16) we find that

$$c = \frac{F_0}{\Omega^2 - p^2 + 2iKp}, \tag{5.18}$$

so that the complex function

$$\frac{F_0 e^{ipt}}{\Omega^2 - p^2 + 2iKp} \tag{5.19}$$

is a particular integral of the complex equation (5.16). A particular integral of the real equation (5.14) is then given by the real part of the complex expression (5.19). It follows that a **particular integral** of equation (5.14) is given by

$$x^D = a\cos(pt - \gamma),$$

where $a = |c|$ and $\gamma = -\arg c$. This particular integral, which is also time harmonic with the same frequency as the applied force, is called the **driven response** of the oscillator to the force $mF_0 \cos pt$; a is the **amplitude** of the driven response and γ $(0 < \gamma \le \pi)$ is the **phase angle** by which the response lags behind the force. From the expression (5.18) for c, it follows that

$$a = \frac{F_0}{\left((\Omega^2 - p^2)^2 + 4K^2p^2\right)^{1/2}}, \qquad \tan\gamma = \frac{2Kp}{\Omega^2 - p^2}. \tag{5.20}$$

The **general solution** of equation (5.14) therefore has the form

$$x = a\cos(pt - \gamma) + x^{CF}, \tag{5.21}$$

where x^{CF} is the complementary function, that is, the general solution of the corresponding *undriven* problem.

The undriven problem has already been solved in the last section. The solution took three different forms depending on whether the damping was supercritical, critical or subcritical. However, all these forms have one feature in common, that is, *they all decay to zero with increasing time*. For this reason, the complementary function for this equation is often called the **transient response** of the oscillator. Any solution of equation (5.21) is therefore the sum of the driven response x^D (which persists) and a transient response x^{CF} (which dies away). Thus, *no matter what the initial conditions, after a suffcently long time we are left with just the driven response*. In many problems, the transient response can be disregarded, but it must be included if inital conditions are to be satisfied.

Example 5.2 *An initial value problem for driven motion*

The equation of motion of a certain driven damped oscillator is

$$\frac{d^2x}{dt^2} + 3\frac{dx}{dt} + 2x = 10\cos t$$

and initially the particle is at rest at the origin. Find the subsequent motion.

Solution

First we find the **driven response** x^D. The complex counterpart of the equation of motion is

$$\frac{d^2x}{dt^2} + 3\frac{dx}{dt} + 2x = 10e^{it}$$

and we seek a solution of this equation of the form $x = ce^{it}$. On substituting in, we find that

$$c = \frac{10}{1+3i} = 1 - 3i.$$

It follows that the **driven response** x^D is given by

$$x^D = \Re\left[(1-3i)e^{it}\right] = \cos t + 3\sin t.$$

Now for the **complementary function** x^{CF}. This is the general solution of the corresponding undriven equation

$$\frac{d^2x}{dt^2} + 3\frac{dx}{dt} + 2x = 0,$$

which is easily found to be

$$x = Ae^{-t} + Be^{-2t},$$

where A and B are arbitrary constants. The **general solution** of the equation of motion is therefore

$$x = \cos t + 3\sin t + Ae^{-t} + Be^{-2t}.$$

It now remains to choose A and B so that the **initial conditions** are satisfied. The condition $x = 0$ when $t = 0$ implies that

$$0 = 1 + A + B,$$

and the condition $\dot{x} = 0$ when $t = 0$ implies that

$$0 = 3 - A - 2B.$$

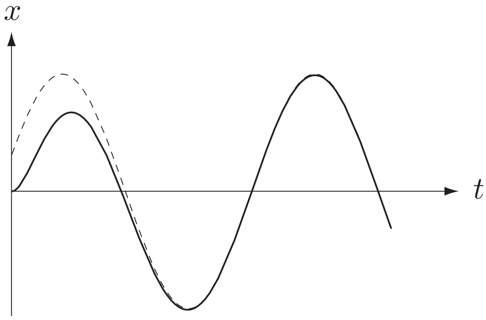


FIGURE 5.5 The solid curve is the actual response and the dashed curve the driven response only.

Solving these simultaneous equations gives $A = -5$ and $B = 4$. The **subsequent motion** of the oscillator is therefore given by

$$x = \underbrace{\cos t + 3\sin t}_{\text{driven response}} \underbrace{-5e^{-t} + 4e^{-2t}}_{\text{transient response}}.$$

This solution is shown in Figure 5.5 together with the driven response only. In this case, the transient response is insignificant after less than one cycle of the driving force. The **amplitude** of the driven reponse is $(1^2 + 3^2)^{1/2} = \sqrt{10}$ and the **phase lag** is $\tan^{-1}(3/1) \approx 72°$. ■

Resonance of an oscillating system

Consider the general formula

$$a = \frac{F_0}{\left((\Omega^2 - p^2)^2 + 4K^2p^2\right)^{1/2}}$$

for the amplitude a of the driven response to the force $mF_0 \cos pt$ (see equation (5.20)). Suppose that the amplitude of the applied force, the spring constant, and the resistance constant are held fixed and that the *angular frequency* p of the applied force is varied. Then a is a function of p only. Which value of p produces the largest driven response? Let

$$f(q) = (\Omega^2 - q)^2 + 4K^2q.$$

Then, since $a = F_0/\sqrt{f(p^2)}$, we need only find the minimum point of the function $f(q)$ lying in $q > 0$. Now

$$f'(q) = -2(\Omega^2 - q) + 4K^2 = 2\left(q - (\Omega^2 - 2K^2)\right)$$

so that $f(q)$ decreases for $q < \Omega^2 - 2K^2$ and increases for $q > \Omega^2 - 2K^2$. Hence $f(q)$ has a unique minimum point at $q = \Omega^2 - 2K^2$. Two cases arise depending on whether this value is positive or not.

Case **1**. When $\Omega^2 > 2K^2$, the minimum point $q = \Omega^2 - 2K^2$ is positive and a has its maximum value when $p = p^R$, where

$$p^R = (\Omega^2 - 2K^2)^{1/2}.$$

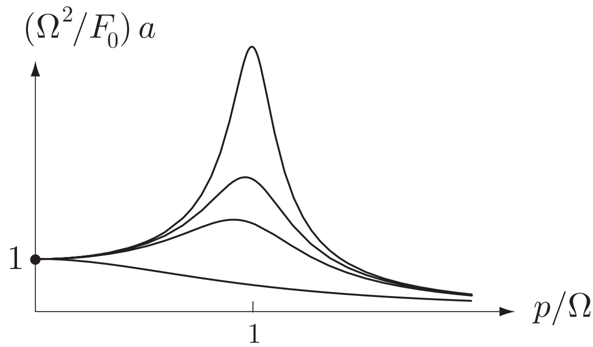


FIGURE 5.6 The dimensionless amplitude $(F_0/\Omega^2)a$ against the dimensionless driving frequency p/Ω for (from the top) $K/\Omega = 0.1, 0.2, 0.3, 1$.

The angular frequency p^R is called the **resonant frequency** of the oscillator. The value of a at the resonant frequency is

$$a_{\text{max}} = \frac{F_0}{2K(\Omega^2 - K^2)^{1/2}}.$$

Case **2**. When $\Omega^2 \le 2K^2$, a is a decreasing funcion of p for $p > 0$ so that a has *no maximum point*.

These results are illustrated in Figure 5.6. They are an example of the general physical phenomenon known as **resonance**, which can be loosely stated as follows:

The phenomenon of resonance

Suppose that, in the absence of damping, a physical system can perform free oscillations with angular frequency Ω. Then a driving force with angular frequency p will induce a large response in the system when p is close to Ω, providing that the damping is not too large.

This principle does not just apply to the mechanical systems we study here. It is a general physical principle that also applies, for example, to the oscillations of electric currents in circuits and to the quantum mechanical oscillations of atoms.

Note that the resonant frequency p^R is always less than Ω, but is *close* to Ω when K/Ω is small. The height of the resonance peak, a_{max}, is given approximately by

$$a_{\text{max}} \sim \frac{F_0}{2\Omega^2}\left(\frac{\Omega}{K}\right)^{-1}$$

in the limit in which K/Ω is small; a_{max} therefore tends to infinity in this limit. In the same limit, the width of the resonance peak is directly proportional to K/Ω and consequently tends to zero.

General periodic driving force

The method we have developed for the time harmonic driving force can be extended to any periodic driving force $mF(t)$. A function $f(t)$ is said to be **periodic** with period τ if the values taken by f in any interval of length τ are then repeated in the next interval of length τ. An example is the 'square wave' function shown in Figure 5.7. The solution method requires that $F(t)$ be expanded as a **Fourier series**.* A textbook on mechanics is not the place to develop the theory of Fourier series. Instead we will simply quote the essential results and then give an example of how the method works. To keep the algebra as simple as possible, we will suppose that the driving force has period 2π.†

Fourier's Theorem

Fourier's theorem states that any function $f(t)$ that is periodic with period 2π can be expanded as a **Fourier series** in the form

$$f(t) = \tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos nt + b_n \sin nt, \qquad (5.22)$$

where the **Fourier coefficients** $\{a_n\}$ and $\{b_n\}$ are given by the formulae

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos nt\, dt, \qquad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin nt\, dt. \qquad (5.23)$$

What this means is that *any* function $f(t)$ with period 2π can be expressed as a sum of *time harmonic* terms, each of which has period 2π. In order to find the driven response of the oscillator when the force $mF(t)$ is applied, we first expand $F(t)$ in a Fourier series. We then find the driven response that would be induced by each of the terms of this Fourier series applied separately, and then simply add these responses together. The method depends on the equation of motion being *linear*.

Example 5.3 ***Periodic non-harmonic driving force***

Find the driven response of the damped linear oscillator

$$\frac{d^2x}{dt^2} + 2K\frac{dx}{dt} + \Omega^2 x = F(t)$$

for the case in which $F(t)$ is periodic with period 2π and takes the values

$$F(t) = \begin{cases} F_0 & (0 < t < \pi), \\ -F_0 & (\pi < t < 2\pi), \end{cases}$$

* After Jean Baptiste Joseph Fourier 1768–1830. The memoir in which he developed the theory of trigonometric series '*On the Propagation of Heat in Solid Bodies*' was submitted for the mathematics prize of the Paris Institute in 1811; the judges included such luminaries as Lagrange, Laplace and Legendre. They awarded Fourier the prize but griped about his lack of mathematical rigour.

† The general case can be reduced to this one by a scaling of the unit of time.

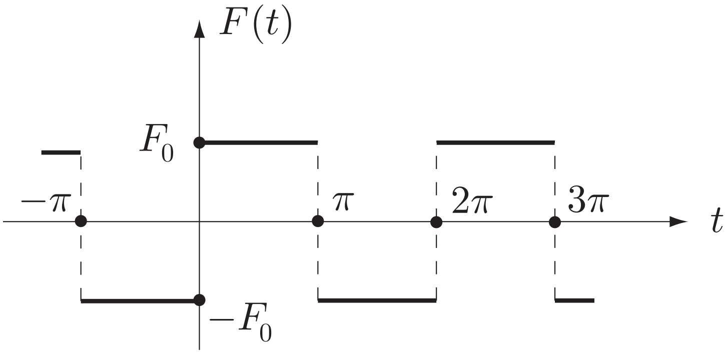


FIGURE 5.7 The 'square wave' input function $F(t)$ is periodic with period 2π. Its value alternates between $\pm F_0$.

in the interval $0 < t < 2\pi$. This function* is shown in Figure 5.7.

Solution

The first step is to find the Fourier series of the function $F(t)$. From the formula (5.23), the coefficient a_n is given by

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} F(t)\cos nt\,dt = \frac{1}{\pi}\int_{-\pi}^{0}(-F_0)\cos nt\,dt + \frac{1}{\pi}\int_{0}^{\pi}(+F_0)\cos nt\,dt \\ &= 0, \end{aligned}$$

since both integrals are zero for $n \ge 1$ and are equal and opposite when $n = 0$. In the same way,

$$\begin{aligned} b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} F(t)\sin nt\,dt = \frac{1}{\pi}\int_{-\pi}^{0}(-F_0)\sin nt\,dt + \frac{1}{\pi}\int_{0}^{\pi}(+F_0)\sin nt\,dt \\ &= \frac{2F_0}{\pi}\int_0^{\pi}\sin nt\,dt, \end{aligned}$$

since this time the two integrals are equal. Hence

$$\begin{aligned} b_n &= \frac{2F_0}{\pi}\left[\frac{-\cos nt}{n}\right]_0^{\pi} = \frac{2F_0}{\pi}\left(\frac{1-\cos n\pi}{n}\right) \\ &= \frac{2F_0}{\pi}\left(\frac{1-(-1)^n}{n}\right). \end{aligned}$$

Hence the **Fourier series** of the function $F(t)$ is

$$F(t) = \sum_{n=1}^{\infty}\frac{2F_0}{\pi}\left(\frac{1-(-1)^n}{n}\right)\sin nt.$$

* This function is the mechanical equivalent of a 'square wave input' in electric circuit theory.

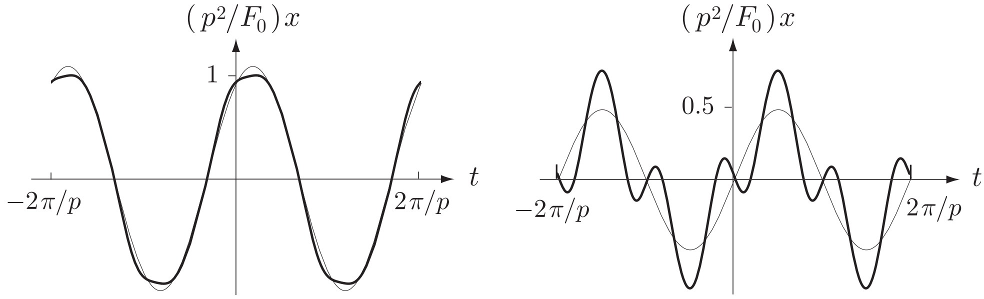


FIGURE 5.8 Driven response of a damped oscillator to the alternating constant force $\pm mF_0$ with angular frequency p : **Left** $\Omega/p = 1.5$, $K/p = 1$. **Right** $\Omega/p = 2.5$, $K/p = 0.1$. The light graphs show the first term of the exapansion series.

The next step is to find the driven response of the oscillator to the force $m(b_n \sin nt)$, that is, the particular integral of the equation

$$\frac{d^2x}{dt^2} + 2K\frac{dx}{dt} + \Omega^2 x = b_n \sin nt. \tag{5.24}$$

The complex counterpart of this equaton is

$$\frac{d^2x}{dt^2} + 2K\frac{dx}{dt} + \Omega^2 x = b_n e^{int}$$

for which the particular integral is ce^{int}, where the complex amplitude c is given by

$$c = \frac{b_n}{\Omega^2 - n^2 + 2iK}.$$

The particular integral of the real equation (5.24) is then given by

$$\Im\left(\frac{b_n e^{int}}{\Omega^2 - n^2 + 2iKn}\right) = b_n\left(\frac{(\Omega^2 - n^2)\sin nt + 2Kn\cos nt}{(\Omega^2 - n^2)^2 + 4K^2n^2}\right).$$

Finally we add together these separate responses to find the **driven response** of the oscillator to the force $mF(t)$. On inserting the value of the coefficient b_n, this gives

$$x = \frac{2F_0}{\pi}\sum_{n=1}^{\infty}\left(\frac{1-(-1)^n}{n}\right)\left(\frac{(\Omega^2 - n^2)\sin nt + 2Kn\cos nt}{(\Omega^2 - n^2)^2 + 4K^2n^2}\right). \tag{5.25}$$

In order to deduce anything from this complicated formula, we must either sum the series numerically or approximate the formula in some way. When Ω and K are both small compared to the forcing frequency p, the series (5.25) converges quite quickly and can be approximated (to within a few percent) by the first term. Even when $\Omega/p = 1.5$ and $K/p = 1$, this is still a reasonable approximation (see Figure 5.8 (left)). However, for larger values of Ω/p, the higher harmonics in the Fourier expansion of $F(t)$ that have frequencies close to Ω produce large contributions (see Figure 5.8 (right)). In this case, the series (5.25) must be summed numerically. ■

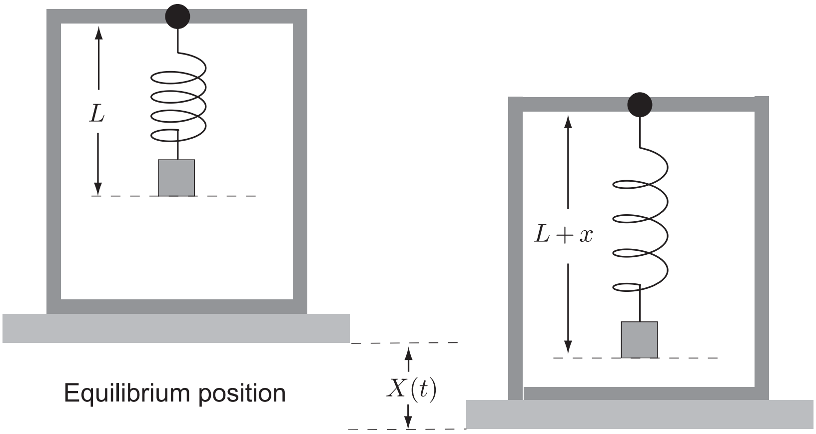


FIGURE 5.9 A simple seismograph for measuring vertical ground motion.

5.5 A SIMPLE SEISMOGRAPH

The seismograph is an instrument that measures the motion of the ground on which it stands. In real earthquakes, the ground motion will generally have both vertical and horizontal components, but, for simplicity, we describe here a device for measuring **vertical motion** only.

Our simple seismograph (see Figure 5.9) consists of a mass which is suspended from a rigid support by a spring; the motion of the mass relative to the support is resisted by a damper. The support is attached to the ground so that the suspension point has the same motion as the ground below it. This motion sets the suspended mass moving and the resulting spring extension is measured as a function of the time. Can we deduce what the ground motion was?

Suppose the ground (and therefore the support) has downward displacement $X(t)$ at time t and that the extension $x(t)$ of the spring is measured from its equilibrium length. Then the *displacement* of the mass is $x + X$, relative to an inertial frame. The equation of motion (5.9) is therefore modified to become

$$m\frac{d^2(x+X)}{dt^2} = -(2mK)\frac{dx}{dt} - (m\Omega^2)x,$$

that is,

$$\frac{d^2x}{dt^2} + 2K\frac{dx}{dt} + \Omega^2 x = -\frac{d^2X}{dt^2}.$$

This means *that the motion of the body relative to the moving support is the same as if the support were fixed and the external driving force* $-m(d^2X/dt^2)$ *were applied to the body.*

First consider the driven response of our seismograph to a train of harmonic waves with amplitude A and angular frequency p, that is,

$$X = A\cos pt.$$

The equation of motion for the spring extension x is then

$$\frac{d^2x}{dt^2} + 2K\frac{dx}{dt} + \Omega^2 x = Ap^2\cos pt.$$

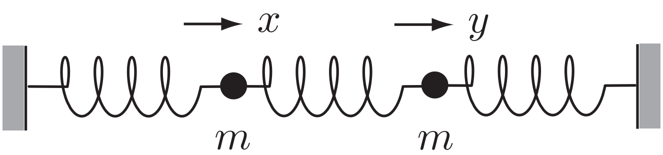


FIGURE 5.10 Two particles are connected between three springs and perform longitudinal oscillations.

The complex amplitude of the driven motion is

$$c = \frac{p^2 A}{-p^2 + 2iKp + \Omega^2},$$

and the real driven motion is

$$x = a\cos(pt - \gamma),$$

where

$$a = |c| = \frac{A}{\left|-1 + 2i(K/p) + (\Omega/p)^2\right|}. \tag{5.26}$$

Thus, providing that the spring and resistance constants are accurately known, the angular frequency p and amplitude A of the incident wave train can be deduced.

In practice, things may not be so simple. In particular, the incident wave train may be a mixture of harmonic waves with different amplitudes and frequencies, and these are not easily disentangled. However, if K and Ω are chosen so that K/p and Ω/p are small compared with unity (for all likely values of p), then $c = -A$ and $X = -x$ approximately. Thus, in this case, the record for $x(t)$ is simply the negative of the ground motion $X(t)$.* Since this result is independent of the incident frequency, it should also apply to complicated inputs such as a pulse of waves.

5.6 COUPLED OSCILLATIONS AND NORMAL MODES

Interesting new effects occur when two or more oscillators are coupled together. Figure 5.10 shows a typical case in which two bodies are connected between three springs and the motion takes place in a straight line. We restrict ourselves here to the classical theory in which the *restoring forces are linear and damping is absent*. If the springs are non-linear, then the displacements of the particles must be small enough so that the linear approximation is adequate.

Let x and y be the displacements of the two bodies from their respective equilibrium positions at time t; because two coordinates are needed to specify the configuration, the system is said to have two *degrees of freedom*. Then, at time t, the extensions of the three springs are x, $y - x$ and $-y$ respectively. Suppose that the strengths of the three springs are α, 2α and

* What is actually happening is that the mass is hardly moving at all (relative to an inertial frame).

4α respectively. Then the three restoring forces are αx, $2\alpha(y - x)$, $-4\alpha y$ and the equations of motion for the two bodies are

$$m\ddot{x} = -\alpha x + 2\alpha(y - x),$$
$$m\ddot{y} = -2\alpha(y - x) - 4\alpha y,$$

which can be written in the form

$$\begin{aligned} \ddot{x} + 3n^2 x - 2n^2 y &= 0, \\ \ddot{y} - 2n^2 x + 6n^2 y &= 0, \end{aligned} \tag{5.27}$$

where the positive constant n is defined by $n^2 = \alpha/m$. These are the **governing equations** for the motion. They are a *pair of simultaneous second order homogeneous linear ODEs* with constant coefficients. The equations are **coupled** in the sense that both unknown functions appear in each equation; thus neither equation can be solved on its own.

The solution procedure: normal modes

The solution procedure is simply an extension of the usual method for finding the complementary function for a single homogeneous linear ODE with constant coefficients. However, rather than seek solutions in exponential form, it is simpler to seek solutions directly in the trigonometric form

$$\begin{aligned} x &= A\cos(\omega t - \gamma), \\ y &= B\cos(\omega t - \gamma), \end{aligned} \tag{5.28}$$

where A, B, ω and γ are constants. A solution of the governing equations (5.27) that has the form (5.28) is called a **normal mode** of the oscillating system. In a normal mode, all the coordinates that specify the configuration of the system vary harmonically in time with the *same frequency* and the *same phase*; however, they generally have *different amplitudes*. On substituting the normal mode form (5.28) into the governing equations (5.27), we obtain

$$-\omega^2 A\cos(\omega t - \gamma) + 3n^2 A\cos(\omega t - \gamma) - 2n^2 B\cos(\omega t - \gamma) = 0,$$
$$-\omega^2 B\cos(\omega t - \gamma) - 2n^2 A\cos(\omega t - \gamma) + 6n^2 B\cos(\omega t - \gamma) = 0,$$

which simplifies to give

$$\begin{aligned} (3n^2 - \omega^2)A - 2n^2 B \quad &= 0, \\ -2n^2 A + (6n^2 - \omega^2)B &= 0, \end{aligned} \tag{5.29}$$

a pair of *simultaneous linear algebraic equations* for the amplitudes A and B. Thus a normal mode will exist if we can find constants A, B and ω so that the equations (5.29) are satisfied. Since the equations are homogeneous, they always have the *trivial solution* $A = B = 0$, whatever the value of ω. However, the trivial solution corresponds to the *equilibrium solution* $x = y = 0$ of the governing equations (5.27), which is not a motion at all. We therefore require the equations (5.29) to have a **non-trivial solution** for A, B. There is a simple condition that this should be so, namely that the determinant of the system of equations should be zero, that is,

$$\det\begin{pmatrix} 3n^2-\omega^2 & -2n^2 \\ -2n^2 & 6n^2-\omega^2 \end{pmatrix} = 0. \tag{5.30}$$

On simplification, this gives the condition

$$\omega^4 - 9n^2\omega^2 + 14n^4 = 0, \tag{5.31}$$

a quadratic equation in the variable ω^2. If this equation has *real positive* roots ω_1^2, ω_2^2, then, for *each* of these values, the linear equations (5.29) will have a non-trivial solution for the amplitudes A, B. In the present case, the equation (5.31) factorises and the roots are found to be

$$\omega_1^2 = 2n^2, \qquad\qquad \omega_2^2 = 7n^2. \tag{5.32}$$

Hence there are **two normal modes** with (angular) frequencies $\sqrt{2}n$ and $\sqrt{7}n$ respectively. These frequencies are known as the **normal frequencies** of the oscillating system.

Slow mode: In the slow mode we have $\omega^2 = 2n^2$ so that the linear equations (5.29) become

$$\begin{aligned} n^2A - 2n^2B &= 0, \\ -2n^2A + 4n^2B &= 0. \end{aligned}$$

These two equations are each equivalent to the single equation $A = 2B$. This is to be expected since, if the equations were linearly independent, then there would be no non-trivial solution for A and B. We have thus found a family of non-trivial solutions $A = 2\delta$, $B = \delta$, where δ can take any (non-zero) value. Thus the *amplitude of the normal mode is not uniquely determined*; this happens because the governing ODEs are linear and homogeneous. The **slow normal mode** therefore has the form

$$\begin{aligned} x &= 2\delta\cos(\sqrt{2}\,nt-\gamma), \\ y &= \ \ \delta\cos(\sqrt{2}\,nt-\gamma), \end{aligned} \tag{5.33}$$

where the amplitude factor δ and phase factor γ can take any values. We see that, in the slow mode, the two bodies always move in the *same* direction with the body on the left having twice the amplitude as the body on the right.

Fast mode: In the fast mode we have $\omega^2 = 7n^2$ and, by following the same procedure, we find that the form of the **fast normal mode** is

$$\begin{aligned} x &= \ \ \delta\cos(\sqrt{7}\,nt-\gamma), \\ y &= -2\delta\cos(\sqrt{7}\,nt-\gamma), \end{aligned} \tag{5.34}$$

where the amplitude factor δ and phase factor γ can take any values. We see that, in the fast mode, the two bodies always move in *opposite* directions with the body on the right having twice the amplitude as the body on the left.

The general motion

Since the governing equations (5.27) are linear and homogeneous, a sum of normal mode solutions is also a solution. Indeed, the general solution can be written as a sum of normal

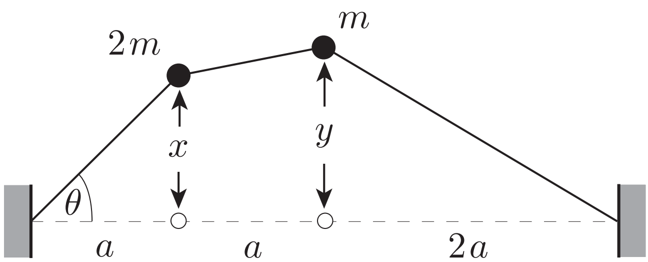


FIGURE 5.11 The two particles are attached to a light stretched string and perform *small* transverse oscillations. The displacements are shown to be large for clarity.

modes. Consider the expression

$$\begin{aligned} x &= 2\delta_1 \cos(\sqrt{2}\, nt - \gamma_1) + \ \delta_2 \cos(\sqrt{7}\, nt - \gamma_2), \\ y &= \ \delta_1 \cos(\sqrt{2}\, nt - \gamma_1) - 2\delta_2 \cos(\sqrt{7}\, nt - \gamma_2). \end{aligned} \tag{5.35}$$

This is simply a sum of the first normal mode (with amplitude factor δ_1 and phase factor γ_1) and the second normal mode (with amplitude factor δ_2 and phase factor γ_2). Since it is possible to choose these *four* arbitrary constants so that x, y, $\dot{x}$, $\dot{y}$ take any set of assigned values when $t = 0$, this must be the **general solution** of the governing equations (5.27).

Question ***Periodicity of the general motion***

Is the general motion periodic?

Answer

The general motion is a sum of normal mode motions with periods τ_1, τ_2 respectively. This sum will be periodic with period τ if (and only if) τ is an integer multiple of both τ_1 and τ_2, that is, if τ_1/τ_2 is a *rational* number. (In this case, the periods are said to be *commensurate*.) This in turn requires that ω_1/ω_2 is a rational number. In the present case, $\omega_1/\omega_2 = (2/7)^{1/2}$, which is irrational. The general motion is therefore **not periodic** in this case. ■

We conclude by solving another typical normal mode problem.

Example 5.4 *Small transverse oscillations*

Two particles P and Q, of masses $2m$ and m, are secured to a light string that is stretched to tension T_0 between two fixed supports, as shown in Figure 5.11. The particles undergo *small transverse oscillations* perpendicular to the equilibrium line of the string. Find the normal frequencies, the forms of the normal modes, and the general motion of this system. Is the general motion periodic?

Solution

First we need to make some simplifying assumptions.* We will assume that the transverse displacements x, y of the two particles are small compared with a; the three sections of the string then make small angles with the equilibrium line. We will also neglect any change in the tensions of the three sections of string.

The left section of string then has constant tension T_0. When the particle P is displaced, this tension force has the transverse component $-T_0 \sin\theta$, which acts as a restoring force on P; since θ is small, this component is approximately $-T_0 x/a$. Similar remarks apply to the other sections of string. The **equations of transverse motion** for P and Q are therefore

$$
\begin{aligned}
2m\ddot{x} &= -\frac{T_0 x}{a} + \frac{T_0(y-x)}{a},\\
m\ddot{y} &= -\frac{T_0(y-x)}{a} - \frac{T_0 y}{2a}.
\end{aligned}
$$

which can be written in the form

$$2\ddot{x} + 2n^2 x - n^2 y = 0, \tag{5.36}$$

$$2\ddot{y} - 2n^2 x + 3n^2 y = 0, \tag{5.37}$$

where the positive constant n is defined by $n^2 = T_0/ma$.

These equations will have **normal mode** solutions of the form

$$
\begin{aligned}
x &= A\cos(\omega t - \gamma),\\
y &= B\cos(\omega t - \gamma),
\end{aligned}
$$

when the simultaneous linear equations

$$
\left.
\begin{aligned}
(2n^2 - 2\omega^2)A - n^2 B \qquad &= 0,\\
-2n^2 A + (3n^2 - 2\omega^2)B &= 0,
\end{aligned}
\right. \tag{5.38}
$$

have a non-trivial solution for the amplitudes A, B. The condition for this is

$$\det\begin{pmatrix} 2n^2 - 2\omega^2 & -n^2 \\ -n^2 & 3n^2 - 2\omega^2 \end{pmatrix} = 0. \tag{5.39}$$

On simplification, this gives

$$2\omega^4 - 5n^2\omega^2 + 2n^4 = 0, \tag{5.40}$$

a quadratic equation in the variable ω^2. This equation factorises and the roots are found to be

$$\omega_1^2 = \tfrac{1}{2}n^2, \qquad\qquad \omega_2^2 = 2n^2. \tag{5.41}$$

Hence there are **two normal modes** with **normal frequencies** $n/\sqrt{2}$ and $\sqrt{2}n$ respectively.

* These assumptions are consistent with the more complete treatment given in Chapter 15.

Slow mode: In the slow mode we have $\omega^2 = n^2/2$ so that the linear equations (5.38) become

$$n^2 A - n^2 B = 0,$$
$$-2n^2 A + 2n^2 B = 0.$$

These two equations are each equivalent to the single equation $A = B$ so that we have the family of non-trivial solutions $A = \delta$, $B = \delta$, where δ can take any (non-zero) value. The **slow normal mode** therefore has the form

$$\begin{aligned} x &= \delta \cos(nt/\sqrt{2} - \gamma), \\ y &= \delta \cos(nt/\sqrt{2} - \gamma), \end{aligned} \tag{5.42}$$

where the amplitude factor δ and phase factor γ can take any values. We see that, in the slow mode, the two particles always have the *same displacement*.

Fast mode: In the fast mode we have $\omega^2 = 2n^2$ and, by following the same procedure, we find that the form of the **fast normal mode** is

$$\begin{aligned} x &= \delta \cos(\sqrt{2}\, nt - \gamma), \\ y &= -2\delta \cos(\sqrt{2}\, nt - \gamma), \end{aligned} \tag{5.43}$$

where the amplitude factor δ and phase factor γ can take any values. We see that, in the fast mode, the two particles always move in *opposite* directions with Q having twice the amplitude of P.

The **general motion** is now the sum of the first normal mode (with amplitude factor δ_1 and phase factor γ_1) and the second normal mode (with amplitude factor δ_2 and phase factor γ_2). This gives

$$\begin{aligned} x &= \delta_1 \cos(nt/\sqrt{2} - \gamma_1) + \delta_2 \cos(\sqrt{2}\, nt - \gamma_2), \\ y &= \delta_1 \cos(nt/\sqrt{2} - \gamma_1) - 2\delta_2 \cos(\sqrt{2}\, nt - \gamma_2). \end{aligned} \tag{5.44}$$

For this system $\tau_1/\tau_2 = \omega_2/\omega_1 = 2$ so that the general motion is **periodic** with period $\tau_1 = 2\sqrt{2}\pi/n$. ■

Problems on Chapter 5

Answers and comments are at the end of the book.

Harder problems carry a star (∗).

Free linear oscillations

5.1 A certain oscillator satisfies the equation

$$\ddot{x} + 4x = 0.$$

Initially the particle is at the point $x = \sqrt{3}$ when it is projected towards the origin with speed 2. Show that, in the subsequent motion,

$$x = \sqrt{3} \cos 2t - \sin 2t.$$

Deduce the amplitude of the oscillations. How long does it take for the particle to first reach the origin?

5.2 When a body is suspended from a fixed point by a certain linear spring, the angular frequency of its vertical oscillations is found to be Ω_1. When a different linear spring is used, the oscillations have angular frequency Ω_2. Find the angular frequency of vertical oscillations when the two springs are used together (i) in parallel, and (ii) in series. Show that the first of these frequencies is at least twice the second.

5.3 A particle of mass m moves along the x-axis and is acted upon by the restoring force $-m(n^2 + k^2)x$ and the resistance force $-2mk\dot{x}$, where n, k are positive constants. If the particle is released from rest at $x = a$, show that, in the subsequent motion,

$$x = \frac{a}{n} e^{-kt} (n \cos nt + k \sin nt).$$

Find how far the particle travels before it next comes to rest.

5.4 An overdamped harmonic oscillator satisfies the equation

$$\ddot{x} + 10\dot{x} + 16x = 0.$$

At time $t = 0$ the particle is projected from the point $x = 1$ towards the origin with speed u. Find x in the subsequent motion.

Show that the particle will reach the origin at some later time t if

$$\frac{u - 2}{u - 8} = e^{6t}.$$

How large must u be so that the particle will pass through the origin?

5.5 A damped oscillator satisfies the equation

$$\ddot{x} + 2K\dot{x} + \Omega^2 x = 0$$

where K and Ω are positive constants with $K < \Omega$ (under-damping). At time $t = 0$ the particle is released from rest at the point $x = a$. Show that the subsequent motion is given by

$$x = ae^{-Kt} \left(\cos \Omega_D t + \frac{K}{\Omega_D} \sin \Omega_D t \right),$$

where $\Omega_D = (\Omega^2 - K^2)^{1/2}$.

Find all the turning points of the function $x(t)$ and show that the ratio of successive maximum values of x is $e^{-2\pi K/\Omega_D}$.

A certain damped oscillator has mass 10 kg, period 5 s and successive maximum values of its displacement are in the ratio 3 : 1. Find the values of the spring and damping constants α and β.

5.6 ***Critical damping*** Find the general solution of the damped SHM equation (5.9) for the special case of critical damping, that is, when $K = \Omega$. Show that, if the particle is initially released from rest at $x = a$, then the subsequent motion is given by

$$x = ae^{-\Omega t}\,(1 + \Omega t)\,.$$

Sketch the graph of x against t.

5.7* ***Fastest decay*** The oscillations of a galvanometer satisfy the equation

$$\ddot{x} + 2K\dot{x} + \Omega^2 x = 0.$$

The galvanometer is released from rest with $x = a$ and we wish to bring the reading permanently within the interval $-\epsilon a \le x \le \epsilon a$ as quickly as possible, where ϵ is a small positive constant. What value of K should be chosen? One possibility is to choose a sub-critical value of K such that the first minimum point of $x(t)$ occurs when $x = -\epsilon a$. [Sketch the graph of $x(t)$ in this case.] Show that this can be acheived by setting the value of K to be

$$K = \Omega\left[1 + \left(\frac{\pi}{\ln(1/\epsilon)}\right)^2\right]^{-1/2}.$$

If K has this value, show that the time taken for x to reach its first minimum is approximately $\Omega^{-1}\ln(1/\epsilon)$ when ϵ is small.

5.8 A block of mass M is connected to a second block of mass m by a linear spring of natural length $8a$. When the system is in equilibrium with the first block on the floor, and with the spring and second block vertically above it, the length of the spring is $7a$. The upper block is then pressed down until the spring has half its natural length and is then resleased from rest. Show that the lower block will leave the floor if $M < 2m$. For the case in which $M = 3m/2$, find when the lower block leaves the floor.

Driven linear oscillations

5.9 A block of mass 2 kg is suspended from a fixed support by a spring of strength $2000\ \mathrm{N\,m^{-1}}$. The block is subject to the vertical driving force $36\cos pt$ N. Given that the spring will yield if its extension exceeds 4 cm, find the range of frequencies that can safely be applied.

5.10 A driven oscillator satisfies the equation

$$\ddot{x} + \Omega^2 x = F_0\cos[\Omega(1+\epsilon)t],$$

where ϵ is a positive constant. Show that the solution that satisfies the initial conditions $x = 0$ and $\dot{x} = 0$ when $t = 0$ is

$$x = \frac{F_0}{\epsilon(1+\frac{1}{2}\epsilon)\Omega^2}\sin\tfrac{1}{2}\epsilon\Omega t\,\sin\Omega(1+\tfrac{1}{2}\epsilon)t.$$

Sketch the graph of this solution for the case in which ϵ is small.

5.11 Figure 5.12 shows a simple model of a car moving with constant speed c along a gently undulating road with profile $h(x)$, where $h'(x)$ is small. The car is represented by a chassis

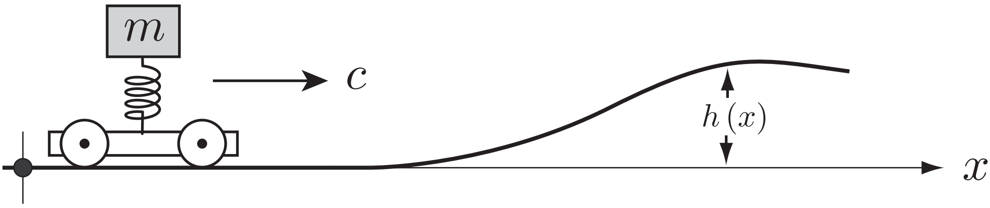


FIGURE 5.12 The car moves along a gently undulating road.

which keeps contact with the road, connected to an upper mass m by a spring and a damper. At time t the upper mass has displacement $y(t)$ above its equilibrium level. Show that, under suitable assumptions, y satisfies a differential equation of the form

$$\ddot{y} + 2K\dot{y} + \Omega^2 y = 2Kch'(ct) + \Omega^2 h(ct)$$

where K and Ω are positive constants.

Suppose that the profile of the road surface is given by $h(x) = h_0 \cos(px/c)$, where h_0 and p are positive constants. Find the amplitude a of the *driven* oscillations of the upper mass.

The vehicle designer adjusts the damper so that $K = \Omega$. Show that

$$a \le \frac{2}{\sqrt{3}} h_0,$$

whatever the values of the consants Ω and p.

5.12 ***Solution by Fourier series*** A driven oscillator satisfies the equation

$$\ddot{x} + 2K\dot{x} + \Omega^2 x = F(t),$$

where K and Ω are positive constants. Find the driven response of the oscillator to the saw tooth' input, that is, when $F(t)$ is given by

$$F(t) = F_0 t \qquad\qquad (-\pi < t < \pi)$$

and $F(t)$ is periodic with period 2π. [It is a good idea to sketch the graph of the function $F(t)$.]

Non-linear oscillations that are piecewise linear

5.13 A particle of mass m is connected to a fixed point O on a smooth horizontal table by a linear elastic string of natural length $2a$ and strength $m\Omega^2$. Initially the particle is released from rest at a point on the table whose distance from O is $3a$. Find the period of the resulting oscillations.

5.14 ***Coulomb friction*** The displacement x of a spring mounted mass under the action of Coulomb friction satisfies the equation

$$\ddot{x} + \Omega^2 x = \begin{cases} -F_0 & \dot{x} > 0 \\ F_0 & \dot{x} < 0 \end{cases}$$

where Ω and F_0 are positive constants. If $|x| > F_0/\Omega^2$ when $\dot{x} = 0$, then the motion continues; if $|x| \leq F_0/\Omega^2$ when $\dot{x} = 0$, then the motion ceases. Initially the body is released from rest with $x = 9F_0/2\Omega^2$. Find where it finally comes to rest. How long was the body in motion?

5.15 A partially damped oscillator satisfies the equation

$$\ddot{x} + 2\kappa\,\dot{x} + \Omega^2 x = 0,$$

where Ω is a positive constant and κ is given by

$$\kappa = \begin{cases} 0 & x < 0 \\ K & x > 0 \end{cases}$$

where K is a positive constant such that $K < \Omega$. Find the period of the oscillator and the ratio of successive maximum values of x.

Normal modes

5.16 A particle P of mass $3m$ is suspended fron a fixed point O by a light linear spring with strength α. A second particle Q of mass $2m$ is in turn suspended from P by a second spring of the same strength. The system moves in the vertical straight line through O. Find the normal frequencies and the form of the normal modes for this system. Write down the form of the general motion.

5.17 Two particles P and Q, each of mass m, are secured at the points of trisection of a light string that is stretched to tension T_0 between two fixed supports a distance $3a$ apart. The particles undergo small *transverse* oscillations perpendicular to the equlilibrium line of the string. Find the normal frequencies, the forms of the normal modes, and the general motion of this system. [Note that the forms of the modes could have been deduced from the symmetry of the system.] Is the general motion periodic?

5.18 A particle P of mass $3m$ is suspended from a fixed point O by a light inextensible string of length a. A second particle Q of mass m is in turn suspended from P by a second string of length a. The system moves in a vertical plane through O. Show that the linearised equations of motion for *small* oscillations near the downward vertical are

$$4\ddot{\theta} + \ddot{\phi} + 4n^2\theta = 0,$$

$$\ddot{\theta} + \ddot{\phi} + n^2\phi = 0,$$

where θ and ϕ are the angles that the two strings make with the downward vertical, and $n^2 = g/a$. Find the normal frequencies and the forms of the normal modes for this system.

Chapter Six

Energy conservation

KEY FEATURES

The key features of this chapter are the **energy principle** for a particle, **conservative fields** of force, **potential energies** and **energy conservation**.

In this Chapter, we introduce the notion of **mechanical energy** and its **conservation**. Although energy methods are never indispensible* for the solution of problems, they do give a greater insight and allow many problems to be solved in a quick and elegant manner. Energy has a fundamental rôle in the Lagrangian and Hamiltonian formulations of mechanics. More generally, the notion of energy has been so widely extended that energy conservation has become the most pervasive and important principle in the whole of physics.

6.1 THE ENERGY PRINCIPLE

Suppose a particle P of mass m moves under the influence of a force $\boldsymbol{F}$. Then its equation of motion is

$$m\frac{d\boldsymbol{v}}{dt} = \boldsymbol{F}, \tag{6.1}$$

where $\boldsymbol{v}$ is the velocity of P at time t. At this stage we place no restrictions on the force $\boldsymbol{F}$. It may depend on the position of P, the velocity of P, the time, or anything else; if more than one force is acting on P, then $\boldsymbol{F}$ means the *vector resultant* of these forces. On taking the scalar product of both sides of equation (6.1) with $\boldsymbol{v}$, we obtain the scalar equation

$$m\boldsymbol{v}\cdot\frac{d\boldsymbol{v}}{dt} = \boldsymbol{F}\cdot\boldsymbol{v}$$

and, since

$$m\boldsymbol{v}\cdot\frac{d\boldsymbol{v}}{dt} = \frac{d}{dt}\left(\tfrac{1}{2}m\boldsymbol{v}\cdot\boldsymbol{v}\right),$$

* Energy is never mentioned in the work of Newton!

this can be written in the form

$$\frac{dT}{dt} = \boldsymbol{F} \cdot \boldsymbol{v}, \tag{6.2}$$

where $T = \frac{1}{2} m\boldsymbol{v} \cdot \boldsymbol{v}$.

Definition 6.1 ***Kinetic energy*** *The scalar quantity* $T = \frac{1}{2} m\boldsymbol{v} \cdot \boldsymbol{v} = \frac{1}{2} m|\boldsymbol{v}|^2$ *is called the* ***kinetic energy*** *of the particle* P.

If we now integrate equation (6.2) over the time interval $[t_1, t_2]$, we obtain

$$\boxed{T_2 - T_1 = \int_{t_1}^{t_2} \boldsymbol{F} \cdot \boldsymbol{v}\, dt} \tag{6.3}$$

where T_1 and T_2 are the kinetic energies of P at times t_1 and t_2 respectively. This is the **energy principle** for a particle moving under a force $\boldsymbol{F}$.

Definition 6.2 ***1-D work done*** *The scalar quantity*

$$W = \int_{t_1}^{t_2} \boldsymbol{F} \cdot \boldsymbol{v}\, dt \tag{6.4}$$

is called the ***work done*** *by the force* $\boldsymbol{F}$ *during the time interval* $[t_1, t_2]$. *The* ***rate of working*** *of* $\boldsymbol{F}$ *at time* t *is thus* $\boldsymbol{F} \cdot \boldsymbol{v}$.
[The SI unit of work is the joule (J) and one joule per second is one watt (W).]

Our result can now be stated as follows:

Energy principle for a particle

In any motion of a particle, the increase in the kinetic energy of the particle in a given time interval is equal to the total work done by the applied forces during this time interval.

The energy principle is a *scalar* equality which is derived by integrating the *vector* equation of motion (6.1). Thus the energy principle will generally contain less information than the equation of motion, so that we have no right to expect the motion of P to be determined from the energy principle *alone*. The situation is simpler when P has **one degree of freedom**, which means that the position of P can be specified by a single scalar variable. In this case the equation of motion and the energy principle are equivalent and the energy principle alone *is* sufficient to determine the motion.

Example 6.1 *Verify the energy principle*

A man of mass 100 kg can pull on a rope with a maximum force equal to two fifths of his own weight. [Take $g = 10\ \mathrm{m\,s^{-2}}$.] In a competition, he must pull a block of mass 1600 kg across a smooth horizontal floor, the block being initially at rest. He is able to apply his maximum force horizontally for 12 seconds before falling exhausted. Find the total work done by the man and confirm that the energy principle is true in this case.

Solution

In this problem, the block is subjected to three forces: the force exerted by the man, uniform gravity, and the vertical reaction of the smooth floor. However, since the last two of these are equal and opposite, they can be ignored.

The man has weight 1000 N so that the force he applies to the block is a constant 200 N. The Second Law then implies that, while the man is pulling on the rope, the block must have constant rectilinear acceleration $200/1600 = 1/8 \text{ m s}^{-2}$. Since the block is initially at rest, its velocity v at time t is therefore $v = t/8 \text{ m s}^{-1}$. The total work W done by the man is then given by the formula (6.4) to be

$$W = \int_0^{12} \boldsymbol{F} \cdot \boldsymbol{v}\, dt = \int_0^{12} 200 \left(\frac{t}{8}\right) dt = 1800 \text{ J}.$$

When $t = 12$ s, the block has velocity $v = 12/8 = 3/2 \text{ m s}^{-1}$, so that the final kinetic energy of the block is $\frac{1}{2}(1600)(3/2)^2 = 1800$ J. Since the initial kinetic energy of the block is zero, the kinetic energy of the block increases by 1800 J, the same as the work done by the man. This confirms the truth of the energy principle. ■

6.2 ENERGY CONSERVATION IN RECTILINEAR MOTION

The energy principle is not normally used in the general form (6.3). When possible, it is transformed into a conservation principle. This is most easily illustrated by the special case of rectilinear motion.

Suppose that the particle P moves along the x-axis under the force F acting in the positive x-direction. In this case, the 'work done' integral (6.4) reduces to

$$W = \int_{t_1}^{t_2} F v\, dt,$$

where $v = \dot{x}$ is the velocity of P in the positive x-direction. For the case in which F is a **force field** (so that $F = F(x)$), the formula for W becomes

$$W = \int_{t_1}^{t_2} F v\, dt = \int_{t_1}^{t_2} F(x) \frac{dx}{dt}\, dt = \int_{x_1}^{x_2} F(x)\, dx,$$

where $x_1 = x(t_1)$ and $x_2 = x(t_2)$. Thus, when P moves over the interval $[x_1, x_2]$ of the x-axis, the work done by the field F is given by

$$\boxed{W = \int_{x_1}^{x_2} F(x)\, dx} \qquad (6.5)$$

(This is a common definition of the work done by a force F. It can be used when $F = F(x)$, but not in general.) It follows that the energy principle for a particle moving in a rectilinear force field can be written

$$T_2 - T_1 = \int_{x_1}^{x_2} F(x)\,dx.$$

Now let $V(x)$ be the indefinite integral of $-F(x)$, so that

$$F = -\frac{dV}{dx} \qquad \text{and} \qquad \int_{x_1}^{x_2} F(x)\,dx = V(x_1) - V(x_2). \tag{6.6}$$

Such a V is called the **potential energy*** function of the force field F. In terms of V, the energy principle in rectilinear motion can be written

$$T_2 + V(x_2) = T_1 + V(x_1),$$

which is equivalent to the **energy conservation** formula

$$\boxed{T + V = E} \tag{6.7}$$

where E is a constant called the **total energy** of the particle. This result can be stated as follows:

Energy conservation in rectilinear motion

When a particle undergoes rectilinear motion in a force field, the sum of its kinetic and potential energies remains constant in the motion.

Example 6.2 *Finding potential energies*

Find the potential energies of (i) the (one-dimensional) SHM force field, (ii) the (one-dimensional) attractive inverse square force field.

Solution

(i) The one-dimensional SHM force field is $F = -\alpha x$, where α is a positive constant. The corresponding V is given by

$$V = -\int_a^x F(x)\,dx = \alpha \int_a^x x\,dx,$$

where a, the lower limit of integration, can be arbitrarily chosen. (This corresponds to the arbitrary choice of the constant of integration.) Note that, by beginning the

* The potential energy corresponding to a given F is uniquely determined apart from a constant of integration; this constant has no physical significance.

integration at $x = a$, we make $V(a) = 0$. In the present case it is conventional to take $a = 0$ so that $V = 0$ at $x = 0$. With this choice, the potential energy is $V = \frac{1}{2}\alpha x^2$.

(ii) The one-dimensional attractive inverse square force field is $F = -K/x^2$, where $x > 0$ and K is a positive constant. The corresponding V is given by

$$V = -\int_a^x F(x)\,dx = K\int_a^x \frac{1}{x^2}\,dx.$$

This time it is not possible to take $a = 0$ (the integral would then be meaningless) and it is conventional to take $a = +\infty$; this makes $V = 0$ when $x = +\infty$. With this choice, the potential energy is $V = -K/x$ $(x > 0)$. ■

Example 6.3 *Rectilinear motion under uniform gravity*

A particle P is projected vertically upwards with speed u and moves under uniform gravity. Find the maximum height achieved and the speed of P when it returns to its starting point.

Solution

Suppose that P is projected from the origin and moves along the z-axis, where Oz points vertically upwards. The force F exerted by uniform gravity is $F = -mg$ and the corresponding potential energy V is given by

$$V = -\int_0^z (-mg)\,dz = mgz.$$

Energy conservation then implies that

$$\tfrac{1}{2}mv^2 + mgz = E,$$

where $v = \dot{z}$, and the constant E is determined from the initial condition $v = u$ when $z = 0$. This gives $E = \frac{1}{2}mu^2$ so that the energy conservation equation for the motion is

$$\tfrac{1}{2}mv^2 + mgz = \tfrac{1}{2}mu^2.$$

Since $v = 0$ when $z = z_{\max}$, it follows that $z_{\max} = u^2/(2g)$. This result was obtained from the Second Law in Chapter 4. When P returns to O, $z = 0$ and so $|v| = u$. Thus P returns to O with speed u, the projection speed. ■

Example 6.4 *Simple harmonic motion*

A particle of mass m is projected from the point $x = a$ with speed u and moves along the x-axis under the SHM force field $F = -m\omega^2 x$. Find the maximum distance from O and the maximum speed achieved by the particle in the subsequent motion.

Solution

The potential energy correspondingto the force field $F = -m\omega^2 x$ is $V = \frac{1}{2}\omega^2 x^2$.

Energy conservation then implies that

$$\tfrac{1}{2}mv^2 + \tfrac{1}{2}m\omega^2 x^2 = E,$$

where $v = \dot{x}$, and the constant E is determined from the initial condition $|v| = u$ when $x = a$. This gives $E = \frac{1}{2}m(u^2+\omega^2 a^2)$ so that the energy conservation equation for the motion becomes

$$v^2 + \omega^2 x^2 = u^2 + \omega^2 a^2.$$

Since $v = 0$ when $|x|$ takes its maximum value, it follows that

$$|x|_{\max} = \left(\frac{u^2}{\omega^2} + a^2\right)^{1/2}.$$

Also, since the left side of the energy conservation equation is the sum of two positive terms, it follows that $|v|$ takes its maximum value when $x = 0$. Hence

$$|v|_{\max} = \left(u^2 + \omega^2 a^2\right)^{1/2}.$$

These results could also be obtained (less quickly) by using the methods described in Chapter 5. ■

6.3 GENERAL FEATURES OF RECTILINEAR MOTION

The energy conservation equation

$$\tfrac{1}{2}mv^2 + V(x) = E \tag{6.8}$$

enables us to deduce the general features of rectilinear motion in a force field. Since $T \ge 0$ (and is equal to zero only when $v = 0$) it follows that the position of the particle is restricted to those values of x that satisfy

$$V(x) \le E,$$

and that equality will occur only when $v = 0$. Suppose that $V(x)$ has the form shown in Figure 6.1 and that E has the value shown. Then the motion of P must take place either (i) in the bounded interval $a \le x \le b$, or (ii) in the unbounded interval $c \le x \le \infty$. Thus, if the particle was situated in the interval $[a, b]$ initially, this is the interval in which the motion will take place.

Bounded motions

Suppose that the motion is started with P in the interval $[a, b]$ and with v positive, so that P is moving to the right. Then, since v can only be zero at $x = a$ and $x = b$, v will remain positive until P reaches the point $x = b$, where it comes to rest*. From equation (6.8), it follows that

* Strictly speaking, we should exclude the possibility that P might approach the point $x = b$ asymptotically as $t \to \infty$, and never actually get there. This can happen, but only in the case in which the line $V = E$ is a *tangent* to the graph of $V(x)$ at $x = b$. In the general case depicted in Figure 6.1, P does arrive at $x = b$ in a finite time.

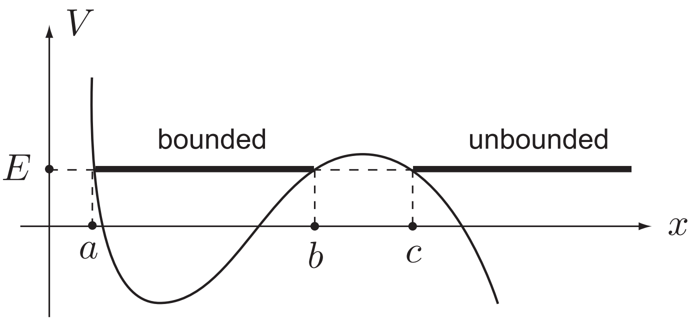


FIGURE 6.1 Bounded and unbounded motions in a rectilinear force field.

the ODE that governs this 'right' part of the motion is

$$\frac{dx}{dt} = +[2(E - V(x))]^{1/2}.$$

At the point $x = b$, $V' > 0$ which implies that $F < 0$. P therefore moves to the left and does not stop until it reaches the point $x = a$. The ODE that governs this 'left' part of the motion is

$$\frac{dx}{dt} = -[2(E - V(x))]^{1/2}.$$

At the point $x = a$, $V' < 0$, which implies that $F > 0$ and that P moves to the right once again. The result is that P performs **periodic oscillations** between the extreme points $x = a$ and $x = b$. Since the 'left' and 'right' parts of the motion take equal times, the period τ of these oscillations can be found by integrating either equation over the interval $a \le x \le b$. Each equation is a separable ODE and integration gives

$$\tau = 2\int_a^b \frac{dx}{[2(E - V(x))]^{1/2}}.$$

It should be noted that these oscillations are generally *not* simple harmonic. In particular, their *period is amplitude dependent*.

Example 6.5 *Periodic oscillations*

A particle P of mass 2 moves on the positive x-axis under the force field $F = (4/x^2) - 1$. Initially P is released from rest at the point $x = 4$. Find the extreme points and the period of the motion.

Solution

The force field F has potential energy $V = (4/x) + x$, so that the energy conservation equation for P is

$$\tfrac{1}{2}(2)v^2 + (4/x) + x = E,$$

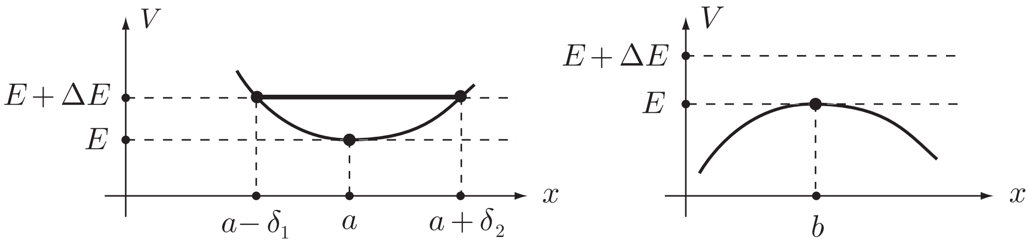


FIGURE 6.2 Positions of stable and unstable equilibrium.

where $v = \dot{x}$ and E is the total energy. The initial condition $v = 0$ when $x = 4$ gives $E = 5$ so that

$$v^2 = 5 - (4/x) - x.$$

The extreme points of the motion occur when $v = 0$, that is, when $x = 1$ and $x = 4$. To find the period τ of the oscillations, write $v = dx/dt$ in the last equation and take square roots. This gives the separable ODEs

$$\frac{dx}{dt} = \pm\left[\frac{(x-1)(4-x)}{x}\right]^{1/2},$$

where the plus and minus signs refer to the motion of P in the positive and negative x-directions respectively. Integration of either equation gives

$$\tau = 2\int_1^4 \left[\frac{x}{(x-1)(4-x)}\right]^{1/2} dx \approx 9.69. \blacksquare$$

Unbounded motions

Suppose now that the motion is started with P in the interval $[c, \infty)$ and with v negative, so that P is moving to the left. Then, since v can only be zero at $x = c$, v will remain negative until P reaches the point $x = c$, where it comes to rest. At the point $x = c$, $V' < 0$ which implies that $F > 0$. P therefore moves to the right and continues to do so indefinitely.

Stable equilibrium and small oscillations

First, we define what we mean by an equilibrium position.

Definition 6.3 *Equilibrium* *The point A is said to be an **equilibrium position** of P if, when P is released from rest at A, P remains at A.*

In the case of rectilinear motion under a force field $F(x)$, the point $x = a$ will be an equilibrium position of P if (and only if) $F(a) = 0$, that is, if $V'(a) = 0$. It follows that *the equilibrium positions of P are the stationary points of the potential energy function $V(x)$*. Consider the equilibrium positions shown in Figure 6.2. These occur at stationary points of V that are a minimum and a maximum respectively. Suppose that P is at rest at the minimum point $x = a$ when it receives an impulse of magnitude J which gives it kinetic energy ΔE $(= J^2/2m)$. The total energy of P is now $E + \Delta E$, and so P will oscillate in the interval $[a - \delta_1, a + \delta_2]$ shown. It is clear from Figure 6.2 that, as the magnitude of J (and therefore

ΔE) tends to zero, the 'amplitude' δ of the resulting motion (the larger of δ_1 and δ_2) also tends to zero. This is the definition of **stable equilibrium**.

Definition 6.4 ***Stable equilibrium*** *Suppose that a particle P is in equilibrium at the point A when it receives an impulse of magnitude J; let δ be the amplitude of the subsequent motion. If $\delta \to 0$ as $J \to 0$, then the point A is said to be a position of* ***stable equilibrium*** *of P.*

On the other hand, if P is at rest at the *maximum* point $x = b$ when it receives an impulse of magnitude J, it is clear that the amplitude of the resulting motion does not tend to zero as J tends to zero, so that a maximum point of $V(x)$ is *not* a position of stable equilibrium of P. The same applies to stationary inflection points.

> **Equilibrium positions of a particle**
>
> The stationary points of the potential energy $V(x)$ are the equilibrium points of P and the minimum points of $V(x)$ are the positions of stable equilibrium. If A is a position of stable equilibrium, then P can execute small-amplitude oscillations about A.

Approximate equation of motion for small oscillations

Suppose that the point $x = a$ is a minimum point of the potential energy $V(x)$. Then, when x is sufficiently close to a, we may approximate $V(x)$ by the first three terms of its Taylor series in powers of the variable $(x - a)$, as follows:

$$\begin{aligned} V(x) &= V(a) + (x-a)V'(a) + \tfrac{1}{2}(x-a)^2 V''(a) \\ &= V(a) + \tfrac{1}{2}(x-a)^2 V''(a), \end{aligned} \tag{6.9}$$

since $V'(a) = 0$. Thus, for small amplitude oscillations about $x = a$, the energy conservation equation is approximately

$$\tfrac{1}{2}mv^2 + V(a) + \tfrac{1}{2}(x-a)^2 V''(a) = E.$$

If we now differentiate this equation with respect to t (and divide by v), we obtain the approximate (linearised) equation of motion

$$m\frac{d^2x}{dt^2} + V''(a)\,(x-a) = 0.$$

Provided that $V''(a) > 0$, this is the equation for **simple harmonic oscillations** with angular frequency $(V''(a)/m)^{1/2}$ about the point $x = a$. The small oscillations of P about $x = a$ are therefore **approximately simple harmonic** with **approximate period** $\tau = 2\pi(m/V''(a))^{1/2}$.

Example 6.6 ***Finding the period of small oscillations***

A particle P of mass 8 moves on the x-axis under the force field whose potential energy is

$$V = \frac{x(x-3)^2}{3}.$$

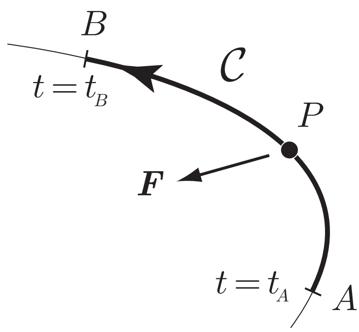


FIGURE 6.3 Particle P is in general motion under the force $\boldsymbol{F}$. The arc $\mathcal{C}$ is the path taken by P between the points A and B.

Show that there is a single position of stable equilibrium and find the approximate period of small oscillations about this point.

Solution

For this V, $V' = x^2 - 4x + 3$ and $V'' = 2x - 4$. The equilibrium positions occur when $V' = 0$, that is when $x = 1$ and $x = 3$. Since $V''(1) = -2$ and $V''(3) = 2$, we deduce that the only position of *stable* equilibrium is at $x = 3$. The approximate period τ of small oscillations about this point is therefore given by $\tau = 2\pi(8/V''(3))^{1/2} = 4\pi$. ■

6.4 ENERGY CONSERVATION IN A CONSERVATIVE FIELD

Suppose now that the particle P is in **general three-dimensional motion** under the force $\boldsymbol{F}$ and that, in the time interval $[t_A, t_B]$, P moves from the point A to the point B along the path $\mathcal{C}$, as shown in Figure 6.3. Then, by the energy principle (6.3),

$$T_B - T_A = \int_{t_A}^{t_B} \boldsymbol{F} \cdot \boldsymbol{v}\, dt, \tag{6.10}$$

where T_A and T_B are the kinetic energies of P when $t = t_A$ and $t = t_B$ respectively. When $\boldsymbol{F}$ is a **force field** $\boldsymbol{F}(\boldsymbol{r})$, the 'work done' integral on the right side of equation (6.10) can be written in the form

$$\int_{t_A}^{t_A} \boldsymbol{F} \cdot \boldsymbol{v}\, dt = \int_{t_A}^{t_B} \boldsymbol{F}(\boldsymbol{r}) \cdot \frac{d\boldsymbol{r}}{dt}\, dt = \int_{\mathcal{C}} \boldsymbol{F}(\boldsymbol{r}) \cdot d\boldsymbol{r},$$

where $\mathcal{C}$ is the path taken by P in the time interval $[t_A, t_B]$. It follows that the energy principle for a particle moving in a 3D force field can be written

$$T_B - T_A = \int_{\mathcal{C}} \boldsymbol{F}(\boldsymbol{r}) \cdot d\boldsymbol{r}. \tag{6.11}$$

Integrals like that on the right side of equation (6.11) are called **line integrals**. They differ from ordinary integrals in that the range of integration is not an interval of the x-axis, but a *path* in three-dimensional space. Line integrals are treated in detail in texts on vector field theory (see for example Schey [11]), but their *physical* meaning in the present context is clear enough. The quantity $\boldsymbol{F} \cdot d\boldsymbol{r}$ is the infinitesimal work done by $\boldsymbol{F}$ when P traverses the element d $\boldsymbol{r}$ of the path $\mathcal{C}$. The line integral sums these contributions to give the *total* work done by $\boldsymbol{F}$.

The line intgral of $\boldsymbol{F}$ along $\mathcal{C}$ is taken to be the *definition* of the work done by the force field $\boldsymbol{F}(\boldsymbol{r})$ when its point of application moves along *any* path $\mathcal{C}$ that connects A and B.

Definition 6.5 ***3-D work done*** *The expression*

$$W[\,A \to B\,;\,\mathcal{C}\,] = \int_{\mathcal{C}} \boldsymbol{F}(\boldsymbol{r}) \cdot d\boldsymbol{r} \tag{6.12}$$

is called the ***work done*** *by the force field $\boldsymbol{F}(\boldsymbol{r})$ when its point of application moves from A to B along any path $\mathcal{C}$.*

The above definition is more than just an alternative definition of the work done by a force field acting on a particle. It defines the quantity $W[\,A \to B\,;\,\mathcal{C}\,]$ *whether or not $\mathcal{C}$ is an actual path traversed by the particle P*. In this wider sense, the concept of 'work done' is purely notional, as is the concept of the 'point of application of the force'. In this sense, W exists for *all* paths joining the points A and B, but W should be regarded as the real work done by $\boldsymbol{F}$ only when $\mathcal{C}$ is an actual path of a particle moving under the field $\boldsymbol{F}(\boldsymbol{r})$.

Energy conservation

In order to develop an energy conservation principle for the general three-dimensional case, we need the right side of equation (6.11) to be expressible in the form

$$\int_{\mathcal{C}} \boldsymbol{F}(\boldsymbol{r}) \cdot d\boldsymbol{r} = V_A - V_B, \tag{6.13}$$

for *some* scalar function of position $V(\boldsymbol{r})$, where V_A and V_B are the values of V at the points A and B. In the rectilinear case, there was no difficulty in finding such a V (it was the indefinite integral of $-F(x)$). In the general case however, it is far from clear that any such V should exist. For, if there does exist a function $V(\boldsymbol{r})$ satisfying equation (6.13), this must mean that the line integral $W[\,A \to B\,;\,\mathcal{C}\,]$ has the same value for *all* paths $\mathcal{C}$ that connect the points A and B. There is no reason why this should be true and, in general, it is *not* true. There is however an important class of fields $\boldsymbol{F}(\boldsymbol{r})$ for which $V(\boldsymbol{r})$ does exist, and it is these fields that we shall consider from now on.

Definition 6.6 ***Conservative field*** *If the field $\boldsymbol{F}(\boldsymbol{r})$ can be expressed in the form**

$$\boldsymbol{F} = -\,\mathrm{grad}\, V, \tag{6.14}$$

where $V(\boldsymbol{r})$ is a scalar function of position, then $\boldsymbol{F}$ is said to be a ***conservative*** *field and the function V is said to be the* ***potential energy*** *function† for $\boldsymbol{F}$.*

* If $\psi(\boldsymbol{r})$ is a *scalar* field then grad ψ is the *vector* field defined by

$$\mathrm{grad}\,\psi = \frac{\partial \psi}{\partial x}\,\boldsymbol{i} + \frac{\partial \psi}{\partial y}\,\boldsymbol{j} + \frac{\partial \psi}{\partial z}\,\boldsymbol{k}.$$

Thus if $\psi = xy^3z^5$, then grad $\psi = y^3z^5\boldsymbol{i} + 3xy^2z^5\boldsymbol{j} + 5xy^3z^4\boldsymbol{k}$. We could omit the minus sign in the definition (6.14), but the potential energy of $\boldsymbol{F}$ would then be $-V$ instead of V.

† If V exists, then it is unique apart from a constant of integration. As in the rectilinear case, this constant has no physical significance.

Example 6.7 ***Conservative or not conservative?***

Show that the field $\boldsymbol{F}_1 = -2x\boldsymbol{i} - 2y\boldsymbol{j} - 2z\boldsymbol{k}$ is conservative but that the field $\boldsymbol{F}_2 = y\boldsymbol{i} - x\boldsymbol{j}$ is not.

Solution

(i) If $\boldsymbol{F}_1$ is conservative, then its potential energy V must satisfy

$$-\frac{\partial V}{\partial x} = -2x, \qquad -\frac{\partial V}{\partial y} = -2y, \qquad -\frac{\partial V}{\partial z} = -2z,$$

and these equations integrate to give

$$V = x^2 + p(y,z), \qquad V = y^2 + q(x,z), \qquad V = z^2 + r(x,y),$$

where p, q and r are 'constants' of integration, which, in this case, are functions of the other variables. If V really exists, then these three representations of V can be made identical by making special choices of the functions p, q and r. In this example it is clear that this can be achieved by taking $p = y^2 + z^2$, $q = x^2 + z^2$, and $r = x^2 + y^2$. Hence $\boldsymbol{F}_1 = -\operatorname{grad}(x^2 + y^2 + z^2)$ and so $\boldsymbol{F}_1$ is conservative.

(ii) If $\boldsymbol{F}_2$ is conservative, then its potential energy V must satisfy

$$-\frac{\partial V}{\partial x} = y, \qquad -\frac{\partial V}{\partial y} = -x, \qquad -\frac{\partial V}{\partial z} = 0.$$

There is no V that satisfies these equations simultaneously. The easiest way to show this is to observe that, from the first equation, $\partial^2 V/\partial y \partial x = -1$ while, from the second equation, $\partial^2 V/\partial x \partial y = +1$. Since these mixed partial derivatives of V should be equal, we have a contradiction. The conclusion is that no such V exists and that $\boldsymbol{F}_2$ is not conservative. ■

Suppose now that the field $\boldsymbol{F}(\boldsymbol{r})$ is conservative with potential energy $V(\boldsymbol{r})$ and let $\mathcal{C}$ be *any* path connecting the points A and B. Then

$$\begin{aligned} W[A \to B\,;\,\mathcal{C}] &= \int_{\mathcal{C}} \boldsymbol{F}(\boldsymbol{r}) \cdot d\boldsymbol{r} = \int_{\mathcal{C}} (-\operatorname{grad} V) \cdot d\boldsymbol{r} \\ &= -\int_{\mathcal{C}} \left(\frac{\partial V}{\partial x}\boldsymbol{i} + \frac{\partial V}{\partial y}\boldsymbol{j} + \frac{\partial V}{\partial z}\boldsymbol{k} \right) \cdot (dx\,\boldsymbol{i} + dy\,\boldsymbol{j} + dz\,\boldsymbol{k}) \\ &= -\int_{\mathcal{C}} \frac{\partial V}{\partial x}dx + \frac{\partial V}{\partial y}dy + \frac{\partial V}{\partial z}dz \\ &= -\int_{\mathcal{C}} dV = V_A - V_B. \end{aligned} \tag{6.15}$$

Thus, when $\boldsymbol{F}$ is conservative with potential energy V,

$$\int_{\mathcal{C}} \boldsymbol{F}(\boldsymbol{r}) \cdot d\boldsymbol{r} = V_A - V_B,$$

for any path $\mathcal{C}$ connecting the points A and B. The energy principle (6.11) can therefore be written

$$T_B + V_B = T_A + V_A$$

which is equivalent to the **energy conservation** formula

$$\boxed{T + V = E} \qquad (6.16)$$

Our result can be summarised as follows:

Energy conservation in 3-D motion

When a particle moves in a **conservative** force field, the sum of its kinetic and potential energies remains constant in the motion.

The condition that $\boldsymbol{F}$ be conservative seems restrictive, but most force fields encountered in mechanics actually are conservative!

Example 6.8 *Finding 3-D potential energies*

(a) Show that the uniform gravity field $\boldsymbol{F} = -mg\boldsymbol{k}$ is conservative with potential energy $V = mgz$.

(b) Show that any force field of the form

$$\boldsymbol{F} = h(r)\,\widehat{\boldsymbol{r}}$$

(a central field) is conservative with potential energy $V = -H(r)$, where $H(r)$ is the indefinite integral of $h(r)$. Use this result to find the potential energies of (i) the 3-D SHM field $\boldsymbol{F} = -\alpha r\,\widehat{\boldsymbol{r}}$, and (ii) the attractive inverse square field $\boldsymbol{F} = -(K/r^2)\,\widehat{\boldsymbol{r}}$, where α and K are positive constants.

Solution

Since the potential energies are given, it is sufficient to evaluate $-\operatorname{grad} V$ in each case and show that this gives the appropriate $\boldsymbol{F}$. Case (a) is immediate. In case (b),

$$\frac{\partial H(r)}{\partial x} = \frac{dH}{dr}\frac{\partial r}{\partial x} = H'(r)\frac{x}{r} = h(r)\frac{x}{r},$$

since $r = \left(x^2 + y^2 + z^2\right)^{1/2}$ and $H'(r) = h(r)$. Thus

$$-\operatorname{grad}\,[-H(r)] = h(r)\left(\frac{x}{r}\,\boldsymbol{i} + \frac{y}{r}\,\boldsymbol{j} + \frac{z}{r}\,\boldsymbol{k}\right) = h(r)\frac{\boldsymbol{r}}{r} = h(r)\,\widehat{\boldsymbol{r}},$$

as required.

In particular then, the potential energy of the SHM field $\boldsymbol{F} = -\alpha r\,\widehat{\boldsymbol{r}}$ is $V = \frac{1}{2}\alpha r^2$, and the potential energy of the attractive inverse square field $\boldsymbol{F} = -(K/r^2)\,\widehat{\boldsymbol{r}}$ is $V = -K/r$. ■

Example 6.9 *Projectile motion*

A body is projected from the ground with speed u and lands on the flat roof of a building of height h. Find the speed with which the projectile lands. [Assume uniform gravity and no air resistance.]

Solution

Since uniform gravity is a conservative field with potential energy mgz, energy conservation applies in the form

$$\tfrac{1}{2}m|\boldsymbol{v}|^2 + mgz = E,$$

where O is the initial position of the projectile and Oz points vertically upwards. From the initial conditions, $E = \frac{1}{2}mu^2$. Hence, when the body lands,

$$\tfrac{1}{2}m|\boldsymbol{v}^L|^2 + mgh = \tfrac{1}{2}mu^2,$$

where $\boldsymbol{v}^L$ is the landing velocity. The **landing speed** is therefore

$$|\boldsymbol{v}^L| = \left(u^2 - 2gh\right)^{1/2}.$$

Thus, energy conservation determines the *speed* of the body on landing, but not its *velocity*. ■

Example 6.10 *Escape from the Moon*

A body is projected from the surface of the Moon with speed u *in any direction*. Show that the body cannot escape from the Moon if $u^2 < 2MG/R$, where M and R are the mass and radius of the Moon. [Assume that the Moon is spherically symmetric.]

Solution

If the Moon is spherically symmetric, then the force $\boldsymbol{F}$ that it exerts on the body is given by $\boldsymbol{F} = -(mMG/r^2)\widehat{\boldsymbol{r}}$, where m is the mass of the body, and $\boldsymbol{r}$ is the position vector of the body relative to an origin at the centre of the Moon. This force is a conservative field with potential energy $V = -mMG/r$. Hence energy conservation applies in the form

$$\tfrac{1}{2}m|\boldsymbol{v}|^2 - \frac{mMG}{r} = E,$$

and, from the initial conditions, $E = \frac{1}{2}mu^2 - (mMG)/R$. Thus the energy conservation equation is

$$|\boldsymbol{v}|^2 = u^2 + 2MG\left(\frac{1}{r} - \frac{1}{R}\right).$$

Since the left side of the above equation is *positive*, the values of r that occur in the motion must satisfy the inequality

$$u^2 + 2MG\left(\frac{1}{r} - \frac{1}{R}\right) \geq 0.$$

If the body is to escape, this inequality must hold for *arbitrarily large* r. This means that the condition

$$u^2 - \frac{2MG}{R} \geq 0$$

is necessary for escape. Hence if $u^2 < 2MG/R$, the body cannot escape. The interesting feature here is that the 'escape speed' is the same for *all directions* of projection from the surface of the Moon. (The special case in which the body is projected vertically upwards was solved in Chapter 4.) ■

Example 6.11 ***Stability of equilibrium in a 3-D conservative field***

A particle P of mass m can move under the gravitational attraction of two particles, of equal mass M, fixed at the points $(0, 0, \pm a)$. Show that the origin O is a position of equilibrium, but that it is not stable. [This illustrates the general result that *no* free-space static gravitational field can provide a position of stable equilibrium.]

Solution

When P is at O, the fixed particles exert equal and opposite forces so that the total force on P is zero. The origin is therefore an equilibrium position for P.

Just as in rectilinear motion, O will be a position of *stable* equilibrium if the potential energy function $V(x, y, z)$ has a minimum at O. This means that the value of V at O must be less than its values at *all* nearby points. But at points on the z-axis between $z = -a$ and $z = a$

$$V(0, 0, z) = -\frac{mMG}{a - z} - \frac{mMG}{a + z} = -\frac{2amMG}{a^2 - z^2},$$

which has a *maximum* at $z = 0$. Hence the equilibrium at O is unstable to disturbances in the z-direction. ■

6.5 ENERGY CONSERVATION IN CONSTRAINED MOTION

Some of the most useful applications of energy conservation occur when the moving particle is subject to geometrical constraints, such as being connected to a fixed point by a light inextensible string, or being required to remain in contact with a fixed rigid surface (see section 4.2). Since constraint forces are not known beforehand one may wonder how to find the work that they do. The answer is that, in the idealised problems that we study, the *work done by the constraint forces is often zero*. In these cases the constraint forces make no contribution to the energy principle and they can be disregarded. Situations in which constraint forces do no work include:

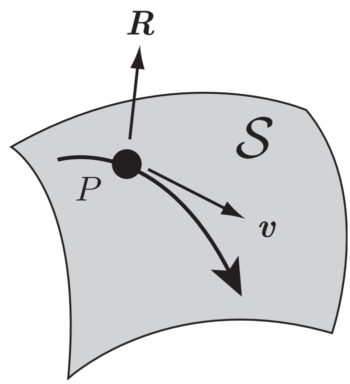


FIGURE 6.4 The particle P slides over the fixed smooth surface $\mathcal{S}$. The reaction $\boldsymbol{R}$ is normal to $\mathcal{S}$ and hence perpendicular to $\boldsymbol{v}$, the velocity of P.

Some constraint forces that do no work

- A particle connected to a fixed point by a light inextensible string; here the *string tension does no work*.
- A particle sliding along a smooth fixed wire; here the *reaction of the wire does no work*.
- A particle sliding over a smooth fixed surface; here the *reaction of the surface does no work*.

Consider for example the case of a particle P sliding over a smooth fixed surface $\mathcal{S}$ as shown in Figure 6.4. Because $\mathcal{S}$ is smooth, any reaction force $\boldsymbol{R}$ that it exerts must always be normal to $\mathcal{S}$. But, since P remains on $\mathcal{S}$, its velocity $\boldsymbol{v}$ must always be tangential to $\mathcal{S}$. Hence $\boldsymbol{R}$ is always perpendicular to $\boldsymbol{v}$ so that $\boldsymbol{R} \cdot \boldsymbol{v} = 0$. Thus the *rate of working* of $\boldsymbol{R}$ is zero and so $\boldsymbol{R}$ makes no contribution to the energy principle. Very similar arguments apply to the other two cases.

We may now extend the use of conservation of energy as follows:

Energy conservation in constrained motion

When a particle moves in a **conservative** force field and is subject to **constraint forces that do no work**, the sum of its kinetic and potential energies remains constant in the motion.

Example 6.12 *The snowboarder*

A snowboarder starts from rest and descends a slope, losing 320 m of altitude in the process. What is her speed at the bottom? [Neglect all forms of resistance and take $g = 10 \text{ m s}^{-2}$.]

Solution

The snowboarder moves under uniform gravity and the reaction force of the smooth hillside. Since this reaction force does no work, energy conservation applies in the form

$$\tfrac{1}{2}m|\boldsymbol{v}|^2 + mgz = E,$$

where m and $\boldsymbol{v}$ are the mass and velocity of the snowboarder, and z is the *altitude* of the snowboarder relative to the bottom of the hill. If the snowboarder starts from rest at altitude h, then $E = 0 + mgh$. Hence, at the bottom of the hill where $z = 0$, her speed is

$$|\boldsymbol{v}| = (2gh)^{1/2},$$

just as if she had fallen down a vertical hole! This speed evaluates to 80 m s^{-1}, about 180 mph. [At such speeds, air resistance would have an important influence.] ■

Our next example concerns a particle constrained to move on a vertical circle. This is one of the classical applications of the energy conservation method. There are two distinct cases: (i) where the particle is constrained always to remain on the circle, or (ii) where the particle is constrained to remain on the circle only while the constraint force has a particular sign.

Example 6.13 *Motion in a vertical circle*

A fixed hollow sphere has centre O and a smooth inner surface of radius b. A particle P, which is inside the sphere, is projected horizontally with speed u from the lowest interior point (see Figure 6.5). Show that, in the subsequent motion,

$$v^2 = u^2 - 2gb(1 - \cos\theta),$$

provided that P remains in contact with the sphere.

Solution

While P remains in contact with the sphere, the motion is as shown in Figure 6.5. The forces acting on P are uniform gravity mg and the constraint force N, which is the normal reaction of the smooth sphere. Since N is always perpendicular to v (the circumferential velocity of P), it follows that N does no work. Hence energy conservation applies in the form

$$\tfrac{1}{2}mv^2 - mgb\cos\theta = E,$$

where m is the mass of P, and the zero level of the potential energy is the horizontal plane through O. Since $v = u$ when $\theta = 0$, it follows that $E = \tfrac{1}{2}mu^2 - mgb$ and the energy conservation equation becomes

$$v^2 = u^2 - 2gb(1 - \cos\theta), \tag{6.17}$$

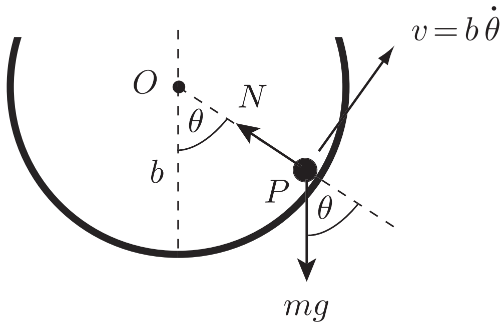


FIGURE 6.5 Particle P slides on the smooth inner surface of a fixed sphere. The motion takes place in a vertical plane through the centre O.

as required. This gives the value of v as a function of θ while P remains in contact with the sphere. ■

Question *The reaction force*

Find the reaction force N as a function of θ.

Answer

Once the motion is determined (for example by equation (6.17)), the unknown constraint forces may be found by using the Second Law in reverse. In the present case, consider the component of the Second Law $\boldsymbol{F} = m\boldsymbol{a}$ in the direction $\overrightarrow{PO}$. This gives

$$N - mg\cos\theta = mv^2/b,$$

where we have made use of the formula (2.17) for the acceleration of a particle in general circular motion. On using the formula for v^2 from equation (6.17), we obtain

$$N = \frac{mu^2}{b} + mg(3\cos\theta - 2). \tag{6.18}$$

This gives the value of N as a function of θ while P remains in contact with the sphere. ■

Question *Does P leave the surface of the sphere?*

For the particular case in which $u = (3gb)^{1/2}$, show that P will leave the surface of the sphere, and find the value of θ at which it does so.

Answer

When $u = (3gb)^{1/2}$, the formulae (6.17), (6.18) for v^2 and N become

$$v^2 = gb(1 + 2\cos\theta), \qquad N = mg(1 + 3\cos\theta).$$

If P remains in contact with the sphere, then it comes to rest when $v = 0$, that is, when $\cos\theta = -1/2$. This first happens when $\theta = 120°$ (a point on the upper half of the sphere, higher than O). If P were threaded on a circular *wire* (from which it could not fall off) this is exactly what would happen; P would perform periodic oscillations in the range $-120° \le \theta \le 120°$. However, in the present case, the reaction N is restricted to be *positive* and this condition will be violated when $\theta > \cos^{-1}(-1/3) \approx 109°$. Since this angle is less than 120°, the conclusion is that P loses contact with the sphere when $\theta = \cos^{-1}(-1/3)$; at this instant, the speed of P is $(gb/3)^{1/2}$. P then moves as a free projectile until it strikes the sphere. ■

Question *Complete circles*

How large must the initial speed be for P to perform complete circles?

Answer

For complete circles to be executed, it is necessary (and sufficient) that $v > 0$ and $N \ge 0$ at all times, that is,

$$u^2 > 2gb(1 - \cos\theta) \qquad \text{and} \qquad u^2 \ge gb(2 - 3\cos\theta)$$

for all values of θ. For these inequalities to hold for *all* θ, the speed u must satisfy $u^2 > 4gb$ and $u^2 \ge 5gb$ respectively. Since the second of these conditions implies the first, it follows that P will execute complete circles if $u^2 \ge 5gb$. ■

Example 6.14 *Small oscillations in constrained motion*

A particle P of mass m can slide freely along a long straight wire. P is connected to a fixed point A, which is at a distance $4a$ from the wire, by a light elastic cord of natural length $3a$ and strength α. Find the approximate period of small oscillations of P about its equilibrium position.

Solution

Suppose P has displacement x from its equilibrium position. In this position, the length of the cord is $(16a^2 + x^2)^{1/2}$ and its potential energy V is

$$\begin{aligned} V &= \tfrac{1}{2}\alpha\left[\left(16a^2 + x^2\right)^{1/2} - 3a\right]^2 \\ &= \tfrac{1}{2}\alpha\left[25a^2 + x^2 - 6a\left(16a^2 + x^2\right)^{1/2}\right]. \end{aligned}$$

The energy conservation equation for P is therefore

$$\tfrac{1}{2}m\dot{x}^2 + \tfrac{1}{2}\alpha\left[25a^2 + x^2 - 6a\left(16a^2 + x^2\right)^{1/2}\right] = E,$$

which, on neglecting powers of x higher than the second, becomes

$$\tfrac{1}{2}m\dot{x}^2 + \tfrac{1}{2}\alpha\left[a^2 + \frac{x^2}{4}\right] = E.$$

On differentiating this equation with respect to t, we obtain the approximate **linearised equation of motion**

$$m\ddot{x} + \frac{\alpha}{4}x = 0.$$

This is the SHM equation with $\omega^2 = \alpha/4m$. It follows that the approximate **period** of small oscillations about $x = 0$ is $4\pi(m/\alpha)^{1/2}$. ■

Energy conservation from a physical viewpoint

Suppose that a particle P of mass m can move on the x-axis and is connected to a fixed post at $x = -a$ by a light elastic spring of natural length a and strength α. Then the force $F(x)$ exerted on P by the spring is given by $F = -\alpha x$, where x is the displacement of P in the positive x-direction. This force field has potential energy $V = \frac{1}{2}\alpha x^2$ and the energy conservation equation for P takes the form

$$\tfrac{1}{2}mv^2 + \tfrac{1}{2}\alpha x^2 = E, \tag{6.19}$$

where $v = \dot{x}$. Here, the spring is regarded merely as an agency that supplies a force field with potential energy $V = \frac{1}{2}\alpha x^2$. However, there is a much more satisfying interpretation of the energy conservation principle (6.19) that can be made.

To see this we consider the spring as described above, but now with no particle attached to its free end. Suppose that the spring is in equilibrium (with its free end at $x = 0$) when an external force $G(t)$ is applied there. This force is initially zero and increases so that, at any time t, the spring has extension $X = G(t)/\alpha$. Suppose that this process continues until the spring has extension Δ. Then the total work done by the force $G(t)$ in producing this extension is given by

$$\int_0^\tau G(t)\dot{X}\,dt = \int_0^\tau \alpha X \frac{dX}{dt} dt = \int_0^\Delta \alpha X\,dX = \tfrac{1}{2}\alpha\Delta^2.$$

Since the force exerted by the fixed post does no work, the *total* work done by the external forces in producing the extension Δ is $\frac{1}{2}\alpha\Delta^2$. Suppose now that the spring is 'frozen' in its extended state (by being propped open, for example) while the particle P is connected to the free end. The system is then released from rest. The energy conservation equation for P is given by equation (6.19), where, from the initial condition $v = 0$ when $x = \Delta$, the total energy $E = \frac{1}{2}\alpha\Delta^2$. This gives

$$\tfrac{1}{2}mv^2 + \tfrac{1}{2}\alpha x^2 = \tfrac{1}{2}\alpha\Delta^2. \tag{6.20}$$

Thus, the total energy in the subsequent motion is equal to the original work done in stretching the spring. The natural physical interpretation of this is that the spring is able to *store* the work that is done upon it as **internal energy**. Then, when the particle is connected and the system released, this stored energy is available to be transferred to the particle in the form of kinetic energy. Equation (6.20) can thus be interpreted as an **energy conservation** principle for the **particle** and **spring** together, as follows: *In any motion of the particle and spring, the sum of the kinetic energy of the particle and the internal energy of the spring remains constant.* In this interpretation, the particle has no potential energy; instead, the spring has internal energy.

In the above example, the particle and the spring can pass energy to each other, but the total of the two energies is conserved. This is the essential nature of energy. It is an entity that can appear in different forms but whose **total is always conserved**. Energy is probably the most important notion in the whole of physics. However, it should be remembered that, in the context of mechanics, it is not usual to take account of forms of energy such as heat or light.

As a result, we will find situations (inelastic collisions, for example) in which energy seems to disappear. There is no contradiction in this; the energy has simply been transferred into forms that we choose not to recognise.

Problems on Chapter 6

Answers and comments are at the end of the book.

Harder problems carry a star ($*$).

Unconstrained motion

6.1 A particle P of mass 4 kg moves under the action of the force $\boldsymbol{F} = 4\boldsymbol{i} + 12t^2\,\boldsymbol{j}$ N, where t is the time in seconds. The initial velocity of the particle is $2\boldsymbol{i} + \boldsymbol{j} + 2\boldsymbol{k}$ m s^{-1}. Find the work done by $\boldsymbol{F}$, and the increase in kinetic energy of P, during the time interval $0 \le t \le 1$. What principle does this illustrate?

6.2 In a competition, a man pushes a block of mass 50 kg with constant speed 2 m s^{-1} up a smooth plane inclined at 30° to the horizontal. Find the rate of working of the man. [Take $g = 10$ m s^{-2}.]

6.3 An athlete putts a shot of mass 7 kg a distance of 20 m. Show that the athlete must do to at least 700 J of work to achieve this. [Ignore the height of the athlete and take $g = 10$ m s^{-2}.]

6.4 Find the work needed to lift a satellite of mass 200 kg to a height of 2000 km above the Earth's surface. [Take the Earth to be spherically symmetric and of radius 6400 km. Take the surface value of g to be 9.8 m s^{-2}.]

6.5 A particle P of unit mass moves on the positive x-axis under the force field

$$F = \frac{36}{x^3} - \frac{9}{x^2} \qquad (x > 0).$$

Show that each motion of P consists of either (i) a periodic oscillation between two extreme points, or (ii) an unbounded motion with one extreme point, depending upon the value of the total energy. Initially P is projected from the point $x = 4$ with speed 0.5. Show that P oscillates between two extreme points and find the period of the motion. [You may make use of the formula

$$\int_a^b \frac{x\,dx}{[(x-a)(b-x)]^{1/2}} = \frac{\pi(a+b)}{2}.\ ]$$

Show that there is a single equilibrium position for P and that it is stable. Find the period of *small* oscillations about this point.

6.6 A particle P of mass m moves on the x-axis under the force field with potential energy $V = V_0(x/b)^4$, where V_0 and b are positive constants. Show that any motion of P consists of a periodic oscillation with centre at the origin. Show further that, when the oscillation has amplitude a, the period τ is given by

$$\tau = 2\sqrt{2}\left(\frac{m}{V_0}\right)^{1/2} \frac{b^2}{a} \int_0^1 \frac{d\xi}{(1-\xi^4)^{1/2}}.$$

[Thus, the larger the amplitude, the shorter the period!]

6.7 A particle P of mass m, which is on the negative x-axis, is moving towards the origin with constant speed u. When P reaches the origin, it experiences the force $F = -Kx^2$, where K is a positive constant. How far does P get along the positive x-axis?

6.8 A particle P of mass m moves on the x-axis under the combined gravitational attraction of two particles, each of mass M, fixed at the points $(0, \pm a, 0)$ respectively (see Figure 3.3). Example 3.4 shows that the force field $F(x)$ acting on P is given by

$$F = -\frac{2mMGx}{(a^2+x^2)^{3/2}}.$$

Find the corresponding potential energy $V(x)$.

Initially P is released from rest at the point $x = 3a/4$. Find the maximum speed achieved by P in the subsequent motion.

6.9 A particle P of mass m moves on the axis Oz under the gravitational attraction of a uniform circular disk of mass M and radius a as shown in Figure 3.6. Example 3.6 shows that the force field $F(z)$ acting on P is given by

$$F = -\frac{2mMG}{a^2}\left[1 - \frac{z}{(a^2+z^2)^{1/2}}\right] \qquad (z > 0).$$

Find the corresponding potential energy $V(z)$ for $z > 0$.

Initially P is released from rest at the point $z = 4a/3$. Find the speed of P when it hits the disk.

6.10 A catapult is made by connecting a light elastic cord of natural length $2a$ and strength α between two fixed supports, which are distance $2a$ apart. A stone of mass m is placed at the center of the cord, which is pulled back a distance $3a/4$ and then released from rest. Find the speed with which the stone is projected by the catapult.

6.11 A light spring of natural length a is placed on a horizontal floor in the upright position. When a block of mass M is resting in equilibrium on top of the spring, the compression of the spring is $a/15$. The block is now lifted to a height $3a/2$ above the floor and released from rest. Find the compression of the spring when the block first comes to rest.

6.12 A particle P carries a charge e and moves under the influence of the static magnetic field $\boldsymbol{B}(\boldsymbol{r})$ which exerts the force $\boldsymbol{F} = e\boldsymbol{v} \times \boldsymbol{B}$ on P, where $\boldsymbol{v}$ is the velocity of P. Show that P travels with constant *speed*.

6.13* A mortar shell is to be fired from level ground so as to clear a flat topped building of height h and width a. The mortar gun can be placed anywhere on the ground and can have any angle of elevation. What is the least projection speed that will allow the shell to clear the building? [*Hint* How is the reqired minimum projection speed changed if the mortar is raised to rooftop level?]

For the special case in which $h = \frac{1}{2}a$, find the optimum position for the mortar and the optimum elevation angle to clear the building.

If you are a star at electrostatics, try the following two problems:

6.14* An *earthed* conducting sphere of radius a is fixed in space, and a particle P, of mass m and charge q, can move freely outside the sphere. Initially P is a distance $b\ (> a)$ from the centre O of the sphere when it is projected directly away from O. What must the projection speed be for P to escape to infinity? [Ignore electro*dynamic* effects. Use the method of images to solve the electro*static* problem.]

6.15* An *uncharged* conducting sphere of radius a is fixed in space and a particle P, of mass m and charge q, can move freely outside the sphere. Initially P is a distance $b\ (> a)$ from the centre O of the sphere when it is projected directly away from O. What must the projection speed be for P to escape to infinity? [Ignore electro*dynamic* effects. Use the method of images to solve the electro*static* problem.]

Constrained motion

6.16 A bead of mass m can slide on a smooth circular wire of radius a, which is fixed in a vertical plane. The bead is connected to the highest point of the wire by a light spring of natural length $3a/2$ and strength λ. Determine the stability of the equilibrium position at the lowest point of the wire in the cases (i) $\alpha = 2mg/a$, and (ii) $\lambda = 5mga$.

6.17 A smooth wire has the form of the helix $x = a\cos\theta$, $y = a\sin\theta$, $z = b\theta$, where θ is a real parameter, and a, b are positive constants. The wire is fixed with the axis Oz pointing vertically upwards. A particle P, which can slide freely on the wire, is released from rest at the point $(a, 0, 2\pi b)$. Find the speed of P when it reaches the point $(a, 0, 0)$ and the time taken for it to do so.

6.18 A smooth wire has the form of the parabola $z = x^2/2b$, $y = 0$, where b is a positive constant. The wire is fixed with the axis Oz pointing vertically upwards. A particle P, which can slide freely on the wire, is performing oscillations with x in the range $-a \le x \le a$. Show that the period τ of these oscillations is given by

$$\tau = \frac{4}{(gb)^{1/2}} \int_0^a \left(\frac{b^2 + x^2}{a^2 - x^2}\right)^{1/2} dx.$$

By making the substitution $x = a\sin\psi$ in the above integral, obtain a new formula for τ. Use this formula to find a two-term approximation to τ, valid when the ratio a/b is small.

6.19* A smooth wire has the form of the cycloid $x = c(\theta + \sin\theta)$, $y = 0$, $z = c(1 - \cos\theta)$, where c is a positive constant and the parameter θ lies in the range $-\pi \le \theta \le \pi$. The wire is fixed with the axis Oz pointing vertically upwards. [Make a sketch of the wire.] A particle

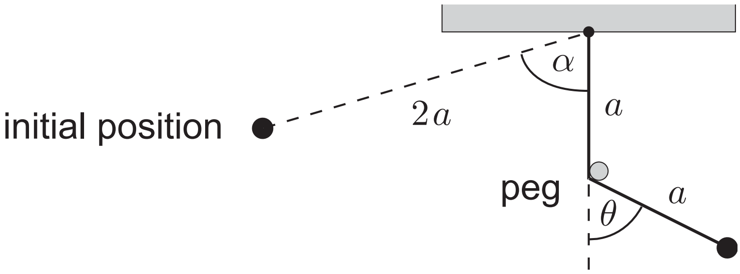


FIGURE 6.6 The swing of the pendulum is obstructed by a fixed peg.

can slide freely on the wire. Show that the energy conservation equation is

$$(1+\cos\theta)\,\dot{\theta}^2 + \frac{g}{c}(1-\cos\theta) = \text{constant}.$$

A new parameter u is defined by $u = \sin\frac{1}{2}\theta$. Show that, in terms of u, the equation of motion for the particle is

$$\ddot{u} + \left(\frac{g}{4c}\right)u.$$

Deduce that the particle performs oscillations with period $4\pi(c/g)^{1/2}$, independent of the amplitude!

6.20 A smooth horizontal table has a vertical post fixed to it which has the form of a circular cylinder of radius a. A light inextensible string is wound around the base of the post (so that it does not slip) and its free end of the string is attached to a particle that can slide on the table. Initially the unwound part of the string is taut and of length $4a/3$. The particle is then projected horizontally at right angles to the string so that the string winds itself *on* to the post. How long does it take for the particle to hit the post? [You may make use of the formula

$$\int (1+\phi^2)^{1/2} d\phi = \tfrac{1}{2}\phi(1+\phi^2)^{1/2} + \tfrac{1}{2}\sinh^{-1}\phi.\,]$$

6.21 A heavy ball is suspended from a fixed point by a light inextensible string of length b. The ball is at rest in the equilibrium position when it is projected horizontally with speed $(7gb/2)^{1/2}$. Find the angle that the string makes with the upward vertical when the ball begins to leave its circular path. Show that, in the subsequent projectile motion, the ball returns to its starting point.

6.22* A new *avant garde* mathematics building has a highly polished outer surface in the shape of a huge hemisphere of radius 40 m. The Head of Department, Prof. Oldfart, has his student, Vita Youngblood, hauled to the summit (to be photographed for publicity purposes) but a small gust of wind causes Vita to begin to slide down. Oldfart's displeasure is increased when Vita lands on (and severely damages) his car which is parked nearby. How far from the outer edge of the building did Oldfart park his car? Did he get what he deserved? (Happily, Vita escaped injury and found a new supervisor.)

6.23 * * A heavy ball is attached to a fixed point O by a light inextensible string of length $2a$. The ball is drawn back until the string makes an acute angle α with the downward vertical and is then released from rest. A thin peg is fixed a distance a vertically below O in the path of the string, as shown in Figure 6.6. In a game of skill, the contestant chooses the value of α and wins a prize if the ball strikes the peg. Show that the winning value of α is approximately 86°.

Chapter Seven

Orbits in a central field

including Rutherford scattering

KEY FEATURES

For motion in *general* central force fields, the key results are the **radial motion equation** and the **path equation**. For motion in the *inverse square* force field, the key formulae are the **E-formula**, the **L-formula** and the **period formula**.

The theory of orbits has a special place in classical mechanics for it was the desire to understand why the planets move as they do which provided the major stimulus in the development of mechanics as a scientific discipline. Early in the seventeenth century, Johannes Kepler * published his 'laws of planetary motion', which he deduced by analysing the accurate experimental observations made by the astronomer Tycho Brahe.†

* The German mathematician and astronomer Johannes Kepler (1571–1630) was a firm believer in the Copernican (heliocentric) model of the solar system. In 1596 he became mathematical assistant to Tycho Brahe, the foremost observational astronomer of the day, and began working on the intractable problem of the orbit of Mars. This work continued after Tycho's death in 1601 and, after much labour, Kepler showed that Tycho's observations of Mars corresponded very precisely to an elliptic orbit with the Sun at a focus. This result, together with the 'law of areas' (the second law) was published in 1609. Kepler then found similar orbits for other planets and his third law was published in 1619.

† Tycho Brahe (1546–1601) was a Danish nobleman. He had a lifelong interest in observational astronomy and developed a succession of new and more accurate instruments. The King of Denmark gave him money to create an observatory and also the island of Hven on which to build it. It was here that Tycho made his accurate observations of the planets from which Kepler was able to deduce his laws of planetary motion. Tycho's other claim to fame is that he had a metal nose. When the original was cut off in a duel, he had an artificial nose made from an alloy of silver and gold. Tycho is perhaps better remembered for his nose job than he is for a lifetime of observations.

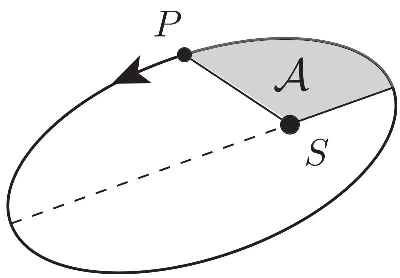


FIGURE 7.1 Each planet P moves on an elliptical path with the Sun S at one focus. The area $\mathcal{A}$ is that referred to in Kepler's second law.

Kepler's laws of planetary motion

First law Each of the planets moves on an elliptical path with the Sun at one focus of the ellipse.

Second law For each of the planets, the straight line connecting the planet to the Sun sweeps out equal areas in equal times.

Third law The squares of the periods of the planets are proportional to the cubes of the major axes of their orbits.

The problem of determining the law of force that causes the motions described by Kepler (and *proving* that it does so) was the most important scientific problem of the seventeenth century. In what must be the finest achievement in the whole history of science, Newton's publication of *Principia* in 1687 not only proved that the inverse square law of gravitation implies Kepler's laws, but also laid down the entire framework of the science of mechanics. Orbit theory is just as important today, the principal fields of application being astronomy, particle scattering and space travel.

In this chapter, we treat the problem of a particle moving in a central force field with a *fixed centre*; this is called the **one-body problem**. The assumption that the centre of force is fixed is an accurate approximation in the context of planetary orbits. The combined mass of all the planets, moons and asteroids is less than 0.2% of the mass of the Sun. We therefore expect the motion of the Sun to be comparatively small, as are inter-planetary influences.* However, we do not confine our interest to motion under the attractive inverse square field. At first, we consider motion in *any* central force field with a fixed centre. This part of the theory will then apply not only to gravitating bodies, but also (for example) to the scattering of neutrons. The important cases of inverse square attraction and repulsion are then examined in greater detail.

* The more general **two-body problem** is treated in Chapter 10. The two-body theory must be used to analyse problems in which the *masses of the two interacting bodies are comparable*, as they are in binary stars.

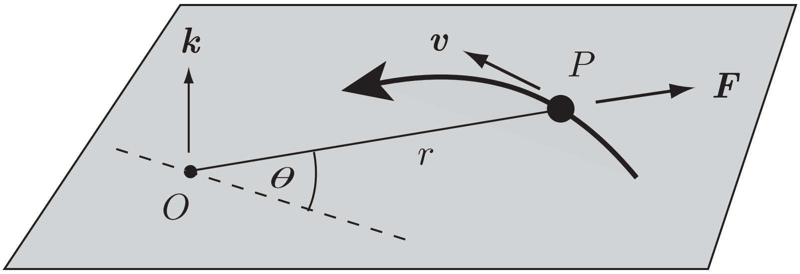


FIGURE 7.2 Each orbit of a particle P in a central force field with centre O takes place in a plane through O. The position of P in the plane of motion is specified by polar coordinates r, θ with centre at O.

7.1 THE ONE-BODY PROBLEM – NEWTON'S EQUATIONS

First we define what we mean by a central force field.

Definition 7.1 ***Central field*** *A force field $\boldsymbol{F}(\boldsymbol{r})$ is said to be a* ***central field*** *with centre O if it has the form*

$$\boldsymbol{F}(\boldsymbol{r}) = F(r)\,\widehat{\boldsymbol{r}},$$

where $r = |\boldsymbol{r}|$ and $\widehat{\boldsymbol{r}} = \boldsymbol{r}/r$. A central field is thus ***spherically symmetric*** *about its centre.*

A good example of a central force is the gravitational force exerted by a *fixed* point mass. Suppose P has mass m and moves under the gravitational attraction of a point mass M fixed at the origin. In this case, the force acting on P is given by the law of gravitation to be

$$\boldsymbol{F}(\boldsymbol{r}) = -\frac{mMG}{r^2}\,\widehat{\boldsymbol{r}},$$

where G is the constant of gravitation. This is a central field with

$$F(r) = -\frac{mMG}{r^2}.$$

Each orbit lies in a plane through the centre of force

The first thing to observe is that, when a particle P moves in a central field with centre O, *each orbit of P takes place in a plane through O*, as shown in Figure 7.2. This is the plane that contains O and the initial position and velocity of P. One may give a vectorial proof of this, but it is quite clear on symmetry grounds that P will never leave this plane. Each motion is therefore two-dimensional and we take polar coordinates r, θ (centred on O) to specify the position of P in the plane of motion. On using the formulae (2.14) for the components of acceleration in polar coordinates, the **Newton equations of motion** for

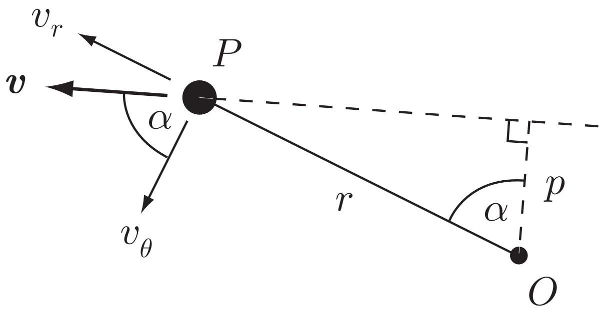


FIGURE 7.3 The angular momentum $mr^2\dot{\theta} = mp\,v$, where $v = |\boldsymbol{v}|$.

P become

$$m\left(\ddot{r} - r\dot{\theta}^2\right) = F(r), \tag{7.1}$$

$$m\left(r\ddot{\theta} + 2\dot{r}\dot{\theta}\right) = 0. \tag{7.2}$$

Angular momentum conservation

Equation (7.2) can be written in the form

$$\frac{1}{r}\frac{d}{dt}\left(mr^2\dot{\theta}\right) = 0,$$

which can be integrated with respect to t to give

$$mr^2\dot{\theta} = \text{constant}.$$

The quantity $mr^2\dot{\theta}$, which is a constant of the motion, is called the **angular momentum*** of P. The general theory of angular momentum (and its conservation) is described in Chapter 11, but for now it is sufficient to regard 'angular momentum' simply as a *name* that we give to the conserved quantity $mr^2\dot{\theta}$. This angular momentum has a simple kinematical interpretation. From Figure 7.3 it follows that

$$\begin{aligned} mr^2\dot{\theta} &= mr(r\dot{\theta}) = mr\,v_\theta = m(r\cos\alpha)\left(\frac{v_\theta}{\cos\alpha}\right) \\ &= mp\,v, \end{aligned}$$

where p is the perpendicular distance of O from the tangent to the path of P, and $v = |\boldsymbol{v}|$. This formula provides the usual way of calculating the constant value the angular momentum from the initial conditions.

* More precisely, it is the angular momentum of the particle about the *axis* $\{O, \boldsymbol{k}\}$, where the unit vector $\boldsymbol{k}$ is perpendicular to the plane of motion (see Figure 7.2). The angular momentum of P about the *point* O is the *vector* quantity $m\boldsymbol{r}\times\boldsymbol{v}$, but the axial angular momentum used in the present chapter is the component of this vector in the $\boldsymbol{k}$-direction.

Newton equations in specific form

It is usual and convenient to eliminate the mass m from the theory. If we write

$$F(r) = mf(r),$$

where $f(r)$ is the outward force *per unit mass*, and let L $(= r^2\dot{\theta})$ be the angular momentum *per unit mass* then the Newton equations (7.1), (7.2) reduce to the **specific form**

$$\ddot{r} - r\dot{\theta}^2 = f(r), \tag{7.3}$$

$$r^2\dot{\theta} = L, \tag{7.4}$$

where L is a constant.* Note that these equations apply to orbits in *any central field*. The second of these equations appears throughout this chapter and we will call it the *angular momentum equation*.

Angular momentum equation

$$r^2\dot{\theta} = L \tag{7.5}$$

Kepler's second law

Angular momentum conservation is equivalent to Kepler's second law. The area $\mathcal{A}$ shown in Figure 7.1 can be expressed (with an obvious choice of initial line) as

$$\mathcal{A} = \tfrac{1}{2}\int_0^{\theta} r^2\, d\theta.$$

Then, by the chain rule,

$$\frac{d\mathcal{A}}{dt} = \frac{d\mathcal{A}}{d\theta} \times \frac{d\theta}{dt} = \tfrac{1}{2}r^2\dot{\theta} = \tfrac{1}{2}L,$$

where L is the constant value of the angular momentum. Thus $\mathcal{A}$ increases at a constant rate, which is what Kepler's second law says. Thus Kepler's second law holds for *all* central force fields, not just the inverse square law.

7.2 GENERAL NATURE OF ORBITAL MOTION

In our first method of solution, we take as our starting point the principles of conservation of **angular momentum** and **energy**.

---

* Without losing generality, we will take L to be positive, that is, we suppose θ is *increasing* with time. (The special case in which $L = 0$ corresponds to rectilinear motion through O.)

Energy conservation

Every central field $\boldsymbol{F} = mf(r)\widehat{\boldsymbol{r}}$ is **conservative** with potential energy $mV(r)$, where

$$f(r) = -\frac{dV}{dr}. \tag{7.6}$$

Energy conservation then implies that

$$T + V = E,$$

where T is the specific kinetic energy, V is the specific potential energy, and the constant E is the specific total energy. On replacing T by its expression in polar coordinates, we obtain

Energy equation

$$\tfrac{1}{2}\left(\dot{r}^2 + (r\dot{\theta})^2\right) + V(r) = E \tag{7.7}$$

as the **energy conservation** equation. The conservation equations (7.5), (7.7) are equivalent to the Newton equations (7.1), (7.2) and are a convenient starting point for investigating the *general* nature of orbital motion.

The radial motion equation

From the angular momentum conservation equation (7.5), we have

$$\dot{\theta} = L/r^2$$

and, on eliminating $\dot{\theta}$ from the energy conservation equation (7.7), we obtain

$$\tfrac{1}{2}\dot{r}^2 + V(r) + \frac{L^2}{2r^2} = E, \tag{7.8}$$

an ODE for the radial distance $r(t)$. We call this the **radial motion equation** for the particle P. Equation (7.8) (together with the initial conditions) is sufficient to determine the variation of r with t, and the angular momentum equation (7.5) then determines the variation of θ with t. Unfortunately, for most laws of force, this procedure cannot be carried through analytically. However, it is still possible to make important deductions about the general nature of the motion.

Equation (7.8) can be written in the form

$$\tfrac{1}{2}\dot{r}^2 + V^*(r) = E, \tag{7.9}$$

where

$$V^*(r) = V(r) + \frac{L^2}{2r^2}. \tag{7.10}$$

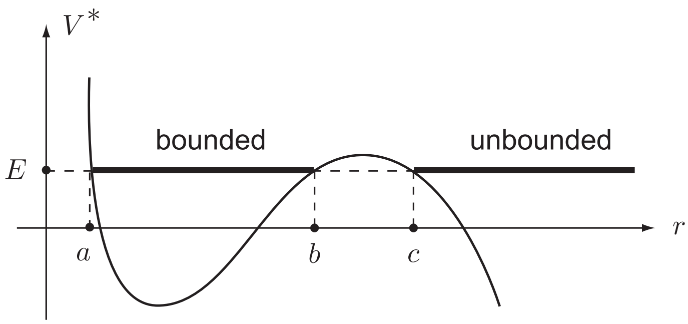


FIGURE 7.4 The effective potential V^* shown admits bounded and unbounded orbits, depending on the initial conditions.

The function $V^*(r)$ is called the **effective potential** of the radial motion and its use reduces the radial motion of P to a rectilinear problem. It must be emphasised though that the whole motion is *two*-dimensional since θ is increasing in accordance with (7.5).

Because r satisfies the radial motion equation (7.9), the variation of r with t can be analysed by using the same methods as were used in Chapter 6 for rectilinear particle motion. In particular, the general nature of the motion depends on the shape of the graph of V^* (which depends on L) and the value of E. The values of the constants L and E depend on the initial conditions.

Suppose for example that the law of force and the initial conditions are such that V^* has the form shown in Figure 7.4 and that E has the value shown. Then, since $T \geq 0$, it follows that the motion is restricted to those values r that satisfy the inequality

$$V^*(r) \leq E,$$

with equality holding when $\dot{r} = 0$. There are two possible motions, in each of which the variation of r with t is governed by the radial motion equation (7.8).

(i) a **bounded motion** in which r oscillates in the range $[a, b]$. In this motion, $r(t)$ is a periodic function.*

(ii) an **unbounded motion** in which r lies in the interval $[c, \infty)$. In this motion r is not periodic but decreases until the minimum value $r = c$ is achieved and then increases without limit.

The bounded orbit. A typical bounded orbit is shown in Figure 7.5 (left). The orbit alternately touches the inner and outer circles $r = a$ and $r = b$, which corresponds to the radial coordinate r oscillating in the interval $[a, b]$. Without losing generality, suppose that P is at the point B_1 when $t = 0$ and that OB_1 is the line $\theta = 0$. Consider the part of

* The fact that $r(t)$ is periodic does *not* mean that the whole motion must be periodic.

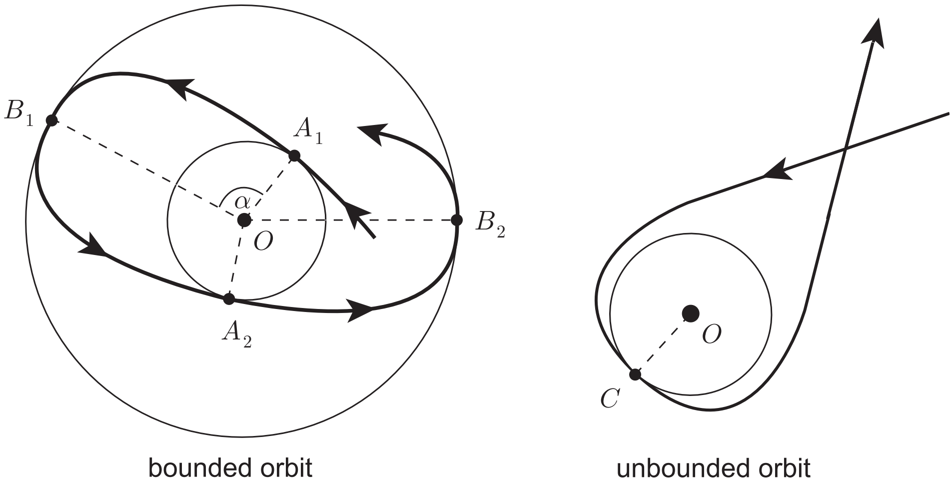


FIGURE 7.5 Typical bounded and unbounded orbits.

the orbit between A_1 and A_2. It follows from the governing equations (7.8), (7.5) that r is an *even* function of t while θ is an *odd* function of t. This means that the segment B_1A_2 of the orbit is just the reflection of the segment A_1B_1 in the line OB_1. This argument can be repeated to show that the segment A_2B_2 is the reflection of the segment B_1A_2 in the line OA_2, and so on. Thus the whole orbit can be constructed from a knowledge of a single segment such as A_1B_1.

It follows from what has been said that the angles $A_1\widehat{O}B_1$, $B_1\widehat{O}A_2$, $A_2\widehat{O}B_2$ (and so on) are all equal. Let α be the common value of these angles. Then the orbit will eventually close itself if some integer multiple of α is equal to some whole number of complete revolutions, that is, if α/π is a *rational number*. There is no reason to expect this condition to hold and, in general, it does not. It follows that these bounded orbits are *not generally closed*. The closed orbits associated with the attractive inverse square field are therefore exceptional, rather than typical!

The unbounded orbit. In the unbounded case there are just two segments both of which are semi-infinite (see Figure 7.5 (right)). The segment in which P recedes from O is the reflection of the segment in which P approaches O in the line OC.

Apses and apsidal distances

The points at which an orbit touches its bounding circles are important and are given a special name:

Definition 7.2 ***Apse, apsidal distance, apsidal angle*** *A point of an orbit at which the distance OP achieves its maximum or minimum value is called an* ***apse*** *of the orbit. These maximum and minimum distances are called the* ***apsidal distances*** *and the angular displacement between successive apses (the angle α in Figure 7.5 (left)) is called the* ***apsidal angle****.*

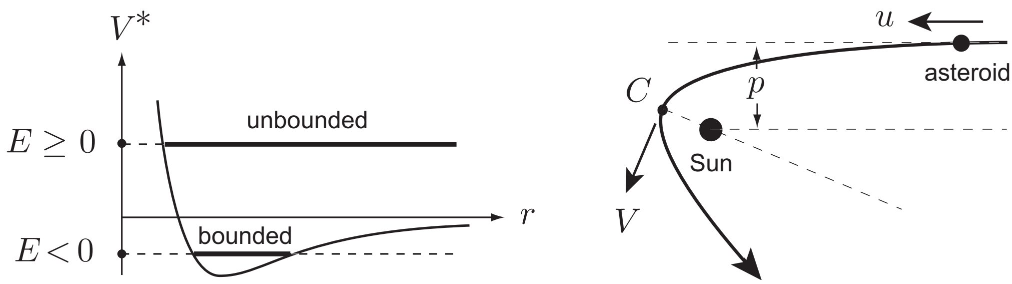


FIGURE 7.6 **Left:** The effective potential V^* for the attractive inverse square force. **Right:** The path of the asteroid around the Sun. C is the point of closest approach.

In the special case of orbits around the Sun, the point of closest approach is called the **perihelion** and the point of maximum distance the **aphelion**. The corresponding terms for orbits around the Earth are **perigee** and **apogee**.

The **apsidal distances**, the maximum and minimum distances of P from O, are easily found from the radial motion equation (7.8). At an apse, $\dot{r} = 0$ and so r must satisfy

$$V(r) + \frac{L^2}{2r^2} = E. \tag{7.11}$$

The positive roots of this equation are the apsidal distances.

Example 7.1 *Asteroid deflected by the Sun*

A particle P of mass m moves in the central force field $-(m\gamma/r^2)\widehat{\boldsymbol{r}}$, where γ is a positive constant. Show that bounded and unbounded orbits are possible depending on the value of E.

An asteroid is approaching the Sun from a great distance. At this time it has constant speed u and is moving in a straight line whose perpendicular distance from the Sun is p. Find the equation satisfied by the apsidal distances of the subsequent orbit. For the special case in which $u^2 = 4M_\odot G/3p$ (where $M_\odot$ is the mass of the Sun), find (i) the distance of closest approach of the asteroid to the Sun, and (ii) the speed of the asteroid at the time of closest approach.

Solution

For this law of force, $V = -\gamma/r$ and the effective potential V^* is

$$V^* = -\frac{\gamma}{r} + \frac{L^2}{2r^2}.$$

This V^* has the form shown in Figure 7.6 (left), from which it is clear that the orbit will be

(i) **bounded** if $E < 0$,

(ii) **unbounded** if $E \geq 0$,

whatever the value of L.

In the asteroid example, the constant $\gamma = M_\odot G$, where $M_\odot$ is the mass of the Sun and G is the constant of gravitation. With the given initial conditions, $L = p\,u$ and $E = u^2/2$, so that $E > 0$ and the orbit is **unbounded**.

The equation (7.11) for the **apsidal distances** becomes

$$-\frac{\gamma}{r} + \frac{p^2u^2}{2r^2} = \tfrac{1}{2}u^2,$$

that is,

$$u^2r^2 + 2\gamma r - p^2u^2 = 0,$$

where $\gamma = M_\odot G$.

For the special case in which $u^2 = 4M_\odot G/3p$, this equation simplifies to

$$2r^2 + 3pr - 2p^2 = 0.$$

The **distance** of closest approach of the asteroid is the *positive* root of this quadratic equation, namely $r = p/2$.

The **speed** V of the asteroid at closest approach is easily deduced from angular momentum conservation. Initially, $L = pu$ and, at closest approach, $L = (p/2)V$. It follows that $V = 2u$. ■

7.3 THE PATH EQUATION

In principle, the method of the last section allows us to determine the complete motion of the orbiting body as a function of the time. However, the procedure is usually too difficult to be carried through analytically. We can make the problem easier (and make more progress) by seeking just the **equation of the path** taken by the body, and not enquiring where the body is on this path at any particular time.

We start from the Newton equation (7.3) and try to eliminate the time by using the angular momentum equation (7.5). In doing this it is helpful to introduce the new dependent variable u, given by

$$u = 1/r. \tag{7.12}$$

This transformation has a magically simplifying effect. We begin by transforming $\dot{r}$ and $\ddot{r}$. By the chain rule,

$$\dot{r} = \frac{d}{dt}\left(\frac{1}{u}\right) = -\frac{1}{u^2} \times \frac{du}{d\theta} \times \frac{d\theta}{dt} = -\left(r^2\dot{\theta}\right)\frac{du}{d\theta}$$

which, on using the angular momentum equation (7.5), gives

$$\dot{r} = -L\frac{du}{d\theta}. \tag{7.13}$$

A second differentiation with respect to t then gives

$$\ddot{r} = -L\frac{d}{dt}\left(\frac{du}{d\theta}\right) = -L\frac{d^2u}{d\theta^2} \times \frac{d\theta}{dt} = -L^2u^2\frac{d^2u}{d\theta^2}, \tag{7.14}$$

on using the angular momentum equation again.

The term $r\dot{\theta}^2 = L^2u^3$ so that the Newton equation (7.3) is transformed into

$$-L^2u^2\frac{d^2u}{d\theta^2} - L^2u^3 = f(1/u),$$

that is,

The path equation

$$\frac{d^2u}{d\theta^2} + u = -\frac{f(1/u)}{L^2u^2} \tag{7.15}$$

This is the **path equation**. Its solutions are the polar equations of the paths that the body can take when it moves under the force field $\boldsymbol{F} = mf(r)\widehat{\boldsymbol{r}}$.

Despite the appearance of the left side of equation (7.15), the path equation is **not linear** in general. This is because the right side is a function of u, the *dependent* variable. Only for the **inverse square** and **inverse cube** laws does the path equation become linear. It is a remarkable piece of good luck that the inverse square law (the most important case by far) is one of only two cases that can be solved easily.

Initial conditions for the path equation

Suitable initial conditions for the path equation are provided by specifying the values of u and $du/d\theta$ when $\theta = \alpha$, say. Since $u = 1/r$, the initial value of u is given directly by the initial data. The value of $du/d\theta$ is not given directly but can be deduced from equation (7.13) in the form

$$\frac{du}{d\theta} = -\frac{\dot{r}}{L}, \tag{7.16}$$

where $\dot{r}$ and L are obtained from the initial data.

Example 7.2 *Path equation for the inverse cube law*

The engines of the starship Enterprise have failed and the ship is moving in a straight line with speed V. The crew calculate that their present course will miss the planet B – Zar by a distance p. However, B – Zar is known to exert the force

$$\boldsymbol{F} = -\frac{m\gamma}{r^3}\widehat{\boldsymbol{r}}$$

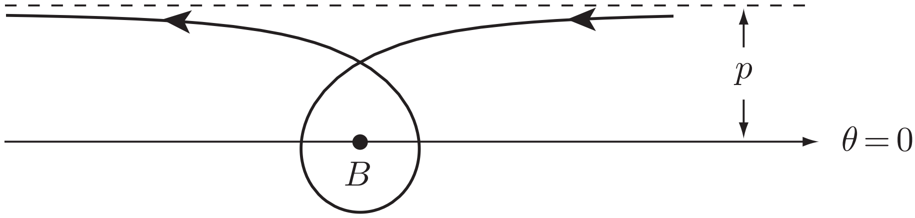


FIGURE 7.7 The path of the Enterprise around the planet B – Zar (B).

on any mass m in its vicinity. A measurement of the constant γ reveals that

$$\gamma = \frac{8p^2V^2}{9}.$$

Show that the crew of the Enterprise will get a free tour around B –Zar before continuing along their *original* path. What is the distance of closest approach and what is the speed of the Enterprise at that instant?

Solution

For the given law of force, $f(r) = -\gamma/r^3$ so that $f(1/u) = -\gamma u^3$. Also, from the initial conditions, $L = pV$. The path equation is therefore

$$\frac{d^2u}{d\theta^2} + u = \frac{\gamma u^3}{p^2V^2u^2},$$

which simplifies to

$$\frac{d^2u}{d\theta^2} + \frac{u}{9} = 0,$$

on using the stated value of γ. The general solution of this equation is

$$u = A\cos(\theta/3) + B\sin(\theta/3).$$

The constants A and B can now be determined from the initial conditions. Take the initial line $\theta = 0$ as shown in Figure 7.7. Then:

(i) The initial condition $r = \infty$ when $\theta = 0$ implies that $u = 0$ when $\theta = 0$. It follows that $A = 0$.

(ii) The initial condition on $du/d\theta$ is given by (7.16) to be

$$\frac{du}{d\theta} = -\frac{\dot{r}}{L} = -\left(\frac{-V}{pV}\right) = \frac{1}{p}$$

when $\theta = 0$. It follows that $B = 3/p$.

The required solution is therefore

$$u = \frac{3}{p}\sin(\theta/3),$$

that is

$$r = \frac{p}{3\sin(\theta/3)}.$$

This is the polar equation of the **path** of the Enterprise, as shown in Figure 7.7. The Enterprise recedes to infinity when $\sin(\theta/3) = 0$ again, that is when $\theta = 3\pi$. Thus the Enterprise makes one circuit of B–Zar before continuing on as before.

The distance of **closest approach** is $p/3$ and is achieved when $\theta = 3\pi/2$. By angular momentum conservation, the **speed** of the Enterprise at that instant is $3V$. ■

7.4 NEARLY CIRCULAR ORBITS

Although the path equation cannot be solved exactly for most laws of force, it is possible to obtain approximate solutions when the body is slightly perturbed from a *known* orbit. In particular, this can always be done when the unperturbed orbit is a circle with centre O.

Suppose that a particle P moves in a circular orbit of radius a under the *attractive* force $f(r)$ per unit mass. This is only possible if its speed v satisfies $v^2/a = f(a)$, in which case its angular momentum L is given by $L^2 = a^3 f(a)$. Suppose that P is now slightly disturbed by a small *radial* impulse. The angular momentum is unchanged but P now moves along some new path

$$u = \frac{1}{a}(1 + \xi(\theta)),$$

where ξ is a **small perturbation**. In terms of ξ, the path equation becomes

$$\frac{d^2\xi}{d\theta^2} + 1 + \xi = + \frac{(1+\xi)^{-2}}{f(a)} f\left(\frac{a}{1+\xi}\right).$$

This exact equation for ξ is non-linear, but we will now approximate it by expanding the right side in powers of ξ. On expanding the function $f(r)$ in a Taylor series about $r = a$ we obtain

$$\begin{aligned} f\left(\frac{a}{1+\xi}\right) &= f\left(a - \frac{a\xi}{1+\xi}\right) \\ &= f(a) - \left(\frac{a\xi}{1+\xi}\right) f'(a) + O\left(\frac{\xi}{1+\xi}\right)^2 \\ &= f(a) - af'(a)\,\xi + O\left(\xi^2\right), \end{aligned}$$

and a simple binomial expansion gives

$$(1+\xi)^{-2} = 1 - 2\xi + O\left(\xi^2\right).$$

On combining these results together, the constant terms cancel and we obtain

$$\frac{d^2\xi}{d\theta^2} + \left(3 + \frac{af'(a)}{f(a)}\right)\xi = 0, \tag{7.17}$$

on neglecting terms of order $O(\xi^2)$. This is the approximate **linearised equation** satisfied by the perturbation $\xi(\theta)$.

The general behaviour of the solutions of equation (7.17) depends on the *sign* of the coefficient of ξ.

(i) If

$$3 + \frac{af'(a)}{f(a)} < 0, \tag{7.18}$$

then the solutions are linear combinations of *real* exponentials, one of which has a positive exponent. In this case, the solution for ξ will not remain small, contrary to assumption. The conclusion is that the original circular orbit is **unstable**.

(ii) Alternatively, if

$$\Omega^2 \equiv 3 + \frac{af'(a)}{f(a)} > 0, \tag{7.19}$$

then the solutions are linear combinations of real cosines and sines, which remain bounded. The conclusion is that the original circular orbit is **stable** (at least to small radial impulses).

Closure of the perturbed orbits

From now on we will assume that the stability condition (7.19) is satisfied. The general solution of equation (7.17) then has the form

$$\xi = A\cos\Omega\theta + B\sin\Omega\theta.$$

We see that the perturbed orbit will **close** itself after one revolution if Ω is a **positive integer**. When the law of force is the **power law**

$$f(r) = kr^{\nu},$$

the perturbed orbit is stable for $\nu > -3$ and will close itself after one revolution if

$$\nu = m^2 - 3,$$

where m is a positive integer. The case $m = 1$ corresponds to inverse square attraction and $m = 2$ corresponds to simple harmonic attraction. The exponents $\nu = 6, 13, \ldots$ are also predicted to give closed orbits. It should be remembered though that these are only the predictions of the approximate linearised theory.* It is possible (but not pretty) to improve on the linear approximation by including quadratic terms in ξ as well as linear ones. The result of this refined theory is that the powers $\nu = -2$ and $\nu = 1$ still give

* It makes no sense to say that an orbit *approximately* closes itself!

closed orbits, but the powers $\nu = 6,\ 13,\ \ldots$ do not. This shows that the power laws with $\nu = 6,\ 13,\ \ldots$ do *not* give perturbed orbits that close after one revolution, but the cases $\nu = -2$ and $\nu = 1$ are still not finally decided. Mercifully, there is no need to carry the approximation procedure any further because all the paths corresponding to both inverse square and simple harmonic attraction can be calculated exactly. It is found that, for these two laws of force, *all bounded orbits close after one revolution*.* There remains the possibility that the perturbed orbits might close themselves after more than one revolution, but a similar analysis shows that this does not happen. We have therefore shown that *the only power laws for which all bounded orbits are closed are the simple harmonic and inverse square laws*. This result is actually true for all central fields (not just power laws) and is known as **Bertrand's theorem**.

Precession of the perihelion of Mercury

The fact that the inverse square law leads to closed orbits, whilst very similar laws do not, provides an extremely sensitive test of the law of gravitation. Suppose for instance that the attractive force experienced by a planet were

$$f(r) = \frac{\gamma}{r^{2+\epsilon}}$$

(per unit mass), where $\gamma > 0$ and $|\epsilon|$ is small. Then the value of Ω for a nearly circular orbit is

$$\Omega = (1-\epsilon)^{1/2} = 1 - \tfrac{1}{2}\epsilon + O(\epsilon^2).$$

This perturbed orbit does not close but has **apsidal angle** α, where

$$\alpha = \frac{\pi}{\Omega} = \frac{\pi}{1 - \frac{1}{2}\epsilon + O(\epsilon^2)} = \pi(1 + \tfrac{1}{2}\epsilon) + O(\epsilon^2).$$

Hence successive perihelions of the planet will not occur at the same point, but the **perihelion will advance** 'annually' by the small angle $\pi\epsilon$. The position of the perihelion of a planet can be measured with great accuracy. For the planet Mercury it is found (after all known perturbations have been subtracted out) that the perihelion advances by 43 (± 0.5) seconds of arc per century, or 5×10^{-7} radians per revolution. This correponds to $\epsilon = 1.6 \times 10^{-7}$ and a power of -2.00000016 instead of -2. Miniscule though this discrepancy from the inverse square law seems, it is considerably greater than the error in the observations and for a considerable time was something of a puzzle.

This puzzle was resolved in a striking fashion by the theory of **general relativity**, published by Einstein in 1915. Einstein showed that one consequence of his theory was that planetary orbits *should* precess slightly and that, in the case of Mercury, the rate of precession should be 43 seconds of arc per century!

---

* In the inverse square case, the bounded orbits are ellipses with a *focus* at O, and, in the simple harmonic case, they are ellipses with the *centre* at O.

7.5 THE ATTRACTIVE INVERSE SQUARE FIELD

Because of its many applications to **astronomy**, the attractive inverse square field is the most important force field in the theory of orbits. The same field occurs in particle scattering when the two particles carry unlike electric charges. Because of these important applications, we will treat the inverse square field in more detail than other fields. In particular, we will obtain formulae that enable inverse square problems to be solved quickly and easily without referring to the equations of motion at all.

The paths

Suppose that $f(r) = -\gamma/r^2$ where $\gamma > 0$. Then $f(1/u) = -\gamma u^2$ and the path equation becomes

$$\frac{d^2u}{d\theta^2} + u = \frac{\gamma}{L^2},$$

where L is the angular momentum of the orbit. This has the form of the SHM equation with a constant on the right. The general solution is

$$u = A\cos\theta + B\sin\theta + \frac{\gamma}{L^2},$$

which can be written in the form

$$\frac{1}{r} = \frac{\gamma}{L^2}\Big(1 + e\cos(\theta - \alpha)\Big), \tag{7.20}$$

where e, α are constants with $e \geq 0$. This is the **polar equation of a conic** of eccentricity e and with one focus at O; α is the angle between the major axis of the conic and the initial line $\theta = 0$. If $e < 1$, then the conic is an **ellipse**; if $e = 1$ then the conic is a **parabola**; and when $e > 1$ the conic is the *near* branch of a **hyperbola**. The neccessary geometry of the ellipse and hyperbola is summarised in Appendix A at the end of the chapter; the special case of the parabolic orbit is of marginal interest and we will make little mention of it.

Kepler's first law

It follows from the above that the only bounded orbits in the attractive inverse square field are **ellipses** with one **focus at the centre of force**. This is Kepler's first law, which is therefore a consequence of inverse square law attraction by the Sun. It would not be true for other laws of force.

The L-formula and the E-formula

By comparing the path formula (7.20) with the standard polar forms given in Appendix A, we see that the angular momentum L of the orbit is related to the conic parameters a,

b by the formula

$$\frac{\gamma}{L^2} = \frac{a}{b^2},$$

that is,

The L-formula

$$L^2 = \gamma b^2/a \tag{7.21}$$

We will call this result the **L-formula**. It applies to both elliptic and hyperbolic orbits. It is the first of two important formulae that relate L, E, the dynamical constants of the motion, to the conic parameters of the resulting orbit.

The second such formula involves the energy E. At the point of closest approach $r = c$,

$$E = \tfrac{1}{2}V^2 - \frac{\gamma}{c},$$

where V is the speed of P when $r = c$. Since P is moving transversely at the point of closest approach, it follows that $cV = L$, so that E may be written

$$E = \frac{L^2}{2c^2} - \frac{\gamma}{c} = \frac{\gamma b^2}{ac^2} - \frac{\gamma}{c}$$

on using the L-formula.

From this point on, the different types of conic must be treated separately. When the orbit is an ellipse, $c = a(1 - e)$, where e is the eccentricity, and a, b and e are related by the formula

$$e^2 = 1 - \frac{b^2}{a^2}.$$

Then E can be written

$$\begin{aligned} E &= \frac{\gamma a^2(1-e^2)}{2a^3(1-e)^2} - \frac{\gamma}{a(1-e)} \\ &= -\frac{\gamma}{2a}. \end{aligned}$$

Thus the total energy E in the orbit is directly connected to a, the semi-major axis of the elliptical orbit. The parabolic and hyperbolic orbits are treated similarly and the full result, which we will call the **E-formula**, is

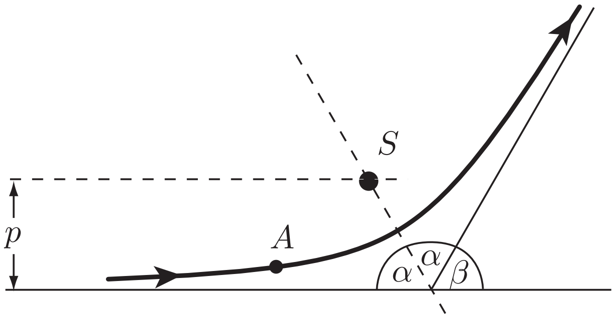


FIGURE 7.8 The asteroid A moves on a hyperbolic orbit around the Sun S as a focus and is deflected through the angle β.

The E-formula

Ellipse:	$E < 0$	$E = -\dfrac{\gamma}{2a}$
Parabola:	$E = 0$	
Hyperbola:	$E > 0$	$E = +\dfrac{\gamma}{2a}$

(7.22)

Note that the type of orbit is determined solely by the *sign* of the total energy E. It follows that that the **escape condition** (the condition that the body should eventually go off to infinity) is simply that $E \geq 0$. This useful result is true only for the inverse square law.

Example 7.3 *Asteroid deflected by the Sun*

An asteroid approaches the Sun with speed V along a line whose perpendicular distance from the Sun is p. Find the angle through which the asteroid is deflected by the Sun.

Solution

In this case we have the attractive inverse square field with $\gamma = M_{\odot}G$, where $M_{\odot}$ is the mass of the Sun. This problem can be solved from first principles by using the path equation, but here we make short work of it by using the L- and E-formulae.

From the initial conditions, $L = pV$ and $E = \frac{1}{2}V^2$. Since $E > 0$, the orbit is the near branch of a **hyperbola** and the L- and E-formulae give

$$p^2V^2 = \frac{M_{\odot}Gb^2}{a} \qquad \text{and} \qquad \tfrac{1}{2}V^2 = +\frac{M_{\odot}G}{2a}.$$

It follows that

$$a = \frac{M_\odot G}{V^2}, \qquad b = p.$$

The semi-angle α between the asymptotes of the hyperbola is then given (see Appendix A) by

$$\tan\alpha = \frac{b}{a} = \frac{pV^2}{M_\odot G}.$$

Let β be the angle through which the asteroid is deflected. Then (see Figure 7.8) $\beta = \pi - 2\alpha$ and

$$\tan(\beta/2) = \tan(\pi/2 - \alpha) = \cot\alpha = \frac{M_\odot G}{pV^2}. \blacksquare$$

Period of the elliptic orbit

Whatever the law of force, once the path of P has been found, the progress of P along that path can be deduced from the angular momentum equation

$$r^2\dot{\theta} = L.$$

If we take $\theta = 0$ when $t = 0$, then the time t taken for P to progress to the point of the orbit with polar coordintes r, θ is given by

$$t = \frac{1}{L}\int_0^\theta r^2\, d\theta, \tag{7.23}$$

where $r = r(\theta)$ is the equation of the path. In particular then, the period τ of the elliptic orbit is given by

$$\tau = \frac{1}{L}\int_0^{2\pi} r^2\, d\theta,$$

where the path $r = r(\theta)$ is given by

$$\frac{1}{r} = \frac{a}{b^2}\left(1 + e\cos\theta\right). \tag{7.24}$$

Fortunately there is no need to evaluate the above integral since, for any path that closes itself after one circuit,

$$\tfrac{1}{2}\int_0^{2\pi} r^2\, d\theta = A,$$

where A is the area enclosed by the path. For the elliptical path, $A = \pi ab$ so that

$$\tau = \frac{2\pi ab}{L},$$

and on using the L-formula, the period of the elliptic orbit is given by:

The period formula

$$\tau = 2\pi \left(\frac{a^3}{\gamma} \right)^{1/2} \qquad (7.25)$$

Kepler's third law

In the case of the planetary orbits, $\gamma = M_\odot G$, where $M_\odot$ is the mass of the Sun. Equation (7.25) can then be written

$$\tau^2 = \left(\frac{4\pi^2}{M_\odot G} \right) a^3. \qquad (7.26)$$

This is Kepler's third law, which is therefore a consequence of inverse square law attraction by the Sun and would not be true for other laws of force.

Masses of celestial bodies

Once the constant of gravitation G is known, the formula (7.26) provides an accurate way to find the mass of the Sun. The same method applies to *any celestial body that has a satellite*. All that is needed is to measure the major axis $2a$ and the period τ of the satellite's orbit.*

Question ***Finding the mass of Jupiter***

The Moon moves in a nearly circular orbit of radius 384,000 km and period 27.32 days. Callisto, the fourth moon of the planet Jupiter, moves in a nearly circular orbit of radius 1,883,000 km and period 16.69 days. Estimate the mass of Jupiter as a multiple of the mass of the Earth.

Answer

$M_J = 316 M_E$.

Astronomical units

For astronomical problems, it is useful to write the period equation (7.26) in **astronomical units**. In these units, the unit of mass is the mass of the Sun ($M_\odot$), the unit of length (the AU) is the semi-major axis of the Earth's orbit, and the unit of time is the (Earth) year. On

* It should be noted that here we are neglecting the motion of the centre of force. We will see later that, when this is taken into account, formula (7.26) actually gives the *sum* of the masses of the body and its satellite. Usually, the satellite has a much smaller mass than the body and its contribution can be disregarded.

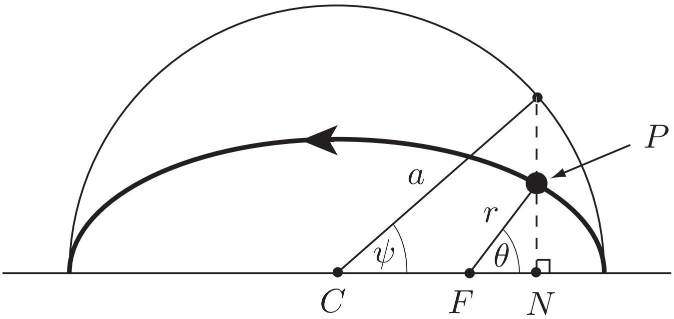


FIGURE 7.9 The eccentric angle ψ corresponding to the polar angle θ.

substituting the data for the Earth and Sun into equation (7.26), we find that $G = 4\pi^2$ in astronomical units. Hence, in **astronomical units** the period formula becomes

$$\tau^2 = \frac{a^3}{M}.$$

Question ***The major axis of the orbit of Pluto***

The period of Pluto is 248 years. What is the semi-major axis of its orbit?

Answer

39.5 AU. ■

Time dependence of the motion – Kepler's equation

The formula (7.23) can be used to find how long it takes for P to progress to a general point of the orbit. However, although the integration with respect to θ can be done in closed form, it is a *very* complicated expression. In order to obtain a manageable formula, we make a cunning change of variable, replacing the polar angle θ by the **eccentric angle** ψ. The relationship between these two angles is shown in Figure 7.9. Since $CN = CF + FN$, it follows that

$$a\cos\psi = ae + r\cos\theta,$$

and, on using the polar equation for the ellipse (7.24) together with the formula $b^2 = a^2(1 - e^2)$, the relation between ψ and θ can be written in the symmetrical form

$$(1 - e\cos\psi)(1 + e\cos\theta) = \frac{b^2}{a^2}. \tag{7.27}$$

Implicit differentiation of equation (7.27) with respect to ψ then gives

$$\frac{d\theta}{d\psi} = \frac{b}{a\,(1 - e\cos\psi)}, \tag{7.28}$$

after more manipulation.

We can now make the change of variable from θ to ψ. From (7.23) and (7.24)

$$
\begin{aligned}
t &= \frac{b^4}{a^2 L} \int_0^{\theta} \frac{d\theta}{(1 + e\cos\theta)^2} \\
&= \frac{b^4}{a^2 L} \int_0^{\psi} \frac{1}{(1 + e\cos\theta)^2} \left(\frac{d\theta}{d\psi} \right) d\psi \\
&= \frac{ab}{L} \int_0^{\psi} (1 - e\cos\psi)\, d\psi, \\
&= \frac{ab}{L} (\psi - e\sin\psi),
\end{aligned}
$$

on using (7.27), (7.28). Finally, on making use of the L-formula $L^2 = \gamma b^2/a$, we obtain

Kepler's equation

$$t = \frac{\tau}{2\pi} (\psi - e\sin\psi) \tag{7.29}$$

where τ (given by (7.25)) is the period of the orbit. This is **Kepler's equation** which gives the time as a function of position on the elliptical orbit.

If one needs to calculate the position of the orbiting body after a *given time*, then equation (7.29) must be solved numerically for the eccentric angle ψ. The corresponding value of θ is then given by equation (7.27) and the r value by equation (7.24) which, in view of (7.27), can be written in the form

$$r = a\,(1 - e\cos\psi)\,. \tag{7.30}$$

The need to solve Kepler's equation for the unknown ψ was a major stimulus in the development of approximate numerical methods for finding roots of equations.

Example 7.4 *Kepler's equation*

A body moving in an inverse square attractive field traverses an elliptical orbit with eccentricity e and period τ. Find the time taken for the body to traverse the half of the orbit that is nearer the centre of force.

Solution

The half of the orbit nearer the centre of force corresponds to the range $-\pi/2 \le \psi \le \pi/2$. The time taken is therefore

$$\frac{\tau}{\pi}\left(\frac{\pi}{2} - e\right) = \tau\left(\frac{1}{2} - \frac{e}{\pi}\right).$$

For example, Halley's comet moves on an elliptic orbit whose eccentricity is almost unity. It therefore spends only about 18% of its time on the half of its orbit that is nearer the Sun.

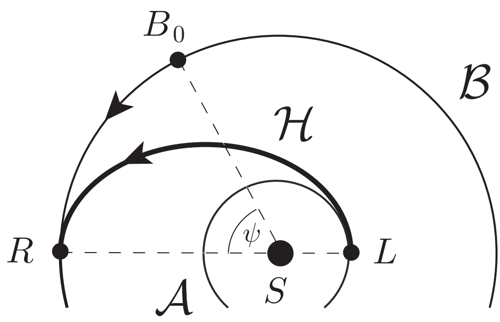


FIGURE 7.10 Two planets move on the circular orbits $\mathcal{A}$ and $\mathcal{B}$. A spacecraft is required to depart from one planet and rendezvous with the other planet at some point of its orbit. The Hohmann orbit $\mathcal{H}$ achieves this with the least expenditure of fuel.

7.6 SPACE TRAVEL – HOHMANN TRANSFER ORBITS

An important problem in space travel, and one that nicely illustrates the preceding theory, is that of transferring a spacecraft from one planet to another (from Earth to Jupiter say). In order to simplify the analysis, we will assume that both of the planetary orbits are circular. We will also suppose that the spacecraft has already effectively been removed from Earth's gravity, but is still in the vicinity of the Earth and is orbiting the Sun on the same orbit as the Earth. The object is to use the rocket motors to transfer the spacecraft to the vicinity of Jupiter, orbiting the Sun on the same orbit as Jupiter. Like everything else on board a spacecraft, fuel has to be transported from Earth at huge cost, so the transfer from Earth to Jupiter must be achieved using the *least mass of fuel*. In our analysis we will neglect the time during which the rocket engines are firing so that the engines are regarded as delivering an impulse to the spacecraft, resulting in a sudden change of velocity. After the initial firing impulse, the spacecraft is assumed to move freely under the Sun's gravitation until it reaches the orbit of Jupiter, when a second firing impulse is required to circularise the orbit. This is called a **two-impulse transfer**.

If the two firings produce velocity changes of $\Delta\boldsymbol{v}^A$ and $\Delta\boldsymbol{v}^B$ respectively, then the quantity Q that must be minimised if the least fuel is to be used is

$$Q = |\Delta\boldsymbol{v}^A| + |\Delta\boldsymbol{v}^B|.$$

The orbit that connects the two planetary orbits and minimises Q is called the **Hohmann transfer orbit*** and is shown in Figure 7.10. It has its perihelion at the lift-off point L and its aphelion at the rendezvous point R. It is not at all obvious that this is the optimal orbit; a proof is given in Appendix B at the end of the chapter. However, it is quite easy to find its properties.

Since the perihelial and aphelial distances in the Hohmann orbit are A and B (the radii of the orbits of Earth and Jupiter), it follows that

$$A = a(1 - e), \qquad B = a(1 + e),$$

* After Walter Hohmann, the German space research pioneer.

so that the geometrical parameters of the orbit are given by

$$a = \tfrac{1}{2}(B + A), \qquad e = \frac{B - A}{B + A}.$$

The angular momentum L of the orbit is then given by the L-formula to be

$$L^2 = \frac{\gamma b^2}{a} = \gamma\left(1 - e^2\right)a = \frac{\gamma B A}{B + A},$$

where $\gamma = M_{\odot}G$.

From L we can find the **speed** V^L of the spacecraft just after the lift-off firing, and the **speed** V^R at the rendezvous point just before the second firing. These are

$$V^L = \left(\frac{2\gamma B}{A(B + A)}\right)^{1/2}, \qquad V^R = \left(\frac{2\gamma A}{B(B + A)}\right)^{1/2}.$$

The **travel time** T, which is half the period of the Hohmann orbit, is given by

$$T^2 = \frac{\pi^2 a^3}{\gamma} = \frac{\pi^2 (B + A)^3}{8\gamma}.$$

Finally, in order to rendezvous with Jupiter, the lift-off must take place when Earth and Jupiter have the correct relative positions, so that Jupiter arrives at the meeting point at the right time. Since the speed of Jupiter is $(\gamma/B)^{1/2}$ and the travel time is now known, the angle ψ in Figure 7.10 must be

$$\psi = \pi\left(\frac{B + A}{2B}\right)^{3/2}.$$

Numerical results for the Earth–Jupiter transfer

In astronomical units, $G = 4\pi^2$, $A = 1$ AU and, for Jupiter, $B = 5.2$ AU. A speed of 1 AU per year is 4.74 km per second. Simple calculations then give:

(i) The travel time is 2.73 years, or 997 days.

(ii) V^L is 8.14 AU per year, which is 38.6 km per second. This is the speed the spacecraft must have after the lift-off firing.

(iii) V^R is 1.56 AU per year, which is 7.4 km per second. This is the speed with which the spacecraft arrives at Jupiter before the second firing.

(iv) The angle ψ at lift-off must be 83°.

The speeds V^L and V^R should be compared with the speeds of Earth and Jupiter in their orbits. These are 29.8 km/sec and 13.1 km/sec respectively. Thus the first firing must boost the speed of the spacecraft from 29.8 to 38.6 km/sec, and the second firing must boost the speed from 7.4 to 13.1 km/sec. The sum of these speed increments, 14.5 km/sec, is greater than the speed increment needed (12.4 km/sec) to escape from the Earth's orbit to infinity. Thus it takes more fuel to transfer a spacecraft from Earth's orbit to Jupiter's orbit than it does to escape from the solar system altogether!

7.7 THE REPULSIVE INVERSE SQUARE FIELD

The force field with $f(r) = +\gamma/r^2$, ($\gamma > 0$), is the **repulsive inverse square field**. It occurs in the interaction of charged particles carrying *like* charges and is required for the analysis of Rutherford scattering. Below we summarise the important properties of orbits in a repulsive inverse square field. These results are obtained in exactly the same way as for the attractive case.

The paths

The path equation is

$$\frac{d^2u}{d\theta^2} + u = -\frac{\gamma}{L^2},$$

where L is the angular momentum of the orbit. Its general solution can be written in the form

$$\frac{1}{r} = \frac{\gamma}{L^2}\left[-1 + e\cos(\theta - \alpha)\right],$$

where e, α are constants with $e \geq 0$. By comparing this path with the standard polar forms of conics given in Appendix A, we see that the path can only be the *far* branch of a **hyperbola** with focus at the centre O.

The L- and E-formulae

The formulae relating L, E, the dynamical constants of the orbit, to the hyperbola parameters are

$$L^2 = \gamma b^2/a, \tag{7.31}$$

$$E = +\gamma/2a. \tag{7.32}$$

7.8 RUTHERFORD SCATTERING

The most celebrated application of orbits in a repulsive inverse square field is Rutherford's* famous experiment in which a beam of alpha particles was scattered by gold nuclei in a sheet of gold leaf. We will analyse Rutherford's experiment in detail, beginning with the basic problem of a single alpha particle being deflected by a single fixed gold nucleus.

Alpha particle deflected by a heavy nucleus

An alpha particle A of mass m and charge q approaches a gold nucleus B of charge Q (see Figure 7.11). B is initially at rest and A is moving with speed V along a line whose

* Ernest Rutherford (1871–1937), a New Zealander, was one of the greatest physicists of the twentieth century. His landmark work on the structure of the nucleus in 1911 (and with Geiger and Marsden in 1913) was conducted at the University of Manchester, England.

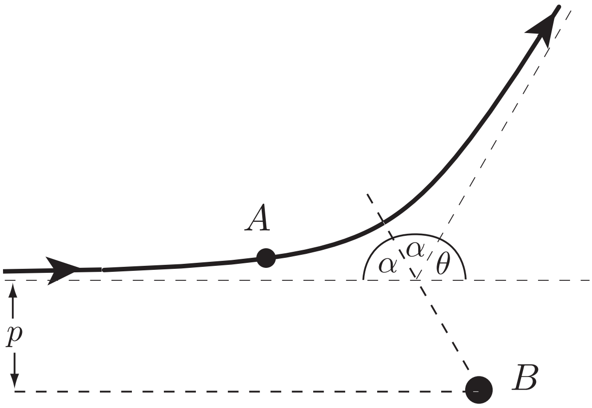


FIGURE 7.11 The alpha particle A of mass m and charge q is repelled by the fixed nucleus B of charge Q and moves on a hyperbolic orbit with the nucleus at the far focus. The alpha particle is deflected through the angle θ.

perpendicular distance from B is p. In the present treatment, we neglect the motion of the gold nucleus. This is justified since the mass of the gold nucleus is about fifty times larger than that of the alpha particle. A then moves in the electrostatic field due to B, which we now suppose to be fixed at the origin O. The force exerted on A is then

$$\boldsymbol{F} = +\frac{qQ}{r^2}\,\widehat{\boldsymbol{r}}$$

in cgs units. This is the **repulsive inverse square** field with $\gamma = qQ/m$.

We wish to find θ, the angle through which the alpha particle is deflected. This is obtained in exactly the same way as that of the asteroid in Example 7.1. From the initial conditions, $L = pV$ and $E = \frac{1}{2}V^2$. The L-formula (7.31) and the E-formula (7.32) then give

$$p^2V^2 = \frac{\gamma b^2}{a}, \qquad \tfrac{1}{2}V^2 = +\frac{\gamma}{2a}.$$

It follows that

$$a = \frac{\gamma}{V^2}, \qquad b = p.$$

The semi-angle α between the asymptotes of the hyperbola is then given (see Appendix A) by

$$\tan\alpha = \frac{b}{a} = \frac{pV^2}{\gamma}.$$

Hence, θ, the angle through which the asteroid is deflected, is given by

$$\tan(\theta/2) = \tan(\pi/2 - \alpha) = \cot\alpha = \frac{\gamma}{pV^2}.$$

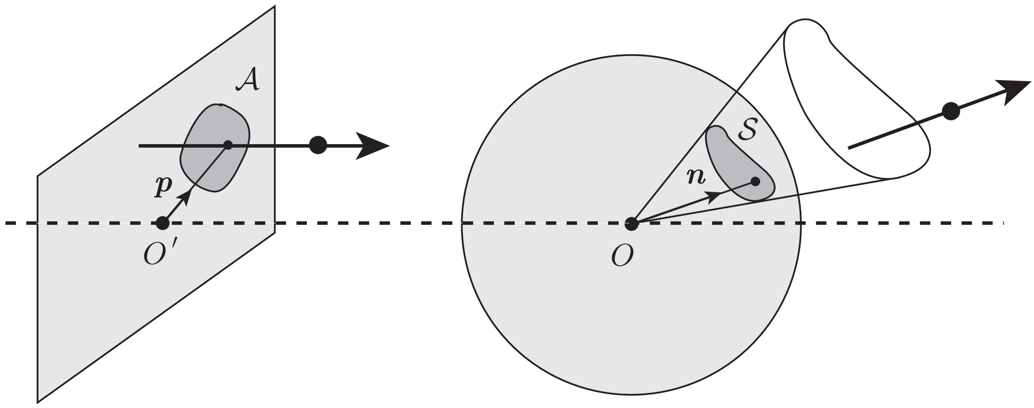


FIGURE 7.12 **General scattering**. A typical particle crosses the reference plane at the point $\boldsymbol{p}$ and finally emerges in the direction of the unit vector $\boldsymbol{n}$. Particles that cross the reference plane within the region $\mathcal{A}$ emerge within the (generalised) cone shown.

On writing $\gamma = qQ/m$, we obtain

$$\tan(\theta/2) = \frac{qQ}{mpV^2}. \tag{7.33}$$

as the formula for the **deflection angle** of the alpha particle. The quantity p, the distance by which the incident particle would miss the scatterer if there were no interaction, is called the **impact parameter** of the particle.

The deflection formula (7.33) cannot be confirmed directly by experiment since this would require the observation of a single alpha particle, a single nucleus, and a knowledge of the impact parameter p. What is actually done is to irradiate a gold target by a uniform beam of alpha particles of the same energy. Thus the target consists of many gold nuclei together with their associated electrons. However, the electrons have masses that are very small compared to that of an alpha particle and so their influence can be disregarded. Also, the gold target is taken in the form of thin foil to minimise the chance of multiple collisions. If multiple collisions are eliminated, then the gold nuclei act as *independent* scatterers and the problem reduces to that of a *single* fixed gold nucleus irradiated by a *uniform beam* of alpha particles. In this problem the alpha particles come in with different values of the impact parameter p and are scattered through different angles in accordance with formula (7.33). What *can* be measured is the **angular distribution** of the scattered alpha particles.

Differential scattering cross-section

The angular distribution of scattered particles is expressed by a function $\sigma(\boldsymbol{n})$, called the **differential scattering cross section**, where the unit vector $\boldsymbol{n}$ specifies the final direction of emergence of a particle from the scatterer O. One may imagine the values of $\boldsymbol{n}$ corresponding to points on the surface of a sphere with centre O and unit radius, as shown in Figure 7.12. Then values of $\boldsymbol{n}$ that lie in the shaded patch $\mathcal{S}$ correspond to particles whose final direction of emergence lies inside the (generalised) cone shown.

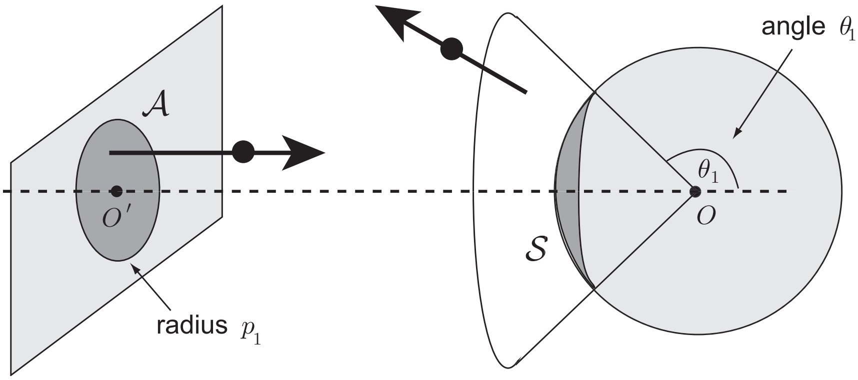


FIGURE 7.13 Axisymmetric scattering. Particles crossing the reference plane within the shaded circular disk are scattered and emerge in directions within the circular cone.

Take a reference plane far to the left of the scatterer and perpendicular to the incident beam, as shown in Figure 7.12. Suppose that there is a uniform flux of incomimg particles crossing the reference plane such that N particles cross any unit area of the reference plane in unit time. When these particles have been scattered, they will emerge in different directions and some of the particles will emerge with directions lying within the (generalised) cone shown in Figure 7.12. The **differential scattering cross section** is defined to be that function $\sigma(\boldsymbol{n})$ such that the flux of particles that emerge with directions lying within the cone is given by the surface integral

$$N \int_{\mathcal{S}} \sigma(\boldsymbol{n})\, dS. \tag{7.34}$$

It is helpful to regard $\sigma(\boldsymbol{n})$ as a **scattering density**, analogous to a probability density, that must be integrated to give the flux of particles scattered within any given solid angle.

The particles that finally emerge within the cone must have crossed the reference plane within some region $\mathcal{A}$ as shown in Figure 7.12. A typical particle crosses the reference plane at the point $\boldsymbol{p}$ (relative to O') and eventually emerges in the direction $\boldsymbol{n}$ lying within the cone. However because the incoming beam is uniform, the flux of these particles across $\mathcal{A}$ is just $N|\mathcal{A}|$, where $|\mathcal{A}|$ is the **area** of the region $\mathcal{A}$. On equating the incoming and outgoing fluxes, we obtain the relation

$$\int_{\mathcal{S}} \sigma(\boldsymbol{n})\, dS = |\mathcal{A}|. \tag{7.35}$$

This is the general relation that *any* differential scattering cross section must satisfy; it simply expresses the equality of incoming and outgoing fluxes of particles. However, Rutherford scattering is axisymmetric and this provides a major simplification.

Axisymmetric scattering and Rutherford's formula

Rutherford scattering is simpler than the general case outlined above in that the problem is **axisymmetric** about the axis $O'O$. Thus σ depends on θ (the angle between $\boldsymbol{n}$ and the axis $O'O$), but is independent of ϕ (the azimuthal angle measured around the axis). In this case $\sigma(\theta)$ can be determined by using the axisymmetric regions shown in Figure 7.13. Particles that cross the reference plane within the *circle* centre O' and radius p_1 emerge within the *circular* cone $\theta_1 \le \theta \le \pi$, where p_1 and θ_1 are related by the deflection formula for a single particle, in our case formula (7.33). On applying equation (7.35) to the present case, we obtain

$$\int_{\mathcal{S}} \sigma(\theta)\, dS = \pi p_1^2.$$

We evaluate the surface integral using θ, ϕ coordinates. The element of surface area on the unit sphere is given by $dS = \sin\theta\, d\theta d\phi$ so that

$$\begin{aligned}\int_{\mathcal{S}} \sigma(\theta)\, dS &= \int_{\theta_1}^{\pi} \left\{ \int_0^{2\pi} \sigma(\theta) \sin\theta\, d\phi \right\} d\theta \\ &= 2\pi \int_{\theta_1}^{\pi} \sigma(\theta) \sin\theta\, d\theta.\end{aligned}$$

Hence

$$\begin{aligned}2\pi \int_{\theta_1}^{\pi} \sigma(\theta) \sin\theta\, d\theta &= \pi p_1^2 \\ &= 2\pi \int_0^{p_1} p\, dp \\ &= -2\pi \int_{\theta_1}^{\pi} p\, \frac{dp}{d\theta}\, d\theta,\end{aligned}$$

on changing the integration variable from p to θ. Here the impact parameter p is regarded as a function of the scattering angle θ. Now the above equality holds for all choices of the integration limit θ_1 and this can only be true if the two *integrands* are equal. Hence:

Axisymmetric scattering cross section

$$\sigma(\theta) = -\left(\frac{p}{\sin\theta}\right) \frac{dp}{d\theta} \tag{7.36}$$

This is the formula for the differential scattering cross section σ in any problem of **axisymmetric scattering**. All that is needed to evaluate it in any particular case is the expression for the impact parameter p in terms of the scattering angle θ.

In the case of **Rutherford scattering**, the expression for p in terms of θ is provided by solving equation (7.33) for p, which gives

$$p = \frac{qQ}{mV^2} \tan(\theta/2).$$

On substituting this function into the formula (7.36), we obtain

Rutherford's scattering cross-section

$$\sigma(\theta) = \frac{q^2 Q^2}{16E^2}\left(\frac{1}{\sin^4(\theta/2)}\right) \tag{7.37}$$

where $E(=\frac{1}{2}mV^2)$ is the energy of the incident alpha particles. This is **Rutherford's formula** for the angular distribution of the scattered alpha particles.

Significance of Rutherford's experiment

In the above description we have used the term 'nucleus' for convenience. What we really mean is '*the positively charged part of the atom that carries most of the mass*'. If this positive charge is distributed in a spherically symmetric manner, then the above results still hold, irrespective of the radius of the charge, provided that the alpha particles *do not penetrate* into the charge itself. What Rutherford found was that, when using alpha particles from a radium source, the formula (7.37) held even for particles that were scattered through angles close to π. These are the particles that get closest to the nucleus, the distance of closest approach being qQ/E. This meant that the nuclear radius of gold must be smaller than this distance, which was about 10^{-12} cm in Rutherford's experiment. The radius of an atom of gold is about 10^{-8} cm. This result completely contradicted the Thompson model, in which the positive charge was distributed over the whole volume of the atom, by showing that the nucleus (as it became known) must be a very small and very dense core at the centre of the atom.

Note on two-body scattering problems

Throughout this section we have neglected the motion of the target nucleus. This will introduce only small errors when the target nucleus is much heavier than the incident particles, as it was in Rutherford's experiment. However, if lighter nuclei are used as the target, then the motion of the nucleus cannot be neglected and we have a **two-body scattering problem**. Such problems are treated in Chapter 10.

Appendix A The geometry of conics

Ellipse

(i) In **Cartesian coordinates**, the standard ellipse with **semi-major axis** a and **semi-minor axis** b ($b \le a$) has the equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

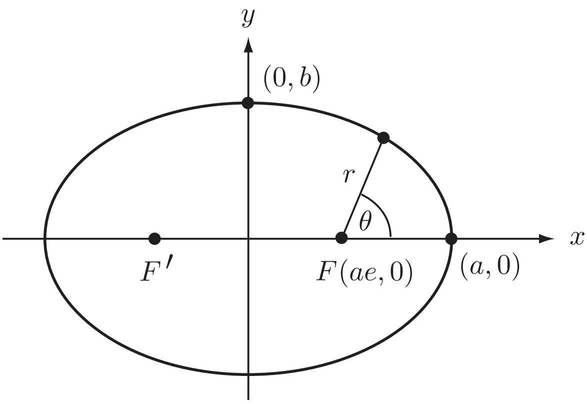


FIGURE 7.14 The standard ellipse $x^2/a^2 + y^2/b^2 = 1$.

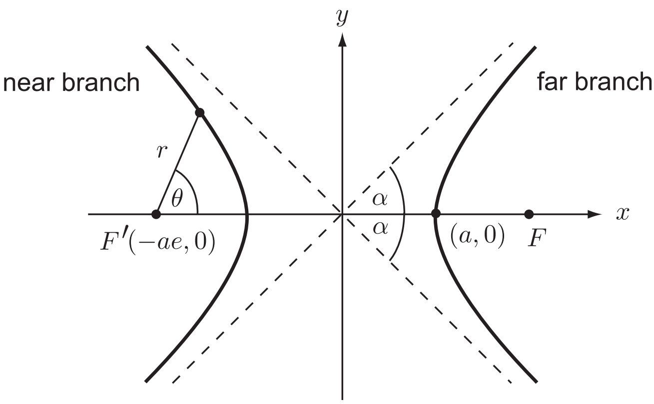


FIGURE 7.15 The standard hyperbola $x^2/a^2 - y^2/b^2 = 1$. The near and far branches are relative to the focus F', which is the origin of polar coordinates.

(ii) The **eccentricity** e of the ellipse is defined by

$$e^2 = 1 - \frac{b^2}{a^2}$$

and lies in the range $0 \leq e < 1$. When $e = 0$, $b = a$ and the ellipse is a circle.

(iii) The **focal points** F, F' of the ellipse lie on the major axis at $(\pm ae, 0)$.

(iv) In **polar coordinates** with origin at the focus F and with initial line in the positive x-direction, the equation of the ellipse is

$$\frac{1}{r} = \frac{a}{b^2}(1 + e\cos\theta).$$

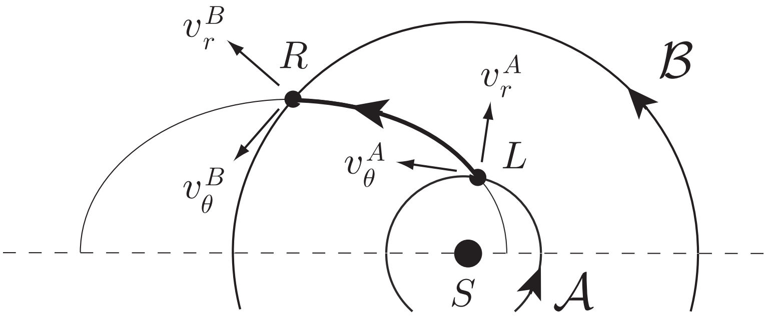


FIGURE 7.16 The circular orbits $\mathcal{A}$ and $\mathcal{B}$ are the orbits of the two planets. The elliptical orbit shown is a possible path for the spacecraft, which travels along the arc LR. The velocities shown are those *after* the first firing at L and *before* the second firing at R.

Hyperbola

(i) In **Cartesian coordinates**, the standard hyperbola has the equation

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \qquad (a, b > 0)$$

so that the angle 2α between the asymptotes is given by

$$\tan\alpha = \frac{b}{a}.$$

(ii) The **eccentricity** e of the hyperbola is defined by

$$e^2 = 1 + \frac{b^2}{a^2}$$

and lies in the range $e > 1$.

(iii) The **focal points** F, F' of the hyperbola lie on the x-axis at $(\pm ae, 0)$.

(iv) In **polar coordinates** with origin at the focus F' and with initial line in the positive x-direction, the equations of the near and far branches of the hyperbola are

$$\frac{1}{r} = \frac{a}{b^2}(1 + e\cos\theta), \qquad \frac{1}{r} = \frac{a}{b^2}(-1 + e\cos\theta),$$

respectively.

Appendix B The Hohmann orbit is optimal

The result that the Hohmann orbit is the connecting orbit that minimises Q is not at all obvious and correct proofs are rare.* Hopefully, the proof given below *is* correct!

* It is sometimes stated that the optimality requirement is to minimise the *energy* of the connecting orbit, which is not true. In any case, the Hohmann orbit is *not* the connecting orbit of minimum energy!

Proof of optimality Consider the general two-impulse transfer orbit LR shown in Figure 7.16, where the orbit is regarded as being generated by the velocity components v_θ^A, v_r^A of the spacecraft after the first impulse. Then, by angular momentum and energy conservation,

$$Av_\theta^A = Bv_\theta^B,$$

$$\left(v_r^A\right)^2 + \left(v_\theta^A\right)^2 - \frac{2\gamma}{A} = \left(v_r^B\right)^2 + \left(v_\theta^B\right)^2 - \frac{2\gamma}{B},$$

where A, B are the radii of the circular orbits of Earth and Jupiter and $\gamma = M_\odot G$. Thus

$$v_\theta^B = \frac{A}{B} v_\theta^A,$$

$$\left(v_r^B\right)^2 = \left(1 - \frac{A^2}{B^2}\right)\left(v_\theta^A\right)^2 + \left(v_r^A\right)^2 - 2\gamma\left(\frac{1}{A} - \frac{1}{B}\right).$$

Since the orbital speeds of Earth and Jupiter are $(\gamma/A)^{1/2}$ and $(\gamma/B)^{1/2}$, it follows that the velocity changes $\Delta \boldsymbol{v}^A$, $\Delta \boldsymbol{v}^B$ required at L and R have magnitudes given by

$$|\Delta \boldsymbol{v}^A|^2 = \left(v_\theta^A - \left(\frac{\gamma}{A}\right)^{1/2}\right)^2 + \left(v_r^A\right)^2,$$

$$\begin{aligned}
|\Delta \boldsymbol{v}^B|^2 &= \left(\left(\frac{\gamma}{B}\right)^{1/2} - v_\theta^B\right)^2 + \left(v_r^B\right)^2 \\
&= \left(\left(\frac{\gamma}{B}\right)^{1/2} - \frac{A}{B} v_\theta^A\right)^2 + \left(1 - \frac{A^2}{B^2}\right)\left(v_\theta^A\right)^2 + \left(v_r^A\right)^2 - 2\gamma\left(\frac{1}{A} - \frac{1}{B}\right) \\
&= \left(v_\theta^A - \frac{\gamma^{1/2} A}{B^{3/2}}\right)^2 + \left(v_r^A\right)^2 + \gamma\left(\frac{3}{B} - \frac{2}{A} - \frac{A^2}{B^3}\right).
\end{aligned}$$

It is evident that, with v_θ^A fixed, both $|\Delta \boldsymbol{v}^A|$ and $|\Delta \boldsymbol{v}^B|$ are *increasing* functions of v_r^A. Thus Q may be reduced by reducing v_r^A provided that the resulting orbit still meets the circle $r = B$. Q can be thus reduced until either

(i) v_r^A is reduced to zero, or
(ii) the orbit shrinks until it touches the circle $r = B$ and any further reduction in v_r^A would mean that the orbit would not meet $r = B$.

In the first case, L becomes the perihelion of the orbit and, in the second case, R becomes the aphelion of the orbit. We will proceed assuming the first case, the second case being treated in a similar manner and with the same result.

Suppose then that L is the perihelion of the connecting orbit. Then $v_r^A = 0$ and, from now on, we will simply write v instead of v_θ^A. The velocity v must be such that the orbit reaches the circle $r = B$, which now means that the major axis of the orbit must not be less than $A + B$. On using the E-formula, this implies that v must satisfy

$$v^2 \geq \frac{2\gamma B}{A(A+B)}.$$

The formulae for $|\Delta \boldsymbol{v}^A|$ and $|\Delta \boldsymbol{v}^B|$ now simplify to

$$|\Delta \boldsymbol{v}^A|^2 = \left(v - \left(\frac{\gamma}{A}\right)^{1/2}\right)^2,$$

$$|\Delta \boldsymbol{v}^B|^2 = \left(v - \frac{\gamma^{1/2} A}{B^{3/2}}\right)^2 + \gamma\left(\frac{3}{B} - \frac{2}{A} - \frac{A^2}{B^3}\right)$$

from which it is evident that, for v in the permitted range, both of $|\Delta \boldsymbol{v}^A|$ and $|\Delta \boldsymbol{v}^B|$ are *increasing* functions v. Hence the minimum value of Q is achieved when v takes its smallest permitted value, namely

$$v = \left(\frac{2\gamma B}{A(A+B)}\right)^{1/2}.$$

With this value of v, the orbit touches the circle $r = B$ and so has its aphelion at R. Hence *the optimum orbit has its perihelion at L and its aphelion at R*. This is precisely the **Hohmann orbit**. ■

Problems on Chapter 7

Answers and comments are at the end of the book.

Harder problems carry a star (∗).

Radial motion equation, apses

7.1 A particle P of mass m moves under the repulsive inverse cube field $\boldsymbol{F} = (m\gamma/r^3)\,\widehat{\boldsymbol{r}}$. Initially P is at a great distance from O and is moving with speed V towards O along a straight line whose perpendicular distance from O is p. Find the equation satisfied by the apsidal distances. What is the distance of closest approach of P to O?

7.2 A particle P of mass m moves under the attractive inverse square field $\boldsymbol{F} = -(m\gamma/r^2)\,\widehat{\boldsymbol{r}}$. Initially P is at a point C, a distance c from O, when it is projected with speed $(\gamma/c)^{1/2}$ in a direction making an acute angle α with the line OC. Find the apsidal distances in the resulting orbit.

Given that the orbit is an ellipse with O at a focus, find the semi-major and semi-minor axes of this ellipse.

7.3 A particle of mass m moves under the attractive inverse square field $\boldsymbol{F} = -(m\gamma/r^2)\,\widehat{\boldsymbol{r}}$. Show that the equation satisfied by the apsidal distances is

$$2Er^2 + 2\gamma r - L^2 = 0,$$

where E and L are the specific total energy and angular momentum of the particle. When $E < 0$, the orbit is known to be an ellipse with O as a focus. By considering the sum and product of the roots of the above equation, establish the elliptic orbit formulae

$$L^2 = \gamma b^2/a, \qquad E = -\gamma/2a.$$

7.4 A particle P of mass m moves under the simple harmonic field $\boldsymbol{F} = -(m\Omega^2 r)\,\widehat{\boldsymbol{r}}$, where Ω is a positive constant. Obtain the radial motion equation and show that all orbits of P are bounded.

Initially P is at a point C, a distance c from O, when it is projected with speed Ωc in a direction making an acute angle α with OC. Find the equation satisfied by the apsidal distances. Given that the orbit of P is an ellipse with centre O, find the semi-major and semi-minor axes of this ellipse.

Path equation

7.5 A particle P moves under the attractive inverse square field $\boldsymbol{F} = -(m\gamma/r^2)\,\widehat{\boldsymbol{r}}$. Initially P is at the point C, a distance c from O, and is projected with speed $(3\gamma/c)^{1/2}$ perpendicular to OC. Find the polar equation of the path make a sketch of it. Deduce the angle between OC and the final direction of departure of P.

7.6 A comet moves under the gravitational attraction of the Sun. Initially the comet is at a great distance from the Sun and is moving towards it with speed V along a straight line whose perpendicular distance from the Sun is p. By using the path equation, find the angle through which the comet is deflected and the distance of closest approach.

7.7 A particle P of mass m moves under the attractive inverse cube field $\boldsymbol{F} = -(m\gamma^2/r^3)\,\widehat{\boldsymbol{r}}$, where γ is a positive constant. Initially P is at a great distance from O and is projected towards O with speed V along a line whose perpendicular distance from O is p. Obtain the path equation for P.

For the case in which

$$V = \frac{15\gamma}{\sqrt{209}\,p},$$

find the polar equation of the path of P and make a sketch of it. Deduce the distance of closest approach to O, and the final direction of departure.

7.8* A particle P of mass m moves under the central field $\boldsymbol{F} = -(m\gamma^2/r^5)\,\widehat{\boldsymbol{r}}$, where γ is a positive constant. Initially P is at a great distance from O and is projected towards O with speed $\sqrt{2}\gamma/p^2$ along a line whose perpendicular distance from O is p. Show that the polar equation of the path of P is given by

$$r = \frac{p}{\sqrt{2}}\coth\left(\frac{\theta}{\sqrt{2}}\right).$$

Make a sketch of the path.

7.9* A particle of mass m moves under the central field

$$\boldsymbol{F} = -m\gamma^2\left(\frac{4}{r^3} + \frac{a^2}{r^5}\right)\widehat{\boldsymbol{r}},$$

where γ and a are positive constants. Initially the particle is at a distance a from the centre of force and is projected at right angles to the radius vector with speed $3\gamma/\sqrt{2}a$. Find the polar equation of the resulting path and make a sketch of it.

Find the time taken for the particle to reach the centre of force.

Nearly circular orbits

7.10 A particle of mass m moves under the central field

$$\boldsymbol{F} = -m\left(\frac{\gamma e^{-\epsilon r/a}}{r^2}\right)\widehat{\boldsymbol{r}},$$

where γ, a and ϵ are positive constants. Find the apsidal angle for a nearly circular orbit of radius a. When ϵ is small, show that the perihelion of the orbit advances by approximately $\pi\epsilon$ on each revolution.

7.11 ***Solar oblateness*** A planet of mass m moves in the equatorial plane of a star that is a uniform oblate spheroid. The planet experiences a force field of the form

$$\boldsymbol{F} = -\frac{m\gamma}{r^2}\left(1 + \frac{\epsilon\, a^2}{r^2}\right)\widehat{\boldsymbol{r}},$$

approximately, where γ, a and ϵ are positive constants and ϵ is small. If the planet moves in a nearly circular orbit of radius a, find an approximation to the 'annual' advance of the perihelion. [It has been suggested that oblateness of the Sun might contribute significantly to the precession of the planets, thus undermining the success of general relativity. This point has yet to be resolved conclusively.]

7.12 Suppose the solar system is embedded in a dust cloud of uniform density ρ. Find an approximation to the 'annual' advance of the perihelion of a planet moving in a nearly circular orbit of radius a. (For convenience, let $\rho = \epsilon M/a^3$, where M is the solar mass and ϵ is small.)

7.13 ***Orbits in general relativity*** In the theory of general relativity, the path equation for a planet moving in the gravitational field of the Sun is, in the standard notation,

$$\frac{d^2u}{d\theta^2} + u = \frac{MG}{L^2} + \left(\frac{3MG}{c^2}\right)u^2,$$

where c is the speed of light. Find an approximation to the 'annual' advance of the perihelion of a planet moving in a nearly circular orbit of radius a.

Scattering

7.14 A uniform flux of particles is incident upon a fixed hard sphere of radius a. The particles that strike the sphere are reflected elastically. Find the differential scattering cross section.

7.15 A uniform flux of particles, each of mass m and speed V, is incident upon a fixed scatterer that exerts the repulsive radial force $\boldsymbol{F} = (m\gamma^2/r^3)\widehat{\boldsymbol{r}}$. Find the impact parameter p as a function of the scattering angle θ, and deduce the differential scattering cross section. Find the total back-scattering cross-section.

Assorted inverse square problems

Some useful **data**:

The radius R of the Earth is 6380 km. To obtain the value of MG, where M is the mass of the Earth, use the formula $MG = R^2 g$, where $g = 9.80\ \mathrm{m\,s^{-2}}$.

1 AU per year is 4.74 km per second. In astronomical units, $G = 4\pi^2$.

7.16 In Yuri Gagarin's first manned space flight in 1961, the perigee and apogee were 181 km and 327 km above the Earth. Find the period of his orbit and his maximum speed in the orbit.

7.17 An Earth satellite has a speed of 8.60 km per second at its perigee 200 km above the Earth's surface. Find the apogee distance above the Earth, its speed at the apogee, and the period of its orbit.

7.18 A spacecraft is orbiting the Earth in a circular orbit of radius c when the motors are fired so as to multiply the speed of the spacecraft by a factor k ($k > 1$), its direction of motion being unaffected. [You may neglect the time taken for this operation.] Find the range of k for which the spacecraft will escape from the Earth, and the eccentricity of the escape orbit.

7.19 A spacecraft travelling with speed V approaches a planet of mass M along a straight line whose perpendicular distance from the centre of the planet is p. When the spacecraft is at a distance c from the planet, it fires its engines so as to multiply its current speed by a factor k ($0 < k < 1$), its direction of motion being unaffected. [You may neglect the time taken for this operation.] Find the condition that the spacecraft should go into orbit around the planet.

7.20 A body moving in an inverse square attractive field traverses an elliptical orbit with major axis $2a$. Show that the time average of the potential energy $V = -\gamma/r$ is $-\gamma/a$. [Transform the time integral to an integral with repect to the eccentric angle ψ.]

Deduce the time average of the kinetic energy in the same orbit.

7.21 A body moving in an inverse square attractive field traverses an elliptical orbit with eccentricity e and major axis $2a$. Show that the time average of the distance r of the body from the centre of force is $a(1 + \frac{1}{2}e^2)$. [Transform the time integral to an integral with respect to the eccentric angle ψ.]

7.22 A spacecraft is 'parked' in a circular orbit 200 km above the Earth's surface. The spacecraft is to be sent to the Moon's orbit by Hohmann transfer. Find the velocity changes Δv^E and Δv^M that are required at the Earth and Moon respectively. How long does the journey take? [The radius of the Moon's orbit is 384,000 km. Neglect the gravitation of the Moon.]

7.23* A spacecraft is 'parked' in an *elliptic* orbit around the Earth. What is the most fuel efficient method of escaping from the Earth by using a single impulse?

7.24 A satellite already in the Earth's heliocentric orbit can fire its engines only once. What is the most fuel efficient method of sending the satellite on a 'flyby' visit to another planet? The satellite can visit either Mars or Venus. Which trip would use less fuel? Which trip would take the shorter time? [The orbits of Mars and Venus have radii 1.524 AU and 0.723 AU respectively.]

7.25 A satellite is 'parked' in a circular orbit 250 km above the Earth's surface. What is the most fuel efficient method of transferring the satellite to an (elliptical) synchronous orbit by using a single impulse? [A synchronous orbit has a period of 23 hr 56 m.] Find the value of $\Delta \boldsymbol{v}$ and apogee distance.

Effect of resistance

7.26 A satellite of mass m moves under the attractive inverse square field $-(m\gamma/r^2)\,\widehat{\boldsymbol{r}}$ and is also subject to the linear resistance force $-mK\,\boldsymbol{v}$, where K is a positive constant. Show that the governing equations of motion can be reduced to the form

$$\ddot{r} + K\dot{r} + \frac{\gamma}{r^2} - \frac{L_0^2\, e^{-2Kt}}{r^3} = 0, \qquad r^2\dot{\theta} = L_0\, e^{-Kt},$$

where L_0 is a constant which will be assumed to be positive.

Suppose now that the effect of resistance is slight and that the satellite is executing a 'circular' orbit of slowly changing radius. By neglecting the terms in $\dot{r}$ and $\ddot{r}$, find an approximate solution for the time variation of r and θ in such an orbit. Deduce that small resistance causes the circular orbit to contract slowly, but that the satellite speeds up!

7.27 Repeat the last problem for the case in which the particle moves under the simple harmonic attractive field $-(m\Omega^2 r)\,\widehat{\boldsymbol{r}}$ with the same law of resistance. Show that, in this case, the body slows down as the orbit contracts. [This problem can be solved exactly in Cartesian coordinates, but do not do it this way.]

Computer assisted problems

7.28 ***See the advance of the perihelion of Mercury*** It is possible to 'see' the advance of the perihelion of Mercury predicted by general relativity by direct numerical solution. Take Einstein's path equation (see Problem 7.13) in the dimensionless form

$$\frac{d^2 \upsilon}{d\theta^2} + \upsilon = \frac{1}{1-e^2} + \eta \upsilon^2,$$

where $\upsilon = au$. Here a and e are the semi-major axis and eccentricity of the non-relativistic elliptic orbit and $\eta = 3MG/ac^2$ is a small dimensionless parameter. For the orbit of Mercury, $\eta = 2.3 \times 10^{-7}$ approximately.

Solve this equation numerically with the initial conditions $r = a(1+e)$ and $\dot{r} = 0$ when $\theta = 0$; this makes $\theta = 0$ an aphelion of the orbit. To make the precession easy to see, use a fairly eccentric ellipse and take η to be about 0.005, which speeds up the precession by a factor of more than 10^4!

7.29 ***Orbit with linear resistance*** Confirm the approximate solution for small resistance obtained in Problem 7.26 by numerical solution of the governing simultaneous ODEs. First write the governing equations in dimensionless form. Suppose that, in the absence of resistance, a circular orbit with $r = a$ and $\dot{\theta} = \Omega$ is possible; then $\gamma = a^3\Omega$ and $L_0 = a^2\Omega$. On taking dimensionless variables ρ, τ defined by $\rho = r/a$ and $\tau = \Omega t$, and taking $L_0 = a^2\Omega$, the governing equations become

$$\frac{d^2\rho}{d\tau^2} + \epsilon\,\frac{d\rho}{d\tau} + \frac{1}{\rho^2} - \frac{e^{-2\epsilon\tau}}{\rho^3} = 0, \qquad \rho^2\frac{d\theta}{d\tau} = e^{-2\epsilon\tau},$$

where $\epsilon = K/\Omega$ is the dimensionless resistance parameter. Solve these equations with the initial conditions $\rho = 1$, $d\rho/d\tau = 0$ and $\theta = 0$ when $\tau = 0$. Choose some small value for ϵ and plot a polar graph of the path.

Chapter Eight

Non-linear oscillations and phase space

KEY FEATURES

The key features of this chapter are the use of **perturbation theory** to solve weakly non-linear problems, the notion of **phase space**, the **Poincaré–Bendixson** theorem, and **limit cycles**.

In reality, most oscillating mechanical systems are governed by **non-linear equations**. The linear oscillation theory developed in Chapter 5 is generally an approximation which is accurate only when the amplitude of the oscillations is small. Unfortunately, non-linear oscillation equations do not have nice exact solutions as their linear counterparts do, and this makes the non-linear theory difficult to investigate analytically.

In this chapter we describe two different analytical approaches, each of which is successful in its own way. The first is to use **perturbation theory** to find successive corrections to the linear theory. This gives a more accurate solution than the linear theory when the non-linear terms in the equation are small. However, because the solution is close to that predicted by the linear theory, new phenomena associated with non-linearity are unlikely to be discovered by perturbation theory! The second approach involves the use of geometrical arguments in **phase space**. This has the advantage that the non-linear effects can be large, but the conclusions are likely to be qualitative rather than quantitative. A particular triumph of this approach is the **Poincaré–Bendixson** theorem, which can be used to prove the existence of **limit cycles**, a new phenomenon that exists only in the non-linear theory.

8.1 PERIODIC NON-LINEAR OSCILLATIONS

Most oscillating mechanical systems are not exactly linear but are approximately linear when the oscillation amplitude is small. In the case of a body on a spring, the restoring force might actually have the form

$$S = m\Omega^2 x + m\Lambda x^3, \tag{8.1}$$

which is approximated by the linear formula $S = m\Omega^2 x$ when the displacement x is small. The new constant Λ is a measure of the strength of the non-linear effect. If $\Lambda < 0$, then

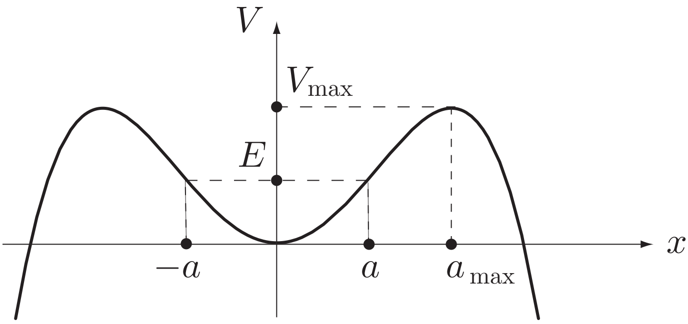


FIGURE 8.1 Existence of periodic oscillations for the quartic potential energy $V = \frac{1}{2}m\Omega^2x^2 + \frac{1}{4}m\Lambda x^4$ with $\Lambda < 0$.

S is less than its linear approximation and the spring is said to be **softening** as x increases. Conversely, if $\Lambda > 0$, then the spring is **hardening** as x increases. The formula (8.1) is typical of non-linear restoring forces that are *symmetrical* about $x = 0$. If the restoring force is unsymmetrical about $x = 0$, the leading correction to the linear case will be a term in x^2.

Existence of non-linear periodic oscillations

Consider the **free undamped oscillations** of a body sliding on a smooth horizontal table* and connected to a fixed point of the table by a spring whose restoring force is given by the cubic formula (8.1). In rectilinear motion, the **governing equation** is then

$$\frac{d^2x}{dt^2} + \Omega^2x + \Lambda x^3 = 0, \tag{8.2}$$

which is Duffing's equation with no forcing term (see section 8.5). The existence of periodic oscillations can be proved by the energy method described in Chapter 6. The restoring force has potential energy

$$V = \tfrac{1}{2}m\Omega^2x^2 + \tfrac{1}{4}m\Lambda x^4,$$

so that the energy conservation equation is

$$\tfrac{1}{2}mv^2 + \tfrac{1}{2}m\Omega^2x^2 + \tfrac{1}{4}m\Lambda x^4 = E,$$

where $v = \dot{x}$. The motion is therefore restricted to those values of x that satisfy

$$\tfrac{1}{2}m\Omega^2x^2 + \tfrac{1}{4}m\Lambda x^4 \le E,$$

* Would the motion be the same (relative to the equilibrium position) if the body were suspended vertically by the same spring?

with equality when $v = 0$. Figure 8.1 shows a sketch of V for a *softening* spring ($\Lambda < 0$). For each value of E in the range $0 < E < V_{\max}$, the particle oscillates in a symmetrical range $-a \le x \le a$ as shown. Thus oscillations of any amplitude less than $a_{\max}$ ($= \Omega/|\Lambda|^{1/2}$) are possible. For a *hardening* spring, oscillations of any amplitude whatsoever are possible.

Solution by perturbation theory

Suppose then that the body is performing periodic oscillations with amplitude a. In order to reduce the number of parameters, we non-dimensionalise equation (8.2). Let the *dimensionless displacement* X be defined by $x = aX$. Then X satisfies the equation

$$\frac{1}{\Omega^2}\frac{d^2X}{dt^2} + X + \epsilon X^3 = 0, \tag{8.3}$$

with the initial conditions $X = 1$ and $dX/dt = 0$ when $t = 0$. The dimensionless parameter ϵ, defined by

$$\epsilon = \frac{a^2\Lambda}{\Omega^2}, \tag{8.4}$$

is a measure of the strength of the non-linearity. Equation (8.3) contains ϵ as a parameter and hence so does the solution. A major feature of interest is how the period τ of the motion varies with ϵ.

The non-linear equation of motion (8.3) cannot be solved explicitly but it reduces to a simple linear equation when the parameter ϵ is zero. In these circumstances, one can often find an *approximate* solution to the non-linear equation *valid when* ϵ *is small*. Equations in which the non-linear terms are small are said to be **weakly non-linear** and the solution technique is called **perturbation theory**. There is a well established theory of such perturbations. The simplest case is as follows:

Regular perturbation expansion

If the parameter ϵ appears as the coefficient of any term of an ODE that is *not* the highest derivative in that equation, then, when ϵ is small, the solution corresponding to fixed initial conditions can be expanded as a **power series** in ϵ.

This is called a **regular perturbation expansion*** and it applies to the equation (8.3). It follows that the solution $X(t, \epsilon)$ can be expanded in the **regular perturbation series**

$$X(t, \epsilon) = X_0(t) + \epsilon X_1(t) + \epsilon^2 X_2(t) + \cdots . \tag{8.5}$$

* The case in which the small parameter multiplies the *highest* derivative in the equation is called a **singular perturbation**. For experts only!

The standard method is to substitute this series into the equation (8.3) and then to try to determine the functions $X_0(t)$, $X_1(t)$, $X_2(t)$, ... In the present case however, this leads to an unsatisfactory result because the functions $X_1(t)$, $X_2(t)$, ... , turn out to be *non-periodic* (and unbounded) even though the exact solution $X(t,\epsilon)$ is periodic!* Also, it is not clear how to find approximations to τ from such a series.

This difficulty can be overcome by replacing t by a new variable s so that the solution $X(s,\epsilon)$ has period 2π in *s whatever the value of ϵ*. Every term of the perturbation series will then also be periodic with period 2π. This trick is known as *Lindstedt's method*.

Lindstedt's method

Let $\omega(\epsilon)$ $(= 2\pi/\tau(\epsilon))$ be the angular frequency of the required solution of equation (8.3). Now introduce a new independent variable s (the *dimensionless time*) by the equation $s = \omega(\epsilon)t$. Then $X(s,\epsilon)$ satisfies the equation

$$\left(\frac{\omega(\epsilon)}{\Omega}\right)^2 X'' + X + \epsilon X^3 = 0 \tag{8.6}$$

with the initial conditions $X = 1$ and $X' = 0$ when $s = 0$. (Here $'$ means d/ds.) We now seek a solution of this equation in the form of the perturbation series

$$X(s,\epsilon) = X_0(s) + \epsilon X_1(s) + \epsilon^2 X_2(s) + \cdots . \tag{8.7}$$

which is possible when ϵ is small. By construction, this solution must have period 2π *for all* ϵ from which it follows that each of the functions $X_0(s)$, $X_1(s)$, $X_2(s)$, ... must also have period 2π. However we have paid a price for this simplification since the *unknown* angular frequency $\omega(\epsilon)$ now appears in the equation (8.6); indeed, the function $\omega(\epsilon)$ is part of the *answer* to this problem! We must therefore also expand $\omega(\epsilon)$ as a perturbation series in ϵ. From equation (8.3), it follows that $\omega(0) = \Omega$ so we may write

$$\frac{\omega(\epsilon)}{\Omega} = 1 + \omega_1\epsilon + \omega_2\epsilon^2 + \cdots , \tag{8.8}$$

where $\omega_1, \omega_2, \ldots$ are unknown constants that must be determined along with the functions $X_0(s)$, $X_1(s)$, $X_2(s)$, ...

On substituting the expansions (8.7) and (8.8) into the governing equation (8.6) and its initial conditions, we obtain:

$$(1 + \omega_1\epsilon + \omega_2\epsilon^2 + \cdots)^2 (X_0'' + \epsilon X_1'' + \epsilon^2 X_2'' + \cdots) +$$
$$(X_0 + \epsilon X_1 + \epsilon^2 X_2 + \cdots) + \epsilon(X_0 + \epsilon X_1 + \epsilon^2 X_2 + \cdots)^3 = 0,$$

* This 'paradox' causes great bafflement when first encountered, but it is inevitable when the period τ of the motion depends on ϵ, as it does in this case. To have a series of non-periodic terms is not *wrong*, as is sometimes stated. However, it is certainly unsatisfactory to have a non-periodic approximation to a periodic function.

with

$$\begin{aligned} X_0 + \epsilon X_1 + \epsilon^2 X_2 + \cdots &= 1, \\ X_0' + \epsilon X_1' + \epsilon^2 X_2' + \cdots &= 0, \end{aligned}$$

when $s = 0$. If we now equate coefficients of powers of ϵ in these equalities, we obtain a succession of ODEs and initial conditions, the first two of which are as follows:

From coefficients of ϵ^0, we obtain the **zero order** equation

$$X_0'' + X_0 = 0, \tag{8.9}$$

with $X_0 = 1$ and $X_0' = 0$ when $s = 0$.
From coefficients of ϵ^1, we obtain the **first order** equation

$$X_1'' + X_1 = -2\omega_1 X_0'' - X_0^3, \tag{8.10}$$

with $X_1 = 0$ and $X_1' = 0$ when $s = 0$.

This procedure can be extended to any number of terms but the equations rapidly become very complicated. The method now is to solve these equations in order; the only sticking point is how to determine the unknown constants $\omega_1, \omega_2, \ldots$ that appear on the right sides of the equations. The solution of the **zero order** equation and initial conditions is

$$X_0 = \cos s \tag{8.11}$$

and this can now be substituted into the first order equation (8.10) to give

$$\begin{aligned} X_1'' + X_1 &= 2\omega_1 \cos s - \cos^3 s \\ &= \tfrac{1}{4}\left(8\omega_1 - 3\right)\cos s + \tfrac{1}{4}\cos 3s, \end{aligned} \tag{8.12}$$

on using the trigonometric identity $\cos 3s = 4\cos^3 s - 3\cos s$. This equation can now be solved by standard methods. The particular integral corresponding to the $\cos 3s$ on the right is $-(1/8)\cos 3s$, but the particular integral corrsponding to the $\cos s$ on the right is $(1/2)s \sin s$, since $\cos s$ is a solution of the equation $X'' + X = 0$. The general solution of the first order equation is therefore

$$X_1 = \left(\omega_1 - \tfrac{3}{8}\right) s \sin s - \tfrac{1}{32}\cos 3s + A\cos s + B \sin s,$$

where A and B are arbitrary constants. Observe that the functions $\cos s$, $\sin s$ and $\cos 3s$ are all periodic with period 2π, but the term $s \sin s$ is *not periodic*. Thus, *the coefficient of $s \sin s$ must be zero, for otherwise $X_1(s)$ would not be periodic, which we know it must be*. Hence

$$\omega_1 = \frac{3}{8}, \tag{8.13}$$

which determines the first unknown coefficient in the expansion (8.8) of $\omega(\epsilon)$. The solution of the **first order** equation and initial conditions is then

$$X_1 = \frac{1}{32}\left(\cos s - \cos 3s\right). \tag{8.14}$$

We have thus shown that, when ϵ is small,

$$\frac{\omega}{\Omega} = 1 + \frac{3}{8}\epsilon + O\left(\epsilon^2\right),$$

and

$$X = \cos s + \frac{\epsilon}{32}\left(\cos s - \cos 3s\right) + O\left(\epsilon^2\right),$$

where $s = \left(1 + \frac{3}{8}\epsilon + O\left(\epsilon^2\right)\right)\Omega t$.

Results

When ϵ ($= a^2\Lambda/\Omega^2$) is small, the **period** τ of the oscillation of equation (8.2) with amplitude a is given by

$$\tau = \frac{2\pi}{\omega} = \frac{2\pi}{\Omega}\left(1 + \frac{3}{8}\epsilon + O\left(\epsilon^2\right)\right)^{-1} = \frac{2\pi}{\Omega}\left(1 - \frac{3}{8}\epsilon + O\left(\epsilon^2\right)\right) \tag{8.15}$$

and the corresponding displacement $x(t)$ is given by

$$x = a\left[\cos s + \frac{\epsilon}{32}\left(\cos s - \cos 3s\right) + O\left(\epsilon^2\right)\right], \tag{8.16}$$

where $s = \left(1 + \frac{3}{8}\epsilon + O\left(\epsilon^2\right)\right)\Omega t$.

This is the *approximate solution correct to the first order in the small parameter* ϵ. More terms can be obtained in a similar way but the effort needed increases exponentially and this is best done with computer assistance (see Problem 8.15).

These formulae apply only when ϵ is small, that is, when the *non- linearity in the equation has a small effect*. Thus we have laboured through a sizeable chunk of mathematics to produce an answer that is only slightly different from the linear case. This sad fact is true of all *regular* perturbation problems. However, in non-linear mechanics, one must be thankful for even modest successes.

8.2 THE PHASE PLANE ((x_1, x_2)–plane)

The second approach that we will describe could not be more different from perturbation theory. It makes use of qualitative geometrical arguments in the phase space of the system.

Systems of first order ODEs

The notion of **phase space** springs from the theory of **systems of first order ODEs**. Such systems are very common and need have no connection with classical mechanics. A standard example is the predator-prey system of equations

$$\dot{x}_1 = ax_1 - bx_1x_2,$$
$$\dot{x}_2 = bx_1x_2 - cx_2,$$

which govern the population density $x_1(t)$ of a prey and the population density $x_2(t)$ of its predator. In the general case there are n unknown functions satisfying n first order ODEs, but here we will only make use of *two* unknown functions $x_1(t)$, $x_2(t)$ that satisfy a *pair* of first order ODEs of the form

$$\begin{aligned} \dot{x}_1 &= F_1(x_1, x_2, t), \\ \dot{x}_2 &= F_2(x_1, x_2, t). \end{aligned} \tag{8.17}$$

Just to confuse matters, a system of ODEs like (8.17) is called a **dynamical system**, whether it has any connection with classical mechanics or not! In the predator-prey dynamical system, the function $F_1 = ax_1 - bx_1x_2$ and the function $F_2 = bx_1x_2 - cx_2$. In this case F_1 and F_2 have no explicit time dependence. Such systems are said to be *autonomous*; as we shall see, more can be said about the behaviour of autonomous systems.

Definition 8.1 ***Autonomous system*** *A system of equations of the form*

$$\begin{aligned} \dot{x}_1 &= F_1(x_1, x_2), \\ \dot{x}_2 &= F_2(x_1, x_2), \end{aligned} \tag{8.18}$$

is said to be ***autonomous***.

The phase plane

The values of the variables x_1, x_2 at any instant can be represented by a point in the (x_1, x_2)-plane. This plane is called the **phase plane*** of the system. A solution of the system of equations (8.17) is then represented by a point moving in the phase plane. The path traced out by such a point is called a **phase path**† of the system and the set of all phase paths is called the **phase diagram**. In the predator–prey problem, the variables x_1, x_2 are positive quantities and so the physically relevant phase paths lie in the first quadrant of the phase plane. It can be shown that they are all closed curves! (See Problem 8.10).

Phase paths of autonomous systems

The problem of finding the phase paths is much easier when the system is **autonomous**. The method is as follows:

---

* In the general case with n unknowns, the phase space is n-dimensional.

† Also called an *orbit* of the system.

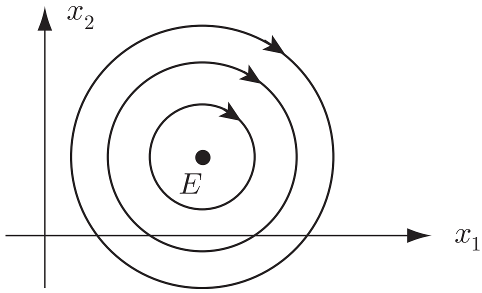


FIGURE 8.2 Phase diagram for the system $dx_1/dt = x_2 - 1,\ dx_2/dt = -x_1 + 2$. The point $E(2, 1)$ is an equilibrium point of the system.

Example 8.1 *Finding phase paths for an autonomous system*

Sketch the phase diagram for the autonomous system of equations

$$\frac{dx_1}{dt} = x_2 - 1,$$

$$\frac{dx_2}{dt} = -x_1 + 2.$$

Solution

The phase paths of an *autonomous* system can be found by eliminating the time derivatives. The path gradient is given by

$$\frac{dx_2}{dx_1} = \frac{dx_2/dt}{dx_1/dt} = -\frac{x_1 - 2}{x_2 - 1}$$

and this is a first order separable ODE satisfied by the phase paths. The general solution of this equation is

$$(x_1 - 2)^2 + (x_2 - 1)^2 = C$$

and each (positive) choice for the constant of integration C corresponds to a phase path. The **phase paths** are therefore circles with centre (2, 1); the **phase diagram** is shown in Figure 8.2.

The direction in which the phase point progresses along a path can be deduced by examining the *signs* of the right sides in equations (8.18). This gives the signs of $\dot{x}_1$ and $\dot{x}_2$ and hence the direction of motion of the phase point. ■

When the system is autonomous, one can say quite a lot about the *general nature* of the phase paths without finding them. The basic result is as follows:

Theorem 8.1 ***Autonomous systems: a basic result*** *Each point of the phase space of an autonomous system has exactly one phase path passing through it.*

Proof. Let (a, b) be any point of the phase space. Suppose that the motion of the phase point (x_1, x_2) satisfies the equations (8.18) and that the phase point is at (a, b) when $t = 0$. The general theory of ODEs then tells us that a solution of the equations (8.18), that satisfies the initial conditions $x_1 = a$, $x_2 = b$ when $t = 0$, exists and is unique. Let this solution be $\{X_1(t), X_2(t)\}$, which we will suppose is defined for all t, both positive and negative. This phase path certainly passes through the point (a, b) and we must now show that there is no other. Suppose then that there is another solution of the equations in which the phase point is at (a, b) when $t = \tau$, say. This motion also exists and is uniquely determined and, in the general case, would not be related to $\{X_1(t), X_2(t)\}$. However, for autonomous systems, the right sides of equations (8.18) are independent of t so that *the two motions differ only by a shift in the origin of time*. To be precise, the new motion is simply $\{X_1(t-\tau), X_2(t-\tau)\}$. Thus, although the two motions are distinct, the two phase points travel along the *same path* with the second point delayed relative to the first by the constant time τ. Hence, although there are infinitely many *motions* of the phase point that pass through the point (a, b), they all follow the same path. This proves the theorem. ■

Some important deductions follow from this basic result.

Phase paths of autonomous systems

- Distinct phase paths of an autonomous system **do not cross** or touch each other.
- **Periodic motions** of an autonomous system correspond to phase paths that are simple* **closed loops**.

Figure 8.2 showsthe phase paths of an autonomous system. For this system, *all* of the phase paths are simple closed loops and so every motion is periodic. An exception occurs if the phase point is started from the point (2, 1). In this case the system has the constant solution $x_1 = 2$, $x_2 = 1$ so that the phase point never moves; for this reason, the point (2, 1) is called an **equilibrium point** of the system. In this case, the 'path' of the phase point consists of the *single point* (2, 1). However, this still qualifies as a path and the above theory still applies. Consequently no 'real' path may pass through an equilibrium point of an autonomous system.†

8.3 THE PHASE PLANE IN DYNAMICS ((x, v)–plane)

The above theory seems unconnected to classical mechanics since dynamical equations of motion are *second order* ODEs. However, *any second order ODE can be expressed as a pair of first order ODEs*. For example, consider the general linear oscillator equation

$$\frac{d^2x}{dt^2} + 2K\frac{dx}{dt} + \Omega^2 x = F(t). \tag{8.19}$$

If we introduce the new variable $v = dx/dt$, then

$$\frac{dv}{dt} + 2kv + \Omega^2 x = F(t).$$

* A simple curve is one that does not cross (or touch) itself (except possibly to close).

† It may appear from diagrams that phase paths *can* pass through equilibrium points. This is not so. Such a path approaches *arbitrarily close* to the equilibrium point in question, but never reaches it!

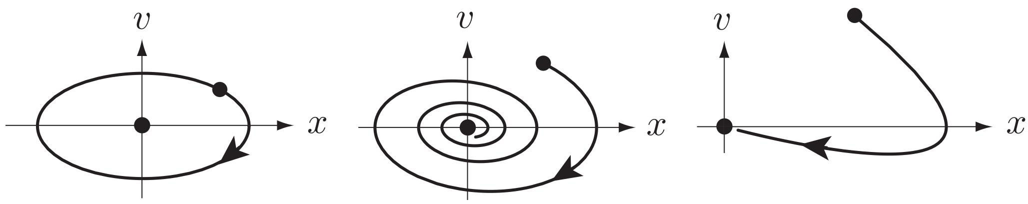


FIGURE 8.3 Typical phase paths for the simple harmonic oscillator equation. **Left**: No damping. **Centre**: Sub-critical damping. **Right**: Super-critical damping.

It follows that the second order equation (8.19) is equivalent to the pair of first order equations

$$\frac{dx}{dt} = v,$$
$$\frac{dv}{dt} = F(t) - 2kv - \Omega^2 x.$$

We may now apply the theory we have developed to this system of first order ODEs, where the **phase plane** is now the (x, v)-plane. It is clear that **driven motion** leads to a **non-autonomous system** because of the presence of the explicit time dependence of $F(t)$; **undriven motion** (in which $F(t) = 0$) leads to an **autonomous** system. It is also clear that equilibrium points in the (x, v)-plane lie on the x-axis and correspond to the ordinary equilibrium positions of the particle.

The form of the phase paths for the *undriven* SHO equation

$$\frac{d^2x}{dt^2} + 2K\frac{dx}{dt} + \Omega^2 x = 0$$

depends on the parameters K and Ω. We could find these paths by the method used in Example 8.1, but there is no point in doing so since we have already solved the equation explicitly in Chapter 5. For instance, when $K = 0$, the general solution is given by

$$x = C\cos(\Omega t - \gamma),$$

from which it follows that

$$v = \frac{dx}{dt} = -C\Omega\sin(\Omega t - \gamma).$$

The phase paths in the (x, v)-plane are therefore similar ellipses centred on the origin, which is an equilibrium point. This, and two typical cases of damped motion, are shown in Figure 8.3. In the presence of damping, the phase point *tends* to the equilibrium point at the origin as $t \to \infty$. Although the equilibrium point is never actually reached, it is convenient to say that these paths 'terminate' at the origin.

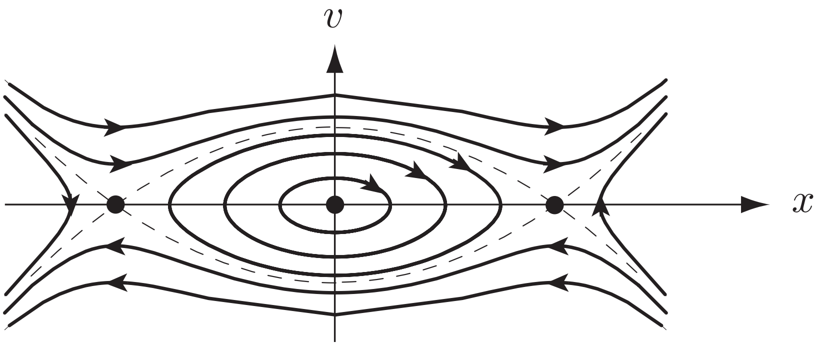


FIGURE 8.4 The phase diagram for the undamped Duffing equation with a softening spring.

Example 8.2 *Phase diagram for equation* $d^2x/dt^2 + \Omega^2 x + \Lambda x^3 = 0$

Sketch the phase diagram for the non-linear oscillation equation

$$d^2x/dt^2 + \Omega^2 x + \Lambda x^3 = 0,$$

when $\Lambda < 0$ (the softening spring).

Solution

This equation is equivalent to the pair of first order equations

$$\frac{dx}{dt} = v,$$
$$\frac{dv}{dt} = -\Omega^2 x - \Lambda x^3,$$

which is an **autonomous** system. The **phase paths** satisfy the equation

$$\frac{dv}{dx} = -\frac{\Omega^2 x + \Lambda x^3}{v},$$

which is a first order separable ODE whose general solution is

$$v^2 = C - \Omega^2 x^2 - \tfrac{1}{2}\Lambda x^4,$$

where C is a constant of integration. Each *positive* value of C corresponds to a phase path. The phase diagram for the case $\Lambda < 0$ is shown in Figure 8.4. There are three **equilibrium points** at $(0, 0)$, $(\pm\Omega/|\Lambda|^{1/2}, 0)$. The closed loops around the origin correspond to **periodic oscillations** of the particle about $x = 0$. Such oscillations can therefore exist for any amplitude less than $\Omega/|\Lambda|^{1/2}$; this confirms the prediction of the energy argument used earlier. Outside this region of closed loops, the paths are unbounded and correspond to unbounded motions of the particle. These two regions of differing behaviour are separated by the dashed paths (known as *separatrices*) that 'terminate' at the equilibrium points $(\pm\Omega/|\Lambda|^{1/2}, 0)$. ■

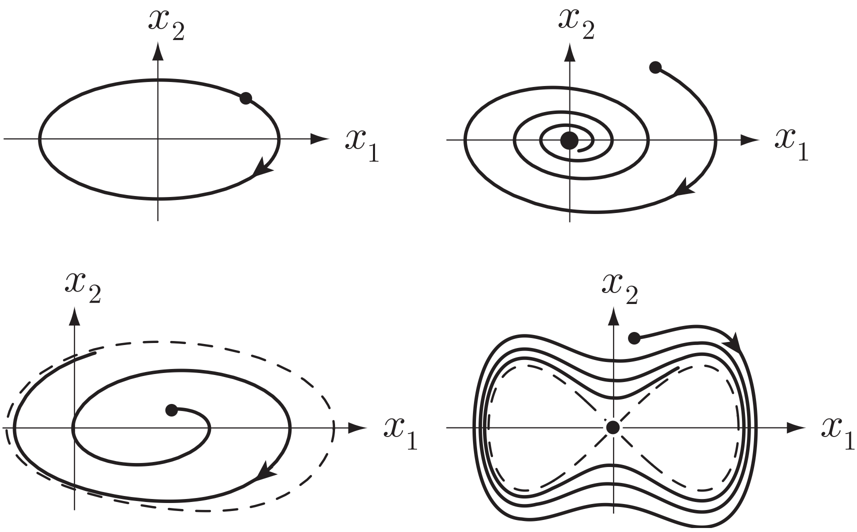


FIGURE 8.5 The Poincaré–Bendixson theorem. Any *bounded* phase path of a plane autonomous system must either close itself (**top left**), terminate at an equilibrium point (**top right**), or tend to a limit cycle (normal case **bottom left**, degenerate case **bottom right**).

8.4 POINCARÉ–BENDIXSON THEOREM: LIMIT CYCLES

In the autonomous systems we have studied so far, those phase paths that are *bounded* either (i) form a closed loop (corresponding to periodic motion), or (ii) 'terminate' at an equilibrium point (so that the motion dies away). Figure 8.3 shows examples of this. The famous Poincaré–Bendixson theorem* which is stated below, says that there is just *one* further possibility.

> **Poincaré–Bendixson theorem**
>
> Suppose that a phase path of a **plane autonomous system** lies in a **bounded** domain of the phase plane for $t > 0$. Then the path must either
>
> - **close** itself, or
> - **terminate** at an equilibrium point as $t \to \infty$, or
> - tend to a **limit cycle** (or a degenerate limit cycle) as $t \to \infty$.

A proper proof of the theorem is long and difficult (see Coddington & Levinson [9]).

* After Jules Henri Poincaré (1854–1912) and Ivar Otto Bendixson (1861–1935). The theorem was first proved by Poincaré but a more rigorous proof was given later by Bendixson.

The third possibility is new and needs explanation. A **limit cycle** is a **periodic motion** of a special kind. It is *isolated* in the sense that nearby phase paths are *not* closed but are attracted towards the limit cycle* ; they spiral around it (or inside it) getting ever closer, as shown in Figure 8.5 (bottom left). The *degenerate* limit cycle shown in Figure 8.5 (bottom right) is an obscure case in which the limiting curve is not a periodic motion but has one or more equilibrium points actually on it. This case is often omitted in the literature, but it definitely exists!

Proving the existence of periodic solutions

The Poincaré–Bendixson theorem provides a way of *proving* that a plane autonomous system has a periodic solution even when that solution cannot be found explicitly. If a phase path can be found that cannot escape from some bounded domain $\mathcal{D}$ of the phase plane, and if $\mathcal{D}$ contains no equilibrium points, then Poincaré–Bendixson implies that the phase path must either be a **closed loop** or tend to a **limit cycle**. In either case, the system must have a **periodic solution** lying in $\mathcal{D}$. The method is illustrated by the following examples.

Example 8.3 *Proving existence of a limit cycle*

Prove that the autonomous system of ODEs

$$\dot{x} = x - y - (x^2 + y^2)x,$$
$$\dot{y} = x + y - (x^2 + y^2)y,$$

has a limit cycle.

Solution

This system clearly has an equilibrium point at the origin $x = y = 0$, and a little algebra shows that there are no others. Although we have not proved this result, it is true that any periodic solution (simple closed loop) in the phase plane must have an equilibrium point lying *inside* it. In the present case, it follows that, if a periodic solution exists, then it *must* enclose the origin. This suggests taking the domain $\mathcal{D}$ to be the annular region between two circles centred on the origin.

It is convenient to express the system of equations in **polar coordinates** r, θ. The transformed equations are (see Problem 8.5)

$$\dot{r} = \frac{x_1\dot{x}_1 + x_2\dot{x}_2}{r}, \qquad \dot{\theta} = \frac{x_1\dot{x}_2 - x_2\dot{x}_1}{r^2},$$

where $x_1 = r\cos\theta$ and $x_2 = r\sin\theta$. In the present case, the polar equations take the simple form

$$\dot{r} = r(1 - r^2), \qquad \dot{\theta} = 1.$$

* This actually describes a *stable* limit cycle, which is the only kind likely to be observed.

These equations can actually be solved explicitly, but, in order to illustrate the method, we will make no use of this fact. Let $\mathcal{D}$ be the annular domain $a < r < b$, where $0 < a < 1$ and $b > 1$. On the circle $r = b$, $\dot{r} = b(1 - b^2) < 0$. Thus a phase point that starts anywhere on the outer boundary $r = b$ *enters* the domain $\mathcal{D}$. Similarly, on the circle $r = a$, $\dot{r} = a(1 - a^2) > 0$ and so a phase point that starts anywhere on the inner boundary $r = a$ also *enters* the domain $\mathcal{D}$. It follows that *any phase path that starts in the annular domain $\mathcal{D}$ can never leave*. Since $\mathcal{D}$ is a *bounded* domain with *no equilibrium points* within it or on its boundaries, it follows from Poincaré–Bendixson that any such path must either be a simple closed loop or tend to a limit cycle. In either case, the system must have a **periodic solution** lying in the annulus $a < r < b$.

We can say more. Phase paths that begin on either *boundary* of $\mathcal{D}$ enter $\mathcal{D}$ and can never leave. These phase paths cannot close themselves (that would mean leaving $\mathcal{D}$) and so can only tend to a limit cycle. It follows that the system must have (at least one) **limit cycle** lying in the domain $\mathcal{D}$. [The explicit solution shows that the circle $r = 1$ is a limit cycle and that there are no other periodic solutions.] ■

Not all examples are as straightforward as the last one. Often, considerable ingenuity has to be used to find a suitable domain $\mathcal{D}$. In particular, the boundary of $\mathcal{D}$ cannot always be composed of circles. Most readers will find our second example rather difficult!

Example 8.4 *Rayleigh's equation has a limit cycle*

Show that **Rayleigh's equation**

$$\ddot{x} + \epsilon \dot{x} \left(\dot{x}^2 - 1 \right) + x = 0,$$

has a limit cycle for any *positive* value of the parameter ϵ.

Solution

Rayleigh's equation arose in his theory of the bowing of a violin string. In the context of particle oscillations however, it corresponds to a simple harmonic oscillator with a strange damping term. When $|\dot{x}| > 1$, we have ordinary (positive) damping and the motion decays. However, when $|\dot{x}| < 1$, we have *negative damping* and the motion grows. The possibility arises then of a periodic motion which is positively damped on some parts of its cycle and negatively damped on others. Somewhat surprisingly, this actually exists.

Rayleigh's equation is equivalent to the autonomous system of ODEs

$$\begin{aligned} \dot{x} &= v, \\ \dot{v} &= -x - \epsilon v \left(v^2 - 1 \right), \end{aligned} \tag{8.20}$$

for which the only equilibrium position is at $x = v = 0$. It follows that, if there is a periodic solution, then it must enclose the origin. At first, we proceed as in the first example. In polar form, the equations (8.20) become

$$\begin{aligned} \dot{r} &= \quad -\epsilon r \sin^2\theta \left(r^2 \sin^2\theta - 1 \right), \\ \dot{\theta} &= -1 - \epsilon \sin^2\theta \left(r^2 \sin^2\theta - 1 \right). \end{aligned} \tag{8.21}$$

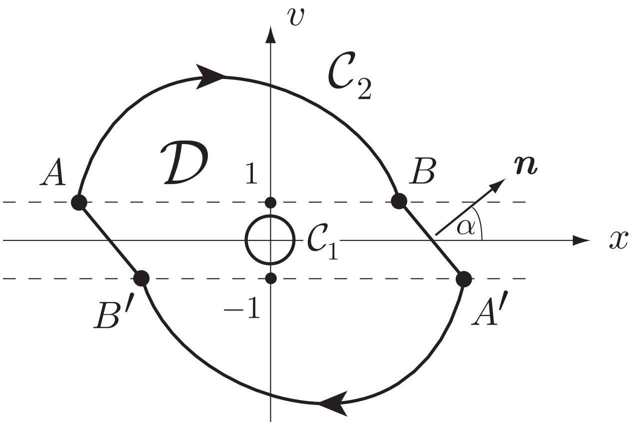


FIGURE 8.6 A suitable domain $\mathcal{D}$ to show that Rayleigh's equation has a limit cycle.

Let $r = c$ be a circle with centre at the origin and radius less than unity. Then $\dot{r} > 0$ everywhere on $r = c$ except at the two points $x = \pm c$, $v = 0$, where it is zero. Hence, except for these two points, we can deduce that a phase point that starts on the circle $r = c$ *enters* the domain $r > c$. Fortunately, these exceptional points can be disregarded. It does not matter if there are a *finite* number of points on $r = c$ where the phase paths go the 'wrong' way, since this provides only a *finite* number of escape routes! The circle $r = c$ thus provides a suitable *inner* boundary $\mathcal{C}_1$ of the domain $\mathcal{D}$.

Sadly, one cannot simply take a large circle to be the outer boundary of $\mathcal{D}$ since $\dot{r}$ has the wrong sign on those segments of the circle that lie in the strip $-1 < v < 1$. This allows any number of phase paths to escape and so invalidates our argument. However, this does not prevent us from choosing a boundary of a different shape. A suitable outer boundary for $\mathcal{D}$ is the contour $\mathcal{C}_2$ shown in Figure 8.6. This contour is made up from four segments. The first segment AB is part of an *actual phase path* of the system which starts at $A(-a, 1)$ and continues as far as $B(b, 1)$. The form of this phase path can be deduced from equations (8.21). When $v(= r\sin\theta) > 1$, $\dot{r} < 0$ and $\dot{\theta} < -1$, so that the phase point moves *clockwise* around the origin with r decreasing. In particular, B must be closer to the origin than A so that $b < a$, as shown. Similarly, the segment $A'B'$ is part of a second actual phase path that begins at $A'(a, -1)$. Because of the symmetry of the equations (8.20) under the transformation $x \to -x$, $v \to -v$, this segment is just the reflection of the segment AB in the origin; the point B' is therefore $(-b, -1)$. The contour is closed by inserting the straight line segments BA' and $B'A$.

We will now show that, when $\mathcal{C}_2$ is made sufficiently large, it is a suitable outer boundary for our domain $\mathcal{D}$. Consider first the segment AB. Since this *is* a phase path, no other phase path may cross it (in either direction); the same applies to the segment $A'B'$. Now consider the straight segment BA'. Because $a > b$, the outward unit normal $\boldsymbol{n}$ shown in Figure 8.6 makes a *positive* acute angle α with the axis Ox. Now the 'phase plane velocity' of a phase point is

$$\dot{x}\boldsymbol{i} + \dot{v}\boldsymbol{j} = v\boldsymbol{i} - \Big(\epsilon v(v^2 - 1) + x\Big)\boldsymbol{j}$$

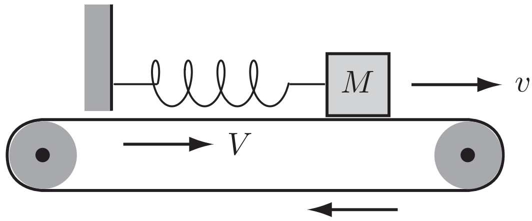


FIGURE 8.7 The body is supported by a rough moving belt and is attached to a fixed post by a light spring.

and the component of this 'velocity' in the $\boldsymbol{n}$-direction is therefore

$$\left(v\boldsymbol{i} - \left(\epsilon v(v^2 - 1) + x\right)\boldsymbol{j}\right) \cdot (\cos\alpha\,\boldsymbol{i} + \sin\alpha\,\boldsymbol{j})$$

$$= v\cos\alpha - \sin\alpha\left(\epsilon v(v^2 - 1) + x\right)$$
$$= -x\sin\alpha + v\left(\cos\alpha + \epsilon\sin\alpha(1 - v^2)\right)$$
$$< -b\sin\alpha + (1 + \epsilon),$$

for (x, v) on BA'. We wish to say that this expression is negative so that phase points that begin on BA' *enter* the domain $\mathcal{D}$. This is true if the contour $\mathcal{C}_2$ is made *large* enough. If we let a tend to infinity, then b also tends to infinity and α tends to $\pi/2$. It follows that, whatever the value of the parameter ϵ, we can make $b\sin\alpha > (1+\epsilon)$ by taking a large enough. A similar argument applies to the segment $B'A$. Thus the contour $\mathcal{C}_2$ is a suitable *outer* boundary for the domain $\mathcal{D}$. It follows that *any phase path that starts in the domain* $\mathcal{D}$ *enclosed by* $\mathcal{C}_1$ *and* $\mathcal{C}_2$ *can never leave*. Since $\mathcal{D}$ is a *bounded* domain with *no equilibrium points* within it or on its boundaries, it follows from Poincaré–Bendixson that any such path must either be a simple closed loop or tend to a limit cycle. In either case, Rayleigh's equation must have a **periodic solution** lying in $\mathcal{D}$.

We can say more. Phase paths that begin on either of the *straight* segments of the outer boundary $\mathcal{C}_2$ enter $\mathcal{D}$ and can never leave. These phase paths cannot close themselves (that would mean leaving $\mathcal{D}$) and so can only tend to a limit cycle. It follows that Rayleigh's equation must have (at least one) **limit cycle** lying in the domain $\mathcal{D}$. [There is in fact only one.] ■

A realistic mechanical system with a limit cycle

Finding *realistic* mechanical systems that exhibit limit cycles is not easy. Driven oscillations are eliminated by the requirement that the system be autonomous. Undamped oscillators have bounded *periodic* motions, and the introduction of damping causes the motions to die away to zero, not to a limit cycle. In order to keep the motion going, the system needs to be *negatively damped* for part of the time. This is an unphysical requirement, but it can be simulated in a physically realistic system as follows.

Consider the system shown in Figure 8.7. A block of mass M is supported by a rough horizontal belt and is attached to a fixed post by a light linear spring. The belt is made to move with constant speed V. Suppose that the motion takes place in a straight line and that $x(t)$ is the extension of the spring beyond its natural length at time t. Then the

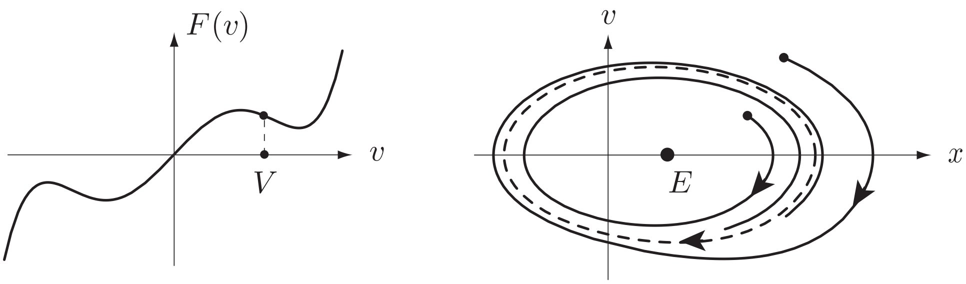


FIGURE 8.8 **Left**: The form of the frictional resistance function $G(v)$. **Right**: The limit cycle in the phase plane; E is the unstable equilibrium point.

equation of motion of the block is

$$M\frac{dv}{dt} = -M\Omega^2 x - F(v - V),$$

where $v = dx/dt$, $M\Omega^2$ is the spring constant, and $F(v)$ is the frictional force that the belt *would* exert on the block if the block had velocity v and the belt were *at rest*; in the actual situation, the argument v is replaced by the relative velocity $v - V$. The function $F(v)$ is supposed to have the form shown in Figure 8.8 (left). Although this choice is unusual ($F(v)$ is not an increasing function of v for all v), it is *not* unphysical!

Under the above conditions, the block has an equilibrium position at $x = F(V)/(M\Omega^2)$. The linearised equation for small motions near this equlibrium position is given by

$$M\frac{d^2x'}{dt^2} = -M\Omega^2 x' - F'(V)\frac{dx'}{dt},$$

where x' is the displacement of the block from the equilibrium position. If we select the belt velocity V so that $F'(V)$ is negative (as shown in Figure 8.8 (left)), then the *effective* damping is negative and small motions will grow. The equilibrium position is therefore *unstable*; oscillations of the block about the equlibrium position then do not die out, but instead tend to a **limit cycle**. This limit cycle is shown in Figure 8.8 (right). The formal proof that such a limit cycle exists is similar to that for Rayleigh's equation. Indeed, this system is essentially Rayleigh's model for the bowing of a violin string, where the belt is the bow, and the block is the string.

Chaotic motions

Another important conclusion from Poincaré–Bendixson is that *no bounded motion of a plane autonomous system can exhibit chaos*. The phase point cannot just wander about in a bounded region of the phase plane for ever. It must either close itself, terminate at an equilibrium point, or tend to a limit cycle and none of these motions is chaotic. In particular, no bounded motion of an undriven non-linear oscillator can be chaotic. As we will see in the next section however, the *driven* non-linear oscillator (a non-autonomous system) *can* exhibit bounded chaotic motions.

It should be remembered that Poincaré–Bendixson applies only to the bounded motion of *plane* autonomous systems. If the phase space has dimension three or more, then other motions, including chaos, are possible.

8.5 DRIVEN NON-LINEAR OSCILLATIONS

Suppose that we now introduce damping and a **harmonic driving force** into equation (8.2). This gives

$$\frac{d^2x}{dt^2} + k\frac{dx}{dt} + \Omega^2 x + \Lambda x^3 = F_0 \cos pt, \tag{8.22}$$

which is known as **Duffing's equation**.

The presence of the driving force $F_0 \cos pt$ makes this system *non-autonomous*. The behaviour of non-autonomous systems is considerably more complex than that of autonomous systems. Phase space is still a useful aid in *depicting* the motion of the system, but little can be said about the general behaviour of the phase paths. In particular, phase paths can cross each other any number of times, and Poincaré–Bendixson does not apply. Our treatment of driven non-linear oscillations is therefore restricted to perturbation theory.

In view of the large number of parameters, it is sensible to non-dimensionalise equation (8.22). The *dimensionless displacement* X is defined by $x = (F_0/p^2)X$ and the *dimensionless time* s by $s = pt$. The function $X(s)$ then satisfies the dimensionless equation

$$X'' + \left(\frac{k}{p}\right) X' + \left(\frac{\Omega}{p}\right)^2 X + \epsilon X^3 = \cos s, \tag{8.23}$$

where the dimensionless parameter ϵ is defined by

$$\epsilon = \frac{F_0^2 \Lambda}{\Omega^6}. \tag{8.24}$$

When $\epsilon = 0$, equation (8.23) reduces to the linear problem. This suggests that, when ϵ is small, we may be able to find approximate solutions by perturbation theory. The linear problem always has a periodic solution for X (the driven motion) that is harmonic with period 2π. Proving the existence of **periodic solutions** of Duffing's equation is an interesting and difficult problem. Here we address this problem for the case in which ϵ is small, a regular perturbation on the linear problem. To simplify the working we will suppose that damping is absent; the general features of the solution remain the same. The governing equation (8.23) then simplifies to

$$X'' + \left(\frac{\Omega}{p}\right)^2 X + \epsilon X^3 = \cos s. \tag{8.25}$$

Initial conditions do not come into this problem. We are simply seeking a family of solutions $X(s, \epsilon)$, parametrised by ϵ, that are (i) periodic, and (ii) reduce to the linear solution when $\epsilon = 0$. We need to consider first the **periodicity** of this family of solutions. In the non-linear problem, we have no right to suppose that the angular frequency of the driven motion is equal to that of the driving force, as it is in the linear problem; it could depend on ϵ. However, suppose that the driving force has minimum period τ_0 and that a family of solutions $X(s, \epsilon))$ of equation (8.25) exists with minimum period $\tau\ (= \tau(\epsilon))$. Then, since the derivatives and powers of X also have period τ, it follows that the left side of equation (8.25) must have period τ. The right side however has period τ_0 and this is known to be the minimum period. It follows that τ *must be an integer multiple of* τ_0; note that τ is not compelled to be *equal* to τ_0.* However, in the present case, the period $\tau(\epsilon)$ is supposed to be a *continuous* function of ϵ with $\tau = \tau_0$ when $\epsilon = 0$. It follows that the only possibility is that $\tau = \tau_0$ for all ϵ. Thus *the period of the driven motion is independent of* ϵ *and is equal to the period of the driving force*. This argument leaves open the possibility that other driven motions may exist that have periods that are integer multiples of τ_0. However, even if they exist, they cannot occur in our perturbation scheme.

We therefore expand $X(s, \epsilon)$ in the **perturbation series**

$$X(t, \epsilon) = X_0(t) + \epsilon X_1(t) + \epsilon^2 X_2(t) + \cdots , \tag{8.26}$$

and seek a solution of equation (8.25) that has period 2π. It follows that the expansion functions $X_0(s)$, $X_1(s)$, $X_2(s)$, ... must also have period 2π. If we now substitute this series into the equation (8.25) and equate coefficients of powers of ϵ, we obtain a succession of ODEs the first two of which are as follows:

From coefficients of ϵ^0:

$$X_0'' + \left(\frac{\Omega}{p}\right)^2 X_0 = \cos s. \tag{8.27}$$

From coefficients of ϵ^1:

$$X_1'' + \left(\frac{\Omega}{p}\right)^2 X_1 = -X_0^3. \tag{8.28}$$

For $p \neq \Omega$, the general solution of the zero order equation (8.27) is

$$X_0 = \left(\frac{p^2}{\Omega^2 - p^2}\right) \cos s + A \cos(\Omega s/p) + B \sin(\Omega s/p),$$

where A and B are arbitrary constants. Since X_0 is known to have period 2π, it follows that A and B must be zero unless Ω is an integer multiple of p; we will assume this is *not*

* The fact that τ is the minimum period of X does not *neccessarily* make it the minimum period of the left side of equation (8.25).

the case. Then the required solution of the **zero order** equation is

$$X_0 = \left(\frac{p^2}{\Omega^2 - p^2}\right) \cos s. \tag{8.29}$$

The **first order** equation (8.28) can now be written

$$\begin{aligned} X_1'' + \left(\frac{\Omega}{p}\right)^2 X_1 &= -\left(\frac{p^2}{\Omega^2 - p^2}\right)^3 \cos^3 s \\ &= -\left(\frac{p^6}{4(\Omega^2 - p^2)^3}\right) (3\cos s + \cos 3s), \end{aligned} \tag{8.30}$$

on using the trigonometric identity $\cos 3s = 4\cos^3 s - 3\cos s$. Since Ω/p is not an integer, the only solution of this equation that has period 2π is

$$X_1 = -\left(\frac{p^8}{4(\Omega^2 - p^2)^3}\right)\left(\frac{3\cos s}{\Omega^2 - p^2} + \frac{\cos 3s}{\Omega^2 - 9p^2}\right). \tag{8.31}$$

Results

When ϵ $(= F_0^3 \Lambda/p^6)$ is small, the **driven response** of the Duffing equation (8.22) (with $k = 0$) is given by

$$x = \frac{F_0}{\Omega^2 - p^2}\left[\cos pt - \left(\frac{3p^6 \cos pt}{(\Omega^2 - p^2)^3} + \frac{p^6 \cos 3pt}{(\Omega^2 - p^2)^2(\Omega^2 - 9p^2)}\right)\epsilon + O\left(\epsilon^2\right)\right]. \tag{8.32}$$

This is the *approximate solution correct to the first order in the small parameter* ϵ. More terms can be obtained in a similar way but this is best done with computer assistance.

The most interesting feature of this formula is the behaviour of the first order correction term when Ω is close to $3p$, which suggests the existence of a *super-harmonic resonance* with frequency $3p$. Similar 'resonances' occur in the higher terms at the frequencies $5p, 7p, \ldots$, and are caused by the presence of the non-linear term Λx^3. It should not however be concluded that large amplitude responses occur at these frequencies.* The critical case in which $\Omega = 3p$ is solved in Problem 8.14 and reveals no infinities in the response.

Sub-harmonic responses and chaos

We have so far left open the interesting question of whether a driving force with minimum period τ can excite a **subharmonic response**, that is, a response whose minimum period is

* This is a subtle point. Like all power series, perturbation series have a certain 'radius of convergence'. When *all* the terms of the perturbation series are included, ϵ is resticted to some range of values $-\epsilon_0 < \epsilon < \epsilon_0$. What seems to happen when Ω approaches $3p$ is that ϵ_0 approaches zero so that the first order correction term never actually gets large.

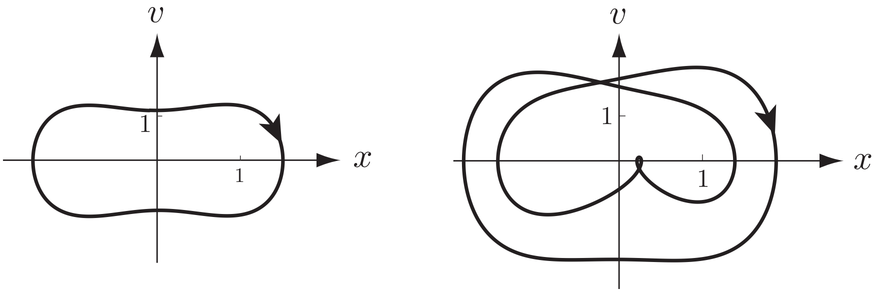


FIGURE 8.9 Two different periodic responses to the same driving force. **Left**: A response of period 2π, **Right**: A sub-harmonic response of period 4π.

an *integer multiple* of τ. This is certainly not possible in the linear case, where the driving force and the induced response always have the same period. One way of investigating this problem would be to expand the (unknown) response $x(t)$ as a Fourier series and to substitute this into the left side of Duffing's equation. One would then require all the odd numbered terms to magically cancel out leaving a function with period 2τ. Unlikely though this may seem, it can happen! *There are ranges of the parameters in Duffing's equation that permit a sub-harmonic response*. Indeed, it is possible for the *same* set of parameters to allow more than one periodic response. Figure 8.9 shows two different periodic responses of the equation $d^2x/dt^2 + kdx/dt + x^3 = A\cos t$, each corresponding to $k = 0.04$, $A = 0.9$. One response has period 2π while the other is a **subharmonic response** with period 4π. Which of these is the steady state response depends on the initial conditions. It is also possible for the motion to be **chaotic** with no steady state ever being reached, even though damping is present.

Problems on Chapter 8

Answers and comments are at the end of the book.

Harder problems carry a star (∗).

Periodic oscillations: Lindstedt's method

8.1 A non-linear oscillator satisfies the equation

$$\left(1 + \epsilon x^2\right)\ddot{x} + x = 0,$$

where ϵ is a small parameter. Use Linstedt's method to obtain a two-term approximation to the oscillation frequency when the oscillation has unit amplitude. Find also the corresponding two-term approximation to $x(t)$. [You will need the identity $4\cos^3 s = 3\cos s + \cos 3s$.]

8.2 A non-linear oscillator satisfies the equation

$$\ddot{x} + x + \epsilon x^5 = 0,$$

where ϵ is a small parameter. Use Linstedt's method to obtain a two-term approximation to the oscillation frequency when the oscillation has unit amplitude. [You will need the identity $16\cos^5 s = 10\cos s + 5\cos 3s + \cos 5s$.]

8.3 ***Unsymmetrical oscillations*** A non-linear oscillator satisfies the equation

$$\ddot{x} + x + \epsilon x^2 = 0,$$

where ϵ is a small parameter. Explain why the oscillations are unsymmetrical about $x = 0$ in this problem.

Use Linstedt's method to obtain a two-term approximation to $x(t)$ for the oscillation in which the *maximum* value of x is unity. Deduce a two-term approximation to the *minimum* value achieved by $x(t)$ in this oscillation.

8.4* ***A limit cycle by perturbation theory*** Use perturbation theory to investigate the limit cycle of **Rayleigh's equation**, taken here in the form

$$\ddot{x} + \epsilon\left(\tfrac{1}{3}\dot{x}^2 - 1\right)\dot{x} + x = 0,$$

where ϵ is a small positive parameter. Show that the zero order approximation to the limit cycle is a circle and determine its centre and radius. Find the frequency of the limit cycle correct to order ϵ^2, and find the function $x(t)$ correct to order ϵ.

Phase paths

8.5 ***Phase paths in polar form*** Show that the system of equations

$$\dot{x}_1 = F_1(x_1, x_2, t), \qquad \dot{x}_2 = F_2(x_1, x_2, t)$$

can be written in polar coordinates in the form

$$\dot{r} = \frac{x_1 F_1 + x_2 F_2}{r}, \qquad \dot{\theta} = \frac{x_1 F_2 - x_2 F_1}{r^2},$$

where $x_1 = r\cos\theta$ and $x_2 = r\sin\theta$.

A dynamical system satisfies the equations

$$\dot{x} = -x + y,$$
$$\dot{y} = -x - y.$$

Convert this system into polar form and find the polar equations of the phase paths. Show that every phase path encircles the origin infinitely many times in the clockwise direction. Show further that every phase path terminates at the origin. Sketch the phase diagram.

8.6 A dynamical system satisfies the equations

$$\dot{x} = x - y - (x^2 + y^2)x,$$
$$\dot{y} = x + y - (x^2 + y^2)y.$$

Convert this system into polar form and find the polar equations of the phase paths that begin in the domain $0 < r < 1$. Show that all these phase paths spiral anti-clockwise and tend to the limit cycle $r = 1$. Show also that the same is true for phase paths that begin in the domain $r > 1$. Sketch the phase diagram.

8.7 A damped linear oscillator satisfies the equation

$$\ddot{x} + \dot{x} + x = 0.$$

Show that the polar equations for the motion of the phase points are

$$\dot{r} = -r\sin^2\theta, \qquad \dot{\theta} = -\left(1 + \tfrac{1}{2}\sin 2\theta\right).$$

Show that every phase path encircles the origin infinitely many times in the clockwise direction. Show further that these phase paths terminate at the origin.

8.8 A non-linear oscillator satisfies the equation

$$\ddot{x} + \dot{x}^3 + x = 0.$$

Find the polar equations for the motion of the phase points. Show that phase paths that begin within the circle $r < 1$ encircle the origin infinitely many times in the clockwise direction. Show further that these phase paths terminate at the origin.

8.9 A non-linear oscillator satisfies the equation

$$\ddot{x} + \left(x^2 + \dot{x}^2 - 1\right)\dot{x} + x = 0.$$

Find the polar equations for the motion of the phase points. Show that any phase path that starts in the domain $1 < r < \sqrt{3}$ spirals clockwise and tends to the limit cycle $r = 1$. [The same is true of phase paths that start in the domain $0 < r < 1$.] What is the period of the limit cycle?

8.10 ***Predator–prey*** Consider the symmetrical predator–prey equations

$$\dot{x} = x - xy, \qquad \dot{y} = xy - y,$$

where $x(t)$ and $y(t)$ are positive functions. Show that the phase paths satisfy the equation

$$\left(xe^{-x}\right)\left(ye^{-y}\right) = A,$$

where A is a constant whose value determines the particular phase path. By considering the shape of the surface

$$z = \left(xe^{-x}\right)\left(ye^{-y}\right),$$

deduce that each phase path is a simple closed curve that encircles the equilibrium point at $(1, 1)$. Hence *every solution* of the equations is periodic! [This prediction can be confirmed by solving the original equations numerically.]

Poincaré–Bendixson

8.11 Use Poincaré–Bendixson to show that the system

$$\dot{x} = x - y - (x^2 + 4y^2)x,$$
$$\dot{y} = x + y - (x^2 + 4y^2)y,$$

has a limit cycle lying in the annulus $\frac{1}{2} < r < 1$.

8.12 ***Van der Pol's equation*** Use Poincaré–Bendixson to show that Van der Pol's equation*

$$\ddot{x} + \epsilon \dot{x}\left(x^2 - 1\right) + x = 0,$$

has a limit cycle for any *positive* value of the constant ϵ. [The method is similar to that used for Rayleigh's equation in Example 8.4.]

Driven oscillations

8.13 A driven non-linear oscillator satisfies the equation

$$\ddot{x} + \epsilon \dot{x}^3 + x = \cos pt,$$

where ϵ, p are positive constants. Use perturbation theory to find a two-term approximation to the driven response when ϵ is small. Are there any restrictions on the value of p?

8.14 ***Super-harmonic resonance*** A driven non-linear oscillator satisfies the equation

$$\ddot{x} + 9x + \epsilon x^3 = \cos t,$$

where ϵ is a small parameter. Use perturbation theory to investigate the possible existence of a superharmonic resonance. Show that the zero order solution is

$$x_0 = \frac{1}{8}\left(\cos t + a_0 \cos 3t\right),$$

where the constant a_0 is a constant that is not known at the zero order stage.

By proceeding to the first order stage, show that a_0 is the unique real root of the cubic equation

$$3a_0^3 + 6a_0 + 1 = 0,$$

which is about -0.164. Thus, when driving the oscillator at this sub-harmonic frequency, the non-linear correction appears in the *zero order* solution. However, there are no infinities to be found in the perturbation scheme at this (or any other) stage.

Plot the graph of $x_0(t)$ and the path of the phase point $(x_0(t), x_0'(t))$.

---

* After the extravagantly named Dutch physicist Balthasar Van der Pol (1889–1959). The equation arose in connection with the current in an electronic circuit. In 1927 Van der Pol observed what is now called *deterministic chaos*, but did not investigate it further.

Computer assisted problems

8.15 ***Linstedt's method*** Use computer assistance to implement Lindstedt's method for the equation

$$\ddot{x} + x + \epsilon x^3 = 0.$$

Obtain a three-term approximation to the oscillation frequency when the oscillation has unit amplitude. Find also the corresponding three-term approximation to $x(t)$.

8.16 ***Van der Pol's equation*** A classic non-linear oscillation equation that has a limit cycle is Van der Pol's equation

$$\ddot{x} + \epsilon \left(x^2 - 1\right) \dot{x} + x = 0,$$

where ϵ is a positive parameter. Solve the equation numerically with $\epsilon = 2$ (say) and plot the motion of a few of the phase points in the (x, v)-plane. All the phase paths tend to the limit cycle. One can see the same effect in a different way by plotting the solution function $x(t)$ against t.

8.17 ***Sub-harmonic and chaotic responses*** Investigate the *steady state* responses of the equation

$$\ddot{x} + k\dot{x} + x^3 = A \cos t$$

for various choices of the parameters k and A and various initial conditions. First obtain the responses shown in Figure 8.9 and then go on to try other choices of the parameters. Some very exotic results can be obtained! For various chaotic responses try $K = 0.1$ and $A = 7$.

Part Two

MULTI-PARTICLE SYSTEMS AND CONSERVATION PRINCIPLES

CHAPTERS IN PART TWO

Chapter Nine

The energy principle

and energy conservation

KEY FEATURES

The key features of this chapter are the **energy principle** for a multi-particle system, the **potential energies** arising from **external** and **internal** forces, and **energy conservation**.

This is the first of three chapters in which we study the mechanics of **multi-particle systems**. This is an important development which greatly increases the range of problems that we can solve. In particular, multi-particle mechanics is needed to solve problems involving the rotation of rigid bodies.

The chapter begins by obtaining the **energy principle** for a multi-particle system. This is the first of the three great principles of multi-particle mechanics* that apply to *every* mechanical system without restriction. We then show that, under appropriate conditions, the total energy of the system is conserved. We apply this **energy conservation** principle to a wide variety of systems. When the system has just one degree of freedom, the energy conservation equation is sufficient to determine the whole motion.

9.1 CONFIGURATIONS AND DEGREES OF FREEDOM

A **multi-particle system** $\mathcal{S}$ may consist of any number of particles P_1, P_2, ..., P_N, with masses m_1, m_2 ..., m_N respectively.† A possible 'position' of the system is called a **configuration**. More precisely, if the particles P_1, P_2, ..., P_N of a system have position vectors $\boldsymbol{r}_1$, $\boldsymbol{r}_2$, ..., $\boldsymbol{r}_N$, then any *geometrically possible* set of values for the position vectors $\{\boldsymbol{r}_i\}$ is a configuration of the system.

If the system is **unconstrained**, then each particle can take up any position in space (independently of the others) and all choices of the $\{\boldsymbol{r}_i\}$ are possible. This would be the case, for instance, if the particles of $\mathcal{S}$ were moving freely under their mutual gravitation. On the other hand, when **constraints** are present, the $\{\boldsymbol{r}_i\}$ are restricted. Suppose for instance that the particles P_1 and P_2 are connected by a light rigid rod of length a.

* The other two are the *linear momentum* and *angular momentum* principles.

† To save space, we will usually express this by saying that $\mathcal{S}$ is the system of particles $\{P_i\}$ with masses $\{m_i\}$, the range of the index number i being understood to be $1 \le i \le N$.

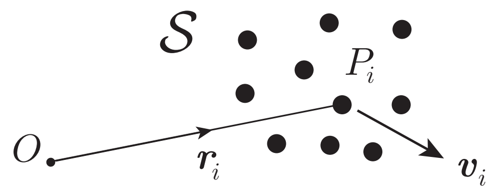


FIGURE 9.1 The multi-particle system $\mathcal{S}$ consists of N particles $P_1, P_2, \ldots, P_N$, of which the typical particle P_i is labelled. The particle P_i has mass m_i, position vector $\boldsymbol{r}_i$, and velocity $\boldsymbol{v}_i$.

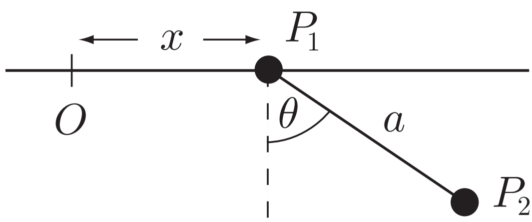


FIGURE 9.2 The generalised coordinates x and θ are sufficient to specify the configuration of this two-particle system in planar motion.

This imposes the geometrical restriction $|\boldsymbol{r}_1 - \boldsymbol{r}_2| = a$ so that not all choices of the $\{\boldsymbol{r}_i\}$ are then possible. This difference is reflected in the number of scalar variables needed to specify the configuration of $\mathcal{S}$. In the unconstrained case, all of the position vectors $\{\boldsymbol{r}_i\}$ must be specified separately. Since each of these vectors may be specified by three Cartesian coordinates, it follows that a total of $3N$ *scalar variables are needed to specify the configuration of an unconstrained* N*-particle system*. When constraints are present, this number is reduced, often dramatically so.

For example, consider the system shown in Figure 9.2, which consists of two particles P_1 and P_2 connected by a light rigid rod of length a. The particle P_1 is also constrained to move along a fixed horizontal rail and the whole system moves in the vertical plane through the rail. The *two* scalar variables x and θ shown are sufficient to specify the configuration of this system. This contrasts with the *six* scalar variables that would be needed if the two particles were in unconstrained motion. The variables x and θ are said to be a set of **generalised coordinates** for this system.* Other choices for the generalised coordinates could be made, but the *number* of generalised coordinates needed is always the same.

Definition 9.1 ***Degrees of freedom*** *The* ***number*** *of generalised coordinates needed to specify the configuration of a system* $\mathcal{S}$ *is called the number of* ***degrees of freedom*** *of* $\mathcal{S}$.

Importance of degrees of freedom

The number of degrees of freedom of a system is important because it is equal to the *number of equations that are needed to determine the motion of the system*. For example,

* Besides being sufficient to specify the configuration of the system, the generalised coordinates are also required to be *independent*, that is, there must be no functional relation between them. The coordinates x, θ in Figure 9.2 are certainly independent variables. If the coordinates were connected by a functional relation, they would not all be needed and one of them could be discarded.

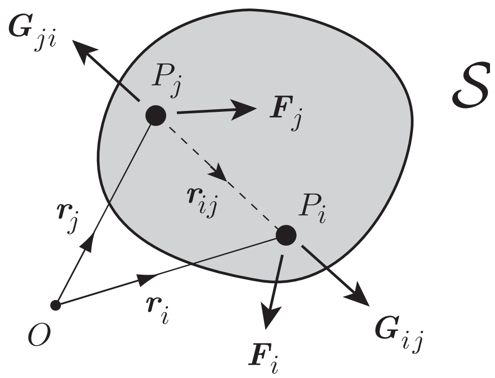


FIGURE 9.3 The multi-particle $\mathcal{S}$ consists of N particles $P_1, P_2, \ldots, P_N$, of which the typical particles P_i and P_j are shown explicitly. The force $\boldsymbol{F}_i$ is the *external* force acting on P_i and the force $\boldsymbol{G}_{ij}$ is the *internal* force exerted on P_i by the particle P_j.

the system shown in Figure 9.2 has *two* degrees of freedom and so needs *two* equations to determine the motion completely.

Example 9.1 ***Degrees of freedom***

Find the number of degrees of freedom of the following mechanical systems: (i) the simple pendulum (moving in a vertical plane), (ii) a door swinging on its hinges, (iii) a bar of soap (a particle) sliding on the inside of a hemispherical basin, (iv) a rigid rod sliding on a flat table, (v) four rigid rods flexibly jointed to form a quadrilateral which can slide on a flat table.

Solution

(i) 1 (ii) 1 (iii) 2 (iv) 3 (v) 4. ■

9.2 THE ENERGY PRINCIPLE FOR A SYSTEM

Let $\mathcal{S}$ be a system of N particles $\{P_i\}$, as shown in Figure 9.3. We classify the forces acting on the particles of $\mathcal{S}$ as being external or internal. **External forces** are those originating from *outside* $\mathcal{S}$. (In the case of a single particle, these are the only forces that act.) Uniform gravity is an example of an external force. However, in multi-particle systems, the particles are also subject to their own *mutual interactions*, that is, the forces that they exert upon each other. These mutual interactions are called the **internal forces** acting on $\mathcal{S}$. The situation is shown in Figure 9.3. $\boldsymbol{F}_i$ is the external force acting on the particle P_i, while $\boldsymbol{G}_{ij}$ is the internal force exerted on P_i by the particle P_j. By the Third Law, the force $\boldsymbol{G}_{ji}$ that P_i exerts on P_j must be equal and opposite to the force $\boldsymbol{G}_{ij}$, and both forces must be parallel to the straight line joining P_i and P_j. In short, the $\{\boldsymbol{G}_{ij}\}$ must satisfy

$$\boldsymbol{G}_{ji} = -\boldsymbol{G}_{ij}, \qquad \text{and} \qquad \boldsymbol{G}_{ij} \parallel (\boldsymbol{r}_i - \boldsymbol{r}_j). \tag{9.1}$$

To obtain the energy principle for the system $\mathcal{S}$, we proceed in the same way as we did for a single particle in section 6.1. The equation of motion for the particle P_i is*

$$m_i \frac{d\boldsymbol{v}_i}{dt} = \boldsymbol{F}_i + \sum_{j=1}^{N} \boldsymbol{G}_{ij}, \tag{9.2}$$

where $\boldsymbol{v}_i$ is the velocity of P_i at time t. On taking the scalar product of both sides of equation (9.2) with $\boldsymbol{v}_i$ and then *summing* the result over all the particles ($1 \le i \le N$), we obtain

$$\frac{dT}{dt} = \sum_{i=1}^{N} \left\{ \boldsymbol{F}_i + \sum_{j=1}^{N} \boldsymbol{G}_{ij} \right\} \cdot \boldsymbol{v}_i, \tag{9.3}$$

where

$$T = \sum_{i=1}^{N} \tfrac{1}{2} m_i |\boldsymbol{v}_i|^2,$$

the **total kinetic energy** of the whole system $\mathcal{S}$. Suppose that, in the time interval $[t_{\mathcal{A}}, t_{\mathcal{B}}]$, the system $\mathcal{S}$ moves from configuration $\mathcal{A}$ to configuration $\mathcal{B}$. On integrating equation (9.3) with respect to t over the time interval $[t_{\mathcal{A}}, t_{\mathcal{B}}]$ we obtain

$$\boxed{T_{\mathcal{B}} - T_{\mathcal{A}} = \sum_{i=1}^{N} \int_{t_{\mathcal{A}}}^{t_{\mathcal{B}}} \boldsymbol{F}_i \cdot \boldsymbol{v}_i \, dt + \sum_{i=1}^{N} \sum_{j=1}^{N} \int_{t_{\mathcal{A}}}^{t_{\mathcal{B}}} \boldsymbol{G}_{ij} \cdot \boldsymbol{v}_i \, dt} \tag{9.4}$$

where $T_{\mathcal{A}}$ and $T_{\mathcal{B}}$ are the kinetic energies of the system $\mathcal{S}$ at times $t_{\mathcal{A}}$ and $t_{\mathcal{B}}$ respectively. This is the **energy principle** for a multi-particle system moving under the external forces $\{\boldsymbol{F}_i\}$ and internal forces $\{\boldsymbol{G}_{ij}\}$. This impressive looking result can be stated quite simply as follows:

Energy principle for a multi-particle system

In any motion of a system, the increase in the total kinetic energy of the system in a given time interval is equal to the total work done by all the external and internal forces during this time interval.

---

* The summation over j in equation (9.2) contains the term $\boldsymbol{G}_{ii}$ which corresponds to the force that the particle P_i exerts upon *itself*. Since such a force is not actually present, we should really say that the summation is over the range $1 \le j \le N$ with $j \ne i$. Since this would make the formulae look messy, we adopt the device of regarding the terms $\boldsymbol{G}_{11}, \boldsymbol{G}_{22}, \ldots, \boldsymbol{G}_{NN}$ (which do not actually exist) as being zero.

9.3 ENERGY CONSERVATION FOR A SYSTEM

In order to develop an energy conservation principle, we need to write the right side of the energy principle (9.4) in the form $V(\mathcal{A}) - V(\mathcal{B})$, where V is the potential energy function for the *whole system*. We first consider *unconstrained* systems.

Unconstrained systems

When the system is unconstrained, all the forces that act on the system are specified directly. We will assume that the **external forces** $\boldsymbol{F}_i$ are *conservative fields*. In this case $\boldsymbol{F}_i = -\operatorname{grad}\phi_i$, where ϕ_i is the potential energy function of the field $\boldsymbol{F}_i$. Then the total work done by the external forces can be written

$$\sum_{i=1}^{N}\int_{t_\mathcal{A}}^{t_\mathcal{B}} \boldsymbol{F}_i \cdot \boldsymbol{v}_i\, dt = \sum_{i=1}^{N}\left(\phi_i(\boldsymbol{r}_\mathcal{A}) - \phi_i(\boldsymbol{r}_\mathcal{B})\right) = \Phi(\mathcal{A}) - \Phi(\mathcal{B}),$$

where

$$\Phi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N) = \phi_1(\boldsymbol{r}_1) + \phi_2(\boldsymbol{r}_2) + \cdots + \phi_N(\boldsymbol{r}_N)$$

is the potential energy of $\mathcal{S}$ arising from the **external** forces.

Example 9.2 *Potential energy under uniform gravity*

Find the potential energy Φ when the external forces on $\mathcal{S}$ arise from uniform gravity.

Solution

Under uniform gravity, the force $\boldsymbol{F}_i$ exerted on particle P_i is $\boldsymbol{F}_i = -m_i g\boldsymbol{k}$, where the unit vector $\boldsymbol{k}$ points vertically upwards. This conservative field has potential energy $\phi_i = m_i g z_i$, where z_i is the z-coordinate of P_i. The total potential energy of $\mathcal{S}$ due to uniform gravity is therefore

$$\Phi = m_1 g z_1 + m_2 g z_2 + \cdots + m_N g z_N.$$

On using the definition of centre of mass given in section 3.5, this can be written in the alternative form

$$\Phi = MgZ,$$

where M is the total mass of $\mathcal{S}$, and Z is the z-cordinate of the centre of mass of $\mathcal{S}$. Thus *the potential energy of any system due to uniform gravity is the same as if all its mass were concentrated at its centre of mass.* ■

We now need to make a similar transformation to show that the work done by the **internal forces** can be written in the form $\Psi(\mathcal{A}) - \Psi(\mathcal{B})$, where Ψ is the internal potential energy. The argument is as follows:

We know from the Third Law that the $\{\boldsymbol{G}_{ij}\}$ satisfy the conditions (9.1), but a little more must be assumed. We further assume that the *magnitude* of $\boldsymbol{G}_{ij}$ depends only on r_{ij}, the distance between P_i and

P_j.* Internal forces that satisfy this conditions will be called **conservative**; mutual gravitation forces are a typical example. Hence, when the internal forces are conservative, $\boldsymbol{G}_{ij}$ must have the form

$$\boldsymbol{G}_{ij} = h_{ij}(r_{ij})\,\widehat{\boldsymbol{r}}_{ij} \tag{9.5}$$

where (see Figure 9.3)

$$\boldsymbol{r}_{ij} = \boldsymbol{r}_i - \boldsymbol{r}_j \qquad r_{ij} = |\boldsymbol{r}_{ij}| \qquad \widehat{\boldsymbol{r}}_{ij} = \boldsymbol{r}_{ij}/r_{ij}. \tag{9.6}$$

Note that h_{ij} is the *repulsive* force that the particles P_i and P_j exert upon each other.

Consider now the rate of working of the *pair* of forces $\boldsymbol{G}_{ij}$ and $\boldsymbol{G}_{ji}$. This is

$$\begin{aligned}\boldsymbol{G}_{ij}\cdot\boldsymbol{v}_i + \boldsymbol{G}_{ji}\cdot\boldsymbol{v}_j = \boldsymbol{G}_{ij}\cdot\left(\boldsymbol{v}_i - \boldsymbol{v}_j\right) = h_{ij}(r_{ij})\,\widehat{\boldsymbol{r}}_{ij}\cdot\frac{d\boldsymbol{r}_{ij}}{dt} &= \left(\frac{h_{ij}(r_{ij})}{r_{ij}}\right)\boldsymbol{r}_{ij}\cdot\frac{d\boldsymbol{r}_{ij}}{dt}\\ &= h_{ij}(r_{ij})\,\frac{dr_{ij}}{dt},\end{aligned}$$

on using equations (9.1), (9.6) and the identity $\boldsymbol{r}_{ij}\cdot\dot{\boldsymbol{r}}_{ij} = r_{ij}\dot{r}_{ij}$. The total work done by the forces $\boldsymbol{G}_{ij}$ and $\boldsymbol{G}_{ji}$ during the time interval $[t_{\mathcal{A}}, t_{\mathcal{B}}]$ is therefore

$$\int_{t_{\mathcal{A}}}^{t_{\mathcal{B}}} h_{ij}(r_{ij})\,\frac{dr_{ij}}{dt}\,dt = \int_{r_{ij}(\mathcal{A})}^{r_{ij}(\mathcal{B})} h_{ij}(r_{ij})\,dr_{ij} = H_{ij}(r_{ij}(\mathcal{A})) - H_{ij}(r_{ij}(\mathcal{B})),$$

where H_{ij} is the indefinite integral of $-h_{ij}$. The function $H_{ij}(r_{ij})$ is called the **mutual potential energy** of the particles P_i and P_j.

It follows that the total work done by all the internal forces in the time interval $[t_{\mathcal{A}}, t_{\mathcal{B}}]$ can be written in the form

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\int_{t_{\mathcal{A}}}^{t_{\mathcal{B}}} \boldsymbol{G}_{ij}\cdot\boldsymbol{v}_i\,dt = \Psi(\mathcal{A}) - \Psi(\mathcal{B}),$$

where

$$\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N) = \sum_{i=1}^{N}\sum_{j=1}^{i-1} H_{ij}(r_{ij})$$

is the **potential energy** of $\mathcal{S}$ arising from the **internal** forces. This potential energy is just the sum of the mutual potential energies of all pairs of particles.

Example 9.3 *Internal energy of three charged particles*

Three particles P_1, P_2, P_3 carry electric charges e_1, e_2, e_3 respectively. Find the internal potential energy Ψ.

Solution

In cgs/electrostatic units, the particles P_1 and P_2 repel each other with the force $h_{12}(r_{12}) = e_1e_2/(r_{12})^2$, where r_{12} is the distance between P_1 and P_2. Their mutual potential energy is therefore

$$H_{12} = -\int h_{12}(r_{12})\,dr_{12} = -\int \frac{e_1e_2}{(r_{12})^2}\,dr_{12} = \frac{e_1e_2}{r_{12}}.$$

* This is equivalent to the very reasonable assumptions that the magnitude of $\boldsymbol{G}_{ij}$ is invariant under spatial translations and rotations of each pair of particles P_i and P_j, and is independent of the time.

The **internal potential energy** of the whole system is therefore

$$\Psi = \frac{e_1 e_2}{r_{12}} + \frac{e_1 e_3}{r_{13}} + \frac{e_2 e_3}{r_{23}}. \blacksquare$$

On combining the above results, the energy principle (9.4) can be written

$$T_{\mathcal{B}} - T_{\mathcal{A}} = V(\mathcal{A}) - V(\mathcal{B}),$$

where $V = \Phi + \Psi$ is the **total potential energy** of the system $\mathcal{S}$. This is equivalent to the **energy conservation** formula

$$\boxed{T + V = E} \tag{9.7}$$

where E is the total energy of the system.This result can be summarised as follows:

Energy conservation for an unconstrained system

When both the external and internal forces acting on a system are *conservative*, the sum of its kinetic and potential energies* remains constant in the motion.

Example 9.4 *A star with two planets*

A star of very large mass M is orbited by two planets P_1 and P_2 of masses m_1 and m_2. Find the energy conservation equation for this system.

Solution

Since the mass of the star is supposed to be very much larger than the planetary masses, we will neglect its motion and suppose that it is fixed at the origin O. We then have a *two-particle* problem in which the planets move under the (external) gravitational attraction of the star and their (internal) mutual gravitational interaction. This is an unconstrained system.

The total potential energy arising from **external forces** is then

$$\Phi = -\frac{Mm_1G}{r_1} - \frac{Mm_2G}{r_2},$$

where r_1,r_2 are the distances OP_1, OP_2.

The particles P_1 and P_2 repel each other with the force $h_{12}(r_{12}) = -m_1m_2G/(r_{12})^2$, where r_{12} is the distance between P_1 and P_2. Their **mutual potential energy** is therefore

$$H_{12} = -\int h_{12}(r_{12})\, dr_{12} = \int \frac{m_1m_2G}{(r_{12})^2}\, dr_{12} = -\frac{m_1m_2G}{r_{12}},$$

and this is the only contribution to the internal potential energy Ψ.

* The potential energy is the total of the potential energies arising from both the external and internal forces.

Since the system is unconstrained and the external and internal forces are conservative, energy conservation applies. The **energy conservation equation** for the system is

$$\tfrac{1}{2}m_1\,|\boldsymbol{v}_1|^2 + \tfrac{1}{2}m_2\,|\boldsymbol{v}_2|^2 - MG\left(\frac{m_1}{r_1} + \frac{m_2}{r_2}\right) - \frac{m_1 m_2 G}{r_{12}} = E,$$

where $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ are the velocities of the planets P_1, P_2, and E is the constant total energy. The value of E is determined from the initial conditions.

Since this system has six degrees of freedom (four if the motions are confined to a plane through O), the energy conservation equation is by no means sufficient to determine the motion! ■

Question ***Can a planet escape?***

If the initial conditions are such that $E < 0$, is it possible for a planet to escape to infinity?

Answer

If $E < 0$, then it is certainly not possible for *both* planets to escape to infinity, since the total energy would then be positive. However, the escape of *one* planet is not prohibited by energy conservation. This does not mean however that such an escape will actually happen.

Constrained systems

When a system is subject to constraints, not all the forces that act on the system are specified. This is because constraints are enforced by **constraint forces** that are not part of the specification of the problem; all we know is that their *effect* is to enforce the given constraints. The work done by constraint forces cannot generally be calculated (or expressed in terms of a potential energy) and we are restricted to those *systems for which the total work done by the constraint forces happens to be zero*.*

The constraint forces acting on the system may be **external** (for example, when a particle of the system is constrained to remain at rest), or **internal** (for example, when two particles of the system are constrained to remain the same distance apart).

A The list of **external** constraint forces that do no work is the same as that given in Section 6.5 for single particle motion.

B The most important result regarding **internal** constraint forces that do no work is this: *The total work done by any pair of mutual interaction forces is zero when the particles on which they act are constrained to remain a fixed distance apart.* The proof is as follows:

Suppose two particles P_i and P_j are constrained to remain a fixed distance apart and that their mutual interaction forces are $\boldsymbol{G}_{ij}$ and $\boldsymbol{G}_{ji}$ (see Figure 9.3). Since the distance between P_i and P_j is constant, it follows that $(\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{r}_i - \boldsymbol{r}_j)$ is constant, which, on differentiating with respect

* *Individual* constraint forces may do work.

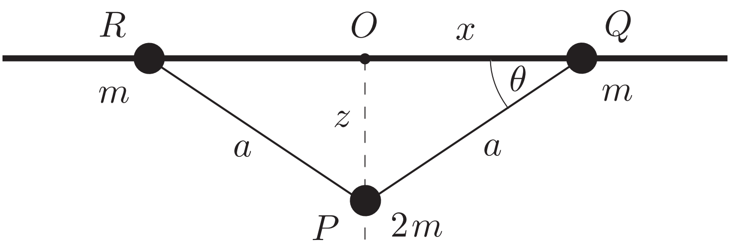


FIGURE 9.4 The particles Q and R slide along a smooth horizontal rail while the particle P moves vertically.

to t, gives

$$(\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{v}_i - \boldsymbol{v}_j) = 0.$$

Thus the vector $(\boldsymbol{v}_i - \boldsymbol{v}_j)$ must be *perpendicular* to the straight line joining P_i and P_j. Hence, the rate of working of the *two* forces $\boldsymbol{G}_{ij}$ and $\boldsymbol{G}_{ji}$ is

$$\boldsymbol{G}_{ij} \cdot \boldsymbol{v}_i + \boldsymbol{G}_{ji} \cdot \boldsymbol{v}_j = \boldsymbol{G}_{ij} \cdot \left(\boldsymbol{v}_i - \cdot \boldsymbol{v}_j\right) = 0,$$

since $\boldsymbol{G}_{ij}$ is known to be *parallel* to the straight line joining P_i and P_j. Thus the internal constraint forces $\boldsymbol{G}_{ij}$ and $\boldsymbol{G}_{ji}$ do no work *in total*.

It follows, for example, that the two tension forces exerted by a light inextensible string do no work in total. It further follows that the *internal forces that enforce rigidity in a rigid body do no work in total*. This important result allows us to solve rigid body problems by energy methods.

Our result for constrained systems can be summarised as follows:

Energy conservation for a constrained system

When the specified external and internal forces acting on a system are *conservative*, and the constraint forces *do no work in total*, the sum of the kinetic and potential energies of the system remains constant in the motion.

Example 9.5 *A constrained three-particle system*

Figure 9.4 shows a ball P of mass $2m$ suspended by light inextensible strings of length a from two sliders Q and R, each of mass m, which can move on a smooth horizontal rail. The system moves symmetrically so that O, the mid-point of Q and R, remains fixed and P moves on the downward vertical through O. Initially the system is released from rest with the three particles in a straight line and with the strings taut. Find the energy conservation equation for the system.

Solution

This is a system with one degree of freedom and we take the angle θ as the generalised coordinate. Let z and x be the displacements of the particles P and Q from the fixed point O. Then, in terms of the generalised coordinate θ, $x = a\cos\theta$ and $z = a\sin\theta$. Differentiating these formulae with respect to t then gives

$$\dot{x} = -(a\sin\theta)\dot{\theta}, \qquad \dot{z} = (a\cos\theta)\dot{\theta}.$$

Hence the total **kinetic energy** of the system is given by

$$T = \tfrac{1}{2}(2m)\dot{z}^2 + \tfrac{1}{2}m\dot{x}^2 + \tfrac{1}{2}m\dot{x}^2 = ma^2\dot{\theta}^2.$$

The only contribution to the **potential energy** comes from uniform gravity, so that

$$V = -(2m)gz + 0 + 0 = -2mga\sin\theta,$$

where we have taken the zero level of potential energy to be at the rail.

We must now show that the constraint forces do no work. The reactions exerted by the smooth rail on the particles Q and R are perpendicular to the rail and therefore perpendicular to the velocities of Q and R; these reactions therefore do no work. Also, the tension forces exerted by the inextensible strings do no work in total. Hence, the **constraint forces** do no work in total.

Energy conservation therefore applies in the form

$$ma^2\dot{\theta}^2 - 2mga\sin\theta = E.$$

From the initial conditions $\theta = \dot{\theta} = 0$ when $t = 0$, it follows that $E = 0$. The **energy conservation equation** for the system is therefore

$$\dot{\theta}^2 - \frac{2g}{a}\sin\theta = 0. \blacksquare$$

Question *When do the sliders collide?*

Find the time that elapses before the sliders collide.

Answer

Since this system has only one degree of freedom, the motion can be found from energy conservation alone. From the energy conservation equation, it follows that

$$\frac{d\theta}{dt} = \pm\left(\frac{2g}{a}\right)^{1/2}(\sin\theta)^{1/2},$$

and, since θ is an *increasing* function of t, we take the *positive* sign. This equation is a first order separable ODE.

Since the sliders collide when $\theta = \pi/2$, the time τ that elapses is given by

$$\tau = \left(\frac{a}{2g}\right)^{1/2}\int_0^{\pi/2}\frac{d\theta}{(\sin\theta)^{1/2}} \approx 1.85\left(\frac{a}{g}\right)^{1/2}. \blacksquare$$

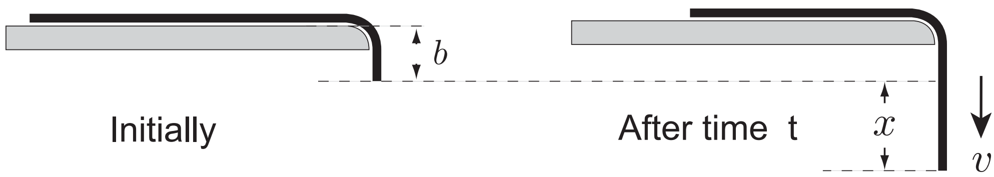


FIGURE 9.5 A uniform rope is released from rest hanging over the edge of a smooth table (left). After time t it has displacement x (right).

Example 9.6 *Rope sliding off a table*

A uniform inextensible rope of mass M and length a is released from rest hanging over the edge of a smooth horizontal table, as shown in Figure 9.5. Find the speed of the rope when it has the displacement x shown.

Solution

A rope is a continuous distribution of mass, unlike the discrete masses that appear in our theory. We regard the rope as being represented by a *light* inextensible string of length a with N particles, each of mass M/N, attached to the string at equally spaced intervals along its length. When N is very large, we expect this discrete set of masses to approximate the behaviour of the rope.

Since each particle of the rope has the same *speed* $v\,(=\dot{x})$, the total **kinetic energy** of the rope is simply

$$T = \tfrac{1}{2}Mv^2.$$

The only contribution to the **potential energy** comes from uniform gravity. If we take the reference state for V to be the initial configuration (Figure 9.5 (left)), then the potential energy in the displaced configuration (right) is the same *as if* a length x of the rope lying on the table were cut off and this piece were then suspended from the hanging end. In the continuous limit (that is, as $N \to \infty$), this piece of rope has mass Mx/a and its centre of mass is lowered a distance $b + (x/2)$ by this operation. The potential energy of the rope in the displaced configuration is therefore

$$V = -\left(\frac{Mx}{a}\right) g \left(b + \tfrac{1}{2}x\right).$$

We must now show that the constraint forces do no work. The reactions exerted by the smooth table on the particles of the rope are always perpendicular to the velocities of these particles; these reactions therefore do no work. Also, the tension forces exerted by each segment of the inextensible string (connecting adjacent particles of the rope) do no work in total. Hence, the **constraint forces** do no work in total.

Energy conservation therefore applies in the form

$$\tfrac{1}{2}Mv^2 - \left(\frac{Mx}{a}\right) g \left(b + \tfrac{1}{2}x\right) = E.$$

The initial condition $v = 0$ when $x = 0$ implies that $E = 0$. The energy equation for the rope is therefore

$$v^2 = \frac{g}{a}\, x(x + 2b).$$

This gives the **speed** of the rope when it has displacement x. This formula holds while there is still some rope left on the the table *top*. ■

Note. In the above solution we have assumed that the rope follows the contour of the table edge and then falls vertically. However, it can be shown that this *cannot* be true when the rope is close to leaving the table. What actually happens is that the end of the rope overshoots the table edge. This is a tricky point which we will not investigate further.

Question *Displacement at time t*

Find the displacement of the rope at time t.

Answer

Since this system has only one degree of freedom, the motion can be found from energy conservation alone. From the energy conservation equation, it follows that

$$\frac{dx}{dt} = \pm n\, x^{1/2}(x + 2b)^{1/2},$$

where $n^2 = g/a$. Since x is an *increasing* function of t, we take the *positive* sign. This equation is a first order separable ODE.

It follows that

$$\begin{aligned} nt &= \int \frac{dx}{x^{1/2}(x + 2b)^{1/2}} \\ &= 2\sinh^{-1}\left(\frac{x}{2b}\right)^{1/2} + C, \end{aligned}$$

on using the substitution $x = 2b\sinh^2 w$. The initial condition $x = 0$ when $t = 0$ implies that $C = 0$ and, after some simplification, we obtain

$$x = b(\cosh nt - 1)$$

as the **displacement** of the rope after time t. As before, this formula holds while there is still some rope left on the the table top. ■

Example 9.7 *Stability of a plank on a log*

A uniform thin rigid plank is placed on top of a rough circular log and can roll without slipping. Show that the equilibrium position, in which the plank rests symmetrically on top of the log, is stable.

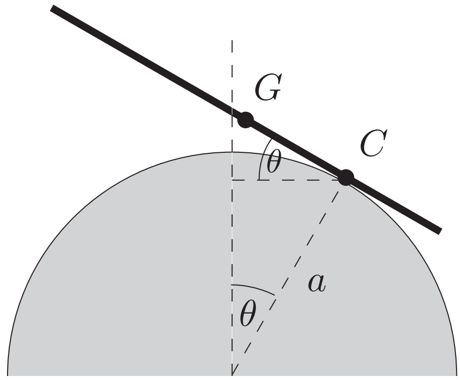


FIGURE 9.6 A thin uniform plank is placed symmetrically on top of a fixed rough circular log. Is the equilibrium position of the plank stable?

Solution

Suppose that the plank is disturbed from its equilibrium position and is tilted by an angle θ as shown in Figure 9.6. The plank is known to *roll* on the log, which means that the distance GC from the centre G of the plank to the contact point C must always be equal to the arc length of the log that has been traversed. If the radius of the log is a, then this arc length is $a\theta$.

We are not yet able to calculate the **kinetic energy** of the plank in terms of the coordinate θ. This is done in the next section. However, we do not need it to investigate stability.

The only contribution to the **potential energy** of the plank comes from uniform gravity. This is given by $V = MgZ$, where Z is the vertical displacement of the centre of mass G of the plank. Elementary trigonometry (see Figure 9.6) shows that $Z = a\cos\theta + a\theta\sin\theta - a$, so that

$$V = Mga(\cos\theta + \theta\sin\theta - 1).$$

We must now show that the constraint forces do no work. The rate of working of the constraint force $\boldsymbol{R}$ that the log exerts on the plank is $\boldsymbol{R}\cdot\boldsymbol{v}^C$, where $\boldsymbol{v}^C$ is the velocity of the particle C of the plank that is *instantaneously* in contact with the log. But, since the plank rolls on the log, $\boldsymbol{v}^C = \boldsymbol{0}$ so that the rate of working of $\boldsymbol{R}$ is zero. Also, the internal constraint forces that enforce the rigidity of the plank do no work in total. Hence, the **constraint forces** do no work in total.

Energy conservation therefore applies in the form

$$T + Mga(\cos\theta + \theta\sin\theta - 1) = E.$$

It follows that the equilibrium position (with the plank on top of the log) will be stable if V has a *minimum* at $\theta = 0$. Now $V' = Mga\theta\cos\theta$ and $V'' = Mga(\cos\theta - \theta\sin\theta)$ so that, when $\theta = 0$, $V' = 0$ and $V'' = 1$. Hence V has a minimum at $\theta = 0$ and so the **equilibrium position** is stable. ■

9.4 KINETIC ENERGY OF A RIGID BODY

The general theory we have presented applies to *any* multi-particle system; in particular, it applies to the rigid array of particles that we call a **rigid body**. However, in

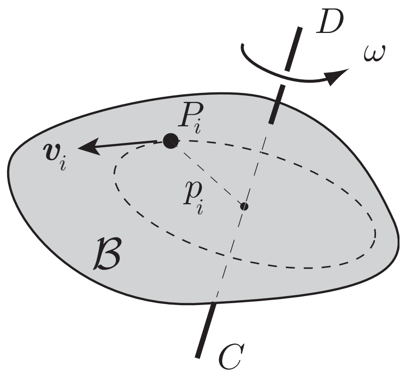


FIGURE 9.7 The rigid body $\mathcal{B}$ rotates about the *fixed* axis CD with angular velocity ω. A typical particle P_i moves on the circular path shown.

order to make use of energy conservation in rigid body dynamics, we need to be able to express the **kinetic energy** T of the body in terms of the generalised coordinates.

Rigid body with a fixed axis

Figure 9.7 shows a rigid body $\mathcal{B}$ which is rotating about the *fixed* axis CD. (Imagine that the body is penetrated by a thin light spindle, which is smoothly pivoted in a fixed position.) A typical particle P_i of the body can move on the circular path shown. This circle has radius p_i, where p_i is the perpendicular distance of P_i from the axis CD. Suppose that, at some instant, the angular velocity of $\mathcal{B}$ about the axis CD is ω. Then the *speed* of particle P_i at this instant is $|\omega| p_i$, and its kinetic energy is $\frac{1}{2}m_i\,(\omega p_i)^2$. The total kinetic energy of $\mathcal{B}$ is therefore

$$T = \sum_{i=1}^{N}\left(\tfrac{1}{2}m_i\,(\omega p_i)^2\right) = \tfrac{1}{2}\left(\sum_{i=1}^{N} m_i\,{p_i}^2\right)\omega^2.$$

Definition 9.2 ***Moment of inertia*** *The quantity*

$$\boxed{I_{CD} = \sum_{i=1}^{N} m_i\,{p_i}^2} \tag{9.8}$$

where p_i is the perpendicular distance of the mass m_i from the axis CD, is called the ***moment of inertia*** *of the body $\mathcal{B}$ about the axis CD.*

The **moment of inertia**, as defined above, does not depend on the motion of the body $\mathcal{B}$. It is a purely *geometrical* quantity (like centre of mass), which describes how the mass in $\mathcal{B}$ is distributed relative to the axis CD. The further the mass in $\mathcal{B}$ lies from the axis, the larger is the moment of inertia of $\mathcal{B}$ about that axis. In the theory of rotating rigid bodies, the moment of inertia plays a similar rôle to that played by mass in the translational motion of a particle.

Our result may be summarised as follows:

Kinetic energy of a rigid body with a fixed axis

Suppose the rigid body $\mathcal{B}$ is rotating about the fixed axis CD with angular velocity ω. Then the kinetic energy of $\mathcal{B}$ is given by

$$T = \tfrac{1}{2} I_{CD}\, \omega^2, \tag{9.9}$$

where I_{CD} is the moment of inertia of $\mathcal{B}$ about the axis CD.

Example 9.8 *Moment of inertia of a hoop*

Find the moment of inertia of a uniform hoop of mass M and radius a about its axis of rotational symmetry.

Solution

This is the easiest case to treat since each particle of the hoop has perpendicular distance a from the specified axis. The required moment of inertia is therefore

$$I = \sum_{i=1}^{N} m_i\, a^2 = \left(\sum_{i=1}^{N} m_i \right) a^2 = Ma^2,$$

where M is the mass of the whole hoop. ■

It is evident that, in order to solve problems that include rotating rigid bodies, we need to know their moments of inertia. These can be worked out from the definition (9.8), or its counterpart for continuous mass distributions. The Appendix at the end of the book contains examples of how to do this and also contains a table of common moments of inertia, including those for the uniform **rod, hoop, disk** and **sphere**. Most readers will find it convenient to remember the moments of inertia in these four cases.

Example 9.9 *Rotational kinetic energy of the Earth*

Estimate the rotational kinetic energy of the Earth, regarded as a rigid uniform sphere rotating about a *fixed* axis through its centre.

Solution

From the Appendix, we find that I, the moment of inertia of a uniform sphere about an axis through its centre is given by $I = 2MR^2/5$, where M is the mass of the sphere and R its radius. The kinetic energy of the Earth is therefore given by

$$T = \tfrac{1}{2} I \omega^2 = \tfrac{1}{5} M R^2 \omega^2,$$

where M is the mass the Earth, R is its radius, and ω is its angular velocity.

On inserting the values $M = 6.0 \times 10^{24}$ kg, $R = 6400$ km and $\omega = 7.3 \times 10^{-5}$ radians per second, $T = 2.6 \times 10^{29}$ J approximately. ■

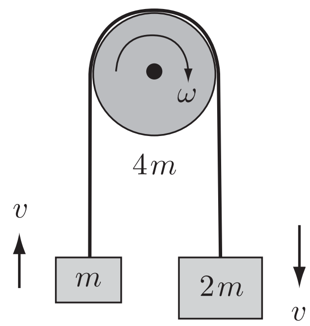


FIGURE 9.8 Two blocks of masses m and $2m$ are connected by a light inextensible string which passes over a circular pulley of mass $4m$ and radius a.

Example 9.10 *Attwood's machine*

Two blocks of masses m and $2m$ are connected by a light inextensible string which passes over a uniform circular pulley of radius a and mass $4m$. Find the upward acceleration of the mass m.

Solution

The system is shown in Figure 9.8. We suppose that the string does not slip on the pulley and that the pulley is smoothly pivoted about its axis of symmetry.

Let z be the upward displacement of the mass m (from some reference configuration) and v $(= \dot{z})$ its upward velocity at time t. Then, since the string is *inextensible*, the mass $2m$ must have the same displacement and velocity, but measured downwards. The angular velocity ω of the pulley is determined from the condition that the string does not slip. In this case, the velocity of the rim of the pulley and the velocity of the string must be the same at each point where they are in contact, that is, $a\omega = v$. Hence $\omega = v/a$. Also, from the table in the Appendix, the moment of inertia of a uniform circular disk of mass M and radius a about its axis of symmetry is $\frac{1}{2}Ma^2$. Hence, the total **kinetic energy** of the system is

$$T = \tfrac{1}{2}mv^2 + \tfrac{1}{2}(2m)v^2 + \tfrac{1}{2}\left(\tfrac{1}{2}(4m)a^2\right)\left(\frac{v}{a}\right)^2 = \tfrac{5}{2}mv^2.$$

The gravitational **potential energy** of the system (relative to the reference configuration) is

$$V = mgz - (2m)gz = -mgz.$$

We must now dispose of the **constraint forces**. (i) At the smooth pivot that supports the pulley, the reactions are perpendicular to the velocities of the particles on which they act. Hence these reactions do no work. (ii) Since there is no slippage between the string and the three material bodies of the system, the total work done by the string on the bodies must be equal and opposite to the total work done by the bodies on the string.* (iii) The internal forces that keep the pulley rigid do no work in total. Hence the constraint forces do no work in total.

* Since this string is massless and inextensible, it can have neither kinetic nor potential energy so that the total work done on the string must actually be zero.

Energy conservation therefore applies in the form

$$\tfrac{5}{2}mv^2 - mgz = E,$$

where E is the total energy. If we now differentiate this equation with respect to t (and cancel by mv), we obtain

$$\frac{dv}{dt} = \tfrac{1}{5}g$$

which is the **equation of motion** of the system. Thus the upward **acceleration** of the mass m is $g/5$. (If the pulley were massless, the result would be $g/3$.) ∎

Rigid body in general motion

We now go on to find the kinetic energy of a rigid body that has translational as well as rotational motion. The method depends on the following theorem.

Theorem 9.1 *Suppose a general system of particles $\mathcal{S}$ has total mass M and that its centre of mass G has velocity $\boldsymbol{V}$. Then the total kinetic energy of $\mathcal{S}$ can be written in the form*

$$T = \tfrac{1}{2}MV^2 + T^G, \tag{9.10}$$

where $V = |\boldsymbol{V}|$ and T^G is the kinetic energy of $\mathcal{S}$ in its motion ***relative to*** *G.*

Proof. By definition,

$$\begin{aligned}
T^G &= \frac{1}{2}\sum_{i=1}^{N} \tfrac{1}{2}m_i\,(\boldsymbol{v}_i - \boldsymbol{V})\cdot(\boldsymbol{v}_i - \boldsymbol{V}) \\
&= \frac{1}{2}\sum_{i=1}^{N} m_i\boldsymbol{v}_i\cdot\boldsymbol{v}_i - \frac{1}{2}\left(\sum_{i=1}^{N} m_i\boldsymbol{v}_i\right)\cdot\boldsymbol{V} - \frac{1}{2}\boldsymbol{V}\cdot\left(\sum_{i=1}^{N} m_i\boldsymbol{v}_i\right) + \frac{1}{2}\left(\sum_{i=1}^{N} m_i\right)\boldsymbol{V}\cdot\boldsymbol{V} \\
&= T - \tfrac{1}{2}(M\boldsymbol{V})\cdot\boldsymbol{V} - \tfrac{1}{2}\boldsymbol{V}\cdot(M\boldsymbol{V}) + \tfrac{1}{2}M(\boldsymbol{V}\cdot\boldsymbol{V}) \\
&= T - \tfrac{1}{2}MV^2,
\end{aligned}$$

as required. ∎

The term $\tfrac{1}{2}MV^2$ can be regarded as the **translational** contribution to T. When the system $\mathcal{S}$ is a **rigid body**, T^G also has a nice physical interpretation. In this case, the motion of $\mathcal{S}$ relative to G is an angular velocity ω about an axis CD passing through G, as shown in Figure 9.9. It then follows from equation (9.9) that $T^G = \tfrac{1}{2}I_{CD}\,\omega^2$. This can be regarded as the **rotational** contribution to T. We therefore have the result:

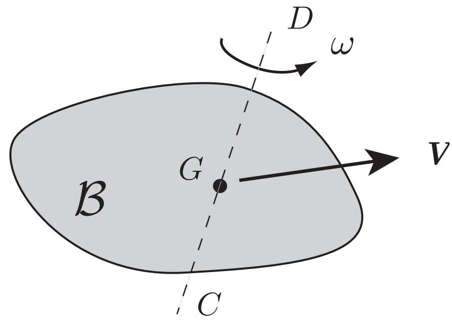


FIGURE 9.9 A rigid body $\mathcal{B}$ in general motion. The centre of mass G has velocity $\boldsymbol{V}$ and $\mathcal{B}$ is also rotating with angular velocity ω about an axis through G.

Kinetic energy of a rigid body in general motion

Let $\mathcal{B}$ be a rigid body of mass M and and let G be its centre of mass. Suppose that G has velocity $\boldsymbol{V}$ and that the body is also rotating with angular velocity ω about an axis CD passing through G. Then the **kinetic energy** of $\mathcal{B}$ is given by

$$T = \tfrac{1}{2}MV^2 + \tfrac{1}{2}I_{CD}\,\omega^2, \tag{9.11}$$

where $V = |\boldsymbol{V}|$ and I_{CD} is the moment of inertia of $\mathcal{B}$ about the axis CD. The term $\frac{1}{2}MV^2$ is called the **translational** kinetic energy and the term $\frac{1}{2}I_{CD}\,\omega^2$ the **rotational** kinetic energy of $\mathcal{B}$.

Example 9.11 *Kinetic energy of a rolling wheel*

Find the kinetic energy of the rolling wheel shown in Figure 2.8.

Solution

Assume the wheel to be uniform with mass M and radius b. Then its centre of mass C has speed u so that the **translational** kinetic energy is $\frac{1}{2}Mu^2$. Because of the rolling condition, the angular velocity of the wheel is given by $\omega = u/b$ so that the **rotational** kinetic energy is $\frac{1}{2}I(u/b)^2$, where $I = \frac{1}{2}Mb^2$. The **total** kinetic energy of the wheel is therefore given by

$$T = \tfrac{1}{2}Mu^2 + \tfrac{1}{2}\left(\tfrac{1}{2}Mb^2\right)\left(\frac{u}{b}\right)^2 = \frac{3Mu^2}{4}. \blacksquare$$

Example 9.12 *Cylinder rolling down a plane*

A uniform hollow circular cylinder is rolling down a *rough* plane inclined at an angle α to the horizontal. Find the acceleration of the cylinder.

Solution

Suppose that, at time t, the cylinder has displacement x down the plane (from some reference configuration) and that the centre of mass G of the cylinder has velocity

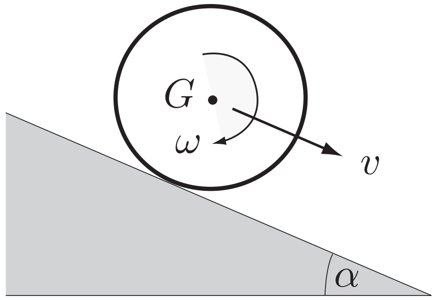


FIGURE 9.10 A hollow circular cylinder rolls down a plane inclined at angle α to the horizontal.

v $(= \dot{x})$ down the plane. The angular velocity ω of the cylinder is then determined by the rolling condition to be $\omega = v/b$. The kinetic energy of the cylinder is therefore

$$T = \tfrac{1}{2}Mv^2 + \tfrac{1}{2}I\omega^2 = \tfrac{1}{2}Mv^2 + \tfrac{1}{2}I\left(\frac{v}{b}\right)^2$$

where M is the mass of the cylinder, and I is its moment of inertia about its axis of symmetry. From the Appendix, we find that $I = Mb^2$ so that the **kinetic energy** of the cylinder is given by $T = Mv^2$.

The gravitational **potential energy** of the cylinder is given by $V = -Mgx \sin\alpha$.

We must now dispose of the constraint forces. The reaction forces that the inclined plane exerts on the cylinder act on particles of the cylinder which, because of the rolling condition, have zero velocity. These reaction forces therefore do no work. Also the internal forces that keep the cylinder rigid do no work in total. Hence the **constraint forces** do no work in total.

Conservation of energy therefore applies in the form

$$Mv^2 - Mgx\sin\alpha = E,$$

where E is the total energy. If we now differentiate this equation with respect to t (and cancel by Mv), we obtain

$$\frac{dv}{dt} = \tfrac{1}{2}g\sin\alpha,$$

which is the **equation of motion** of the cylinder. Thus the **acceleration** of the cylinder down the plane is $\frac{1}{2}g\sin\alpha$. (A block sliding down a *smooth* plane would have acceleration $g\sin\alpha$.) ■

Example 9.13 *The sliding ladder*

A uniform ladder of length $2a$ is supported by a smooth horizontal floor and leans against a smooth vertical wall.* The ladder is released from rest in a position making an angle of 60° with the downward vertical. Find the energy conservation equation for the ladder.

* Don't try this at home!

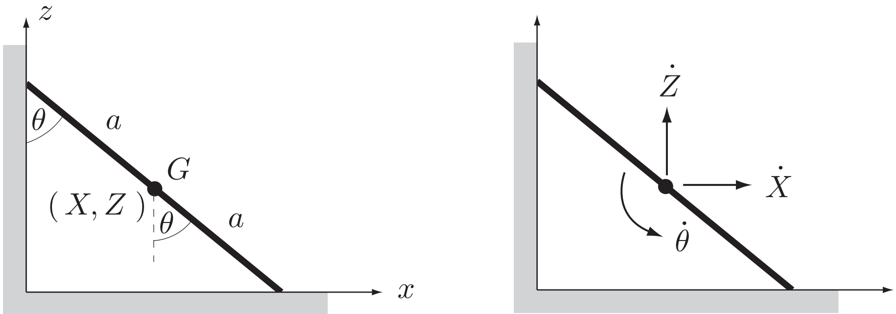


FIGURE 9.11 A uniform ladder of mass M and length $2a$ is supported by a smooth horizontal floor and leans against a smooth vertical wall. At time t, its centre of mass G has (x, z)-coordinates (X, Z) and the ladder makes an angle θ with the downward vertical.

Solution

Let θ be the angle that the ladder makes with the downward vertical after time t. The (x, z)-coordinates of the centre of mass G are then given by

$$X = a\sin\theta, \qquad Z = a\cos\theta,$$

and the corresponding velocity components by

$$\dot{X} = (a\cos\theta)\dot{\theta}, \qquad \dot{Z} = -(a\sin\theta)\dot{\theta}.$$

The angular velocity of the ladder at time t is simply $\dot{\theta}$ (see Figure 9.11). The **kinetic energy** of the ladder is therefore given by

$$T = \tfrac{1}{2}M\left(\dot{X}^2 + \dot{Y}^2\right) + \tfrac{1}{2}I\dot{\theta}^2 = \tfrac{1}{2}Ma^2\dot{\theta}^2 + \tfrac{1}{2}I\dot{\theta}^2,$$

where M is the mass of the ladder and I is its moment of inertia about the horizontal axis through G. From the Appendix, we find that $I = Ma^2/3$ so that the **kinetic energy** of the ladder is given by $T = (2Ma^2/3)\dot{\theta}^2$.

The gravitational **potential energy** of the ladder is given by $V = MgZ = Mga\cos\theta$.

We must now dispose of the constraint forces. The reaction forces that the smooth floor and wall exert on the ladder are both perpendicular to the particles of the ladder on which they act. These reaction forces therefore do no work. Also, the internal forces that keep the ladder rigid do no work in total. Hence the **constraint forces** do no work in total.

Conservation of energy therefore applies in the form

$$\tfrac{2}{3}Ma^2\,\dot{\theta}^2 + Mga\cos\theta = E,$$

where E is the total energy. From the initial conditions $\dot{\theta} = 0$ and $\theta = \pi/3$ when $t = 0$, it follows that $E = \frac{1}{2}Mga$. The **energy conservation equation** for the ladder

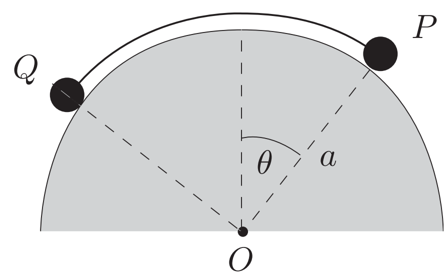


FIGURE 9.12 Two particles P and Q are connected by a light inextensible string and can move, with the string taut, on the surface of a smooth horizontal cylinder.

is therefore

$$\dot{\theta}^2 = \frac{3g}{4a}(1 - 2\cos\theta).$$

Since the system has only one degree of freedom, this equation is sufficient to determine the motion.

A curious feature of this problem (not proved here) is that the ladder does not maintain contact with the wall all the way down, but leaves the wall when θ becomes equal to $\cos^{-1}(1/3) \approx 71°$. ■

Problems on Chapter 9

Answers and comments are at the end of the book.

Harder problems carry a star (∗).

Potential energy and stability

9.1 Figure 9.12 shows two particles P and Q, of masses M and m, that can move on the smooth outer surface of a fixed horizontal cylinder. The particles are connected by a light inextensible string of length $\pi a/2$. Find the equilibrium configuration and show that it is unstable.

9.2 A uniform rod of length $2a$ has one end smoothly pivoted at a fixed point O. The other end is connected to a fixed point A, which is a distance $2a$ vertically above O, by a light elastic spring of natural length a and modulus $\frac{1}{2}mg$. The rod moves in a vertical plane through O. Show that there are two equilibrium positions for the rod, and determine their stability. [The vertically upwards position for the rod would compress the spring to zero length and is excluded.]

9.3 The internal potential energy function for a diatomic molecule is approximated by the **Morse potential**

$$V(r) = V_0\left(1 - e^{-(r-a)/b}\right)^2 - V_0,$$

where r is the distance of separation of the two atoms, and V_0, a, b are positive constants. Make a sketch of the Morse potential.

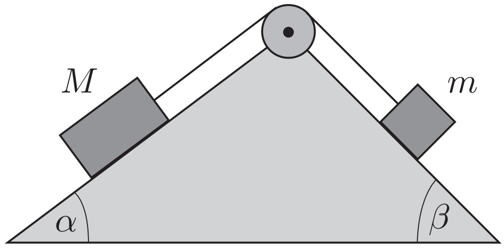


FIGURE 9.13 Two blocks of masses M and m slide on smooth planes inclined at angles α and β to the horizontal. The blocks are connected by a light inextensible string that passes over a light frictionless pulley.

Suppose the molecule is restricted to *vibrational* motion in which the centre of mass G of the molecule is fixed, and the atoms move on a fixed straight line through G. Show that there is a single equilibrium configuration for the molecule and that it is stable. If the atoms each have mass m, find the angular frequency of small vibrational oscillations of the molecule.

9.4∗ The internal gravitational potential energy of a system of masses is sometimes called the **self energy** of the system. (The reference configuration is taken to be one in which the particles are all a great distance from each other.) Show that the self energy of a uniform sphere of mass M and radius R is $-3M^2G/5R$. [Imagine that the sphere is built up by the addition of successive thin layers of matter brought in from infinity.]

Particles only

9.5 Figure 9.13 shows two blocks of masses M and m that slide on smooth planes inclined at angles α and β to the horizontal. The blocks are connected by a light inextensible string that passes over a light frictionless pulley. Find the acceleration of the block of mass m up the plane, and deduce the tension in the string.

9.6 Consider the system shown in Figure 9.12 for the special case in which the particles P, Q have masses $2m$, m respectively. The system is released from rest in a symmetrical position with θ, the angle between OP and the upward vertical, equal to $\pi/4$. Find the energy conservation equation for the subsequent motion in terms of the coordinate θ.

∗ Find the normal reactions of the cylinder on each of the particles. Show that P is first to leave the cylinder and that this happens when $\theta = 70°$ approximately.

Ropes

9.7 A heavy uniform rope of length $2a$ is draped symmetrically over a *thin* smooth horizontal peg. The rope is then disturbed slightly and begins to slide off the peg. Find the speed of the rope when it finally leaves the peg.

9.8 A uniform heavy rope of length a is held at rest with its two ends close together and the rope hanging symmetrically below. (In this position, the rope has two long vertical segments connected by a small curved segment at the bottom.) One of the ends is then released. Find the velocity of the free end when it has descended by a distance x.

Deduce a similar formula for the acceleration of the free end and show that it always *exceeds* g. Find how far the free end has fallen when its acceleration has risen to $5g$.

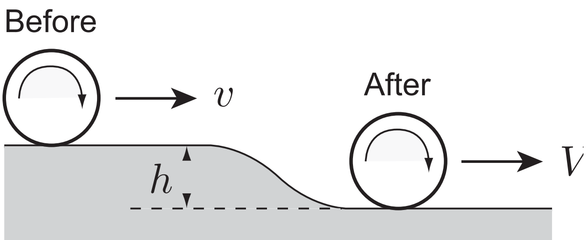


FIGURE 9.14 The circular hoop *rolls* down the slope from one level to another.

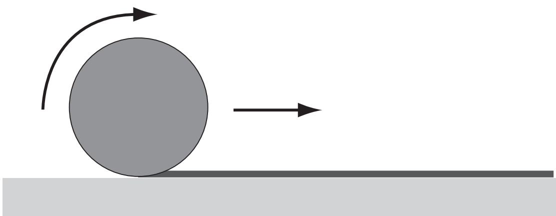


FIGURE 9.15 The roll of paper moves to the right and the free paper is gathered on to the roll.

9.9 A heavy uniform rope of mass M and length $4a$ has one end connected to a fixed point on a smooth horizontal table by light elastic spring of natural length a and modulus $\frac{1}{2}Mg$, while the other end hangs down over the edge of the table. When the spring has its natural length, the free end of the rope hangs a distance a vertically below the level of the table top. The system is released from rest in this position. Show that the free end of the rope executes simple harmonic motion, and find its period and amplitude.

Rigid bodies

9.10 A circular hoop is rolling with speed v along level ground when it encounters a slope leading to more level ground, as shown in Figure 9.14. If the hoop loses altitude h in the process, find its final speed.

9.11 A uniform ball is rolling in a straight line down a *rough* plane inclined at an angle α to the horizontal. Assuming the ball to be in planar motion, find the energy conservation equation for the ball. Deduce the acceleration of the ball.

9.12 A uniform circular cylinder (a yo-yo) has a light inextensible string wrapped around it so that it does not slip. The free end of the string is secured to a fixed point and the yo-yo descends in a vertical straight line with the straight part of the string also vertical. Explain why the string does no work on the yo-yo. Find the energy conservation equation for the yo-yo and deduce its acceleration.

9.13 Figure 9.15 shows a partially unrolled roll of paper on a horizontal floor. Initially the paper on the roll has radius a and the free paper is laid out in a straight line on the floor. The roll is then projected horizontally with speed V in such a way that the free paper is gathered up on to the roll. Find the speed of the roll when its radius has increased to b. [Neglect the bending stiffness of the paper.] Deduce that the radius of the roll when it comes to rest is

$$a\left(\frac{3V^2}{4ga}+1\right)^{1/3}.$$

9.14 A rigid body of general shape has mass M and can rotate freely about a fixed horizontal axis. The centre of mass of the body is distance h from the rotation axis, and the moment of inertia of the body about the rotation axis is I. Show that the period of small oscillations of the body about the downward equilibrium position is

$$2\pi\left(\frac{I}{Mgh}\right)^{1/2}.$$

Deduce the period of small oscillations of a uniform rod of length $2a$, pivoted about a horizontal axis perpendicular to the rod and distance b from its centre.

9.15 A uniform ball of radius a can roll without slipping on the *outside* surface of a fixed sphere of (outer) radius b and centre O. Initially the ball is at rest at the highest point of the sphere when it is slightly disturbed. Find the speed of the centre G of the ball in terms of the variable θ, the angle between the line OG and the upward vertical. [Assume planar motion.]

9.16 A uniform ball of radius a and centre G can roll without slipping on the *inside* surface of a fixed hollow sphere of (inner) radius b and centre O. The ball undergoes planar motion in a vertical plane through O. Find the energy conservation equation for the ball in terms of the variable θ, the angle between the line OG and the downward vertical. Deduce the period of small oscillations of the ball about the equilibrium position.

9.17* Figure 9.6 shows a uniform thin rigid plank of length $2b$ which can roll without slipping on top of a rough circular log of radius a. The plank is initially in equilibrium, resting symmetrically on top of the log, when it is slightly disturbed. Find the period of small oscillations of the plank.

Chapter Ten

The linear momentum principle

and linear momentum conservation

KEY FEATURES

The key features of this chapter are the **linear momentum principle**; its equivalent form, the **centre of mass equation**; and **conservation of linear momentum**. These principles are applied to **rocket propulsion**, **collision theory**, the **two-body problem** and **two-body scattering**.

This chapter is essentially based on the **linear momentum principle** and its consequences. The linear momentum principle is the second of the three great principles of multi-particle mechanics* that apply to *every* mechanical system without restriction. Under appropriate conditions, the linear momentum of a system (or one of its components) is **conserved**. Important applications include rocket propulsion, collision theory, the two-body problem and two-body scattering.

10.1 LINEAR MOMENTUM

We begin with the definition of linear momentum for a single particle and for a system of particles.

Definition 10.1 ***Linear momentum*** *If a particle has mass m and velocity $\boldsymbol{v}$, then $\boldsymbol{p}$, its* ***linear momentum****, is defined to be*

$$\boldsymbol{p} = m\boldsymbol{v}. \tag{10.1}$$

For a multi-particle system $\mathcal{S}$ consisting of particles $P_1, P_2, \ldots, P_N$, with masses $m_1, m_2, \ldots, m_N$ and velocities $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_N$ (see Figure 9.1), $\boldsymbol{P}$, the ***linear momentum*** *of $\mathcal{S}$, is defined to be the* ***vector sum*** *of the linear momenta of the individual particles, that is,*

$$\boldsymbol{P} = \sum_{i=1}^{N} \boldsymbol{p}_i = \sum_{i=1}^{N} m_i \boldsymbol{v}_i. \tag{10.2}$$

* The other two are the energy and angular momentum principles.

Newton's Second Law can be written in terms of linear momentum in the form

$$\frac{d\boldsymbol{p}}{dt} = \boldsymbol{F}.$$

Although this offers no advantage in the mechanics of a single particle, we will find that this type of formulation is very useful in multi-particle mechanics. The expression (10.2) can be written simply in terms of the motion of G, the centre of mass of $\mathcal{S}$. Since the position vector $\boldsymbol{R}$ of G is given by

$$\boldsymbol{R} = \frac{\sum_{i=1}^{N} m_i \boldsymbol{r}_i}{\sum_{i=1}^{N} m_i},$$

where $\boldsymbol{r}_i$ is the position vector of the particle P_i, it follows that $\boldsymbol{V}$, the velocity of G is given by

$$\boldsymbol{V} = \frac{\sum_{i=1}^{N} m_i \boldsymbol{v}_i}{\sum_{i=1}^{N} m_i} = \frac{\boldsymbol{P}}{M},$$

where M $(= \sum m_i)$ is the total mass of the system $\mathcal{S}$. Hence

$$\boldsymbol{P} = M\boldsymbol{V}. \tag{10.3}$$

Thus the *linear momentum of any system is the same as if all its mass were concentrated at its centre of mass*.

Although true for all systems, this result is most useful when finding the linear momentum of a moving **rigid body**. Note that the rotational motion of the rigid body does not contribute to its linear momentum; this contrasts with the corresponding calculation of the kinetic energy of a rigid body (see Chapter 9).

10.2 THE LINEAR MOMENTUM PRINCIPLE

We now derive the fundamental result which relates the linear momentum of any system to the external forces that act upon it: **the linear momentum principle**.

Suppose that the system $\mathcal{S}$ is acted upon by the **external** forces $\{\boldsymbol{F}_i\}$ and **internal** forces $\{\boldsymbol{G}_{ij}\}$, as shown in Figure 9.3. Then the equation of motion for the particle P_i is

$$m_i \frac{d\boldsymbol{v}_i}{dt} = \boldsymbol{F}_i + \sum_{j=1}^{N} \boldsymbol{G}_{ij}, \tag{10.4}$$

where, as in Chapter 9, we take $\boldsymbol{G}_{ij} = \boldsymbol{0}$ when $i = j$. Then the rate of increase of the linear momentum of the system $\mathcal{S}$ can be written

$$\frac{d\boldsymbol{P}}{dt} = \frac{d}{dt}\left(\sum_{i=1}^{N} m_i \boldsymbol{v}_i\right) = \sum_{i=1}^{N} m_i \frac{d\boldsymbol{v}_i}{dt}, \tag{10.5}$$

which, on using the equation of motion (10.4), gives

$$\frac{d\boldsymbol{P}}{dt} = \sum_{i=1}^{N}\left\{\boldsymbol{F}_i + \sum_{j=1}^{N}\boldsymbol{G}_{ij}\right\} = \sum_{i=1}^{N}\boldsymbol{F}_i + \sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{G}_{ij}$$
$$= \sum_{i=1}^{N}\boldsymbol{F}_i + \sum_{i=1}^{N}\left(\sum_{j=1}^{i-1}\left(\boldsymbol{G}_{ij} + \boldsymbol{G}_{ji}\right)\right),$$

where the terms of the double sum have been grouped in pairs and those terms known to be zero have been omitted. Now the internal forces $\{\boldsymbol{G}_{ij}\}$ satisfy the Third Law, so that $\boldsymbol{G}_{ji} = -\boldsymbol{G}_{ij}$. Hence, each term of the double sum in equation (10.5) is zero and we obtain

Linear momentum principle

$$\frac{d\boldsymbol{P}}{dt} = \boldsymbol{F} \qquad (10.6)$$

where $\boldsymbol{F}$ is the **total external force** acting on $\mathcal{S}$. This is the **linear momentum principle**. This fundamental principle can be expressed as follows:

Linear momentum principle

In any motion of a system, the rate of increase of its linear momentum is equal to the total *external* force acting upon it.

It should be noted that only the external forces appear in the linear momentum principle so that the *internal forces need not be known*. It is this fact which gives the linear momentum principle its power.

10.3 MOTION OF THE CENTRE OF MASS

The linear momentum principle can be written in an alternative form called the centre of mass equation, which is more useful for some purposes. If we substitute the expession (10.3) for $\boldsymbol{P}$ into the linear momentum principle (10.6) we obtain

Centre of mass equation

$$M\frac{d\boldsymbol{V}}{dt} = \boldsymbol{F} \qquad (10.7)$$

which is called the **centre of mass equation**. It has the form of an equation of motion for a *fictitious* particle of mass M situated at the centre of mass, which moves under the total of the external forces acting on the system $\mathcal{S}$. This important result can be simply expressed as follows:

Motion of the centre of mass

The centre of mass of any system moves as if it were a particle of mass the total mass, and all the *external* forces acted upon it.

Example 10.1 *Jumping cat*

A cat leaps off a table and lands on the floor. Show that, while the cat is in the air, its centre of mass moves on a parabolic path.

Solution

While the cat is in the air, the total external force on its body is due to uniform gravity, that is, $\boldsymbol{F} = -Mg\boldsymbol{k}$. The centre of mass equation for the cat is therefore

$$M\frac{d\boldsymbol{V}}{dt} = -Mg\boldsymbol{k},$$

which is precisely the equation of projectile motion for a single particle. The path of the centre of mass of the cat is therefore the same as if it were a particle of mass M moving freely under uniform gravity. This path is known (see Chapter 4) to be a parabola. ■

In previous examples, we have often used the Second Law to find an unknown constraint force acting on a particle, once the motion of a system has been found by other means (see, for instance, Example 6.13). The centre of mass equation allows us to do the same thing when the unknown constraint force acts on a rigid body. The following examples illustrate the method.

Example 10.2 *Cylinder rolling down an inclined plane*

Consider again a hollow cylinder of mass M rolling down a rough inclined plane as shown in Figure 9.10. In Example 9.12, energy conservation was used to show that the acceleration of the cylinder down the plane is $\frac{1}{2}g\sin\alpha$. Deduce the reaction force exerted by the plane on the cylinder.

Solution

Suppose that the component of the reaction force normal to the plane is N, while the component of the reaction force up the plane is F. (The plane is rough so both components are present.) The cylinder is therefore subject to these 'two' external forces together with uniform gravity. The **centre of mass equation** for the cylinder (when resolved into components tangential and normal to the plane) is given by

$$M\frac{dv}{dt} = Mg\sin\alpha - F, \qquad 0 = N - Mg\cos\alpha,$$

where $dv/dt = \frac{1}{2}g\sin\alpha$. It follows that the required **reactions** are given by

$$F = \tfrac{1}{2}Mg\sin\alpha, \qquad N = Mg\cos\alpha.$$

Thus, if F and N are restricted by the 'law of friction' $F/N < \mu$, then the supposed rolling motion of the cylinder cannot take place if $\tan\alpha > 2\mu$. ■

Example 10.3 *Sliding ladder*

Consider again the uniform ladder of length $2a$ supported by a smooth horizontal floor and leaning against a smooth vertical wall, as shown in Figure 9.11. The ladder is released from rest with θ, the angle between the ladder and the downward vertical, equal to 60°. In Example 9.13, we used energy conservation to show that, in the subsequent motion, θ satisfies the differential equation

$$\dot{\theta}^2 = \frac{3g}{4a}(1 - 2\cos\theta),$$

provided that the ladder maintains contact with the wall. Deduce that the ladder loses contact with the wall when $\theta = \cos^{-1}(1/3)$.

Solution

Let the normal reactions exerted on the ladder by the smooth floor and wall be N^F and N^W respectively. Then the centre of mass equation for the ladder, resolved into horizontal and vertical components, is given by

$$M\ddot{X} = N^W, \qquad M\ddot{Z} = N^F - Mg,$$

where (X, Z) are the coordinates of the centre of mass of the ladder (see Figure 9.11). Hence

$$N^F = M\ddot{Z} + Mg, \qquad N^W = M\ddot{X}.$$

Now, in terms of the angle θ, $X = a\sin\theta$ and $Z = a\cos\theta$. On differentiating twice with respect to t, we obtain the corresponding acceleration components

$$\ddot{X} = -a(\sin\theta)\,\dot{\theta}^2 + a(\cos\theta)\,\ddot{\theta},$$

$$\ddot{Z} = -a(\cos\theta)\,\dot{\theta}^2 - a(\sin\theta)\,\ddot{\theta}.$$

Hence

$$N^F = -Ma\left((\cos\theta)\,\dot{\theta}^2 - (\sin\theta)\,\ddot{\theta}\right) + Mg,$$
$$N^W = Ma\left(-(\sin\theta)\,\dot{\theta}^2 + (\cos\theta)\,\ddot{\theta}\right).$$

In order to express these reactions in terms of θ alone, we need to know $\dot{\theta}^2$ and $\ddot{\theta}$ as functions of θ. From the previously derived equation of motion, we already have

$$\dot{\theta}^2 = \frac{3g}{4a}(1 - 2\cos\theta)$$

and, if we differentiate this equation with respect to t (and cancel by $\dot{\theta}$), we obtain

$$\ddot{\theta} = \frac{3g}{4a}\sin\theta.$$

On making use of the above expressions for $\dot{\theta}^2$ and $\ddot{\theta}$, the required **reactions** are found to be

$$N^F = \frac{Mg}{4}(1 - 3\cos\theta + 9\cos^2\theta), \quad N^W = \frac{3Mg}{4}\sin\theta(3\cos\theta - 1).$$

We observe that the predicted value of N^W becomes zero when $\theta = \cos^{-1}(1/3)$ and is *negative* thereafter. Since negative values of N^W cannot occur (the wall can only *push*), we conclude that the condition that the ladder maintains contact with the wall is violated when $\theta > \cos^{-1}(1/3)$. Therefore, the **ladder leaves the wall** when $\theta = \cos^{-1}(1/3)$. ■

10.4 CONSERVATION OF LINEAR MOMENTUM

Suppose that $\mathcal{S}$ is an **isolated** system, meaning that **no external force** acts on any of its particles. Then $\boldsymbol{F}$, the total external force acting on $\mathcal{S}$, is obviously zero. The linear momentum principle (10.6) for $\mathcal{S}$ then takes the form $d\boldsymbol{P}/dt = \boldsymbol{0}$, which implies that $\boldsymbol{P}$ must remain constant. This simple but important result can be stated as follows:

Conservation of linear momentum

In any motion of an isolated system, the total linear momentum is conserved.

It follows from equation (10.3) that the above result can also be stated in the alternative form '*In any motion of an isolated system, the centre of mass of the system moves with constant velocity*'. Clearly the same result applies to any system for which the *total* external force is zero, whether isolated or not.

It is also possible for a particular component of $\boldsymbol{P}$ to be conserved while other components are not. Let $\boldsymbol{n}$ be a *constant* unit vector and suppose that $\boldsymbol{F}\cdot\boldsymbol{n} = 0$ at all times. Then

$$\frac{d}{dt}(\boldsymbol{P}\cdot\boldsymbol{n}) = \frac{d\boldsymbol{P}}{dt}\cdot\boldsymbol{n} + \boldsymbol{P}\cdot\frac{d\boldsymbol{n}}{dt} = \frac{d\boldsymbol{P}}{dt}\cdot\boldsymbol{n} = \boldsymbol{F}\cdot\boldsymbol{n} = 0.$$

Hence the component $\boldsymbol{P}\cdot\boldsymbol{n}$ is conserved. This result can be stated as follows:

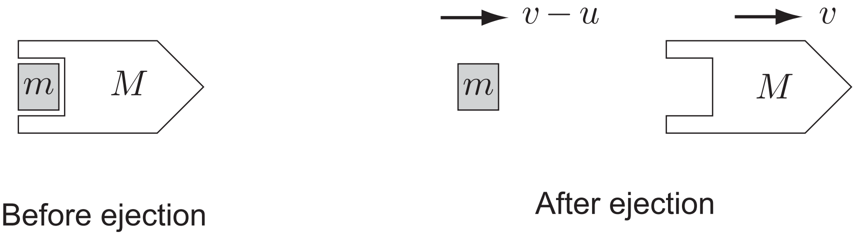


FIGURE 10.1 A rigid body of mass M (the rocket) contains a removable rigid block of mass m (the fuel). An internal source of energy causes the fuel block to be ejected backwards with speed u *relative to the rocket* and the rocket is projected forwards.

Conservation of a component of linear momentum

If the total force acting on a system has zero component in a *fixed* direction, then, in any motion of the system, the component of the total linear momentum in that direction is conserved.

Conservation of linear momentum is an important property of a system and the sections that follow rely heavily upon it. Two examples of momentum conservation are as follows:

- The **solar system** is an example of an *isolated* system, being extremely remote from any other masses. It follows that the total linear momentum of the solar system is conserved. Thus the centre of mass of the solar system moves with constant velocity.
- On the other hand, a **grasshopper** trying to move on a perfectly smooth horizontal table is *not isolated*, being subject to gravity and the vertical reaction of the table. However, since the grasshopper is not subject to any external *horizontal* force, it follows that, whatever the grasshopper tries to do, his component of total linear momentum in any *horizontal* direction is conserved. His vertical component of linear momentum is not conserved; he can leap into the air if he wishes.

10.5 ROCKET MOTION

An important application of linear momentum conservation is **rocket propulsion**. Figure 10.1 shows a rigid body of mass M (the rocket) which contains a removable rigid block of mass m (the fuel). The system is at rest when an internal source of energy causes the fuel block to be ejected backwards with speed u *relative to the rocket*. If the system is isolated, then its **total linear momentum is conserved**, which implies that

$$Mv + m(v - u) = 0$$

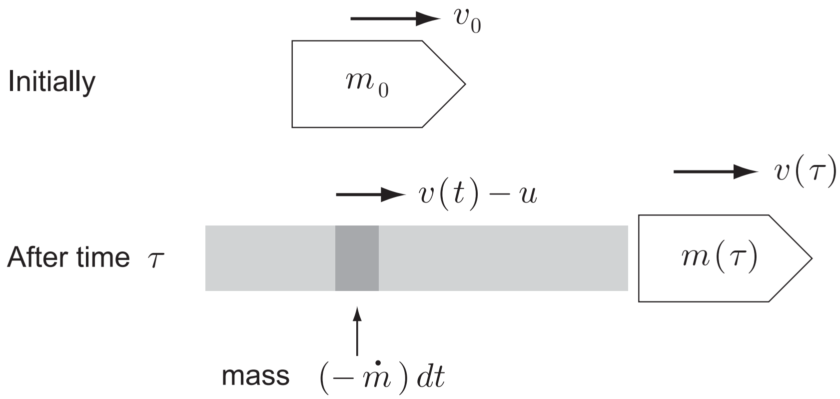


FIGURE 10.2 The rocket and its fuel at times $t = 0$ and $t = \tau$. The element of fuel ejected in the time interval $[t, t + dt]$ has mass $(-\dot{m})dt$ and (forward) velocity $v(t) - u$.

where v is the forward velocity of the rocket after the ejection of the fuel. As a result of this process the rocket acquires the forward velocity

$$v = \left(\frac{m}{M + m}\right) u.$$

This is the basic principle of **rocket propulsion**. The only mechanically significant difference between the simple example above and real rocket propulsion is that, in the case of the real rocket, the fuel mass is ejected continuously over a period of time and not in a single lump. In practice, the fuel is burned continuously and the combustion products eject themselves due to their rapid expansion.

Rocket motion in free space

Figure 10.2 shows a more realistic situation. Initially the rocket and its fuel have combined mass m_0 and are moving with constant velocity v_0. At time $t = 0$ the motors are started and fuel products are ejected backwards with speed u *relative to the rocket*. The fuel 'burn' continues for a time T, at the end of which the rocket and unburned fuel have mass m_1. Let $m = m(t)$ be the mass of the rocket and its unburned fuel after time t. Then m is a decreasing function of t and the rate of ejection of mass at time t is $-\dot{m}$. Let the system $\mathcal{S}$ consist of the rocket together with its fuel at time $t = 0$. After some time τ into the burn, the mass of $\mathcal{S}$ is distributed as shown in Figure 10.2. The rocket and unburned fuel have mass $m(\tau)$ and the remaining mass has been ejected as expended fuel. We will suppose that, once an element of fuel is ejected, it continues to move with the velocity it had at the instant of ejection.* Since we are assuming $\mathcal{S}$ to be an isolated system, its *total linear momentum is conserved*. The initial linear momentum in this one-dimensional problem is $m_0 v_0$ and the final linear momentum of the rocket and *unburned* fuel is $m(\tau)v(\tau)$,

* This assumption simplifies our derivation but, as we will see, it is not essential.

where $v\,(=\,v(t))$ is the velocity of the rocket at time t. It remains to take account of the linear momentum of the ejected fuel. Consider the element of fuel that was ejected in the time interval $[t, t+dt]$. This has mass $(-\dot{m}(t))\,dt$ and its forward velocity at the instant of ejection was $v(t)-u$. The linear momentum of this fuel element is therefore $(-\dot{m})(v-u)\,dt$ and the total linear momentum of the fuel expended in the time interval $[0, \tau]$ is

$$-\int_0^\tau \dot{m}(v-u)\,dt.$$

Linear momentum conservation for the system $\mathcal{S}$ therefore requires that

$$m_0 v_0 = m(\tau)v(\tau) - \int_0^\tau \dot{m}(v-u)\,dt,$$

which can be written in the form

$$\int_0^\tau \left[\frac{d}{dt}(mv) - \dot{m}(v-u)\right] dt = 0.$$

Since this equality must hold for *any choice* of τ during the burn, it follows that the integrand must be zero, that is

$$\frac{d}{dt}(mv) - \dot{m}(v-u) = 0$$

for $0 \le t \le T$. This simplifies to give

Rocket equation in free space

$$m\frac{dv}{dt} = (-\dot{m})u \qquad (10.8)$$

the **rocket equation**, which holds for $0 < t < T$. The rocket equation can be interpreted physically as the Second Law applied to a *system of variable mass** $m(t)$, namely the rocket and its *unburned* fuel. In this interpretation, the term on the right, $-\dot{m}u$, plays the rôle of force and is called the **thrust** supplied by the motors.

Note. In our derivation, we assumed that, once an element fuel is ejected, it continues to move with the velocity it had at the instant of ejection. This is equivalent to assuming that each element of ejected

* This terminology is undesirable since, in classical mechanics, a 'system' means a fixed set of masses (or, at the very least, fixed total mass). No standard mechanical principle applies to a 'system' whose total mass is changing with time.

fuel is isolated from other fuel and from the rocket. It clearly makes no difference to the momentum of the ejected fuel if momentum is exchanged between *elements of itself* so that this assumption is actually unnecessary. However we must retain the assumption that ejected fuel has no further interaction *with the rocket*. This seems likely to be true in free space, but whether it is true just after take off from solid ground is questionable.

Providing that the ejection speed u is constant, the rocket equation (10.8) can easily be solved for any mass ejection rate. On dividing through by m and integrating with respect to t, we obtain

$$\int dv = \int \frac{(-\dot{m})u}{m} dt = -u \int \frac{dm}{m} = -u \ln m + \text{constant}$$

and, on applying the initial condition $v = v_0$ when $t = 0$, we obtain

$$v(t) = v_0 + u \ln\left(\frac{m_0}{m(t)}\right).$$

This gives the **rocket velocity** at time t. In particular, at the end of the fuel burn, the rocket velocity has increased by

$$\Delta v = v_1 - v_0 = u \ln\left(\frac{m_0}{m_1}\right), \tag{10.9}$$

where m_1 and v_1 are the final mass and velocity of the rocket. One can make some interesting deductions from this solution.

(i) Δv, the increase in the rocket velocity, is directly proportional to u, the fuel ejection speed. Thus it pays to make u as large as possible. Chemical processes can produce values of u as high as 5000 m s^{-1}.

(ii) If the fuel were all ejected in a single lump, Δv would never exceed the ejection speed u. But when the fuel is ejected over a period of time, it is possible for the rocket to attain any velocity by making the mass ratio m_0/m_1 large enough. For example, if we wish to make $\Delta v = 3u$, then we need $m_0/m_1 = e^3 \approx 20$. This means that 19 kg of fuel would be required for every kilogram of payload. The amount of fuel needed to achieve higher velocities quickly makes the process impractical. To achieve $\Delta v = 10u$ takes 22 metric tons of fuel for every kilogram of payload!

Rocket motion under gravity

Suppose now that the rocket is moving vertically under gravity. If we regard the governing equation as the equation of motion for the variable mass $m(t)$, then, when gravity is introduced, the equation of motion becomes

Rocket equation including gravity

$$m\frac{dv}{dt} = (-\dot{m})u - mg \tag{10.10}$$

where v is measured vertically upwards, and the weight force mg means $m(t)g$. In this case, the effective force on the right is the sum of the **thrust** $(-\dot{m})u$ acting upwards and the **weight force** mg acting downwards.

When the gravity is uniform and the ejection speed u is constant, the new rocket equation (10.10) can also be solved easily for any mass ejection rate. On dividing through by m and integrating with respect to t, we obtain

$$\int dv = \int \left[\frac{(-\dot{m})u}{m} - g \right] dt = -u \ln m + gt + \text{constant}$$

and, on applying the initial condition $v = v_0$ when $t = 0$, we obtain

$$v(t) = v_0 + u \ln \left(\frac{m_0}{m(t)} \right) - gt.$$

This gives the **rocket velocity** at time t. In particular, at the end of the fuel burn, the rocket velocity has increased by

$$\Delta v = v_1 - v_0 = u \ln \left(\frac{m_0}{m_1} \right) - gT, \tag{10.11}$$

where m_1 and v_1 are the final mass and velocity of the rocket and T is the time taken to burn all the fuel.

It will be noticed that, if T is too large, then Δv will be negative, which is hardly possible for a rocket standing on the ground. The reason for this paradox is that, if the fuel is burned too slowly then the thrust will be less than the initial weight of the rocket, which will not take off until its weight has become less than the thrust. We will therefore assume that $(-\dot{m})u > mg$ at all times during the burn so that the rocket has positive upward acceleration and achieves its maximum speed when $t = T$. If the rocket starts from rest, it then follows that the **maximum speed** achieved is

$$v_{\max} = u \ln \left(\frac{m_0}{m_1} \right) - gT.$$

In this and the zero gravity case, the *distance* travelled during the burn depends on the functional form of $m(t)$.

10.6 COLLISION THEORY

Another important application of linear momentum conservation occurs when we have an isolated system of **two particles**, and one particle is in **collision** with the other.

Collision processes

It is important to understand the meaning of the term ‘collision’. Suppose that the mutual interaction between the two particles tends to zero as the distance between them tends to

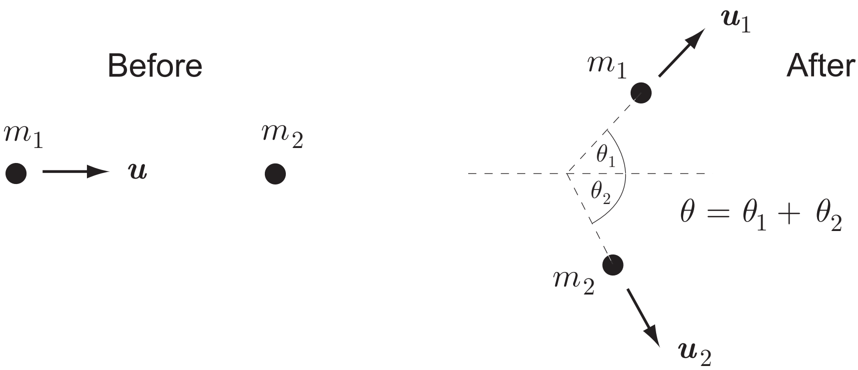


FIGURE 10.3 A collision between two particles viewed from the **laboratory frame**. A particle of mass m_1 and initial velocity $\boldsymbol{u}$ collides with a 'target' particle of mass m_2, which is initially at rest. After the collision, the particles have velocities $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ respectively. θ_1 is the **scattering angle** of the mass m_1, θ_2 is the **recoil angle** of the mass m_2, and θ $(= \theta_1 + \theta_2)$ is the **opening angle** between the emerging paths.

infinity, so that, if the particles are initially a great distance apart, each must be moving with constant velocity. If the particles approach each other, then there follows a period during which their mutual interaction causes their straight line motions to be disturbed. If the particles finally retreat to a great distance from each other, then they will move with constant velocities again, and these final velocities will generally be different to the initial velocities. This is what we mean by a **collision process**. Note that a collision process is not restricted to those cases in which the particles make physical contact with each other. This can of course happen, as in the 'real' collision of two pool balls. However, the deflection suffered by an alpha particle in passing close to a nucleus is also a 'collision', even though the alpha particle and the nucleus never made contact. Collision processes are particularly important in **nuclear** and **particle physics**, where they are the major source of experimental information.

General collisions

Consider the collision shown in Figure 10.3. A particle of mass m_1 and initial velocity $\boldsymbol{u}$ is incident upon a 'target' particle of mass m_2 which is initially at rest.* This is typical of the collisions observed in nuclear physics. After the collision we will suppose the particles retain their identities (and therefore their masses) and emerge with velocities $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ respectively. How are the final and initial motions of the particles related? Clearly we cannot 'solve the problem' since we have not even said what the mutual interaction between the particles is. However, it is surprising how much can be deduced simply from conservation laws without any detailed knowledge of the interaction.

Since the two particles form an isolated system, their total **linear momentum** is **conserved**, that is,

* This means 'at rest in the *laboratory reference frame*'.

$$\boxed{m_1\boldsymbol{u} = m_1\boldsymbol{u}_1 + m_2\boldsymbol{u}_2} \qquad (10.12)$$

This linear relation between the vectors $\boldsymbol{u}$, $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ implies that these three velocities must lie in the same plane so that scattering processes are *two-dimensional*.

Generally, collisions are **not energy preserving**. The energy principle for the collision has the form

$$\tfrac{1}{2}m_1u^2 + Q = \tfrac{1}{2}m_1u_1^2 + \tfrac{1}{2}m_2u_2^2,$$

where $u = |\boldsymbol{u}|$, $u_1 = |\boldsymbol{u}_1|$, $u_2 = |\boldsymbol{u}_2|$, and Q is the energy gained in the collision. In 'real' collisions between large bodies, energy is usually lost in the form of heat, so that Q is negative. However, in nuclear collisions in which the particles change their identities, it is perfectly possible for energy to be gained.

Example 10.4 *Making Kraptons*

A little known particle physicist has proposed the existence of a new particle, with charge +2 and mass 2, which he has named the Krapton. He has calculated that this can be produced by the collision of two protons in the reaction*

$$p^+ + p^+ + 10\text{ MeV} \to K^{++}$$

Having failed to obtain funding to verify his theory, he has built his own equipment with which he accelerates protons to an energy of 16 MeV and uses them to bombard a stationary target of hydrogen. Could he succeed in making a Krapton?

Solution

Suppose a proton with kinetic energy E collides with proton at rest. Then this system has initial linear momentum $(2mE)^{1/2}$, where m is the mass of a proton. This *linear momentum is preserved* by the collision so that, if a Krapton of mass $2m$ *were* produced, it would have linear momentum $(2mE)^{1/2}$ and therefore kinetic energy $E/2$. Hence, only 8 MeV of the initial energy is available for Krapton building and, according to the physicist's own calculation, this is not enough. (On the other hand, a head-on collision between two 5 MeV protons *would* be enough. Why?) ■

Elastic collisions

The linear momentum equation (10.12) holds whether the collision is between pool balls, protons or peaches. Much more can be said if the collision is also **energy preserving**.

Definition 10.2 ***Elastic collision*** *A collision between particles is said to be* ***elastic*** *if the* ***total kinetic energy*** *of the particles is* ***conserved*** *in the collision.*

---

* The electron volt (eV) is a unit of energy equal to 1.6×10^{-19} J approximately.

Frame invariance In order that the above definition be physically meaningful, it is necessary that a collision observed to be elastic in one inertial frame should also be elastic when observed from any other. This is not obviously true, since kinetic energy is not a linear quantity. However, since the total kinetic energy of the system can be written in the form $T = T^{CM} + T^G$ (see Theorem 9.10), where T^{CM} is preserved in the collision and T^G is frame independent, it follows that *any gain or loss of kinetic energy in the collision is independent of the inertial reference frame used to observe the event.*

Elastic collisions are very common and extremely important. For example, *any collision in which the mutual interaction force is conservative is elastic.* In particular, the collisions that occur in Rutherford scattering are elastic. In elastic collisions, we have **energy conservation** in the form

$$\boxed{\tfrac{1}{2}m_1u^2 = \tfrac{1}{2}m_1u_1^2 + \tfrac{1}{2}m_2u_2^2} \tag{10.13}$$

and, together with linear momentum conservation (10.12), we can make some interesting deductions. If we take the scalar product of each side of the linear momentum equation (10.12) with itself, we obtain

$$m_1^2u^2 = m_1^2u_1^2 + 2m_1m_2\boldsymbol{u}_1 \cdot \boldsymbol{u}_2 + m_2^2u_2^2,$$

and, if we now eliminate the term in u^2 between this equation and the energy conservation equation (10.13), we obtain, after simplification,

$$2m_1\boldsymbol{u}_1 \cdot \boldsymbol{u}_2 = (m_1 - m_2)\, u_2^2. \tag{10.14}$$

Since $\boldsymbol{u}_1 \cdot \boldsymbol{u}_2 = u_1u_2\cos\theta$, where θ is the **opening angle** between the paths of the emerging particles, the formula (10.14) can also be written

$$\cos\theta = \frac{(m_1 - m_2)\, u_2}{2m_1u_1}, \tag{10.15}$$

provided that $u_1 \neq 0$, that is, provided that the incident particle is not brought to rest by the collision.* This formula holds for **all elastic collisions**, whatever the nature of the particles and the interaction. It therefore applies equally well to pool balls† and protons, but not peaches. Given the mass ratio of the two particles, formula (10.15) relates the speeds of the particles and the opening angle between their paths after the collision.

* The incident particle can be brought to rest in a head-on collision with a particle of equal mass.

† Collisions between pool balls are very nearly elastic. However, in the present treatment, we are disregarding the rotation of the balls.

Example 10.5 ***Finding the final energies***

A ball of mass m and (kinetic) energy E is in an *elastic* collision with a second ball of mass $4m$ that is initially at rest. The two balls depart in directions making an angle of 120° with each other. What are the final energies of the two balls?

Solution

On substituting the given data into the formula (10.15), we find that $u_1/u_2 = 3$. It follows that

$$\frac{E_1}{E_2} = \frac{\frac{1}{2}mu_1^2}{\frac{1}{2}(4m)u_2^2} = \frac{1}{4}\left(\frac{u_1}{u_2}\right)^2 = \frac{9}{4}.$$

Hence $E_1 = \frac{9}{13}E$ and $E_2 = \frac{4}{13}E$. ■

An important special case occurs when the two particles have equal masses. In this case, formula (10.15) shows that the opening angle must always be a right angle. Thus, *in an elastic collision between particles of equal mass, the particles depart in directions at right angles*. Note that this result applies only when the target particle is initially at rest.

Example 10.6 ***Elastic collision between two electrons***

In an elastic collision between an electron with kinetic energy E and an electron at rest, the incoming electron is observed to be deflected through an angle of 30°. What are the energies of the two electrons after the collision?

Solution

Since the collision is elastic and the electrons have equal mass, the opening angle between the emerging paths must be 90°. The target electron must therefore recoil at an angle of 60° to the initial direction of the incoming electron. Let the speed of the incoming electron be u and speeds of the electrons after the collision be u_1 and u_2 respectively. Then conservation of linear momentum implies that

$$mu = mu_1 \cos 30° + mu_2 \cos 60°,$$
$$0 = mu_1 \sin 30° - mu_2 \sin 60°,$$

which gives $u_1 = \frac{1}{2}\sqrt{3}\,u$ and $u_2 = \frac{1}{2}u$. Hence, after the collision, the electrons have **energies** $\frac{3}{4}E$ and $\frac{1}{4}E$ respectively. ■

10.7 COLLISION PROCESSES IN THE ZERO-MOMENTUM FRAME

We have so far supposed that the inertial reference frame from which the scattering process is observed is the one occupied by the experimental observer. This is called the **laboratory frame** (or lab frame) since it is the frame in which measurements (of scattering angles, for instance) are actually taken. In the lab frame, the target particle is initially at rest.

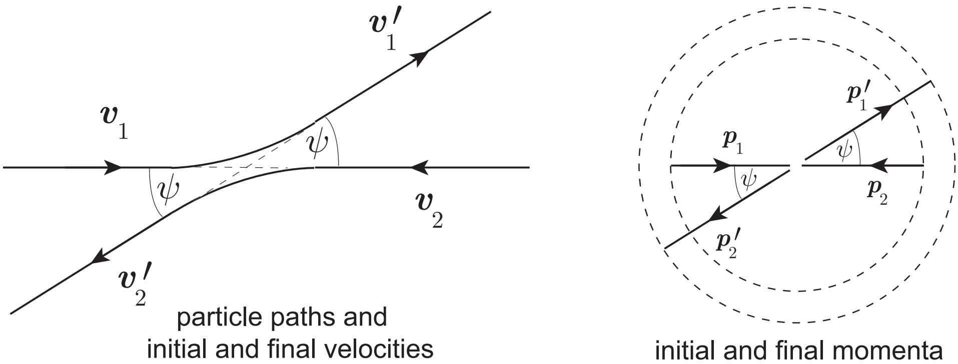


FIGURE 10.4 A collision between two particles viewed from the **zero-momentum frame**. The **initial momenta** $\boldsymbol{p}_1$, $\boldsymbol{p}_2$ are equal and opposite, as are the **final momenta** $\boldsymbol{p}'_1$, $\boldsymbol{p}'_2$. The angle ψ is the angle through which each of the masses is scattered.

However, it is very convenient to 'view' the scattering process from a different inertial frame. Since the two particles form an *isolated* system, their centre of mass G moves with constant velocity and so the frame* in which G is at rest is inertial. In this frame, the total linear momentum of the two particles is zero and, for this reason, we call it the **zero-momentum frame**† or **ZM frame**.

Consider, for example, the scattering problem which, in the lab frame, is shown in Figure 10.3. Then the total linear momentum $\boldsymbol{P} = m_1\boldsymbol{u}$ and the velocity $\boldsymbol{V}$ of the centre of mass of the two particles is

$$\boxed{\boldsymbol{V} = \frac{m_1\boldsymbol{u}}{m_1 + m_2}.} \tag{10.16}$$

This therefore is the **velocity of the ZM frame** relative to the lab frame for this collision process.

Collisions viewed from the ZM frame

Two-particle collisions look simple when viewed from the ZM frame. This is because, since the total linear momentum is now zero, the initial linear momenta $\boldsymbol{p}_1$, $\boldsymbol{p}_2$ of the two particles and the final momenta $\boldsymbol{p}'_1$, $\boldsymbol{p}'_2$ of the two particles must satisfy

$$\boldsymbol{p}_1 + \boldsymbol{p}_2 = \boldsymbol{0}, \qquad \boldsymbol{p}'_1 + \boldsymbol{p}'_2 = \boldsymbol{0}. \tag{10.17}$$

Thus, when a two-particle collision is **viewed from the ZM frame, the initial momenta are equal and opposite and so are the final momenta**. Figure 10.4 shows what a two-particle collision looks like when viewed from the ZM frame. Because of the relations

* This frame has the same velocity as G, and no rotation, relative to the lab frame.

† The term 'centre of mass frame' is also used. However, 'zero-momentum frame' is preferable since this notion holds good in relativistic mechanics.

(10.17), the particles both arrive and depart in opposite directions, so that **each particle is deflected through that same angle** ψ. All this follows solely from conservation of linear momentum. We can say more if we also have an energy principle of the form

$$\tfrac{1}{2}m_1|\boldsymbol{v}_1|^2 + \tfrac{1}{2}m_2|\boldsymbol{v}_2|^2 + Q = \tfrac{1}{2}m_1|\boldsymbol{v}_1'|^2 + \tfrac{1}{2}m_2|\boldsymbol{v}_2'|^2,$$

where $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\boldsymbol{v}_1'$, $\boldsymbol{v}_2'$ are the initial and final velocities of the particles (as shown in Figure 10.4), and Q is the kinetic energy gained as a result of the collision.* Let p be the common *magnitude* of the initial momenta $\boldsymbol{p}_1$, $\boldsymbol{p}_2$, and p' be the common *magnitude* of the final momenta $\boldsymbol{p}_1'$, $\boldsymbol{p}_2'$. Then the energy balance equation can be re-written in the form

$$\frac{p^2}{2m_1} + \frac{p^2}{2m_2} + Q = \frac{p'^2}{2m_1} + \frac{p'^2}{2m_2},$$

that is,

$$p'^2 = p^2 + \left(\frac{2Qm_1m_2}{m_1 + m_2}\right). \tag{10.18}$$

Thus the magnitudes of the initial and final momenta are related through Q, the energy gained in the collision. This is depicted in the momentum diagram in Figure 10.4. The magnitudes of the initial and final momenta (p and p') are represented by the radii of the two dashed circles. The diagram shows the case in which $p' > p$, which corresponds to $Q > 0$. For an **elastic collision**, the circles are coincident and **all four momenta have equal magnitudes**.

In a typical scattering problem, the masses m_1, m_2 and the initial momenta $\boldsymbol{p}_1$, $\boldsymbol{p}_2$ are known. For the scattering problem shown (in the lab frame) in Figure 10.3, $\boldsymbol{v}_1 = \boldsymbol{u} - \boldsymbol{V}$ and $\boldsymbol{v}_2 = -\boldsymbol{V}$, where $\boldsymbol{V}$ is given by equation (10.16). It follows that the initial momentum magnitude in the ZM frame are given by

$$p = \frac{m_1m_2u}{m_1 + m_2}, \tag{10.19}$$

where $u = |\boldsymbol{u}|$. The scattering process is now entirely determined by the parameters Q (the energy gain) and ψ (the ZM scattering angle). Given p and Q, p' is determined from equation (10.18). Together with ψ, this determines the final momenta $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$. The parameters Q and ψ depend on the physics of the actual collision. For instance, the collision may be known to be elastic, in which case $Q = 0$. The question of how the scattering angle ψ is related to the actual interaction and initial conditions is addressed in section 10.9.

Returning to the lab frame (elastic collisions only)

Although the scattering process looks simpler in the ZM frame, we usually need to know the details of the scattering actually observed by the experimenter in the lab frame. This

* As remarked earlier, Q is frame independent and so is the same as the energy gain measured in the lab frame.

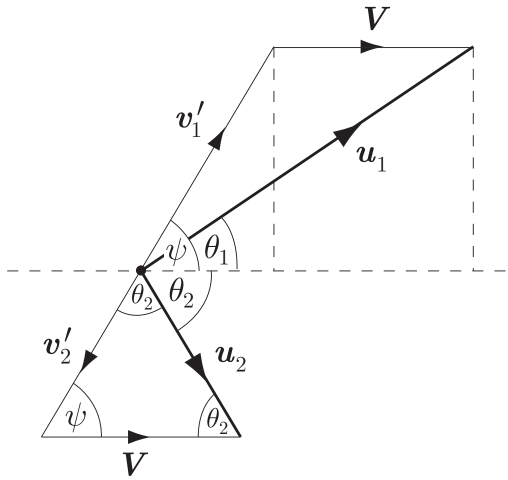


FIGURE 10.5 The final particle velocities $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ in the lab frame are obtained from the final velocities $\boldsymbol{v}_1'$, $\boldsymbol{v}_2'$ in the ZM frame by the relations $\boldsymbol{u}_1 = \boldsymbol{v}_1' + \boldsymbol{V}$, $\boldsymbol{u}_2 = \boldsymbol{v}_2' + \boldsymbol{V}$. The diagram shows the elastic case, in which the velocity triangle for $\boldsymbol{u}_2$ is isosceles.

entails transforming the properties of the final state (velocities, momenta and kinetic energies) from the ZM frame back to the lab frame. Since the ZM frame has velocity $\boldsymbol{V}$ (given by (10.16)) relative to the lab frame, the final velocities $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ observed in the lab frame are related to the final velocities $\boldsymbol{v}_1'$, $\boldsymbol{v}_2'$ in the ZM frame by

$$\boldsymbol{u}_1 = \boldsymbol{v}_1' + \boldsymbol{V}, \qquad \boldsymbol{u}_2 = \boldsymbol{v}_2' + \boldsymbol{V}. \tag{10.20}$$

Any other properties can then be found from $\boldsymbol{u}_1$, $\boldsymbol{u}_2$. The transformations (10.20) are depicted geometrically in Figure 10.5. The transformation formulae become rather complicated in the general case, but simplify nicely when the collision is elastic. From now on we will restrict ourselves to **elastic collisions only**. In this case, $Q = 0$ and the collisions are parametrised by ψ alone. The energy equation (10.18) then implies that $p' = p$ so that the four momentum magnitudes are equal, and given by equation (10.19). The final *speeds* of the particles in the ZM frame are therefore given by

$$v_1' = \frac{m_2 u}{m_1 + m_2}, \qquad v_2' = \frac{m_1 u}{m_1 + m_2} = V, \tag{10.21}$$

where $V = |\boldsymbol{V}|$. We may now deduce the required information from Figure 10.5. The lab **scattering angle** θ_1 can be expressed in terms of the parameter angle ψ by

$$\tan\theta_1 = \frac{v_1' \sin\psi}{v_1' \cos\psi + V} = \frac{\sin\psi}{\cos\psi + (V/v_1')} = \frac{\sin\psi}{\cos\psi + (m_1/m_2)},$$

on using equations (10.21). The lab **recoil angle** θ_2 is easily found since, in an elastic collision, the velocity triangle for $\boldsymbol{u}_2$ is isosceles with angles ψ, θ_2 and θ_2, as shown in Figure 10.5. It follows that

$$\theta_2 = \tfrac{1}{2}(\pi - \psi).$$

The expression for the lab **opening angle** θ $(= \theta_1 + \theta_2)$ is therefore given by

$$\tan\theta = \tan(\theta_1 + \theta_2) = \frac{\tan\theta_1 + \tan\theta_2}{1 + \tan\theta_1 \tan\theta_2} = \left(\frac{m_1 + m_2}{m_1 - m_2}\right)\cot(\tfrac{1}{2}\psi),$$

after some simplification.

To find the final energies, we observe that

$$u_2 = 2V\,\sin(\tfrac{1}{2}\psi)$$

so that E_2, the final lab **energy** of the mass m_2 is given by

$$\frac{E_2}{E_0} = \frac{\frac{1}{2}m_2\left(2V\,\sin(\frac{1}{2}\psi)\right)^2}{\frac{1}{2}m_1 u^2} = \frac{4m_1 m_2}{(m_1 + m_2)^2}\sin^2(\tfrac{1}{2}\psi),$$

where E_0 $(= \frac{1}{2}m_1 u^2)$ is the lab energy of the incident mass m_1. Since the collision is elastic, the final lab energy of the mass m_1 is simply deduced from the energy conservation formula $E_1 + E_2 = E_0$.

The above formulae give the properties of the final state following an elastic two-particle collision in terms of the ZM scattering angle ψ. We will call them the **elastic collision formulae** and they are summarised below:

Elastic collision formulae

A. $\tan\theta_1 = \dfrac{\sin\psi}{\cos\psi + \gamma}$ **B**. $\theta_2 = \frac{1}{2}(\pi - \psi)$

C. $\tan\theta = \left(\dfrac{\gamma + 1}{\gamma - 1}\right)\cot(\frac{1}{2}\psi)$ **D**. $\dfrac{E_2}{E_0} = \dfrac{4\gamma}{(\gamma + 1)^2}\sin^2(\frac{1}{2}\psi)$

(10.22)

ψ is the scattering angle in the ZM frame, and $\gamma = m_1/m_2$, the mass ratio of the two particles.

Using the elastic collision formulae A word of advice about the use of these formulae may be helpful. Most questions on this topic tell you some property of the scattering in the lab frame and ask you to find another property of the scattering in the lab frame; the ZM frame is never mentioned. *It is inadvisable to start manipulating the elastic scattering formulae*. This is almost guaranteed to cause errors. The simplest method is as follows: (i) Use the given data to find ψ by using the appropriate formula 'backwards', and then (ii) use this value of ψ to find the required scattering property. In short, the advice is '*go via* ψ'.

Example 10.7 ***Using the elastic scattering formulae***

In an experiment, particles of mass m and energy E are used to bombard stationary target particles of mass $2m$.

Q. The experimenters wish to select particles that, after scattering, have energy $E/3$. At what scattering angle will they find such particles?

A. If $E_1/E_0 = 1/3$, then by energy conservation $E_2/E_0 = 2/3$. First use formula **D** to find ψ. Since the mass ratio $\gamma = 1/2$, this gives

$$\frac{2}{3} = \frac{8}{9}\sin^2(\tfrac{1}{2}\psi),$$

so that $\psi = 120°$. Now use formula **A** to find the scattering angle θ_1. This gives $\tan\theta_1 = \infty$ so that $\theta_1 = 90°$. Particles scattered with energy $E/3$ will therefore be found emerging at right angles to the incident beam.

Q. In one collision, the opening angle was measured to be 45°. What were the individual scattering and recoil angles?

A. First use formula **C** to find ψ. This gives

$$\cot(\tfrac{1}{2}\psi) = \frac{1}{3},$$

so that $\frac{1}{2}\psi = 72°$, to the nearest degree. Now use formula **B** to find the recoil angle θ_2. This gives $\theta_2 = 90° - 72° = 18°$. The scattering angle θ_1 must therefore be $\theta_1 = \theta - \theta_2 = 45° - 18° = 27°$.

Q. In another collision, the scattering angle was measured to be 45°. What was the recoil angle?

A. First use formula **A** to find ψ. This shows that ψ satisfies the equation

$$2\cos\psi - 2\sin\psi = 1,$$

which can be written in the form*

$$\sqrt{8}\cos\left(\psi + 45°\right) = 1.$$

This gives $\psi = 24°$, to the nearest degree. Formula **B** now gives the recoil angle θ_2 to be $\theta_2 = 78°$, to the nearest degree.

10.8 THE TWO-BODY PROBLEM

The problem of determining the motion of two particles, moving solely under their mutual interaction, is called the **two-body problem**. Strictly speaking, all of the orbit

* Recall that equations of the form $a\cos\psi + b\sin\psi = c$ are solved by writing the left side in the 'polar form' $R\cos(\psi - \alpha)$, where $R^2 = a^2 + b^2$ and $\tan\alpha = b/a$.

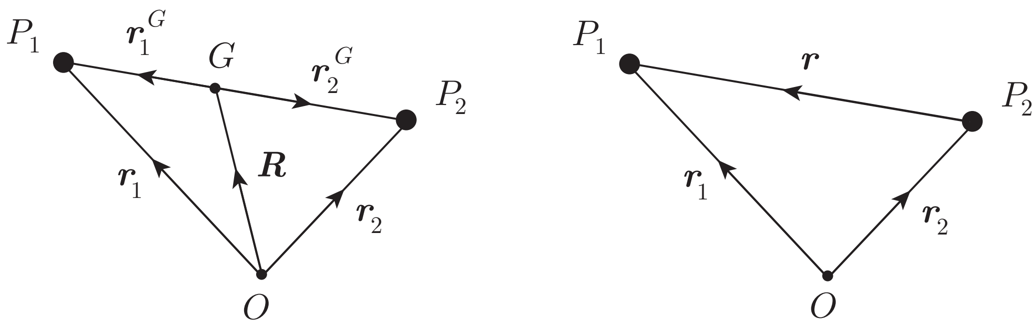


FIGURE 10.6 The motion of P_1 and P_2 relative to their centre of mass (left), and the motion of P_1 relative to P_2 (right).

problems considered in Chapter 7 should have been treated as two-body problems since centres of force are never actually fixed. The one-body theory is a good approximation when one particle is much more massive than the other. When the two particles have similar masses, the problem must be treated by two-body theory, in which neither particle is assumed to be fixed.

Let P_1 and P_2 be two particles moving under their mutual interaction. By the Third Law, the forces that they exert on each other are equal in magnitude, opposite in direction, and act along the line joining them. We will further suppose that the magnitude of these interaction forces depends only on r, the distance separating P_1 from P_2. The forces $\boldsymbol{F}_1$, $\boldsymbol{F}_2$, acting on P_1, P_2, then have the form

$$\boldsymbol{F}_1 = F(r)\,\widehat{\boldsymbol{r}}, \qquad\qquad \boldsymbol{F}_2 = -F(r)\,\widehat{\boldsymbol{r}},$$

where $\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2$, $r = |\boldsymbol{r}_1 - \boldsymbol{r}_2|$ and $\widehat{\boldsymbol{r}} = \boldsymbol{r}/r$ (see Figure 10.6). The **equations of motion** for P_1, P_2 are therefore

$$m_1\ddot{\boldsymbol{r}}_1 = F(r)\,\widehat{\boldsymbol{r}}, \qquad\qquad m_2\ddot{\boldsymbol{r}}_2 = -F(r)\,\widehat{\boldsymbol{r}}. \tag{10.23}$$

This is a generalisation of central force motion in which each particle moves under a force centred upon the other particle. Although this problem appears to be complicated, it can be quickly reduced to an **equivalent one-body problem**.

We first observe that the two particles form an isolated system so that their total linear momentum is conserved, or (equivalently) their centre of mass G moves with constant velocity. The motion of G is therefore determined from the initial conditions and it remains to find the motion of each particle *relative* to G, that is, their motions in the ZM frame. It turns out however that it is easier to find the motion of one particle *relative to the other*. The motion of each particle relative to G can then be easily deduced.

The equation of relative motion

It follows from the equations of motion (10.23) that

$$\ddot{\boldsymbol{r}}_1 - \ddot{\boldsymbol{r}}_2 = \frac{F(r)\,\widehat{\boldsymbol{r}}}{m_1} + \frac{F(r)\,\widehat{\boldsymbol{r}}}{m_2} = \left(\frac{m_1 + m_2}{m_1 m_2}\right) F(r)\,\widehat{\boldsymbol{r}},$$

so that $\boldsymbol{r}$, the position vector of P_1 relative to P_2, satisfies the equation

Relative motion equation

$$\left(\frac{m_1 m_2}{m_1 + m_2}\right)\ddot{\boldsymbol{r}} = F(r)\,\widehat{\boldsymbol{r}}, \tag{10.24}$$

which we call the **relative motion equation**.

Definition 10.3 ***Reduced mass*** *The quantity* μ*, defined by*

$$\mu = \frac{m_1 m_2}{m_1 + m_2}. \tag{10.25}$$

is called the ***reduced mass****.*

Our result can be expressed as follows:

Two-body problem – the relative motion

In the two-body problem, the motion of P_1 relative to P_2 is the same *as if* P_2 were held fixed and P_1 had the reduced mass μ instead of its actual mass m_1.

This rule* allows us to replace the problem of the motion of P_1 relative to P_2 by an **equivalent one-body problem** in which P_2 is fixed. The solution of such problems is fully described in Chapter 7.

Example 10.8 *Escape from a free gravitating body*

Two particles P_1 and P_2, with masses m_1 and m_2, can move freely under their mutual gravitation. Initially both particles are at rest and separated by a distance c. With what speed must P_1 be projected so as to escape from P_2?

Solution

Since this is a mutual gravitation problem, we take our rule in the form: *The motion of* P_1 *relative to* P_2 *is the same as if* P_2 *were held fixed and the constant of gravitation* G *replaced by* G'*, where*

$$G' = \left(\frac{m_1 + m_2}{m_2}\right) G.$$

* The rule is ambiguous when the force F also depends on m_1, as in mutual gravitation. Do you also replace *this* m_1 by μ? The answer is no, but the easiest way to avoid this glitch in the mutual gravitation problem is to make the transformation $G \to (m_1 + m_2)G/m_2$ instead. This has the correct effect and is not ambiguous.

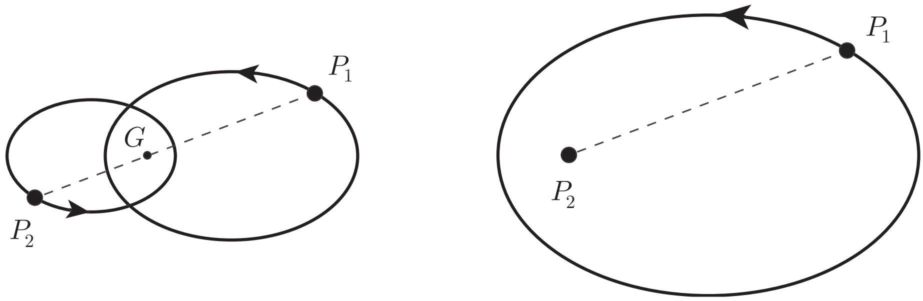


FIGURE 10.7 Particles P_1 and P_2 move under their mutual gravitation. In the zero momentum frame, the orbits are similar conics, each with a focus at G (left). The orbit of P_1 relative to P_2 is a third similar conic with P_2 at a focus (right).

From the one-body theory in Chapter 7, we know that P_1 will escape from a *fixed* P_2 if it has positive energy, that is if

$$\tfrac{1}{2}m_1V^2 - \frac{m_1m_2G}{c} \geq 0.$$

Hence, when P_2 is not fixed, P_1 will escape if

$$\tfrac{1}{2}m_1V^2 - \frac{m_1m_2G'}{c} \geq 0,$$

that is, if

$$V^2 \geq \frac{2(m_1+m_2)G}{c}.$$

This is the required **escape condition**. ■

Once the *relative* motion of the particles has been found, one may easily deduce the motion of each particle in the ZM frame since (see Figure 10.6)

$$\boldsymbol{r}_1^G = \left(\frac{m_2}{m_1+m_2}\right)\boldsymbol{r}, \qquad \boldsymbol{r}_2^G = -\left(\frac{m_1}{m_1+m_2}\right)\boldsymbol{r}.$$

It follows that the orbits of P_1, P_2 in the ZM frame are **geometrically similar** to the orbit in the relative motion. For instance, suppose that the mutual interaction of P_1 and P_2 is gravitational attraction, and that the orbit of P_1 relative to P_2 has been found to be an ellipse. Then the orbits of P_1 and P_2 in the ZM frame are *similar ellipses*, as shown in Figure 10.7. The ratio of the major axes of these orbits is $m_2 : m_1$, and the sum of their major axes is equal to the major axis of the orbit of P_1 relative to P_2. All three orbits have the same period τ given by

$$\tau^2 = \frac{4\pi^2a^3}{G\,(m_1+m_2)}, \tag{10.26}$$

where a is the semi-major axis of the *relative* orbit. This formula is simply obtained from the one-body period formula (7.26) by replacing G by G'.

Formula (10.26) shows that, in the **approximate treatment** in which P_2 is regarded as fixed, the value of the period is overestimated by the factor

$$\left(1 + \frac{m_1}{m_2}\right)^{1/2},$$

which is a small correction when m_1/m_2 is small. In the Solar system, the largest value of m_1/m_2 for a planetary orbit is that for Jupiter, which is about 1/1000.

Binary stars

It is probable that over half of the 'stars' in our galaxy are not single stars, like the Sun, but occur in pairs* that move under their mutual gravitation. Such a pair is called a **binary star**.

Binary stars are important in astronomy and also provide a nice application of our two-body theory. In particular, the two components of the binary must orbit their centre of mass on similar ellipses, as shown in Figure 10.7; the orbit of either component relative to the other is a third similar ellipse; and the period of all three motions is given by formula (10.26), where a is the semi-major axis of the *relative* orbit.

One reason why binary stars are important in astronomy is that the masses of their component stars can be found by direct measurement; indeed they are the only stars for which this can be done. Suppose that the star is an *optical* binary, which means that both components are visible through a suitably large telescope. Then the period of the binary can be measured by direct observation. It is also possible to measure the major axis of the relative orbit. Once τ and a are known, formula (10.26) tells us the **sum of the masses** of the two components of the binary.

Example 10.9 *Sirius A and B*

A typical example of a binary is **Sirius** in the constellation *Canis Major*, the brightest star in the night sky. The large bright component is called Sirius A and its small dim companion Sirius B. The period of their mutual orbital motion is 50 years and the value of a is 20 AU. (This is about the distance from the Sun to the planet Uranus.) Find the sum of the masses of the two components of Sirius.

Solution

In terms of astronomical units, in which $G = 4\pi^2$, formula (10.26) gives

$$M_A + M_B = \frac{20^3}{50^2} = 3.2\ M_\odot \blacksquare$$

* Groups of three or more also occur.

In order to determine the **individual masses** of the components by optical means, it is necessary to find the a values for one of the individual components in its motion *relative to the centre of mass*. The procedure is essentially the same as before, but much more difficult observationally since the motion of the chosen component must be measured absolutely, that is, relative to background stars. In the case of Sirius, it is found that $M_A = 2.1\ M_\odot$ and $M_B = 1.1\ M_\odot$.

10.9 TWO-BODY SCATTERING

An important application of two-body motion is the **two-body scattering** problem. In our treatment of collision theory, we considered the whole class of possible collisions between two particles that were consistent with momentum and energy conservation. These collisions were parametrised by the ZM scattering angle ψ. We now consider the problem in more detail. Given the interaction between the particles and the impact parameter p, what *is* the resulting ZM scattering angle? This question can be answered by using two-body theory. We break up the process into a number of steps:

1. Find the $\{p, \theta\}$–relation for the one-body problem

First consider the **one-body problem** in which the particle P_2 is held *fixed*, and work out (or look up) the relation between the impact parameter p and the scattering angle θ. For example, the $\{p, \theta\}$–relation for **Rutherford scattering** was derived in Chapter 7 and was found to be

$$\tan \tfrac{1}{2}\theta = \frac{q_1 q_2}{m_1 p u^2}. \qquad (10.27)$$

2. Find the $\{p, \phi\}$–relation for the relative motion problem

The next step is to find the relation between the impact parameter p and the scattering angle ϕ observed in the **relative motion problem**. This is easily obtained from the one-body formula (10.27) by replacing m_1 by μ (the reduced mass) and replacing θ by ϕ. This gives

$$\tan \tfrac{1}{2}\phi = \frac{q_1 q_2 (1 + \gamma)}{m_1 p u^2}, \qquad (10.28)$$

where γ $(= m_1/m_2)$ is the ratio of the two masses.

3. Find the $\{p, \psi\}$–relation observed in the ZM frame

The angle ϕ that appears in the formula (10.28) is the scattering angle in the motion of m_1 *relative to* m_2. However, by an amazing stroke of good fortune, it is actually the same angle as the ZM scattering angle ψ that we used in collision theory.* Hence, the

* The reason is as follows: The relative motion in the lab frame must be the same as the relative motion in the ZM frame. In this frame, the initial relative velocity of P_1 is equal to $(\boldsymbol{p}_1/m_1) - (\boldsymbol{p}_2/m_2)$, which has the same *direction* as $\boldsymbol{p}_1$. Likewise, the final relative velocity of P_1 is equal to $(\boldsymbol{p}_1'/m_1) - (\boldsymbol{p}_2'/m_2)$, which has the same *direction* as $\boldsymbol{p}_1'$. Hence the scattering angle in the relative motion is the same as that in the ZM frame.

$\{p, \psi\}$–relation when **two-body scattering** is observed from the **ZM frame** is obtained by simply replacing ϕ in formula (10.28) by ψ, that is,

$$\tan \tfrac{1}{2}\psi = \frac{q_1 q_2 (1+\gamma)}{m_1 p u^2}. \tag{10.29}$$

As always, u means the speed of the incident particle observed in the lab frame.

4. Find θ_1 and θ_2 in terms of p from the elastic collision formulae

Since the $\{p, \psi\}$–relation (10.29) gives the ZM scattering angle ψ in terms of p, this expression for ψ can now be substituted into the elastic scattering formulae (10.22**A**) and (10.22**B**) to give expressions for the **two-body scattering angle** θ_1, and **recoil angle** θ_2 in terms of p. For Rutherford scattering, this gives, after some simplification,

Two-body Rutherford scattering formulae

$$\tan \theta_1 = \frac{4 q_1 q_2 p E}{4 p^2 E^2 - (1-\gamma^2) q_1^2 q_2^2} \qquad \tan \theta_2 = \frac{2 p E}{q_1 q_2 (1+\gamma)} \tag{10.30}$$

where $E\ (= \frac{1}{2} m_1 u^2)$ is the energy of the incident particle and $\gamma = m_1/m_2$.

These formulae simplify further when the particles have **equal masses**. In this special case, $\gamma = 1$ and the scattering and recoil angles are given by

$$\tan \theta_1 = \frac{q_1 q_2}{pE}, \qquad \tan \theta_2 = \frac{pE}{q_1 q_2}. \tag{10.31}$$

(As expected, $\theta_1 + \theta_2 = \frac{1}{2}\pi$.) These formulae would apply, for example, to the scattering of alpha particles by helium nucei.

Two-body scattering cross section

Having found the $\{p, \theta_1\}$–relation for the two-body scattering problem, the **two-body scattering cross section** σ^{TB} is given, in principle, by the formula

$$\sigma^{TB}(\theta_1) = -\frac{p}{\sin \theta_1} \frac{dp}{d\theta_1}.$$

However, this requires that the $\{p, \theta_1\}$–relation be solved to give p as a function of θ_1 and the resulting algebra is formidable.

The following method has the advantage that σ^{TB} is determined directly from the corresponding one-body scattering cross section. The trick is to introduce the ZM scattering angle ψ. By the chain rule,

$$\frac{dp}{d\theta_1} = \frac{dp}{d\psi} \times \frac{d\psi}{d\theta_1},$$

and so σ^{TB} can be written

$$\begin{aligned}\sigma^{TB} &= -\frac{p}{\sin\theta_1}\frac{dp}{d\theta_1} = -\frac{p}{\sin\theta_1}\left(\frac{dp}{d\psi}\times\frac{d\psi}{d\theta_1}\right)\\ &= \left(\frac{\sin\psi}{\sin\theta_1}\right)\left(\frac{d\psi}{d\theta_1}\right)\left(-\frac{p}{\sin\psi}\frac{dp}{d\psi}\right)\\ &= \left(\frac{\sin\psi}{\sin\theta_1}\right)\left(\frac{d\psi}{d\theta_1}\right)\sigma^{ZM}(\psi),\end{aligned}$$

where σ^{ZM} is defined by

$$\sigma^{ZM}(\psi) = -\frac{p}{\sin\psi}\frac{dp}{d\psi}. \tag{10.32}$$

Now $\sigma^{ZM}(\psi)$ is easily obtained from the one-body cross-section $\sigma(\theta)$ by replacing m_1 by μ and θ by ψ. The two-body cross section is then given by

Two-body scattering cross section

$$\sigma^{TB}(\theta_1) = \left(\frac{\sin\psi}{\sin\theta_1}\right)\left(\frac{d\psi}{d\theta_1}\right)\sigma^{ZM}(\psi) \tag{10.33}$$

In this formula, we have yet to replace ψ by its expression in terms of θ_1. To do this, we must invert the formula (10.22 **A**) to obtain ψ as a function of θ_1. Formula (10.22 **A**) can be rearranged in the form

$$\sin(\psi - \theta_1) = \gamma\sin\theta_1,$$

from which we obtain

$$\psi = \theta_1 + \sin^{-1}(\gamma\sin\theta_1) \tag{10.34}$$

and, by differentiation,

$$\frac{d\psi}{d\theta_1} = 1 + \frac{\gamma\cos\theta_1}{(1-\gamma^2\sin^2\theta_1)^{1/2}}. \tag{10.35}$$

These expressions for ψ and $d\psi/d\theta_1$ in terms of θ_1 must now be substituted into equation (10.33) to obtain the final formula for the **two-body scattering cross section** $\sigma^{TB}(\theta_1)$. These operations can be done with computer assistance.

For example, in **Rutherford scattering**, we first obtain $\sigma^{ZM}(\psi)$ by replacing m_1 by μ (and θ by ψ) in the one-body cross section formula (7.37) obtained in Chapter 7. This gives

$$\sigma^{ZM}(\psi) = \frac{q_1^2 q_2^2(1+\gamma)^2}{4m_1^2u^4}\left(\frac{1}{\sin^4\frac{1}{2}\psi}\right). \tag{10.36}$$

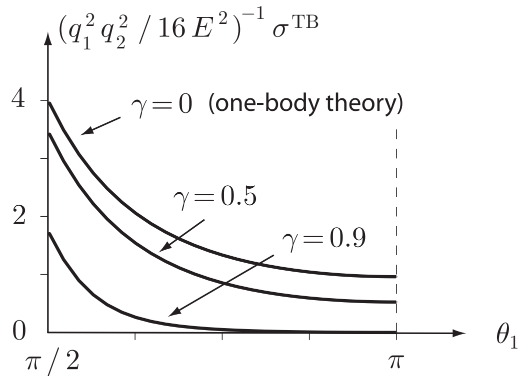


FIGURE 10.8 The **Rutherford two-body scattering cross section** σ^{TB} plotted against the scattering angle θ_1 $(\pi/2 \le \theta_1 \le \pi)$ for various values of the mass ratio γ $(= m_1/m_2)$. E is the kinetic energy of the incident particles.

The **two-body Rutherford scattering cross section** is now obtained by substituting the expression (10.36) into the general formula (10.33) and then replacing ψ and $d\psi/d\theta_1$ by the expressions (10.34), (10.35). After much manipulation, the answer is found to be

$$\sigma^{TB} = \frac{q_1^2 q_2^2}{16E^2}\left(\frac{4(1+\gamma)^2(\gamma\cos\theta_1 + S)^2}{S(1+\gamma\sin^2\theta_1 - \cos\theta_1 S)^2}\right), \tag{10.37}$$

where

$$S = \left(1 - \gamma^2\sin^2\theta_1\right)^{1/2}$$

and E $(= \frac{1}{2}m_1 u^2)$ is the energy of the incident particle.

Figure 10.8 shows graphs of $\sigma^{TB}(\theta_1)$ in Rutherford scattering for various choices of the mass ratio γ. In Rutherford's actual experiment with alpha particles and gold nucleii, the value of γ was about 0.02 and the error in the scattering cross section caused by using the one-body theory was less than 0.1%. However, as the graphs show, larger values of γ can give rise to a substantial deviation from the one-body theory.

When the mass ratio γ $(= m_1/m_2)$ is *small*, the formula (10.37) is approximated by

$$\sigma^{TB} = \frac{q_1^2 q_2^2}{16E^2}\left(\frac{1}{\sin^4(\theta_1/2)} - 2\gamma^2 + O\left(\gamma^4\right)\right).$$

Thus, when γ is small, the leading correction to the one-body approximation is a constant.

Equal masses

The whole process of finding σ^{TB} simplifies wonderfully when the two particles have equal masses. In this case, $\psi = 2\theta_1$, $d\psi/d\theta_1 = 2$, and the general formula (10.33) becomes

$$\sigma^{TB}(\theta_1) = 4\cos\theta_1\,\sigma^{ZM}(2\theta_1) \qquad (0 < \theta_1 \le \pi/2).$$

For example, in Rutherford scattering where the particles have equal masses, σ^{TB} has the simple form

$$\sigma^{TB}(\theta_1) = \frac{q_1^2 q_2^2}{E^2}\left(\frac{\cos\theta_1}{\sin^4\theta_1}\right) \qquad (0 \le \theta_1 \le \pi/2).$$

This formula would apply, for example, to the scattering of protons by protons.

10.10 INTEGRABLE MECHANICAL SYSTEMS

A mechanical system is said to be **integrable** if *its equations of motion are soluble in the sense that they can be reduced to integrations*.* The most important class of integrable systems are those that satisfy as many conservation principles as they have degrees of freedom. Suppose that a mechanical system $\mathcal{S}$ has n degrees of freedom and that it satisfies n conservation principles. Then it is certainly true that the n conservation equations are sufficient to *determine* the motion of the system, in the sense that no more equations are needed. More importantly though, it can be shown† that *these equations can always be reduced to integrations*. The system $\mathcal{S}$ is therefore **integrable**.

Before we can apply this method to particular systems, there is a **kinematical problem** to be overcome, namely: how does one find the velocities (and angular velocities) of the elements‡ of $\mathcal{S}$ when there are two or more generalised coordinates which vary simultaneously? The answer is by drawing a **velocity diagram** for $\mathcal{S}$ as described below:

Drawing a velocity diagram

- Draw the system in general position and select a set of generalised coordinates.
- Let the first generalised coordinate vary (with the other coordinates held constant) and mark in the velocity of each element.
- Now let the second generalised coordinate vary (with the other coordinates held constant) and, on the same diagram, mark in the velocity of each element. Continue in this way through all the generalised coordinates.
- Then, when all the generalised coordinates are varying simultaneously, the **velocity** of each element of $\mathcal{S}$ is the vector sum of the velocities given to that element when the coordinates vary individually.

In the above, 'velocity' means 'velocity and/or angular velocity'.

---

* The system is still said to be integrable even when the integrals cannot be evaluated in terms of standard functions!

† This is Liouville's theorem on integrable systems (see Problem 14.15)

‡ The elements of $\mathcal{S}$ are the particles and/or rigid bodies of which $\mathcal{S}$ is made up. One needs to find (i) the velocity of each particle, (ii) the velocity of the centre of mass of each rigid body, and (iii) the angular velocity of each rigid body, in each case in terms of the chosen coordinates and their time derivatives.

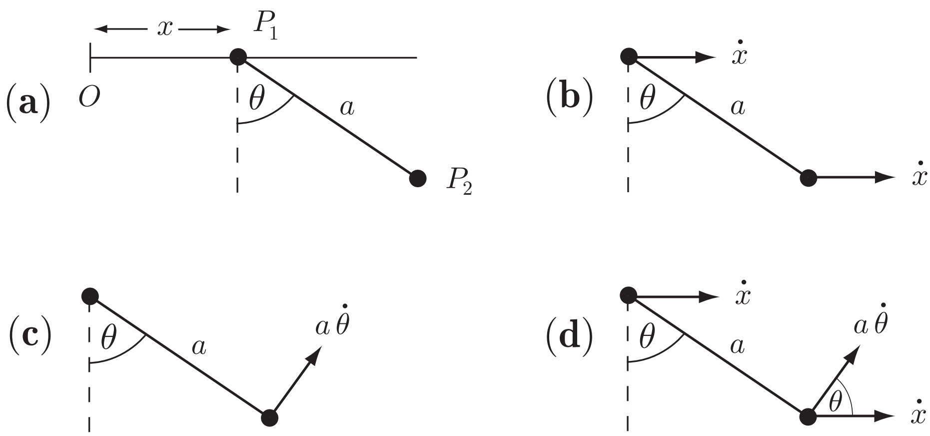


FIGURE 10.9 Constructing a **velocity diagram**. Figure (a) shows the system and the coordinates x and θ. Figure (b) shows the velocities generated when x varies with θ held constant. Figure (c) shows the velocities generated when θ varies with x held constant. Figure (d) is the **velocity diagram** which is formed by superposing the velocities in diagrams (b) and (c). Note that the velocity of P_2 is the *vector* sum of the two contributions shown.

Example 10.10 *Drawing a velocity diagram 1*

The system shown in Figure 10.9 consists of two particles P_1 and P_2 connected by a light inextensible string of length a. The particle P_1 is also constrained to move along a fixed horizontal rail and the whole system moves in the vertical plane through the rail. Take the variables x and θ shown as generalised coordinates and draw the velocity diagram.

Solution

The construction of the velocity diagram is shown in Figure 10.9. ■

Example 10.11 *Drawing a velocity diagram 2*

Two rigid rods CD and DE, of lengths $2a$ and $2b$, are flexibly jointed at D and can move freely on a horizontal table. Choose generalised coordinates and draw a velocity diagram for this system.

Solution

Let Oxy be a system of Cartesian coordinates in the plane of the table. Let (X, Y) be the Cartesian coordinates of the centre of the rod CD, and let θ and ϕ be the angles that the two rods make with positive x-axis. Then X, Y, θ, ϕ are a set of generalised coordinates for this system. These coordinates, and the corresponding velocity diagram are shown in Figure 10.10. There are *four* contributions to the velocity of the centre of the rod DE. Also, each rod has an angular velocity. ■

We will now solve the system shown in Figure 10.9 by using conservation principles.

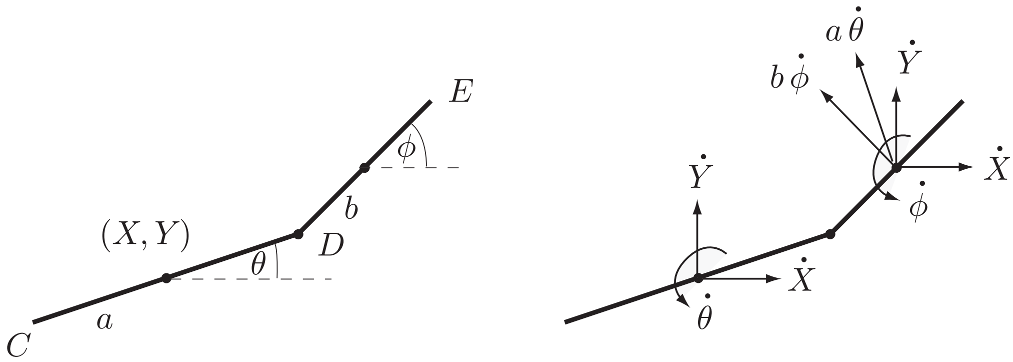


FIGURE 10.10 The **velocity diagram** for a system with four degrees of freedom. The figure on the left shows the system and the generalised coordinates X, Y, θ, ϕ. The figure on the right is the completed **velocity diagram**.

Example 10.12 *Solving an integrable system*

Consider the system shown in Figure 10.9 for the case in which P_1 and P_2 have masses $3m$ and m, the rail is smooth, and the system moves under uniform gravity. Initially, the system is released from rest with the string making an angle of $\pi/3$ with the downward vertical. Use conservation principles to obtain two equations for the subsequent motion.

Solution

Let $\boldsymbol{i}$ be the unit vector parallel to the rail (in the direction of increasing x). Since the rail is smooth, all the *external* forces on the system are vertical which means that $\boldsymbol{F} \cdot \boldsymbol{i} = 0$. This implies that $\boldsymbol{P} \cdot \boldsymbol{i}$, the *horizontal component of the total linear momentum, is conserved*. From the velocity diagram, the value of $\boldsymbol{P} \cdot \boldsymbol{i}$ at time t is given by

$$\boldsymbol{P} \cdot \boldsymbol{i} = 3m\dot{x} + m\left(\dot{x} + (a\dot{\theta})\cos\theta\right) = 4m\dot{x} + ma\dot{\theta}\cos\theta.$$

Also, since the motion is started from rest, $\boldsymbol{P} \cdot \boldsymbol{i} = 0$ initially. Hence, **conservation** of $\boldsymbol{P} \cdot \boldsymbol{i}$ implies that

$$4\dot{x} + a\dot{\theta}\cos\theta = 0, \qquad (10.38)$$

on cancelling by m. This is our first equation for the subsequent motion.

Since the rail is smooth, the constraint force exerted by the rail does no work and the tensions in the inextensible string do no total work. Hence **energy is conserved**.

From the velocity diagram, the kinetic energy of the system at time t is given by*

$$\begin{aligned} T &= \tfrac{1}{2}(3m)\dot{x}^2 + \tfrac{1}{2}m\left(\dot{x}^2 + (a\dot{\theta})^2 + 2\dot{x}(a\dot{\theta})\cos\theta\right) \\ &= \tfrac{1}{2}m\left(4\dot{x}^2 + a^2\dot{\theta}^2 + 2a\dot{x}\dot{\theta}\cos\theta\right). \end{aligned}$$

The gravitational potential energy of the system at time t is given by

$$V = 0 - mga\cos\theta.$$

Since the system was released from rest with $\theta = 60°$, the initial value of T is zero, while the initial value of $V = -\tfrac{1}{2}mga$. Hence, **conservation of energy** implies that

$$\tfrac{1}{2}m\left(4\dot{x}^2 + a^2\dot{\theta}^2 + 2a\dot{x}\dot{\theta}\cos\theta\right) - mga\cos\theta = -\tfrac{1}{2}mga,$$

which simplifies to give

$$4\dot{x}^2 + a^2\dot{\theta}^2 + 2a\dot{x}\dot{\theta}\cos\theta = ga\,(2\cos\theta - 1)\,. \tag{10.39}$$

This is our second equation for the subsequent motion.

Since this system has *two* degrees of freedom and satisfies *two* conservation principles, it must be **integrable**. Hence, the conservation equations (10.38), (10.39) must be soluble in the sense described above. ■

Question *Equation for* θ

Deduce an equation satisfied by θ alone and find the speeds of P_1 and P_2 when the string becomes vertical.

Answer

From the linear momentum equation (10.38),

$$\dot{x} = -\tfrac{1}{4}a\dot{\theta}\cos\theta$$

and, if we now eliminate $\dot{x}$ from the energy equation (10.39), we obtain, after simplification,

$$\dot{\theta}^2 = \frac{4g}{a}\left(\frac{2\cos\theta - 1}{4 - \cos^2\theta}\right), \tag{10.40}$$

which is an equation for θ alone.

It follows from this equation that, when the string becomes vertical (that is, when $\theta = 0$), $\dot{\theta}^2 = 4g/3a$. Hence, at this instant, $\dot{\theta} = -(4g/3a)^{1/2}$ and (from the momentum conservation equation) $\dot{x} = +(g/12a)^{1/2}$. Hence, the **speed** of P_1 is $(ag/12)^{1/2}$ and the **speed** of P_2 is $|\dot{x} + a\dot{\theta}| = (3ag/4)^{1/2}$. ■

---

* Suppose a velocity $\boldsymbol{V}$ is the sum of two contributions, $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$, so that $\boldsymbol{V} = \boldsymbol{v}_1 + \boldsymbol{v}_2$. Then

$$|\boldsymbol{V}|^2 = \boldsymbol{V}\cdot\boldsymbol{V} = (\boldsymbol{v}_1 + \boldsymbol{v}_2)\cdot(\boldsymbol{v}_1 + \boldsymbol{v}_2) = \boldsymbol{v}_1\cdot\boldsymbol{v}_1 + \boldsymbol{v}_2\cdot\boldsymbol{v}_2 + 2\boldsymbol{v}_1\cdot\boldsymbol{v}_2 = |\boldsymbol{v}_1|^2 + |\boldsymbol{v}_2|^2 + 2\boldsymbol{v}_1\cdot\boldsymbol{v}_2.$$

This formula was used to find the kinetic energy of particle P_2.

Question ***Period of oscillation***

Find the period of oscillation of the system.

Answer

From the equation (10.40), it follows that the motion is restricted to those values of θ that make the right side *positive*, and that $\dot{\theta} = 0$ when the right side is zero. Hence, θ oscillates periodically in the range $-\pi/3 < \theta < \pi/3$. Consider the first half-oscilliation. In this part of the motion, $\dot{\theta} < 0$ and so θ satisfies the equation

$$\dot{\theta} = -\left(\frac{4g}{a}\right)^{1/2}\left(\frac{2\cos\theta - 1}{4 - \cos^2\theta}\right)^{1/2},$$

a first order separable ODE. On separating, we find that τ, the **period** of a full oscillation, is given by

$$\tau = \left(\frac{a}{g}\right)^{1/2}\int_{-\pi/3}^{\pi/3}\left(\frac{4 - \cos^2\theta}{2\cos\theta - 1}\right)^{1/2} \approx 6.23\left(\frac{a}{g}\right)^{1/2}.$$

Thus the determination of $\theta(t)$ has been reduced to an integration, and, with $\theta(t)$ 'known', the equation(10.38) can be solved to give $x(t)$ as an integral. This confirms that the system is **integrable**. ■

Appendix A Modelling bodies by particles

When can a large body, such as a tennis ball, a spacecraft, or the Earth, be modelled by a particle?

The answer commonly given is that '*a body may be modelled by a particle if its size is small compared with the extent of its motion*'. For example, since the radius of the Earth is small compared with the radius of its solar orbit, it is argued that the Earth may be modelled by a particle, at least in respect of its translational motion. This argument sounds reasonable enough, but it is derived only from intuition and, although it often gives the correct answer, it is not the correct condition at all!

We can make some more definite statements on this quite tricky question by using the **centre of mass equation**. This states that '*the centre of mass of any system moves as if it were a particle of mass the total mass, and all the external forces acted upon it*'. It might appear that this principle enables us to predict the motion of the centre of mass of any system, but this is not so. The reason is that, in general, the *total external force acting on a system does not depend solely on the motion of its centre of mass*; it may depend on the positions of the *individual* particles and also other factors such as the particle velocities. Suppose, for example, that the system is a rigid body of *general shape* moving under the gravitational attraction of a fixed mass. Then the total gravitational force acting on the body is only *approximately* given by supposing all the mass to be concentrated at the centre of mass G. The exact force depends on the *orientation* of the body as well as the position of G. The centre of mass equation tells us nothing about this orientation and so the total force on the body is not known and the motion of G cannot be determined.

There are however some important exceptions:

- Consider a **rigid body** moving **without rotation**. In this case the motion of G determines the motion of every particle of the body. Then the total external force on the body is known and the motion of G can be determined. This, in turn, determines the motion of the whole body. For example, the problem of a block sliding without rotation on a table can be completely solved by particle mechanics.
- Consider any system moving solely under **uniform gravity**. In this case, the total external force on the system is a known constant and the motion of G can be determined. This does not however determine the motion of the individual particles. For example, if the system were a brick thrown through the air, then particle mechanics can calculate exactly where its centre of mass will go, but not which particle of the brick will hit the ground first.

In the general case however, we must use approximations. For example, suppose that the particles of the system move in the force fields $\boldsymbol{F}_i(\boldsymbol{r})$ so that the total force on the system is

$$\sum_{i=1}^{N} \boldsymbol{F}_i(\boldsymbol{r}_i).$$

In general, this is not equal to $\sum \boldsymbol{F}_i(\boldsymbol{R})$. We can however *approximate* $\sum \boldsymbol{F}_i(\boldsymbol{r}_i)$ by $\sum \boldsymbol{F}_i(\boldsymbol{R})$, in which case we are assuming that the ratio

$$\frac{|\boldsymbol{F}_i(\boldsymbol{r}_i) - \boldsymbol{F}_i(\boldsymbol{R})|}{|\boldsymbol{F}_i(\boldsymbol{R})|} \ll 1 \tag{10.41}$$

for all i. In the following argument we investigate when this condition can be expected to be hold.

Let $\boldsymbol{\delta}_i$ be the position vector of the particle P_i of the system *relative to* G. Then

$$\boldsymbol{F}_i(\boldsymbol{R} + \boldsymbol{\delta}_i) - \boldsymbol{F}_i(\boldsymbol{R}) = \left(\left. \frac{d\boldsymbol{F}_i}{ds} \right|_{\boldsymbol{r}=\boldsymbol{R}} \right) |\boldsymbol{\delta}_i| + O\left(|\boldsymbol{\delta}_i|^2\right),$$

where $d\boldsymbol{F}_i/ds$ means the *directional derivative* of $\boldsymbol{F}_i$ in the direction of the displacement $\boldsymbol{\delta}_i$. The condition (10.41) therefore requires that

$$\frac{\left(\left. \frac{d\boldsymbol{F}_i}{ds} \right|_{\boldsymbol{r}=\boldsymbol{R}} \right) |\boldsymbol{\delta}_i|}{|\boldsymbol{F}_i(\boldsymbol{R})|} \ll 1$$

for all i and for all values of $\boldsymbol{R}$ that are attained in the motion of the system. This will hold if

$$\Delta \ll \frac{|\boldsymbol{F}_i(\boldsymbol{R})|}{\max \left(\left. \frac{d\boldsymbol{F}_i}{ds} \right|_{\boldsymbol{r}=\boldsymbol{R}} \right)} \tag{10.42}$$

for all i and for all points on the path of the centre of mass. Here 'max' means the maximum over all directions, and Δ is the 'radius' of the system (the maximum distance of any particle of the system from the centre of mass). Thus the radius of the system is required to be small compared with the quantities above, not the lateral extent of the motion.

Although the condition (10.42) looks formidable, its physical meaning is quite simple: *the radius of the system is required to be small compared with a length scale over which any of the force fields vary significantly*.

Consider for example a body moving under the gravitational attraction of a mass M_0 which is fixed at the origin O. In this field, the particle P_i of the system moves under the force field

$$\boldsymbol{F}_i(\boldsymbol{r}) = -\frac{m_i M_0 G}{r_i^2}\widehat{\boldsymbol{r}}_i.$$

For this field, the right side of the condition (10.42) evaluates to give $R/2$, where $R\ (= |\boldsymbol{R}|)$ is the distance of the centre of mass of the body from O. Therefore the total gravitational force on the body will be accurately approximated by the force

$$\boldsymbol{F}(\boldsymbol{R}) = -\frac{M M_0 G}{R^2}\widehat{\boldsymbol{R}}$$

(where M is the total mass of the body), if $\Delta \ll R$ at each point on the path of the centre of mass. This means that the *radius of the body must be small compared with its distance of closest approach to the centre O*. This condition has no direct connection with the 'extent of the motion'. Indeed, on a hyperbolic orbit, the path is infinite, but the condition (10.42) will not hold if the path passes too close to the centre of force. Similar remarks apply to motion of a body in any central field governed by a *power* law. If however the field were that corresponding to the Yukawa potential

$$V = -k\frac{e^{-r/a}}{r},$$

where k, a are positive constants, then Δ is required to be small compared with the length scale a as well as the distance of closest approach to O.

Problems on Chapter 10

Answers and comments are at the end of the book.

Harder problems carry a star (∗).

Linear momentum principle & centre of mass equation

10.1 Show that, if a system moves from one state of rest to another over a certain time interval, then the average of the total external force over this time interval must be zero.

An hourglass of mass M stands on a fixed platform which also measures the apparent weight of the hourglass. The sand is at rest in the upper chamber when, at time $t = 0$, a tiny disturbance causes the sand to start running through. The sand comes to rest in the lower chamber after a time $t = \tau$. Find the time average of the appararent weight of the hourglass over the time interval $[0, \tau]$. [The apparent weight of the hourglass is however *not constant* in time. One can advance an argument that, when the sand is steadily running through, the apparent weight of the hourglass *exceeds* the real weight!]

10.2 Show that, if a system moves periodically, then the average of the total external force over a period of the motion must be zero.

A juggler juggles four balls of masses M, $2M$,$3M$ and $4M$ in a periodic manner. Find the time average (over a period) of the total force he applies to the balls. The juggler wishes to cross a shaky bridge that cannot support the combined weight of the juggler and his balls. Would it help if he juggles his balls while he crosses?

10.3* A boat of mass M is at rest in still water and a man of mass m is sitting at the bow. The man stands up, walks to the stern of the boat and then sits down again. If the water offers a resistance to the motion of the boat proportional to the velocity of the boat, show that the boat will *eventually* come to rest at its orginal position. [This remarkable result is independent of the resistance constant and the details of the man's motion.]

10.4 A uniform rope of mass M and length a is held at rest with its two ends close together and the rope hanging symmetrically below. (In this position, the rope has two long vertical segments connected by a small curved segment at the bottom.) One of the ends is then released. It can be shown by energy conservation (see Problem 9.8) that the velocity of the free end when it has descended by a distance x is given by

$$v^2 = \left(\frac{x(2a-x)}{a-x}\right) g.$$

Find the reaction R exerted by the support at the *fixed* end when the free end has descended a distance x. The support will collapse if R exceeds $\frac{3}{2}Mg$. Find how far the free end will fall before this happens.

10.5 A fine uniform chain of mass M and length a is held at rest hanging vertically downwards with its lower end just touching a fixed horizontal table. The chain is then released. Show that, while the chain is falling, the force that the chain exerts on the table is always *three times* the weight of chain actually lying on the table. [Assume that, before hitting the table, the chain falls freely under gravity.]

* When all the chain has landed on the table, the loose end is pulled upwards with the constant force $\frac{1}{3}Mg$. Find the height to which the chain will first rise. [This time, assume that the force exerted on the chain by the table is *equal* to the weight of chain lying on the table.]

10.6 A uniform ball of mass M and radius a can roll without slipping on the rough outer surface of a fixed sphere of radius b and centre O. Initially the ball is at rest at the highest point of the sphere when it is slightly disturbed. Find the speed of the centre G of the ball in terms of the variable θ, the angle between the line OG and the upward vertical. [Assume planar motion.] Show that the ball will leave the sphere when $\cos\theta = \frac{10}{17}$.

Rocket motion

10.7 A rocket of initial mass M, of which $M - m$ is fuel, burns its fuel at a constant rate in time τ and ejects the exhaust gases with constant speed u. The rocket starts from rest and moves vertically under uniform gravity. Show that the maximum speed acheived by the rocket is $u \ln\gamma$ and that its height at burnout is

$$u\tau\left(1 - \frac{\ln\gamma}{\gamma - 1}\right),$$

where $\gamma = M/m$. [Assume that the thrust is such that the rocket takes off immediately.]

10.8 ***Saturn V rocket*** In first stage of the Saturn V rocket, the initial mass was 2.8×10^6 kg, of which 2.1×10^6 kg was fuel. The fuel was burned at a constant rate over 150 s and the exhaust speed was 2, 600 m s^{-1}. Use the results of the last problem to find the speed and height of the Saturn V at first stage burnout. [Take g to be constant at 9.8 m s^{-2} and neglect air resistance.]

10.9 ***Rocket in resisting medium*** A rocket of initial mass M, of which $M - m$ is fuel, burns its fuel at a constant rate k and ejects the exhaust gases with constant speed u. The rocket starts from rest and moves through a medium that exerts the resistance force $-\epsilon k v$, where v is the forward velocity of the rocket, and ϵ is a small positive constant. Gravity is absent. Find the maximum speed V achieved by the rocket. Deduce a two term approximation for V, valid when ϵ is small.

10.10 ***Two-stage rocket*** A two-stage rocket has a first stage of initial mass M_1, of which $(1 - \eta)M_1$ is fuel, a second stage of initial mass M_2, of which $(1 - \eta)M_2$ is fuel, and an inert payload of mass m_0. In each stage, the exhaust gases are ejected with the same speed u. The rocket is initially at rest in free space. The first stage is fired and, on completion, the first stage carcass (of mass ηM_1) is discarded. The second stage is then fired. Find an expression for the final speed V of the rocket and deduce that V will be maximised when the mass ratio $\alpha = M_2/(M_1 + M_2)$ satisfies the equation

$$\alpha^2 + 2\beta\alpha - \beta = 0,$$

where $\beta = m_0/(M_1 + M_2)$. [Messy algebra.]

Show that, when β is small, the optimum value of α is approximatelely $\beta^{1/2}$ and the maximum velocity reached is approximately $2u \ln \gamma$, where $\gamma = 1/\eta$.

10.11* A raindrop falls vertically through stationary mist, collecting mass as it falls. The raindrop remains spherical and the rate of mass accretion is proportional to its speed and the square of its radius. Show that, if the drop starts from rest with a negligible radius, then it has constant acceleration $g/7$. [Tricky ODE.]

Collisions

10.12 A body of mass $4m$ is at rest when it explodes into *three* fragments of masses $2m$, m and m. After the explosion the two fragments of mass m are observed to be moving with the same speed in directions making 120° with each other. Find the proportion of the total kinetic energy carried by each fragment.

10.13 Show that, in an elastic head-on collision between two spheres, the relative velocity of the spheres after impact is the negative of the relative velocity before impact.

A tube is fixed in the vertical position with its lower end on a horizontal floor. A ball of mass M is released from rest at the top of the tube followed closely by a second ball of mass m. The first ball bounces off the floor and immediately collides with the second ball coming down. Assuming that both collisions are elastic, show that, when m/M is small, the second ball will be projected upwards to a height nearly nine times the length of the tube.

10.14 Two particles with masses m_1, m_2 and velocities $\boldsymbol{v}_1, \boldsymbol{v}_2$ collide and stick together. Find the velocity of this composite particle and show that the loss in kinetic energy due to the collision is

$$\frac{m_1 m_2}{2(m_1+m_2)}\,|\boldsymbol{v}_1-\boldsymbol{v}_2|^2 .$$

10.15 In an elastic collision between a proton moving with speed u and a helium nucleus at rest, the proton was scattered through an angle of 45°. What proportion of its initial energy did it lose? What was the recoil angle of the helium nucleus?

10.16 In an elastic collision between an alpha particle and an unknown nucleus at rest, the alpha particle was deflected through a right angle and lost 40% of its energy. Identify the mystery nucleus.

10.17 ***Some inequalities in elastic collisions*** Use the elastic scattering formulae to show the following inequalities:

(i) When $m_1 > m_2$, the scattering angle θ_1 is restricted to the range $0 \le \theta_1 \le \sin^{-1}(m_2/m_1)$.

(ii) If $m_1 < m_2$, the opening angle is obtuse, while, if $m_1 > m_2$, the opening angle is acute.

(iii)

$$\frac{E_1}{E_0} \ge \left(\frac{m_1-m_2}{m_1+m_2}\right)^2, \qquad \frac{E_2}{E_0} \le \frac{4m_1m_2}{(m_1+m_2)^2}.$$

10.18 ***Equal masses*** Show that, when the particles are of equal mass, the elastic scattering formulae take the simple form

$$\theta_1=\tfrac{1}{2}\psi \qquad \theta_2=\tfrac{1}{2}\pi-\tfrac{1}{2}\psi \qquad \theta=\tfrac{1}{2}\pi \qquad \frac{E_1}{E_0}=\cos^2\tfrac{1}{2}\psi \qquad \frac{E_2}{E_0}=\sin^2\tfrac{1}{2}\psi$$

where ψ is the scattering angle in the ZM frame.

In the scattering of neutrons of energy E by neutrons at rest, in what directions should the experimenter look to find neutrons of energy $\frac{1}{4}E$? What other energies would be observed in these directions?

10.19 Use the elastic scattering formulae to express the energy of the scattered particle as a function of the scattering angle, and the energy of the recoiling particle as a function of the recoil angle, as follows:

$$\frac{E_1}{E_0}=\frac{1+\gamma^2\cos 2\theta_1+2\gamma\cos\theta_1\left(1-\gamma^2\sin^2\theta_1\right)^{1/2}}{(\gamma+1)^2}, \qquad \frac{E_2}{E_0}=\frac{4\gamma}{(\gamma+1)^2}\cos^2\theta_2.$$

Make polar plots of E_1/E_0 as a function of θ_1 for the case of neutrons scattered by the nuclei of hydrogen, deuterium, helium and carbon.

Two-body problem and two-body scattering

10.20 ***Binary star*** The observed period of the binary star Cygnus X-1 (of which only one component is visible) is 5.6 days, and the semi-major axis of the orbit of the visible component is about 0.09 AU. The mass of the visible component is believed to be about $20M_\odot$. Estimate the mass of its dark companion. [Requires the numerical solution of a cubic equation.]

10.21 In two-body elastic scattering, show that the angular distribution of the *recoiling* particles is given by

$$4\cos\theta_2\,\sigma^{ZM}(\pi - 2\theta_2),$$

where $\sigma^{ZM}(\psi)$ is defined by equation (10.32).

In a Rutherford scattering experiment, alpha particles of energy E were scattered by a target of ionised helium. Find the angular distribution of the emerging particles.

10.22* Consider two-body elastic scattering in which the incident particles have energy E_0. Show that the energies of the *recoiling* particles lie in the interval $0 \le E \le E_{\max}$, where $E_{\max} = 4\gamma E_0/(1+\gamma)^2$. Show further that the energies of the recoiling particles are distributed over the interval $0 \le E \le E_{\max}$ by the frequency distribution

$$f(E) = \left(\frac{4\pi}{E_{\max}}\right)\sigma^{ZM}(\psi),$$

where σ^{ZM} is defined by equation (10.32), and

$$\psi = 2\sin^{-1}\left(\frac{E}{E_{\max}}\right)^{1/2}.$$

In the elastic scattering of neutrons of energy E_0 by protons at rest, the energies of the recoiling protons were found to be uniformly distributed over the interval $0 \le E \le E_0$, the total cross section being A. Find the *angular* distribution of the recoiling protons and the scattering cross section of the incident neutrons.

Integrable systems

10.23 A particle Q has mass $2m$ and two other particles P, R, each of mass m, are connected to Q by light inextensible strings of length a. The system is free to move on a smooth horizontal table. Initially P, Q R are at the points $(0, a)$, $(0, 0)$, $(0, -a)$ respectively so that they lie in a straight line with the strings taut. Q is then projected in the positive x-direction with speed u. Express the conservation of linear momentum and energy for this system in terms of the coordinates x (the displacement of Q) and θ (the angle turned by each of the strings).

Show that θ satisfies the equation

$$\dot{\theta}^2 = \frac{u^2}{a^2}\left(\frac{1}{2-\cos^2\theta}\right)$$

and deduce that P and R will collide after a time

$$\frac{a}{u}\int_0^{\pi/2}\left[2-\cos^2\theta\right]^{\frac{1}{2}}\,d\theta.$$

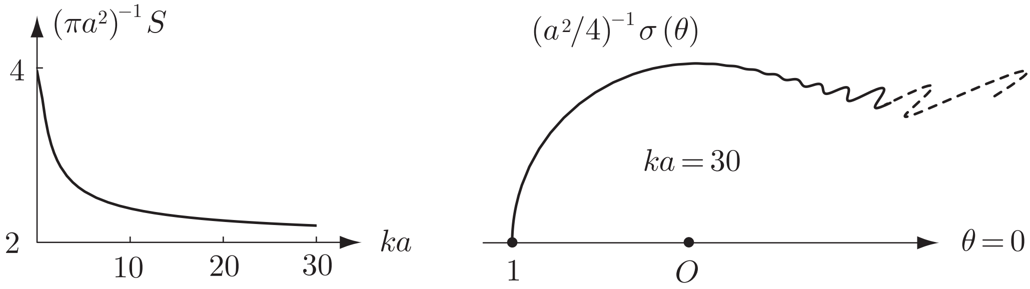


FIGURE 10.11 The quantum mechanical solution of the problem in which a uniform beam of particles, each with momentum $\hbar k$, is scattered by an impenetrable sphere of radius a. **Left**: The (dimensionless) total cross section $(\pi a^2)^{-1} S$ against ka. **Right**: A polar graph of the (dimensionless) scattering cross section $(a^2/4)^{-1}\sigma(\theta)$ against θ when $ka = 30$.

10.24 A uniform rod of length $2a$ has its lower end in contact with a smooth horizontal table. Initially the rod is released from rest in a position making an angle of 60° with the upward vertical. Express the conservation of linear momentum and energy for this system in terms of the coordinates x (the horizontal displacement of the centre of mass of the rod) and θ (the angle between the rod and the upward vertical). Deduce that the centre of mass of the rod moves in a vertical straight line, and that θ satisifies the equation

$$\dot{\theta}^2 = \frac{3g}{a}\left(\frac{1 - 2\cos\theta}{4 - 3\cos^2\theta}\right).$$

Find how long it takes for the rod to hit the table.

Computer assisted problems

10.25 ***Two-body Rutherford scattering*** Calculate the two-body scattering cross section σ^{TB} for Rutherford scattering and obtain the graphs shown in Figure 10.8. Obtain also an approximate formula for σ^{TB} valid for small γ $(= m_1/m_2)$, and correct to order $O(\gamma^2)$.

10.26 ***Comparison with quantum scattering*** A uniform flux of particles is incident upon a fixed hard sphere of radius a. The particles that strike the sphere are reflected elastically. Show that the differential scattering cross section is $\sigma(\theta) = a^2/4$ and that the total cross section is $S = \pi a^2$.

The solution of the same problem given by quantum mechanics is

$$\sigma(\theta) = \frac{a^2}{(ka)^2}\left|\sum_{l=0}^{\infty}\frac{(2l+1)\,j_l(ka)P_l(\cos\theta)}{h_l(ka)}\right|^2, \qquad S = \frac{4\pi a^2}{(ka)^2}\sum_{l=0}^{\infty}\left|\frac{(2l+1)\,j_l(ka)}{h_l(ka)}\right|^2,$$

where $P_l(z)$ is the Legendre polynomial of degree l, and $j_l(z)$, $h_l(z)$ are spherical Bessel functions order l. (Stay cool: these special functions should be available on your computer package.) The parameter k is related to the particle momentum p by the formula $p = \hbar k$, where $\hbar$ is the modified Planck constant. When ka is large, one would expect the quantum mechanical values for $\sigma(\theta)$ and S to approach the classical values. Calculate the quantum mechanical values numerically for ka up to about 30 (the calculation becomes increasingly difficult as ka increases), using about 100 terms of the series.

The author's results are shown in Figure 10.11. The quantum mechanical value for $\sigma(\theta)$ does approach the classical value for larger scattering angles, but behaves very erratically for small scattering angles. Also, the value of S tends to *twice* the value expected! Your physics lecturer will be pleased to explain these interesting anomalies.

Chapter Eleven

The angular momentum principle

and angular momentum conservation

KEY FEATURES

The key features of this chapter are the **angular momentum principle** and **conservation of angular momentum**. Together, the linear and angular momentum principles provide the governing equations of **rigid body motion**.

This chapter is essentially based on the **angular momentum principle** and its consequences. The angular momentum principle is the last of the three great principles of multi-particle mechanics* that apply to *every* mechanical system without restriction. Under appropriate conditions, the angular momentum of a system (or one of its components) is **conserved**, and we use this conservation principle to solve a variety of problems.

Together, the linear and angular momentum principles provide the governing equations of **rigid body motion**; the linear momentum principle determines the *translational* motion of the centre of mass, while the angular momentum principle determines the *rotational* motion of the body relative to the centre of mass. In this chapter, we restrict our attention to the special case of **planar rigid body motion**. Three-dimensional motion of rigid bodies is considered in Chapter 19.

11.1 THE MOMENT OF A FORCE

We begin with the definition of the moment of a force about a *point*, which is a vector quantity. The moment of a force about an *axis*, a scalar quantity, is the component along the axis of the corresponding vector moment.

Definition 11.1 ***Moment of a force about a point*** *Suppose a force* $\boldsymbol{F}$ *acts on a particle* P *with position vector* $\boldsymbol{r}$ *relative to an origin* O. *Then* $\boldsymbol{k}_O$, *the* ***moment***† *of the force* $\boldsymbol{F}$ *about the point* O *is defined to be*

$$\boldsymbol{k}_O = \boldsymbol{r} \times \boldsymbol{F}, \tag{11.1}$$

* The other two are the energy and linear momentum principles.

† Also called **torque**, especially in the engineering literature.

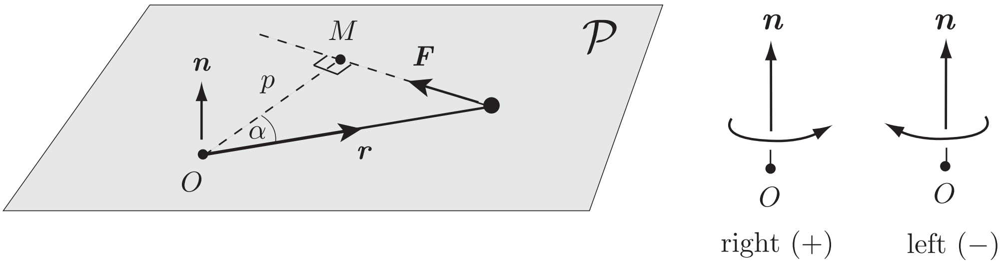


FIGURE 11.1 **Left**: Geometrical interpretation of the vector moment $\boldsymbol{k}_O = \boldsymbol{r}\times\boldsymbol{F}$. **Right**: The right- and left-handed senses around the 'axis' $\{O, \boldsymbol{n}\}$.

*a **vector** quantity. If the system of particles $P_1, P_2, \ldots, P_N$, with position vectors $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N$ are acted upon by the system of forces $\boldsymbol{F}_1, \boldsymbol{F}_2, \ldots, \boldsymbol{F}_N$ respectively, then $\boldsymbol{K}_O$, the **total moment** of the system of forces about O is defined to be the **vector sum** of the moments of the individual forces, that is,*

$$\boldsymbol{K}_O = \sum_{i=1}^{N} \boldsymbol{r}_i \times \boldsymbol{F}_i. \tag{11.2}$$

Since any fixed point can be taken to be the origin O, there is no loss of generality in the above definitions. However, there are occasions on which it is convenient to take moments about a general point A whose position vector is $\boldsymbol{a}$. To find $\boldsymbol{K}_A$, we simply replace $\boldsymbol{r}_i$ in the above definitions by the position vector of P_i *relative to* A, namely, $\boldsymbol{r}_i - \boldsymbol{a}$. This gives

$$\boldsymbol{K}_A = \sum_{i=1}^{N} (\boldsymbol{r}_i - \boldsymbol{a}) \times \boldsymbol{F}_i. \tag{11.3}$$

It follows that $\boldsymbol{K}_A$ and $\boldsymbol{K}_O$ are simply related by

$$\boldsymbol{K}_A = \boldsymbol{K}_O - \boldsymbol{a}\times\boldsymbol{F},$$

where $\boldsymbol{F}$ is the resultant force. Hence, if $\boldsymbol{F}$ is zero, the total moment of the forces $\{\boldsymbol{F}_i\}$ is the *same about every point*. Such a force system is said to be a **couple** with moment $\boldsymbol{K}$.

Geometrical interpretation of vector moment

The formula (11.1) has a nice geometrical interpretation. Let $\mathcal{P}$ be the plane that contains the origin and the force $\boldsymbol{F}$, as shown in Figure 11.1. Let $\boldsymbol{n}$ be a unit vector normal to $\mathcal{P}$, and suppose that $\boldsymbol{F}$ acts in the right-handed (or positive) sense around the 'axis' $\{O, \boldsymbol{n}\}$. (This is the case shown in Figure 11.1.) Then, from the definition (1.4) of the vector product,

$$\boldsymbol{K}_O = \boldsymbol{r}\times\boldsymbol{F} = \Big(|\boldsymbol{r}|\,|\boldsymbol{F}|\sin(\tfrac{1}{2}\pi + \alpha)\Big)\,\boldsymbol{n} = F(r\cos\alpha)\,\boldsymbol{n} = (F\times p)\,\boldsymbol{n},$$

where F is the magnitude of $\boldsymbol{F}$ and $p\,(= OM)$ is the perpendicular distance of O from the 'line of action' of $\boldsymbol{F}$. Thus, $\boldsymbol{K}_O$ has magnitude $F\times p$ and points in the $\boldsymbol{n}$-direction. If $\boldsymbol{F}$ has the left-handed (or negative) sense around $\{O, \boldsymbol{n}\}$, then $\boldsymbol{K}_O = -(F\times p)\,\boldsymbol{n}$.

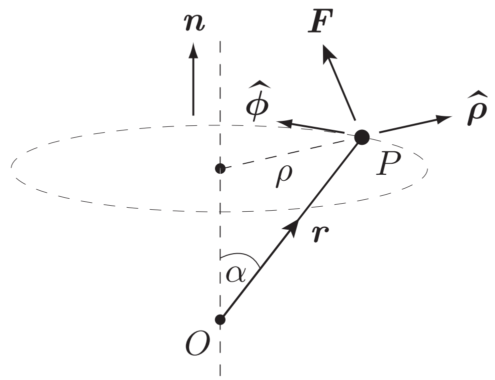


FIGURE 11.2 The moment of the force $\boldsymbol{F}$ about the axis $\{O, \boldsymbol{n}\}$ is $\rho \times (\boldsymbol{F} \cdot \widehat{\boldsymbol{\phi}})$.

Motion in a plane

Suppose we have a **system** of particles that lie in a plane, and the forces acting on the particles also lie in this plane. Such a system is said to be **two-dimensional**. Then the **total moment $\boldsymbol{K}_O$** of these forces about a point O of the plane is given by

$$\boldsymbol{K}_O = \sum_{i=1}^{N} \pm F_i \, p_i \, \boldsymbol{n} = \left(\sum_{i=1}^{N} \pm F_i \, p_i \right) \boldsymbol{n},$$

where the plus (or minus) sign is taken when the sense of $\boldsymbol{F}_i$ around the axis $\{O, \boldsymbol{n}\}$ is right- (or left-) handed. This formula explains why, in *two-dimensional* mechanics, the moment of a force can be represented by the *scalar* quantity $\pm F \times p$. In the two-dimensional case, the directions of all the moments are parallel, so that they add like scalars. However, in three dimensional mechanics, the moments have general directions and must be summed as vectors.

Moments about an axis

Definition 11.2 ***Moment of a force about an axis*** *The component of the moment* $\boldsymbol{K}_O$ *in the direction of a unit vector* $\boldsymbol{n}$ *is called the* ***moment*** *of* $\boldsymbol{F}$ *about the* ***axis**** $\{O, \boldsymbol{n}\}$*; it is the scalar quantity* $\boldsymbol{K}_O \cdot \boldsymbol{n}$.

This axial moment can be written (see Figure 11.2)

$$\begin{aligned} \boldsymbol{K}_O \cdot \boldsymbol{n} &= (\boldsymbol{r} \times \boldsymbol{F}) \cdot \boldsymbol{n} = (\boldsymbol{n} \times \boldsymbol{r}) \cdot \boldsymbol{F} = \big((r \sin\alpha)\, \widehat{\boldsymbol{\phi}}\big) \cdot \boldsymbol{F} \\ &= \rho \left(\boldsymbol{F} \cdot \widehat{\boldsymbol{\phi}}\right), \end{aligned}$$

where ρ is the distance of P from the axis $\{O, \boldsymbol{n}\}$ and ϕ is measured around the axis. The direction of the unit vector $\widehat{\boldsymbol{\phi}}$ is called the **azimuthal** direction around the axis $\{O, \boldsymbol{n}\}$. Thus $\boldsymbol{F} \cdot \widehat{\boldsymbol{\phi}}$ is the **azimuthal component** of $\boldsymbol{F}$.

---

* This ‘axis’ is merely a directed line in space. It does not neccessarily correspond to the rotation of any rigid body.

Example 11.1 ***Finding moments (numerical example)***

A force $\boldsymbol{F} = 2\boldsymbol{i} - \boldsymbol{j} - 2\boldsymbol{k}$ acts on a particle located at the point $P(0, 3, -1)$. Find the moment of $\boldsymbol{F}$ about the origin O and about the point $A(-2, 4, -3)$. Find also the moment of $\boldsymbol{F}$ about the axis through O in the direction of the vector $3\boldsymbol{i} - 4\boldsymbol{k}$.

Solution

The moment $\boldsymbol{K}_O$ is given by

$$\boldsymbol{K}_O = \boldsymbol{r}\times\boldsymbol{F} = (3\boldsymbol{j} - \boldsymbol{k})\times(2\boldsymbol{i} - \boldsymbol{j} - 2\boldsymbol{k}) = -7\boldsymbol{i} - 2\boldsymbol{j} - 6\boldsymbol{k}.$$

Similarly,

$$\boldsymbol{K}_A = (\boldsymbol{r} - \boldsymbol{a})\times\boldsymbol{F} = (2\boldsymbol{i} - \boldsymbol{j} + 2\boldsymbol{k})\times(2\boldsymbol{i} - \boldsymbol{j} - 2\boldsymbol{k}) = 4\boldsymbol{i} + 8\boldsymbol{j}.$$

The required axial moment is $\boldsymbol{K}_O \cdot \boldsymbol{n}$, where $\boldsymbol{n}$ is the unit vector in the direction of $3\boldsymbol{i} - 4\boldsymbol{k}$, namely

$$\boldsymbol{n} = \frac{3\boldsymbol{i} - 4\boldsymbol{k}}{|3\boldsymbol{i} - 4\boldsymbol{k}|} = \frac{3\boldsymbol{i} - 4\boldsymbol{k}}{5}.$$

Hence

$$\boldsymbol{K}_O \cdot \boldsymbol{n} = (-7\boldsymbol{i} - 2\boldsymbol{j} - 6\boldsymbol{k}) \cdot \left(\frac{3\boldsymbol{i} - 4\boldsymbol{k}}{5}\right) = \frac{3}{5}. \blacksquare$$

Example 11.2 ***Total moment of gravity forces***

A system $\mathcal{S}$ moves under uniform gravity. Show that the total moment of the gravity forces about any point is the same as if all the mass of $\mathcal{S}$ were concentrated at its centre of mass.

Solution

Without losing generality, let the point about which moments are taken be the origin O. Under uniform gravity, $\boldsymbol{F}_i = -m_i g\boldsymbol{k}$, where the unit vector $\boldsymbol{k}$ points vertically upwards, so that

$$\begin{aligned}\boldsymbol{K}_O &= \sum_{i=1}^{N} \boldsymbol{r}_i \times(-m_i g\boldsymbol{k}) = \left(\sum_{i=1}^{N} m_i \boldsymbol{r}_i\right)\times(-g\boldsymbol{k}) = (M\boldsymbol{R})\times(-g\boldsymbol{k}) \\ &= \boldsymbol{R}\times(-Mg\boldsymbol{k}),\end{aligned}$$

where M is the total mass of $\mathcal{S}$ and $\boldsymbol{R}$ is the position vector of its centre of mass. This is the required result. Note that it is only true for *uniform* gravity. ■

11.2 ANGULAR MOMENTUM

We begin with the definition of the angular momentum of a particle about a fixed point. The old name for angular momentum is ‘moment of momentum’ and that is exactly what it is - the moment of the linear momentum of the particle about the chosen point.

Definition 11.3 ***Angular momentum about a point*** *Suppose a particle P of mass m has position vector $\boldsymbol{r}$ and velocity $\boldsymbol{v}$. Then $\boldsymbol{l}_O$, the* ***angular momentum*** *of P about O is defined to be*

$$\boldsymbol{l}_O = \boldsymbol{r} \times (m\boldsymbol{v}), \tag{11.4}$$

a ***vector*** *quantity. If the system of particles $P_1, P_2, \dots, P_N$, with masses $m_1, m_2, \dots, m_N$, have position vectors $\boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_N$ and velocities $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_N$ respectively, then $\boldsymbol{L}_O$, the* ***angular momentum*** *of the system about O, is defined to be the* ***vector sum*** *of the angular momenta of the individual particles, that is,*

$$\boldsymbol{L}_O = \sum_{i=1}^{N} \boldsymbol{r}_i \times (m_i \boldsymbol{v}_i). \tag{11.5}$$

The corresponding formula for angular momentum about a general point A is therefore

$$\boldsymbol{L}_A = \sum_{i=1}^{N} (\boldsymbol{r}_i - \boldsymbol{a}) \times (m_i \boldsymbol{v}_i),$$

from which it follows that $\boldsymbol{L}_A$ and $\boldsymbol{L}_O$ are simply related by

$$\boldsymbol{L}_A = \boldsymbol{L}_O - \boldsymbol{a} \times \boldsymbol{P},$$

where $\boldsymbol{P}$ is the total *linear* momentum of the system.

The geometrical interpretation of the angular momentum of a particle is similar to that of moment of a force (see Figure 11.1). Let $\mathcal{P}$ be the plane that contains O, P and the velocity $\boldsymbol{v}$, and let $\boldsymbol{n}$ be a unit vector normal to $\mathcal{P}$. Then

$$\boldsymbol{L}_O = \pm(mv \times p)\,\boldsymbol{n},$$

where v is the magnitude of $\boldsymbol{v}$ and p is the perpendicular distance of O from the line through P parallel to $\boldsymbol{v}$. The $\pm$ sign is decided by the sense of $\boldsymbol{v}$ around the axis $\{O, \boldsymbol{n}\}$, as shown in Figure 11.1.

Example 11.3 *Calculating the angular momentum of a particle*

The position of a particle P of mass m at time t is given by $x = a\theta^2$, $y = 2a\theta$, $z = 0$, where $\theta = \theta(t)$. Find the angular momentum of P about the point $B(a, 0, 0)$ at time t.

Solution

The position vector of the particle relative to B at time t is

$$\boldsymbol{r} - \boldsymbol{b} = \left(a\theta^2 \boldsymbol{i} + 2a\theta\,\boldsymbol{j}\right) - a\boldsymbol{i} = a\left[(\theta^2 - 1)\boldsymbol{i} + 2\theta\,\boldsymbol{j}\right]$$

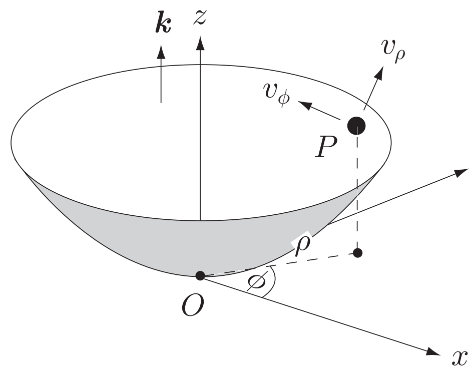


FIGURE 11.3 The particle P slides on the inside surface of the axially symmetric bowl $z = f(\rho)$.

and the velocity of the particle at time t is

$$\boldsymbol{v} = \frac{d\boldsymbol{r}}{dt} = \frac{d\boldsymbol{r}}{d\theta} \times \frac{d\theta}{dt} = 2a(\theta\boldsymbol{i} + \boldsymbol{j})\dot{\theta}.$$

The angular momentum of the particle about B at time t is therefore

$$\begin{aligned}\boldsymbol{L}_B &= (\boldsymbol{r} - \boldsymbol{b})\times(m\boldsymbol{v}) = 2ma^2\dot{\theta}\left[(\theta^2 - 1)\boldsymbol{i} + 2\theta\boldsymbol{j}\right]\times[\theta\boldsymbol{i} + \boldsymbol{j}] \\ &= -2ma^2(\theta^2 + 1)\dot{\theta}\boldsymbol{k}. \blacksquare\end{aligned}$$

Angular momentum about an axis

Definition 11.4 ***Angular momentum about an axis*** *The component of the angular momentum $\boldsymbol{L}_O$ in the direction of a unit vector $\boldsymbol{n}$ is called the* ***angular momentum*** *of P about the* ***axis*** *$\{O, \boldsymbol{n}\}$; it is the scalar quantity $\boldsymbol{L}_O \cdot \boldsymbol{n}$.*

By that same argument as was used for moments about an axis, the angular momentum of a particle of mass m and velocity $\boldsymbol{v}$ about the axis $\{O, \boldsymbol{n}\}$ can be written in the form

$$\boldsymbol{L}_O \cdot \boldsymbol{n} = m\rho\left(\boldsymbol{v}\cdot\widehat{\boldsymbol{\phi}}\right), \tag{11.6}$$

where ρ is the perpendicular distance of the particle from the axis and $\boldsymbol{v}\cdot\widehat{\boldsymbol{\phi}}$ is the azimuthal component of $\boldsymbol{v}$ around the axis.

Example 11.4 *Particle sliding inside a bowl*

A particle P of mass m slides on the inside surface of an axially symmetric bowl. Find its angular momentum about its axis of symmetry in terms of the coordinates ρ, ϕ shown in Figure 11.3.

Solution

In order to express $\boldsymbol{L}_O \cdot \boldsymbol{k}$ in terms of coordinates, we draw a velocity diagram for the system as explained in section 10.10. The velocities v_ρ and v_ϕ, corresponding to the coordinates ρ and ϕ, have the directions shown in Figure 11.3. These two velocities are perpendicular, with the v_ϕ contribution in the *azimuthal* direction around the

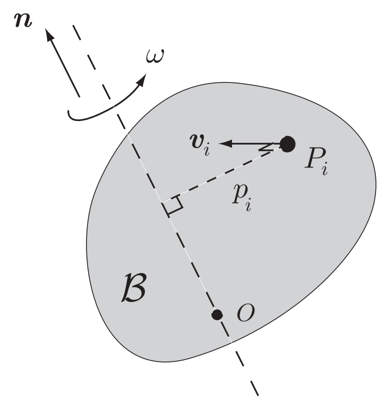


FIGURE 11.4 The rigid body $\mathcal{B}$ rotates about the fixed axis $\{O, \boldsymbol{n}\}$ with angular velocity ω.

vertical axis $\{O, \boldsymbol{k}\}$. It follows that $\boldsymbol{v} \cdot \widehat{\boldsymbol{\phi}} = v_\phi = \rho\dot{\phi}$. The required axial angular momentum is therefore

$$\boldsymbol{L}_O \cdot \boldsymbol{k} = m\rho(\boldsymbol{v} \cdot \widehat{\boldsymbol{\phi}}) = m\rho(\rho\dot{\phi}) = m\rho^2\dot{\phi}.$$

Just for the record, the velocity v_ρ is the (vector) sum of $\dot{\rho}$ *radially outwards* and $\dot{z}$ *vertically upwards*. Note that $\dot{z}$ is not an independent quantity. If the equation of the bowl is $z = f(\rho)$, then $\dot{z} = f'(\rho)\dot{\rho}$. In particular then, the kinetic energy of P is given by

$$T = \tfrac{1}{2}m\left[\dot{\rho}^2 + \left(f'(\rho)\dot{\rho}\right)^2 + \left(\rho\dot{\phi}\right)^2\right].$$

and its potential energy by $V = mgf(\rho)$. ■

11.3 ANGULAR MOMENTUM OF A RIGID BODY

The probem of finding the angular momentum of a moving rigid body in the general three-dimensional case is tricky and is deferred until Chapter 19. In the present chapter we essentially restrict ourselves to the case of planar rigid body motion, for which it is sufficient to find the angular momentum of the body *about its axis of rotation*. This axis may be fixed (as in the armature of a motor) or, more generally, may be the instantaneous rotation axis through the centre of mass of the body (as in the case of a rolling penny). In this section we consider only the case of rotation about a fixed axis; the case of planar motion is treated in section 11.6.

Consider a rigid body $\mathcal{B}$ rotating with angular velocity ω about the *fixed* axis $\{O, \boldsymbol{n}\}$, as shown in Figure 11.4. Then the angular momentum of the body about this axis is

$$\boldsymbol{L}_O \cdot \boldsymbol{n} = \left(\sum_{i=1}^{N} \boldsymbol{l}_O^{(i)}\right) \cdot \boldsymbol{n} = \sum_{i=1}^{N} (\boldsymbol{l}_O^{(i)} \cdot \boldsymbol{n}) = \sum_{i=1}^{N} m_i p_i (\boldsymbol{v}_i \cdot \widehat{\boldsymbol{\phi}}),$$

where p_i is the perpendicular distance of m_i from the axis, and $\boldsymbol{v}_i \cdot \widehat{\boldsymbol{\phi}}$ is the azimuthal component of $\boldsymbol{v}_i$ around the axis (see formula (11.6)). But, since the body is rigid, the velocity of m_i is entirely azimuthal and is equal to ωp_i. Hence

$$\boldsymbol{L}_O \cdot \boldsymbol{n} = \left(\sum_{i=1}^{N} m_i p_i^2\right) \omega = I\,\omega, \tag{11.7}$$

where I is the moment of inertia of $\mathcal{B}$ about the rotation axis $\{O, \boldsymbol{n}\}$. We have thus proved that:

Angular momentum of a rigid body about its rotation axis

If a rigid body is rotating with angular velocity ω about the fixed axis $\{A, \boldsymbol{n}\}$, then the angular momentum of the body about this axis is given by

$$\boldsymbol{L}_A \cdot \boldsymbol{n} = I\,\omega, \tag{11.8}$$

where I is the moment of inertia of the body about the axis $\{O, \boldsymbol{n}\}$.

It should be remembered that, if a rigid body of general shape is rotating about the fixed axis $\{O, \boldsymbol{n}\}$, then $\boldsymbol{L}_O$, the angular momentum of the body about O, is not generally parallel to the rotation axis. If the rotation axis happens to be an axis of **rotational symmetry** of the body, then $\boldsymbol{L}_O$ *will* be parallel to the rotation axis and $\boldsymbol{L}_O$ is simply given by

$$\boldsymbol{L}_O = (I\omega)\boldsymbol{n}. \tag{11.9}$$

Example 11.5 *Axial angular momentum of a hollow sphere*

A hollow sphere of inner radius a and outer radius b is made of material of uniform density ρ. The sphere is spinning with angular velocity Ω about a fixed axis through its centre. Find the angular momentum of the sphere about its rotation axis.

Solution

From equation (SysA:L=Iomega), the angular momentum of the sphere about its rotation axis is given by $L = I\,\omega$, where I is its moment of inertia and ω is its angular velocity about this axis. In the present case,

$$I = \tfrac{2}{5}Mb^2 - \tfrac{2}{5}ma^2,$$

where $M = 4\rho b^3/3$ and $m = 4\rho a^3/3$, giving

$$I = \frac{8\rho}{15}\left(b^5 - a^5\right).$$

The angular momentum of the sphere about its rotation axis is therefore

$$L = \frac{8\rho}{15}\left(b^5 - a^5\right)\Omega. \blacksquare$$

11.4 THE ANGULAR MOMENTUM PRINCIPLE

We now derive the fundamental result which relates the angular momentum of any system to the external forces that act upon it – **the angular momentum principle**.

Consider the general multi-particle system $\mathcal{S}$ which consists of particles $P_1, P_2, \dots, P_N$, with masses $m_1, m_2, \dots, m_N$ and velocities $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_N$, as shown in Figure 9.1. Suppose that $\mathcal{S}$ is acted upon by **external** forces $\boldsymbol{F}_i$ and **internal** forces $\boldsymbol{G}_{ij}$, as shown in Figure 9.3. Then the equation of motion for the particle P_i is

$$m_i \frac{d\boldsymbol{v}_i}{dt} = \boldsymbol{F}_i + \sum_{j=1}^{N} \boldsymbol{G}_{ij}, \tag{11.10}$$

where, as in Chapter 9, we take $\boldsymbol{G}_{ij} = \boldsymbol{0}$ when $i = j$. Then the rate of increase of the angular momentum of the system $\mathcal{S}$ about the origin O can be written

$$\begin{aligned}\frac{d\boldsymbol{L}_O}{dt} &= \frac{d}{dt}\left(\sum_{i=1}^{N} \boldsymbol{r}_i \times (m_i \boldsymbol{v}_i)\right) = \sum_{i=1}^{N}\left\{\boldsymbol{r}_i \times \left(m_i \frac{d\boldsymbol{v}_i}{dt}\right) + \dot{\boldsymbol{r}}_i \times (m_i \boldsymbol{v}_i)\right\} \\ &= \sum_{i=1}^{N} \boldsymbol{r}_i \times \left(m_i \frac{d\boldsymbol{v}_i}{dt}\right),\end{aligned}$$

since $\dot{\boldsymbol{r}}_i \times (m_i \boldsymbol{v}_i) = m_i \boldsymbol{v}_i \times \boldsymbol{v}_i = \boldsymbol{0}$. On using the equation of motion (11.10), we obtain

$$\begin{aligned}\frac{d\boldsymbol{L}_O}{dt} &= \sum_{i=1}^{N} \boldsymbol{r}_i \times \left\{\boldsymbol{F}_i + \sum_{j=1}^{N} \boldsymbol{G}_{ij}\right\} = \sum_{i=1}^{N} \boldsymbol{r}_i \times \boldsymbol{F}_i \; + \; \sum_{i=1}^{N}\sum_{j=1}^{N} \boldsymbol{r}_i \times \boldsymbol{G}_{ij} \\ &= \boldsymbol{K}_O \; + \; \sum_{i=2}^{N}\left(\sum_{j=1}^{i-1}\left(\boldsymbol{r}_i \times \boldsymbol{G}_{ij} + \boldsymbol{r}_j \times \boldsymbol{G}_{ji}\right)\right),\end{aligned} \tag{11.11}$$

where $\boldsymbol{K}_O$ is the **total moment** about O of the *external* forces. We have also grouped the terms of the double sum in pairs and omitted those terms known to be zero. Now the internal forces $\{\boldsymbol{G}_{ij}\}$ satisfy the Third Law, which means that $\boldsymbol{G}_{ij}$ must be equal and opposite to $\boldsymbol{G}_{ji}$, and that $\boldsymbol{G}_{ij}$ must be parallel to the line $P_i P_j$. It follows that

$$\boldsymbol{r}_i \times \boldsymbol{G}_{ij} + \boldsymbol{r}_j \times \boldsymbol{G}_{ji} = \boldsymbol{r}_i \times \boldsymbol{G}_{ij} - \boldsymbol{r}_j \times \boldsymbol{G}_{ij} = (\boldsymbol{r}_i - \boldsymbol{r}_j) \times \boldsymbol{G}_{ij} = \boldsymbol{0},$$

since $\boldsymbol{G}_{ij}$ is parallel to the vector $\boldsymbol{r}_i - \boldsymbol{r}_j$. Thus each pair of terms of the double sum in equation (11.11) is zero and we obtain

$$\frac{d\boldsymbol{L}_O}{dt} = \boldsymbol{K}_O,$$

which is the **angular momentum principle**. Since *any fixed point* can be taken to be the origin, this proves that:

Angular momentum principle about fixed points

$$\frac{d\boldsymbol{L}_A}{dt} = \boldsymbol{K}_A \tag{11.12}$$

for *any fixed point* A. This fundamental principle can be stated as follows:

Angular momentum principle about a fixed point

In any motion of a system $\mathcal{S}$, the rate of increase of the angular momentum of $\mathcal{S}$ about any fixed point is equal to the total moment about that point of the external forces acting on $\mathcal{S}$.

It should be noted that only the external forces appear in the angular momentum principle so that the **internal forces need not be known**. It is this fact which gives the principle its power.

Question ***Overusing the angular momentum principle***

The angular momentum principle can be applied about any point. Are all the resulting equations independent of each other?

Answer

The short answer is obviously no. The long answer is as follows: From the definitions of $\boldsymbol{K}$ and $\boldsymbol{L}$, we have already shown that

$$\boldsymbol{K}_A = \boldsymbol{K}_O - \boldsymbol{a}\times\boldsymbol{F},$$

and

$$\boldsymbol{L}_A = \boldsymbol{L}_O - \boldsymbol{a}\times\boldsymbol{P},$$

where $\boldsymbol{F}$ is the total force acting on the system $\mathcal{S}$, and $\boldsymbol{P}$ is its linear momentum. It follows that, for any fixed point A,

$$\boldsymbol{K}_A - \dot{\boldsymbol{L}}_A = \left(\boldsymbol{K}_O - \dot{\boldsymbol{L}}_O\right) - \boldsymbol{a}\times\left(\boldsymbol{F} - \dot{\boldsymbol{P}}\right).$$

Hence, if the linear momentum principle $\dot{\boldsymbol{P}} = \boldsymbol{F}$ and the angular momentum principle $\dot{\boldsymbol{L}}_O = \boldsymbol{K}_O$ have already been used, then nothing new is obtained by applying the angular momentum principle about another point A. ■

Angular momentum principle about the centre of mass

The angular momentum principle in the form (11.12) does not generally apply if A is a *moving* point. However, the standard form does apply when moments and angular momenta are taken about the **centre of mass** G, even though G may be accelerating. This follows from the theorem below. The corresponding result for kinetic energy appeared in Chapter 9.

Theorem 11.1 *Suppose a general system of particles $\mathcal{S}$ has total mass M and that its centre of mass G has position vector $\boldsymbol{R}$ and velocity $\boldsymbol{V}$. Then the angular momentum of $\mathcal{S}$ about O can be written in the form*

$$\boldsymbol{L}_O = \boldsymbol{R}\times(M\boldsymbol{V}) + \boldsymbol{L}_G, \tag{11.13}$$

where $\boldsymbol{L}_G$ is the angular momentum of $\mathcal{S}$ about G in its motion **relative** *to G.*

Proof. By definition,

$$\begin{aligned}
\boldsymbol{L}_G &= \sum_{i=1}^{N} m_i\,(\boldsymbol{r}_i - \boldsymbol{R})\times(\boldsymbol{v}_i - \boldsymbol{V})\\
&= \sum_{i=1}^{N} m_i \boldsymbol{r}_i\times\boldsymbol{v}_i - \left(\sum_{i=1}^{N} m_i\boldsymbol{r}_i\right)\times\boldsymbol{V} - \boldsymbol{R}\times\left(\sum_{i=1}^{N} m_i\boldsymbol{v}_i\right) + \left(\sum_{i=1}^{N} m_i\right)\boldsymbol{R}\times\boldsymbol{V}\\
&= \boldsymbol{L}_O - (M\boldsymbol{R})\times\boldsymbol{V} - \boldsymbol{R}\times(M\boldsymbol{V}) + M(\boldsymbol{R}\times\boldsymbol{V})\\
&= \boldsymbol{L}_O - \boldsymbol{R}\times(M\boldsymbol{V}),
\end{aligned}$$

as required. ■

The two terms on the right of equation (11.13) have a nice physical interpretation. The term $\boldsymbol{R}\times(M\boldsymbol{V})$ is the **translational** contribution to $\boldsymbol{L}_O$ while the term $\boldsymbol{L}_G$ is the contribution from the motion of $\mathcal{S}$ **relative** to G. If $\mathcal{S}$ is a **rigid body**, then the motion of $\mathcal{S}$ relative to G is an angular velocity about some axis through G, and the term $\boldsymbol{L}_G$ then represents the **rotational contribution** to $\boldsymbol{L}_O$.

The angular momentum principle for $\mathcal{S}$ about O can therefore be written

$$\begin{aligned}
\boldsymbol{K}_O &= \frac{d}{dt}(M\boldsymbol{R}\times\boldsymbol{V}) + \frac{d\boldsymbol{L}_G}{dt}\\
&= M\boldsymbol{R}\times\dot{\boldsymbol{V}} + \frac{d\boldsymbol{L}_G}{dt}.
\end{aligned}$$

Furthermore, since $\boldsymbol{K}_O = \boldsymbol{K}_G + \boldsymbol{R}\times\boldsymbol{F}$, it follows that

$$\begin{aligned}
\boldsymbol{K}_G &= \frac{d\boldsymbol{L}_G}{dt} + \boldsymbol{R}\times\left(M\dot{\boldsymbol{V}} - \boldsymbol{F}\right)\\
&= \frac{d\boldsymbol{L}_G}{dt},
\end{aligned}$$

on using the *linear* momentum principle. We therefore obtain:

Angular momentum principle about G

$$\frac{d\boldsymbol{L}_G}{dt} = \boldsymbol{K}_G \tag{11.14}$$

Thus *the standard form of the angular momentum principle applies to the motion of $\mathcal{S}$ relative to the centre of mass G.*

The rigid body equations

The linear and angular momentum principles provide sufficient equations to determine the motion of a **single rigid body** moving under **known forces**. The standard form of the **rigid body equations** is

Rigid body equations

$$M\frac{d\boldsymbol{V}}{dt} = \boldsymbol{F} \qquad \frac{d\boldsymbol{L}_G}{dt} = \boldsymbol{K}_G \tag{11.15}$$

in which we have taken both the linear and angular momentum principles in their centre of mass form. The linear momentum principle thus determines the **translational motion** of G (as if it were a particle), and the angular momentum principle determines the **rotational motion** of the body about G.

We will use a subset of these equations later in this chapter to solve problems of **planar** rigid body motion. The delights of general **three-dimensional** rigid body motion* are revealed in Chapter 19.

Example 11.6 *Rigid body moving under uniform gravity*

A rigid body is moving in any manner under uniform gravity. Show that its motion *relative to its centre of mass* is the same as if gravity were absent.

Solution

Under uniform gravity, the total moment of the gravity forces about any point is the same as if they all acted at G, the centre of mass of the body (see Example 11.2). It follows that $\boldsymbol{K}_G = \boldsymbol{0}$.

The rigid body equations (11.15) therefore take the form

$$M\frac{d\boldsymbol{V}}{dt} = -Mg\boldsymbol{k}, \qquad \frac{d\boldsymbol{L}_G}{dt} = \boldsymbol{0}.$$

Hence, when a rigid body moves under uniform gravity, G undergoes projectile motion (which we already knew), and the equation for the motion of the body relative to G is the same as if the body were moving in free space. ■

* The difficulty in the three-dimensional case is the calulation of $\boldsymbol{L}$.

11.5 CONSERVATION OF ANGULAR MOMENTUM

Isolated systems

Suppose that $\mathcal{S}$ is an **isolated** system, and let A be any fixed point. Then $\boldsymbol{L}_A$, the total moment about A of the external forces acting on $\mathcal{S}$, is obviously zero. The angular momentum principle (11.12) then implies that $d\boldsymbol{L}_A/dt = \boldsymbol{0}$, which implies that $\boldsymbol{L}_A$ remains constant. The same argument holds for $\boldsymbol{L}_G$. This simple but important result can be stated as follows:

> **Conservation of angular momentum about a point**
>
> In any motion of an isolated system, the angular momentum of the system about any fixed point is conserved. The angular momentum of the system about its centre of mass is also conserved.

For example, the angular momentum of the **solar system** about any fixed point (or about its centre of mass) is conserved. The same is true for an astronaut floating freely in space (irrespective of how he moves his body). The angular momentum of a system about its centre of mass may still be conserved even when external forces are present. For any system moving under **uniform gravity** (a falling cat trying to land on its feet, say) $\boldsymbol{K}_G = \boldsymbol{0}$ which implies that $\boldsymbol{L}_G$ is conserved.

Angular momentum in central field orbits

In the case of a particle P moving in a **central field** with centre O,

$$\boldsymbol{K}_O = \boldsymbol{r} \times \boldsymbol{F} = \boldsymbol{0},$$

since $\boldsymbol{r}$ and $\boldsymbol{F}$ are parallel. This implies that $\boldsymbol{L}_O$ is conserved. (Angular momentum about other points is not conserved.) By symmetry, each possible motion of P must take place in a plane through O and we may take polar coordinates r, θ (centred on O) to specify the position of P in the plane of motion. In terms of these coordinates,

$$\boldsymbol{L}_O = \boldsymbol{r} \times (m\boldsymbol{v}) = m(r\,\widehat{\boldsymbol{r}}) \times \left(\dot{r}\,\widehat{\boldsymbol{r}} + (r\dot{\theta})\,\widehat{\boldsymbol{\theta}}\right) = mr^2\dot{\theta}\,\boldsymbol{n},$$

where the constant unit vector $\boldsymbol{n}$ $(= \widehat{\boldsymbol{r}} \times \widehat{\boldsymbol{\theta}})$ is perpendicular to the plane of motion. Hence, in this case, conservation of $\boldsymbol{L}_O$ is equivalent to conservation of $\boldsymbol{L}_O \cdot \boldsymbol{n}$, the angular momentum of P about the axis $\{O, \boldsymbol{n}\}$. The conclusion is then that the quantity

$$\boldsymbol{L}_O \cdot \boldsymbol{n} = mr^2\dot{\theta} = L,$$

where L is a constant. This important result was obtained in Chapter 7 by integrating the azimuthal equation of motion for P. We now see that it is a consequence of angular momentum conservation and that the constant L is the angular momentum* of P about the axis $\{O, \boldsymbol{n}\}$.

* The constant L used in Chapter 7 was actually the angular momentum *per unit mass*.

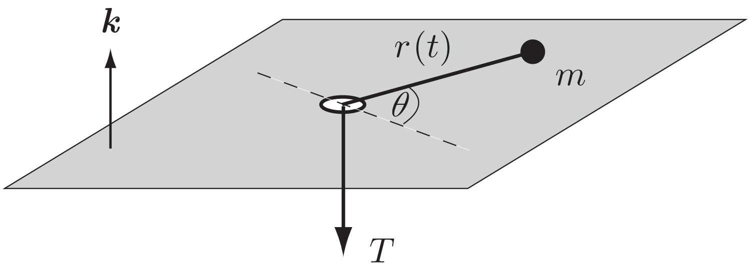


FIGURE 11.5 The particle slides on the table while the string is pulled down through the hole.

Conservation of angular momentum about an axis

Even when $\boldsymbol{K}_A \neq \boldsymbol{0}$ it is still possible for angular momentum to be conserved about a *particular axis* through A. Let $\boldsymbol{n}$ be a fixed unit vector and A a fixed point so that $\{A, \boldsymbol{n}\}$ is a *fixed axis* through the point A. Then

$$\frac{d}{dt}(\boldsymbol{L}_A \cdot \boldsymbol{n}) = \frac{d\boldsymbol{L}_A}{dt} \cdot \boldsymbol{n} + \boldsymbol{L}_A \cdot \frac{d\boldsymbol{n}}{dt} = \frac{d\boldsymbol{L}_A}{dt} \cdot \boldsymbol{n} = \boldsymbol{K}_A \cdot \boldsymbol{n}$$

Hence, if $\boldsymbol{K}_A \cdot \boldsymbol{n} = 0$ at all times, it follows that $\boldsymbol{L}_A \cdot \boldsymbol{n}$ is conserved. This result can be stated as follows:

Conservation of angular momentum about an axis

If the external forces acting on a system have no total moment about a fixed axis, then the angular momentum of the system about that axis is conserved. The same applies for a moving axis which passes through G and maintains a constant direction.

In our first example, angular momentum conservation is sufficient to determine the entire motion.

Example 11.7 *Pulling a particle through a hole*

A particle P of mass m can slide on a smooth horizontal table. P is connected to a light inextensible string which passes through a small smooth hole O in the table, so that the lower end of the string hangs vertically below the table while P moves on top with the string taut (see figure 11.5). Initially the lower end of the string is held fixed with P moving with speed u on a circle of radius a. The string is now pulled down from below in such a way that the string above the table has the length $r(t)$ at time t. Find the velocity of P and the tension in the string at time t.

Solution

We must first establish that some component of angular momentum is conserved in this motion. The forces acting on P are gravity, the normal reaction of the smooth table, and the tension in the string. Since the first two are equal and opposite and the tension force points towards O, it follows that $\boldsymbol{K}_O = \boldsymbol{0}$. Thus, however the string is pulled, $\boldsymbol{L}_O$ is **conserved** in the motion of P.

Now we must **calculate** $\boldsymbol{L}_O$. As in the case of central field orbits, $\boldsymbol{L}_O$ is perpendicular to the plane of motion and conservation of $\boldsymbol{L}_O$ is equivalent to conservation of the axial angular momentum $\boldsymbol{L}_O \cdot \boldsymbol{k}$. Hence, as in orbital motion,

$$\boldsymbol{L}_O \cdot \boldsymbol{k} = mr^2\dot{\theta} = L,$$

where the constant L is given by the initial conditions to be $L = mau$. Hence, in the motion of P, the **conservation equation**

$$mr^2\dot{\theta} = mau$$

is satisfied. Since $r(t)$ is given, this equation is sufficient to determine the motion of P. In particular, the velocity of P at time t is given by

$$\boldsymbol{v} = \dot{r}\,\widehat{\boldsymbol{r}} + (r\dot{\theta})\,\widehat{\boldsymbol{\theta}} = \dot{r}\,\widehat{\boldsymbol{r}} + \left(\frac{au}{r}\right)\widehat{\boldsymbol{\theta}},$$

from which we see that the transverse velocity of P tends to infinity as r tends to zero.

The **string tension** T can be found from the radial equation of motion for P, namely,

$$m\left(\ddot{r} - r\dot{\theta}^2\right) = -T,$$

which gives

$$T = m\left(r\dot{\theta}^2 - \ddot{r}\right) = m\left(\frac{a^2u^2}{r^3} - \ddot{r}\right).$$

For example, in order to pull the string down with constant speed, the applied tension must be

$$T = \frac{ma^2u^2}{r^3}.$$

This tends rapidly to infinity as r tends to zero, making it impossible to pull the particle through the hole! ■

Our second example belongs to a class of problems that could be called '*before* and *after* problems'. We have encountered the same notion before. In elastic collision problems, the linear momentum and energy of the system are conserved and these conservation laws are used to relate the initial state of the system (*before*) to the final state (*after*). This provides information about the final state that is independent of the nature of the particle interaction. Conservation of **angular momentum** can be exploited in the same way. In the following example, angular momentum conservation is sufficient to determine the final state uniquely.

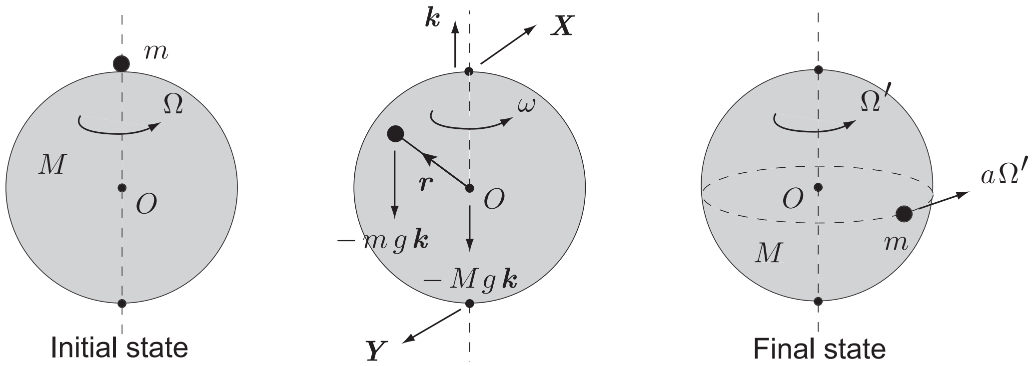


FIGURE 11.6 The beetle and the ball: the ball is smoothly pivoted about a vertical diameter and the beetle crawls on the surface of the ball.

Example 11.8 *The beetle and the ball*

A uniform ball of mass M and radius a is pivoted so that it can turn freely about one of its diameters which is fixed in a vertical position. A beetle of mass m can crawl on the surface of the ball. Initially the ball is rotating with angular speed Ω with the beetle at the 'North pole' (see Figure 11.6 (left)). The beetle then walks (in any manner) to the 'equator' of the ball and sits down. What is the angular speed of the ball now?

Solution

We must first establish that some component of angular momentum is conserved. The external forces acting on the system of 'beetle and ball' are shown in Figure 11.6 (centre). The forces $\boldsymbol{X}$ and $\boldsymbol{Y}$ are the constraint forces exerted by the pivots. The total moment of the external forces about O is therefore

$$\boldsymbol{K}_O = \boldsymbol{0}\times(-Mg\boldsymbol{k}) + \boldsymbol{r}\times(-mg\boldsymbol{k}) + (a\boldsymbol{k})\times\boldsymbol{X} + (-a\boldsymbol{k})\times\boldsymbol{Y}.$$

It follows that

$$\boldsymbol{K}_O \cdot \boldsymbol{k} = 0,$$

since all the resulting triple scalar products contain two $\boldsymbol{k}$'s. Hence $\boldsymbol{L}_O \cdot \boldsymbol{k}$, the angular mometum of the system about the rotation axis, is conserved, irrespective of the wanderings of the beetle. It follows that this axial angular momentum is the same *after* as it was *before*.

In the **initial** state, the angular momentum of the ball about its rotation axis is given by $I\,\Omega$, where $I = 2Ma^2/5$. Initially the beetle has zero velocity so its angular momentum is zero. Hence the **initial value** of the axial angular momentum is

$$\boldsymbol{L}_O \cdot \boldsymbol{k} = \left(\tfrac{2}{5}Ma^2\right)\Omega.$$

In the **final** state the ball has an unknown angular velocity Ω' and axial angular momentum $2Ma^2\Omega'/5$. The velocity of the beetle is entirely azimuthal and is equal to $\Omega' a$. Hence, on using the formula (11.6), the axial angular momentum of the beetle is given by $m\rho(\boldsymbol{v}\cdot\widehat{\boldsymbol{\phi}}) = ma(\Omega a)$. The **final value** of the axial angular momentum is therefore

$$\boldsymbol{L}_O\cdot\boldsymbol{k} = \left(\tfrac{2}{5}Ma^2\right)\Omega' + ma(\Omega' a) = \tfrac{1}{5}(2M+5m)a^2\Omega'.$$

Since $\boldsymbol{L}_O\cdot\boldsymbol{k}$ is known to be conserved it follows that

$$\tfrac{1}{5}(2M+5m)a^2\Omega' = \tfrac{2}{5}Ma^2\Omega,$$

and hence the **final angular velocity** of the ball is

$$\Omega' = \left(\frac{2M}{2M+5m}\right)\Omega. \blacksquare$$

Question ***Change in kinetic energy***

Find the change in kinetic energy of the system caused by the beetle's journey.

Answer

The initial and final kinetic energies of the system are

$$\tfrac{1}{2}\left(\tfrac{2}{5}Ma^2\right)\Omega^2 \quad\text{and}\quad \tfrac{1}{2}\left(\tfrac{2}{5}Ma^2\right)\Omega'^2 + \tfrac{1}{2}m(a\Omega')^2$$

respectively. On using the value of Ω' found above and simplifying, the **kinetic energy** of the system is found to *decrease* by

$$\frac{mMa^2\Omega^2}{2M+5m}. \blacksquare$$

Question ***Red hot beetle***

Does this loss of energy mean that the beetle arrives in a red-hot condition?

Answer

Your mechanics lecturer will be pleased to answer this question. ■

Our last example, the spherical pendulum, is a system with two degrees of freedom. By using both angular momentum and energy conservation, a complete solution can be found.

Example 11.9 ***The spherical pendulum: an integrable system***

A particle P of mass m is suspended from a fixed point O by a light inextensible string of length a and moves with the string taut in three-dimensional space (the spherical pendulum). Show that angular momentum about the vertical axis through O is conserved and express this conservation law in terms of the generalised coordinates θ, ϕ, as shown in Figure 11.7. Obtain also the corresponding equation for conservation of energy.

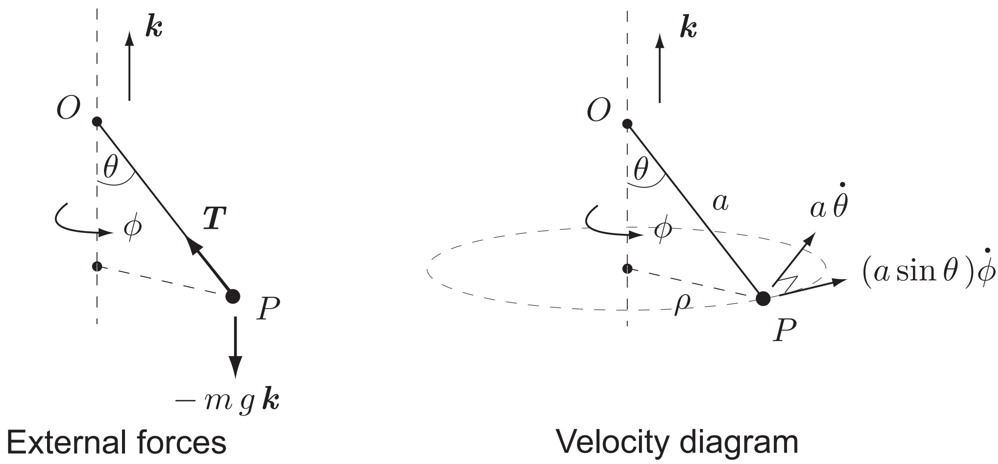


FIGURE 11.7 The spherical pendulum with generalised coordinates θ and ϕ. **Left**: the external forces. **Right**: the velocity diagram.

Initially the string makes an acute angle α with the downward vertical and the particle is projected with speed u in a horizontal direction at right angles to the string. Determine the constants of the motion, and deduce an equation satisfied by $\theta(t)$ in the subsequent motion.

Solution

The external forces on the particle are gravity and the tension in the string (see Figure 11.7 (left)). Hence,

$$\boldsymbol{K}_O = \boldsymbol{r}\times(-mg\boldsymbol{k}) + \boldsymbol{r}\times\boldsymbol{T} = -mg\boldsymbol{r}\times\boldsymbol{k},$$

the second term being zero since $\boldsymbol{r}$ and $\boldsymbol{T}$ are parallel. It follows that

$$\boldsymbol{K}_O\cdot\boldsymbol{k} = -mg(\boldsymbol{r}\times\boldsymbol{k})\cdot\boldsymbol{k} = 0,$$

since the triple scalar product has two $\boldsymbol{k}$'s. Hence $\boldsymbol{L}_O\cdot\boldsymbol{k}$ is **conserved**.

In order to express this conservation law in terms of coordinates, we draw a velocity diagram for the system as explained in section 10.10. The velocities corresponding to the coordinates θ and ϕ are $a\dot{\theta}$ and $\rho\dot{\phi}$ ($=(a\sin\theta)\dot{\phi}$) respectively in the directions shown in Figure 11.7 (right). These two velocities are perpendicular, with the $(a\sin\theta)\dot{\phi}$ contribution in the *azimuthal* direction around the vertical axis $\{O, \boldsymbol{k}\}$. It follows that $\boldsymbol{v}\cdot\widehat{\boldsymbol{\phi}} = (a\sin\theta)\dot{\phi}$. The required axial angular momentum is therefore

$$\boldsymbol{L}_O\cdot\boldsymbol{k} = m\rho(\boldsymbol{v}\cdot\widehat{\boldsymbol{\phi}}) = m(a\sin\theta)(a\sin\theta\,\dot{\phi}) = ma^2\sin^2\theta\,\dot{\phi}$$

and **conservation** of $\boldsymbol{L}_O\cdot\boldsymbol{k}$ is expressed by

$$ma^2\sin^2\theta\,\dot{\phi} = L,$$

where the axial angular momentum L is a constant of the motion.

The diagram also shows that the kinetic energy of P is given by

$$\tfrac{1}{2}m\left((a\dot{\theta})^2 + (a\sin\theta\,\dot{\phi})^2\right)$$

and the potential energy by $V = -mg(a\cos\theta)$. **Conservation of energy** therefore requires that

$$\tfrac{1}{2}m\left((a\dot{\theta})^2 + (a\sin\theta\,\dot{\phi})^2\right) - mg(a\cos\theta) = E,$$

where the total energy E is a constant of the motion.

From the prescribed **initial conditions**,

$$L = m(a\sin\alpha)u, \qquad E = \tfrac{1}{2}mu^2 - mga\cos\alpha,$$

so that the subsequent motion of P satisfies the **conservation equations**

$$ma^2\sin^2\theta\,\dot{\phi} = ma\sin\alpha\,u, \tag{11.16}$$

$$\tfrac{1}{2}m\left(a^2\dot{\theta}^2 + a^2\sin^2\theta\,\dot{\phi}^2\right) - mga\cos\theta = \tfrac{1}{2}mu^2 - mga\cos\alpha. \tag{11.17}$$

Since the spherical pendulum has two degrees of freedom, these two conservation equations are **sufficient to determine the motion**. Moreover, the system is **integrable** (see section (10.10) so that it must be possible to reduce the solution of the problem to integrations.

The equations (11.16), (11.17) are a pair of *coupled* first order ODEs for the unknown functions $\theta(t)$, $\phi(t)$. However, because the coordinate ϕ only appears as $\dot{\phi}$ in both equations, ϕ can be eliminated (θ can not!). From equation (11.16) we have

$$\dot{\phi} = \frac{u\sin\alpha}{a\sin^2\theta} \tag{11.18}$$

and this can now be substituted into equation (11.17) to obtain an equation for $\theta(t)$ alone. After some algebra we find that

$$\dot{\theta}^2 = \frac{u^2}{a^2}(\cos\alpha - \cos\theta)\left(\frac{\cos\alpha + \cos\theta}{\sin^2\theta} - \frac{2ag}{u^2}\right), \tag{11.19}$$

which is the **required equation** satisfied by $\theta(t)$. On taking square roots, this equation becomes a first order *separable* ODE whose solution can be written as an integral. Now that $\theta(t)$ is 'known', $\phi(t)$ can be found (as another integral) from equation (11.18). Thus, as predicted, the solution has thus been reduced to integrations. ■

Question *Form of the motion*

This is all very well, but what does the motion actually look like?

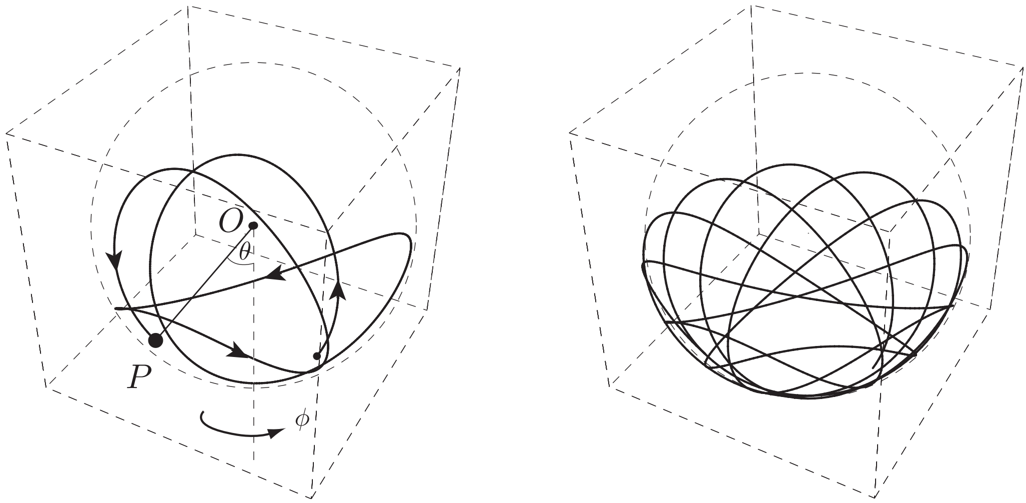


FIGURE 11.8 The calculated path of the spherical pendulum for the case $\alpha = \pi/6$ and $u^2/ag = 1.9$. **Left**: After four oscillations of θ. **Right**: After ten oscillations of θ. The surrounding boxes show the perspective.

Answer

Despite the problem being called integrable, the integrals arising from the separation procedure cannot be evaluated and no explicit solution is possible. However, equation (11.19) has the form of an energy equation for a system with one degree of freedom. We have met this situation before with the radial motion equation in orbit theory and the deductions we can make are the same. Because the left side of (11.19) is positive, it follows that the motion is restricted to those values of θ that make the function

$$F = (\cos\alpha - \cos\theta)\left(\frac{\cos\alpha + \cos\theta}{\sin^2\theta} - \frac{2ag}{u^2}\right)$$

positive. Moreover, maximium and minimum values of θ can only occur when $F(\theta) = 0$.

Since $F(\alpha) = 0$, $\theta = \alpha$ is one extremum* and any other extremum must be a root of the equation $G(\theta) = 0$, where

$$G = \frac{\cos\alpha + \cos\theta}{\sin^2\theta} - \frac{2ag}{u^2}.$$

Whether α is a maximum or minimum point of θ depends on the value of the initial projection speed u. On differentiating equation (11.19) with respect to t, we find that the initial value of $\ddot{\theta}$ is given by

$$\ddot{\theta}\big|_{\theta=\alpha} = \frac{u^2}{a^2}\left(\frac{\cos\alpha}{\sin^2\alpha} - \frac{ag}{u^2}\right),$$

* This is because of the form of the initial conditions.

so that θ initially *increases* if $u^2/ag > \sin^2\alpha/\cos\alpha$, and θ initially *decreases* if $u^2/ag < \sin^2\alpha/\cos\alpha$. (The critical case corresponds to the special case of conical motion.) Suppose that the first condition holds. Then $\theta_{\min} = \alpha$ and $\theta_{\max}$ must be a root of the equation $G(\theta) = 0$. Since $G(\alpha) > 0$ and $G(\pi - \alpha) < 0$, such a root does exist and is less than $\pi - \alpha$.

For example, consider the particular case in which $\alpha = \pi/3$ and $u^2 = 4ag$. Then the equation $G(\theta) = 0$ simplifies to give

$$\cos\theta(\cos\theta + 2) = 0,$$

from which it follows that $\theta_{\max} = \pi/2$. Hence, in this case, θ oscillates periodically in the range $\pi/3 \le \theta \le \pi/2$.

At the same time as the coordinate θ oscillates, the coordinate ϕ increases in accordance with equation (11.18). Hence, during each oscillation period τ of the coordinate θ, ϕ increases by

$$\Phi = \frac{u\sin\alpha}{a}\int_0^{\tau}\frac{dt}{\sin^2\theta}.$$

This *pattern* of motion repeats itself with period τ, but the motion is only truely periodic if it eventually links up with itself; this occurs only when the initial conditions are such that Φ/π is a *rational number*.

Figure 11.8 shows an actual path of the spherical pendulum, corresponding to the initial conditions $\alpha = \pi/6$ and $u^2/g = 1.9$. The results are entirely consistent with the theory above. ■

11.6 PLANAR RIGID BODY MOTION

What is planar motion?

Planar motion is a generalisation of two-dimensional motion in which two-dimensional methods are still valid. The complications of the full three-dimensional theory (represented by the two *vector* equations (11.15)) melt away to leave three *scalar* equations, which often have a very simple form. This enables a variety of fascinating problems to be solved in simple closed form. Planar rigid body motion is good value for money!

A system is said to be in **planar motion** if each of its particles moves in a plane and all of these planes are parallel to a fixed plane $\mathcal{P}$ called the **plane of motion**. For example, any *purely translational* motion of a rigid body in which G moves in the plane $\mathcal{P}$ is a planar motion, as is any *purely rotational* motion about an axis through G that is perpendicular to $\mathcal{P}$. The same is true when both of these motions are present together. For example, a cylinder (of any cross-section) rolling down a rough inclined plane is in planar motion. It is not necessary for the bodies that make up our system to be cylinders, nor even to be bodies of revolution. We merely suppose that *each constituent of the system should have reflective symmetry in the plane of motion*,* as shown in Figure 11.9.

* The reason for this symmetry restriction is that, if a body of completely general shape (a potato, say) were started in planar motion and moved under realistic forces (uniform gravity, say), then *the motion would*

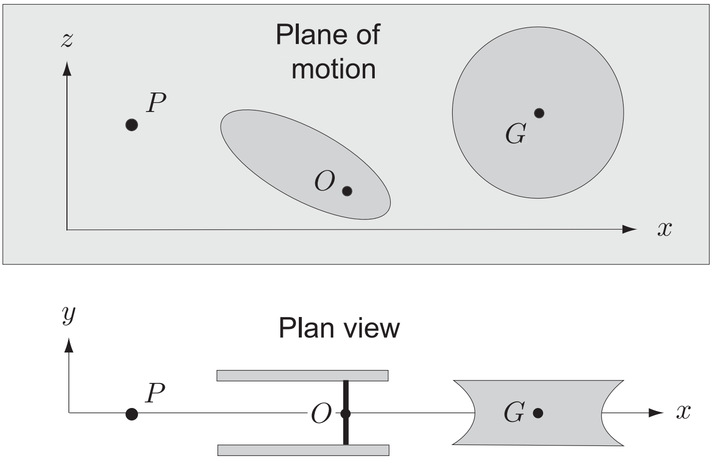


FIGURE 11.9 Three typical elements of a system in planar motion. The particle P moves in the plane of motion $y = 0$; the elliptical crank rotates about the fixed axis $\{O, \boldsymbol{j}\}$; and the circular pulley is in general planar motion. In the last case, G moves in the plane of motion, and the pulley also rotates about the axis $\{G, \boldsymbol{j}\}$.

Question ***Bodies in planar motion***

Decide whether the following rigid bodies are in planar motion: (i) a cotton reel rolling on a table, (ii) the Earth, and (iii) a snooker ball rolling after being struck with 'side'. [If you don't know what this means, get a player to show you.]

Answer

(i) Yes. (ii) No, because the Earth's rotation axis is not perpendicular to the plane of its orbit. (iii) No, because the rotation axis of the ball is not horizontal when the ball is struck with 'side'. ∎

Angular momentum in planar motion

The fact that makes planar motion so special is that the total **angular momentum** of each rigid body in the system has a **constant direction** normal to the plane of motion. This follows from the reflective symmetry that each body has in the plane of motion. In other words, if A is any point lying in the plane of motion, then the angular momentum of each rigid body about A has the simple form

$$\boldsymbol{L}_A = L_A\,\boldsymbol{j},$$

where, as shown in Figure 11.9, the plane of motion has been taken to be $y = 0$, and L_A is a short form for the axial angular momentum $\boldsymbol{L}_A \cdot \boldsymbol{j}$. Similar remarks apply to the total moment about A of the external forces acting on each rigid body. This follows from the

not remain planar. However, if the system and the external forces have reflective symmetry in the plane of motion, then a motion that is initially planar will remain planar.

supposed symmetry of these forces about the plane of motion. Hence the total moment about A of the external forces acting on each rigid body has the form

$$\boldsymbol{K}_A = K_A\, \boldsymbol{j},$$

where K_A is a short form for the axial moment $\boldsymbol{K}_A \cdot \boldsymbol{j}$. If A is a fixed point, or the centre of mass of the body, the angular momentum principle for each rigid body then takes the form

$$\frac{d}{dt}\left(L_A\, \boldsymbol{j}\right) = K_A\, \boldsymbol{j}$$

which, **since $\boldsymbol{j}$ is a constant vector**, reduces to the scalar equation

$$\frac{dL_A}{dt} = K_A.$$

Since the axis of rotation of the body in its motion relative to A is also normal to the plane of motion, it follows from equation (11.8) that

$$L_A = \boldsymbol{L}_A \cdot \boldsymbol{j} = I_A\, \omega,$$

where ω is the angular velocity of the body, and I_A its moment of inertia, about the axis $\{A, \boldsymbol{j}\}$. We therefore obtain the planar angular momentum principle in the form

$$\frac{d}{dt}(I_A\, \omega) = K_A, \qquad (11.20)$$

where A is either some fixed point in the plane of motion, or the centre of mass of the body. In applications, the moment of inertia I_A is usually constant.

Planar rigid body equations

We are now in a position to reduce the full rigid body equations (11.15) to planar form. Since there is no motion in the $\boldsymbol{j}$-direction, only the $\boldsymbol{i}$- and $\boldsymbol{k}$-components of the linear momentum principle survive, and, as we showed in the last section, only the $\boldsymbol{j}$-component of the angular momentum principle survives. Thus, each rigid body in planar motion satisfies the *three scalar equations of motion*:

Planar rigid body equations

$$M\frac{dV_x}{dt} = F_x \qquad M\frac{dV_z}{dt} = F_z \qquad I_G\frac{d\omega}{dt} = K_G \qquad (11.21)$$

where we have taken the angular momentum principle about the centre of mass G; I_G is then constant. These are the **planar rigid body equations**.

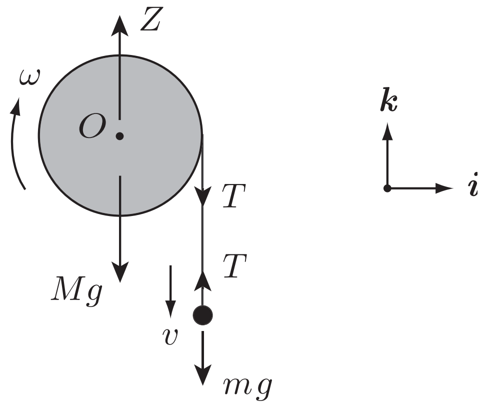


FIGURE 11.10 The pulley rotates about the fixed horizontal axis $\{O, \boldsymbol{j}\}$ and the suspended particle moves vertically.

A special case arises when the body is rotating about a *fixed* axis like the elliptical crank shown in Figure 11.9. Although the equations (11.21) could still be used, this is not the quickest way. Let the fixed axis be $\{O, \boldsymbol{j}\}$, where O lies in the plane of motion (see Figure 11.9). If the angular momentum principle is now applied about O (instead of G), then the unknown reactions exerted by the pivots make no contributuion to K_O and *do not appear* in the third equation of (11.21). The first two equations in (11.21) serve only to determine these reactions, once the motion has been calculated. Hence, unless the pivot reactions are actually required, it is sufficient to use the single equation:

Rigid body equation – fixed axis

$$I_O \frac{d\omega}{dt} = K_O \tag{11.22}$$

In this case also, the moment of inertia I_O is constant.

Example 11.10 *Planar motion: mass hanging from a pulley*

A circular pulley of mass M and radius a is smoothly pivoted about the axis $\{O, \boldsymbol{j}\}$, as shown in Figure 11.10. A light inextensible string is wrapped round the pulley so that it does not slip, and a particle of mass m is suspended from the free end. The system undergoes planar motion with the particle moving vertically. Find the downward acceleration of the particle.

Solution

This problem is most easily solved by using energy conservation, but it is instructive to solve it as a planar motion problem to illustrate the difference between the two approaches. In the energy conservation approach, the *whole system* is considered to be a single entity, and in this case, the string tensions do no *total* work and need not be considered. In the setting of planar motion however, the system consists of **two elements**, (i) a **particle** moving in a vertical straight line, and (ii) a **rigid pulley** rotating about a fixed horizontal axis. For each constituent, the tension force exerted by the string is **external**. The string tensions (which are equal in this problem) therefore appear in the equations of motion. For this reason, the conservation method is simpler, but there are many problems where there is no useful conservation principle and the planar motion approach is essential.

Let the particle have downward vertical velocity v, and the pulley have angular velocity ω (in the sense shown) at time t. Since the vector $\boldsymbol{j}$ points *into* the page, this is the positive sense around the axis $\{O, \boldsymbol{j}\}$. Thus, in the notation used here, *the **positive sense** for angular velocity is **clockwise**. The same applies to moments and angular momenta.*

First consider the motion of the **particle**. Since the motion is in a vertical straight line, the only surviving euation is

$$m\frac{dv}{dt} = mg - T,$$

where T is the string tension at time t.

Now consider the motion of the **pulley**. Since this is rotating about a *fixed* axis, the equation of motion is $I_O d\omega/dt = K_O$, that is,

$$I_O\frac{d\omega}{dt} = aT,$$

since the weight force Mg and the pivot reaction Z have zero moment about O.

Hence, the unknown tension T can be eliminated to give

$$\left(ma\frac{dv}{dt} + I_O\frac{d\omega}{dt}\right) = mga.$$

This equation applies whether or not the string slips on the pulley. However, since we are given that the string does not slip, v and ω must be related by the **no-slip condition** $v = \omega a$. On using this condition in the last equation, we obtain

$$\frac{dv}{dt} = \left(\frac{ma^2}{ma^2 + I_O}\right)g.$$

This is the downward **acceleration** of the particle. If the moment of inertia of the pulley is $\frac{1}{2}Ma^2$, then the value of this acceleration is $[2m/(2m + M)]g$. ■

Our next example, the cotton reel problem, is a famous problem in planar mechanics. The mathematics is elementary, but the solution needs to be interpreted carefully.

Example 11.11 *The cotton reel problem*

A cotton reel is at rest on a rough horizontal table when the free end of the thread is pulled horizontally with a constant force T, as shown in Figure 11.11. Given that the reel undergoes planar motion,* how does it move?

* In practice, it is impossible to maintain planar motion in the problem as described (try it). However, the problem is the same if the thread is replaced by a broad flat tape for which planar motion is easier to achieve.

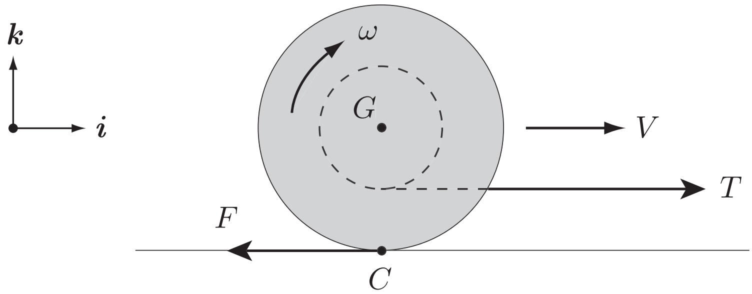


FIGURE 11.11 The reel is initially at rest on a rough horizontal table, when the free end of the thread is pulled with a constant force.

Solution

Suppose the ends of the reel have radius a, the axle wound with thread has radius b (with $b < a$), and the whole reel together with its thread has mass M. (We will neglect the mass of any thread pulled from the reel.) Let the reel have horizontal velocity V and angular velocity ω in the directions shown at time t. Note that we are not presuming the variables V, ω (or F) must take positive values. The signs of these variables will be deduced in the course of solving the problem.

The external forces on the reel are the string tension T and the fricton force F at the table. (The weight force and the normal reaction at the table cancel.) Hence, the equations of motion for the reel are

$$M\frac{dV}{dt} = T - F, \tag{11.23}$$

$$Mk^2\frac{d\omega}{dt} = aF - bT, \tag{11.24}$$

where Mk^2 is the moment of inertia of the reel about $\{G, \boldsymbol{j}\}$. We are not making any prior assumption about whether the reel slides or rolls. We will simply presume that the friction force F is bounded in magnitude by some maximum $F^{\max}$, that is,

$$-F^{\max} \le F \le F^{\max}, \tag{11.25}$$

and that $F = +F^{\max}$ (or $-F^{\max}$) when the reel is sliding forwards (or backwards).

Whether the reel slides or rolls depends on v^C, the velocity of the contact particle C. Since $v^C = V - \omega a$, it follows by manipulating the equations (11.23), (11.24) that v^C satisfies the equation

$$\left(\frac{Mk^2}{k^2+ab}\right)\frac{dv^C}{dt} = T - \gamma F, \tag{11.26}$$

where the constant γ is given by

$$\gamma = \frac{k^2+a^2}{k^2+ab}. \tag{11.27}$$

Different cases arise depending on how hard one pulls on the thread.

Strong pull $T > \gamma F^{\max}$

In this case, the right side of equation (11.26) is certain to be positive so that $dv^C/dt > 0$ for all t. Since the system starts from rest, $v^C = 0$ initially and so $v^C > 0$ for all $t > 0$. In other words, the **reel slides forwards**. This in turn implies that $F = F^{\max}$ so that the equations of motion (11.23), (11.24) become

$$\frac{dV}{dt} = \frac{T - F^{\max}}{M},$$
$$\frac{d\omega}{dt} = \frac{aF^{\max} - bT}{Mk^2}.$$

These equations imply that the reel slides forwards with **constant acceleration** and **constant angular acceleration**. Note that ω is positive for $\gamma F^{\max} < T < (a/b)F^{\max}$ and negative for $T > (a/b)F^{\max}$.

Gentle pull $T < \gamma F^{\max}$

In this case, the reel must **roll**. The proof of this is by contradiction, as follows.

Suppose that the reel were to *slide forwards* at any time in the subsequent motion. Then there must be a time τ, at which v^C and dv^C/dt are both positive. The condition $v^C > 0$ implies that $F = F^{\max}$ when $t = \tau$, and the condition $dv^C/dt > 0$ then implies that $T > \gamma F^{\max}$ when $t = \tau$. This is contrary to asumption and so forward sliding can never take place. A similar argument excludes backward sliding and so the only possibility is that the reel must roll.

The reel must therefore satisfy the rolling condition $V = \omega a$ and this, together with the equations of motion (11.23), (11.24), implies that the reel must **roll forwards** with constant acceleration

$$\frac{dV}{dt} = \frac{a(a-b)T}{M(k^2 + a^2)}. \blacksquare$$

Example 11.12 *A circus trick*

In a circus trick, a performer of mass m causes a large ball of mass M and radius a to accelerate to the right (see Figure 11.12) by running to the left on the upper surface of the ball. The man does not fall off the ball because he maintains this motion in such a way that the angle α shown remains constant. Find the conditions neccessary for such a motion to take place.

Solution

Suppose the motion is planar and that, at time t, the ball has velocity V in the $\boldsymbol{i}$-direction and angular velocity ω $(= V/a)$ around the axis $\{O, \boldsymbol{j}\}$. If the man is to maintain his position on the ball, then he must run up the surface of the ball (towards the highest point) with velocity V. This maintains his vertical height and his acceleraton is then the same as that of the ball, namely $(dV/dt)\boldsymbol{i}$.

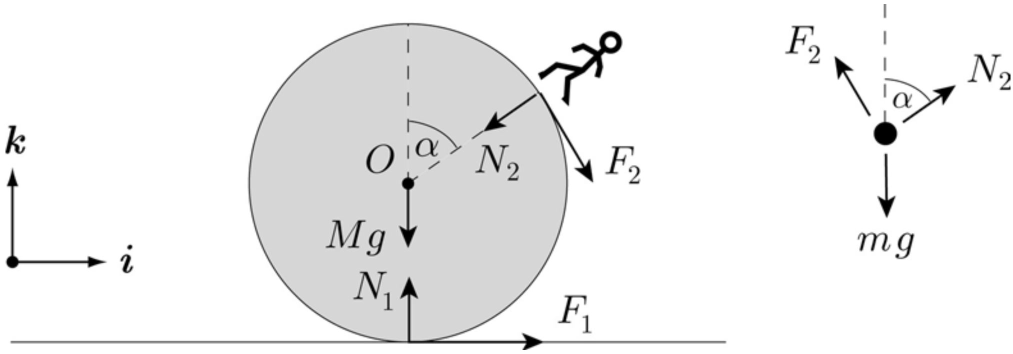


FIGURE 11.12 The circus trick: forces on the ball and the performer.

The equations of motion for the **man** are therefore

$$m\frac{dV}{dt} = N_2 \sin\alpha - F_2 \cos\alpha$$
$$0 = N_2 \cos\alpha + F_2 \sin\alpha - mg$$

and the equations of motion for the **ball** are

$$M\frac{dV}{dt} = F_1 - N_2 \sin\alpha + F_2 \cos\alpha$$
$$0 = N_1 - N_2 \cos\alpha - F_2 \sin\alpha - Mg$$
$$I_O\frac{d\omega}{dt} = aF_2 - aF_1$$

where I_O is the moment of inertia of the ball about $\{O, \boldsymbol{j}\}$.

These five equations, together with the rolling condition $V = \omega a$, are sufficient to determine the six unknowns dV/dt, $d\omega/dt$, F_1, N_1, F_2 and N_2. After some algebra, the solution for the forward **acceleration** of the ball turns out to be

$$\frac{dV}{dt} = \frac{mg \sin\alpha}{M + (I_O/a^2) + m(1 + \cos\alpha)}.$$

Hence the motion is possible for any acute angle α provided that the performer can accelerate relative to the ball with this acceleration. For the case in which the ball is hollow, the masses of the man and the ball are equal, and $\alpha = 45°$, the acceleration required is approximately $0.21\ g$. ■

11.7 RIGID BODY STATICS IN THREE DIMENSIONS

Although we are not yet able to attempt problems in which a rigid body undergoes three-dimensional motion, we *are* able to solve problems in which a rigid body is in *equilibrium* under a three-dimensional system of forces. In equilibrium, the linear and angular momentum of the body are known to be zero, so that the rigid body equations become

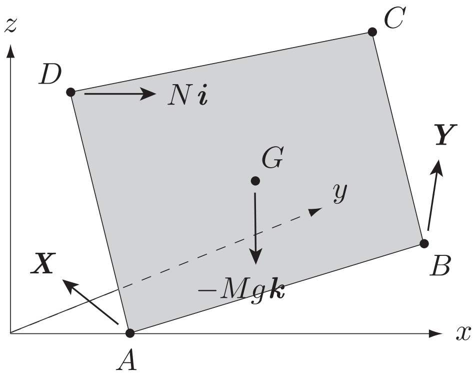


FIGURE 11.13 The rectangular panel $ABCD$ rests on the rough floor $z = 0$ and leans against the smooth wall $x = 0$.

Equations of rigid body statics

$$\boldsymbol{F} = \boldsymbol{0} \qquad \boldsymbol{K}_A = \boldsymbol{0} \tag{11.28}$$

where A is any fixed point of space.* In other words, *when a system is in equlibrium, the resultant force and the resultant moment of the external forces must be zero*.

Since there is no motion, one may wonder what there is left to calculate in statical problems. However, the body is usually supported or restrained in some prescribed way, and it is the unknown constraint forces that are to be determined. If these constraint forces can be determined solely from the equilibrium equations (11.28), then the problem is said to be **statically determinate**.†

Example 11.13 *Leaning panel*

A rectangular panel $ABCD$ of mass M is (rather carelessly) placed with its edge AB on the rough horizontal floor $z = 0$ and with the vertex D resting against the smooth wall $x = 0$, as shown in Figure 11.13. The four vertices of the panel are at the points $A(2, 0, 0)$, $B(6, 4, 0)$, $C(4, 6, 6)$ and $D(0, 2, 6)$ respectively. Given that the panel does not slip on the floor, find the reaction force exerted by the wall.

Solution

The external forces acting on the panel are the normal reaction of the wall $P\boldsymbol{i}$, the weight force $-Mg\boldsymbol{k}$, and the reaction of the floor on the edge AB. Now the reaction of the floor is distributed along the edge AB and, although we could treat it as such,

* It is unneccessary and inconvenient to restrict moments to be taken about G.

† Not all problems are statically determinate by any means. In the two-dimensional theory, if a heavy rigid plank is resting on three or more supports, then the individual reactions at the supports cannot be found from the equlibrium equations. What this means is that modelling the plank as a *rigid* body is not appropriate in such a problem. One should instead model the plank as a *deformable* body, solve the problem using the theory of elasticity, and then pass to the rigid limit.

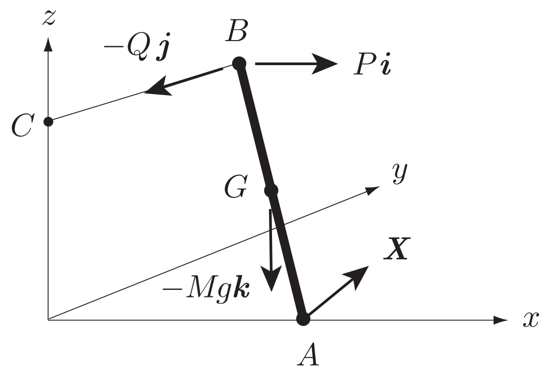


FIGURE 11.14 The rod AB is in equilibrium with A on a rough floor and B resting against a smooth wall. The rod is prevented from falling by the string BC.

this is an irrelevant complication. We will therefore suppose that (in order to avoid damage to the floor) the panel has been supported on two small pads beneath the corners A and B, in which case the reaction of the floor consists of the forces $\boldsymbol{X}$ and $\boldsymbol{Y}$ as shown.

We now apply the equilibrium conditions. The condition $\boldsymbol{F} = \boldsymbol{0}$ yields

$$N\boldsymbol{i} + \boldsymbol{X} + \boldsymbol{Y} - Mg\boldsymbol{k} = \boldsymbol{0}. \tag{11.29}$$

If we take moments about the corner A, the reaction $\boldsymbol{X}$ makes no contribution and the condition $\boldsymbol{K}_A = \boldsymbol{0}$ becomes

$$\overrightarrow{AD} \times (N\boldsymbol{i}) + \overrightarrow{AB} \times \boldsymbol{Y} + \overrightarrow{AG} \times (-Mg\boldsymbol{k}) = \boldsymbol{0}.$$

On inserting the given numbers, this equation becomes

$$N(6\boldsymbol{j} - 2\boldsymbol{k}) + (4\boldsymbol{i} + 4\boldsymbol{j}) \times \boldsymbol{Y} - Mg\,(3\boldsymbol{i} - \boldsymbol{j}) = \boldsymbol{0}. \tag{11.30}$$

The six scalar equations in (11.29), (11.30) contain the seven scalar unknowns $\boldsymbol{X}$, $\boldsymbol{Y}, N$ which means that the problem is *not statically determinate*. However, this does not stop us from finding N, since $\boldsymbol{Y}$ can be eliminated from equation (11.30) by taking the scalar product with the vector $4\boldsymbol{i}+4\boldsymbol{j}$. (This is equivalent to taking moments about the *axis* AB so that both $\boldsymbol{X}$ and $\boldsymbol{Y}$ disappear to leave N as the only unknown.) This gives the **reaction exerted** by the wall to be $N = Mg/3$. ■

Example 11.14 *Leaning rod*

A rough floor lies in the horizontal plane $z = 0$ and the plane $x = 0$ is occupied by a smooth vertical wall. A uniform rod of mass M has its lower end on the floor at the point $(a, 0, 0)$ and its upper end rests in contact with the wall at the point $(0, b, c)$. The rod is prevented from falling by having its upper end connected to the point $(0, 0, c)$ by a light inextensible string. Given that the rod does not slip, find the tension in the string and the reaction exerted by the wall.

Solution

The external forces acting on the rod are the normal reaction $N\boldsymbol{i}$ of the wall, the tension force $-Q\boldsymbol{j}$ in the string, the weight force $-Mg\boldsymbol{k}$, and reaction $\boldsymbol{X}$ of the floor. The equilibrium equations are therefore

$$P\boldsymbol{i} - Q\boldsymbol{j} - Mg\boldsymbol{k} + \boldsymbol{X} = \boldsymbol{0}, \tag{11.31}$$

$$(2L\boldsymbol{n})\times(P\boldsymbol{i} - Q\boldsymbol{j}) + (L\boldsymbol{n})\times(-Mg\boldsymbol{k}) = \boldsymbol{0}, \tag{11.32}$$

where we have taken moments about A to eliminate the reaction $\boldsymbol{X}$. Here, $2L$ is the length of the rod, and $\boldsymbol{n}$ is the unit vector in the direction $\overrightarrow{AB}$.

Equation (11.31) serves only to determine the reaction $\boldsymbol{X}$ once P and Q are known. To extract P and Q from the vector equation (11.32), we take components in any two directions other than the $\boldsymbol{n}$-direction; the easiest choices are the $\boldsymbol{i}$- and $\boldsymbol{j}$-directions. On taking the scalar product of equation (11.32) with $\boldsymbol{i}$, we obtain

$$\begin{aligned} 0 &= 2\,[\boldsymbol{n}, P\boldsymbol{i} - Q\boldsymbol{j}, \boldsymbol{i}] + [\boldsymbol{n}, -Mg\boldsymbol{k}, \boldsymbol{i}] \\ &= 2P\,[\boldsymbol{n}, \boldsymbol{i}, \boldsymbol{i}] - 2Q\,[\boldsymbol{n}, \boldsymbol{j}, \boldsymbol{i}] - Mg\,[\boldsymbol{n}, \boldsymbol{k}, \boldsymbol{i}] \\ &= 0 - 2Q\,\boldsymbol{n}\cdot(\boldsymbol{j}\times\boldsymbol{i}) - Mg\,\boldsymbol{n}\cdot(\boldsymbol{k}\times\boldsymbol{i}) \\ &= 2Q\,(\boldsymbol{n}\cdot\boldsymbol{k}) - Mg\,(\boldsymbol{n}\cdot\boldsymbol{j}), \end{aligned}$$

where we have used the notation $[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}]$ to mean the triple scalar product of the vectors $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$. Hence

$$Q = \frac{Mg\,(\boldsymbol{n}\cdot\boldsymbol{j})}{2\,(\boldsymbol{n}\cdot\boldsymbol{k})},$$

and, by taking the scalar product of equation (11.32) with $\boldsymbol{j}$ and proceed in the same way, we obtain

$$P = -\frac{Mg\,(\boldsymbol{n}\cdot\boldsymbol{i})}{2\,(\boldsymbol{n}\cdot\boldsymbol{k})}.$$

Finally, we need to express these answers in terms of the data given in the question. Since the unit vector $\boldsymbol{n}$ is given by

$$\boldsymbol{n} = \frac{-a\boldsymbol{i} + b\boldsymbol{j} + c\boldsymbol{k}}{2L},$$

it follows that the **reaction** exerted by the wall, and the **tension** in the string, are

$$P = \frac{Mga}{2c}, \qquad Q = \frac{Mgb}{2c}. \blacksquare$$

Problems on Chapter 11

Answers and comments are at the end of the book.

Harder problems carry a star (∗).

11.1 ***Non-standard angular momentum principle*** If A is a generally moving point of space and $\boldsymbol{L}_A$ is the angular momentum of a system $\mathcal{S}$ about A *in its motion relative to* A, show that the angular momentum principle for $\mathcal{S}$ about A takes the non-standard form

$$\frac{d\boldsymbol{L}_A}{dt} = \boldsymbol{K}_A - M(\boldsymbol{R} - \boldsymbol{a}) \times \frac{d^2\boldsymbol{a}}{dt^2}.$$

[Begin by expanding the expression for $\boldsymbol{L}_A$.]

When does this formula reduce to the standard form? [This non-standard version of the angular momentum principle is rarely needed. However, see Problem 11.9.]

Problems soluble by conservation principles

11.2 A fairground target consists of a uniform circular disk of mass M and radius a that can turn freely about a diameter which is fixed in a vertical position. Initially the target is at rest. A bullet of mass m is moving with speed u along a horizontal straight line at right angles to the target. The bullet embeds itself in the target at a point distance b from the rotation axis. Find the final angular speed of the target. [The moment of inertia of the disk about its rotation axis is $Ma^2/4$.]

Show also that the energy lost in the impact is

$$\frac{1}{2}mu^2\left(\frac{Ma^2}{Ma^2 + 4mb^2}\right).$$

11.3 A uniform circular cylinder of mass M and radius a can rotate freely about its axis of symmetry which is fixed in a vertical position. A light string is wound around the cylinder so that it does not slip and a particle of mass m is attached to the free end. Initially the system is at rest with the free string taut, horizontal and of length b. The particle is then projected horizontally with speed u at right angles to the string. The string winds itself around the cylinder and eventually the particle strikes the cylinder and sticks to it. Find the final angular speed of the cylinder.

11.4 ***Rotating gas cloud*** A cloud of interstellar gas of total mass M can move freely in space. Initially the cloud has the form of a uniform sphere of radius a rotating with angular speed Ω about an axis through its centre. Later, the cloud is observed to have changed its form to that of a thin uniform circular disk of radius b which is rotating about an axis through its centre and perpendicular to its plane. Find the angular speed of the disk and the increase in the kinetic energy of the cloud.

11.5 ***Conical pendulum with shortening string*** A particle is suspended from a support by a light inextensible string which passes through a small fixed ring vertically below the support. Initially the particle is performing a conical motion of angle 60°, with the moving part of the string of a. The support is now made to move slowly upwards so that the motion remains nearly conical. Find the angle of this conical motion when the support has been raised by a distance $a/2$. [Requires the numerical solution of a trigonometric equation.]

11.6 ***Baseball bat*** A baseball bat has mass M and moment of inertia Mk^2 about any axis through its centre of mass G that is perpendicular to the axis of symmetry. The bat is at rest when a ball of mass m, moving with speed u, is normally incident along a straight line through the axis of symmetry at a distance b from G. Show that, whether the impact is elastic or not, there is a point on the axis of symmetry of the bat that is instantaneously at rest after the impact and that the distance c of this point from G is given by $bc = k^2$. In the elastic case, find the speed of the ball after the impact. [Gravity (and the batter!) should be ignored throughout this question.]

11.7 ***Hoop mounting a step*** A uniform hoop of mass M and radius a is rolling with speed V along level ground when it meets a step of height h ($h < a$). The particle C of the hoop that makes contact with the step is suddenly brought to rest. Find the instantaneous speed of the centre of mass, and the instantaneous angular velocity of the hoop, immediately after the impact. Deduce that the particle C cannot remain at rest on the edge of the step if

$$V^2 > (a-h)g\left(1-\frac{h}{2a}\right)^{-2}.$$

Suppose that the particle C *does* remain on the edge of the step. Show that the hoop will go on to mount the step if

$$V^2 > hg\left(1-\frac{h}{2a}\right)^{-2}.$$

Deduce that the hoop cannot mount the step in the manner described if $h > a/2$.

11.8 ***Particle sliding on a cone*** A particle P slides on the smooth inner surface of a circular cone of semi-angle α. The axis of symmetry of the cone is vertical with the vertex O pointing downwards. Show that the vertical component of angular momentum about O is conserved in the motion. State a second dynamical quantity that is conserved.

Initially P is a distance a from O when it is projected horizontally along the inside surface of the cone with speed u. Show that, in the subsequent motion, the distance r of P from O satisfies the equation

$$\dot{r}^2 = (r-a)\left[\frac{u^2(r+a)}{r^2} - 2g\cos\alpha\right].$$

Case A For the case in which gravity is absent, find r and the azimuthal angle ϕ explicitly as functions of t. Make a sketch of the path of P (as seen from 'above') when $\alpha = \pi/6$.

Case B For the case in which $\alpha = \pi/3$, find the value of u such that r oscillates between a and $2a$ in the subsequent motion. With this value of u, show that r will first return to the value $r = a$ after a time

$$2\sqrt{3}\left(\frac{a}{g}\right)^{1/2}\int_1^2 \frac{\xi\, d\xi}{[(\xi-1)(2-\xi)(2+3\xi)]^{1/2}}.$$

11.9* ***Bug running on a hoop*** A uniform circular hoop of mass M can slide freely on a smooth horizontal table, and a bug of mass m can run on the hoop. The system is at rest when the bug starts to run. What is the angle turned through by the hoop when the bug has completed one lap of the hoop? [This is a classic problem, but difficult. Apply the angular momentum principle about the centre of the hoop, using the *non-standard* version given in Problem 11.1]

Planar rigid body motion

11.10 ***General rigid pendulum*** A rigid body of general shape has mass M and can rotate freely about a fixed horizontal axis. The centre of mass of the body is distance h from the rotation axis, and the moment of inertia of the body about the rotation axis is I. Show that the period of small oscillations of the body about the downward equilibrium position is

$$2\pi\left(\frac{I}{Mgh}\right)^{1/2}.$$

Deduce the period of small oscillations of a uniform rod of length $2a$, pivoted about a horizontal axis perpendicular to the rod and distance b from its centre.

11.11 ***From sliding to rolling*** A snooker ball is at rest on the table when it is projected forward with speed V and no angular velocity. Find the speed of the ball when it eventually begins to roll. What proportion of the original kinetic energy is lost in the process?

11.12 ***Rolling or sliding?*** A uniform ball is released from rest on a rough plane inclined at angle α to the horizontal. The coefficient of friction between the ball and the plane is μ. Will the ball roll or slide down the plane? Find the acceleration of the ball in each case.

11.13 A circular disk of mass M and radius a is smoothly pivoted about its axis of symmetry which is fixed in a horizontal position. A bug of mass m runs with constant speed u around the rim of the disk. Initially the disk is held at rest and is released when the bug reaches its lowest point. What is the condition that the bug will reach the highest point of the disk?

11.14 ***Yo-yo with moving support*** A uniform circular cylinder (a yo-yo) has a light inextensible string wrapped around it so that it does not slip. The free end of the string is fastened to a support and the yo-yo moves in a vertical straight line with the straight part of the string also vertical. At the same time the support is made to move vertically having upward displacement $Z(t)$ at time t. Find the acceleration of the yo-yo. What happens if the system starts from rest and the support moves upwards with acceleration $2g$?

11.15 ***Supermarket belt*** A circular cylinder, which is axially symmetric but not uniform, has mass M and moment of inertia Mk^2 about its axis of symmetry. The cylinder is placed on a rough horizontal belt at right angles to the direction in which the belt can move. Initially the cylinder and the belt are both at rest when the belt begins to move with velocity $V(t)$. Given that there is no slipping, find the velocity of the cylinder at time t.

Explain why drinks bottles tend to spin on a supermarket belt (instead of moving forwards) if they are placed at right-angles to the belt.

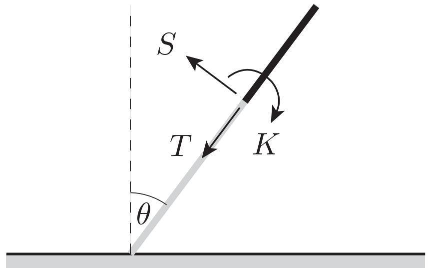


FIGURE 11.15 The tension force T, the shear force S and the couple K exerted on the the upper part of the rod (black) by the lower part (grey).

11.16* ***Falling chimney*** A uniform rod of length $2a$ has one end on a rough table and is balanced in the vertically upwards position. The rod is then slightly disturbed. Given that its lower end does not slip, show that, in the subsequent motion, the angle θ that the rod makes with the upward vertical satisfies the equation

$$2a\dot{\theta}^2 = 3g(1 - \cos\theta).$$

Consider now the the *upper part* of the rod of length $2\gamma a$, as shown in Figure 11.15. Let T, S and K be the tension force, the shear force and the couple exerted on the upper part of the rod by the lower part. By considering the upper part of the rod to be a rigid body in planar motion, find expressions for S and K in terms of θ.

If a tall thin chimney begins to fall, at what point along its length would you expect it to break first?

Rigid body statics

11.17 ***Leaning triangular panel*** A rough floor lies in the horizontal plane $z = 0$ and the planes $x = 0$, $y = 0$ are occupied by smooth vertical walls. A rigid uniform triangular panel ABC has mass m. The vertex A of the panel is placed on the floor at the point (2, 2, 0) and the vertices B, C rest in contact with the walls at the points (0, 1, 6), (1, 0, 6) respectively. Given that the vertex A does not slip, find the reactions exerted by the walls. Deduce the reaction exerted by the floor.

11.18 ***Triangular coffee table*** A trendy swedish coffee table has an unsymmetrical triangular glass top supported by a leg at each vertex. Show that, whatever the shape of the triangular top, each leg bears one third of its weight.

11.19 ***Pile of balls*** Three identical balls are placed in contact with each other on a horizontal table and a fourth identical ball is placed on top of the first three. Show that the four balls cannot be in equilibrium unless (i) the coefficient of friction between the balls is at least $\sqrt{3} - \sqrt{2}$, and (ii) the coefficient of friction between each ball and the table is at least $\frac{1}{4}(\sqrt{3} - \sqrt{2})$.

Part Three

ANALYTICAL MECHANICS

CHAPTERS IN PART THREE

Chapter Twelve

Lagrange's equations and conservation principles

KEY FEATURES

The key features of this chapter are **generalised coordinates** and **configuration space**, the derivation and use of **Lagrange's equations**, the **Lagrangian**, and the connection between **symmetry** of the Lagrangian and **conservation principles**.

Lagrange's equations mark a change in direction in our development of mechanics. Building on the work of d'Alembert, Lagrange* devised a general method for obtaining the **equations of motion** for a very wide class of mechanical systems. In earlier chapters we have used conservation principles for this purpose, but there is no guarantee that enough conservation principles exist. In contrast, Lagrange's method is completely general and is not restricted to problems soluble by conservation principles. The method is so simple to apply that it is quite possible to solve complex mechanical problems whilst knowing very little about mechanics! However, the supporting theory has its subtleties.

Lagrange's equations also mark the beginning of **analytical mechanics** in which general principles, such as the connection between symmetry and conservation principles, begin to take over from actual problem solving.

12.1 CONSTRAINTS AND CONSTRAINT FORCES

A **general mechanical system** $\mathcal{S}$ consists of any number of particles P_1, P_2, ..., P_N. The particles of $\mathcal{S}$ may have interconnections of various kinds (light strings, springs and so on) and also be subject to external connections and constraints. These could include features such as a particle being forced to remain on a fixed surface or suspended from a

---

* Joseph-Louis Lagrange (Giuseppe Lodovico Lagrangia), (1736–1813). Although Lagrange is often considered to be French, he was in fact born in Turin, Italy and did not move to Paris until 1787. Lagrange had a long career in Turin and Berlin during which time he made major contributions to mechanics, fluid mechanics and the calculus of variations. His famous book *Mécanique Analitique*, published in Paris in 1788, is a definitive account of his contributions to mechanics. This work transformed mechanics into a branch of mathematical analysis. Perhaps to emphasise this, there is not a single diagram in the whole book!

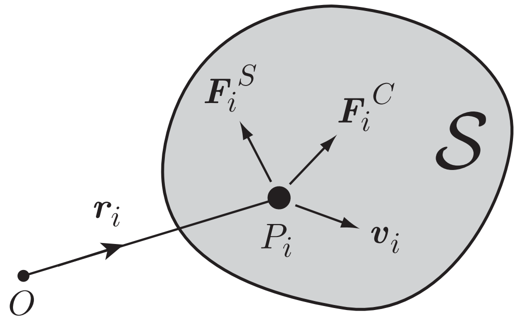


FIGURE 12.1 The general mechanical system $\mathcal{S}$ consists of any number of particles $\{P_i\}$ $(i = 1\ldots, N)$. The typical particle P_i has mass m_i, position vector $\boldsymbol{r}_i$ and velocity $\boldsymbol{v}_i$. $\boldsymbol{F}_i^S$ is the specified force and $\boldsymbol{F}_i^C$ the constraint force acting on P_i.

fixed point by a light inextensible string. The pendulum, the spinning top, the bicycle and the solar system are examples of mechanical systems.

Unconstrained systems

If the particles of $\mathcal{S}$ are free to move anywhere in space *independently of each other* then $\mathcal{S}$ is said to be an **unconstrained system**. In this special case, the equations of motion for $\mathcal{S}$ are simply Newton's equations for the N individual particles. Suppose that the typical particle P_i has mass m_i, position vector $\boldsymbol{r}_i$ and velocity $\boldsymbol{v}_i$. Then the **equations of motion** for the system $\mathcal{S}$ are

$$m_i \dot{\boldsymbol{v}}_i = \boldsymbol{F}_i \qquad (i = 1\ldots, N),$$

where $\boldsymbol{F}_i$ is the force acting on the particle P_i.

Example 12.1 *Two-body problem*

Write down the Newton equations for the two-body gravitation problem.

Solution

In this problem $\mathcal{S}$ consists of two particles moving solely under their mutual gravitational attraction. There are no constraints. The motion of the system is therefore governed by the two Newton equations

$$m_1 \dot{\boldsymbol{v}}_1 = m_1 m_2 G \frac{\boldsymbol{r}_2 - \boldsymbol{r}_1}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|^3}, \qquad m_2 \dot{\boldsymbol{v}}_2 = m_1 m_2 G \frac{\boldsymbol{r}_1 - \boldsymbol{r}_2}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|^3},$$

where G is the constant of gravitation. These equations, together with the initial conditions, are sufficient to determine the motion of the two particles. ■

Constrained systems

Unconstrained mechanical systems are relatively rare. Indeed many of the problems solved in earlier chapters involve mechanical systems that are subject to **geometrical** or **kinematical constraints**. Geometrical constraints are those that involve only the position vectors $\{\boldsymbol{r}_i\}$; kinematical constraints involve the $\{\boldsymbol{v}_i\}$ as well. Some **typical constraints** are as follows:

- The bob of a pendulum *must* remain a fixed distance from the point of support.
- The particles of a rigid body *must* maintain fixed distances from each other.
- A particle sliding on a wire *must not* leave the wire.
- The contact particle of a body rolling on a fixed surface *must* be at rest.

The rolling condition is a *kinematical constraint* since it involves the *velocity* of a particle. All the other constraints are geometrical.

These, and all other constraints, are enforced by **constraint forces**. Constraint forces are not part of the specification of a system and are therefore *unknown*. For example, when a particle is constrained to slide on a wire, it is prevented from leaving the wire by the force that the wire exerts upon it. This constraint force (which would commonly be called the reaction of the wire on the particle) is unknown; we know only that it is sufficient to keep the particle on the wire.

For **constrained systems** the straightforward approach of using the Newton equations runs into the following difficulties:

A The equations of motion do not incorporate the constraints
The Newton equations (in Cartesian coordinates) do not incorporate the constraints. These must therefore be included in the form of additional conditions to be *solved simultaneously* with the dynamical equations.

B The constraint forces are unknown
For constrained systems, the Newton equations have the form

$$m_i \dot{\boldsymbol{v}}_i = \boldsymbol{F}_i^S + \boldsymbol{F}_i^C \qquad (1 \le i \le N), \tag{12.1}$$

where $\boldsymbol{F}_i^S$ is the **specified force** and $\boldsymbol{F}_i^C$ is the **constraint force** acting on the particle P_i. The $\boldsymbol{F}_i^S$ are known but the $\boldsymbol{F}_i^C$ are not.

Because of these two difficulties, only the simplest problems of constrained motion are tackled this way. In the following sections we show how these difficulties can be overcome. The first difficulty is overcome by using a new (reduced) set of coordinates called **generalised coordinates**, while the second is overcome by using **Lagrange's equations** instead of Newton's.

12.2 GENERALISED COORDINATES

Suppose that the system is subject to **geometrical constraints** only. Then the position vectors $\{\boldsymbol{r}_i\}$ of its particles are not independent variables, but are related to each other by these constraints. A possible 'position' of such a system is called a **configuration**. More precisely, a set of values for the position vectors $\{\boldsymbol{r}_i\}$ that is *consistent with the geometrical constraints* is a configuration of the system.

The trick is to select new 'coordinates' that *are* independent of each other but are still sufficient to specify the configuration of the system. These new coordinates are called **generalised coordinates** and their official definition is as follows:

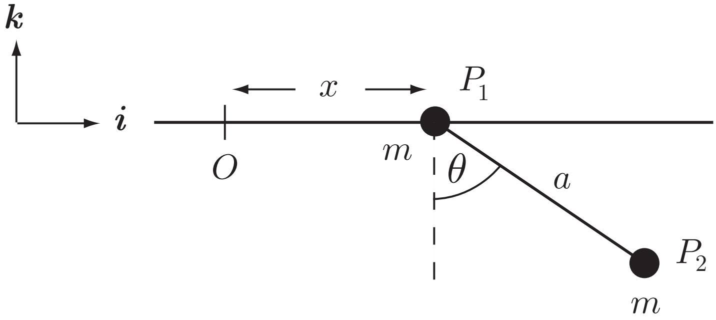


FIGURE 12.2 The variables x and θ are a set of generalised coordinates for this system.

Definition 12.1 ***Generalised coordinates*** *If the configuration of a system $\mathcal{S}$ is determined by the values of a set of independent variables $q_1, \dots, q_n$, then $\{q_1, \dots, q_n\}$ is said to be a set of* ***generalised coordinates*** *for $\mathcal{S}$.*

This definition deserves some explanation.

(i) When we say the generalised coordinates must be **independent variables**, we mean that there must be *no functional relation connecting them*. If there were, one of the coordinates could be removed and the remaining $n - 1$ coordinates would still determine the configuration of the system. The set of generalised coordinates must not be reducible in this way.

(ii) When we say the generalised coordinates $q_1, \dots, q_n$ **determine the configuration** of the system $\mathcal{S}$, we mean that, when the values of the coordinates $q_1, \dots, q_n$ are given, the position of every particle of $\mathcal{S}$ is determined. In other words, the position vectors $\{\boldsymbol{r}_i\}$ of the particles must be known functions of the independent variables $q_1, \dots, q_n$, that is,

$$\boldsymbol{r}_i = \boldsymbol{r}_i(q_1, \dots, q_n) \qquad (i = 1, \dots, N). \tag{12.2}$$

Abstract though this concept may seem, generalised coordinates are remarkably easy to use. In practice, *they are chosen to be displacements or angles that appear naturally in the problem*. This is illustrated by the following examples.

Example 12.2 *Choosing generalised coordinates*

Let $\mathcal{S}$ be the system shown in Figure 12.2 which consists of two particles P_1 and P_2 connected by a light rigid rod of length a. The particle P_1 is constrained to move along a fixed horizontal rail and the system moves in the vertical plane through the rail. Select generalised coordinates for this system and obtain expressions for the position vectors $\boldsymbol{r}_1$, $\boldsymbol{r}_2$ in terms of these coordinates.

Solution

Consider the variables x, θ shown. These are certainly independent variables (they are not connected by any functional relation) and, when they are given, the configuration of $\mathcal{S}$ is determined. Thus $\{x, \theta\}$ is a set of **generalised coordinates** for the system $\mathcal{S}$.

In terms of the coordinates x and θ, the positions of the particles P_1 and P_2 are given by

$$\begin{aligned} \boldsymbol{r}_1 &= x\,\boldsymbol{i}, \\ \boldsymbol{r}_2 &= (x + a\sin\theta)\,\boldsymbol{i} - (a\cos\theta)\,\boldsymbol{k}, \end{aligned}$$

which are the expressions (12.2) for this system and this choice of coordinates. ■

Example 12.3 *Choosing more generalised coordinates*

Choose generalised coordinates for the system consisting of three particles P_1, P_2, P_3 where P_1, P_2 are connected by a light rigid rod of length a and P_2, P_3 are connected by a light rigid rod of length b. The system slides on a horizontal table. [Make a sketch of the system.]

Solution

Many choices of generalised coordinates are possible. Let $Oxyz$ be a system of rectangular coordinates with O on the table and Oz pointing vertically upwards. One set of generalised coordinates consists of (i) the x and y coordinates of the particle P_1, (ii) the angle θ between the line P_1P_2 and the x-axis, (iii) the angle ϕ between the line P_2P_3 and the x-axis. A second set of generalised coordinates consists of (i) the x and y coordinates of the particle P_2, (ii) the angle θ between the line P_1P_2 and the y-axis, (iii) the angle ϕ between the line P_1P_2 and the line P_2P_3. ■

Degrees of freedom

It is evident from the above example that the configuration of a system can be specified by many different sets of generalised coordinates. However the *number of coordinates needed is always the same*. In the last example, the number of generalised coordinates needed is *always* three.

Definition 12.2 ***Degrees of freedom*** *Let $\mathcal{S}$ be a mechanical system subject to geometrical constraints. Then the* ***number*** *of generalised coordinates needed to specify the configuration of $\mathcal{S}$ is called the number of* ***degrees of freedom*** *of $\mathcal{S}$.*

The number of degrees of freedom is an important property of a mechanical system. Suppose, for example, that we have a system with *three* degrees of freedom and generalised coordinates q_1, q_2, q_3. Suppose also that the system is in some given configuration when it is started into motion in some given way. This means that we know the initial values of the coordinates q_1, q_2, q_3, and their time derivatives $\dot{q}_1, \dot{q}_2, \dot{q}_3$. How many equations of motion (second order ODEs) do we need to determine the functions $q_1(t)$, $q_2(t)$, $q_3(t)$ that describe the subsequent motion of the system? The answer is provided by the general theory of ODEs. If the three functions $q_1(t)$, $q_2(t)$, $q_3(t)$ satisfy *three* (independent) second order ODEs, then the general theory guarantees that there is precisely one solution that satisfies the prescribed initial conditions. If there are fewer equations, the solution is not uniquely determined; if there are more, the equations are not independent. This gives us the following important result:

Degrees of freedom and equations of motion

The number of degrees of freedom of a system is equal to the number of equations of motion (second order ODEs) that are needed to determine the motion of the system.

Example 12.4 ***Degrees of freedom***

State the number of degrees of freedom of the following mechanical systems: (i) the simple pendulum, (ii) the spherical pendulum, (iii) a door swinging on its hinges, (iv) a bar of soap (a particle) sliding on the inside of a basin, (v) four rigid rods flexibly jointed to form a quadrilateral which can move on a flat table, (vi) a ball rolling on a rough table.

Solution

(i) 1 (ii) 2 (iii) 1 (iv) 2 (v) 4 (vi) Not defined! This system has a *kinematical* constraint, namely, the rolling condition at the contact point. ■

Kinematical constraints

So far we have not discussed kinematical constraints such as the rolling condition. We can now handle geometrical constraints since they are automatically taken into account by using generalised coordinates. But kinematical constraints involve the particle velocities which in turn depend not only on the coordinates $q_1, \ldots, q_n$, but also their time derivatives $\dot{q}_1, \ldots, \dot{q}_n$. *In general, kinematical constraints cannot be incorporated by selecting some new set of generalised coordinates*. As a result, such constraints have to remain as additional ODEs that must be solved along with the equations of motion.

All is not lost however since, in some special but important cases, the ODE representing the kinematical constraint can be immediately integrated to yield an *equivalent geometrical constraint*. Such a constraint is said to be **integrable**.

Example 12.5 ***An integrable kinematical constraint***

A circular cylinder rolls down a rough inclined plane. Show that, in this problem, the rolling condition is an integrable constraint.

Solution

In the absence of the rolling condition, this system has two degrees of freedom; take as generalised coordinates x (the displacement of the cylinder axis down the plane) and θ (the rotation angle of the cylinder). The **rolling condition** is then given by the first order ODE

$$\dot{x} = a\dot{\theta}, \qquad (12.3)$$

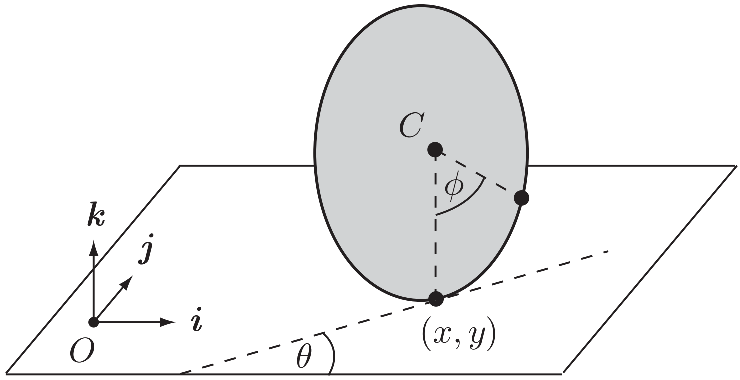


FIGURE 12.3 For the generally rolling wheel, the rolling conditions are non-integrable.

where a is the radius of the cylinder. But this constraint can be integrated (without solving the problem!) to give

$$x = a\theta, \tag{12.4}$$

on taking $x = \theta = 0$ in the reference configuration. Thus the kinematical constraint (12.3) is equivalent to the geometrical constraint (12.4). This geometrical constraint can now be incorporated by selecting a new (reduced) set of generalised coordinates. In this example, only one generalised coordinate is finally required (either x or θ) so that the rolling cylinder has **one degree of freedom**. ■

Example 12.6 *A non-integrable kinematical constraint*

Figure 12.3 shows a circular disk of radius a which is constrained to roll on a horizontal floor with its plane vertical. Show that, in this problem, the rolling conditions are not integrable.

Solution

In the absence of the rolling condition, this system has four degrees of freedom. Let $Oxyz$ be a fixed system of rectangular coordinates with O on the floor and Oz pointing vertically upwards. Then a set of generalised coordinates is given by

(i) the x and y coordinates of the centre C of the disk,

(ii) the angle θ between the plane of the disc and the x-axis,

(iii) the angle ϕ that the disk has rotated about its axis (relative to some reference position).

Now we impose the **rolling condition**, namely, that the contact particle should have zero velocity. In terms of the chosen coordinates, this gives

$$\dot{x} + a\dot{\phi}\cos\theta = 0, \qquad \dot{y} + a\dot{\phi}\sin\theta = 0,$$

a *pair* of first order ODEs. These equations cannot be integrated since θ is an unknown function of the time and $\dot{\theta}$ is absent from both equations. It follows that, in this problem, the rolling conditions are **not integrable** and cannot be replaced by equivalent geometrical constraints. ■

Holonomic and non-holonomic systems

Mechanical systems are classified according as to whether or not they have non-integrable kinematical constraints.

Definition 12.3 ***Holonomic systems*** *If a system has only geometrical or integrable kinematical constraints, then it is said to be* ***holonomic****. If it has non-integrable kinematical constraints, then it is* ***non-holonomic****.*

Non-holonomic systems are the bad guys. In particular, *non-holonomic systems do not satisfy Lagrange's equations* (as presented later in this chapter). It is beyond the scope of this book to proceed any further with the *analytical* mechanics of such systems and, from now on, we will deal only with holonomic systems. (A way of extending the Lagrange method to non-holonomic systems is described by Goldstein [4].) Such systems can still be treated by standard Newtonian methods however. The problem of the rolling wheel is solved in this way in Chapter 19.

12.3 CONFIGURATION SPACE (q–space)

Let $\mathcal{S}$ be a **holonomic mechanical system** with generalised coordinates $q_1 \ldots, q_n$. It is convenient to regard the list of values $q_1, \ldots, q_n$ as the coordinates of a 'point' $\boldsymbol{q}$ in a space of n dimensions, that is,

$$\boldsymbol{q} = (q_1, \ldots, q_n). \tag{12.5}$$

Mathematicians call such a space $\mathbb{E}_n$ (the Euclidean space of n dimensions), but we will denote it by $\mathcal{Q}$ (the space to which $\boldsymbol{q}$ belongs) and call it **configuration space**. Since the values of $q_1, \ldots, q_n$ determine the configuration of the system $\mathcal{S}$, it follows that *the configuration of $\mathcal{S}$ is determined by the 'position' of the point $\boldsymbol{q}$ in configuration space*, that is,

$$\boldsymbol{r}_i = \boldsymbol{r}_i(\boldsymbol{q}) \qquad\qquad (i = 1, \ldots, N).$$

This abstract view becomes much clearer when applied to a particular example. Let $\mathcal{S}$ be the two-particle system shown in Figure 12.4. This system has two degrees of freedom and generalised coordinates x, θ. In this case the configuration space $\mathcal{Q}$ is the (x, θ)-plane. Each point $\boldsymbol{q} = (x, \theta)$ lying in $\mathcal{Q}$ corresponds to a configuration of the mechanical system $\mathcal{S}$. Moreover, as the configuration of the system changes with time, the point $\boldsymbol{q}$ moves through the configuration space as shown.

Generalised velocities

When the configuration of $\mathcal{S}$ changes with time, the point $\boldsymbol{q}$ moves through the configuration space $\mathcal{Q}$ so that $\boldsymbol{q} = \boldsymbol{q}(t)$. This leads to the notion of **generalised velocities**.

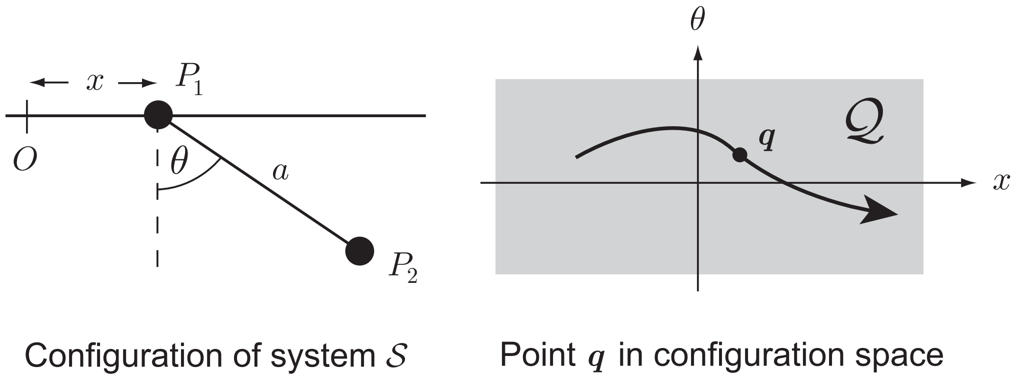


FIGURE 12.4 The configuration of the system $\mathcal{S}$ is represented by the point $\boldsymbol{q} = (x, \theta)$ in the configuration space $\mathcal{Q}$. As the configuration of $\mathcal{S}$ changes with time, the point $\boldsymbol{q}$ moves on a path lying in the configuration space $\mathcal{Q}$.

Definition 12.4 ***Generalised velocities*** *The time derivatives $\dot{q}_1, \ldots, \dot{q}_n$ of the generalised coordinates $q_1 \ldots, q_n$ are called the* ***generalised velocities*** *of the system $\mathcal{S}$.*

The n-dimensional vector $(\dot{q}_1, \ldots, \dot{q}_n)$, formed from the $\{\dot{q}_j\}$, is just the time derivative of the vector $\boldsymbol{q}$, that is,

$$\dot{\boldsymbol{q}} = (\dot{q}_1, \ldots, \dot{q}_n). \tag{12.6}$$

The vector $\dot{\boldsymbol{q}}$ can be regarded as the 'velocity' of the point $\boldsymbol{q}$ as it moves through the configuration space $\mathcal{Q}$.

Particle velocities

The values of $\boldsymbol{q}$ and $\dot{\boldsymbol{q}}$ determine the position and velocity of every particle of the system $\mathcal{S}$. For, since $\boldsymbol{r}_i = \boldsymbol{r}_i(\boldsymbol{q})$ and $\boldsymbol{q} = \boldsymbol{q}(t)$, it follows from the chain rule that

$$\boldsymbol{v}_i = \frac{\partial \boldsymbol{r}_i}{\partial q_1}\dot{q}_1 + \cdots + \frac{\partial \boldsymbol{r}_i}{\partial q_n}\dot{q}_n = \sum_{j=1}^{n} \frac{\partial \boldsymbol{r}_i}{\partial q_j}\dot{q}_j. \tag{12.7}$$

This expression for $\boldsymbol{v}_i$ is **linear** in the variables $\dot{q}_1, \ldots, \dot{q}_n$ with coefficients that depend on $\boldsymbol{q}$.

Example 12.7 *Rule for finding particle velocities*

What is the connection between the formula (12.7) and the rule we have often used to find particle velocities?

Solution

Formula (12.7) says that $\boldsymbol{v}_i$ is the vector sum of n contributions, each one arising from the variation of a particular q_j. This therefore *justifies* our rule for finding particle velocities. ■

Example 12.8 *Finding the kinetic energy*

Find the particle velocities for the two-particle system shown in Figure 12.2, and deduce the formula for the kinetic energy.

Solution

The velocities of the particles P_1, P_2 are given by

$$\boldsymbol{v}_1 = \dot{x}\,\boldsymbol{i}, \qquad \boldsymbol{v}_2 = \dot{x}\,\boldsymbol{i} + (a\cos\theta\,\boldsymbol{i} + a\sin\theta\,\boldsymbol{k})\,\dot{\theta}.$$

The kinetic energy of the system is therefore given by

$$\begin{aligned} T &= \tfrac{1}{2}m\,(\boldsymbol{v}_1\cdot\boldsymbol{v}_1) + \tfrac{1}{2}m\,(\boldsymbol{v}_2\cdot\boldsymbol{v}_2) \\ &= \tfrac{1}{2}m\dot{x}^2 + \tfrac{1}{2}m\left(\dot{x}^2 + (a\dot{\theta})^2 + 2\dot{x}(a\dot{\theta})\cos\theta\right) \\ &= m\,\dot{x}^2 + (\tfrac{1}{2}ma^2)\,\dot{\theta}^2 + (ma\cos\theta)\,\dot{x}\dot{\theta}. \blacksquare \end{aligned}$$

Example 12.9 *General form of the kinetic energy*

Show that the kinetic energy of *any holonomic mechanical system* has the form

$$T = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}(\boldsymbol{q})\,\dot{q}_j\dot{q}_k$$

that is, a **homogeneous quadratic form** in the variables $\dot{q}_1, \dots, \dot{q}_n$, with coefficients depending on $\boldsymbol{q}$.

Solution

Let P be a typical particle of $\mathcal{S}$ with position vector $\boldsymbol{r}$ and velocity $\boldsymbol{v}$. Then

$$\boldsymbol{v} = \frac{\partial\boldsymbol{r}}{\partial q_1}\dot{q}_1 + \cdots + \frac{\partial\boldsymbol{r}}{\partial q_n}\dot{q}_n = \sum_{j=1}^{n}\frac{\partial\boldsymbol{r}}{\partial q_j}\dot{q}_j$$

and so

$$\begin{aligned} \boldsymbol{v}\cdot\boldsymbol{v} &= \left(\frac{\partial\boldsymbol{r}}{\partial q_1}\dot{q}_1 + \cdots + \frac{\partial\boldsymbol{r}}{\partial q_n}\dot{q}_n\right)\cdot\left(\frac{\partial\boldsymbol{r}}{\partial q_1}\dot{q}_1 + \cdots + \frac{\partial\boldsymbol{r}}{\partial q_n}\dot{q}_n\right) \\ &= \left(\sum_{j=1}^{n}\frac{\partial\boldsymbol{r}}{\partial q_j}\dot{q}_j\right)\cdot\left(\sum_{k=1}^{n}\frac{\partial\boldsymbol{r}}{\partial q_k}\dot{q}_k\right) = \sum_{j=1}^{n}\sum_{k=1}^{n}\left(\frac{\partial\boldsymbol{r}}{\partial q_j}\cdot\frac{\partial\boldsymbol{r}}{\partial q_k}\right)\dot{q}_j\dot{q}_k, \end{aligned}$$

which is a homogeneous quadratic form in the variables $\dot{q}_1, \dots, \dot{q}_n$, with coefficients depending on $\boldsymbol{q}$. The kinetic energy of $\mathcal{S}$ is then given by

$$T = \tfrac{1}{2}\sum_{i=1}^{N} m_i\,(\boldsymbol{v}_i\cdot\boldsymbol{v}_i) = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}(\boldsymbol{q})\,\dot{q}_j\dot{q}_k,$$

where

$$a_{jk}(\boldsymbol{q}) = \tfrac{1}{2} \sum_{i=1}^{N} m_i \left(\frac{\partial \boldsymbol{r}_i}{\partial q_j} \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_k} \right).$$

It follows that T is also a **homogeneous quadratic form** in the variables $\dot{q}_1, \ldots, \dot{q}_n$, with coefficients depending on $\boldsymbol{q}$. ■

12.4 D'ALEMBERT'S PRINCIPLE

For a holonomic system, we can overcome the problem that the position vectors $\{\boldsymbol{r}_i\}$ are not independent variables by using generalised coordinates. We must now overcome the problem that the **constraint forces are unknown**.

The Newton equations of motion for the general mechanical system $\mathcal{S}$ are

$$m_i \dot{\boldsymbol{v}}_i = \boldsymbol{F}_i^S + \boldsymbol{F}_i^C \qquad (1 \leq i \leq N), \tag{12.8}$$

where $\boldsymbol{F}_i^S$ is the **specified force** and $\boldsymbol{F}_i^C$ is the **constraint force** acting on the particle P_i. The $\{\boldsymbol{F}_i^S\}$ are known while the $\{\boldsymbol{F}_i^C\}$ are unknown. The trick is to construct linear combinations of the equations (12.8) so as to eliminate the $\{\boldsymbol{F}_i^C\}$.

Let $\boldsymbol{a}_1(t)$, $\boldsymbol{a}_2(t)$, …, $\boldsymbol{a}_N(t)$ be *any* vector functions of the time. Then, by taking the scalar product of equation (12.8) with $\boldsymbol{a}_i$ and summing over i, we obtain the scalar equation

$$\sum_{i=1}^{N} m_i \dot{\boldsymbol{v}}_i \cdot \boldsymbol{a}_i = \sum_{i=1}^{N} \boldsymbol{F}_i^S \cdot \boldsymbol{a}_i + \sum_{i=1}^{N} \boldsymbol{F}_i^C \cdot \boldsymbol{a}_i. \tag{12.9}$$

The question now is whether we can make

$$\sum_{i=1}^{N} \boldsymbol{F}_i^C \cdot \boldsymbol{a}_i = 0 \tag{12.10}$$

by a cunning choice of the functions $\{\boldsymbol{a}_i\}$. More precisely, since $\mathcal{S}$ has n degrees of freedom, we need n linearly independent choices of the $\{\boldsymbol{a}_i\}$ that make the equation (12.10) true.

Actually, we already know one choice of the $\{\boldsymbol{a}_i\}$ that makes the equation (12.10) true. Suppose that the total rate of working of the constraint forces is zero, which is true for many constraints (see Chapter 6). This condition can be written

$$\sum_{i=1}^{N} \boldsymbol{F}_i^C \cdot \boldsymbol{v}_i = 0,$$

where $\boldsymbol{v}_i$ is the velocity of the particle P_i at time t. Thus the condition (12.10) certainly holds for such a system if the $\{\boldsymbol{a}_i\}$ are chosen to be the particle velocities $\{\boldsymbol{v}_i\}$. With

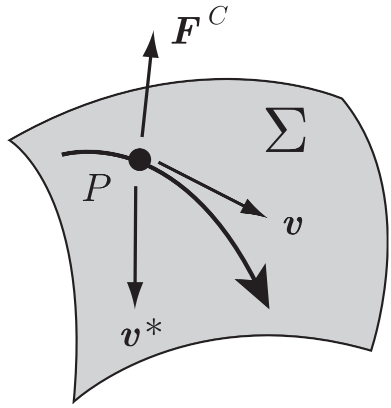


FIGURE 12.5 The particle P belonging to the system $\mathcal{S}$ is constrained to slide on the smooth fixed surface Σ.

this choice, the $\{\boldsymbol{F}_i^C\}$ are eliminated from equation (12.9). The result of this operation is actually well known to us; it leads to the energy principle for the system! This is *not* quite what we are looking for, but it does suggest what the correct choices of the $\{\boldsymbol{a}_i\}$ might be.

Now comes the clever bit. For all the usual constraints that do no work, it is also true that the stronger condition

$$\sum_{i=1}^{N} \boldsymbol{F}_i^C \cdot \boldsymbol{v}_i^* = 0 \qquad (12.11)$$

holds, where the $\{\boldsymbol{v}_i^*\}$ are *any kinematically possible set of particle velocities at time t*. The $\{\boldsymbol{v}_i^*\}$ need not be the *actual* particle velocities at time t. For example, suppose a particle P of $\mathcal{S}$ is constrained to move on a *smooth* fixed surface Σ. Let $\boldsymbol{v}$ be the actual velocity of P as shown in Figure 12.5. Since Σ is a smooth surface, the constraint force $\boldsymbol{F}^C$ that it exerts must be *normal* to Σ. Moreover, any kinematically possible motion of $\mathcal{S}$ at time t gives P a velocity $\boldsymbol{v}^*$ that is *tangential* to Σ. It follows that

$$\boldsymbol{F}^C \cdot \boldsymbol{v}^* = 0,$$

for any choice of $\boldsymbol{v}^*$ that is kinematically possible. Although there is no *theorem* to this effect, a similar conclusion can be drawn for all the usual constraint forces that do no work.

A set of velocities $\{\boldsymbol{v}_i^*\}$ that is kinematically possible at time t is called a **virtual motion** of the system. The condition (12.11) is therefore equivalent to the statement that the total rate of working of the constraint forces is zero in all virtual motions, or, more briefly, that the constraint forces do **no virtual work**. We have therefore obtained the following result, known as **d' Alembert's principle**.*

* After Jean le Rond d'Alembert (1717–1783). He was baptised Jean le Rond after being found abandoned on the steps of a Paris church of that name. His principle was published in 1743 in his *Traite de Dynamique*.

D'Alembert's principle

If the constraint forces on a system do **no virtual work**, then

$$\sum_{i=1}^{N} m_i \dot{\boldsymbol{v}}_i \cdot \boldsymbol{v}_i^* = \sum_{i=1}^{N} \boldsymbol{F}_i^S \cdot \boldsymbol{v}_i^*, \qquad (12.12)$$

where $\{\boldsymbol{v}_i^*\}$ is any virtual motion of the system at time t.

Differential form of d'Alembert's principle
D'Alembert's principle is often quoted in the equivalent differential form

$$\sum_i m_i \dot{\boldsymbol{v}}_i \cdot \boldsymbol{dr}_i = \sum_i \boldsymbol{F}_i^S \cdot \boldsymbol{dr}_i,$$

where the $\{\boldsymbol{dr}_i\}$ are any *kinematically possible set of infinitesimal displacements* of the particles $\{P_i\}$ at time t. This form is also known as the **principle of virtual work**.

D'Alembert's principle is not often applied directly, except in statical problems. We have obtained it because it leads to Lagrange's equations.

12.5 LAGRANGE'S EQUATIONS

From now on, we will suppose that our mechanical system is holonomic *and* that its constraint forces do no virtual work.

Definition 12.5 ***Standard system*** *If a mechanical system is holonomic and its constraint forces do no virtual work, we will call it a* ***standard system****.*

Consider then a standard mechanical system with n degrees of freedom and generalised coordinates $\boldsymbol{q} = (q_1, q_2, \ldots, q_n)$. Consider first the virtual motion $\{\boldsymbol{v}_i^*\}$ generated by *prescribing* the generalised velocities at time t to be

$$\dot{q}_1 = 1, \; \dot{q}_2 = \cdots = \dot{q}_n = 0.$$

From (12.7) it follows that the corresponding particle velocities are given by

$$\boldsymbol{v}_i^* = \frac{\partial \boldsymbol{r}_i}{\partial q_1} \qquad (1, \ldots, N).$$

Since we are assuming our system to be *holonomic*, these $\{\boldsymbol{v}_i^*\}$ *are* a kinematically possible set of velocities.* Furthermore, since we are assuming that the constraint forces do *no*

* This is the only point in the derivation of Lagrange's equations where it is *essential* that the system be holonomic.

virtual work, d'Alembert's principle holds. It therefore follows that

$$\sum_{i=1}^{N} m_i \dot{\boldsymbol{v}}_i \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_1} = \sum_{i=1}^{N} \boldsymbol{F}_i^S \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_1}.$$

A similar argument holds when the $\{\dot{q}_j\}$ are prescribed to be $\dot{q}_1 = 0,\ \dot{q}_2 = 1, \dot{q}_3 = \cdots = \dot{q}_n = 0$, and so on. We thus obtain the system of equations

$$\sum_{i=1}^{N} m_i \dot{\boldsymbol{v}}_i \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_j} = \sum_{i=1}^{N} \boldsymbol{F}_i^S \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_j} \qquad (j = 1, \ldots, n). \tag{12.13}$$

These are essentially Lagrange's equations. It remains only to put them into a form that is easy to use. In fact, the left sides of equations (12.13) can be constructed simply from the **kinetic energy** of the system. The result is as follows:

Suppose the holonomic system $\mathcal{S}$ has kinetic energy $T(\boldsymbol{q}, \dot{\boldsymbol{q}})$. Then the left sides of equations (12.13) can be written in the form

$$\sum_{i} m_i \dot{\boldsymbol{v}}_i \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_j} = \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} \tag{12.14}$$

$(1 \le j \le n)$*, where, for the purpose of calculating the partial derivatives, T is considered to be a function of the $2n$ independent variables $q_1, \ldots, q_n, \dot{q}_1, \ldots \dot{q}_n$.*

Lagrange partial derivatives

The partial derivatives of T that appear in equations (12.14) are peculiar to Lagrange's equations. In the expression for T, the coordinate velocities $\dot{q}_1, \ldots \dot{q}_n$ are considered to be *independent variables* in addition to the coordinates $q_1, \ldots, q_n$. Consider, for example, the two particle system in Example 12.2. For this system

$$T = m\,\dot{x}^2 + (\tfrac{1}{2}ma^2)\,\dot{\theta}^2 + (ma\cos\theta)\,\dot{x}\dot{\theta}$$

and this expression is considered to be a function of the *four* independent variables $x, \theta, \dot{x}, \dot{\theta}$ (x is absent). The Lagrange partial derivatives of T are therefore

$$\frac{\partial T}{\partial x} = 0, \quad \frac{\partial T}{\partial \dot{x}} = 2m\dot{x} + (ma\cos\theta)\dot{\theta}, \quad \frac{\partial T}{\partial \theta} = -(ma\sin\theta)\dot{x}\dot{\theta}, \quad \frac{\partial T}{\partial \dot{\theta}} = ma^2\dot{\theta} + (ma\cos\theta)\dot{x}.$$

The proof of the formula (12.14) is straightforward (once Lagrange had found the answer!) but a bit messy because of the many suffices.

Proof of the formula (12.14)

Since

$$\boldsymbol{v}_i = \frac{\partial \boldsymbol{r}_i}{\partial q_1}\dot{q}_1 + \cdots + \frac{\partial \boldsymbol{r}_i}{\partial q_n}\dot{q}_n,$$

it follows that

$$\frac{\partial}{\partial \dot{q}_j}\left(\tfrac{1}{2}\boldsymbol{v}_i \cdot \boldsymbol{v}_i\right) = \boldsymbol{v}_i \cdot \frac{\partial \boldsymbol{v}_i}{\partial \dot{q}_j} = \boldsymbol{v}_i \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_j}.$$

The last step follows since, in the formula for $\boldsymbol{v}_i$, $\boldsymbol{q}$ and $\dot{\boldsymbol{q}}$ are regarded as independent variables*. Then

$$\begin{aligned}\frac{d}{dt}\left[\frac{\partial}{\partial \dot{q}_j}\left(\tfrac{1}{2}\boldsymbol{v}_i \cdot \boldsymbol{v}_i\right)\right] &= \dot{\boldsymbol{v}}_i \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_j} + \boldsymbol{v}_i \cdot \frac{d}{dt}\left(\frac{\partial \boldsymbol{r}_i}{\partial q_j}\right) \\ &= \dot{\boldsymbol{v}}_i \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_j} + \boldsymbol{v}_i \cdot \sum_{k=1}^{n} \frac{\partial^2 \boldsymbol{r}_i}{\partial q_k \partial q_j}\dot{q}_k,\end{aligned}$$

after a further application of the chain rule. In a similar way,

$$\frac{\partial}{\partial q_j}\left(\tfrac{1}{2}\boldsymbol{v}_i \cdot \boldsymbol{v}_i\right) = \boldsymbol{v}_i \cdot \frac{\partial \boldsymbol{v}_i}{\partial q_j} = \boldsymbol{v}_i \cdot \frac{\partial}{\partial q_j}\left(\sum_{k=1}^{n} \frac{\partial \boldsymbol{r}_i}{\partial q_k}\dot{q}_k\right) = \boldsymbol{v}_i \cdot \sum_{k=1}^{n} \frac{\partial^2 \boldsymbol{r}_i}{\partial q_j \partial q_k}\dot{q}_k,$$

where, in the formula for $\partial \boldsymbol{v}_i / \partial q_j$, we have regarded $\boldsymbol{q}$ and $\dot{\boldsymbol{q}}$ as independent variables. Combining these two results gives

$$\frac{d}{dt}\left[\frac{\partial}{\partial \dot{q}_j}\left(\tfrac{1}{2}\boldsymbol{v}_i \cdot \boldsymbol{v}_i\right)\right] - \frac{\partial}{\partial q_j}\left(\tfrac{1}{2}\boldsymbol{v}_i \cdot \boldsymbol{v}_i\right) = \dot{\boldsymbol{v}}_i \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_j}.$$

If we now multiply by m_i and sum over i we obtain

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = \sum_i m_i \dot{\boldsymbol{v}}_i \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_j},$$

which is the required result. ■

For **general specified forces** $\{\boldsymbol{F}_i\}$ there is no simplification for the right sides of the equations (12.14), but we do give them names:

Definition 12.6 ***Generalised force*** *The quantity Q_j, defined by*

$$Q_j = \sum_i \boldsymbol{F}_i^S \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_j}$$

is called the ***generalised force*** *corresponding to the coordinate q_j.*

We have therefore proved that:

---

* The formula

$$\frac{\partial \dot{\boldsymbol{r}}_i}{\partial \dot{q}_j} = \frac{\partial \boldsymbol{r}_i}{\partial q_j}$$

is sometimes facetiously referred to as 'cancelling the dots'. Only mathematicians find this amusing.

Lagrange's equations for a general standard system

Let $\mathcal{S}$ be a **standard system** with generalised coordinates $\boldsymbol{q}$, kinetic energy $T(\boldsymbol{q}, \dot{\boldsymbol{q}})$ and generalised forces $\{Q_j\}$. Then, in any motion of $\mathcal{S}$, the coordinates $\boldsymbol{q}(t)$ must satisfy the system of equations

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = Q_j \qquad (1 \le j \le n). \tag{12.15}$$

This is the **form of Lagrange's equations** that applies to **any standard system**.

Conservative systems

When the standard system is also **conservative**, the generalised forces $\{Q_j\}$ can be written in terms of the **potential energy** $V(\boldsymbol{q})$ as

$$Q_j = -\frac{\partial V}{\partial q_j}. \tag{12.16}$$

[This result is simply a generalisation of the formula $\boldsymbol{F} = -\operatorname{grad} V$.]

Proof of the formula (12.16)
Let $\boldsymbol{q}^A$, $\boldsymbol{q}^B$ be any two points of configuration space that can be joined by a straight line parallel to the q_j-axis. Then

$$\begin{aligned}\int_{\boldsymbol{q}^A}^{\boldsymbol{q}^B} Q_j\, dq_j &= \int_{\boldsymbol{q}^A}^{\boldsymbol{q}^B} \left(\sum_i \boldsymbol{F}_i^S \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_j}\right) dq_j = \sum_i \int_{C_i} \boldsymbol{F}_i^S \cdot d\boldsymbol{r} \\ &= V(\boldsymbol{q}^A) - V(\boldsymbol{q}^B) = -\int_{\boldsymbol{q}^A}^{\boldsymbol{q}^B} \frac{\partial V}{\partial q_j}\, dq_j.\end{aligned}$$

This equality holds for all $\boldsymbol{q}^A$, $\boldsymbol{q}^B$ chosen as described, which implies that the two integrands must be equal. Hence

$$Q_j = -\frac{\partial V}{\partial q_j},$$

as required. ■

We have therefore proved that:

Lagrange's equations for a conservative standard system

Let $\mathcal{S}$ be a **conservative standard system** with generalised coordinates $\boldsymbol{q}$, kinetic energy $T(\boldsymbol{q}, \dot{\boldsymbol{q}})$ and potential energy $V(\boldsymbol{q})$. Then, in any motion of $\mathcal{S}$, the coordinates $\boldsymbol{q}(t)$ must satisfy the system of equations

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = -\frac{\partial V}{\partial q_j} \qquad (1 \le j \le n). \qquad (12.17)$$

These are **Lagrange's equations** for a **conservative standard system**. This is by far the most important case; most of analytical mechanics deals with conservative systems. It is remarkable that *all one needs to obtain the equations of motion for a conservative system are the expressions for the* ***kinetic*** *and* ***potential energies***.

Sufficiency of the Lagrange equations

We have shown that if $\mathcal{S}$ is a conservative standard system then Lagrange's equations (12.17) must hold. Thus Lagrange's equations are *necessary* conditions for $\boldsymbol{q}(t)$ to be a motion of $\mathcal{S}$. It does not seem possible to reverse this argument to show that the Lagrange equations are also *sufficient*. (Where does the reverse argument break down?) However, from the general theory of ODEs, we are assured that there is a unique solution of the Lagrange equations corresponding to each set of initial values for $\boldsymbol{q}, \dot{\boldsymbol{q}}$. Thus the Lagrange equations actually *are* sufficient to determine the motion of $\mathcal{S}$.

The Lagrange method for finding the equations of motion of a conservative system is summarised below:

Lagrange's method for conservative systems

- Confirm that the system is standard and that the specified forces are conservative.
- Select generalised coordinates.
- Evaluate the expressions for T and V in terms of the chosen coordinates*.
- Substitute these expressions into the Lagrange equations (12.17) and turn the handle. It's a piece of cake!

Example 12.10 *Using Lagrange's equations: I*

Consider a block of mass m sliding on a smooth wedge of mass M and angle α which itself slides on a smooth horizontal floor, as shown in Figure 12.6. The whole motion

* See Chapter 9 for the details of how to find T.

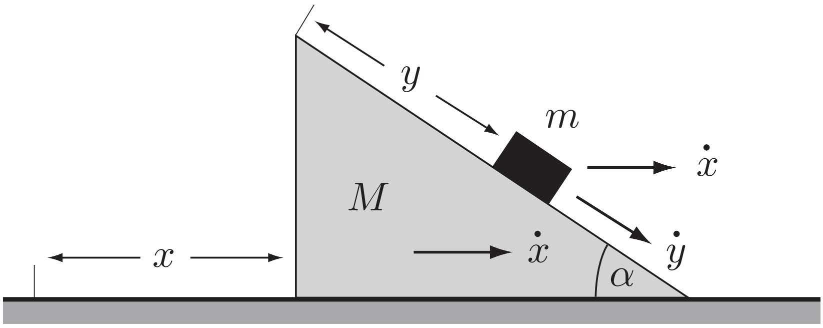


FIGURE 12.6 The block slides on the smooth surface of the wedge which slides on a smooth horizontal floor.

is planar. Find Lagrange's equations for this system and deduce (i) the acceleration of the wedge, and (ii) the acceleration of the block relative to the wedge.

Solution

This is a standard conservative system with two degrees of freedom. Take as generalised coordinates x, the displacement of the wedge from a fixed point on the floor, and y, the displacement of the block from a fixed point on the wedge. The calculation of the **kinetic** and **potential energies** in terms of x, y is performed exactly as in Chapter 9 and gives

$$T = \tfrac{1}{2}M\dot{x}^2 + \tfrac{1}{2}m\left(\dot{x}^2 + \dot{y}^2 + 2\dot{x}\dot{y}\cos\alpha\right),$$
$$V = -mgy\sin\alpha.$$

The required partial derivatives of T and V are then given by

$$\frac{\partial T}{\partial x} = 0, \qquad \frac{\partial T}{\partial \dot{x}} = (M+m)\dot{x} + (m\cos\alpha)\dot{y}, \qquad \frac{\partial V}{\partial x} = 0.$$

$$\frac{\partial T}{\partial y} = 0, \qquad \frac{\partial T}{\partial \dot{y}} = (m\cos\alpha)\dot{x} + m\dot{y}, \qquad \frac{\partial V}{\partial y} = -mg\sin\alpha.$$

We can now form up the **Lagrange equations**. The equation corresponding to the coordinate x is

$$\frac{d}{dt}\left[(M+m)\dot{x} + (m\cos\alpha)\dot{y}\right] - 0 \;=\; 0, \qquad (12.18)$$

and the equation corresponding to the coordinate y is

$$\frac{d}{dt}\left[(m\cos\alpha)\dot{x} + m\dot{y}\right] - 0 \;=\; mg\sin\alpha. \qquad (12.19)$$

If we now perform the time derivatives in equations (12.18), (12.19) and solve for the unknowns $\ddot{x}$, $\ddot{y}$ we obtain

$$\ddot{x} = -\frac{mg\sin\alpha\cos\alpha}{M + m\sin^2\alpha}, \qquad \ddot{y} = \frac{(M+m)g\sin\alpha}{M + m\sin^2\alpha},$$

which are the required **accelerations**. They are both constant. ■

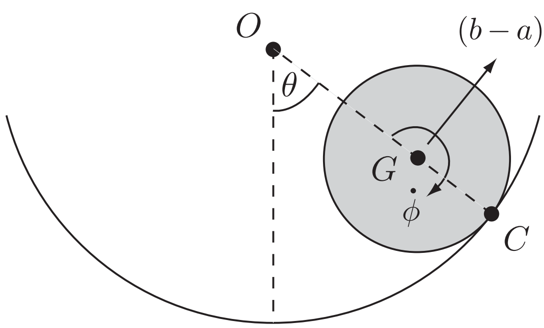


FIGURE 12.7 The small solid cylinder rolls on the inside surface of the large fixed cylinder.

These results can of course be obtained by more elementary means. For instance we could solve this problem by appealing to conservation of horizontal linear momentum and energy. However the Lagrange method does have the advantage that less physical insight is needed to solve the problem. If the system is a standard one and T and V can be calculated, then turning the handle produces the equations of motion.

Example 12.11 *Using Lagrange's equations: II*

Figure 12.7 shows a solid cylinder with centre G and radius a rolling on the rough inside surface of a *fixed* cylinder with centre O and radius $b > a$. Find the Lagrange equation of motion and deduce the period of small oscillations about the equilibrium position.

Solution

If the cylinder were not obliged to roll, the system would have two degrees of freedom with generalised coordinates θ (the angle between OG and the downward vertical) and ϕ (the rotation angle of the cylinder measured from some reference position). The **rolling condition** imposes the kinematical constraint

$$(b-a)\dot{\theta} - a\dot{\phi} = 0.$$

This constraint is **integrable** and is equivalent to the geometrical constraint

$$(b-a)\theta - a\phi = 0$$

on taking $\phi = 0$ when $\theta = 0$. Thus the rolling cylinder is a standard conservative system with *one* degree of freedom.

Take θ as the generalised coordinate. Then the **kinetic energy** is given by

$$\begin{aligned} T &= \tfrac{1}{2}m\left((b-a)\dot{\theta}\right)^2 + \tfrac{1}{2}\left(\tfrac{1}{2}ma^2\right)\dot{\phi}^2 \\ &= \tfrac{1}{2}m\left((b-a)\dot{\theta}\right)^2 + \tfrac{1}{2}\left(\tfrac{1}{2}ma^2\right)\left(\frac{b-a}{a}\right)^2\dot{\theta}^2 \\ &= \tfrac{3}{4}m(b-a)^2\dot{\theta}^2 \end{aligned}$$

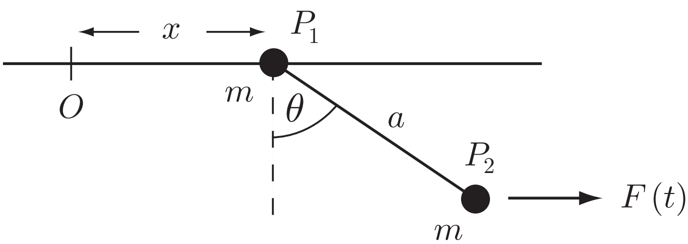


FIGURE 12.8 The system moves under the prescribed force $F(t)$.

and the **potential energy** by

$$V = -mg(b-a)\cos\theta.$$

There is only one **Lagrange equation**, namely

$$\frac{d}{dt}\left[\tfrac{3}{2}m(b-a)^2\dot{\theta}\right] - 0 = -mg(b-a)\sin\theta$$

which simplifies to give

$$\ddot{\theta} + \frac{2g}{3(b-a)}\sin\theta = 0.$$

Interestingly, this equation is identical to the exact equation for the oscillations of a simple pendulum of length $3(b-a)/2$ as obtained in Chapter 6.

The linearised equation governing small oscillations of the cylinder about $\theta = 0$ is

$$\ddot{\theta} + \frac{2g}{3(b-a)}\theta = 0$$

so that the **period** τ of small oscillations is given by

$$\tau = 2\pi\left(\frac{3(b-a)}{2g}\right)^{1/2}. \blacksquare$$

Example 12.12 *Using Lagrange's equations: III*

Let $\mathcal{S}$ be the system shown in Figure 12.8. The rail is smooth and the prescribed force $F(t)$ acts on the particle P_2 as shown. Gravity is absent. Find the Lagrange equations for $\mathcal{S}$.

Solution

$\mathcal{S}$ is a standard system with two degrees of freedom. The new feature is the prescribed external force $F(t)$ acting on P_2. This time dependent force cannot be represented by a potential energy and so the generalised forces $\{Q_j\}$ must be evaluated direct from the definition (12.16).

Take generalised coordinates x, θ as shown and let the corresponding generalised forces be called Q_x, Q_θ. Then, since $\mathcal{S}$ has just two particles,

$$Q_x = \boldsymbol{F}_1^S \cdot \frac{\partial \boldsymbol{r}_1}{\partial x} + \boldsymbol{F}_2^S \cdot \frac{\partial \boldsymbol{r}_2}{\partial x},$$
$$Q_\theta = \boldsymbol{F}_1^S \cdot \frac{\partial \boldsymbol{r}_1}{\partial \theta} + \boldsymbol{F}_2^S \cdot \frac{\partial \boldsymbol{r}_2}{\partial \theta},$$

where

$$\boldsymbol{F}_1^S = \boldsymbol{0}, \qquad \boldsymbol{F}_2^S = F(t)\,\boldsymbol{i},$$

and

$$\boldsymbol{r}_1 = x\,\boldsymbol{i}, \qquad \boldsymbol{r}_2 = (x + a\sin\theta)\,\boldsymbol{i} - (a\cos\theta)\,\boldsymbol{k}.$$

The **generalised forces** Q_x, Q_θ are therefore given by

$$Q_x = 0 + (F(t)\,\boldsymbol{i}) \cdot \boldsymbol{i} = F(t)$$

and

$$Q_\theta = 0 + (F(t)\,\boldsymbol{i}) \cdot (a\cos\theta\,\boldsymbol{i} + a\sin\theta\,\boldsymbol{k}) = (a\cos\theta)\,F(t).$$

The **kinetic energy** is given by

$$T = m\dot{x}^2 + (ma\cos\theta)\dot{x}\dot{\theta} + \tfrac{1}{2}m\dot{\theta}^2$$

and so the **Lagrange equations** are

$$\frac{d}{dt}\left[2m\dot{x} + (ma\cos\theta)\dot{\theta}\right] = F(t),$$
$$\frac{d}{dt}\left[(ma\cos\theta)\dot{x} + m\dot{\theta}\right] - \left[-(ma\sin\theta)\dot{x}\dot{\theta}\right] = (a\cos\theta)F(t). \blacksquare$$

Question ***Incorporating extra forces***

How would you incorporate gravity into the last example?

Answer

Since the expression (12.16) for the $\{Q_j\}$ is linear in the $\{\boldsymbol{F}_i\}$, the extra forces are incorporated by just *adding* in their respective contributions to the $\{Q_j\}$. Thus when gravity is present in the last example, Q_x, Q_θ become

$$Q_x = F(t) + 0, \qquad Q_\theta = (a\cos\theta)F(t) - mga\sin\theta. \blacksquare$$

Many more examples of the use of Lagrange's equations are given in the problems at the end of the chapter.

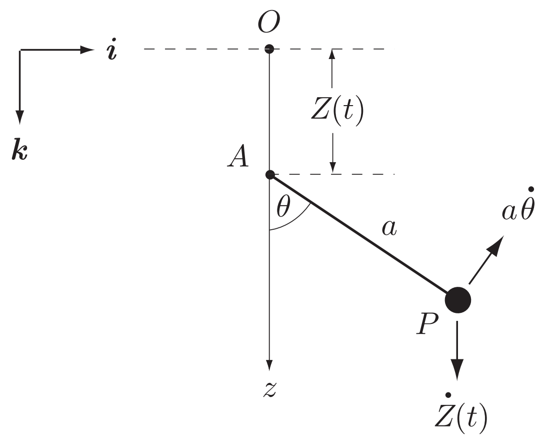


FIGURE 12.9 The pendulum with a moving support.

12.6 SYSTEMS WITH MOVING CONSTRAINTS

The theory of Lagrange's equations can be extended to include a fascinating class of problems in which the constraints are time dependent. Consider the system shown in Figure 12.9 which is a simple pendulum in which the support point A is made to move vertically so that its downward displacement from the *fixed* origin O at time t is some *specified* function $Z(t)$. For example it could be made to oscillate so that $Z(t) = Z_0 \cos pt$. With this constraint, the coordinate θ is no longer sufficient to specify the position of the particle P. In fact, relative to the origin O, the position vector of P at time t is given by

$$\boldsymbol{r} = (a \sin\theta)\,\boldsymbol{i} + (Z(t) + a\cos\theta)\,\boldsymbol{k},$$

so that $\boldsymbol{r}$ is a function of θ and t, not just θ. Constraints which cause the $\{\boldsymbol{r}_i\}$ to depend on $\boldsymbol{q}$ and t (and not just $\boldsymbol{q}$) are called **time dependent constraints**, or simply **moving constraints**. Systems that have moving constraints include:

Systems with moving constraints

- Systems in which particles are forced to move *in a prescribed manner*.
- Systems in which particles are forced to remain on boundaries that move *in a prescribed manner*.
- Systems in which the motion is viewed from a frame of reference that is accelerating or rotating *in a prescribed manner*.
- Systems in which beetles, mice (or lions!) move around *in a prescribed manner*. (These creatures are highly trained!)

We assume our systems are such that they *would* be standard *if* the constraints were fixed. We will refer to such systems as **standard systems with moving constraints**.

Kinematics of systems with moving constraints

The configuration $\{\boldsymbol{r}_i\}$ of a system with moving constraints is specified by

$$\boldsymbol{r}_i = \boldsymbol{r}_i(\boldsymbol{q}, t) \qquad (1 \le i \le N). \tag{12.20}$$

Here, the time t has the rôle of an 'additional coordinate'. However, it is not a true coordinate and we will still regard the system as being holonomic with n degrees of freedom. The corresponding particle velocities are given by

$$\boldsymbol{v}_i = \frac{\partial \boldsymbol{r}_i}{\partial q_1}\dot{q}_1 + \cdots + \frac{\partial \boldsymbol{r}_i}{\partial q_n}\dot{q}_n + \frac{\partial \boldsymbol{r}_i}{\partial t}. \tag{12.21}$$

This expression is still a linear form in the variables $\{\dot{q}_j\}$ but it is not homogeneous; there is now a 'constant' term (a function of $\boldsymbol{q}$ and t).

Question ***Form of the kinetic energy***

What is the form of the kinetic energy when moving constraints are present?

Answer

It follows from the above expression for the particle velocities that T has the form

$$T(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}(\boldsymbol{q}, t)\dot{q}_j\dot{q}_k + \sum_{j=1}^{n} b_j(\boldsymbol{q}, t)\dot{q}_j + c(\boldsymbol{q}, t),$$

which is still a quadratic form in the variables $\{\dot{q}_j\}$, but it is not homogeneous; there are now linear terms and a constant term. ■

Energy not conserved with moving constraints

Another feature of systems with moving constraints is that the **constraint forces do work**. This is quite obvious from the driven pendulum example above. The constraint force that causes the specified displacement of the support point A will generally have a vertical component and, since A is *moving* vertically, this force will do work. So, even when the *specified* forces are conservative (as gravity is in the driven pendulum example), the total energy $T + V$ is not a constant because the constraint force does work. Hence, systems with moving constraints are generally **not conservative**.

Lagrange's equations with moving constraints

There are good reasons to expect that systems with moving constraints do not satisfy Lagrange's equations. In general, constraint forces that enforce moving constraints do work. Since virtual motions include the special case of real motion, surely such constraints must also do *virtual* work; then d'Alemberts principle and Lagrange's equations will not hold. Compelling though this argument seems, it is false. Systems with moving constraints *do* satisfy Lagrange's equations! To see why this is so, one must identify the crucial steps in the derivation of Lagrange's equations. There are actually only three:

- Are the equations $\sum_{i=1}^{N} \boldsymbol{F}_i^C \cdot (\partial \boldsymbol{r}_i/\partial q_j) = 0$ still true?
  This question could be posed in the form '*do the constraint forces do virtual work?*' and we have presented a plausible argument that they do. Consider however the meaning of the partial derivatives $\partial \boldsymbol{r}_i/\partial q_j$. Since we now have $\boldsymbol{r}_i = \boldsymbol{r}_i(\boldsymbol{q}, t)$, $\partial \boldsymbol{r}_i/\partial q_1$ means the derivative of $\boldsymbol{r}_i$ with respect to q_1 keeping q_2, q_3,..., q_n *and the time* t constant. Thus these derivatives are calculated at constant t. It follows that the virtual motion defined by the $\{\partial \boldsymbol{r}_i/\partial q_1\}$ is kinematically consistent with constraints that are *fixed* at time t, not with the actual moving constraints. Hence, since $\sum_{i=1}^{N} \boldsymbol{F}_i^C \cdot (\partial \boldsymbol{r}_i/\partial q_j)$ would be zero if the constraints were fixed, it is still zero when the constraints are moving!*

- Is the formula (12.14) still true?
  The point here is that the formula (12.21) for the particle velocities now has the extra term $\partial \boldsymbol{r}_i/\partial t$ which might upset the formula (12.14). However, it does not. The proof of this is left as an exercise.

- When the specified forces are conservative, is the formula $\sum_{i=1}^{N} \boldsymbol{F}_i^s \cdot (\partial \boldsymbol{r}_i/\partial q_j) = \partial V/\partial q_j$ still true?
  The potential energy V is a function of the configuration of the system, but, since the configuration is now specified by $\boldsymbol{q}$ and t, $V = V(\boldsymbol{q}, t)$. However, since the partial derivatives $\partial \boldsymbol{r}_i/\partial q_j$ and $\partial V/\partial q_j$ are evaluated at constant t, the proof of the formula is unchanged.

We have therefore obtained the following result:

Lagrange's equations with moving constraints

Lagrange's equations still hold when moving constraints are present provided that, in the expressions for T and V, the time t is regarded as an *independent* variable.

Example 12.13 *Pendulum with an oscillating support*

Find the Lagrange equation for the driven pendulum for the case in which the displacement function $Z(t) = Z_0 \cos pt$. [Assume that the 'string' is a light rigid rod that cannot go slack.]

Solution

This system has one degree of freedom and a moving constraint at A. Take θ as the generalised coordinate. It follows from Figure 12.9 that the **kinetic energy** T is given by

$$T = \tfrac{1}{2}m\left(a^2\dot{\theta}^2 + \dot{Z}^2 - 2a\dot{\theta}\dot{Z}\sin\theta\right)$$

* The situation is sometimes loosely expressed by the mysterious statement that '*moving constraints do real work but no virtual work*'.

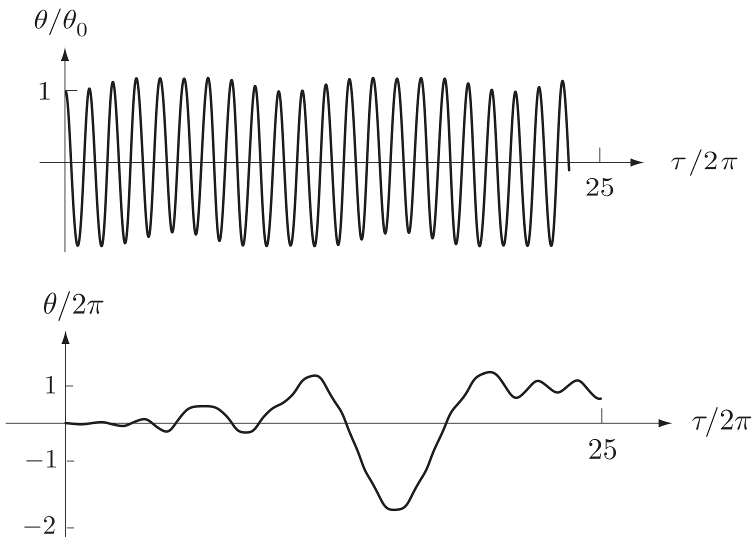


FIGURE 12.10 Motions of the driven pendulum. **Top:** $p/\Omega = 1.1$, $Z_0/a = 0.2$ and $\theta_0 = 0.1$. **Botton:** $p/\Omega = 1.9$, $Z_0/a = 0.2$ and $\theta_0 = 0.1$.

and the **potential energy** V by

$$V = -mg(Z + a\cos\theta).$$

The required partial derivatives are therefore

$$\frac{\partial T}{\partial \dot{\theta}} = m\left(a^2\dot{\theta} - a\dot{Z}\sin\theta\right), \qquad \frac{\partial T}{\partial \theta} = -ma\,\dot{\theta}\dot{Z}\cos\theta, \qquad \frac{\partial V}{\partial \theta} = mga\sin\theta.$$

The Lagrange equation corresponding to the coordinate θ is therefore

$$\frac{d}{dt}ma\left(a\dot{\theta} + \dot{Z}\sin\theta\right) - ma\dot{\theta}\dot{Z}\cos\theta = -mga\sin\theta,$$

which simplifies to give

$$\ddot{\theta} + (\Omega^2 - a^{-1}\ddot{Z})\sin\theta = 0,$$

where $\Omega^2 = g/a$. Hence, for the case in which $Z = Z_0\cos pt$, the **Lagrange equation** is

$$\ddot{\theta} + \left(\Omega^2 + \frac{Z_0 p^2}{a}\cos pt\right)\sin\theta = 0. \blacksquare \qquad (12.22)$$

Question *Motions of the driven pendulum*

What do the pendulum motions look like?

Answer

The equation (12.22) has some fascinating solutions, but they can only be found numerically. First we will reduce the number of parameters by putting the equation in dimensionless form. If we define the dimensionless time τ by $\tau = pt$, then the equation becomes

$$\frac{d^2\theta}{d\tau^2} + \left(\frac{\Omega^2}{p^2} + \left(\frac{Z_0}{a}\right)\cos\tau\right)\sin\theta = 0.$$

We can now see that the solutions depend on the dimensionless driving frequency p/Ω and the dimensionless driving amplitude Z_0/a.

One interesting question is whether the small oscillations of the pendulum about $\theta = 0$ are destabilised by the motion of the support. The answer is that it depends on the dimensionless parameters p/Ω and Z_0/a in a complicated way. Figure 12.10 shows results obtained by numerical solution of the equation with initial conditions of the form $\theta = \theta_0$, $\dot{\theta} = 0$ when $t = 0$. The **top graph** shows the motion for the case $p/\Omega = 1.1$, $Z_0/a = 0.2$ and $\theta_0 = 0.1$. In this graph, θ/θ_0 is plotted against $\tau/2\pi$ (the number of oscillations of the support). The motion turns out to be **stable** with the amplitude of the oscillations remaining close to their initial value. The **bottom graph** shows the motion for the case $p/\Omega = 1.9$, $Z_0/a = 0.2$ and $\theta_0 = 0.1$. In this graph, $\theta/2\pi$ (the number of *revolutions* of the pendulum) is plotted against $\tau/2\pi$ (the number of oscillations of the support). This motion turns out to be **unstable**. The amplitude of the oscillations grows until the pendulum performs complete circles; it then stops and goes the opposite way. Numerical results suggest that this chaotic motion continues indefinitely. ■

12.7 THE LAGRANGIAN

The Lagrange equations of motion (12.17) for a conservative standard system are expressed in terms of the kinetic energy $T = T(\boldsymbol{q}, \dot{\boldsymbol{q}})$ and the potential energy $V = V(\boldsymbol{q})$. They can however be written in terms of the single function $T - V$. Since $\partial V/\partial \dot{q}_j = 0$, the equations can be written

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = \frac{d}{dt}\left(\frac{\partial V}{\partial \dot{q}_j}\right) - \frac{\partial V}{\partial q_j} \qquad (1 \le j \le n),$$

that is,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0 \qquad (1 \le j \le n),$$

where $L(\boldsymbol{q}, \dot{\boldsymbol{q}}) = T(\boldsymbol{q}, \dot{\boldsymbol{q}}) - V(\boldsymbol{q})$ is called the **Lagrangian** of the system. The same operation can be applied to systems with moving constraints whose specified forces are conservative. The only difference is that, in this case, $L = L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$.

Writing the Lagrange equations in this form makes no difference whatever to problem solving. However, any system of equations that can be written in this way has special properties. In particular, it is equivalent to a **stationary principle** (see Chapter 13), and can also be written in **Hamiltonian form** (see Chapter 14). This is the form most suitable for advanced developments and for making the transition to quantum mechanics. There is therefore a strong interest in *any* physical system whose equations can be written in Lagrangian form.

Definition 12.7 ***Lagrangian form*** *If the equations of motion of a holonomic system with generalised coordinates* $\boldsymbol{q}$ *can be written in the form*

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0 \qquad (1 \le j \le n), \tag{12.23}$$

for some function $L = L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$*, then* L *is called the* ***Lagrangian*** *of the system and the equations are said to have* ***Lagrangian form****.*

For example, the Lagrangian for the driven pendulum is

$$L(\theta, \dot{\theta}, t) = \tfrac{1}{2}m\left(a^2\dot{\theta}^2 + \dot{Z}^2 - 2a\dot{\theta}\dot{Z}\sin\theta\right) + mg(Z + a\cos\theta),$$

where $Z = Z(t)$ is the displacement of the support point.

Velocity dependent potential

There are systems whose specified forces are **not conservative** (so that V does not exist), but their equations of motion can still be written in Lagrangian form. Any standard system with generalised forces $\{Q_j\}$ satisfies the Lagrange equations (12.15). If it happens that the generalised forces can be written in the form

$$Q_j = \frac{d}{dt}\left(\frac{\partial U}{\partial \dot{q}_j}\right) - \frac{\partial U}{\partial q_j} \qquad (1 \le j \le n), \tag{12.24}$$

for some function $U(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$, then clearly the equations (12.15) can be written in Lagrangian form by taking

$$L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) = T(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) - U(\boldsymbol{q}, \dot{\boldsymbol{q}}.t).$$

The function $U(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$ is called the **velocity dependent potential** of the system.

This seems to be a mathematical artifice that has no importance in practice. It is true that there is only *one* important case in which a velocity dependent potential exists, but that case is very important; it is the case of a charged particle moving in electromagnetic fields. The following example proves this to be so for *static* fields. The corresponding result for electro*dynamic* fields is the subject of Problem 12.15.

Example 12.14 *Charged particle in static EM fields*

A particle P of mass m and charge e can move freely in the static electric field $\boldsymbol{E} = \boldsymbol{E}(\boldsymbol{r})$ and the static magnetic field $\boldsymbol{B} = \boldsymbol{B}(\boldsymbol{r})$. The electric and magnetic fields exert a force on P given by the **Lorentz force** formula

$$\boldsymbol{F} = e\,\boldsymbol{E} + e\,\boldsymbol{v}\times\boldsymbol{B},$$

where $\boldsymbol{v}$ is the velocity of P. Show that this force can be represented by a velocity dependent potential $U(\boldsymbol{r}, \dot{\boldsymbol{r}})$ and find the Lagrangian of the system.

Solution

In the static case, Maxwell's equations for the electromagnetic field reduce to

$$\text{div}\,\boldsymbol{D} = \rho, \qquad \text{curl}\,\boldsymbol{E} = \boldsymbol{0}, \qquad \text{curl}\,\boldsymbol{H} = \boldsymbol{j}, \qquad \text{div}\,\boldsymbol{B} = 0.$$

In particular, the equation curl $\boldsymbol{E} = \boldsymbol{0}$ implies that $\boldsymbol{E}(\boldsymbol{r})$ is a conservative field and can be written in the form

$$\boldsymbol{E} = -\,\text{grad}\,\phi$$

where $\phi = \phi(\boldsymbol{r})$ is the **electrostatic potential**. The equation div $\boldsymbol{B} = 0$ implies that $\boldsymbol{B}(\boldsymbol{r})$ can be written in the form

$$\boldsymbol{B} = \text{curl}\,\boldsymbol{A},$$

where $\boldsymbol{A} = \boldsymbol{A}(\boldsymbol{r})$ is the **magnetic vector potential**.*

Take the generalised coordinates to be the Cartesian coordinates x, y z of the particle and, from now on, let $\boldsymbol{r}$ mean (x, y, z). What we are looking for is a velocity dependent potential $U(\boldsymbol{r}, \dot{\boldsymbol{r}})$ that yields the correct generalised forces when substituted into the equations (12.24). In the present case the generalised forces Q_x, Q_y, Q_z are simply the x- y- and z-components of the Lorentz force $\boldsymbol{F}$. The **electric** part of the force, $e\boldsymbol{E}$ is easily dealt with since it is conservative and can be represented by the ordinary potential energy $V = e\,\phi(\boldsymbol{r})$. It is the **magnetic** part of the force that needs the velocity dependent potential. One wonders how anyone found the correct U, but they did, and it turns out to be $-e\,\dot{\boldsymbol{r}}\cdot\boldsymbol{A}(\boldsymbol{r})$. All *we* need to do is to check that this potential is correct.

Consider therefore the potential

$$U = \dot{\boldsymbol{r}}\cdot\boldsymbol{A}(\boldsymbol{r}) = \dot{x}A_x(\boldsymbol{r}) + \dot{y}A_y(\boldsymbol{r}) + \dot{z}A_z(\boldsymbol{r}).$$

* The potential ϕ is unique to within an added constant, but, for any fixed $\boldsymbol{B}$, there are many possibilities for $\boldsymbol{A}$. Adding the grad of any scalar function to $\boldsymbol{A}$ does not change the value of $\boldsymbol{B}$. This ambiguity in $\boldsymbol{A}$ makes no difference in the present context; *any* choice of $\boldsymbol{A}$ such that $\boldsymbol{B} = \text{curl}\,\boldsymbol{A}$ will do. The actual determination of vector potentials is described in textbooks on vector field theory (see Schey [11]) for example.

Then

$$\frac{\partial U}{\partial \dot{x}} = A_x, \qquad \frac{\partial U}{\partial x} = \dot{x}\frac{\partial A_x}{\partial x} + \dot{y}\frac{\partial A_y}{\partial x} + \dot{z}\frac{\partial A_z}{\partial x}$$

and so

$$\begin{aligned}
\frac{d}{dt}\left(\frac{\partial U}{\partial \dot{x}}\right) - \frac{\partial U}{\partial x} &= \frac{d}{dt}(A_x) - \dot{x}\frac{\partial A_x}{\partial x} - \dot{y}\frac{\partial A_y}{\partial x} - \dot{z}\frac{\partial A_z}{\partial x} \\
&= \left(\frac{\partial A_x}{\partial x}\dot{x} + \frac{\partial A_x}{\partial y}\dot{y} + \frac{\partial A_x}{\partial z}\dot{z}\right) - \dot{x}\frac{\partial A_x}{\partial x} - \dot{y}\frac{\partial A_y}{\partial x} - \dot{z}\frac{\partial A_z}{\partial x} \\
&= -\dot{y}\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) + \dot{z}\left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right) \\
&= -\dot{y}\,(\text{curl}\,\boldsymbol{A})_z + \dot{z}\,(\text{curl}\,\boldsymbol{A})_y = -\dot{y}B_z + \dot{z}B_y \\
&= -\left(\dot{\boldsymbol{r}}\times\boldsymbol{B}\right)_x .
\end{aligned}$$

When multiplied by $-e$ this is Q_x for the magnetic part of the force. The values of Q_y and Q_z are confirmed in the same way.

We have therefore proved that the **Lorentz force** is derivable from the **velocity dependent potential**

$$U = e\,\phi(\boldsymbol{r}) - e\,\dot{\boldsymbol{r}}\cdot\boldsymbol{A}(\boldsymbol{r}) \tag{12.25}$$

and the **Lagrangian** of the particle is therefore

$$\boxed{L = \tfrac{1}{2}m\,\dot{\boldsymbol{r}}\cdot\dot{\boldsymbol{r}} - e\,\phi(\boldsymbol{r}) + e\,\dot{\boldsymbol{r}}\cdot\boldsymbol{A}(\boldsymbol{r})} \tag{12.26}$$

Question ***Why bother?***

Why should we find the Lagrangian for this system when we already know that the equation of motion is

$$m\frac{d\boldsymbol{v}}{dt} = e\,\boldsymbol{E} + e\,\boldsymbol{v}\times\boldsymbol{B}\ ?$$

Answer

The interest in this Lagrangian is that, from it, one can find the **Hamiltonian**, and this is what is needed to formulate the corresponding problem in **quantum mechanics**. This problem has important applications to the spectra of atoms in magnetic fields. ■

12.8 THE ENERGY FUNCTION h

Let $\mathcal{S}$ be any holonomic mechanical system with Lagrangian $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$. Then the equations of motion for $\mathcal{S}$ take the form (12.23). On multiplying the j-th equation by $\dot{q}_j$

and summing over j we obtain

$$0 = \sum_{j=1}^{n} \left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} \right] \dot{q}_j$$
$$= \sum_{j=1}^{n} \left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\dot{q}_j\right) - \frac{\partial L}{\partial q_j}\dot{q}_j - \frac{\partial L}{\partial \dot{q}_j}\ddot{q}_j \right]$$
$$= \frac{d}{dt}\left[\sum_{j=1}^{n} \left(\frac{\partial L}{\partial \dot{q}_j}\dot{q}_j\right) - L \right] + \frac{\partial L}{\partial t}.$$

Note that $\partial L/\partial t$ means the partial derivative of $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$ with respect to its final argument, holding $\boldsymbol{q}$ and $\dot{\boldsymbol{q}}$ constant. Thus

$$\boxed{\frac{dh}{dt} + \frac{\partial L}{\partial t} = 0} \tag{12.27}$$

where

$$\boxed{h = \sum_{j=1}^{n} \left(\frac{\partial L}{\partial \dot{q}_j}\dot{q}_j\right) - L} \tag{12.28}$$

Definition 12.8 ***Energy function*** *The function h defined by equation (12.28) is called the* ***energy function*** *of the system $\mathcal{S}$.*

The energy function h is a *generalisation of the notion of energy*. For conservative systems, we will show that it is identical with the total energy $E = T + V$. However, for non-conservative systems, V may not exist and, even if it does, h and E are not generally equal. There are three typical cases:

Case A If $L = L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$, then $\partial L/\partial t \neq 0$ and h is **not conserved**.

Example 12.15 *h for the driven pendulum*

Find the energy function h for the driven pendulum problem.

Solution

In the driven pendulum problem,

$$L = \tfrac{1}{2}m\left(a^2\dot{\theta}^2 + \dot{Z}^2 - 2a\dot{\theta}\dot{Z}\sin\theta\right) + mg(Z + a\cos\theta),$$

and so

$$h = \dot{\theta}\,\frac{\partial T}{\partial \dot{\theta}} - L = \tfrac{1}{2}m\left(a^2\dot{\theta}^2 - \dot{Z}^2\right) - mg(Z + a\cos\theta).$$

This is not the same as the total energy

$$T + V = \tfrac{1}{2}m\left(a^2\dot{\theta}^2 + \dot{Z}^2 - 2a\dot{\theta}\dot{Z}\sin\theta\right) - mg(Z + a\cos\theta),$$

and neither quantity is conserved. ■

Case B If $L = L(\boldsymbol{q}, \dot{\boldsymbol{q}})$ then $\partial L/\partial t = 0$ so that h is a **constant**. The conservation formula

$$\boxed{\sum_{j=1}^{n}\left(\frac{\partial L}{\partial \dot{q}_j}\dot{q}_j\right) - L = \text{constant}} \qquad (12.29)$$

is called the **energy integral** of the system $\mathcal{S}$.

Systems for which $L = L(\boldsymbol{q}, \dot{\boldsymbol{q}})$ are said to be **autonomous**. The above result can therefore be expressed in the form:

> **Autonomous systems conserve h**
>
> In any motion of an autonomous system, the energy function $h(\boldsymbol{q}, \dot{\boldsymbol{q}})$ is conserved.

Example 12.16 *A charge moving in a magnetic field*

Find the energy integral for a particle of mass m and charge e moving in the *static* magnetic field $\boldsymbol{B}(\boldsymbol{r})$.

Solution

For this problem

$$L = \tfrac{1}{2}m\,\dot{\boldsymbol{r}}\cdot\dot{\boldsymbol{r}} + e\,\dot{\boldsymbol{r}}\cdot\boldsymbol{A}(\boldsymbol{r}),$$

where $\boldsymbol{A}$ is the magnetic vector potential. Since $\partial L/\partial t = 0$, the energy integral exists and has the form

$$\dot{x}\,\frac{\partial L}{\partial \dot{x}} + \dot{y}\,\frac{\partial L}{\partial \dot{y}} + \dot{z}\,\frac{\partial L}{\partial \dot{z}} - L = h,$$

where h is a constant. On using the formula for L, this becomes

$$m\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) + e\,\dot{\boldsymbol{r}}\cdot\boldsymbol{A} - \tfrac{1}{2}m\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) - e\,\dot{\boldsymbol{r}}\cdot\boldsymbol{A} = h,$$

that is

$$\tfrac{1}{2}m\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) = h.$$

This is the required **energy integral**. In this case, the constant h is the **kinetic energy** of the particle.

This result is well known. When a charged particle moves in a magnetic field, the force is perpendicular to the velocity of the charge. Thus no work is done by the magnetic field and so the *kinetic energy of the particle is conserved*. For this system, V does not exist since the force exerted by the magnetic field is velocity dependent; the total energy E is therefore not defined. ■

Case C If $\mathcal{S}$ is a **conservative standard system**, then $\mathcal{S}$ is autonomous and so h is **conserved**. In addition, the energy integral can be written in a more familiar form. In this case, $L = T - V$, where T has the form

$$T = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}(\boldsymbol{q})\,\dot{q}_j\dot{q}_k$$

(see Example 12.9), and $V = V(\boldsymbol{q})$. Hence

$$\frac{\partial L}{\partial \dot{q}_j} = \frac{\partial T}{\partial \dot{q}_j} - 0 = 2\sum_{k=1}^{n} a_{jk}(\boldsymbol{q})\,\dot{q}_k$$

and so

$$\sum_{j=1}^{n} \frac{\partial L}{\partial \dot{q}_j}\dot{q}_j = 2\sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}(\boldsymbol{q})\,\dot{q}_j\dot{q}_k = 2T.$$

The **energy integral** therefore becomes

$$2T - (T - V) = \text{constant},$$

that is

$$\boxed{T + V = \text{constant}} \tag{12.30}$$

which is the classical form of **conservation of energy**. In this case, the constant is the **total energy** E of the system.

12.9 GENERALISED MOMENTA

The generalised momenta of a mechanical system are defined in a different way to conventional linear and angular momentum.

Definition 12.9 ***Generalised momenta*** *Consider a holonomic mechanical system with Lagrangian $L = L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$. Then the scalar quantity p_j defined by*

$$p_j = \frac{\partial L}{\partial \dot{q}_j}$$

*is called the **generalised momentum** corresponding to the coordinate q_j. It is also called the momentum **conjugate** to q_j.*

Example 12.17 *Finding generalised momenta*

Consider the problem in Example 12.10 whose Lagrangian is

$$L = \tfrac{1}{2}M\dot{x}^2 + \tfrac{1}{2}m\left(\dot{x}^2 + \dot{y}^2 + 2\dot{x}\dot{y}\cos\alpha\right) + mgy\sin\alpha.$$

Find the generalised momenta.

Solution

With this Lagrangian, the momenta p_x and p_y are given by

$$\begin{aligned} p_x &= \frac{\partial L}{\partial \dot{x}} = M\dot{x} + m\left(\dot{x} + \dot{y}\cos\alpha\right), \\ p_y &= \frac{\partial L}{\partial \dot{y}} = m\left(\dot{y} + \dot{x}\cos\alpha\right). \blacksquare \end{aligned}$$

Generalised momenta are often recognisable as components of linear or angular momentum of the system. In the above example, p_x is the horizontal component of the linear momentum of $\mathcal{S}$, but p_y is *not* a component of linear momentum.

Conservation of generalised momenta

In terms of the generalised momentum p_j, the j-th Lagrange equation can be written

$$\frac{dp_j}{dt} = \frac{\partial L}{\partial q_j}.$$

It follows that *if $\partial L/\partial q_j = 0$ (that is, if the coordinate q_j is absent from the Lagrangian), then the generalised momentum p_j is constant in any motion*. Such 'absent' coordinates are said to be **cyclic**. We have therefore shown that:

> **Conservation of momentum**
>
> If q_j is a cyclic coordinate (in the sense that it does not appear in the Lagrangian), then p_j, the generalised momentum conjugate to q_j, is constant in any motion.

In the last example, the coordinate x is cyclic but y is not. It follows that p_x is conserved but p_y is not.

Example 12.18 *A cyclic coordinate for the spherical pendulum*

Consider the spherical pendulum shown in Figure 11.7. The Lagrangian L is given by

$$L = \tfrac{1}{2}ma^2\left[\dot{\theta}^2 + (\sin\theta\,\dot{\phi})^2\right] + mga\cos\theta,$$

where θ, ϕ are the polar angles shown. Verify that ϕ is a cyclic coordinate and find the corresponding conserved momentum.

Solution

Since $\partial L/\partial \phi = 0$, the coordinate ϕ is **cyclic**. It follows that the conjugate momentum p_ϕ is conserved, where

$$p_\phi = \frac{\partial L}{\partial \dot{\phi}} = ma^2 \sin^2 \theta \, \dot{\phi}.$$

This generalised momentum is actually the angular momentum of the pendulum about the polar axis. ■

12.10 SYMMETRY AND CONSERVATION PRINCIPLES

The existence of a cyclic coordinate is not the only reason why a generalised momentum (or momentum-like quantity) may be conserved. Indeed, whether a cyclic coordinate is present depends not only on the system, but also on *which coordinates are chosen*; if the 'wrong' coordinates are chosen then the conserved quantity will be missed. The existence of conserved quantities of the form $F(\boldsymbol{q}, \dot{\boldsymbol{q}})$ is in fact closely linked with **symmetries of the system**. We illustrate this by the following two results, which are the most important of such cases.

Theorem 12.1 ***Invariance of*** V ***under translation*** *Let $\mathcal{S}$ be a conservative standard system with potential energy V. Then if $\mathcal{S}$ can be translated (as if rigid) parallel to a constant vector $\boldsymbol{n}$ without violating any constraints, and if V is unchanged by this translation, then, in any motion of $\mathcal{S}$, the component of **linear momentum** in the $\boldsymbol{n}$-direction is **conserved**.*

Proof. Let $\{\boldsymbol{r}_i\}$ be any configuration of $\mathcal{S}$ and let the corresponding point in configuration space be $\boldsymbol{q}$. Then a (rigid) displacement λ in the $\boldsymbol{n}$-direction will have the effect

$$\boldsymbol{r}_i \to \boldsymbol{r}_i^\lambda,$$

where

$$\boldsymbol{r}_i^\lambda = \boldsymbol{r}_i + \lambda \boldsymbol{n}.$$

Since this displacement is consistent with the system constraints, $\{\boldsymbol{r}_i^\lambda\}$ is also a configuration of $\mathcal{S}$ and corresponds to some point $\boldsymbol{q}^\lambda$ in configuration space. Thus, in configuration space, the displacement has the effect

$$\boldsymbol{q} \to \boldsymbol{q}^\lambda,$$

where

$$\boldsymbol{r}_i^\lambda = \boldsymbol{r}_i(\boldsymbol{q}^\lambda).$$

Note that $\lambda = 0$ corresponds to the undisplaced state so that $\boldsymbol{r}_i^\lambda = \boldsymbol{r}_i$ and $\boldsymbol{q}^\lambda = \boldsymbol{q}$ when $\lambda = 0$.

Suppose now that $\boldsymbol{q}(t)$ is a motion of $\mathcal{S}$ under the potential $V(\boldsymbol{q})$. Then $\boldsymbol{q}$ satisfies Lagrange's equations which we choose to take in the form

$$\sum_i m_i \dot{\boldsymbol{v}}_i \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_j} = -\frac{\partial V}{\partial q_j} \qquad (j = 1, \ldots, n).$$

On multiplying the j-th Lagrange equation by

$$\left[\frac{\partial q_j^\lambda}{\partial \lambda}\right]_{\lambda=0}$$

and summing over j we obtain

$$\sum_i m_i \dot{\boldsymbol{v}}_i \cdot \left(\sum_{j=1}^n \frac{\partial \boldsymbol{r}_i}{\partial q_j}\left[\frac{\partial q_j^\lambda}{\partial \lambda}\right]_{\lambda=0}\right) = -\sum_{j=1}^n \frac{\partial V}{\partial q_j}\left[\frac{\partial q_j^\lambda}{\partial \lambda}\right]_{\lambda=0}.$$

Now

$$\sum_{j=1}^n \frac{\partial \boldsymbol{r}_i}{\partial q_j}\left[\frac{\partial q_j^\lambda}{\partial \lambda}\right]_{\lambda=0} = \sum_{j=1}^n \left[\left(\frac{\partial}{\partial q_j^\lambda}\boldsymbol{r}_i(\boldsymbol{q}^\lambda)\right)\frac{\partial q_j^\lambda}{\partial \lambda}\right]_{\lambda=0} = \left[\frac{d}{d\lambda}\boldsymbol{r}_i(\boldsymbol{q}^\lambda)\right]_{\lambda=0}, \qquad (12.31)$$

by the chain rule. Furthermore

$$\frac{d}{d\lambda}\boldsymbol{r}_i(\boldsymbol{q}^\lambda) = \frac{\partial \boldsymbol{r}_i^\lambda}{\partial \lambda} = \boldsymbol{n},$$

since $\boldsymbol{r}_i^\lambda = \boldsymbol{r}_i(\boldsymbol{q}) + \lambda \boldsymbol{n}$ in the given displacement.

In the same way,

$$\sum_{j=1}^n \frac{\partial V}{\partial q_j}\left[\frac{\partial q_j^\lambda}{\partial \lambda}\right]_{\lambda=0} = \sum_{j=1}^n \left[\left(\frac{\partial}{\partial q_j^\lambda}V(\boldsymbol{q}^\lambda)\right)\frac{\partial q_j^\lambda}{\partial \lambda}\right]_{\lambda=0} = \left[\frac{d}{d\lambda}V(\boldsymbol{q}^\lambda)\right]_{\lambda=0} = 0,$$

since V is unchanged by the displacement, that is, $V(\boldsymbol{q}^\lambda) = V(\boldsymbol{q})$. On combining these results together, we obtain

$$\sum_i m_i \dot{\boldsymbol{v}}_i \cdot \boldsymbol{n} = 0.$$

Finally, since $\boldsymbol{n}$ is a constant vector, we may integrate with respect to t to obtain

$$\left(\sum_i m_i \boldsymbol{v}_i\right) \cdot \boldsymbol{n} = C,$$

where C is a constant. Thus the component of **linear momentum in the *n*-direction is conserved**. ■

For example, this theorem applies to the system shown in Figure 12.6 (the wedge and block). This system can be translated in the x-direction without violating any constraints, and this translation leaves the potential energy unchanged. The conserved quantity is the component of linear momentum in the x-direction.

Theorem 12.2 ***Invariance of*** V ***under rotation*** *Let $\mathcal{S}$ be a conservative standard system with potential energy V. Then if $\mathcal{S}$ can be rotated (as if rigid) about the fixed axis $\{O, \boldsymbol{k}\}$ without violating any constraints, and if V is unchanged by this rotation, then, in any motion of $\mathcal{S}$, the* **angular momentum** *about the axis $\{O, \boldsymbol{k}\}$ is* **conserved**.

Proof. The proof closely follows that in the last theorem. Let λ be the angle turned in a rotation of $\mathcal{S}$ about the fixed axis $\{O, \boldsymbol{k}\}$, where O is also the origin of position vectors. Then by following the same steps, we obtain, as before,

$$\sum_i m_i \dot{\boldsymbol{v}}_i \cdot \left[\frac{\partial \boldsymbol{r}_i^\lambda}{\partial \lambda}\right]_{\lambda=0} = 0.$$

This time the λ-derivative means the rate of change with respect to the *rotation* angle λ so that

$$\frac{\partial \boldsymbol{r}_i^\lambda}{\partial \lambda} = \boldsymbol{k} \times \boldsymbol{r}_i^\lambda.$$

Since $\boldsymbol{r}_i^\lambda = \boldsymbol{r}_i$ when $\lambda = 0$, it follows that

$$\sum_i m_i \dot{\boldsymbol{v}}_i \cdot (\boldsymbol{k} \times \boldsymbol{r}_i) = 0.$$

Finally, since $\boldsymbol{k}$ is a constant vector, we may integrate with respect to t to obtain

$$\left(\sum_i m_i \boldsymbol{r}_i \times \boldsymbol{v}_i\right) \cdot \boldsymbol{k} = C,$$

where C is a constant. Thus the **angular momentum about the axis** $\{O, \boldsymbol{k}\}$ **is conserved**. ∎

For example, this theorem applies to the spherical pendulum. The pendulum can be rotated about the axis $\{O, \boldsymbol{k}\}$ (where O is the support and $\boldsymbol{k}$ points vertically upwards) without violating any constraints, and this rotation leaves the potential energy unchanged. The conserved quantity is the angular momentum of the pendulum about the vertical axis through O.

These two theorems are powerful tools for identifying conserved components of linear or angular momentum even when the system is very complex. For example, any conservative standard system whose potential energy is invariant under *all* translations and rotations conserves all three components of linear and angular momentum, as well as the total energy, making seven conserved quantities in all.

Noether's theorem

The two theorems above are particular instances of an abstract result known as **Noether's theorem**.* In each of the above cases, there is a one-parameter family of mappings $\{\mathfrak{M}^\lambda\}$,

* After the German mathematician Emmy Amalie Noether (1882–1935). Despite the obstacles placed in the way of women academics at the time, she made fundamental contributions to pure mathematics in the areas of invariance theory and abstract algebra. The result now known as Noether's theorem was published in 1918.

parametrised by a real variable λ, that act on the configuration space $\mathcal{Q}$, that is,

$$\boldsymbol{q} \xrightarrow{\mathfrak{M}^\lambda} \boldsymbol{q}^\lambda. \tag{12.32}$$

In each case $\lambda = 0$ corresponds to the identity mapping (that is, $\boldsymbol{q} \to \boldsymbol{q}$), and in each case the potential energy $V(\boldsymbol{q})$ is invariant under $\{\mathfrak{M}^\lambda\}$, that is,

$$V(\boldsymbol{q}^\lambda) = V(\boldsymbol{q}).$$

From these facts, we were able to prove that, in each case, a certain momentum component was a constant of the motion.

This idea was generalised by Noether to apply to *any* Lagrangian system and *any* family of mappings $\{\mathfrak{M}^\lambda\}$, provided that the **Lagrangian** L is **invariant** under $\{\mathfrak{M}^\lambda\}$ in the sense that

$$L(\boldsymbol{q}^\lambda, \dot{\boldsymbol{q}}^\lambda, t) = L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$$

for all λ. In this formula, $\boldsymbol{q}^\lambda$ is a known function of the variables $\boldsymbol{q}$ and λ (as defined by the mapping $\mathfrak{M}^\lambda$), but we have not yet said what we mean by $\dot{\boldsymbol{q}}^\lambda$. This is however defined in the following commonsense way: let λ be fixed and let $\boldsymbol{q}^\lambda$ be the image point of a typical point $\boldsymbol{q}$. Suppose now that the point $\boldsymbol{q}$ has velocity $\dot{\boldsymbol{q}}$ in the configuration space $\mathcal{Q}$. This motion of $\boldsymbol{q}$ *imparts* a velocity to the image point $\boldsymbol{q}^\lambda$, and it is this velocity that we call $\dot{\boldsymbol{q}}^\lambda$. This definition is expressed by the formula

$$\dot{\boldsymbol{q}}^\lambda = \sum_{j=1}^{n} \frac{\partial \boldsymbol{q}^\lambda}{\partial q_j} \dot{q}_j \tag{12.33}$$

from which we see that $\dot{\boldsymbol{q}}^\lambda$ is a known function of the variables $\boldsymbol{q}$, $\dot{\boldsymbol{q}}$ and λ.

The formal statement and proof of Noether's theorem are as follows:

Theorem 12.3 ***Noether's theorem*** *Let $\mathcal{S}$ be a holonomic mechanical system with* ***Lagrangian*** *$L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$ and let $\{\mathfrak{M}^\lambda\}$ be a one-parameter family of* ***mappings*** *that have the action*

$$\boldsymbol{q} \xrightarrow{\mathfrak{M}^\lambda} \boldsymbol{q}^\lambda \tag{12.34}$$

where $\boldsymbol{q}^\lambda = \boldsymbol{q}$ when $\lambda = 0$. If the mappings $\{\mathfrak{M}^\lambda\}$ leave L ***invariant*** *in the sense that*

$$L(\boldsymbol{q}^\lambda, \dot{\boldsymbol{q}}^\lambda, t) = L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) \tag{12.35}$$

for all λ, then the quantity

$$\sum_{j=1}^{n} p_j \left[\frac{\partial q_j^\lambda}{\partial \lambda} \right]_{\lambda=0} \tag{12.36}$$

*is **conserved** in any motion of* $\mathcal{S}$. [Note that the conserved quantity is not generally one of the momenta $\{p_j\}$ but a linear combination of all of them with coefficients depending on $\boldsymbol{q}$.]

Proof. Let $\boldsymbol{q}(t)$ be any physical motion of the system $\mathcal{S}$, that is, a solution of the Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0 \qquad (1 \le j \le n),$$

where $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$ is the Lagrangian of the system $\mathcal{S}$. Now consider the expression

$$\begin{aligned}
\frac{d}{dt}\left(p_j\left[\frac{\partial q_j^\lambda}{\partial \lambda}\right]_{\lambda=0}\right) &= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\left[\frac{\partial q_j^\lambda}{\partial \lambda}\right]_{\lambda=0}\right)\\
&= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right)\left[\frac{\partial q_j^\lambda}{\partial \lambda}\right]_{\lambda=0} + \frac{\partial L}{\partial \dot{q}_j}\left(\frac{d}{dt}\left[\frac{\partial q_j^\lambda}{\partial \lambda}\right]_{\lambda=0}\right)\\
&= \frac{\partial L}{\partial q_j}\left[\frac{\partial q_j^\lambda}{\partial \lambda}\right]_{\lambda=0} + \frac{\partial L}{\partial \dot{q}_j}\left(\frac{d}{dt}\left[\frac{\partial q_j^\lambda}{\partial \lambda}\right]_{\lambda=0}\right)
\end{aligned}$$

on using the j-th Lagrange equation. Now, by the chain rule,

$$\frac{d}{dt}\left(\frac{\partial q_j^\lambda}{\partial \lambda}\right) = \sum_{k=1}^{n} \frac{\partial}{\partial q_k}\left(\frac{\partial q_j^\lambda}{\partial \lambda}\right)\dot{q}_k = \frac{\partial}{\partial \lambda}\left(\sum_{k=1}^{n} \frac{\partial q_j^\lambda}{\partial q_k}\dot{q}_k\right) = \frac{\partial \dot{q}_j^\lambda}{\partial \lambda}$$

by the definition (12.33) of $\dot{\boldsymbol{q}}^\lambda$. It follows that

$$\begin{aligned}
\frac{d}{dt}\left(p_j\left[\frac{\partial q_j^\lambda}{\partial \lambda}\right]_{\lambda=0}\right) &= \frac{\partial L}{\partial q_j}\left[\frac{\partial q_j^\lambda}{\partial \lambda}\right]_{\lambda=0} + \frac{\partial L}{\partial \dot{q}_j}\left[\frac{\partial \dot{q}_j^\lambda}{\partial \lambda}\right]_{\lambda=0}\\
&= \left[\frac{\partial}{\partial q_j^\lambda}L(\boldsymbol{q}^\lambda, \dot{\boldsymbol{q}}^\lambda, t)\,\frac{\partial q_j^\lambda}{\partial \lambda} + \frac{\partial}{\partial \dot{q}_j^\lambda}L(\boldsymbol{q}^\lambda, \dot{\boldsymbol{q}}^\lambda, t)\,\frac{\partial \dot{q}_j^\lambda}{\partial \lambda}\right]_{\lambda=0}
\end{aligned}$$

since $\boldsymbol{q}^\lambda = \boldsymbol{q}$ and $\dot{\boldsymbol{q}}^\lambda = \dot{\boldsymbol{q}}$ when $\lambda = 0$. On summing this result over j, we obtain

$$\begin{aligned}
\frac{d}{dt}\left(\sum_{j=1}^{n} p_j\left[\frac{\partial q_j^\lambda}{\partial \lambda}\right]_{\lambda=0}\right) &= \left[\sum_{j=1}^{n} \frac{\partial}{\partial q_j^\lambda}L(\boldsymbol{q}^\lambda, \dot{\boldsymbol{q}}^\lambda, t)\,\frac{\partial q_j^\lambda}{\partial \lambda} + \sum_{j=1}^{n} \frac{\partial}{\partial \dot{q}_j^\lambda}L(\boldsymbol{q}^\lambda, \dot{\boldsymbol{q}}^\lambda, t)\,\frac{\partial \dot{q}_j^\lambda}{\partial \lambda}\right]_{\lambda=0}\\
&= \left[\frac{d}{d\lambda}L(\boldsymbol{q}^\lambda, \dot{\boldsymbol{q}}^\lambda, t)\right]_{\lambda=0}
\end{aligned}$$

by the chain rule. Finally, we appeal to the invariance of L under the mappings $\{\mathfrak{M}^\lambda\}$. In this case, $L(\boldsymbol{q}^\lambda, \dot{\boldsymbol{q}}^\lambda, t) = L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$ and so

$$\frac{d}{d\lambda}L(\boldsymbol{q}^\lambda, \dot{\boldsymbol{q}}^\lambda, t) = \frac{d}{d\lambda}L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) = 0.$$

It follows that

$$\frac{d}{dt}\left(\sum_{j=1}^{n} p_j \left[\frac{\partial q_j^{\lambda}}{\partial \lambda}\right]_{\lambda=0}\right) = 0$$

and this proves the theorem. ■

The importance of Noether's theorem lies in the general notion that **an invariance of the Lagrangian gives rise to a constant of the motion**. Such invariance properties are of great importance when the Lagrangian formalism is extended to continuous systems and fields. For more details, see Goldstein [4] who will tell you more about Noether's theorem than you wish to know!

Problems on Chapter 12

Answers and comments are at the end of the book.
Harder problems carry a star ($*$).

Conservative systems

12.1 A bicycle chain consists of N freely jointed links forming a closed loop. The chain can slide freely on a smooth horizontal table. How many degrees of freedom has the chain? How many conserved quantities are there in the motion? What is the maximum number of links the chain can have for its motion to be determined by conservation principles alone?

12.2 ***Attwood's machine*** A uniform circular pulley of mass $2m$ can rotate freely about its axis of symmetry which is fixed in a horizontal position. Two masses m, $3m$ are connected by a light inextensible string which passes over the pulley without slipping. The whole system undergoes planar motion with the masses moving vertically. Take the rotation angle of the pulley as generalised coordinate and obtain Lagrange's equation for the motion. Deduce the upward acceleration of the mass m.

12.3 ***Double Attwood machine*** A light pulley can rotate freely about its axis of symmetry which is fixed in a horizontal position. A light inextensible string passes over the pulley. At one end the string carries a mass $4m$, while the other end supports a second light pulley. A second string passes over this pulley and carries masses m and $4m$ at its ends. The whole system undergoes planar motion with the masses moving vertically. Find Lagrange's equations and deduce the acceleration of each of the masses.

12.4 ***The swinging door*** A uniform rectangular door of width $2a$ can swing freely on its hinges. The door is misaligned and the line of the hinges makes an angle α with the upward vertical. Take the rotation angle of the door from its equilibrium position as generalised coordinate and obtain Lagrange's equation for the motion. Deduce the period of small oscillations of the door about the equilibrium position.

12.5 A uniform solid cylinder $\mathcal{C}$ with mass m and radius a rolls on the rough outer surface of a fixed horizontal cylinder of radius b. In the motion, the axes of the two cylinders remain parallel to each other. Let θ be the angle between the plane containing the cylinder axes and the upward vertical. Taking θ as generalised coordinate, obtain Lagrange's equation and verify that it is equivalent to the energy conservation equation.

Initially the cylinder $\mathcal{C}$ is at rest on top of the fixed cylinder when it is given a very small disturbance. Find, as a function of θ, the normal component of the reaction force exerted on $\mathcal{C}$. Deduce that $\mathcal{C}$ will leave the fixed cylinder when $\theta = \cos^{-1}(4/7)$. Is the assumption that rolling persists up to this moment realistic?

12.6 A uniform disk of mass M and radius a can roll along a rough horizontal rail. A particle of mass m is suspended from the centre C of the disk by a light inextensible string of length b. The whole system moves in the vertical plane through the rail. Take as generalised coordinates x, the horizontal displacement of C, and θ, the angle between the string and the downward vertical. Obtain Lagrange's equations. Show that x is a cyclic coordinate and find the corresponding conserved momentum p_x. Is p_x the horizontal linear momentum of the system?

Given that θ remains small in the motion, find the period of small oscillations of the particle.

12.7 A uniform ball of mass m rolls down a rough wedge of mass M and angle α, which itself can slide on a smooth horizontal table. The whole system undergoes planar motion. How many degrees of freedom has this system? Obtain Lagrange's equations. For the special case in which $M = 3m/2$, find (i) the acceleration of the wedge, and (ii) the acceleration of the ball relative to the wedge.

12.8 A rigid rod of length $2a$ has its lower end in contact with a smooth horizontal floor. Initially the rod is at an angle α to the upward vertical when it is released from rest. The subsequent motion takes place in a vertical plane. Take as generalised coordinates x, the horizontal displacement of the centre of the rod, and θ, the angle between the rod and the upward vertical. Obtain Lagrange's equations. Show that x remains constant in the motion and verify that the θ-equation is equivalent to the energy conservation equation.

* Find, in terms of the angle θ, the reaction exerted on the rod by the floor.

Moving constraints

12.9 A particle P is connected to one end of a light inextensible string which passes through a small hole O in a smooth horizontal table and extends below the table in a vertical straight line. P slides on the upper surface of the table while the string is pulled downwards from below in a prescribed manner. (Suppose that the length of the horizontal part of the string is $R(t)$ at time t.) Take θ, the angle between OP and some fixed reference line in the table, as generalised coordinate and obtain Lagrange's equation. Show that θ is a cyclic coordinate and find (and identify) the corresponding conserved momentum p_θ. Why is the kinetic energy not conserved?

If the constant value of p_θ is mL, find the tension in the string at time t.

12.10 A particle P of mass m can slide along a smooth rigid straight wire. The wire has one of its points fixed at the origin O, and is made to rotate in the (x, y)-plane with angular speed Ω. Take r, the distance of P from O, as generalised coordinate and obtain Lagrange's equation.

Initially the particle is a distance a from O and is at rest relative to the wire. Find its position at time t. Find also the energy function h and show that it is conserved even though there is a time dependent constraint.

12.11 ***Yo-yo with moving support*** A uniform circular cylinder (a yo-yo) has a light inextensible string wrapped around it so that it does not slip. The free end of the string is fastened to a support and the yo-yo moves in a vertical straight line with the straight part of the string also vertical. At the same time the support is made to move vertically having upward displacement $Z(t)$ at time t. Take the rotation angle of the yo-yo as generalised coordinate and obtain Lagrange's equation. Find the acceleration of the yo-yo. What upwards acceleration must the support have so that the centre of the yo-yo can remain at rest?

Suppose the whole system starts from rest. Find an expression for the total energy $E = T + V$ at time t.

12.12 ***Pendulum with a shortening string*** A particle is suspended from a support by a light inextensible string which passes through a small fixed ring vertically below the support. The particle moves in a vertical plane with the string taut. At the same time the support is made to move vertically having an upward displacement $Z(t)$ at time t. The effect is that the particle oscillates like a simple pendulum whose string length at time t is $a - Z(t)$, where a is a positive constant. Take the angle between the string and the downward vertical as generalised coordinate and obtain Lagrange's equation. Find the energy function h and the total energy E and show that $h = E - m\dot{Z}^2$. Is either quantity conserved?

12.13* ***Bug on a hoop*** A uniform circular hoop of mass M can slide freely on a smooth horizontal table, and a bug of mass m can run on the hoop. The system is at rest when the bug starts to run. What is the angle turned through by the hoop when the bug has completed one lap of the hoop?

Velocity dependent potentials and Lagrangians

12.14 Suppose a particle is subjected to a time dependent force of the form $\boldsymbol{F} = f(t)\,\mathrm{grad}\,W(\boldsymbol{r})$. Show that this force can be represented by the time dependent potential $U = -f(t)W(\boldsymbol{r})$. What is the value of U when $\boldsymbol{F} = f(t)\,\boldsymbol{i}$?

12.15 ***Charged particle in an electrodynamic field*** Show that the velocity dependent potential

$$U = e\,\phi(\boldsymbol{r}, t) - e\,\dot{\boldsymbol{r}} \cdot \boldsymbol{A}(\boldsymbol{r}, t)$$

represents the Lorentz force $\boldsymbol{F} = e\,\boldsymbol{E} + e\,\boldsymbol{v} \times \boldsymbol{B}$ that acts on a charge e moving with velocity $\boldsymbol{v}$ in the general *electrodynamic* field $\{\boldsymbol{E}(\boldsymbol{r}, t), \boldsymbol{B}(\boldsymbol{r}, t)\}$. Here $\{\phi, \boldsymbol{A}\}$ are the *electrodynamic* *potentials* that generate the field $\{\boldsymbol{E}, \boldsymbol{B}\}$ by the formulae

$$\boldsymbol{E} = -\operatorname{grad}\phi - \frac{\partial \boldsymbol{A}}{\partial t}, \qquad \boldsymbol{B} = \operatorname{curl}\boldsymbol{A}.$$

Show that the potentials $\phi = 0$, $\boldsymbol{A} = tz\,\boldsymbol{i}$ generate a field $\{\boldsymbol{E}, \boldsymbol{B}\}$ that satisfies all four Maxwell equations in free space. A particle of mass m and charge e moves in this field. Find the Lagrangian of the particle in terms of Cartesian coordinates. Show that x and y are cyclic coordinates and find the conserved momenta p_x, p_y.

12.16* ***Relativistic Lagrangian*** The relativistic Lagrangian for a particle of rest mass m_0 moving along the x-axis under the simple harmonic potential field $V = \frac{1}{2}m_0\Omega^2x^2$ is given by

$$L = m_0c^2\left(1 - \left(1 - \frac{\dot{x}^2}{c^2}\right)^{1/2}\right) - \tfrac{1}{2}m_0\Omega^2x^2.$$

Obtain the energy integral for this system and show that the period of oscillations of amplitude a is given by

$$\tau = \frac{4}{\Omega}\int_0^{\pi/2} \frac{1 + \frac{1}{2}\epsilon^2\cos^2\theta}{\left(1 + \frac{1}{4}\epsilon^2\cos^2\theta\right)^{1/2}}\,d\theta,$$

where the dimensionless parameter $\epsilon = \Omega a/c$.

Deduce that

$$\tau = \frac{2\pi}{\Omega}\left[1 + \tfrac{3}{16}\epsilon^2 + O\left(\epsilon^4\right)\right],$$

when ϵ is small.

Conservation principles and symmetry

12.17 A particle of mass m moves under the gravitational attraction of a fixed mass M situated at the origin. Take polar coordinates r, θ as generalised coordinates and obtain Lagrange's equations. Show that θ is a cyclic coordinate and find (and identify) the conserved momentum p_θ.

12.18 A particle P of mass m slides on the smooth inner surface of a circular cone of semi-angle α. The axis of symmetry of the cone is vertical with the vertex O pointing downwards. Take as generalised coordinates r, the distance OP, and ϕ, the azimuthal angle about the vertical through O. Obtain Lagrange's equations. Show that ϕ is a cyclic coordinate and find (and identify) the conserved momentum p_ϕ.

12.19 A particle of mass m and charge e moves in the magnetic field produced by a current I flowing in an infinite straight wire that lies along the z-axis. The vector potential $\boldsymbol{A}$ of the induced magnetic field is given by

$$A_r = A_\theta = 0, \qquad A_z = -\left(\frac{\mu_0 I}{2\pi}\right)\ln r,$$

where r, θ, z are cylindrical polar coordinates. Find the Lagrangian of the particle. Show that θ and z are cyclic coordinates and find the corresponding conserved momenta.

12.20 A particle moves freely in the gravitational field of a fixed mass distribution. Find the conservation principles that correspond to the symmetries of the following fixed mass distributions: (i) a uniform sphere, (ii) a uniform half plane, (iii) two particles, (iv) a uniform right circular cone, (v) an infinite uniform circular cylinder.

12.21* ***Helical symmetry*** A particle moves in a conservative field whose potential energy V has *helical symmetry*. This means that V is invariant under the *simultaneous* operations (i) a rotation through any angle α about the axis Oz, and (ii) a translation $c\alpha$ in the z-direction. What conservation principle corresponds to this symmetry?

Computer assisted problem

12.22 ***Upside-down pendulum*** A particle P is attached to a support S by a light rigid rod of length a, which is freely pivoted at S. P moves in a vertical plane through S and at the same time the support S is made to oscillate vertically having upward displacement $Z = \epsilon a \cos pt$ at time t. Take θ, the angle between SP and the *upward* vertical, as generalised coordinate and show that Lagrange's equation is

$$\ddot{\theta} - \left(\Omega^2 + \epsilon p^2 \cos pt\right) \sin\theta = 0,$$

where $\Omega^2 = g/a$. The object is to show that, for suitable choices of the parameters, the pendulum is stable in the vertically *upwards* position!

First write the equation in dimensionless form by introducing the dimensionless time $\tau = pt$. Then $\theta(\tau)$ satisfies

$$\frac{d^2\theta}{d\tau^2} - \left(\frac{\Omega^2}{p^2} + \epsilon \cos\tau\right) \sin\theta = 0.$$

Solve this equation numerically with initial conditions in which the pendulum starts from rest near the upward vertical. Plot the solution $\theta(\tau)$ as a function of τ for about twenty oscillations of the support. Try $\epsilon = 0.3$ with increasing values of the parameter p/Ω in the range $1 \le p/\Omega \le 10$. You will know that the upside-down pendulum is stable when θ remains small in the subsequent motion.

Even more surprisingly, it is possible to stabilise the double pendulum (or any multiple pendulum) in the upside-down position by vibrating the support. See Acheson [1] for photographs of a triple pendulum (and even a length of floppy wire) stabilised in the upside-down position by vibrating the support. However, the famous but elusive 'Indian Rope Trick', in which a small boy climbs up a self-supporting vertical rope, has yet to be demonstrated!

Chapter Thirteen

The calculus of variations and Hamilton's principle

KEY FEATURES

The key features of this chapter are **integral functionals** and the functions that make them **stationary**, the **Euler–Lagrange** equation and **extremals**, and the importance of **variational principles**.

The notion that physical processes are governed by **minimum principles** is older than most of science. It is based on the long held belief that nature arranges itself in the most 'economical' way. Actually, many 'minimum' principles have, on closer inspection, turned out to make their designated quantity *stationary*, but not necessarily a minimum. As a result, they are now known to be **variational principles**, but they are no less important because of this. A good example of a variational principle is **Fermat's principle** of geometrical optics, which was proposed in 1657 as *Fermat's principle of least time* in the form:

Of all the possible paths that a light ray might take between two fixed points, the actual path is the one that minimises the travel time of the ray.

Fermat showed that the laws of reflection and refraction could be derived from his principle, and proposed that the principle was true in general. Not only did Fermat's principle 'explain' the known laws of optics, it was simple and elegant, and was capable of extending the laws of optics far beyond the results that led to its conception. This example explains why variational principles continue to be sought; it is because of their innate simplicity and elegance, and the generality of their application.

The variational principle on which it is possible to base the whole of classical mechanics was discovered by Hamilton* and is known as **Hamilton's principle**.† In its original form, it stated that:

Of all the kinematically possible motions that take a mechanical system from one given configuration to another within a given time interval, the actual motion is the one that minimises the time integral of the Lagrangian of the system.

Lagrange's equations of motion can be derived from Hamilton's principle, which can therefore be taken as the basic postulate of classical mechanics, instead of Newton's laws. More importantly however, Hamilton's principle has had a far reaching influence on many areas of physics, where apparently non-mechanical systems (fields, for example) can be described in the language of classical mechanics, and their behaviour characterised by 'Lagrangians'. Hamilton's principle is generally regarded as one of the most elegant and far reaching principles in physics.

In order to get concrete results from a variational principle, it is usually necessary to convert it to a differential equation. This can be done by using the **calculus of variations**, which is concerned with minimising or maximising the value of an integral functional. The calculus of variations is a large subject and we develop only those aspects most relevant to interesting physical problems, and to the understanding and use of variational principles.

13.1 SOME TYPICAL MINIMISATION PROBLEMS

The **calculus of variations** arose from attempts to solve minimimisation and maximisation problems that occur naturally in physics and mathematics, but the scope of applications has since widened greatly. We begin by describing three minimisation problems taken from geometry, physics and economics respectively. Maximisation problems also occur, but these can be converted into minimisation problems merely by reversing the sign of the quantity to be maximised. Thus we lose no generality by presenting the theory for minimisation problems only.

1. Shortest paths – geodesics A basic problem of the calculus of variations is that of finding the **path of shortest length** that connects two given points A and B. If the path has no constraints to satisfy, such as having to go round obstacles or lie on a given curved

* Sir William Rowan Hamilton (1805–1865), was a great genius but an unhappy man. He was appointed Professor of Astronomy at Trinity College Dublin at the age of twenty one, whilst still an undergraduate. Much of his early work is on optics where he introduced the notion of the characteristic function. His paper *On a General Method in Dynamics*, which contains what is now called Hamilton's principle, was presented to the Royal Irish Academy in 1834. He was knighted in 1835. However, his personal life was as chaotic as his academic achievments were brilliant. He was frustrated in love, frequently depressed and a heavy drinker; this culminated in his making an exhibition of himself at a meeting of the Irish Geological Society. He spent the later years of his life working on the theory of quaternions, but they were never the great discovery he had hoped for.

† Hamilton's principle is sometimes called the *principle of least action*. The terminology in this area is confusing, since another variational principle of mechanics, Maupertuis's principle, is also referred to as the principle of least action.

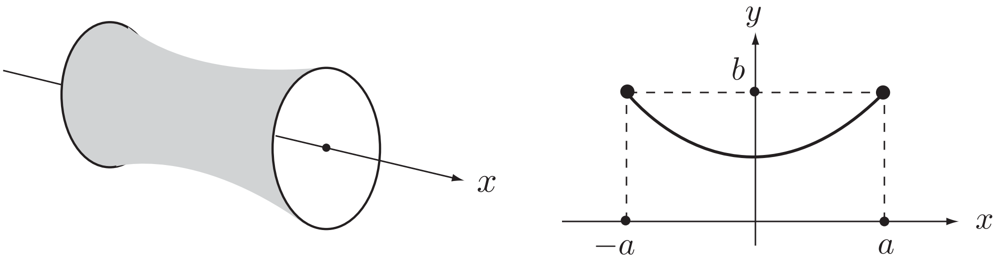


FIGURE 13.1 A soap film is stretched between two circular wires. It has the form of a surface of revolution generated by rotating the curve $y = y(x)$ about the x-axis.

surface, the answer is well known; it is the straight line joining A and B. However, it is still instructive to formulate this problem and to check that the calculus of variations does yield the expected result.

Suppose that $A = (0, 0)$, $B = (1, 0)$ and that the general path in the (x, y)-plane connecting these points is $y = y(x)$. Then the total length of the path is

$$L[y] = \int_0^1 \left[1 + \left(\frac{dy}{dx}\right)^2\right]^{1/2} dx. \tag{13.1}$$

The **problem** is to find the function $y(x)$, satisfying the end conditions $y(0) = 0$, $y(1) = 0$, that minimises the length L. This is the subject of Problem 13.3; the answer is indeed the straight line $y = 0$.

In general, paths of shortest length are called **geodesics**. For example, geodesics on the surface of a sphere are great circles. Some surfaces (such as the cylinder and cone) are *developable*, which means that they can be rolled out flat without changing any lengths. The geodesic can be drawn while the surface is flat (it is now a straight line) and the surface can then be rolled back up again. In general however, surfaces are not developable and geodesics have to be found by the calculus of variations.

2. The soap film problem Two rigid circular wires each of radius b have the same axis of symmetry and are fixed at a distance $2a$ from each other. A soap film is created which spans the two wires as shown in Figure 13.1. The soap film has the form of a surface of revolution with the two circular ends open. What is the shape of the soap film?

Fortunately, we can formulate this problem without resorting to the theory of thin membranes! Since the air pressure is the same on either side of the film, and since the effect of gravity is negligible, surface tension is the dominant effect. The condition that the total energy be a minimum in equilibrium is therefore equivalent to the condition that the *area of the film be a minimum*.

Let the film be the surface generated by rotating the curve $y = y(x)$ about the x-axis. Then the surface area A of the film is

$$A[y] = 2\pi \int_{-a}^{a} y \left\{1 + \dot{y}^2\right\}^{1/2} dx, \tag{13.2}$$

where $\dot{y}$ means dy/dx. The **problem** is to find the function $y(x)$, satisfying the end conditions $y(-a) = b$, $y(a) = b$, that minimises the area A. This is the subject of Problem 13.7.

3. A mimimum cost strategy A manufacturer must produce a volume X of a product in time T. Let $x = x(t)$ be the volume produced after time t and suppose that there is a production cost $\alpha + \beta\dot{x}$ per unit volume of product and a storage cost γx per unit time, where α, β and γ are positive constants. The term $\beta\dot{x}$ is a simple model of the increased costs associated with faster production. Then the total cost C of the production run is

$$C[x] = \int_0^X (\alpha + \beta\dot{x})\, dx + \int_0^T \gamma x\, dt,$$

which can be written in the form

$$C[x] = \int_0^T \{(\alpha + \beta\dot{x})\dot{x} + \gamma x\}\, dt. \tag{13.3}$$

The **problem** is to find the function $x(t)$, satisfying the end conditions $x(0) = 0$, $x(T) = X$, that minimises the cost C. This is the subject of Problem 13.6; it is found that producing the goods at a uniform rate is *not* the best strategy.

Integral functionals

The expressions $L[y]$, $A[y]$ and $C[x]$ are examples of **integral functionals**. Functionals differ from ordinary functions in that the independent variable is a *function*, not a number; however, the dependent variable is a number, as usual. *The calculus of variations is concerned with minimising or maximising integral functionals.*

13.2 THE EULER–LAGRANGE EQUATION

Before we begin the theory proper, it is useful to recall the procedure for finding the value of x that minimises an *ordinary function* $f(x)$ on the interval $a \le x \le b$. The procedure is as follows:

1. First find the values of x (in the range $a < x < b$) that satisfy the equation $f'(x) = 0$. These are the **stationary points** of $f(x)$. They are so called because, if x^* is a stationary point, then

$$f(x^* + h) - f(x^*) = O\left(h^2\right) \tag{13.4}$$

for all sufficiently small h. (That is, at a stationary point, the change in f due to a small change h in x is of order h^2.)

2. Now determine the nature of each stationary point, that is, whether it is a minimum point, a maximum point, or neither. This can usually be done by examining the sign of $f''(x^*)$. The minimum points are the **local minima** of f. They are so called because

$$f(x^* + h) \ge f(x^*) \tag{13.5}$$

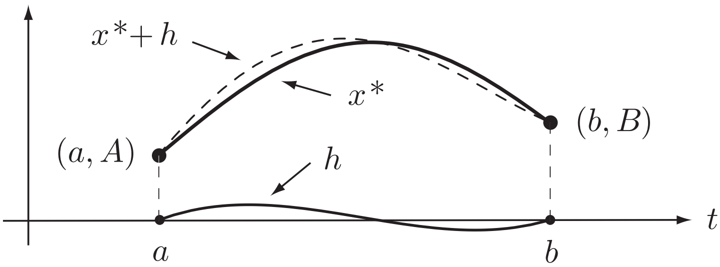


FIGURE 13.2 The minimising function x^* is perturbed by the admissible variation h.

for *sufficiently small* h, but not necessarily for all h. (In other words, the inequality $f(x) \geq f(x^*)$ is true when x is close enough to x^*.)

3. Determine the values of f at the extreme points $x = a$, $x = b$.
4. The **global minimum** of f is then the least of the local minima of f and the extreme values of f.

Each of these steps has its counterpart in the calculus of variations. However, since this material is large enough to fill a book by itself, we will mainly be concerned with the first step. This will still be enough to narrow down the search for the minimising function $x^*(t)$ to a finite number of possibilies and often one ends up with only *one* possibility. Thus, if it is 'known' (rigorously or otherwise!) that a minimising function does exist, then the problem is solved.

The **general problem** in the calculus of variations is that of finding a function $x^*(t)$ that minimises an integral functional of the form

$$J[x] = \int_a^b F(x, \dot{x}, t)\, dt, \tag{13.6}$$

where F is a given function of *three independent variables*.* Suppose that the function $x^*(t)$ minimises the functional $J[x]$. This means that

$$J[x] \geq J[x^*] \tag{13.7}$$

for all *admissible* functions $x(t)$. Here **admissible** means that x must satisfy whatever end conditions are prescribed at $t = a$ and $t = b$. We will always assume that these

* This means that, despite the fact that x, $\dot{x}$ and t are clearly *not* independent of each other (x is a function of t and $\dot{x}$ is the derivative of x), the *partial derivatives* of F are evaluated as if x, $\dot{x}$ and t were three independent variables. For example, in the case of the cost functional C given by equation (13.3), $F = (\alpha + \beta\dot{x})\dot{x} + \gamma x$ and the partial derivatives of F are

$$\frac{\partial F}{\partial x} = \gamma \qquad \frac{\partial F}{\partial \dot{x}} = \alpha + 2\beta\dot{x} \qquad \frac{\partial F}{\partial t} = 0.$$

conditions have the form $x(a) = A$ and $x(b) = B$, where A, B are given. It is convenient to regard the function $x(t)$ that appears in (13.7) as being composed of $x^*(t)$ together with a **variation*** $h(t)$ so that we may alternatively write

$$J[x^* + h] \geq J[x^*] \tag{13.8}$$

for all admissible variations $h(t)$. Since x must satisfiy the same end conditions as x^*, the **admissible variations** are those for which $h(a) = h(b) = 0$ (see Figure 13.2).

Most readers will find the theory that follows quite difficult. To aid understanding, the argument is broken down into three separate steps.

The variation in J and the meaning of 'stationary'

The first step is to give a meaning to the statement that a function $x(t)$ makes the functional $J[x]$ stationary.

Let $x^*(t)$ be *any* admissible function and $h(t)$ an admissible variation. When h is a *small* variation, we can estimate the corresponding variation in $J[x]$ by ordinary calculus, as follows:

Let t have any fixed value. Then x and $\dot{x}$ are just real numbers and the variation in F due to the variation h in x is†

$$F(x^* + h, \dot{x}^* + \dot{h}, t) - F(x^*, \dot{x}^*, t) = h\frac{\partial F}{\partial x}(x^*, \dot{x}^*, t) + \dot{h}\frac{\partial F}{\partial \dot{x}}(x^*, \dot{x}^*, t) + O\left(h^2 + \dot{h}^2\right),$$

when h and $\dot{h}$ are both small. On integrating both sides of this equation with respect to t over the interval $[a, b]$, the corresponding variation in J is given by

$$J[x^* + h] - J[x^*] = \int_a^b \left[h\frac{\partial F}{\partial x}(x^*, \dot{x}^*, t) + \dot{h}\frac{\partial F}{\partial \dot{x}}(x^*, \dot{x}^*, t) \right] dt + O\left(||h||^2\right), \tag{13.9}$$

for small $||h||$, where $||h||$ is defined by

$$||h|| = \max_{a \leq t \leq b} |h(t)| + \max_{a \leq t \leq b} |\dot{h}(t)|$$

and is called the **norm**‡ of h. (When $||h||$ is small, both $|h(t)|$ and $|\dot{h}(t)|$ are small throughout the interval $[a, b]$.) The second term in the integrand of equation (13.9) can be integrated by parts to give

$$\int_a^b \dot{h}\frac{\partial F}{\partial \dot{x}}(x, x^*, t)\, dt = \left[h\frac{\partial F}{\partial \dot{x}}(x^*, \dot{x}^*, t) \right]_{t=a}^{t=b} - \int_a^b h\frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}}(x^*, \dot{x}^*, t) \right) dt$$

---

* These are the variations that give the 'calculus of variations' its name.

† The formula we are using is that, if $G(u, v)$ is a function of the independent variables u and v, then the variation in G caused by the variations $u_0 \to u_0 + h$ and $v_0 \to v_0 + k$ is given by

$$G(u_0 + h, v_0 + k) - G(u_0, v_0) = h\frac{\partial G}{\partial u}(u_0, v_0) + k\frac{\partial G}{\partial v}(u_0, v_0) + O\left(h^2 + k^2\right),$$

for small h, k. Note that the partial derivatives are evaluated at the 'starting point' (u_0, v_0).

‡ Do not be too concerned over the exact definition of $||h||$. It must be written in some such way for mathematical correctness, but the only property that we will need is that $||h||$ is *proportional* to h in the sense that, if h is multiplied by a constant λ, then so is $||h||$.

and, since h is an *admissible* variation satisfying $h(a) = h(b) = 0$, the integrated term evaluates to zero. We thus obtain

$$J[x^* + h] - J[x^*] = \int_a^b \left[\frac{\partial F}{\partial x}(x^*, \dot{x}^*, t) - \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}}(x^*, \dot{x}^*, t) \right) \right] h\, dt + O\left(||h||^2 \right). \tag{13.10}$$

This is the variation in J caused by the admissible variation h in x when $||h||$ is small. The variation in J is therefore linear in to h with an error term of order $||h||^2$. By analogy with the case of ordinary functions, we say that x^* makes $J[x]$ stationary if the linear term is zero, leaving only the error term.

Definition 13.1 ***Stationary*** J *The function $x^*(t)$ is said to make the functional $J[x]$* ***stationary*** *if*

$$J[x^* + h] - J[x^*] = O\left(||h||^2 \right) \tag{13.11}$$

when $||h||$ is small.

It follows that the condition that x^* makes $J[x]$ stationary is equivalent to the condition that

$$\int_a^b \left[\frac{\partial F}{\partial x}(x^*, \dot{x}^*, t) - \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}}(x^*, \dot{x}^*, t) \right) \right] h\, dt = 0 \tag{13.12}$$

for all admissible variations h.

Minimising functions make J stationary

Suppose now that $x^*(t)$ provides a **local minimum** for $J[x]$ in the sense that

$$J[x^* + h] \geq J[x^*] \tag{13.13}$$

when $||h||$ is small. We will now show that such an x^* makes $J[x]$ stationary.

If we substitute equation (13.10) into the inequality (13.13), we obtain

$$\int_a^b \left[\frac{\partial F}{\partial x}(x^*, \dot{x}^*, t) - \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}}(x^*, \dot{x}^*, t) \right) \right] h\, dt + O\left(||h||^2 \right) \geq 0 \tag{13.14}$$

for small $||h||$. It follows from this inequality that the integral term must be zero. The proof is as follows:

In the inequality (13.14), let h be replaced by λh, where λ is a positive constant. On dividing through by λ, this gives

$$\int_a^b \left[\frac{\partial F}{\partial x}(x^*, \dot{x}^*, t) - \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}}(x^*, \dot{x}^*, t) \right) \right] h\, dt + \lambda O\left(||h||^2 \right) \geq 0.$$

On letting $\lambda \to 0$, we find that

$$\int_a^b \left[\frac{\partial F}{\partial x}(x^*, \dot{x}^*, t) - \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}}(x^*, \dot{x}^*, t) \right) \right] h\, dt \geq 0$$

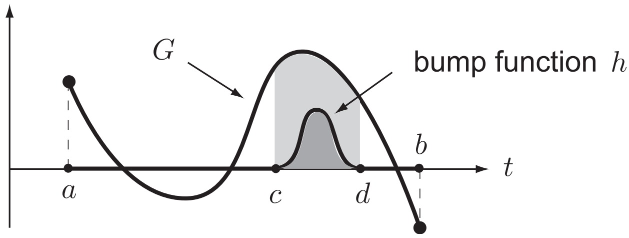


FIGURE 13.3 An interval (c, d) in which $G(t) > 0$ and a corresponding bump function $h(t)$. The integral of $G \times h$ must be positive.

for *all* admissible variations h. In particular, this inequality must remain true if h is replaced by $-h$, but this is only possible if the integral has the value zero.

Equation (13.10) therefore reduces to

$$J[x^* + h] - J[x^*] = O\left(||h||^2\right) \tag{13.15}$$

for small $||h||$. By definition, this means that x^* makes $J[x]$ stationary. The same result applies to functions that provide a local maximum for $J[x]$. Our result is summarised as follows:

Functions that minimise or maximise J make J stationary

If the function x^* provides a local mimimum for the integral functional $J[x]$, then x^* makes J stationary. The same applies to functions that provide a local maximum for J.

Euler–Lagrange equation

We will now obtain the differential equation that must be satified by a function $x^*(t)$ that makes $J[x]$ stationary. This is the counterpart of the elementary condition $f'(x) = 0$. To find this, we return to equation (13.12). In the integrand, the function inside the square brackets looks complicated but it is just a function of t and, if we denote it by $G(t)$, then

$$\int_a^b G(t)\, h(t)\, dt = 0 \tag{13.16}$$

for *all* admissible variations h. In fact, the only function $G(t)$ for which this is possible is the *zero function*, that is, $G(t) = 0$ for $a < t < b$. The proof is as follows:

Suppose that $G(t)$ is *not* the zero function. Then there must exist some interval (c, d), lying inside the interval (a, b) in which $G(t) \neq 0$ and thus has *constant sign* (positive, say). Take $h(t)$ to be a 'bump function' (as shown in Figure 13.3), which is zero outside the interval (c, d) and positive inside. For such a choice of h,

$$\int_a^b G(t)\, h(t)\, dt = \int_c^d G(t)\, h(t)\, dt > 0,$$

since the integral of a positive function must be positive. This contradicts equation (13.16) and so G must be the zero function.

We have therefore shown that

$$\frac{\partial F}{\partial x}(x^*, \dot{x}^*, t) - \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}}(x^*, \dot{x}^*, t)\right) = 0$$

for $a < t < b$, which is the same as saying that x^* must satisfy the **Euler–Lagrange** differential equation

$$\frac{\partial F}{\partial x} - \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}}\right) = 0.$$

The above argument is reversible and so the converse result is also true. Our result is summarised as follows:

Euler–Lagrange equation

If the function x^* makes the integral functional

$$J[x] = \int_a^b F(x, \dot{x}, t)\, dt, \tag{13.17}$$

stationary, then x^* must satisfy the **Euler–Lagrange** differential equation

$$\frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}}\right) - \frac{\partial F}{\partial x} = 0. \tag{13.18}$$

The converse result is also true.

It is very convenient to give solutions of the Euler–Lagrange equation a special name.

Definition 13.2 ***Extremals*** *Any solution of the Euler–Lagrange equation is called an* ***extremal**** *of the functional* $J[x]$.

Key results of the calculus of variations

- If the function x^* minimises or maximises the functional J, then x^* makes J stationary and so x^* must be an extremal of J.
- If x^* is an extremal of J, then x^* makes J stationary, but it may not minimise or maximise J.

* The term *extremal* should not be confused with *extremum* (plural: *extrema*). Extremum means maximum or minimum. Thus a function that provides an extremum of J must make J stationary and so must be an extremal of J (that is, it must satisfy the Euler–Lagrange equation). The converse is not true. An extremal may not provide a mimimum or maximum of J. (Try reading this again!)

Fortunately, solving problems with the E–L equation is much easier than the preceding theory! The E–L equation is a second order *non-linear* ODE, and so one needs a measure of luck when it comes to finding solutions. Nevertheless, many interesting cases can be solved in closed form.

Example 13.1 *Finding extremals 1*

Find the extremal of the functional

$$J[x] = \int_1^2 \frac{\dot{x}^2}{4t}\, dt$$

that satisfies the end conditions $x(1) = 5$ and $x(2) = 11$.

Solution

By definition, extremals are solutions of the E–L equation. In the present case, $F = \dot{x}^2/4t$ so that

$$\frac{\partial F}{\partial x} = 0, \qquad \frac{\partial F}{\partial \dot{x}} = \frac{\dot{x}}{2t},$$

and the E–L equation takes the form

$$\frac{d}{dt}\left(\frac{\dot{x}}{2t}\right) - 0 = 0.$$

On integrating, we obtain

$$x = ct^2 + d,$$

where c and d are constants of integration. The **extremals** of J are therefore a family of parabolas in the (t, x)-plane. The *admissible* extremals are those that satisify the prescribed end conditions $x(1) = 5$ and $x(2) = 11$. On applying these conditions, we find that $c = 2$ and $d = 3$ so that the only **admissible extremal** of $J[x]$ is given by

$$\widehat{x} = 2t^2 + 3. \blacksquare$$

Question *Maximum, minimum, or neither?*

Does the extremal $\widehat{x} = 2t^2 + 3$ maximise or minimise J?

Answer

The admissible extremal $\widehat{x}$ is known to make J *stationary*. It may minimise J, maximise J, or do neither. With the theory that we have at our disposal, we cannot generally decide what happens. However, in a few simple cases (including this one), we can decide very easily.

Let h be *any* admissible variation (not necessarily small) and consider the variation in J that it produces, namely,

$$\begin{aligned}
J[\widehat{x}+h]-J[\widehat{x}] &= \int_1^2 \frac{(4t+\dot{h})^2}{4t}\,dt - \int_1^2 \frac{(4t)^2}{4t}\,dt \\
&= \int_1^2 \left(4t+2\dot{h}+\frac{\dot{h}^2}{4t}\right) dt - \int_1^2 4t\,dt \\
&= 2\Big[\,h\,\Big]_{t=1}^{t=2} + \int_1^2 \frac{\dot{h}^2}{4t}\,dt \\
&= \int_1^2 \frac{\dot{h}^2}{4t}\,dt
\end{aligned}$$

since h is an admissible extremal satisifying $h(1) = h(2) = 0$. Hence

$$J[\widehat{x}+h]-J[\widehat{x}] = \int_1^2 \frac{\dot{h}^2}{4t}\,dt \geq 0,$$

since the integral of a positive function must be positive. Thus $\widehat{x}$ actually provides the **global minimum** of $J[x]$. The global minimum *value* of J is therefore $J[2t^2+3] = 6$. ■

A useful integral of the Euler–Lagrange equation

Not all examples are as easy as the last one and the E–L equation often has a complicated form. However, for the case in which the function $F(x, \dot{x}, t)$ has no *explicit* dependence on t (that is, $F = F(x, \dot{x})$), the second order E–L equation can always be integrated once to yield a first order ODE. This offers a great simplification in many important problems.

Suppose that $F = F(x, \dot{x})$. Then it follows from the product rule and the chain rule that

$$\begin{aligned}
\frac{d}{dt}\left(\dot{x}\frac{\partial F}{\partial \dot{x}} - F\right) &= \ddot{x}\frac{\partial F}{\partial \dot{x}} + \dot{x}\frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}}\right) - \left(\ddot{x}\frac{\partial F}{\partial \dot{x}} + \dot{x}\frac{\partial F}{\partial x}\right) \\
&= \dot{x}\left(\frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}}\right) - \frac{\partial F}{\partial x}\right).
\end{aligned} \tag{13.19}$$

Thus, if x satisfies the E–L equation

$$\frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}}\right) - \frac{\partial F}{\partial x} = 0,$$

it follows that x satisfies the first order equation

$$\dot{x}\frac{\partial F}{\partial \dot{x}} - F = \text{constant}. \tag{13.20}$$

for some choice of the constant c. Conversely, if x is any *non-constant* solution of equation (13.20) then it satisfies the E–L equation.

It should be noted that equation (13.20) always has solutions of the form x = constant, but these solutions usually do *not* satisfy the corresponding E–L equation. They occur because of the factor $\dot{x}$ that appears on the right in equation (13.19). When overlooked, this glitch can give baffling results. *Do not believe that constant solutions of equation (13.20) satisfy the E–L equation unless you have checked it directly!*

Our result is summarised as follows:

A first integral of the E–L equation

Suppose that $F = F(x, \dot{x})$. Then any function that satisfies the E–L equation

$$\frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}}\right) - \frac{\partial F}{\partial x} = 0$$

also satisfies the first order differential equation

$$\dot{x}\frac{\partial F}{\partial \dot{x}} - F = c, \tag{13.21}$$

for some value of the constant c. Conversely, any *non-constant* solution of equation (13.21) satisfies the E–L equation. Constant solutions of equation (13.21) may or may not satisfy the E–L equation.

Example 13.2 *Finding extremals 2*

Find the extremal of the functional

$$J[x] = \int_0^7 \frac{(1+\dot{x}^2)^{1/2}}{x}\, dt$$

that lies in $x > 0$ and satisfies the end conditions $x(0) = 4$ and $x(7) = 3$. [The restriction $x > 0$ ensures that the integrand does not become singular.]

Solution

By definition, extremals are solutions of the E–L equation and, since t is not explicitly present in this functional, we can use the integrated form (13.21). On substituting $F = (1+\dot{x}^2)^{1/2}/x$ into (13.21) and simplifying, we obtain

$$x(1+\dot{x}^2)^{1/2} = C,$$

where C is a constant; since x is assumed positive, C must be positive. This equation can be rearranged in the form

$$\dot{x} = \pm\frac{(C^2 - x^2)^{1/2}}{x},$$

a pair of first order separable ODEs.

The solutions are*

$$\pm\left(C^2 - x^2\right)^{1/2} = t + D,$$

where D is a constant of integration. Hence the **extremals** of J are (the upper halves of) the family of circles

$$x^2 + (t+D)^2 = C^2$$

in the (t, x)-plane. On applying the given end conditions, we find that $C = 5$ and $D = -3$, so that the only **admissible extremal** is an arc of the circle with centre $(3, 0)$ and radius 5, namely

$$\widehat{x} = +\sqrt{16 + 6t - t^2} \qquad (0 \le t \le 7).$$

Since there is only one admissible extremal, it follows that, if it were *known* that a minimising (or maximising) function existed, then this must be it. However, we have no such knowledge and no means of deciding whether $\widehat{x}$ provides a minimum or maximum for J, or neither. (It actually provides the global minimum of J.) ■

Our final example is the famous brachistochrone problem.†

Example 13.3 *The brachistochrone (shortest time) problem*

Two fixed points P and Q are connected by a smooth wire lying in the vertical plane that contains P and Q. A particle is released from rest at P and slides, under uniform gravity, along the wire to Q. What shape should the wire be so that the transfer is completed in the shortest time?

Solution

Suppose that the wire lies in (x, z)-plane, with Oz pointing vertically *downwards*, with P at the origin, and Q at the point (a, b). Let the shape of the wire be given by the curve $z = z(x)$. Then, since the particle is released from rest when $z = 0$, energy conservation implies that the speed of the particle when its downward displacement is z is $(2gz)^{1/2}$. The time T taken for the particle to complete the transfer is therefore

$$T[z] = (2g)^{-1/2} \int_0^a \frac{\left\{1 + \dot{z}^2\right\}^{1/2}}{z^{1/2}}\, dx, \qquad (13.22)$$

* It is evident that these equations also admit the constant solution $x = C$. However, it may be verified that the E–L equation for this problem has no constant solutions.

† This famous minimisation problem was posed in 1696 by Johann Bernoulli (who had already found the solution) as a not-so-friendly challenge to his mathematical contemporaries. Solutions were found by Jacob Bernoulli, de l'Hôpital, Leibnitz, and Newton, who (according to his publicity manager) had the answer within a day. Newton published his solution anonymously, but Johann Bernoulli identified Newton as the author declaring '*one can recognise the lion by the marks of his claw*'. The last word goes to Newton. He complained *'I do not love to be pestered and teased by foreigners about mathematical things ...'*.

where $\dot{z}$ means dz/dx. The problem is to find the function $z(x)$, satisfying the end conditions $z(0) = 0$, $z(a) = b$, that minimises T.

If x^* minimises T, then it must make T stationary and so be an extremal of T. Since x is not explicitly present in this functional, we can use the integrated form (13.21) of the E–L equation. On substituting in $F = (1+\dot{z}^2)^{1/2}/z^{1/2}$ and simplifying, we obtain

$$z\left(1 + \dot{z}^2\right) = 2C,$$

where C is a positive constant. (The constant is called $2C$ for later convenience.) This equation can be arranged in the form

$$\dot{z} = \pm\left(\frac{2C - z}{z}\right)^{1/2},$$

a pair of first order separable ODEs. (The constant solution $z = 2C$ is not an extremal of J and can be disregarded.) Integration gives

$$x = \pm\int\left(\frac{z}{2C - z}\right)^{1/2} dz.$$

To perform the integral, we make the substitution* $z = C(1 - \cos\psi)$, in which case

$$\begin{aligned} x &= \pm C\int\left(\frac{1 - \cos\psi}{1 + \cos\psi}\right)^{1/2} \sin\psi\, d\psi \\ &= \pm C\int 2\sin^2 \tfrac{1}{2}\psi\, d\psi \\ &= \pm C(\psi - \sin\psi) + D, \end{aligned}$$

where D is a constant of integration. Thus the **extremals** of J have the *parametric* form

$$x = \pm C(\psi - \sin\psi) + D, \qquad z = C(1 - \cos\psi),$$

with ψ as parameter. Since the two choices of sign correspond only to a change of sign of the parameter, we may assume the positive choice. These curves are a family of cycloids† with 'radius' C and shift D in the x-direction.

Now we find the **admissible extremals**. The condition that $z = 0$ when $x = 0$ implies that the shift constant $D = 0$. The radius C of the cycloid is then determined from the second end condition $z = b$ when $x = a$. In general, C must be determined numerically but, in special cases, C may be found analytically. For example, if Q is the point $(a, 0)$ (so that P and Q are on the same horizontal level) it is found that $C = a/2\pi$. A more typical case is shown in Figure 13.4.

---

* It's easy to spot the smart substitution when you already know the answer!

† The cycloid is the path traced out by a point on the rim of a disk rolling on a plane; the 'radius' referred to is the radius of this disk. Since the E–L equation cannot be satisfied at a cusp, each extremal must lie on a single loop of the cycloid.

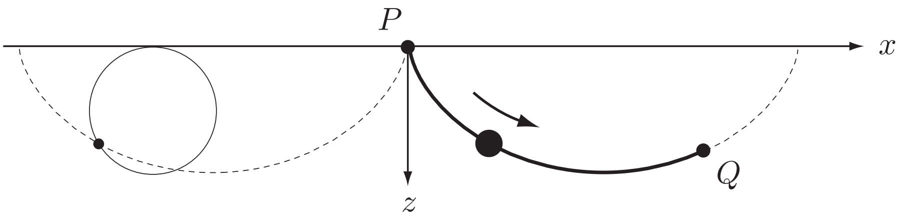


FIGURE 13.4 The curve that minimises $T[z]$ is an arc of a cycloid.

Since there is only one admissible extremal, it must be the **minimising curve** for $T[z]$, *provided that a minimising curve exists at all*. In view of the physical origin of the problem, most of us are happy to take this for granted, but purists may sleep more soundly in the knowledge that this can also be proved mathematically. ∎

13.3 VARIATIONAL PRINCIPLES

The laws of physics are usually formulated in terms of variables or fields that satisfy differential equations. Thus, the generalised coordinates $\boldsymbol{q}(t)$ of a mechanical system satisfy Lagrange's equations (a system of ODEs), the electromagnetic field satisfies Maxwell's equations (a system of PDEs), the wave function of quantum mechanics satisfies Schrödinger's wave equation, and so on. But there is an alternative way of expressing these laws in terms of **variational principles**. *In the variational approach, the actual physical behaviour of the system is distinguished by the fact that it makes a certain integral functional stationary*. Thus all of the physics is somehow contained in the integrand of this functional!

Expressing physical laws in variational form does not make it any easier to solve problems. Indeed, problems will continue to be solved by using differential equations. The virtue of the variational formulation is that it is much easier to extend existing theory to new situations. For example, the theory of fields can be developed in the language of classical mechanics by using the variational formulation.

Fermat's principle

These ideas are nicely illustrated by an example most readers will be familiar with: the paths of **light rays** in geometrical optics. When travelling through a **homogeneous medium**, light rays travel in straight lines. But, on meeting a plane interface between two different homogeneous media, the ray is either reflected or suffers a sharp change of direction called refraction, as shown in Figure 13.5. This change of direction is governed by **Snell's law** of refraction $n_1 \sin\theta_1 = n_2 \sin\theta_2$, where n_1 and n_2 are the refractive indices of the two media, and θ_1 and θ_2 are the angles that the ray makes with the *normal* to the interface. In terms of the angle ψ used in Figure 13.5, Snell's law takes the form

$$n_1 \cos\psi_1 = n_2 \cos\psi_2 = n_3 \cos\psi_3.$$

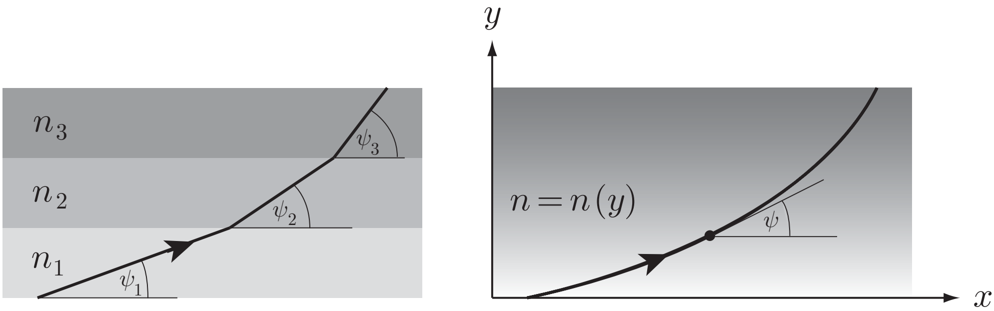


FIGURE 13.5 When a light ray passes between homogeneous media (left), it satisfies Snell's law $n_1 \cos\psi_1 = n_2 \cos\psi_2 = n_3 \cos\psi_3$. In the continuous case with $n = n(y)$ (right), Snell's law becomes $n \cos\psi = \text{constant}$.

In the more general case of an **inhomogeneous medium** in which n varies continuously in the y-direction, one would then expect curved rays that satisfy Snell's law in the form

$$n \cos\psi = \text{constant}.$$

A variational principle consistent with these rules was proposed by Fermat in 1657 and became known as **Fermat's principle** of least time. This stated that:

Of all the possible paths that a light ray might take between two fixed points, the actual path is the one that minimises the travel time of the ray.

Fermat showed that his principle implied the truth of the laws of reflection and refraction (as well as predicting straight rays in a homogeneous medium). Fermat's original principle is a beautifully simple and general statement about the paths taken by light rays but, sadly, it is not quite correct. The correct version is as follows:

Fermat's principle

The actual path taken by a light ray between two fixed points makes the travel time of the ray **stationary**.

The difference between the original and correct versions is that the path taken by the ray does not neccessarily make the travel time a minimum, but it does make the travel time stationary.* In practice, the travel time is usually a minimum, but there are exceptional cases where it is not.

If the free-space speed of light is c, then the speed of light at a point of a medium where the refractive index is n is c/n. A (hypothetical) path $\mathcal{P}$ in the medium would

* Surprisingly, incorrect statements about Fermat's principle abound in the literature. It is often claimed that '*the path of a light ray makes the travel time a minimum or (occasionally) a maximum*'. This is untrue. The path of a ray can *never* make T a maximum. It usually makes T a minimum, occasionally it provides neither a minimum nor a maximum, but it *never* provides a maximum.

therefore be traversed in time T given by the line integral

$$T[\mathcal{P}] = c^{-1} \int_{\mathcal{P}} n\, ds. \tag{13.23}$$

Since paths that make T stationary are extremals of T we can restate Fermat's pinciple in the elegant form:

Fermat's principle – Classy version

The paths of light rays in a medium are the same as the **extremals** of the functional T for that medium.

Suppose that the refractive index n in the medium depends only on y (as in Figure 13.5) and consider rays that lie in the (x, y)-plane. Then a ray that connects the points (x_0, y_0) and (x_1, y_1) must be an extremal of the functional T which, in Cartesian coordinates, takes the form

$$T[y] = c^{-1} \int_{x_0}^{x_1} n \left(1 + \dot{y}^2\right)^{1/2} dx, \tag{13.24}$$

where $\dot{y}$ means dy/dx, and $n = n(y)$. Since n does not depend upon x, we may use the integrated form (13.21) of the E–L equation, which gives

$$\frac{n}{(1 + \dot{y}^2)^{1/2}} = \text{constant}.$$

If we write $\dot{y} = \tan\psi$, where ψ is the angle between the tangent to the ray and the x-axis (see Figure 13.5), then this equation becomes

$$n \cos\psi = \text{constant},$$

exactly as anticipated from Snell's law for layered media.

Question ***A puzzle***

When $n = n(y)$, it is easy to verify that the straight lines y = constant are *not* extremals of $T[y]$ and are therefore *not* rays (although they do satisfy Snell's law!). But since such a 'ray' would experience a constant value of n, how does the ray know that it must bend?

Answer

Your physics lecturer will be pleased to answer this question. ■

The variational approach really comes into its own however when we extend our theory to other inhomogeneous media, where the correct generalisation of Snell's law is difficult to spot. There is no such difficulty with the variational approach; Fermat's principle still holds. The case of a light ray propagating in an axially symmetric medium is solved in Problem 13.9. From Fermat's principle, this is quite straightforward, but, starting from Snell's law, one would probably guess the wrong formula!

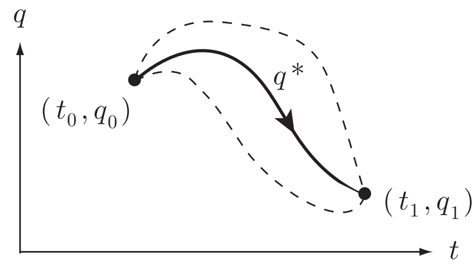


FIGURE 13.6 Hamilton's principle Of all the kinematically possible trajectories of a system that connect the configurations $q = q_1$ and $q = q_2$ in the time interval $[t_1, t_2]$, the actual motion $q^*(t)$ makes the action functional of the system stationary.

13.4 HAMILTON'S PRINCIPLE

Hamilton's principle is the variational principle that is equivalent to Lagrange's equations of motion. The comparison with geometrical optics is that Hamilton's principle corresponds to Lagrange's equations as Fermat's principle corresponds to Snell's law. We consider first the special case of systems with one degree of freedom.

Systems with one degree of freedom

Consider a Lagrangian system with a single generalised coordinate q and Lagrangian $L(q, \dot{q}, t)$. Then the trajectory $q^*(t)$ is an actual motion of the system if, and only if, it satisfies the Lagrange equation

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} = 0. \qquad (13.25)$$

It is impossible not to notice that equation (13.25) is the **Euler–Lagrange** equation one would get by making stationary the functional $S[q]$ defined by

$$S[q] = \int_{t_0}^{t_1} L(q, \dot{q}, t)\, dt. \qquad (13.26)$$

The scalar quantity S is called the **action** and the functional $S[q]$ is called the **action functional** corresponding to the Lagrangian L (for the time interval $[t_0, t_1]$).

From this simple observation, it follows that $q^*(t)$ *is an actual motion of the system if, and only if, it makes the action functional* $S[q]$ *stationary.* The situation is as shown in Figure 13.6. This is **Hamilton's principle** for a mechanical system with one degree of freedom.

Example 13.4 *Hamilton's principle*

A certain oscillator with generalised coordinate q has Lagrangian

$$L = \tfrac{1}{2}\dot{q}^2 - \tfrac{1}{2}q^2.$$

Verify that $q^* = \sin t$ is a motion of the oscillator, and show directly that it makes the action functional $S[q]$ satationary in any time interval $[0, \tau]$.

Solution

Lagrange's equation corresponding to the Lagrangian $L = \frac{1}{2}\dot{q}^2 - \frac{1}{2}q^2$ is

$$\ddot{q} + q = 0.$$

Since q^* satisfies this equation, it is a motion of the oscillator.

Let $h(t)$ be an admissible variation. Then

$$\begin{aligned} S[q^* + h] - S[q^*] &= \tfrac{1}{2}\int_0^\tau \left[\left(\cos t + \dot{h}\right)^2 - (\sin t + h)^2 - \cos^2 t + \sin^2 t\right] dt \\ &= \tfrac{1}{2}\int_0^\tau \left(2\dot{h}\cos t + \dot{h}^2 - 2h\sin t - h^2\right) dt \\ &= \Big[h\cos t\Big]_0^\tau + \tfrac{1}{2}\int_0^\tau \left(\dot{h}^2 - h^2\right) dt \\ &= \tfrac{1}{2}\int_0^\tau \left(\dot{h}^2 - h^2\right) dt, \end{aligned}$$

since $h(0) = h(\tau) = 0$. It follows that

$$\begin{aligned} |\, S[q^* + h] - S[q^*]\,| &\le \tfrac{1}{2}\tau \left(\max_{0\le t\le\tau} |h(t)|^2 + \max_{0\le t\le\tau} |\dot{h}(t)|^2\right) \\ &\le \tfrac{1}{2}\tau \left(\max_{0\le t\le\tau} |h(t)| + \max_{0\le t\le\tau} |\dot{h}(t)|\right)^2 \\ &= \tfrac{1}{2}\tau\, ||h||^2. \end{aligned}$$

Hence

$$S[q^* + h] - S[q^*] = O\left(||h||^2\right),$$

which, by definition, means that q^* makes the action functional $S[q]$ **stationary**. [It may *not* make $S[q]$ a minimum.] ■

Systems with many degrees of freedom

Hamilton's principle can be extended to systems with any number of degrees of freedom. In this more general case, the system has generalised coordinates $\boldsymbol{q} = (q_1, q_2, \ldots, q_n)$, the Lagrangian has the form $L = L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$, and Lagrange's equations of motion are the *n simultaneous equations*

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0 \qquad (1 \le j \le n). \tag{13.27}$$

The action functional is now defined to be:

Definition 13.3 ***Action functional*** *The functional*

$$S[\boldsymbol{q}] = \int_{t_0}^{t_1} L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)\, dt \tag{13.28}$$

is called the ***action functional*** *corresponding to the Lagrangian $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$ (for the time interval $[t_0, t_1]$).*

The notation $S[\boldsymbol{q}]$ is really a shorthand form for $S[q_1, q_2, \ldots, q_n]$ so that now there are n functions that can be varied by the n independent variations $h_1, h_2, \ldots, h_n$ respectively. In the vector notation, such a variation is denoted by $\boldsymbol{h}$, where $\boldsymbol{h} = (h_1, h_2, \ldots, h_n)$. The theory that we have developed does not cover the case where the functional has more than one 'independent variable', but it can be extended to do so. An outline of this extension is as follows:

Consider the general situation in which

$$J[\boldsymbol{x}] = \int_a^b F(\boldsymbol{x}, \dot{\boldsymbol{x}}, t)\, dt,$$

where the vector function $\boldsymbol{x}(t) = (x_1(t), x_2(t), \ldots, x_n(t))$. By using the same argument as before, the variation in J caused by the admissible* variation $\boldsymbol{h}$ in $\boldsymbol{x}^*$ is found to be

$$J[\boldsymbol{x}^* + \boldsymbol{h}] - J[\boldsymbol{x}^*] = \sum_{j=1}^{n} \int_a^b \left[\frac{\partial F}{\partial x_j}(\boldsymbol{x}^*, \dot{\boldsymbol{x}}^*, t) - \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}_j}(\boldsymbol{x}^*, \dot{\boldsymbol{x}}^*, t) \right) \right] h_j\, dt + O\left(||\boldsymbol{h}||^2 \right),$$

where $||\boldsymbol{h}||^2 = ||h_1||^2 + ||h_2||^2 + \cdots + ||h_n||^2$. This variation is linear in $\boldsymbol{h}$ with an error term of order $||\boldsymbol{h}||^2$. As before we say that $\boldsymbol{x}^*$ makes $J[\boldsymbol{x}]$ stationary if the linear term is zero, leaving only the error term.

Definition 13.4 ***Stationary*** J *The vector function $\boldsymbol{x}^*(t)$ is said to make the functional $J[\boldsymbol{x}]$* ***stationary*** *if*

$$J[\boldsymbol{x}^* + \boldsymbol{h}] - J[\boldsymbol{x}^*] = O\left(||\boldsymbol{h}||^2 \right)$$

when $||\boldsymbol{h}||$ is small.

If $\boldsymbol{x}^*$ makes the functional $J[\boldsymbol{x}]$ stationary then

$$\sum_{j=1}^{n} \int_a^b \left[\frac{\partial F}{\partial x_j}(\boldsymbol{x}^*, \dot{\boldsymbol{x}}^*, t) - \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}_j}(\boldsymbol{x}^*, \dot{\boldsymbol{x}}^*, t) \right) \right] h_j\, dt = 0,$$

for all admissible variations $\boldsymbol{h}$. By allowing each of the $\{x_j\}$ to vary separately (while the others remain constant), the 'bump function' argument can be applied exactly as before to show that

$$\frac{\partial F}{\partial x_j}(\boldsymbol{x}^*, \dot{\boldsymbol{x}}^*, t) - \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}_j}(\boldsymbol{x}^*, \dot{\boldsymbol{x}}^*, t) \right) = 0 \qquad (1 \le j \le n).$$

This is the same as saying that $\boldsymbol{x}^*$ must satisfy the *simultaneous* **Euler–Lagrange** equations

$$\frac{\partial F}{\partial x_j} - \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}_j} \right) = 0 \qquad (1 \le j \le n).$$

Our result is summarised as follows:

* The vector variation $\boldsymbol{h}$ is admissible if $\boldsymbol{h}(a) = \boldsymbol{h}(b) = \boldsymbol{0}$, that is, if $h_1, h_2, \ldots, h_n$ are all admissible.

Euler–Lagrange equations with many variables

The vector function $\boldsymbol{x}^*$ makes the integral functional

$$J[\boldsymbol{x}] = \int_a^b F\left(\boldsymbol{x}, \dot{\boldsymbol{x}}, t\right)\, dt$$

stationary if, and only if, $\boldsymbol{x}^*$ satisfies the simultaneous **Euler–Lagrange** differential equations

$$\frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}_j}\right) - \frac{\partial F}{\partial x_j} = 0 \qquad (1 \le j \le n).$$

This is a natural generalisation of the single variable theory and corresponds to the elementary result that a function of n variables $f(x_1, x_2, \ldots, x_n)$ has a stationary point if, and only if, all its first partial derivatives vanish at that point.

The statement of Hamilton's principle for systems with many degrees of freedom is therefore:

Hamilton's principle

The trajectory $\boldsymbol{q}^*(t)$ is an actual motion of a mechanical system if, and only if, $\boldsymbol{q}^*$ makes the action functional of the system stationary.

The only essential difference between this correct version of Hamilton's principle and the original version (quoted at the beginning of the chapter) is that an actual motion of the system does not neccessarily make the action functional a *minimum*, but it always makes the action functional *stationary*.* In practice, the action functional is usually minimised, but there are exceptional cases where it is not (see Problem 13.11).

As with Fermat's principle, there is a classy version of Hamilton's principle, which is less wordy and more satisfactory generally. It makes use of the concept of the **extremals** of J, which are simply solutions of the n simultaneous Euler–Lagrange equations.

Hamilton's principle – Classy version

The actual motions of a mechanical system are the same as the extremals of its action functional.

---

* Incorrect statements about Hamilton's principle also abound in the literature. It is often claimed that '*an actual motion the system makes the action functional a minimum or (occasionally) a maximum*'. This is untrue. It is not possible to make S a maximum. The actual motion of the system usually makes S a minimum, occasionally it provides neither a minimum nor a maximum, but it *never* provides a maximum.

Significance of Hamilton's principle

Since Hamilton's principle is equivalent to Lagrange's equations, it can be regarded as the fundamental postulate of classical mechanics, instead of Newton's laws,* for any mechanical system that has a Lagrangian. It should be emphasised that this is *not* a new theory – the Newtonian theory is correct – but an alternative route to the same results. Thus we can derive Lagrange's equations of motion from the Newtonian theory (as we did) or, more directly, from Hamilton's principle. Because Hamilton's principle can be extended to apply to a wide range of physical phenomena while the Newtonian theory can not, Hamilton's principle is regarded as the more fundamental.

The problem with taking Hamilton's principle as the fundamental postulate of classical mechanics is that, had one not been exposed to the traditional treatment, one would have no idea what the Lagrangian ought to be for any particular system. To convince oneself of the difficulties involved, it is instructive to read Landau's [6] 'derivation' of the Lagrangian for the simplest system imaginable – a single particle moving in free space. Indeed, it seems difficult to introduce the concept of mass convincingly at all. Nevertheless, this is the route that must be followed when Hamilton's principle is extended, for example, to particle physics. The Lagrangian has to be found by intelligent guesswork, and, in particular, by taking account of all the symmetries that are known to exist.

Within classical mechanics itself, it may appear that Hamilton's principle has told us nothing new. It says that the motions of a mechanical system are the same as the extremals of the action functional, that is, the motions satisfy Lagrange's equations; this we already knew. However, because the equations of motion have the special form associated with variational principles, they can be shown to possess important properties that would be very difficult to prove directly. One example of this is the effect on the equations of motion of choosing a new set of generalised coordinates $\boldsymbol{q}' = (q_1', q_2', \ldots, q_n')$. The $\boldsymbol{q}'$ are known functions of the old generalised coordinates $\boldsymbol{q}$ and *vice versa*. The direct approach would be to subject the Lagrange equations (13.27) to this general transformation of the coordinates and see what happens; the result would be a complicated mess. However, in the variational approach, one simply expresses the Lagrangian L as a function of the new variables, that is, $L = L(\boldsymbol{q}', \dot{\boldsymbol{q}}', t)$. Although L has a different *functional form* in terms of the coordinates $\boldsymbol{q}'$, its *values* are the same as before, so that the new action functional

$$S[\boldsymbol{q}'] = \int_{t_0}^{t_1} L(\boldsymbol{q}', \dot{\boldsymbol{q}}', t)\, dt,$$

takes the *same values* as the old, provided that $\boldsymbol{q}'(t)$ and $\boldsymbol{q}(t)$ refer to the same trajectory of the mechanical system. It follows that, if the trajectory $\boldsymbol{q}(t)$ makes $S[\boldsymbol{q}]$ stationary, then the corresponding trajectory $\boldsymbol{q}'(t)$ makes $S[\boldsymbol{q}']$ stationary. Hence the extremals of $S[\boldsymbol{q}]$ map into the extremals of $S[\boldsymbol{q}']$, and *vice versa*. It follows that the transformed equations of motion are just the same as the old ones with $\boldsymbol{q}$ replaced by $\boldsymbol{q}'$. This fact is expressed

* Actually, instead of the Second and Third Laws. The First Law is needed to ensure that the motion is observed from an inertial reference frame.

by saying that the Lagrange equations of motion are **invariant under transformations** of the generalised coordinates. This remarkable result clearly applies to *any* system of equations derived from a variational principle. This provides a general way of ensuring that any proposed set of governing equations should be invariant under a particular group of transformations (the Lorentz transformations, for instance). This will be so if the equations are derivable from a variational principle whose 'Lagrangian' is invariant under the same group of transformations.

Problems on Chapter 13

Answers and comments are at the end of the book.

Harder problems carry a star (∗).

Euler–Lagrange equation

13.1 Find the extremal of the functional

$$J[x] = \int_1^2 \frac{\dot{x}^2}{t^3}\, dt$$

that satisfies $x(1) = 3$ and $x(2) = 18$. Show that this extremal provides the global minimum of J.

13.2 Find the extremal of the functional

$$J[y] = \int_0^\pi (2x \sin t - \dot{x}^2)\, dt$$

that satisfies $x(0) = x(\pi) = 0$. Show that this extremal provides the global maximum of J.

13.3 Find the extremal of the path length functional

$$L[y] = \int_0^1 \left[1 + \left(\frac{dy}{dx}\right)^2\right]^{1/2} dx$$

that satisfies $y(0) = y(1) = 0$ and show that it does provide the global minimum for L.

13.4 An aircraft flies in the (x, z)-plane from the point $(-a, 0)$ to the point $(a, 0)$. ($z = 0$ is ground level and the z-axis points vertically upwards.) The cost of flying the aircraft at height z is $\exp(-kz)$ per unit *distance* of flight, where k is a positive constant. Find the extremal for the problem of minimising the total cost of the journey. [Assume that $ka < \pi/2$.]

13.5∗ ***Geodesics on a cone*** Solve the problem of finding a shortest path over the surface of a cone of semi-angle α by the calculus of variations. Take the equation of the path in the form $\rho = \rho(\theta)$, where ρ is distance from the vertex O and θ is the cylindrical polar angle measured around the axis of the cone. Obtain the general expression for the path length and find the extremal that satisfies the end conditions $\rho(-\pi/2) = \rho(\pi/2) = a$.

Verify that this extremal is the same as the shortest path that would be obtained by developing the cone on to a plane.

13.6 ***Cost functional*** A manufacturer wishes to minimise the cost functional

$$C[x] = \int_0^4 \Big((3 + \dot{x})\dot{x} + 2x\Big)\, dt$$

subject to the conditions $x(0) = 0$ and $x(4) = X$, where X is volume of goods to be produced. Find the extremal of C that satisfies the given conditions and prove that this function provides the global minimum of C.

Why is this solution not applicable when $X < 8$?

13.7 ***Soap film problem*** Consider the soap film problem for which it is required to minimise

$$J[y] = \int_{-a}^{a} y\left(1 + \dot{y}^2\right)^{\frac{1}{2}}\, dx$$

with $y(-a) = y(a) = b$. Show that the extremals of J have the form

$$y = c\cosh\left(\frac{x}{c} + d\right),$$

where c, d are constants, and that the end conditions are satisfied if (and only if) $d = 0$ and

$$\cosh\lambda = \left(\frac{b}{a}\right)\lambda,$$

where $\lambda = a/c$. Show that there are *two* admissible extremals provided that the aspect ratio b/a exceeds a certain critical value and *none* if b/a is less than this cririical value. Sketch a graph showing how this critical value is determined.

The remainder of this question requires computer assistance. Show that the critical value of the aspect ratio b/a is about 1.51. Choose a value of b/a larger than the critical value ($b/a = 2$ is suitable) and find the two values of λ. Plot the two admissible extremals on the same graph. Which one looks like the actual shape of the soap film? Check your guess by perturbing each extremal by small admissible variations and finding the change in the value of the functional $J[y]$.

Fermat's principle

13.8 A sugar solution has a refractive index n that increases with the depth z according to the formula

$$n = n_0\left(1 + \frac{z}{a}\right)^{1/2},$$

where n_0 and a are positive constants. A particular ray is horizontal when it passes through the origin of coordinates. Show that the path of the ray is not the straight line $z = 0$ but the parabola $z = x^2/4a$.

13.9 Consider the propagation of light rays in an axially symmetric medium, where, in a system of cylindrical polar co-ordinates (r, θ, z), the refractive index $n = n(r)$ and the rays lie in the plane $z = 0$. Show that Fermat's time functional has the form

$$T[r] = c^{-1} \int_{\theta_0}^{\theta_1} n \left(r^2 + \dot{r}^2 \right)^{1/2} d\theta,$$

where $r = r(\theta)$ is the equation of the path, and $\dot{r}$ means $dr/d\theta$.

(i) Show that the extremals of T satisfy the ODE

$$\frac{n r^2}{(r^2 + \dot{r}^2)^{1/2}} = \text{constant}.$$

Show further that, if we write $\dot{r} = r \tan \psi$, where ψ is the angle between the tangent to the ray and the local cylindrical surface $r =$ constant, this equation becomes

$$r\, n \cos \psi = \text{constant},$$

which is the form of Snell's law for this case. Deduce that circular rays with centre at the origin exist only when the refractive index $n = a/r$, where a is a positive constant.

Hamilton's principle

13.10 A particle of mass 2 kg moves under uniform gravity along the z-axis, which points verically downwards. Show that (in SI units) the action functional for the time interval [0, 2] is

$$S[z] = \int_0^2 \left(\dot{z}^2 + 20z \right) dt,$$

where g has been taken to be $10\ \text{m}\,\text{s}^{-2}$.

Show directly that, of all the functions $z(t)$ that satisfy the end conditions $z(0) = 0$ and $z(2) = 20$, the actual motion $z = 5t^2$ provides the *least* value of S.

13.11 A certain oscillator with generalised coordinate q has Lagrangian

$$L = \dot{q}^2 - 4q^2.$$

Verify that $q^* = \sin 2t$ is a motion of the oscillator, and show directly that it makes the action functional $S[q]$ satationary in any time interval $[0, \tau]$.

For the time interval $0 \le t \le \pi$, find the variation in the action functional corresponding to the variations (i) $h = \epsilon \sin 4t$, (ii) $h = \epsilon \sin t$, where ϵ is a small parameter. Deduce that the motion $q^* = \sin 2t$ does not make S a minimum or a maximum.

13.12 A particle is constrained to move over a smooth fixed surface under *no forces* other than the force of constraint. By using Hamilton's principle and energy conservation, show that the

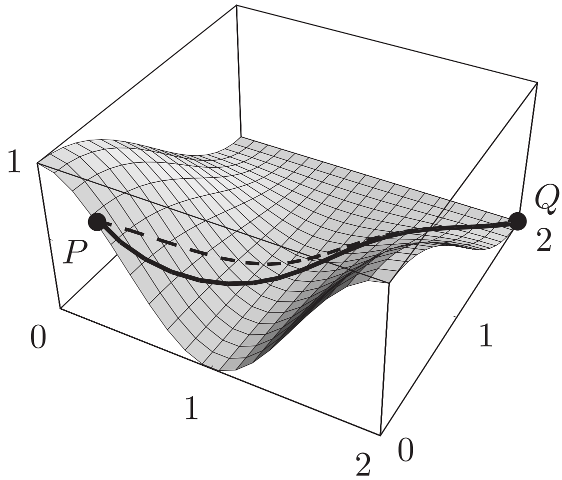


FIGURE 13.7 The path of quickest descent from P to Q in Cosine Valley. Those who lose their nerve at the summit can walk down by the *shortest* path (shown dashed).

path of the particle must be a geodesic of the surface. (The term geodesic has been extended here to mean those paths that make the length functional *stationary*).

This result has a counterpart in the theory of general relativity, where the concept of force does not exist and particles move along the geodesics of a curved space-time.

13.13 By using Hamilton's principle, show that, if the Lagrangian $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$ is modified to L' by any transformation of the form

$$L' = L + \frac{d}{dt}\, g(\boldsymbol{q}, t),$$

then the equations of motion are unchanged.

Computer assisted problems

13.14 ***Geodesics on a paraboloid*** Solve the problem of finding the shortest path between two points $P(0, 1, 1)$ and $Q(0, -1, 1)$ on the surface of the paraboloid $z = x^2 + y^2$.

Let $\mathcal{C}$ be a path lying in the surface that connects P and Q. Show that the length of $\mathcal{C}$ is given by

$$L[r] = \int_{-\pi/2}^{\pi/2} \left[r^2 + \left(1 + 4r^2\right) \dot{r}^2 \right]^{1/2} d\theta,$$

where $r = r(\theta)$ is the *polar equation* of the projection of $\mathcal{C}$ on to the plane $z = 0$, and $\dot{r}$ means $dr/d\theta$. Now find the function $r(\theta)$ that minimises L. It is easier to work directly with the second order E–L-equation (which can be found with computer assistance). Solve the E–L-equation numerically with the initial conditions $r(0) = \lambda$, $\dot{r}(0) = 0$ and choose λ so that the path passes through P (and, by symmetry, Q). Plot the shortest path using 3D graphics.

13.15* ***The downhill skier*** Solve the problem of finding the path of quickest descent for a skier from the point $P(x_0, y_0, z_0)$ to the point $Q(x_1, y_1, z_1)$ on a snow covered mountain whose profile is given by $z = G(x, y)$, where G is a known function. [Assume that the skier starts from rest and that the total energy of the skier is conserved in the descent.]

Let $\mathcal{C}$ be a path connecting P and Q. Show that the time taken to descend by this route is given by

$$T[y] = (2g)^{-1/2} \int_{x_0}^{x_1} \left(\frac{1 + \dot{y}^2 + \left(G,_x + G,_y \dot{y}\right)^2}{G(x_0, y_0) - G(x, y)} \right)^{1/2} dx$$

where $y = y(x)$ is the projection of $\mathcal{C}$ on to the plane $z = 0$, $\dot{y}$ means dy/dx, and $G,_x$, $G,_y$ are the partial derivatives of G with respect to x, y. Obtain the E–L equation with computer assistance and solve it numerically with the initial conditions $y(x_1) = y_1$, $y'(x_1) = \lambda$ and choose λ so that the path passes through the P. [The numerical ODE integrator finds it easier to integrate the equation starting from the bottom. Why?] Plot the quickest route using 3D graphics. The author used the profile of the Cosine Valley resort, for which $G(x, y) = \cos^2(\pi x/2)\cos^2(\pi y/4)$. The skier had to descend from $P(1/3, 0, 3/4)$ to $Q(2, 2, 0)$. The computued quickest route down the valley is shown in Figure 13.7. Those who lose their nerve at the summit can walk down by the *shortest* route (shown dashed). You may make up your own mountain profile, but keep it simple.

Chapter Fourteen

Hamilton's equations and phase space

KEY FEATURES

The key features of this chapter are the equivalence of Lagrange's equations and **Hamilton's equations**, Hamiltonian **phase space**, **Liouville's theorem** and **recurrence**.

In this chapter we show how Lagrange's equations can be reformulated as a set of *first order* differential equations known as **Hamilton's equations**. Nothing new is added to the physics and the Hamilton formulation is not superior to that of Lagrange when it comes to problem solving. The value of Hamilton's supremely elegant formulation lies in providing a foundation for theoretical extensions both within and outside classical mechanics. Within classical mechanics, it is the basis for most further developments such as the Hamilton-Jacobi theory and chaos. Elsewhere, Hamiltonian mechanics provides the best route to statistical mechanics, and the notion of the Hamiltonian is at the heart of quantum mechanics. As applications of the Hamiltonian formulation, we prove Liouville's theorem and Poincaré's recurrence theorem and explore some of their remarkable consequences.

14.1 SYSTEMS OF FIRST ORDER ODES

The standard form for a system of first order ODEs in the n unknown functions $x_1(t), x_2(t), \dots, x_n(t)$ is

$$\begin{aligned}
\dot{x}_1 &= F_1(x_1, x_2, \dots, x_n, t),\\
\dot{x}_2 &= F_2(x_1, x_2, \dots, x_n, t),\\
\vdots \quad & \qquad \vdots\\
\dot{x}_n &= F_n(x_1, x_2, \dots, x_n, t),
\end{aligned} \tag{14.1}$$

where $F_1, F_2, \dots, F_n$ are given functions of the variables $x_1, x_2, \dots, x_n, t$. This can also be written in the compact vector form

$$\dot{\boldsymbol{x}} = \boldsymbol{F}(\boldsymbol{x}, t), \tag{14.2}$$

where $\boldsymbol{x}$ and $\boldsymbol{F}$ are the n-dimensional vectors $\boldsymbol{x} = (x_1, x_2, \dots, x_n)$ and $\boldsymbol{F} = (F_1, F_2, \dots, F_n)$. If the value of $\boldsymbol{x}$ is given when $t = t_0$, the equations $\dot{\boldsymbol{x}} = \boldsymbol{F}(\boldsymbol{x}, t)$ determine the unknowns $\boldsymbol{x}(t)$ at all subsequent times.

A typical example is the **predator-prey** system of equations

$$\dot{x}_1 = ax_1 - bx_1x_2,$$
$$\dot{x}_2 = bx_1x_2 - cx_2,$$

which govern the population density $x_1(t)$ of the prey and the population density $x_2(t)$ of the predator. In this case, $F_1 = ax_1 - bx_1x_2$, and $F_2 = bx_1x_2 - cx_2$.

Converting higher order equations to first order

Higher order ODEs can always be converted into equivalent systems of first order ODEs. For example, consider the **damped oscillator** equation

$$\ddot{x} + 3\dot{x} + 4x = 0. \tag{14.3}$$

If we introduce the new variable v, defined by $v = \dot{x}$, then the second order equation (14.3) for $x(t)$ can be converted into the pair of first order equations

$$\dot{x} = v$$
$$\dot{v} = -3x - 4v,$$

in the unknowns $\{x(t), v(t)\}$. Since this step is reversible, this pair of first order equations is *equivalent* to the original second order equation (14.3). More generally:

Any system of n second order ODEs in n unknowns can be converted into an equivalent system of $2n$ first order ODEs in $2n$ unknowns.

Consider, for example, the **orbit equations** for a particle of mass m attracted by the gravitation of a mass M fixed at the origin O. In terms of polar coordinates centred on O, the Lagrangian is

$$L = \tfrac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right) + \frac{mMG}{r}.$$

and the corresponding Lagrange equations are

$$\ddot{r} - r\dot{\theta}^2 = -\frac{MG}{r^2}, \qquad 2r\dot{r}\dot{\theta} + r^2\ddot{\theta} = 0,$$

a pair of second order ODEs in the unknowns $\{r(t), \theta(t)\}$. If we now introduce the new variables v_r and v_θ, defined by

$$v_r = \dot{r}, \qquad v_\theta = \dot{\theta},$$

the *two second order* Lagrange equations can be converted into

$$\dot{r} = v_r, \qquad \dot{v}_r = rv_\theta^2 - MG/r^2,$$
$$\dot{\theta} = v_\theta, \qquad \dot{v}_\theta = -2v_r v_\theta/r,$$

an equivalent system of *four first order* ODEs in the four unknowns $\{r, \theta, v_r, v_\theta\}$.

Hamilton form

In the above examples, we have performed the conversion to first order equations by introducing the **coordinate velocities** as new variables. This is not the only choice however. When transforming a system of Lagrange equations to a first order system, we could instead take the **congugate momenta**, defined by

$$p_j = \frac{\partial L}{\partial \dot{q}_j}, \tag{14.4}$$

as the new variables. This seems quite attractive since the Lagrange equations already have the form

$$\dot{p}_j = \frac{\partial L}{\partial q_j} \qquad (1 \le j \le n).$$

The downside is that the right sides $\partial L/\partial q_j$ are functions of $\boldsymbol{q}$, $\dot{\boldsymbol{q}}$, t, and must be transformed to functions of $\boldsymbol{q}$, $\boldsymbol{p}$, t. This is achieved by inverting the equations (14.4) to express the $\dot{\boldsymbol{q}}$ as functions of $\boldsymbol{q}$, $\boldsymbol{p}$, t.

For the Lagrangian in the orbit problem, the conjugate momenta are

$$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r}, \qquad p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta},$$

and these equations are easily inverted* to give

$$\dot{r} = \frac{p_r}{m}, \qquad \dot{\theta} = \frac{p_\theta}{mr^2}.$$

The *two second order* Lagrange equations for the orbit problem are therefore equivalent to the system of *four first order* ODEs

$$\dot{r} = \frac{p_r}{m}, \qquad \dot{p}_r = \frac{p_\theta^2}{mr^3} - \frac{mMG}{r^2},$$

$$\dot{\theta} = \frac{p_\theta}{mr^2}, \qquad \dot{p}_\theta = 0,$$

in the four unknowns $\{r, \theta, p_r, p_\theta\}$. This is the **Hamilton form** of Lagrange's equations for the orbit problem.

Why bother?

The reader is probably wondering what is the point of converting Lagrange's equations into a system of first order ODEs. For the purpose of finding solutions to particular problems, like the orbit problem above, there is no point. Indeed, the new system of first order

* This step is difficult in the general case where there are *n coupled* linear equations to solve for $\dot{q}_1, \ldots, \dot{q}_n$.

equations may be harder to solve than the original second order equations. The real interest lies in the structure of the general theory. When Lagrange's equations are expressed in Hamilton form* *in a general manner*, the result is a system of first order equations of great simplicity and elegance, now known as **Hamilton's equations**. These equations are the foundation of further developments in analytical mechanics, such as the Hamilton-Jacobi theory and chaos. Also, the **Hamiltonian function**, which appears in Hamilton's equations, is at the heart of quantum mechanics.

14.2 LEGENDRE TRANSFORMS

The general problem of converting Lagrange's equations into Hamilton form hinges on the inversion of the equations that define $\boldsymbol{p}$, namely,

$$p_j = \frac{\partial}{\partial \dot{q}_j} L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) \qquad (1 \le j \le n), \tag{14.5}$$

so as to express $\dot{\boldsymbol{q}}$ in terms of $\boldsymbol{q}$, $\boldsymbol{p}$, t. This inversion is made easier by the fact that the $\{p_j\}$ are not general functions of $\boldsymbol{q}$, $\dot{\boldsymbol{q}}$ and t, but are the *first partial derivatives of a scalar function*, the Lagrangian $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$. It is a remarkable consequence that the inverse formulae can be written in a similar way.† The details of the argument follow below and the results are summarised at the end of the section.

The two-variable case

We develop the transformation theory for the case of functions of two variables. This has all the important features of the general case but is much easier to follow. Suppose that v_1 and v_2 are defined as functions of the variables u_1 and u_2 by the formulae

$$v_1 = \frac{\partial F}{\partial u_1}, \qquad v_2 = \frac{\partial F}{\partial u_2}, \tag{14.6}$$

where $F(u_1, u_2)$ is a given function of u_1 and u_2. Is it possible to write the inverse formulae‡ in the form

$$u_1 = \frac{\partial G}{\partial v_1}, \qquad u_2 = \frac{\partial G}{\partial v_2}, \tag{14.7}$$

---

* This is the form in which the new variables are the conjugate momenta $p_1, \ldots, p_n$. The form in which the new variables are the generalised velocities $\dot{q}_1, \ldots, \dot{q}_n$ does not lead to an elegant theory, and is therefore not used. It is often claimed that it is not *possible* to take the generalised velocities as new variables because 'they are the time derivatives of the generalised coordinates and therefore cannot be independent variables'. This objection is baseless, as the previous examples show. Indeed, if this objection had any substance, the conjugate momenta would be disqualified as well!

† There is a neat way of seeing that this must be true, which may appeal to mathematicians. If $\boldsymbol{v} = \text{grad}_{\boldsymbol{u}} F(\boldsymbol{u})$, then the Jacobian matrix of the transformation from $\boldsymbol{u}$ to $\boldsymbol{v}$ is symmetric. It follows that the Jacobian matrix of the inverse tranformation must also be symmetric, which is precisely the condition that the inverse transformation has the form $\boldsymbol{u} = \text{grad}_{\boldsymbol{v}} G(\boldsymbol{v})$.

‡ We will always suppose that the inverse transformation does exist.

for some function $G(v_1, v_2)$? In the simplest cases one can answer the question by grinding through the details directly. For example, suppose $F = 2u_1^2 + 3u_1u_2 + u_2^2$. Then

$$v_1 = 4u_1 + 3u_2,$$
$$v_2 = 3u_1 + 2u_2.$$

The inverse formulae are easily obtained by solving these equations for u_1, u_2, which gives

$$u_1 = -2v_1 + 3v_2,$$
$$u_2 = 3v_1 - 4v_2.$$

There is no prior reason to expect that these formulae for u_1, u_2 can be expressed in terms of a single function $G(v_1, v_2)$ in the form (14.7), but it *is* true because the right sides of these equations happen to satisfy the necessary **consistency condition**.* Simple integration then shows that (to within a constant)

$$G = -v_1^2 + 3v_1v_2 - 2v_2^2.$$

This result is not a coincidence. Let $F(u_1, u_2)$ now be *any* function of the variables u_1, u_2, and suppose that a function $G(v_1, v_2)$ satisfying equation (14.7) *does* exist. Consider the expression

$$X = F(u_1, u_2) + G(v_1, v_2) - (u_1v_1 + u_2v_2),$$

which, as it stands, is a function of the four independent variables u_1, u_2, v_1, v_2. Suppose now that, in this formula, we imagine† that v_1 and v_2 are replaced by their expressions in terms of u_1 and u_2. Then X becomes a function of the variables u_1 and u_2 only. Its partial derivative with respect to u_1, holding u_2 constant, is then given by

$$\begin{aligned}\frac{\partial X}{\partial u_1} &= \frac{\partial F}{\partial u_1} + \left(\frac{\partial G}{\partial v_1} \times \frac{\partial v_1}{\partial u_1} + \frac{\partial G}{\partial v_2} \times \frac{\partial v_2}{\partial u_1}\right) - \left(v_1 + u_1\frac{\partial v_1}{\partial u_1} + u_2\frac{\partial v_2}{\partial u_1}\right)\\ &= \left(\frac{\partial F}{\partial u_1} - v_1\right) + \left(\frac{\partial G}{\partial v_1} - u_1\right)\frac{\partial v_1}{\partial u_1} + \left(\frac{\partial G}{\partial v_2} - u_2\right)\frac{\partial v_2}{\partial u_1}\\ &= 0 + 0 + 0 = 0,\end{aligned}$$

* If $u_1 = f_1(v_1, v_2)$ and $u_2 = f_2(v_1, v_2)$ then it is possible to express u_1, u_2 in the form (14.7) only if the functions f_1 and f_2 are related by the formula

$$\frac{\partial f_1}{\partial v_2} = \frac{\partial f_2}{\partial v_1},$$

which is called the **consistency condition**.

† Pure mathematicians strongly object to such feats of imagination. Unfortunately, the alternative is to introduce a welter of functional notation which obscures the essential simplicity of the argument. We will make frequent use of such 'imagined' substitutions.

on using first the chain rule and then the formulae (14.6) and (14.7). Hence X is independent of the variable u_1. In exactly the same way we may show that $\partial X/\partial u_2 = 0$ so that X is also independent of u_2. It follows that X must be a constant! This constant can be absorbed into the function G without disturbing the formulae (14.7), in which case $X = 0$. We have therefore shown that if F and G are related by the equations (14.6) and (14.7), then they must* satisfy the relation

$$F(u_1, u_2) + G(v_1, v_2) = u_1 v_1 + u_2 v_2. \tag{14.8}$$

The above argument is reversible so the converse result is also true. We have therefore shown that:

The required function $G(v_1, v_2)$ ***always exists*** *and can be generated from the function $F(u_1, u_2)$ by the formula*

$$\boxed{G(v_1, v_2) = (u_1 v_1 + u_2 v_2) - F(u_1, u_2)} \tag{14.9}$$

where u_1 and u_2 are to be replaced by their expressions in terms of v_1 and v_2.

It is evident that the relationship between the functions F and G is a symmetrical one. Each function is said to be the **Legendre transform** of the other.

Example 14.1 *Finding a Legendre transform*

Find the Legendre transform of the function $F(u_1, u_2) = 2u_1^2 + 3u_1u_2 + u_2^2$ by using the formula (14.9).

Solution

For this F, $v_1 = \partial F/\partial u_1 = 4u_1 + 3u_2$ and $v_2 = \partial F/\partial u_2 = 3u_1 + 2u_2$. The inverse formulae are $u_1 = -2v_1 + 3v_2$ and $u_2 = 3v_1 - 4v_2$. From equation (14.9), the function G is given by

$$\begin{aligned}
G &= u_1 v_1 + u_2 v_2 - F(u_1, u_2) \\
&= (-2v_1 + 3v_2)v_1 + (3v_1 - 4v_2)v_2 - F(-2v_1 + 3v_2, 3v_1 - 4v_2) \\
&= -2v_1^2 + 6v_1v_2 - 4v_2^2 - \\
&\qquad \left(2(-2v_1 + 3v_2)^2 + 3(-2v_1 + 3v_2)(3v_1 - 4v_2) + (3v_1 - 4v_2)^2\right) \\
&= -v_1^2 + 3v_1v_2 - 2v_2^2,
\end{aligned}$$

the same as was obtained directly. This is the **Legendre transform** of the given function F. ■

* As we have seen, this may require a constant to be added to the function G.

Active and passive variables

The variables $\boldsymbol{u} = (u_1, u_2)$ and $\boldsymbol{v} = (v_1, v_2)$ are called **active variables** because they are the ones that are actually transformed. However, the functions F and G may also depend on additional variables that are not part of the transformation as such, but have the status of parameters. These are called **passive variables**. In the dynamical problem, $\dot{\boldsymbol{q}}$ and $\boldsymbol{p}$ are the active variables and $\boldsymbol{q}$ is the passive variable. We need to find how partial derivatives of F and G with respect to the passive variables are related.

Suppose then that $F = F(u_1, u_2, w)$ and $G = G(v_1, v_2, w)$ satisfy the formulae (14.6) and (14.7), where w is a passive variable. Then (14.6) defines v_1, v_2 as functions of u_1, u_2 and w, and (14.7) defines u_1, u_2 as functions of v_1, v_2 and w. The argument leading to the formula (14.8) still holds so that

$$F(u_1, u_2, w) + G(v_1, v_2, w) = u_1 v_1 + u_2 v_2. \tag{14.10}$$

In this formula, imagine that v_1 and v_2 are replaced by their expressions in terms of u_1, u_2 and w; then differentiate the resulting identity with respect to w, holding u_1 and u_2 constant. On using the chain rule, this gives

$$\frac{\partial F}{\partial w} + \left(\frac{\partial G}{\partial v_1} \times \frac{\partial v_1}{\partial w} + \frac{\partial G}{\partial v_2} \times \frac{\partial v_2}{\partial w} + \frac{\partial G}{\partial w} \right) = u_1 \frac{\partial v_1}{\partial w} + u_2 \frac{\partial v_2}{\partial w},$$

which can be written

$$\frac{\partial F}{\partial w} + \frac{\partial G}{\partial w} = \left(u_1 - \frac{\partial G}{\partial v_1} \right) \frac{\partial v_1}{\partial w} + \left(u_2 - \frac{\partial G}{\partial v_2} \right) \frac{\partial v_2}{\partial w} = 0 + 0 = 0,$$

on using the relations (14.7). Hence the partial derivatives of $F(u_1, u_2, w)$ and $G(v_1, v_2, w)$ with respect to w are related by

$$\boxed{\frac{\partial F}{\partial w} = -\frac{\partial G}{\partial w}} \tag{14.11}$$

This is the required result; it holds for *each* passive variable w.

The general case with many variables

The preceding theory can be extended to any number of variables. The results are exactly what one would expect and are summarised in the box below. This summary is presented in a compact vector form using the n-dimensional 'grad'.* It is a good idea to write these results out in expanded form.

* If $F = F(\boldsymbol{u}, \boldsymbol{w})$, where $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$ and $\boldsymbol{w} = (w_1, w_2, \ldots, w_m)$, then $\text{grad}_{\boldsymbol{u}} F$ and $\text{grad}_{\boldsymbol{w}} F$ mean

$$\text{grad}_{\boldsymbol{u}} F(\boldsymbol{u}, \boldsymbol{w}) = \left(\frac{\partial F}{\partial u_1}, \frac{\partial F}{\partial u_2}, \cdots, \frac{\partial F}{\partial u_n} \right), \qquad \text{grad}_{\boldsymbol{w}} F(\boldsymbol{u}, \boldsymbol{w}) = \left(\frac{\partial F}{\partial w_1}, \frac{\partial F}{\partial w_2}, \cdots, \frac{\partial F}{\partial w_m} \right).$$

Legendre transforms

Suppose that the variables $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$ are defined as functions of the active variables $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$ and passive variables $\boldsymbol{w} = (w_1, w_2, \ldots, w_m)$ by the formula

$$\boldsymbol{v} = \text{grad}_{\boldsymbol{u}} F(\boldsymbol{u}, \boldsymbol{w}), \tag{14.12}$$

where F is a given function of $\boldsymbol{u}$ and $\boldsymbol{w}$. Then the inverse formula can always be written in the form

$$\boldsymbol{u} = \text{grad}_{\boldsymbol{v}} G(\boldsymbol{v}, \boldsymbol{w}), \tag{14.13}$$

where the function $G(\boldsymbol{v}, \boldsymbol{w})$ is related to the function $F(\boldsymbol{u}, \boldsymbol{w})$ by the formula

$$G(\boldsymbol{v}, \boldsymbol{w}) = \boldsymbol{u} \cdot \boldsymbol{v} - F(\boldsymbol{u}, \boldsymbol{w}), \tag{14.14}$$

where $\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$.

Furthermore, the derivatives of F and G with respect to the passive variables $\{w_j\}$ are related by

$$\text{grad}_{\boldsymbol{w}} F(\boldsymbol{u}, \boldsymbol{w}) = -\,\text{grad}_{\boldsymbol{w}} G(\boldsymbol{v}, \boldsymbol{w}). \tag{14.15}$$

The relationship between the functions F and G is symmetrical and each is said to be the **Legendre transform** of the other.

14.3 HAMILTON'S EQUATIONS

Let $\mathcal{S}$ be a Lagrangian mechanical system with n degrees of freedom and generalised **coordinates** $\boldsymbol{q} = (q_1, q_1, \ldots, q_n)$. Then the Lagrange equations of motion for $\mathcal{S}$ are

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0 \qquad (1 \le j \le n), \tag{14.16}$$

where $L = L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$ is the Lagrangian of $\mathcal{S}$. This is a set of n second order ODEs in the unknowns $\boldsymbol{q}(t) = (q_1(t), q_2(t), \ldots, q_n(t))$. We now wish to convert these equations into **Hamilton form**, that is, an equivalent set of $2n$ first order ODEs in the $2n$ unknowns $\boldsymbol{q}(t)$, $\boldsymbol{p}(t)$, where $\boldsymbol{p}(t) = (p_1(t), p_2(t), \ldots, p_n(t))$, where the $\{p_j\}$ are the generalised **momenta** of $\mathcal{S}$. The $\{p_j\}$ are defined by

$$p_j = \frac{\partial L}{\partial \dot{q}_j} \qquad (1 \le j \le n), \tag{14.17}$$

which can be written in the vector form

$$\boldsymbol{p} = \operatorname{grad}_{\dot{\boldsymbol{q}}} L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t). \tag{14.18}$$

The first step is to eliminate the coordinate velocities $\dot{\boldsymbol{q}}$ from the Lagrange equations in favour of the momenta $\boldsymbol{p}$. This in turn requires that the formula (14.18) must be inverted so as to express $\dot{\boldsymbol{q}}$ in terms of $\boldsymbol{q}$, $\boldsymbol{p}$ and t. *This is precisely what Legendre transforms do.* It follows from the theory of the last section that the inverse formula to (14.18) can be written in the form

$$\dot{\boldsymbol{q}} = \operatorname{grad}_{\boldsymbol{p}} H(\boldsymbol{q}, \boldsymbol{p}, t), \tag{14.19}$$

where the function $H(\boldsymbol{q}, \boldsymbol{p}, t)$ is the **Legendre transform** of the Lagrangian function $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$. Here, $\dot{\boldsymbol{q}}$ and $\boldsymbol{p}$ are the active variables and $\boldsymbol{q}$ is the passive variable.

Definition 14.1 ***Hamiltonian function*** *The function $H(\boldsymbol{q}, \boldsymbol{p}, t)$, which is the Legendre transform of the Lagrangian function $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$, is called the **Hamiltonian function** of the system $\mathcal{S}$.*

Since the functions H and L are Legendre transforms of each other, they satisfy the relations

$$\boxed{H(\boldsymbol{q}, \boldsymbol{p}, t) = \dot{\boldsymbol{q}} \cdot \boldsymbol{p} - L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)} \tag{14.20}$$

which can be used to generate H from L, and

$$\operatorname{grad}_{\boldsymbol{q}} L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) = -\operatorname{grad}_{\boldsymbol{q}} H(\boldsymbol{q}, \boldsymbol{p}, t), \tag{14.21}$$

which connects the derivatives of L and H with respect to the passive variables.

It is now quite easy to perform the transformation of Lagrange's equations. The Lagrange equations (14.16) can be written in terms of the generalised momenta $\{p_j\}$ in the form

$$\dot{p}_j = \frac{\partial L}{\partial q_j} \qquad (1 \le j \le n),$$

which is equivalent to the vector form

$$\dot{\boldsymbol{p}} = \operatorname{grad}_{\boldsymbol{q}} L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t). \tag{14.22}$$

The right sides of these equations still involve $\dot{\boldsymbol{q}}$, but, on using the formula (14.21), we obtain

$$\dot{\boldsymbol{p}} = -\operatorname{grad}_{\boldsymbol{q}} H(\boldsymbol{q}, \boldsymbol{p}, t). \tag{14.23}$$

These are the transformed Lagrange equations! *The **Hamilton form** of the Lagrange equations therefore consists of equations (14.23) together with equations (14.19), which effectively define the generalised momentum $\boldsymbol{p}$.*

All of the above argument is reversible and so the Hamilton form and the Lagrange form are equivalent. Our results are summarised below:

Hamilton's equations

The n Lagrange equations (14.16) are equivalent to the system of $2n$ first order ODEs

$$\dot{\boldsymbol{q}} = \text{grad}_{\boldsymbol{p}} H(\boldsymbol{q}, \boldsymbol{p}, t), \qquad \dot{\boldsymbol{p}} = -\text{grad}_{\boldsymbol{q}} H(\boldsymbol{q}, \boldsymbol{p}, t), \tag{14.24}$$

where the **Hamiltonian function** $H(\boldsymbol{q}, \boldsymbol{p}, t)$ is the Legendre transform of the Lagrangian $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$ and is generated by the formula (14.20). This is the vector form of **Hamilton's equations**.* The expanded form is

$$\dot{q}_j = \frac{\partial H}{\partial p_j}, \qquad \dot{p}_j = -\frac{\partial H}{\partial q_j} \qquad (1 \le j \le n).$$

We have shownthat the n second order Lagrange equations in the n unknowns $\boldsymbol{q}(t)$ are mathematically equivalent to the $2n$ first order Hamilton equations in the $2n$ unknowns $\boldsymbol{q}(t)$, $\boldsymbol{p}(t)$. In each of these formulations of mechanics, the motion of the system is determined by the form of a single function, the **Lagrangian** $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$ in the Lagrange formulation, and the **Hamiltonian** $H(\boldsymbol{q}, \boldsymbol{p}, t)$ in the Hamilton formulation. Hamilton's equations are a particularly elegant first order system in which the functions $F_1, F_2, \ldots$ that appear on the right are simply the first partial derivatives of a *single* function, the Hamiltonian H. Moreover these right hand sides also satisfy the special condition† div $\boldsymbol{F} = 0$, which allows Liouville's theorem to be applied to Hamiltonian mechanics.‡

Explicit time dependence

One final note. When the Lagrangian has an *explicit* time dependence, this t has the status of an extra passive variable. It follows that we then have the additional relation

$$\frac{\partial L}{\partial t} = -\frac{\partial H}{\partial t}, \tag{14.25}$$

* After Sir William Rowan Hamilton, whose paper *Second Essay on a General Method in Dynamics* was published in 1835. Hamilton's equations are sometimes called the *canonical equations*; no one seems to know the reason why.

† The scalar quantity div $\boldsymbol{F}$ is defined by

$$\text{div}\, \boldsymbol{F} = \frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} + \cdots + \frac{\partial F_n}{\partial x_n}.$$

‡ None of these statements is true if Lagrange's equations are expressed as a first order system by taking the coordinate *velocities* $\dot{\boldsymbol{q}}$ as the new variables.

which shows that if either of L or H has an explicit time dependence, then so does the other.

Example 14.2 *Finding a Hamiltonian and Hamilton's equations*

Find the Hamiltonian and Hamilton's equations for the simple pendulum.

Solution

The **Lagrangian** for the simple pendulum is

$$L = \tfrac{1}{2}ma^2\dot{\theta}^2 + mga\cos\theta,$$

where θ is the angle between the string and the downward vertical, m is the mass of the bob, and a is the string length. The momentum p_θ conjugate to the coordinate θ is given by

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = ma^2\dot{\theta}$$

and this formula is easily inverted to give

$$\dot{\theta} = \frac{p_\theta}{ma^2}. \tag{14.26}$$

The Hamiltonian H is then given by

$$H = \dot{\theta}\, p_\theta - L,$$

where $\dot{\theta}$ is given by equation (14.26). This gives

$$\begin{aligned} H &= \left(\frac{p_\theta}{ma^2}\right) p_\theta - \tfrac{1}{2}ma^2\left(\frac{p_\theta}{ma^2}\right)^2 - mga\cos\theta \\ &= \frac{p_\theta^2}{2ma^2} - mga\cos\theta. \end{aligned}$$

This is the **Hamiltonian** for the simple pendulum. From H we can find Hamilton's equations. They are

$$\begin{aligned} \dot{\theta} &= \frac{\partial H}{\partial p_\theta} = \frac{p_\theta}{ma^2}, \\ \dot{p_\theta} &= -\frac{\partial H}{\partial \dot{\theta}} = -mga\sin\theta. \end{aligned}$$

These are **Hamilton's equations** for the simple pendulum.

This simple example illustrates clearly why Lagrange's equations are preferred over Hamilton's equations for the *practical* solution of problems. To solve Hamilton's equations in this case, we would differentiate the first equation with respect to t and then use the second equation to eliminate the unknown p_θ. This gives

$$\ddot{\theta} + \frac{g}{a}\sin\theta = 0,$$

which is precisely the Lagrange equation for the system! ■

Properties of the Hamiltonian H

The Hamiltonian function $H(\boldsymbol{q}, \boldsymbol{p}, t)$ has been defined as the Legendre transform of $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$ and, as such, it can be generated by the formula (14.20). We have met this expression before. It is identical to the **energy function** h

$$h = \sum_{j=1}^{n} \dot{q}_j p_j - L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t), \tag{14.27}$$

defined in section 12.8. The only difference between h and H is that the functional form of H is vital. H *must* be be expressed in terms of the variables $\boldsymbol{q}$, $\boldsymbol{p}$, t. On the other hand, since only the *values* taken by h are significant, its functional form is unimportant and it may be expressed in terms of any variables. However, since the values taken by H and h are the same, the results that we obtained in section 12.8 concerning h must also be true for the Hamiltonian H. In particular, when H has **no explicit time dependence**, H is a **constant of the motion**.* This result can also be proved independently, as follows.

Suppose that $H = H(\boldsymbol{q}, \boldsymbol{p})$ and that $\{\boldsymbol{q}(t), \boldsymbol{p}(t)\}$ is a motion of the system. Then, in this motion,

$$\begin{aligned}\frac{dH}{dt} &= \sum_{j=1}^{n} \frac{\partial H}{\partial q_j}\,\dot{q}_j + \sum_{j=1}^{n} \frac{\partial H}{\partial p_j}\,\dot{p}_j \\ &= \sum_{j=1}^{n} \frac{\partial H}{\partial q_j}\left(\frac{\partial H}{\partial p_j}\right) + \sum_{j=1}^{n} \frac{\partial H}{\partial p_j}\left(-\frac{\partial H}{\partial q_j}\right) \\ &= 0,\end{aligned}$$

where the first step follows from the chain rule and the second from Hamilton's equations. Hence H remains constant in the motion.

Systems for which $H = H(\boldsymbol{q}, \boldsymbol{p})$ are said to be **autonomous**. (This term was previously applied to systems for which $L = L(\boldsymbol{q}, \dot{\boldsymbol{q}})$, but equation (14.25) shows that these two classes of systems are the same.) The above result can therefore be expressed in the form:

> **Autonomous systems conserve H**
>
> In any motion of an autonomous system, the Hamiltonian $H(\boldsymbol{q}, \boldsymbol{p})$ is conserved.

In addition, when $\mathcal{S}$ is a **conservative standard system**, the Hamiltonian H can be expressed in the simpler form

$$\boxed{H(\boldsymbol{q}, \boldsymbol{p}) = T(\boldsymbol{q}, \boldsymbol{p}) + V(\boldsymbol{q})} \tag{14.28}$$

* As we remarked earlier, H has an explicit time dependence when L does; the circumstances under which this occurs are listed in section 12.6.

where $T(\boldsymbol{q}, \boldsymbol{p})$ is the kinetic energy of the system expressed in terms of the variables $\boldsymbol{q}$, $\boldsymbol{p}$. *In this case, H is simply the **total energy** of the system, expressed in terms of the variables $\boldsymbol{q}$, $\boldsymbol{p}$.* This is the quickest way of finding H when the system is conservative.

Example 14.3 *Finding a Hamiltonian 2*

Find the Hamiltonian for the inverse square orbit problem considered earlier and deduce Hamilton's equations for this system.

Solution

This is a **conservative** system so that $H = T + V$. With the polar coordinates r and θ as generalised coordinates, T and V are given by

$$T = \tfrac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right) \qquad V = -\frac{mMG}{r}.$$

and the generalised momenta are given by

$$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r}, \qquad p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta}.$$

These equations are easily inverted to give

$$\dot{r} = \frac{p_r}{m}, \qquad \dot{\theta} = \frac{p_\theta}{mr^2}$$

so that the Hamiltonian is given by

$$\begin{aligned} H = T + V &= \tfrac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right) - \frac{mMG}{r} \\ &= \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} - \frac{mMG}{r}. \end{aligned}$$

This is the required **Hamiltonian**. Hamilton's equations are now found by using this Hamiltonian in the general equations (14.25). The partial derivatives of H are

$$\frac{\partial H}{\partial p_r} = \frac{p_r}{m}, \qquad \frac{\partial H}{\partial p_\theta} = \frac{p_\theta}{mr^2}, \qquad \frac{\partial H}{\partial r} = -\frac{p_\theta^2}{mr^3}, \qquad \frac{\partial H}{\partial \theta} = 0$$

and **Hamilton's equations** for the orbit problem are therefore

$$\begin{aligned} \dot{r} &= \frac{p_r}{m}, & \dot{p}_r &= \frac{p_\theta^2}{mr^3} - \frac{mMG}{r^2}, \\ \dot{\theta} &= \frac{p_\theta}{mr^2}, & \dot{p}_\theta &= 0. \end{aligned}$$

Naturally, these are the same equations as were obtained earlier by 'manual' transformation of Lagrange's equations. As in the last example, solution of the Hamilton equations by eliminating the momenta simply leads back to Lagrange's equations. ■

Momentum conservation

From the Hamilton equation $\dot{p}_j = -\partial H/\partial q_j$, it follows that:

If $\partial H/\partial q_j = 0$ *(that is, if the coordinate* q_j *is absent from the Hamiltonian), then the generalised momentum* p_j *is constant in any motion.*

The corresponding result in the Lagrangian formulation is that:

If $\partial L/\partial q_j = 0$ *(that is, if the coordinate* q_j *is absent from the Lagrangian), then the generalised momentum* p_j *is constant in any motion.*

These two results seem slightly different, but they *are* equivalent, since, from equation (14.21), $\partial H/\partial q_j = -\partial L/\partial q_j$. This means that the term **cyclic coordinate**, by which we previously meant a coordinate that did not appear in the *Lagrangian*, can be applied without ambiguity to mean that the coordinate does not appear in the *Hamiltonian*. Our result is then:

Conservation of momentum

If q_j is a cyclic coordinate (in the sense that it does not appear in the Hamiltonian), then p_j, the generalised momentum conjugate to q_j, is constant in any motion.

14.4 HAMILTONIAN PHASE SPACE (($\boldsymbol{q}$, $\boldsymbol{p}$)–space)

Suppose the mechanical system $\mathcal{S}$ has generalised coordinates $\boldsymbol{q}$, conjugate momenta $\boldsymbol{p}$, and Hamiltonian $H(\boldsymbol{q}, \boldsymbol{p}, t)$. If the initial values of $\boldsymbol{q}$ and $\boldsymbol{p}$ are known,* then the subsequent motion of $\mathcal{S}$, described by the functions $\{\boldsymbol{q}(t), \boldsymbol{p}(t)\}$, is uniquely determined by Hamilton's equations. This motion can be represented geometrically by the motion of a 'point' (called a **phase point**) in Hamiltonian **phase space**. Hamiltonian phase space is a real space of $2n$ dimensions in which a 'point' is a set of values $(q_1, q_2, \ldots, q_n, p_1, p_2, \ldots, p_n)$ of the independent variables $\{\boldsymbol{q}, \boldsymbol{p}\}$. (Note that a point in Hamiltonian phase space represents not only the configuration of the system $\mathcal{S}$ but also its instantaneous momenta. This is the distinction between Hamiltonian phase space, which has $2n$ dimensions, and Lagrangian configuration space, which has n dimensions.) Each motion of the system $\mathcal{S}$ then corresponds to the motion of a phase point through the phase space.†

The only case in which we can actually draw the phase space is when $\mathcal{S}$ has just one degree of freedom. Then the phase space is two-dimensional and can be drawn on paper.

* We are more familiar with initial conditions in which $\boldsymbol{q}$ and $\dot{\boldsymbol{q}}$ are prescribed. However, these conditions are equivalent to those in which the initial values of $\boldsymbol{q}$ and $\boldsymbol{p}$ are prescribed.

† It should be noted that *Hamiltonian* phase space is generally not the same as the phase space introduced in Chapter 8, which, in the present notation, is $(\boldsymbol{q}, \dot{\boldsymbol{q}})$-space. In particular, our next result (Liouville's theorem) does not apply in $(\boldsymbol{q}, \dot{\boldsymbol{q}})$-space.

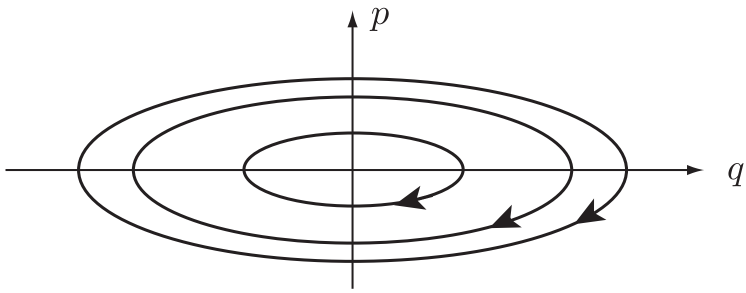


FIGURE 14.1 Typical paths in the phase space (q, p) corresponding to motions of a system $\mathcal{S}$ with Hamiltonian $H = p^2 + q^2/9$. The arrows show the direction that the phase point moves along each path as t increases.

Example 14.4 *Paths in phase space*

Suppose that $\mathcal{S}$ has the single coordinate q and Lagrangian

$$L = \frac{\dot{q}^2}{4} - \frac{q^2}{9}.$$

Find the paths in Hamiltonian phase space that correspond to the motions of $\mathcal{S}$.

Solution

The conjugate momentum $p = \partial L/\partial \dot{q} = \frac{1}{2}\dot{q}$, and the Hamiltonian is

$$H = \dot{q}\,p - L = (2p)\,p - \frac{1}{4}(2p)^2 + \frac{q^2}{9} = p^2 + \frac{q^2}{9}.$$

Hamilton's equations for $\mathcal{S}$ are therefore

$$\dot{q} = 2p, \qquad \dot{p} = -2q/9.$$

On eliminating p, we find that q satisfies the SHM equation

$$\ddot{q} + (4q/9) = 0.$$

The general solution of the Hamilton equations for $\mathcal{S}$ is therefore

$$q = 3A\cos((2t/3) + \alpha), \qquad p = -A\sin((2t/3) + \alpha),$$

where A and α are arbitrary constants. These are the parametric equations of the **paths** in phase space, the parameter being the time t; each path corresponds to a possible motion of the system $\mathcal{S}$. Some typical paths are shown in Figure 14.1. For this system , every motion is periodic so that the paths are *closed* curves in the (q, p)-plane. (They are actually concentric similar ellipes.) The arrows show the direction that the phase point moves along each path as t increases. ■

Of course, most mechanical systems have more than one degree of freedom so that the corresponding phase space has dimension four or more and cannot be drawn. If the system consists of a mole of gas molecules, the dimension of the phase space is six times Avogadro's number! Nevertheless, the notion of phase space is still valuable for we can still apply **geometrical reasoning** to spaces of higher dimension. This will not *solve* Hamilton's equations of motion, but it does enable us to make valuable predictions about the nature of the motion.

The phase fluid

The paths of phase points have a simpler structure when the system is **autonomous**, that is, $H = H(\boldsymbol{q}, \boldsymbol{p})$. In this case, H is a constant of the motion, so that each phase path must lie on a 'surface'* of constant energy† within the phase space. Thus the phase space is filled with the non-intersecting level surfaces of H, like layers in a multi-dimensional onion, and each phase path is restricted to one of these level surfaces.

For autonomous systems, *there can only be one phase path passing through any point of the phase space*. The reason is as follows: suppose that one phase point is at the point $(\boldsymbol{q}_0, \boldsymbol{p}_0)$ at time t_1, and another phase point is at $(\boldsymbol{q}_0, \boldsymbol{p}_0)$ at time t_2. Then, since H is independent of t, the second motion can be obtained from the first by simply making the substitution $t \to t + t_1 - t_2$, a shift in the origin of time. Therefore the two phase points travel along the *same path* with the second point delayed relative to the first by the constant time $t_2 - t_1$. Hence **phase paths cannot intersect**. This means that the phase space is filled with non-intersecting phase paths like the **streamlines** of a fluid in steady flow. Each motion of the system $\mathcal{S}$ corresponds to a phase point moving along one of these paths, just as the real particles of a fluid move along the fluid streamlines. The $2n$-dimensional vector quantity $\boldsymbol{u} = (\dot{\boldsymbol{q}}, \dot{\boldsymbol{p}})$ has the rôle of the fluid velocity field $\boldsymbol{u}(\boldsymbol{r})$‡ and Hamilton's equations serve to specify what this velocity is at the point $(\boldsymbol{q}, \boldsymbol{p})$ of the phase space. Because of this analogy with fluid mechanics, the motion of phase points in phase space is called the **phase flow**.

14.5 LIOUVILLE'S THEOREM AND RECURRENCE

Consider those phase points that, at some instant, occupy the region $\mathcal{R}_0$ of the phase space, as shown in Figure 14.2. As t increases, these points move along their various phase paths in accordance with Hamilton's equations and, after time t, will occupy some new region $\mathcal{R}_t$ of the phase space. This new region will have a different shape§

* If the phase space has dimension six, then a 'surface' of constant H has dimension five. This is therefore a generalisation of the notion of surface, which normally has dimension two.

† This 'energy' is the generalised energy $H(\boldsymbol{q}, \boldsymbol{p})$. For a conservative system, it is equal to the actual total energy $T + V$.

‡ $\boldsymbol{u}(\boldsymbol{r})$ is the velocity of the fluid particle instantaneously at the point with position vector $\boldsymbol{r}$.

§ Since the motion of many systems is sensitive to the initial conditions, the shape of $\mathcal{R}_t$ can become very weird indeed!

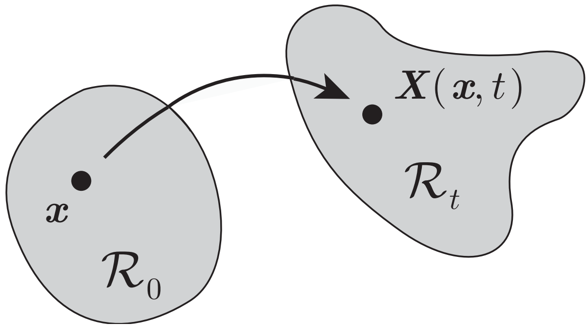


FIGURE 14.2 Liouville's theorem: the Hamiltonian phase flow preserves volume.

to $\mathcal{R}_0$, but **Liouville's theorem*** states that the volumes† of the two regions are equal. This remarkable result is expressed by saying that the *Hamiltonian phase flow preserves volume*. The theorem is easy to apply, but the proof is rather difficult.

Proof of Liouville's theorem

The proof is easier to follow if we use $x_1, x_2, \ldots, x_{2n}$ as the names of the variables (instead of $\boldsymbol{q}$, $\boldsymbol{p}$), and also call the right sides of Hamilton's equations $F_1, F_2, \ldots, F_{2n}$. Then, in vector notation, the equations of motion are $\dot{\boldsymbol{x}} = \boldsymbol{F}(\boldsymbol{x}, t)$. We will give the details for the case when the phase space is two-dimensional; the method in the general case is the same but uglier.

Consider a set of phase points moving in the (x_1, x_2)-plane, which, at some instant in time, occupies the region $\mathcal{R}_0$, as shown in Figure 14.2. Without losing generality, we may suppose that this occurs at time $t = 0$. After time t, a typical point $\boldsymbol{x}$ of $\mathcal{R}_0$ has moved on to position $\boldsymbol{X} = \boldsymbol{X}(\boldsymbol{x}, t)$ and the set as a whole now occupies the region $\mathcal{R}_t$. In this two-dimensional case, the 'volume' $v(t)$ of $\mathcal{R}_t$ is the *area* of this region in the (x_1, x_2)-plane. Now

$$v(t) = \int_{\mathcal{R}_t} dX_1 dX_2 = \int_{\mathcal{R}_0} J\, dx_1 dx_2,$$

where J is the Jacobian of the transformation $\boldsymbol{X} = \boldsymbol{X}(\boldsymbol{x}, t)$, that is,

$$J = \begin{vmatrix} \partial X_1/\partial x_1 & \partial X_1/\partial x_2 \\ \partial X_2/\partial x_1 & \partial X_2/\partial x_2 \end{vmatrix}. \tag{14.29}$$

Now, for small t, $\boldsymbol{X}$ may be approximated by

$$\begin{aligned} \boldsymbol{X}(\boldsymbol{x}, t) &= \boldsymbol{X}(\boldsymbol{x}, 0) + t\, \frac{\partial \boldsymbol{X}}{\partial t}(\boldsymbol{x}, 0) + O(t^2) \\ &= \boldsymbol{x} + t\, \boldsymbol{F}(\boldsymbol{x}, 0) + O(t^2), \end{aligned}$$

on using the equation of motion $\dot{\boldsymbol{x}} = \boldsymbol{F}(\boldsymbol{x}, t)$. The corresponding approximation for J is

$$J = 1 + t \left[\frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} \right]_{t=0} + \ O(t^2) = 1 + t \operatorname{div} \boldsymbol{F}(\boldsymbol{x}, 0) + O(t^2).$$

* After the French mathematician Joseph Liouville (1809–1882).

† Since the dimension of the phase space can be any (even) number, this is a generalisation of the notion of volume.

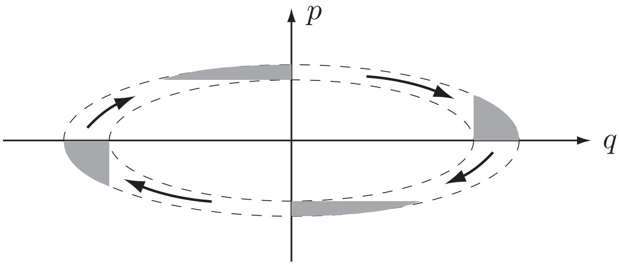


FIGURE 14.3 An instance of Liouville's theorem with the Hamiltonian $H = p^2 + q^2/9$. The shaded region moves through the phase space. Its shape changes but its area remains the same.

Hence the volume of $\mathcal{R}_t$ is approximated by

$$v(t) = \int_{\mathcal{R}_0} \Big(1 + t \operatorname{div} \boldsymbol{F}(\boldsymbol{x}, 0)\Big)\, dx_1 dx_2 + O(t^2)$$

when t is small. It follows that

$$\left.\frac{dv}{dt}\right|_{t=0} = \lim_{t\to 0}\left(\frac{v(t) - v(0)}{t}\right) = \int_{\mathcal{R}_0} \operatorname{div} \boldsymbol{F}(\boldsymbol{x}, 0)\ dx_1 dx_2.$$

Finally, since the initial instant $t = 0$ was arbitrarily chosen, this result must apply for any t, that is,

$$\frac{dv}{dt} = \int_{\mathcal{R}_t} \operatorname{div} \boldsymbol{F}(\boldsymbol{x}, t)\ dx_1 dx_2$$

at any time t. We see that, for *general* systems of equations, the phase flow does *not* preserve volume. However, if $\operatorname{div} \boldsymbol{F}(\boldsymbol{x}, t) = 0$, then volume *is* preserved. For the case of Hamilton's equations with one degree of freedom,

$$\begin{aligned}\operatorname{div} \boldsymbol{F} &= \frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} \\ &= \frac{\partial}{\partial q}\left(\frac{\partial H}{\partial p}\right) + \frac{\partial}{\partial p}\left(-\frac{\partial H}{\partial q}\right) = 0.\end{aligned}$$

Hence the Hamiltonian phase flow satisfies the condition $\operatorname{div} \boldsymbol{F} = 0$ and so preserves volume. This completes the proof. ∎

Liouville's theorem

The motions of a Hamiltonian system preserve volume in $(\boldsymbol{q}, \boldsymbol{p})$-space.

A particular instance of Liouville's theorem is shown in Figure 14.3. The phase paths of the Hamiltonian $H = p^2 + q^2/9$, shown in Figure 14.1, are concentric similar ellipses. Figure 14.3 shows the progress of a region of the phase space lying between two such elliptical paths. The region changes shape but its area remains the same.

Liouville's theorem has many applications and is particularly important in statistical mechanics. The following is a simple example.

Example 14.5 ***No limit cycles in Hamiltonian mechanics***

In the theory of dynamical systems, a periodic solution is said to be an *asymptotically stable limit cycle* if it 'attracts' points in nearby volumes of the phase space (see Chapter 8). Show that limit cycles cannot occur in the dynamics of Hamiltonian systems.

Solution

Suppose there were a closed path $\mathcal{C}$ in the phase space that attracts points in a nearby region $\mathcal{R}$. Then eventually the points that lay in $\mathcal{R}$ must lie in a narrow 'tube' of *arbitrarily small* 'radius' enclosing the path $\mathcal{C}$. The volume of this tube tends to zero with increasing time so that the original volume of $\mathcal{R}$ cannot be be preserved. This is contrary to Liouville's theorem and so asymptotically stable limit cycles cannot exist. ■

Poincaré's theorem and recurrence

Many Hamiltonian systems have the property that each path is confined to some **bounded region** within the phase space. Typically this is a consequence of **energy conservation**, where the energy surfaces happen to be bounded. Liouville's theorem has startling implications concerning the motion of such systems. First, we need to prove a result known as **Poincaré's recurrence theorem**. Poincaré's theorem is actually a result from ergodic theory and has many applications outside classical mechanics. However, since we are going to apply it to phase space, we will prove it in that context.

Theorem 14.1 ***Poincaré's recurrence theorem*** *Let $\mathcal{S}$ be an* ***autonomous*** *Hamiltonian system and consider the motion of the phase points that initially lie in a bounded region $\mathcal{R}_0$ of the phase space. If the paths of all of these points lie within a fixed* ***bounded*** *region of phase space for all time, then some of the points must eventually return to $\mathcal{R}_0$.*

Proof. Let $\mathcal{R}_1$ be the region occupied by the points after time τ. (We will suppose that $\mathcal{R}_1$ does not overlap $\mathcal{R}_0$ so that all the points that lay in $\mathcal{R}_0$ at time $t = 0$ have left $\mathcal{R}_0$ at time $t = \tau$.) We must show that some of them eventually return to $\mathcal{R}_0$. Let $\mathcal{R}_2, \mathcal{R}_3, \ldots, \mathcal{R}_n$ be the regions occupied by the same points after times $2\tau, 3\tau, \ldots, n\tau$. By Liouville's theorem, *all of these regions have the same volume*. Therefore, if they never overlap, their total volume will increase without limit. But, by assumption, all these regions lie within some *finite* volume, so that eventually one of them must overlap a previous one. This much is obvious, but we must now show that an overlap takes place with the *original* region $\mathcal{R}_0$.

Suppose it is $\mathcal{R}_m$ that overlaps $\mathcal{R}_k$ $(0 \le k < m)$. Each point of this overlap region corresponds to an intersection of the paths of two phase points that started out at some points $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ of $\mathcal{R}_0$ at time $t = 0$. In the same notation used in the proof of Liouville's theorem, it means that $\boldsymbol{X}(\boldsymbol{x}_1, m\tau) = \boldsymbol{X}(\boldsymbol{x}_2, k\tau)$. But, since the system is **autonomous**, the two solutions $\boldsymbol{X}(\boldsymbol{x}_1, t)$ and $\boldsymbol{X}(\boldsymbol{x}_2, t)$ must therefore differ only by a shift $(m - k)\tau$ in the origin of time. It follows that $\boldsymbol{X}(\boldsymbol{x}_1, (m-k)\tau) = \boldsymbol{X}(\boldsymbol{x}_2, 0) = \boldsymbol{x}_2$. Thus the phase point that was at $\boldsymbol{x}_1$ when $t = 0$ is at $\boldsymbol{x}_2$ when $t = (m - k)\tau$. This phase point has therefore returned to $\mathcal{R}_0$ after time $(m - k)\tau$ and this completes the proof. ■

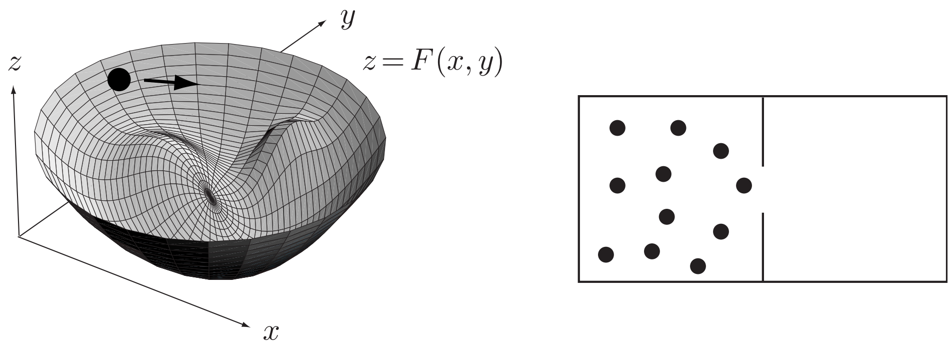


FIGURE 14.4 Consequences of Poincaré's recurrence theorem. **Left**: the particle sliding inside the smooth irregular bowl will eventually almost reassume its initial state. **Right**: The mole of gas molecules, initially all in the left compartment will eventually all be found there again.

Since the recurrence theorem holds for any sub-region of $\mathcal{R}_0$, it follows that, throughout $\mathcal{R}_0$, there are phase points that pass *arbitrarily close* to their original positions. Thus if the system $\mathcal{S}$ has one of these points as its initial state, then $\mathcal{S}$ will eventually become arbitrarily close to reassuming that state. Actually, such points are typical rather than exceptional. To show this requires a stronger version of Poincaré's theorem* than we have proved here, namely: *The path of* ***almost every***† *point in $\mathcal{R}_0$ passes* ***arbitrarily close*** *to its starting point*. This implies the remarkable result that *for amost every choice of the initial conditions, the system $\mathcal{S}$ becomes arbitrarily close to reassuming those conditions at later times*.

An example of this phenomenon is the motion of a single particle P sliding under gravity on the smooth inner surface of a bowl of some irregular shape $z = F(x, y)$, as shown in Figure 14.4. This is an autonomous Hamiltonian system with two degrees of freedom. Take the Cartesian coordinates (x, y) of P to be generalised coordinates. Then the Lagrangian is given by

$$\begin{aligned} L &= \tfrac{1}{2}m\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) - mgz \\ &= \tfrac{1}{2}m\left(\dot{x}^2 + \dot{y}^2 + (F,_x\,\dot{x} + F,_y\,\dot{y})^2\right) - mgF, \end{aligned}$$

where $F,_x = \partial F/\partial x$ and $F,_y = \partial F/\partial y$. The conjugate momenta are

$$p_x = m\left(\dot{x} + (F,_x\,\dot{x} + F,_y\,\dot{y})F,_x\right), \qquad p_y = m\left(\dot{y} + (F,_x\,\dot{x} + F,_y\,\dot{y})F,_y\right).$$

Just because energy is conserved, it does not necessarily mean that Poincaré's theorem applies. We must show that the energy surfaces in phase space are bounded. The proof of this is as follows:

* See Walters [12].

† This means that the set of exceptional points has measure zero. For example, in a two-dimensional phase space, a curve has zero measure.

In this case, $\mathcal{R}_0$ is some bounded region of the phase space (x, y, p_x, p_y). Suppose that z_0 is the maximum value of z and that T_0 is the maximum kinetic energy associated with points of $\mathcal{R}_0$. Then, by energy conservation, the maximum value of z in the subsequent motions cannot exceed $z^{\max} = z_0 + (T_0/mg)$. Hence, providing the bowl rises to at least this height, the motions are confined to values of (x, y) that satisfy $F(x, y) \leq z^{\max}$. It follows that both x and y are bounded in the subsequent motions. Also, if the lowest point of the bowl is at $z = 0$, the value of T in the subsequent motions cannot exceed $T^{\max} = T_0 + mgz_0$. It follows that $\dot{x}$ and $\dot{y}$ are bounded in the subsequent motions and this implies that the same is true for p_x and p_y. Hence x, y, p_x, and p_y are all bounded in the subsequent motions. This means that that the paths of the phase points that lie in $\mathcal{R}_0$ when $t = 0$ are confined to a bounded region of the phase space for all time. **Poincaré's theorem therefore applies**.

Hence, if the particle P is released from rest (say) at some point A on the surface of the bowl, then, whatever the shape of the bowl, P will become arbitrarily close to being at rest at A at later times.

In the same way, it follows that if a compartment containing a mole of gas molecules is separated from an empty compartment by a partition, and the partition is suddenly punctured, then at (infinitely many!) later times the molecules will *all be found in the first compartment again*. This remarkable result, which seems to be in contradiction to the second law of thermodynamics, appears less paradoxical when one realises that 'later times' may mean 10^{20} years later!

Question ***Exceptional points***

How do you know that the initial conditions you have chosen do not correspond to an 'exceptional point' for which Poincaré's theorem does not hold?

Answer

You don't know, but you would be very unlucky if this happened! ■

Problems on Chapter 14

Answers and comments are at the end of the book.

Harder problems carry a star (∗).

Finding Hamiltonians

14.1 Find the Legendre transform $G(v_1, v_2, w)$ of the function

$$F(u_1, u_2, w) = 2u_1^2 - 3u_1u_2 + u_2^2 + 3wu_1,$$

where w is a passive variable. Verify that $\partial F/\partial w = -\partial G/\partial w$.

14.2 A smooth wire has the form of the helix $x = a\cos\theta$, $y = a\sin\theta$, $z = b\theta$, where θ is a real parameter, and a, b are positive constants. The wire is fixed with the axis Oz pointing vertically upwards. A particle P of mass m can slide freely on the wire. Taking θ as generalised coordinate, find the Hamiltonian and obtain Hamilton's equations for this system.

14.3 ***Projectile*** Using Cartesian coordinates, find the Hamiltonian for a projectile of mass m moving under uniform gravity. Obtain Hamilton's equations and identify any cyclic coordinates.

14.4 ***Spherical pendulum*** The spherical pendulum is a particle of mass m attached to a fixed point by a light inextensible string of length a and moving under uniform gravity. It differs from the simple pendulum in that the motion is not restricted to lie in a vertical plane. Show that the Lagrangian is

$$L = \tfrac{1}{2}ma^2\left(\dot{\theta}^2 + \sin^2\theta\,\dot{\phi}^2\right) + mga\cos\theta,$$

where the polar angles θ, ϕ are shown in Figure 11.7. Find the Hamiltonian and obtain Hamilton's equations. Identify any cyclic coordinates.

14.5 The system shown in Figure 10.9 consists of two particles P_1 and P_2 connected by a light inextensible string of length a. The particle P_1 is constrained to move along a fixed smooth horizontal rail, and the whole system moves under uniform gravity in the vertical plane through the rail. For the case in which the particles are of equal mass m, show that the Lagrangian is

$$L = \tfrac{1}{2}m\left(2\dot{x}^2 + 2a\dot{x}\dot{\theta} + a^2\dot{\theta}^2\right) + mga\cos\theta,$$

where x and θ are the coordinates shown in Figure 10.9.

Find the Hamiltonian and verify that it satisfies the equations $\dot{x} = \partial H/\partial p_x$ and $\dot{\theta} = \partial H/\partial p_\theta$. [Messy algebra.]

14.6 ***Pendulum with a shortening string*** A particle is suspended from a support by a light inextensible string which passes through a small fixed ring vertically below the support. The particle moves in a vertical plane with the string taut. At the same time, the support is made to move vertically having an upward displacement $Z(t)$ at time t. The effect is that the particle oscillates like a simple pendulum whose string length at time t is $a - Z(t)$, where a is a positive constant. Show that the Lagrangian is

$$L = \tfrac{1}{2}m\left((a-Z)^2\dot{\theta}^2 + \dot{Z}^2\right) + mg(a-Z)\cos\theta,$$

where θ is the angle between the string and the downward vertical.

Find the Hamiltonian and obtain Hamilton's equations. Is H conserved?

14.7 ***Charged particle in an electrodynamic field*** The Lagrangian for a particle with mass m and charge e moving in the general electrodynamic field $\{\boldsymbol{E}(\boldsymbol{r},t), \boldsymbol{B}(\boldsymbol{r},t)\}$ is given in Cartesian coordinates by

$$L(\boldsymbol{r}, \dot{\boldsymbol{r}}, t) = \tfrac{1}{2}m\,\dot{\boldsymbol{r}}\cdot\dot{\boldsymbol{r}} - e\,\phi(\boldsymbol{r},t) + e\,\dot{\boldsymbol{r}}\cdot\boldsymbol{A}(\boldsymbol{r},t),$$

where $\boldsymbol{r} = (x, y, z)$ and $\{\phi, \boldsymbol{A}\}$ are the electrodynamic potentials of field $\{\boldsymbol{E}, \boldsymbol{B}\}$. Show that the corresponding Hamiltonian is given by

$$H(\boldsymbol{r}, \boldsymbol{p}, t) = \frac{(\boldsymbol{p} - e\boldsymbol{A})\cdot(\boldsymbol{p} - e\boldsymbol{A})}{2m} + e\,\phi,$$

where $\boldsymbol{p} = (p_x, p_y, p_x)$ are the generalised momenta conjugate to the coordinates (x, y, z). [Note that $\boldsymbol{p}$ is *not* the ordinary linear momentum of the particle.] Under what circumstances is H conserved?

14.8 ***Relativistic Hamiltonian*** The relativistic Lagrangian for a particle of rest mass m_0 moving along the x-axis under the potential field $V(x)$ is given by

$$L = m_0 c^2 \left(1 - \left(1 - \frac{\dot{x}^2}{c^2}\right)^{1/2}\right) - V(x).$$

Show that the corresponding Hamiltonian is given by

$$H = m_0 c^2 \left(1 + \left(\frac{p_x}{m_0 c}\right)^2\right)^{1/2} - m_0 c^2 + V(x),$$

where p_x is the generalised momentum conjugate to x.

14.9 ***A variational principle for Hamilton's equations*** Consider the functional

$$J[\boldsymbol{q}(t), \boldsymbol{p}(t)] = \int_{t_0}^{t_1} \Big(H(\boldsymbol{q}, \boldsymbol{p}, t) + \boldsymbol{q} \cdot \dot{\boldsymbol{p}} - \dot{\boldsymbol{q}} \cdot \boldsymbol{p}\Big)\, dt$$

of the $2n$ independent functions $q_1(t), \dots, q_n(t),\ p_1(t), \dots, p_n(t)$. Show that the extremals of J satisfy Hamilton's equations with Hamiltonian H.

Liouville's theorem and recurrence

14.10 In the theory of dynamical systems, a point is said to be an *asymptotically stable equilibrium point* if it 'attracts' points in a nearby volume of the phase space. Show that such points cannot occur in Hamiltonian dynamics.

14.11 A one dimensional damped oscillator with coordinate q satisfies the equation $\ddot{q} + 4\dot{q} + 3q = 0$, which is equivalent to the first order system

$$\dot{q} = v, \qquad \dot{v} = -3q - 4v.$$

Show that the area $a(t)$ of any region of points moving in (q, v)-space has the time variation

$$a(t) = a(0)\, e^{-4t}.$$

Does this result contradict Liouville's theorem?

14.12 ***Ensembles in statistical mechanics*** In statistical mechanics, a macroscopic property of a system $\mathcal{S}$ is calculated by averaging that property over a set, or *ensemble*, of points moving in the phase space of $\mathcal{S}$. The number of ensemble points in any volume of phase space is represented by a *density function* $\rho(\boldsymbol{q}, \boldsymbol{p}, t)$. If the system is autonomous and in *statistical equilibrium*, it is required that, even though the ensemble points are moving (in accordance with

Hamilton's equations), their density function should remain the same, that is, $\rho = \rho(\boldsymbol{q}, \boldsymbol{p})$. This places a restriction on possible choices for $\rho(\boldsymbol{q}, \boldsymbol{p})$. Let $\mathcal{R}_0$ be any region of the phase space and suppose that, after time t, the points of $\mathcal{R}_0$ occupy the region $\mathcal{R}_t$. Explain why statistical equilibrium requires that

$$\int_{\mathcal{R}_0} \rho(\boldsymbol{q}, \boldsymbol{p})\, dv = \int_{\mathcal{R}_t} \rho(\boldsymbol{q}, \boldsymbol{p})\, dv$$

and show that the *uniform* density function $\rho(\boldsymbol{q}, \boldsymbol{p}) = \rho_0$ satisfies this condition. [It can be proved that the above condition is also satisfied by any density function that is constant along the streamlines of the phase flow.]

14.13 Decide if the energy surfaces in phase space are bounded in the following cases:

(i) The two-body gravitation problem with $E < 0$.

(ii) The two-body gravitation problem viewed from the zero momentum frame and with $E < 0$.

(iii) The three-body gravitation problem viewed from the zero momentum frame and with $E < 0$. Does the solar system have the recurrence property?

Poisson brackets

14.14 ***Poisson brackets*** Suppose that $u(\boldsymbol{q}, \boldsymbol{p})$ and $v(\boldsymbol{q}, \boldsymbol{p})$ are any two functions of position in the phase space $(\boldsymbol{q}, \boldsymbol{p})$ of a mechanical system $\mathcal{S}$. Then the **Poisson bracket** $[u, v]$ of u and v is defined by

$$[u, v] = \text{grad}_{\boldsymbol{q}}\, u \cdot \text{grad}_{\boldsymbol{p}}\, v - \text{grad}_{\boldsymbol{p}}\, u \cdot \text{grad}_{\boldsymbol{q}}\, v = \sum_{j=1}^{n} \left(\frac{\partial u}{\partial q_j} \frac{\partial v}{\partial p_j} - \frac{\partial u}{\partial p_j} \frac{\partial v}{\partial q_j} \right).$$

The *algebraic* behaviour of the Poisson bracket of two functions resembles that of the cross product $\boldsymbol{U} \times \boldsymbol{V}$ of two vectors or the commutator $\boldsymbol{UV} - \boldsymbol{VU}$ of two matrices. The Poisson bracket of two functions is closely related to the commutator of the corresponding operators in quantum mechanics.*

Prove the following properties of Poisson brackets.

Algebraic properties

$$[u, u] = 0, \qquad [v, u] = -[u, v], \qquad [\lambda_1 u_1 + \lambda_2 u_2, v] = \lambda_1 [u_1, v] + \lambda_2 [u_2, v]$$

$$[[u, v], w] + [[w, u], v] + [[v, w], u] = 0.$$

This last formula is called *Jacobi's identity*. It is quite important, but there seems to be no way of proving it apart from crashing it out, which is very tedious. Unless you can invent a smart method, leave this one alone.

* The commutator $[\mathbf{U}, \mathbf{V}]$ of two quantum mechanical operators $\mathbf{U}$, $\mathbf{V}$ corresponds to $i\hbar[u, v]$, where $\hbar$ is the modified Planck constant, and $[u, v]$ is the Poisson bracket of the corresponding classical variables u, v.

Fundamental Poisson brackets

$$[q_j, q_k] = 0, \qquad [p_j, p_k] = 0, \qquad [q_j, p_k] = \delta_{jk},$$

where δ_{jk} is the Kroneker delta.

Hamilton's equations

Show that Hamilton's equations for $\mathcal{S}$ can be written in the form

$$\dot{q}_j = [q_j, H], \qquad \dot{p}_j = [p_j, H], \qquad (1 \le j \le n).$$

Constants of the motion

(i) Show that the *total* time derivative of $u(\boldsymbol{q}, \boldsymbol{p})$ is given by

$$\frac{du}{dt} = [u, H]$$

and deduce that u is a constant of the motion of $\mathcal{S}$ if, and only if, $[u, H] = 0$.

(ii) If u and v are constants of the motion of $\mathcal{S}$, show that the Poisson bracket $[u, v]$ is another constant of the motion. [Use Jacobi's identity.] Does this mean that you can keep on finding more and more constants of the motion ?

14.15 ***Integrable systems and chaos*** A mechanical system is said to be **integrable** if its equations of motion are soluble in the sense that they can be reduced to integrations. (You do not need to be able to evaluate the integrals in terms of standard functions.) A theorem due to Liouville states that *any Hamiltonian system with n degrees of freedom is integrable if it has n independent constants of the motion, and all these quantities commute in the sense that all their mutual Poisson brackets are zero.** The qualitative behaviour of integrable Hamiltonian systems is well investigated (see Goldstein [4]). In particular, *no integrable Hamiltonian system can exhibit chaos.*

Use Liouville's theorem to show that any autonomous system with n degrees of freedom and $n - 1$ cyclic coordinates must be integrable.

Computer assisted problem

14.16 *The three body problem*

There is no general solution to the problem of determining the motion of three or more bodies moving under their mutual gravitation. Here we consider a restricted case of the three-body problem in which the mass of one of the bodies, P, is *much smaller* than that of the other two masses, which are called the primaries. In this case we neglect the effect of P on the primaries which therefore move in known fixed orbits. The body P moves in the time dependent, gravitational field of the primaries.

* This result is really very surprising. A *general* system of first order ODEs in $2n$ variables needs $2n$ integrals in order to be integrable in the Liouville sense. Hamiltonian systems need only half that number. The theorem does not rule out the possibility that that there could be other classes of integrable systems. However, according to Arnold [2], every system that has ever been integrated is of the Liouville kind!

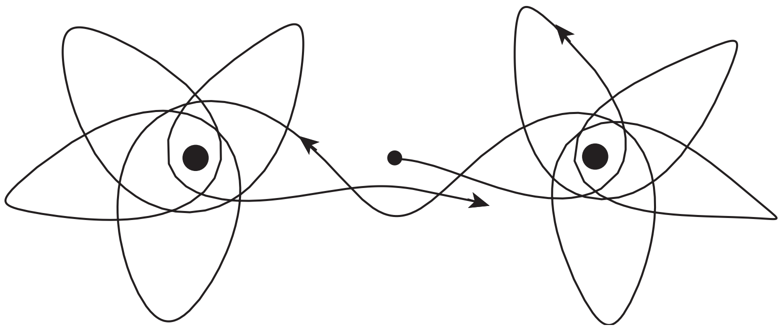


FIGURE 14.5 There is no such thing as a typical orbit in the three-body problem. The orbit shown corresponds to the initial conditions $x = 0$, $y = 0$, $p_x = 1.03$, $p_y = 0$ and is viewed from axes *rotating with the primaries*.

Suppose the primaries each have mass M and move under their mutual gravitation around a fixed circle of radius a, being at the opposite ends of a rotating diameter. The body P moves under the gravitational attraction of the primaries in the same plane as their circular orbit. Using Cartesian coordinates, write code to set up Hamilton's equations for this system and solve them with general initial conditions. [Take M as the unit of mass, a as the unit of length, and take the unit of time so that the speed of the primaries is unity. With this choice of units, the gravitational constant $G = 4$.]

By experimenting with different initial conditions, some very weird orbits can be found for P. It is interesting to plot these relative to fixed axes and also relative to axes rotating with the primaries, as in Figure 14.5. Some fascinating cases are shown by Acheson [] and you should be able to reproduce these. [Acheson used a different normalisation however and his initial data needs to be doubled to be used in your code.]

Part Four

FURTHER TOPICS

CHAPTERS IN PART FOUR

Chapter Fifteen

The general theory of small oscillations

KEY FEATURES

The key features of this chapter are the existence of **small oscillations** near a position of stable equilibrium and the matrix theory of **normal modes**. A simpler account of the basic principles is given in Chapter 5.

Any mechanical system can perform oscillations in the neighbourhood of a position of stable equilibrium. These oscillations are an extremely important feature of the system whether they are intended to occur (as in a pendulum clock), or whether they are undesirable (as in a suspension bridge!). Analogous oscillations occur in continuum mechanics and in quantum mechanics. Here we present the theory of such oscillations for **conservative systems** under the assumption that the amplitude of the oscillations is small enough so that the **linear approximation** is adequate. A simpler account of the theory is given in Chapter 5. This treatment is restricted to systems with two degrees of freedom and does not make use of Lagrange's equations. Although the material in the present chapter is self-contained, it is helpful to have solved a few simple normal mode problems before.

The best way to develop the theory of small oscillations is to use Lagrange's equations. We will show that it is possible to approximate the expressions for T and V from the start so that the linearized equations of motion are obtained immediately. The theory is presented in an elegant matrix form which enables us to make use of concepts from linear algebra, such as eigenvalues and eigenvectors. We prove that fundamental result that a system with n degrees of freedom always has n harmonic motions known as **normal modes**, whose frequencies are generally different. These **normal frequencies** are the most important characteristic of the oscillating system. One important application of the theory is to the internal vibrations of molecules. Although this should really be treated by quantum mechanics, the classical model is extremely valuable in making qualitative predictions and classifying the vibrational modes of the molecule.

15.1 STABLE EQUILIBRIUM AND SMALL OSCILLATIONS

Let $\mathcal{S}$ be a standard mechanical system with n degrees of freedom and with generalised coordinates $\boldsymbol{q} = (q_1, q_2, \ldots, q_n)$. Suppose also that $\mathcal{S}$ is **conservative**. Then the motion of $\mathcal{S}$ is determined by the classical Lagrange equations of motion

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = -\frac{\partial V}{\partial q_j} \qquad (1 \le j \le n), \tag{15.1}$$

where $T(\boldsymbol{q}, \dot{\boldsymbol{q}})$ and $V(\boldsymbol{q})$ are the kinetic and potential energies of $\mathcal{S}$. In particular, these equations determine the **equilibrium positions** of $\mathcal{S}$. The point $\boldsymbol{q}^{(0)}$ in configuration space is an equilibrium position of $\mathcal{S}$ if (and only if) the constant function $\boldsymbol{q} = \boldsymbol{q}^{(0)}$ satisfies the equations (15.1). For a standard system, T has the form

$$T = \sum_{j=1}^{n}\sum_{k=1}^{n} t_{jk}(\boldsymbol{q})\, \dot{q}_j \dot{q}_k, \tag{15.2}$$

that is, a homogeneous quadratic form in the variables $\dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n$, with coefficients depending on $\boldsymbol{q}$. It follows that the left side of the j-th Lagrange equation has the form

$$2\sum_{k=1}^{n}\left(\frac{dt_{jk}}{dt}\,\dot{q}_k + t_{jk}\,\ddot{q}_k\right) - \sum_{j=1}^{n}\sum_{k=1}^{n}\left(\frac{\partial t_{jk}}{\partial q_j}\,\dot{q}_j\dot{q}_k\right),$$

which takes the value zero when the constant function $\boldsymbol{q} = \boldsymbol{q}^{(0)}$ is substituted in. It follows that $\boldsymbol{q} = \boldsymbol{q}^{(0)}$ will satisfy the equations (15.1) if (and only if)

$$\boxed{\frac{\partial V}{\partial q_j} = 0 \qquad (1 \le j \le n)} \tag{15.3}$$

when $\boldsymbol{q} = \boldsymbol{q}^{(0)}$. In other words, we have the result:

Stationary points of V

The **equilibrium positions** of a conservative system are the **stationary points** of its potential energy function $V(\boldsymbol{q})$.

Stable equilibrium

The stability of equilibrium is most easily understood in terms of the motion of the phase point of $\mathcal{S}$ in the Hamilton phase space $(\boldsymbol{q},\, \boldsymbol{p})$ (see Chapter 14). If $\boldsymbol{q}^{(0)}$ is an equilibrium point of $\mathcal{S}$ in *configuration* space, then $(\boldsymbol{q}^{(0)}, \boldsymbol{0})$ is the corresponding equilibrium point of $\mathcal{S}$ in *phase* space; if the phase point starts at $(\boldsymbol{q}^{(0)}, \boldsymbol{0})$, then it remains there, its trajectory consisting of the single point $(\boldsymbol{q}^{(0)}, \boldsymbol{0})$.

Now consider phase paths that begin at points that lie inside the sphere S_δ in phase space which has centre $(\boldsymbol{q}^{(0)}, \boldsymbol{0})$ and radius δ. When δ is small, this corresponds to starting

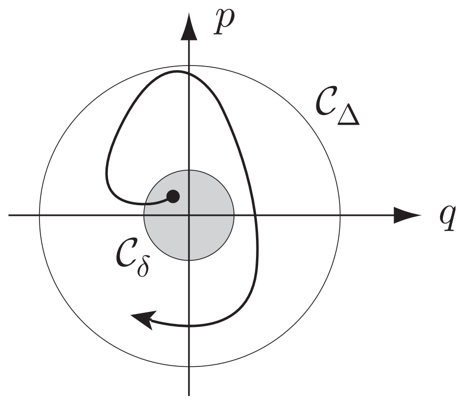


FIGURE 15.1 The circle $\mathcal{C}_\delta$ must be chosen small enough so that all the phase paths that start within it remain within the given circle $\mathcal{C}_\Delta$.

the system from a configuration close to the equilibrium configuration with a small kinetic energy. Let S_Δ be the smallest sphere in phase space with centre $(\boldsymbol{q}^{(0)}, \boldsymbol{0})$ *that contains all the phase paths that begin within* S_δ.

Definition 15.1 ***Stable equilibrium*** *If the radius* Δ *tends to zero as the radius* δ *tends to zero, then the equilibrium point at* $(\boldsymbol{q}^{(0)}, \boldsymbol{0})$ *is said to be* ***stable***.*

This means that, if $\mathcal{S}$ is given a small nudge from a configuration close to a position of *stable* equilibrium, then the subsequent motion of $\mathcal{S}$ (in configuration space) is restricted to a small neighbourhood of the equilibrium point. Thus *any mechanical system can perform small motions near a position of stable equilibrium*. These motions are generally called **small oscillations**.

We know that the equilibrium positions of $\mathcal{S}$ correspond to the stationary points of the potential energy $V(\boldsymbol{q})$, but we have yet to identify which of these points correspond to *stable* equilibrium. In fact it is quite easy to prove the following important result:

Minimum points of *V*

The **minimum points** of the potential energy function $V(\boldsymbol{q})$ are positions of **stable equilibrium** of the system $\mathcal{S}$.

Proof. Without loss of generality, suppose that the minimum point of the function $V(\boldsymbol{q})$ is at $\boldsymbol{q} = \boldsymbol{0}$ and that $V(\boldsymbol{0}) = \boldsymbol{0}$. Take any $\Delta > 0$. Then we must show that we can find a sphere S_δ in phase space such that all the paths that begin within it remain inside the sphere S_Δ. This is illustrated in Figure 15.1 for the only case that can be drawn, namely, when the phase space is two-dimensional; in this case, the 'spheres' are circles.

The result follows from **energy conservation**. Let $T(\boldsymbol{q}, \boldsymbol{p})$ be the kinetic energy of the system. The total energy is then

$$E(\boldsymbol{q}, \boldsymbol{p}) = T(\boldsymbol{q}, \boldsymbol{p}) + V(\boldsymbol{q}).$$

* In the dynamical systems literature, this is known as **Liapunov stability**.

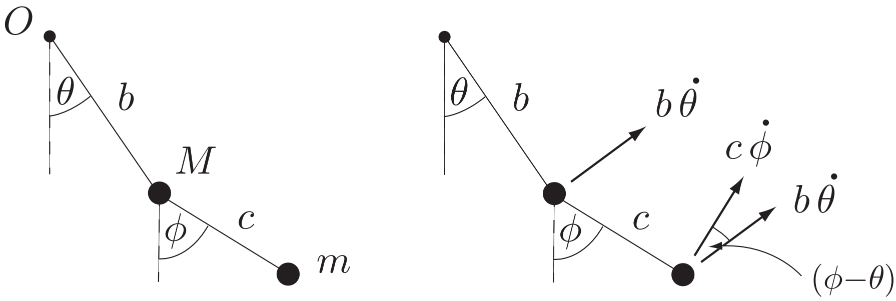


FIGURE 15.2 The double pendulum. **Left**: The generalised coordinates θ, ϕ. **Right**: The velocity diagram.

Since $\boldsymbol{q} = \boldsymbol{0}$ is a minimum point of $V(\boldsymbol{q})$ and $T(\boldsymbol{q}, \boldsymbol{p})$ is positive for $\boldsymbol{p} \neq \boldsymbol{0}$, it follows that $(\boldsymbol{0}, \boldsymbol{0})$, the origin of phase space, is a minimum point of the function $E(\boldsymbol{q}, \boldsymbol{p})$; the value of E at the minimum point is zero. The value of E on the sphere S_Δ must therefore be greater than some *positive* number E_Δ. On the other hand, by making the radius δ small enough, it follows by continuity that the value of E within the sphere S_δ can be made as close to $E(\boldsymbol{0}, \boldsymbol{0})$ $(= 0)$ as we wish; we can certainly make it less than E_Δ. Consider now any phase path starting within the circle $\mathcal{C}_\delta$. Then, by energy conservation, E is constant along this path and is *less* than E_Δ. Such a path cannot cross the sphere S_Δ, for, if it did, the value of E at the crossing point would be *greater* than E_Δ, which is not true. Hence, any phase path starting within the sphere S_δ must remain within the sphere S_Δ and this completes the proof. ■

Example 15.1 *Stability of equilibrium*

Consider the double pendulum shown in Figure 15.2 which moves under uniform gravity. Show that the vertically downwards configuration is a position of stable equilibrium.

Solution

The system is assumed to move in a vertical plane through the suspension point O. In terms of the generalised coordinates θ, ϕ shown, the vertically downwards configuration corresponds to the point $\theta = \phi = 0$ in configuration space. The gravitational potential energy is given by

$$V = (M + m)gb(1 - \cos\theta) + mgc(1 - \cos\phi)$$

so that $\partial V/\partial\theta = (M+m)gb\sin\theta = 0$ when $\theta = \phi = 0$. The same is true for $\partial V/\partial\phi$ and so the point $\theta = \phi = 0$ is a **stationary point** of the function $V(\theta, \phi)$. It follows that the downwards configuration is a position of **equilibrium**.

We could determine the nature of this stationary point by looking at the second derivatives of $V(\theta, \phi)$, but there is no need because it is evident that $V(\theta, \phi) > V(0, 0)$ unless $\theta = \phi = 0$. Thus $(0, 0)$ is a minimum point of V and so the downwards configuration is a position of **stable equilibrium**. ■

15.2 THE APPROXIMATE FORMS OF T AND V

Now that we know small oscillations can take place about any minimum point of V, we can go on to find approximate equations that govern such motions. The obvious (but not the best!) way of doing this is as follows: Take the example of the double pendulum. In this case, T and V are given (see Figure 15.2) by

$$T = \tfrac{1}{2}M(b\dot{\theta})^2 + \tfrac{1}{2}m\left((b\dot{\theta})^2 + (c\dot{\phi})^2 + 2(b\dot{\theta})(c\dot{\phi})\cos(\theta-\phi)\right), \tag{15.4}$$

$$V = (M+m)gb(1-\cos\theta) + mgc(1-\cos\phi). \tag{15.5}$$

If these expressions are substituted into the Lagrange's equations, we obtain (after some simplification) the **exact equations of motion**

$$(M+m)b\ddot{\theta} + mc\cos(\theta-\phi)\ddot{\phi} + mc\sin(\theta-\phi)\dot{\phi}^2 + (M+m)g\sin\theta = 0,$$
$$b\cos(\theta-\phi)\ddot{\theta} + c\ddot{\phi} - b\sin(\theta-\phi)\dot{\theta}^2 + g\sin\phi = 0.$$

This formidable pair of *coupled, second order, non-linear ODEs* govern the large oscillations of the double pendulum. However, for small oscillations about $\theta = \phi = 0$, these equations can be approximated by neglecting everything except linear terms in θ, ϕ and their time derivatives. On carrying out this approximation, the equations simplify dramatically to give

$$(M+m)b\,\ddot{\theta} + mc\,\ddot{\phi} + (M+m)g\,\theta = 0, \tag{15.6}$$

$$b\,\ddot{\theta} + c\,\ddot{\phi} + g\,\phi = 0. \tag{15.7}$$

These are the **linearised equations** governing small oscillations of the double pendulum about the downward vertical. They are a pair of *coupled, second order, linear ODEs with constant coefficients*. An explicit solution is therefore possible.

While the above method of finding the linearised equations of motion is perfectly correct, it is wasteful of effort and is also unsuitable when presenting the general theory. What we did was to obtain the *exact* expressions for T and V, derive the *exact* equations of motion, and then linearise. In the linearisation process, many of the terms we took pains to find were discarded. It makes far better sense to approximate the expressions for T and V from the start so that, when these approximations are used in Lagrange's equations, the linearised equations of motion are produced immediately. The saving in labour is considerable and this is also a nice way to present the general theory.

Consider the double pendulum for example. The exact expression for V is given by equation (15.5) and when θ, ϕ are small, this is given approximately by

$$V = \tfrac{1}{2}(M+m)gb\,\theta^2 + \tfrac{1}{2}mgc\,\phi^2 + \cdots,$$

where the neglected terms have power four or higher. Similarly, when θ, ϕ and their time derivatives are small, T is given approximately by

$$T = \tfrac{1}{2}M(b\dot{\theta})^2 + \tfrac{1}{2}m\left((b\dot{\theta})^2 + (c\dot{\phi})^2 + 2(b\dot{\theta})(c\dot{\phi})(1+\cdots)\right),$$
$$= \tfrac{1}{2}(M+m)b^2\dot{\theta}^2 + mbc\,\dot{\theta}\dot{\phi} + \tfrac{1}{2}mc^2\dot{\phi}^2 + \cdots,$$

where the neglected terms have power four (or higher) in small quantities. If these approximate forms for T and V are now substituted into Lagrange's equations, the linearised equations of motion (15.6), (15.7) are obtained immediately. [Check this.] This is clearly superior to our original method.

The general approximate form of V

In the general case, suppose that the potential energy $V(\boldsymbol{q})$ of the system $\mathcal{S}$ has a minimum at $\boldsymbol{q} = \mathbf{0}$ and that $V(\mathbf{0}) = 0$. (If the minimum point of V is not at $\boldsymbol{q} = \mathbf{0}$, it can always be made so by a simple change of coordinates.) Then, for $\boldsymbol{q}$ near $\mathbf{0}$, $V(\boldsymbol{q})$ can be expanded as an (n-dimensional) Taylor series in the variables $q_1, q_2, \ldots, q_n$. For the special case when $\mathcal{S}$ has two degrees of freedom, this series has the form

$$V(q_1, q_2) = V(0,0) + \left(\frac{\partial V}{\partial q_1} q_1 + \frac{\partial V}{\partial q_1} q_2\right) + \left(\frac{\partial^2 V}{\partial q_1^2} q_1^2 + 2\frac{\partial^2 V}{\partial q_1 \partial q_2} q_1 q_2 + \frac{\partial^2 V}{\partial q_2^2} q_2^2\right) + \cdots,$$

where all partial derivatives of V are evaluated at the expansion point $q_1 = q_1 = 0$. Now V has been selected so that $V(0,0) = 0$. Also, since $(0,0)$ is a stationary point of $V(q_1, q_2)$, it follows that $\partial V/\partial q_1 = \partial V/\partial q_2 = 0$ there. Thus the constant and linear terms are absent from the Taylor expansion of V. It follows that V can be approximated by

$$V^{\text{app}}(q_1, q_2) = v_{11} q_1^2 + 2 v_{12} q_1 q_2 + v_{22} q_2^2,$$

where v_{11}, v_{12}, v_{22} are constants given by

$$v_{11} = \frac{\partial^2 V}{\partial q_1^2}(0,0) \quad v_{12} = \frac{\partial^2 V}{\partial q_1 \partial q_2}(0,0) \quad v_{22} = \frac{\partial^2 V}{\partial q_2^2}(0,0)$$

and the neglected terms have power three (or higher) in the small quantities q_1, q_2. The corresponding approximation to $V(\boldsymbol{q})$ in the case when $\mathcal{S}$ has n-degrees of freedom is

$$\boxed{V^{\text{app}}(\boldsymbol{q}) = \sum_{j=1}^{n} \sum_{k=1}^{n} v_{jk}\, q_j q_k} \tag{15.8}$$

where the $\{v_{jk}\}$ are constants given by

$$v_{jk} = v_{kj} = \left.\frac{\partial^2 V}{\partial q_j \partial q_k}\right|_{\boldsymbol{q}=\mathbf{0}},$$

and the neglected terms have power three (or higher) in the small quantities $q_1, q_2, \ldots, q_n$. This is the general form of the **approximate potential energy** $V^{\text{app}}(\boldsymbol{q})$. It is a homogeneous quadratic form in the variables $q_1, q_2, \ldots, q_n$.

In the theory that follows, we will always assume that $\boldsymbol{q} = \mathbf{0}$ is also a minimum point of the *approximate* potential energy $V^{\text{app}}(\boldsymbol{q})$.* This condition is equivalent to requiring that the quadratic form (15.8) should be **positive definite**. This simply means that it takes positive values except when $\boldsymbol{q} = \mathbf{0}$.

The general approximate form of T

For any standard mechanical system with generalised coordinates $\boldsymbol{q}$, the kinetic energy T has the form

$$T(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \sum_{j=1}^{n}\sum_{k=1}^{n} t_{jk}(\boldsymbol{q})\, \dot{q}_j \dot{q}_k,$$

a quadratic form in the variables $\dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n$ with coefficients that depend on $\boldsymbol{q}$. If we expand each of these coefficients as a Taylor series about $\boldsymbol{q} = \mathbf{0}$, the constant term is simply $t_{jk}(\mathbf{0})$ and

$$T = \sum_{j=1}^{n}\sum_{k=1}^{n} t_{jk}(\mathbf{0})\, \dot{q}_j \dot{q}_k \ + \ \cdots .$$

It follows that T can be approximated by

$$\boxed{T^{\text{app}} = \sum_{j=1}^{n}\sum_{k=1}^{n} t_{jk}\, \dot{q}_j \dot{q}_k} \tag{15.9}$$

where the constants $\{t_{jk}\}$ are what we previously called $\{t_{jk}(\mathbf{0})\}$, and the neglected terms have power three (or higher) in the small quantities $q_1, q_2, \ldots, q_n, \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n$. This is the general form of the **approximate kinetic energy** $V^{\text{app}}(\boldsymbol{q})$. It is a homogeneous quadratic form in the variables $\dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n$. Since $T(\boldsymbol{q}, \dot{\boldsymbol{q}}) > 0$ except when $\dot{\boldsymbol{q}} = \mathbf{0}$, it follows that the quadratic form (15.9) must also be **positive definite**.

Useful tip: Since $T^{\text{app}}(\dot{\boldsymbol{q}}) = T(\mathbf{0}, \dot{\boldsymbol{q}})$, it follows that T^{app} can be be found directly by calculating T *when the system is passing through the equilibrium position*; the general formula for T need never be found!

---

* It might appear that this follows the fact that $\boldsymbol{q} = \mathbf{0}$ is known to be a minimum point of the *exact* V, but this is not necessarily so. For example, if $V(q_1, q_2) = q_1^2 + q_2^4$, then $V^{\text{app}} = q_1^2$, which does not have a *strict* minimum at $q_1 = q_2 = 0$. The general theory of small oscillations does not apply to such cases, and we exclude them.

The *V*-matrix and the *T*-matrix

In order to express the general theory concisely, we introduce the $n \times n$ matrices **V** and **T** as follows:

Definition 15.2 ***The V-matrix and the T-matrix*** *The symmetric $n \times n$ matrix* **V** *whose elements are the coefficients $\{v_{jk}\}$ that appear in the formula (15.8) is called the V* ***-matrix****. The symmetric $n \times n$ matrix* **T** *whose elements are the coefficients $\{t_{jk}\}$ that appear in the formula (15.9) is called the T* ***-matrix****.*

In terms of **V** and **T**, the approximate potential and kinetic energies of $\mathcal{S}$ can be written in compact matrix notation:*

Quadratic forms for V^{app} and T^{app}

$$
\begin{aligned}
V^{\text{app}} &= \sum_{j=1}^{n}\sum_{k=1}^{n} v_{jk}\, q_j q_k = \mathbf{q}' \cdot \mathbf{V} \cdot \mathbf{q} \\
T^{\text{app}} &= \sum_{j=1}^{n}\sum_{k=1}^{n} t_{jk}\, \dot{q}_j \dot{q}_k = \dot{\mathbf{q}}' \cdot \mathbf{T} \cdot \dot{\mathbf{q}}
\end{aligned}
\tag{15.10}
$$

where **q** is the column vector with elements $\{q_j\}$, and $\dot{\mathbf{q}}$ is the column vector with elements $\{\dot{q}_j\}$.

Example 15.2 *Finding* V *and* T *for the double pendulum*

Find the matrices **V** and **T** for the double pendulum.

Solution

For the double pendulum, V^{app} is given by

$$
\begin{aligned}
V^{\text{app}} &= \tfrac{1}{2}(M+m)gb\,\theta^2 + \tfrac{1}{2}mgc\,\phi^2, \\
&= \begin{pmatrix}\theta & \phi\end{pmatrix} \begin{pmatrix} \tfrac{1}{2}(M+m)gb & 0 \\ 0 & \tfrac{1}{2}mgc \end{pmatrix} \begin{pmatrix}\theta \\ \phi\end{pmatrix}
\end{aligned}
$$

and T^{app} is given by

$$
\begin{aligned}
T^{\text{app}} &= \tfrac{1}{2}(M+m)b^2\dot{\theta}^2 + mbc\,\dot{\theta}\dot{\phi} + \tfrac{1}{2}mc^2\dot{\phi}^2 \\
&= \begin{pmatrix}\dot{\theta} & \dot{\phi}\end{pmatrix} \begin{pmatrix} \tfrac{1}{2}(M+m)b^2 & \tfrac{1}{2}mbc \\ \tfrac{1}{2}mbc & \tfrac{1}{2}mc^2 \end{pmatrix} \begin{pmatrix}\dot{\theta} \\ \dot{\phi}\end{pmatrix}.
\end{aligned}
$$

* The notation $\mathbf{x}'$ means the *transpose* of the column vector **x**. The alternative notation $\mathbf{x}^T$ would cause confusion here.

Hence, $\mathbf{V}$ and $\mathbf{T}$ are the 2×2 matrices

$$\mathbf{V} = \begin{pmatrix} \frac{1}{2}(M+m)gb & 0 \\ 0 & \frac{1}{2}mgc \end{pmatrix}, \qquad \mathbf{T} = \begin{pmatrix} \frac{1}{2}(M+m)b^2 & \frac{1}{2}mbc \\ \frac{1}{2}mbc & \frac{1}{2}mc^2 \end{pmatrix}. \blacksquare$$

15.3 THE GENERAL THEORY OF NORMAL MODES

In this section, we develop the general theory of normal modes for any oscillating system. This extends the method described in Chapter 5, which was restricted to two degrees of freedom.

The small oscillation equations

The first step is to obtain the general form of the small oscillation equations. This is done by substituting the approximate potential and kinetic energies V^{app} and T^{app} into Lagrange's equations. Now

$$\frac{\partial T^{\text{app}}}{\partial \dot{q}_j} = 2\sum_{k=1}^{n} t_{jk}\,\dot{q}_k, \qquad \frac{\partial T^{\text{app}}}{\partial q_j} = 0, \qquad \frac{\partial V^{\text{app}}}{\partial q_j} = 2\sum_{k=1}^{n} v_{jk}\,q_k,$$

so that Lagrange's equations become

Small oscillation equations

Expanded form: $$\sum_{k=1}^{n} \left(t_{jk}\,\ddot{q}_k + v_{jk}\,q_k\right) = 0 \qquad (1 \le j \le n)$$

Matrix form: $$\mathbf{T} \cdot \ddot{\mathbf{q}} + \mathbf{V} \cdot \mathbf{q} = \mathbf{0}$$

(15.11)

in the expanded and matrix forms respectively. These are the **linearised equations** for the **small oscillations** of $\mathcal{S}$ about the point $\boldsymbol{q} = \mathbf{0}$. They are a set of *n coupled second order linear ODEs* satisfied by the unknown functions $q_1(t), q_2(t), \ldots, q_n(t)$.

Normal modes

The next step is to find a special class of solutions of the small oscillation equations known as **normal modes**. We will show later that the general solution of the small oscillation equations can be expressed as a sum of normal modes.

Definition 15.3 ***Normal mode*** *A solution of the small oscillation equations that has the special form*

Expanded form:	$q_j = a_j \cos(\omega t - \gamma)$ $(1 \le j \le n)$	(15.12)
Matrix form:	$\mathbf{q} = \mathbf{a} \cos(\omega t - \gamma)$	

where the $\{a_j\}$, ω *and* γ *are constants, is called a* ***normal mode*** *of the system* $\mathcal{S}$.

Notes. In a normal mode, the coordinates $q_1, q_2, \ldots, q_n$ all vary harmonically in time with the *same frequency* ω and the *same phase* γ; however, they generally have *different amplitudes* $a_1, a_2, \ldots, a_n$. The n-dimensional quantity $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$ is called the **amplitude vector** of the mode and, when considered to be a column vector, will be written $\mathbf{a}$. Without losing generality, the angular frequency ω can be assumed to be *positive*.

On substituting the normal mode form (15.12) into the small oscillation equations (15.11), we obtain, on cancelling by the common factor $\cos(\omega t - \gamma)$,

Equations for the amplitude vector

Expanded form:	$\sum_{k=1}^{n} \left(v_{jk} - \omega^2 t_{jk} \right) a_k = 0$ $(1 \le j \le n)$	(15.13)
Matrix form:	$\left(\mathbf{V} - \omega^2 \mathbf{T} \right) \cdot \mathbf{a} = \mathbf{0}$	

This is an $n \times n$ system *simultaneous linear algebraic equations* for the coordinate amplitudes $\{a_k\}$. A normal mode will exist if we can find constants $\{a_k\}$, ω so that the equations (15.13) are satisfied. Since the equations are homogeneous, they always have the *trivial solution* $a_1 = a_2 = \cdots = a_n = 0$, whatever the value of ω. However, the trivial solution corresponds to the *equilibrium solution* $q_1 = q_2 = \cdots = q_n = 0$ of the governing equations (15.11), which is not a motion at all. We therefore need the equations (15.13) to have a **non-trivial solution** for the $\{a_k\}$. There is a simple condition that this should be so, namely that the **determinant** of the system of equations should be **zero**, that is,

Determinantal equation for ω

$$\det \left(\mathbf{V} - \omega^2 \mathbf{T} \right) = 0 \qquad (15.14)$$

This is the equation satisfied by the angular frequency ω in any normal mode of the system $\mathcal{S}$. When expanded, this is a polynomial equation of degree n in the variable ω^2. *If this* *equation has any* ***real positive roots*** $\omega_1^2, \omega_2^2, \ldots$, *then, for each of these values of* ω, *the linear equations (15.13) will have a non-trivial solution for the amplitudes* $\{a_k\}$ *and a normal mode will exist.*

Definition 15.4 ***Normal frequencies*** *The angular frequencies* $\omega_1, \omega_2, \ldots$ *of the normal modes are called the* ***normal frequencies*** *of the system* $\mathcal{S}$.

The normal frequencies are a very important characteristic of an oscillating system. They are found by solving the determinantal equation (15.14) for ω. In the example that follows, we find the normal frequencies of the double pendulum, and three further worked examples are given in section 15.5.

Example 15.3 *Normal frequencies of the double pendulum*

Find the normal frequencies of the double pendulum for the case in which $M = 3m$ and $c = b$.

Solution

With these special values, the matrices **V** and **T** become

$$\mathbf{V} = \tfrac{1}{2}mgb\begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \mathbf{T} = \tfrac{1}{2}mb^2\begin{pmatrix} 4 & 1 \\ 1 & 1 \end{pmatrix}.$$

The **determinantal equation** for ω is therefore

$$\det\left[\tfrac{1}{2}mgb\begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix} - \tfrac{1}{2}mb^2\omega^2\begin{pmatrix} 4 & 1 \\ 1 & 1 \end{pmatrix}\right] = 0,$$

which can be simplified into the form

$$\begin{vmatrix} 4n^2 - 4\omega^2 & -\omega^2 \\ -\omega^2 & n^2 - \omega^2 \end{vmatrix} = 0,$$

where $n^2 = g/b$. On expanding this determinant, we obtain

$$3\omega^4 - 8n^2\omega^2 + 4n^4 = 0,$$

which is a quadratic equation in the variable ω^2. This is the equation satisfied by the normal frequencies. This quadratic factorises and has **two real positive roots** ω_1^2, ω_2^2 for ω^2, where

$$\omega_1^2 = \frac{2n^2}{3}, \qquad \omega_2^2 = 2n^2,$$

where $n^2 = g/b$. The double pendulum therefore has the **two normal frequencies** $(2g/3b)^{1/2}$ and $(2g/b)^{1/2}$. ■

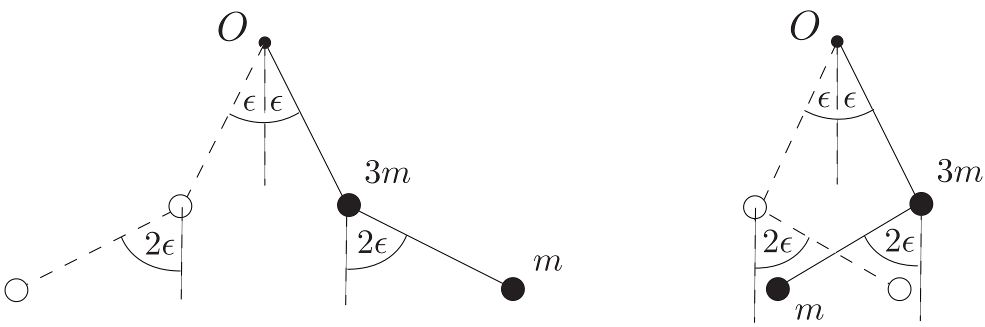


FIGURE 15.3 Normal modes of the double pendulum. **Left:** The slow mode. **Right:** The fast mode. (The angle ϵ, which should be small, is made large for clarity.)

Question *Form of the normal modes*

What do the normal mode motions of the double pendulum look like?

Answer

To answer this we need to find the coordinate amplitudes in each of the normal modes. If the amplitudes of θ and ϕ are a_1 and a_2 respectively, then these amplitudes satisfy the linear equations

$$\begin{pmatrix} 4n^2 - 4\omega^2 & -\omega^2 \\ -\omega^2 & n^2 - \omega^2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = 0.$$

Slow mode: When $\omega^2 = \omega_1^2 = 2n^2/3$, the equations for the amplitudes a_1, a_2 become, after simplification,

$$\begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = 0.$$

Each of these equations is equivalent to the single equation $2a_1 = a_2$ so that we have the family of non-trivial solutions $a_1 = \epsilon$, $a_2 = 2\epsilon$, where ϵ can take any (non-zero) value. There is therefore just **one slow normal mode**. It has the form

$$\begin{aligned} \theta &= \epsilon \cos(\sqrt{2/3}\, nt - \gamma), \\ \phi &= 2\epsilon \cos(\sqrt{2/3}\, nt - \gamma), \end{aligned} \tag{15.15}$$

where the amplitude factor ϵ and phase factor γ can take any values.* This mode is shown in Figure 15.3 (left).

Fast mode: In the fast mode, we have $\omega^2 = 2n^2$ and, by following the same procedure, we find that there is also **one fast normal mode**. It has the form

$$\begin{aligned} \theta &= \epsilon \cos(\sqrt{2}\, nt - \gamma), \\ \phi &= -2\epsilon \cos(\sqrt{2}\, nt - \gamma), \end{aligned} \tag{15.16}$$

where the amplitude factor ϵ and phase factor γ can take any values. This mode is shown in Figure 15.3 (right). ■

* However, the linearised theory is a good approximation only when ϵ is small.

15.4 EXISTENCE THEORY FOR NORMAL MODES

So far we have not said anything general about the *number* of normal mode motions that a system posesses. This is related to the number of *real positive* roots of the equation

$$\det(\mathbf{V} - \lambda\,\mathbf{T}) = 0. \tag{15.17}$$

When expanded, this is a polynomial equation of degree n in the variable λ, where n is the number of degrees of freedom of the system $\mathcal{S}$. Such an equation always has n roots in the *complex* plane, but there seems to be no reason why any of them should be real, let alone positive. In fact, *all of the roots of equation (15.17) are real and positive*. This follows from **generalised eigenvalue theory**, which we will now develop.

Definition 15.5 ***Generalised eigenvalues and eigenvectors*** *Let* $\mathbf{K}$ *and* $\mathbf{L}$ *be real* $n \times n$ *matrices. If there exists a number* λ *and a (non-zero) column vector* $\mathbf{x}$ *such that*

$$\mathbf{K} \cdot \mathbf{x} = \lambda\,\mathbf{L} \cdot \mathbf{x}, \tag{15.18}$$

then λ *is said to be a* ***generalised eigenvalue*** *of the matrix* $\mathbf{K}$ *(with respect to the matrix* $\mathbf{L}$*) and* $\mathbf{x}$ *is a corresponding* ***generalised eigenvector***.*

The defining equation (15.18) can also be written

$$(\mathbf{K} - \lambda\,\mathbf{L}) \cdot \mathbf{x} = \mathbf{0}, \tag{15.19}$$

which has a non-zero solution for $\mathbf{x}$ only when

$$\det(\mathbf{K} - \lambda\,\mathbf{L}) = 0. \tag{15.20}$$

This is the equation satisfied by the eigenvalues. Provided that $\mathbf{L}$ is a non-singular matrix, the eigenvalue equation is a polynomial equation of degree n in λ, which has n roots in the complex plane. The more one knows about the matrices $\mathbf{K}$ and $\mathbf{L}$, the more one can say about their eigenvalues and eigenvectors.

Theorem 15.1 ***Eigenvalues of symmetric, positive definite matrices*** *If* $\mathbf{K}$ *and* $\mathbf{L}$ *are* ***real symmetric*** *matrices and* $\mathbf{L}$ *is* ***positive definite***,† *then all the eigenvalues are* ***real***. *If the matrix* $\mathbf{K}$ *is also* ***positive definite***, *then all the eigenvalues are* ***positive***.

Proof. Let $\mathbf{x}$ be *any* complex column vector and consider the scalar quantity $\overline{\mathbf{x}}' \cdot \mathbf{K} \cdot \mathbf{x}$, where $\overline{\mathbf{x}}$ is the complex conjugate of $\mathbf{x}$. Then, since $\mathbf{K}$ is real and symmetric,

$$\begin{aligned}\overline{\overline{\mathbf{x}}' \cdot \mathbf{K} \cdot \mathbf{x}} &= \mathbf{x}' \cdot \overline{\mathbf{K}} \cdot \overline{\mathbf{x}} = \mathbf{x}' \cdot \mathbf{K} \cdot \overline{\mathbf{x}} = \left(\mathbf{x}' \cdot \mathbf{K} \cdot \overline{\mathbf{x}}\right)' = \overline{\mathbf{x}}' \cdot \mathbf{K}' \cdot \mathbf{x}\\ &= \overline{\mathbf{x}}' \cdot \mathbf{K} \cdot \mathbf{x}.\end{aligned}$$

* Ordinary eigenvalues and eigenvectors correspond to the special case when $\mathbf{L} = \mathbf{1}$, the identity matrix.

† A *matrix* $\mathbf{A}$ is called **positive definite** if its associated quadratic form $\mathbf{x}' \cdot \mathbf{A} \cdot \mathbf{x}$ is positive definite. Since this condition is known to hold for $\mathbf{V}$ and $\mathbf{T}$, both these matrices must be positive definite.

Hence $\overline{\mathbf{x}}' \cdot \mathbf{K} \cdot \mathbf{x}$ must be real, and, by a similar argument, $\overline{\mathbf{x}}' \cdot \mathbf{L} \cdot \mathbf{x}$ must also be real.

Also, if we write $\mathbf{x}$ in terms of its real and imaginary parts in the form $\mathbf{x} = \mathbf{u} + i\,\mathbf{v}$, then

$$\begin{aligned}\overline{\mathbf{x}}' \cdot \mathbf{L} \cdot \mathbf{x} &= (\mathbf{u} - i\,\mathbf{v})' \cdot \mathbf{L} \cdot (\mathbf{u} + i\,\mathbf{v})\\ &= \mathbf{u}' \cdot \mathbf{L} \cdot \mathbf{u} + \mathbf{v}' \cdot \mathbf{L} \cdot \mathbf{v} + i\,(\mathbf{u}' \cdot \mathbf{L} \cdot \mathbf{v} - \mathbf{v}' \cdot \mathbf{L} \cdot \mathbf{u})\\ &= \mathbf{u}' \cdot \mathbf{L} \cdot \mathbf{u} + \mathbf{v}' \cdot \mathbf{L} \cdot \mathbf{v}\end{aligned}$$

since $\overline{\mathbf{x}}' \cdot \mathbf{L} \cdot \mathbf{x}$ is known to be real. Since $\mathbf{L}$ is a positive definite matrix, it follows that $\mathbf{u}' \cdot \mathbf{L} \cdot \mathbf{u}$ is positive except when $\mathbf{u} = \mathbf{0}$, and $\mathbf{v}' \cdot \mathbf{L} \cdot \mathbf{v}$ is positive except when $\mathbf{v} = \mathbf{0}$. Hence $\overline{\mathbf{x}}' \cdot \mathbf{L} \cdot \mathbf{x}$ is positive except when $\mathbf{x} = \mathbf{0}$.

Now suppose that $\mathbf{x}$ is a complex eigenvector corresponding to the complex eigenvalue λ. Then

$$\begin{aligned}\overline{\mathbf{x}}' \cdot \mathbf{K} \cdot \mathbf{x} &= \overline{\mathbf{x}}' \cdot (\mathbf{K} \cdot \mathbf{x}) = \overline{\mathbf{x}}' \cdot (\lambda\, \mathbf{L} \cdot \mathbf{x})\\ &= \lambda\, \left(\overline{\mathbf{x}}' \cdot \mathbf{L} \cdot \mathbf{x}\right).\end{aligned}$$

But $\overline{\mathbf{x}}' \cdot \mathbf{K} \cdot \mathbf{x}$ is known to be real and $\overline{\mathbf{x}}' \cdot \mathbf{L} \cdot \mathbf{x}$ is known to be real and positive (since the complex eigenvector $\mathbf{x}$ is not zero). Hence the eigenvalue λ must be real.

The eigenvalues are now known to be real, and we may therefore restrict the eigenvectors to be real too. Suppose that the real eigenvalue λ has real eigenvector $\mathbf{x}$ and that the matrix $\mathbf{K}$ is now also positive definite. Then

$$\mathbf{x}' \cdot \mathbf{K} \cdot \mathbf{x} = \lambda\, \left(\mathbf{x}' \cdot \mathbf{L} \cdot \mathbf{x}\right).$$

But, since $\mathbf{K}$ and $\mathbf{L}$ are both positive definite matrices, the quantities $\mathbf{x}' \cdot \mathbf{K} \cdot \mathbf{x}$ and $\mathbf{x}' \cdot \mathbf{L} \cdot \mathbf{x}$ are both *positive*. It follows that λ must also be positive. ■

Since the matrices $\mathbf{V}$ and $\mathbf{T}$ are both symmetric and positive definite, the above theorem applies to normal mode theory. It follows that the roots of the determinantal equation (15.17) are all real and positive. If these roots are distinct (the most common case), then there are n distinct **normal frequencies** $\omega_1, \omega_2, \dots, \omega_n$. It is however possible for the determinantal equation (15.14) to have repeated roots, so that there are fewer than n distinct normal frequencies. This usually happens when the system has symmetry; the spherical pendulum oscillating about the downward vertical is a simple example.

The number of normal modes associated with a particular normal frequency, ω_1 (say), depends on whether ω_1^2 is a simple or repeated root of the eigenvalue equation (15.17). It can be proved that, if ω_1^2 is a *simple root*, then the equations (15.13) for the amplitude vector $\boldsymbol{a}$ have a non-trivial solution that is *unique* to within a multiplied constant. There is therefore only *one normal mode* associated with the normal frequency ω_1. More generally, it can be proved that, if the root is repeated k times, then the equations (15.13) for the amplitude vector $\boldsymbol{a}$ have k linearly independent solutions.* The normal frequency then has k normal modes associated with it instead of one. It follows that, in all cases, we have the **fundamental result** that *the total number of normal modes is always equal to n, the number of degrees of freedom of the system.*

* It follows from the orthogonality relations (see section 15.6) that the amplitude vectors of the normal modes must be linearly independent and therefore cannot exceed n in number. Hence, when there are n *distinct* normal frequencies, each frequency must have exactly *one normal mode* associated with it. The corresponding result in the degenerate case is not easy to prove and is beyond the scope of a mechanics text. (See Anton [7] and Lang [10].)

Suppose for example that the oscillating system has six degrees of freedom and that the determinantal equation (15.14) is

$$(\Omega^2 - \omega^2)^2\,(4\Omega^2 - \omega^2)^3\,(25\Omega^2 - \omega^2) = 0,$$

after factorisation, where Ω is a positive constant. The normal frequencies are then $\omega_1 = \Omega$ (double root), $\omega_2 = 2\Omega$ (triple root) and $\omega_3 = 5\Omega$ (simple root). There are therefore two normal modes associated with the normal frequency ω_1, three normal modes associated with the normal frequency ω_2, and one normal mode associated with the normal frequency ω_3. The total number of normal modes is six, which is equal to the number of degrees of the system.

Definition 15.6 ***Degenerate frequencies*** *If a normal frequency has more than one normal mode associated with it, then that frequency is said to be* ***degenerate****.*

In the example above, the normal frequencies ω_1 and ω_2 are degenerate, but ω_3 is not. The notion of degeneracy is important in **quantum mechanics**, where normal frequencies correspond to the energies of stationary states. An unperturbed atom may have an energy level E that is (say) five-fold degenerate. When the atom is perturbed (by a magnetic field, for example) the energies of five states may be changed by differing amounts so that the energy level is 'split' into five nearly equal levels. This is an important effect in the theory of atomic spectra.

Existence of normal modes

- For any oscillating system, the roots of the eigenvalue equation

$$\det\,(\mathsf{V} - \lambda\,\mathsf{T}) = 0$$

  are all real and positive and their values are the *squares* of the normal frequencies $\{\omega_j\}$ of the system.
- If ω_1^2 is a *simple* root of the eigenvalue equation, then the equations

$$\left(\mathsf{V} - \omega_1^2\,\mathsf{T}\right)\cdot \mathbf{a} = \mathbf{0}$$

  for the amplitude vector $\boldsymbol{a}$ have a non-trivial solution that is *unique* to within a multiplied constant. There is therefore only *one normal mode* associated with the normal frequency ω_1.
- More generally, if the root ω_1^2 is *repeated* k times, then the equations for the amplitude vector $\boldsymbol{a}$ have k linearly independent solutions. The normal frequency ω_1 is then degenerate with k normal modes associated with it.
- In all cases, **an oscillating system with n degrees of freedom has a total of n normal modes**.

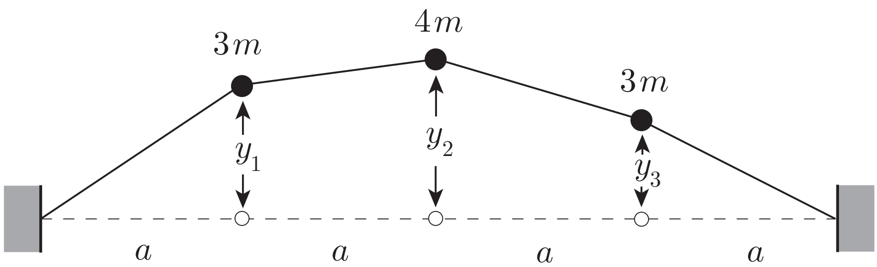


FIGURE 15.4 Transverse oscillations of three particles attached to a stretched string. (The particle displacments, which should be small, are shown large for clarity.)

15.5 SOME TYPICAL NORMAL MODE PROBLEMS

The determination of normal modes of vibration is an important subject with applications in physics, chemistry and mechanical engineering. The following three problems are typical. The first involves the transverse oscillations of a loaded stretched string; such problems make popular examination questions!

Example 15.4 *Transverse oscillations of a loaded stretched string*

A light string is stretched to a tension T_0 between two fixed points a distance $4a$ apart and particles of masses $3m$, $4m$ and $3m$ are attached to the string at equal intervals, as shown in Figure 15.4. The system performs small plane oscillations in which the particles move *transversely*, that is, at right angles to the equilibrium line of the string. Find the frequencies and forms of the normal modes.

Solution

Although it is clear by symmetry that *purely longitudinal* modes exist, it is not obvious that *purely transverse* modes exist. This question is investigated in Problem 15.5, where it is shown that the longitudinal and transverse modes uncouple in the linear theory. In this setting, purely transverse modes do exist and can be found by setting the longitudinal displacements equal to zero, which is what we will do here.

Let the transverse displacements of the three particles be y_1, y_2, y_3, as shown. Then the extension Δ_1 of the first section of string is given by

$$\begin{aligned}\Delta_1 &= \left(a^2 + y_1^2\right)^{1/2} - a = a\left(1 + \frac{y_1^2}{a^2}\right)^{1/2} - a\\ &= \frac{y_1^2}{2a} + \cdots ,\end{aligned}$$

where the neglected terms have power four (or higher) in the small quantity y_1. Consider now the potential energy V_1 of this section of string. If the string had no initial tension then V_1 would be given by $V_1 = \frac{1}{2}\alpha\Delta_1^2$, where α is the 'spring constant' of the first section of the string. However, since there is an initial tension T_0, the formula is modified to

$$V_1 = T_0\Delta_1 + \tfrac{1}{2}\alpha\Delta_1^2 = T_0\left(\frac{y_1^2}{2a} + \cdots\right) + \tfrac{1}{2}\alpha\left(\frac{y_1^2}{2a} + \cdots\right)^2$$
$$= \frac{T_0 y_1^2}{2a} + \cdots ,$$

where the neglected terms have power four (or higher) in the small quantity y_1. Note that, in the quadratic approximation, the spring constant α does not appear so that the increase in tension of the string is negligible.

In the same way, the potential energies of the other three sections of string are given by

$$V_2 = \frac{T_0(y_2 - y_1)^2}{2a} + \cdots , \quad V_3 = \frac{T_0(y_3 - y_2)^2}{2a} + \cdots , \quad V_4 = \frac{T_0 y_3^2}{2a} + \cdots ,$$

and the total **approximate potential energy** is given by

$$V^{\text{app}} = \frac{T_0 y_1^2}{2a} + \frac{T_0(y_2 - y_1)^2}{2a} + \frac{T_0(y_3 - y_2)^2}{2a} + \frac{T_0 y_3^2}{2a}$$
$$= \frac{T_0}{2a}\left(2y_1^2 + 2y_2^2 + 2y_3^2 - 2y_1 y_2 - 2y_2 y_3\right).$$

The V-matrix is therefore

$$\mathbf{V} = \frac{T_0}{2a}\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}.$$

In this problem, the exact and approximate **kinetic energies** are the same, namely

$$T = T^{\text{app}} = \tfrac{1}{2}(3m)\dot{y}_1^2 + \tfrac{1}{2}(4m)\dot{y}_2^2 + \tfrac{1}{2}(3m)\dot{y}_3^2,$$

so that the T-matrix is

$$\mathbf{T} = \tfrac{1}{2}m\begin{pmatrix} 3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

The **eigenvalue equation** $\det(\mathbf{V} - \lambda\mathbf{T}) = 0$ can therefore be written

$$\begin{vmatrix} 2 - 3\mu & -1 & 0 \\ -1 & 2 - 4\mu & -1 \\ 0 & -1 & 2 - 3\mu \end{vmatrix} = 0,$$

where $\mu = ma\omega^2/T_0$. When expanded, this is the **cubic equation**

$$18\mu^3 - 33\mu^2 + 17\mu - 2 = 0$$

for the parameter μ. Such an equation would have to be solved numerically in general, but problems that appear in textbooks (and in examinations!) are usually contrived so that an *exact factorisation* is possible; this is true in the present problem. Sometimes a factor can be spotted while the cubic is still in determinant form. In the present case, one can see that, by subtracting the third row of the determinant from the first, the cubic has the factor $2 - 3\mu$. If no factor can be spotted in this way, then one must try to spot that the expanded cubic equation has a (hopefully small) integer root. In the present case, one would have to spot that $\mu = 1$ is a root of the expanded cubic. By using either method, our cubic equation factorises into

$$(6\mu - 1)(3\mu - 2)(\mu - 1) = 0,$$

and its roots are $\mu_1 = 1/6$, $\mu_2 = 2/3$, $\mu_3 = 1$. Since $\mu = ma\omega^2/T_0$, the **normal frequencies** are given by

$$\omega_1^2 = \frac{T_0}{6ma}, \qquad \omega_2^2 = \frac{2T_0}{3ma}, \qquad \omega_3^2 = \frac{T_0}{ma}.$$

Since the normal frequencies are non-degenerate, the corresponding amplitude vectors are unique to within multiplied constants. In the **slow mode**, $\mu = 1/6$ and the equations $(\mathbf{V} - \lambda\mathbf{T}) \cdot \mathbf{a} = \mathbf{0}$ for the amplitude vector $\mathbf{a}$ become

$$\begin{pmatrix} 9 & -6 & 0 \\ -6 & 8 & -6 \\ 0 & -6 & 9 \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

on clearing fractions. It is evident that $a_1 = 2$, $a_2 = 3$, $a_3 = 2$ is a solution so that the amplitude vector for the mode with frequency ω_1 is $\boldsymbol{a}_1 = (2, \ 3, \ 2)$. The other modes are treated in a similar way and the amplitude vectors are given by

$$\mathbf{a}_1 = \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix}, \qquad \mathbf{a}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \qquad \mathbf{a}_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

These are the **forms** of the three normal modes. (It is a good idea to sketch the shapes of the three modes.) ■

Our second example is concerned with the **internal vibrations of molecules**, an important subject in physical chemistry. Although such problems should really be treated using quantum mechanics, the classical theory of normal modes gives much valuable information with far less effort. It would also be very difficult to understand the quantum treatment of molecular vibrations without first having studied the classical theory. The simplest case in which there is more than one frequency is the **linear triatomic molecule**. In this case the three atoms lie in a straight line and can perform *rectilinear* oscillations. The classic example of a linear triatomic molecule is carbon dioxide.

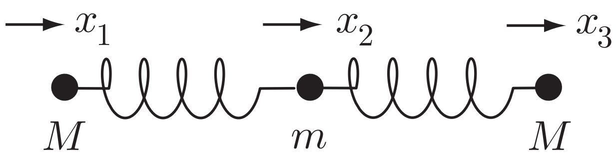


FIGURE 15.5 A simple classical model of a linear symmetric molecule.

Example 15.5 *The linear triatomic molecule*

A symmetric linear triatomic molecule is modelled by three particles connected by two springs, arranged as shown in Figure 15.5. Find the frequencies of rectilinear vibration of the molecule and the forms of the normal modes.

Solution

Let the centre atom have mass m, the outer atoms have mass M, and the springs have constant α. In this context, the spring constant is a measure of the 'strength' of the chemical bond between the two atoms. (We will suppose that the interaction between the two outer atoms is negligible.) Let the displacements of the three atoms from their equilibrium positions be x_1, x_2, x_3 as shown in Figure 15.5. Then the kinetic and potential energies of the molecule are given by

$$T = T^{\text{app}} = \tfrac{1}{2}M\dot{x}_1^2 + \tfrac{1}{2}m\dot{x}_2^2 + \tfrac{1}{2}M\dot{x}_3^2,$$

$$\begin{aligned} V = V^{\text{app}} &= \tfrac{1}{2}\alpha(x_2 - x_1)^2 + \tfrac{1}{2}\alpha(x_3 - x_2)^2 \\ &= \tfrac{1}{2}\alpha\left(x_1^2 + 2x_2^2 + x_3^2 - 2x_1x_2 - 2x_2x_3.\right) \end{aligned}$$

The T- and V- matrices are therefore

$$\mathbf{T} = \tfrac{1}{2}\begin{pmatrix} M & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & M \end{pmatrix}, \qquad \mathbf{V} = \tfrac{1}{2}\alpha\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}.$$

Although everything looks normal, this problem has a non-standard feature in that the potential energy of the molecule does not have a true minimum* at $x_1 = x_2 = x_3 = 0$. This is actually clear from the start since the potential energy is unchanged if the whole molecule is translated to the right or left. Strictly speaking then, our theory does *not* apply to this problem, but, fortunately, only minor modifications are needed. In cases like this, it turns out that one or more of the normal frequencies is zero. These *zero frequencies do not correspond to true normal modes*. They correspond to uniform translational (or rotational) motions of the molecule as a whole. In the present case, the only uniform motion allowed is translational motion along the line of the molecule, so that we expect just *one* of the normal frequencies to be zero.

* As a result, the V-matrix is not positive definite.

The **eigenvalue equation** $\det(\mathbf{V} - \lambda\mathbf{T}) = 0$ can be written

$$\begin{vmatrix} 1-\mu & -1 & 0 \\ -1 & 2-\gamma^{-1}\mu & -1 \\ 0 & -1 & 1-\mu \end{vmatrix} = 0,$$

where $\mu = M\omega^2/\alpha$ and $\gamma = M/m$. The roots of this cubic equation are easily found to be $\mu = 0$, $\mu = 1$ and $\mu = 1+2\gamma$. The zero root corresponds to uniform translation and the other two are genuine oscillatory modes with **vibrational frequencies** given by

$$\omega_1^2 = \frac{\alpha}{M}, \qquad \omega_2^2 = \frac{(1+2\gamma)\alpha}{M}.$$

The **amplitude vectors** corresponding to the vibrational modes are

$$\mathbf{a}_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \qquad \mathbf{a}_2 = \begin{pmatrix} 1 \\ -2\gamma \\ 1 \end{pmatrix},$$

respectively. We see that, in the slow ω_1-mode, the centre particle remains at rest while the outer pair oscillate symmetrically about the centre. This is called the **symmetric stretching mode** of the molecule. The fast ω_2-mode, in which all three particles move with the outer atoms remaining a constant distance apart, is called the **antisymmetric stretching mode**.

Comparison with experiment

Since the value of the constant α is unspecified, it can be chosen to fit any observed frequency. However, the *frequency ratio* ω_2/ω_1 $\left(= (1+2\gamma)^{1/2}\right)$ is independent of α and therefore affords a check on the theory.

The vibrational frequencies of real molecules can be measured with great accuracy by infrared and Raman spectroscopy. Spectroscopists measure the wavelength λ of radiation that excites each vibrational mode. The mode frequency is proportional to the reciprocal wavelength λ^{-1}, the standard units being cm^{-1}. Table 2 compares the observed and theoretical results for **carbon dioxide** and **carbon disulphide**,* both of which have linear symmetric molecules.

The theoretical values of the frequency ratio ω_2/ω_1 are within about 8% of those measured experimentally, which, considering the simplicity of the theory, is very good agreement.

Some more examples on vibrating molecules are given in the problems at the end of the chapter. The standard reference on this subject is the monumental work of Herzberg [13], Volume II. ■

Our third example involves a **rigid body suspended by three strings**. Problems of this type tend to be difficult because the constraints make it difficult to calculate the

* The atomic weights of carbon, oxygen and sulphur are C = 12, O = 16, S = 32.

Molecule	λ_1^{-1} (cm^{-1})	λ_2^{-1} (cm^{-1})	$\lambda_2^{-1}/\lambda_1^{-1}$	$(1+2\gamma)^{1/2}$
O–C–O	1337	2349	1.76	1.91
S–C–S	657	1532	2.33	2.52

Table 2 Vibrational frequencies of linear triatomic molecules; comparison of theory and experiment.

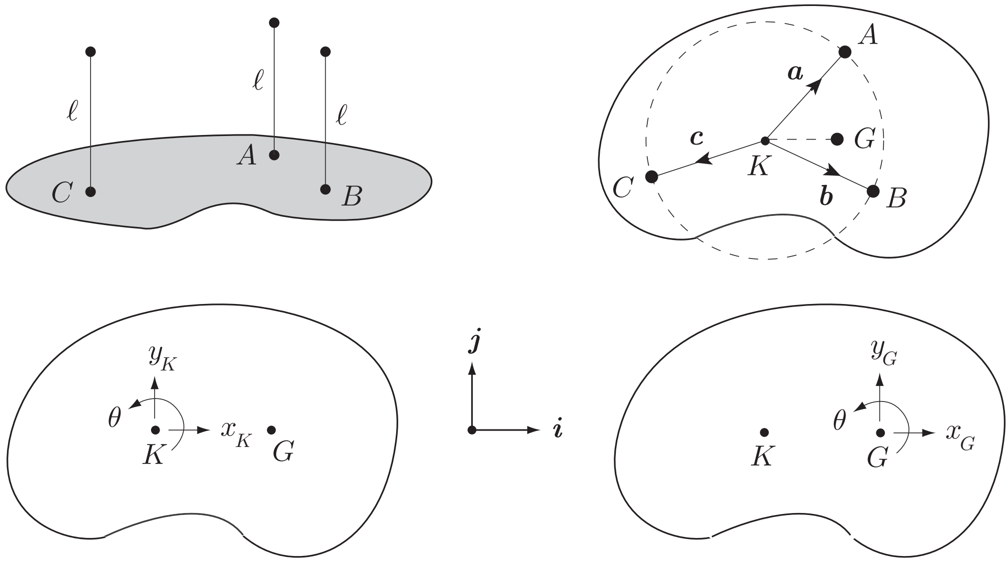


FIGURE 15.6 A flat plate of general shape is suspended by three equal strings.

potential energy. It is all the more surprising then that the following problem has a simple exact solution.

Example 15.6 *Plate supported by three strings*

A flat plate of general shape and general mass distribution is suspended in a horizontal position by three vertical strings of equal length. Find the normal frequencies of small oscillation.

Solution

The system is shown in its equilibrium position in Figure 15.6, top left. The three strings are of length ℓ and are attached to the points A, B, C of the plate. The point K (see Figure 15.6, top right) is the centre of the circle that passes through A, B, C. Our initial choice of generalised coordinates is shown in Figure 15.6, bottom left. X_K, Y_K are the horizontal Cartesian* displacement components of the point K from

* These can be any Cartesian axes K_0xyz with K_0z pointing vertically upwards; $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ are the corresponding unit vectors.

its equilibrium position K_0, and θ is the rotation angle of the plate about the vertical axis through K_0.* Three coordinates are sufficient since the three string constraints reduce the number of degrees of freedom of the plate from six to three.

We will now calculate the **potential energy** of the plate in terms of the coordinates X_K, Y_K and θ. This is the tricky step since the plate does *not* remain horizontal and this complicates the geometry. However, the vertical displacement of any point of the plate is *quadratic* in the small quantities X_K, Y_K, θ and, providing care is taken, this enables us to use approximations. Consider first $\Delta \boldsymbol{a}$, the displacement of the point A. This is given approximately by

$$\Delta \boldsymbol{a} = X_K \boldsymbol{i} + Y_K \boldsymbol{j} + (\theta \boldsymbol{k}) \times \boldsymbol{a} + \cdots ,$$

correct to the *first* order in small quantities, where $\boldsymbol{a}$ is the position vector of A relative to K_0 in the equilibrium position. As expected, this displacement is horizontal, correct to the first order in small quantities. The square of the magnitude of this horizontal displacement is therefore

$$\begin{aligned}(X_K \boldsymbol{i} + Y_K \boldsymbol{j} + \theta \boldsymbol{k} \times \boldsymbol{a})^2 &\\ &= X_K^2 + Y_K^2 + \theta^2 |\boldsymbol{a}|^2 + 2X_K\theta\, (\boldsymbol{i} \cdot (\boldsymbol{k} \times \boldsymbol{a})) + 2Y_K\theta\, (\boldsymbol{j} \cdot (\boldsymbol{k} \times \boldsymbol{a}))\\ &= X_K^2 + Y_K^2 + R^2\theta^2 + 2Y_K\theta(\boldsymbol{a} \cdot \boldsymbol{i}) - 2X_K\theta(\boldsymbol{a} \cdot \boldsymbol{j}),\end{aligned}$$

correct to the *second* order in small quantities, where R $(= |\boldsymbol{a}|)$ is the radius the circle passing through A, B and C. Since A is one of the points that is suspended by a string of length ℓ, an application of Pythagoras shows that the *vertical* displacement of the point A is given by

$$z_A = \left(\frac{X_K^2 + Y_K^2 + R^2\theta^2}{2\ell}\right) + \left(\frac{2Y_K\theta}{2\ell}\right)(\boldsymbol{a} \cdot \boldsymbol{i}) - \left(\frac{2X_K\theta}{2\ell}\right)(\boldsymbol{a} \cdot \boldsymbol{j}),$$

correct to the *second* order in small quantities. Similar expressions exist for z_B and z_C with $\boldsymbol{a}$ replaced by $\boldsymbol{b}$ and $\boldsymbol{c}$ respectively. Now for the clever bit. Since the plate is flat, it follows that a *general* point of the plate with position vector $\boldsymbol{r}$ relative to K_0 in the equilibrium position must have vertical displacement†

$$z = \left(\frac{X_K^2 + Y_K^2 + R^2\theta^2}{2\ell}\right) + \left(\frac{2Y_K\theta}{2\ell}\right)(\boldsymbol{r} \cdot \boldsymbol{i}) - \left(\frac{2X_K\theta}{2\ell}\right)(\boldsymbol{r} \cdot \boldsymbol{j}) \qquad (15.21)$$

in the displaced position, correct to the second order in small quantities. This expression confirms that the plate is not generally horizontal in the displaced position.

* As we will see, the plate does *not* remain horizontal and so the angle θ ought to be defined more carefully. Let KP be any line fixed in the plate and let the projection of this line on to the equilibrium plane of the plate be $K'P'$. The angle θ can be properly defined as the angle turned through by the line $K'P'$. This is not quite the same as the angle turned through by the projection of some other line lying in the plate, but the differences are quadratic in the small quantities X_K, Y_K and θ and will turn out to be immaterial.

† This is because the expression (15.21) is linear in $\boldsymbol{r}$ and gives the correct z-values at A, B, C.

The purpose of all this is to calculate the potential energy $V = Mgz_G$, where M is the mass of the plate and z_G is the vertical displacement of the centre of mass G. With no loss of generality we may take the axis K_0x to point in the direction $\overrightarrow{K_0G_0}$ so that $\boldsymbol{r}_G = D\boldsymbol{i}$, where D is the distance KG. The general formula (15.21) then shows that

$$\begin{aligned} z_G &= \left(\frac{X_K^2 + Y_K^2 + R^2\theta^2}{2\ell}\right) + \left(\frac{2Y_K\theta}{2\ell}\right) D \\ &= \frac{X_K^2 + (Y_K + D\,\theta)^2 + (R^2 - D^2)\theta^2}{2\ell}, \end{aligned}$$

correct to the second order in small quantities. Hence the approximate **potential energy** is given by

$$V^{\text{app}} = \frac{Mg}{2\ell}\left[X_K^2 + (Y_K + D\,\theta)^2 + (R^2 - D^2)\theta^2\right].$$

This formula can be simplified further by a change of generalised coordinates. Let X_G, Y_G be the horizontal Cartesian displacement components of the centre of mass G from its equilibrium position G_0, and θ_G be the rotation angle of the plate about the vertical axis through G_0. Then, correct to the *first* order in small quantities,

$$X_G = X_K, \qquad Y_G = Y_K + D\,\theta, \qquad \theta_G = \theta.$$

In terms of the generalised coordinates X_G, Y_G, θ (we will drop the subscript from θ_G from now on) the expression for V^{app} is

$$V^{\text{app}} = \frac{Mg}{2\ell}\left(X_G^2 + Y_G^2 + (R^2 - D^2)\,\theta^2\right).$$

and the V-matrix is

$$\mathbf{V} = \frac{Mg}{2\ell}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & R^2 - D^2 \end{pmatrix},$$

a remarkably simple result in the end.

The approximate **kinetic energy** is calculated when the plate is passing through its equilibrium position. This is simply

$$T^{\text{app}} = \tfrac{1}{2}M\dot{X}_G^2 + \tfrac{1}{2}M\dot{Y}_G^2 + \tfrac{1}{2}I\,\dot{\theta}^2,$$

where I is the moment of inertia of the plate about the axis through G perpendicular to its plane. If we write $I = Mk^2$, then the T-matrix becomes

$$\mathbf{T} = \tfrac{1}{2}M\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & k^2 \end{pmatrix}.$$

The **normal frequencies** are therefore given by

$$\omega_1^2 = \frac{g}{\ell} \quad \text{(doubly degenerate)}, \qquad \omega_2^2 = \left(\frac{R^2 - D^2}{k^2}\right)\frac{g}{\ell},$$

which correspond to two translational modes and one rotational mode.

In particular, if the lamina is a uniform circular **ring** of radius a, then $R = a$, $D = 0$ and $k = a$. Then $\omega_2 = \omega_1$ and the system has the single triply degenerate normal frequency $(g/\ell)^{1/2}$. In this case, any small motion of the system is periodic with period $2\pi(\ell/g)^{1/2}$. ■

15.6 ORTHOGONALITY OF NORMAL MODES

In this section we will show that the n **amplitude vectors** of the normal modes of an oscillating system are mutually orthogonal, in a sense that we will make clear. This is an important theoretical result, but it is not needed in the practical solution of normal mode problems. We will make use of orthogonality in our treatment of normal coordinates in section 15.8.

The basic theorem on orthogonality of eigenvectors is as follows:

Theorem 15.2 ***Orthogonality of eigenvectors*** *Suppose* K *and* L *are* ***symmetric*** $n \times n$ *matrices and that* $\mathbf{x}_1$ *and* $\mathbf{x}_2$ *are generalised eigenvectors of* K *(with respect to the matrix* L*) belonging to* ***distinct eigenvalues****. Then* $\mathbf{x}_1$ *and* $\mathbf{x}_2$ *are* ***mutually orthogonal*** *(with respect to the matrix* L*) in the sense that they satisfy the relation*

$$\mathbf{x}_1' \cdot \mathsf{L} \cdot \mathbf{x}_2 = 0.$$

Proof. Suppose that λ_1 and λ_2 are distinct eigenvalues with the corresponding eigenvectors $\mathbf{x}_1$ and $\mathbf{x}_2$ respectively. Consider the scalar quantity

$$\begin{aligned}\mathbf{x}_1' \cdot \mathsf{K} \cdot \mathbf{x}_2 &= \mathbf{x}_1' \cdot (\mathsf{K} \cdot \mathbf{x}_2) = \mathbf{x}_1' \cdot (\lambda_2 \mathsf{L} \cdot \mathbf{x}_2)\\ &= \lambda_2\,(\mathbf{x}_1' \cdot \mathsf{L} \cdot \mathbf{x}_2).\end{aligned}$$

However, the same quantity can also be written

$$\begin{aligned}\mathbf{x}_1' \cdot \mathsf{K} \cdot \mathbf{x}_2 &= (\mathsf{K}' \cdot \mathbf{x}_1)' \cdot \mathbf{x}_2 = (\mathsf{K} \cdot \mathbf{x}_1)' \cdot \mathbf{x}_2 = (\lambda_1 \mathsf{L} \cdot \mathbf{x}_1)' \cdot \mathbf{x}_2 = \lambda_1(\mathbf{x}_1' \cdot \mathsf{L}' \cdot \mathbf{x}_2)\\ &= \lambda_1\,(\mathbf{x}_1' \cdot \mathsf{L} \cdot \mathbf{x}_2),\end{aligned}$$

since K and L are both symmetric. It follows that

$$(\lambda_1 - \lambda_2)\left(\mathbf{x}_1' \cdot \mathsf{L} \cdot \mathbf{x}_2\right) = 0,$$

and, since $\lambda_1 \neq \lambda_2$, that

$$\mathbf{x}_1' \cdot \mathsf{L} \cdot \mathbf{x}_2 = 0. \;■$$

Since the matrices V and T of an oscillating system are both real and symmetric, the above theorem applies to normal mode theory. It follows that if $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are the amplitude vectors of two normal modes of an oscillating system with *distinct frequencies*, then

$$\mathbf{a}_1' \cdot \mathsf{T} \cdot \mathbf{a}_2 = 0.$$

This result is not necessarily true for amplitude vectors that belong to the same (degenerate) frequency, but they can always be *chosen* to do so. If this has been done, then the full set of amplitude vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots \boldsymbol{a}_n$ are mutually orthogonal.

Orthogonality of normal modes

The amplitude vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots \boldsymbol{a}_n$ of the normal modes of an oscillating system satisfy (or can be chosen to satisfy) the **orthogonality relations**

$$\mathbf{a}'_j \cdot \mathbf{T} \cdot \mathbf{a}_k = 0 \qquad (j \neq k). \tag{15.22}$$

For theoretical purposes, it is also convenient to *normalise* the amplitude vectors. Since $\mathbf{T}$ is a positive definite matrix, the quantities $\mathbf{a}'_1 \cdot \mathbf{T} \cdot \mathbf{a}_1, \mathbf{a}'_2 \cdot \mathbf{T} \cdot \mathbf{a}_2, \dots, \mathbf{a}'_n \cdot \mathbf{T} \cdot \mathbf{a}_n$ are all *positive*. It follows that the amplitude vectors can be scaled so that

$$\mathbf{a}'_1 \cdot \mathbf{T} \cdot \mathbf{a}_1 = \mathbf{a}'_2 \cdot \mathbf{T} \cdot \mathbf{a}_2 = \cdots = \mathbf{a}'_n \cdot \mathbf{T} \cdot \mathbf{a}_n = 1, \tag{15.23}$$

in which case they are said to be **normalised**. The orthogonality and normalisation relations (15.23), (15.23) can then be combined into the single set of relations

$$\boxed{\mathbf{a}'_j \cdot \mathbf{T} \cdot \mathbf{a}_k = \begin{cases} 0 & (j \neq k) \\ 1 & (j = k) \end{cases}} \tag{15.24}$$

called the **orthonormality relations**.

Rayleigh's minimum principle

As an application of the orthogonality relations, we will now prove a far reaching result known as Rayleigh's minimum principle. Suppose an oscillating system $\mathcal{S}$ with n degrees of freedom has potential and kinetic energy matrices $\mathbf{V}$ and $\mathbf{T}$. Consider the function

$$F(\mathbf{x}) = \frac{\mathbf{x}' \cdot \mathbf{V} \cdot \mathbf{x}}{\mathbf{x}' \cdot \mathbf{T} \cdot \mathbf{x}} \tag{15.25}$$

where $\mathbf{x}$ is *any* non-zero column vector of dimension n. The function $F(\mathbf{x})$ is called **Rayleigh's function** for the system $\mathcal{S}$ and it has some interesting properties. To keep things simple we will suppose that $\mathcal{S}$ has *no degenerate normal frequencies*.

Theorem 15.3 ***Rayleigh's minimum principle*** *Suppose that an oscillating system $\mathcal{S}$ has Rayleigh function $F(\mathbf{x})$. Then*

$$F(\mathbf{x}) \geq \omega_1^2 \tag{15.26}$$

for all non-zero column vectors $\mathbf{x}$*, where* ω_1 *is the fundamental frequency** *of* $\mathcal{S}$*. The minimum value is achieved when* $\mathbf{x}$ *is a multiple of the amplitude vector the fundamental mode.*

Proof. Let the n normal frequencies be ordered so that $\omega_1 < \omega_2 < \cdots < \omega_n$ and let the corresponding amplitude vectors be $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$. We will suppose that the amplitude vectors have been normalised so that they satisfy the orthonormality relations (15.24).

Now let $\mathbf{x}$ be any column vector. Since the n amplitude vectors form a basis set,† $\mathbf{x}$ can be expanded in the form $\mathbf{x} = \alpha_1\mathbf{a}_1 + \alpha_2\mathbf{a}_2 + \cdots + \alpha_n\mathbf{a}_n$. Then

$$
\begin{aligned}
\mathbf{x}' \cdot \mathbf{V} \cdot \mathbf{x} &= \mathbf{x}' \cdot \mathbf{V} \cdot (\alpha_1\mathbf{a}_1 + \alpha_2\mathbf{a}_2 + \cdots + \alpha_n\mathbf{a}_n) \\
&= \alpha_1 \left(\mathbf{x}' \cdot \mathbf{V} \cdot \mathbf{a}_1\right) + \alpha_2 \left(\mathbf{x}' \cdot \mathbf{V} \cdot \mathbf{a}_2\right) + \cdots + \alpha_n \left(\mathbf{x}' \cdot \mathbf{V} \cdot \mathbf{a}_n\right) \\
&= \alpha_1\omega_1^2 \left(\mathbf{x}' \cdot \mathbf{T} \cdot \mathbf{a}_1\right) + \alpha_2\omega_2^2 \left(\mathbf{x}' \cdot \mathbf{T} \cdot \mathbf{a}_2\right) + \cdots + \alpha_n\omega_n^2 \left(\mathbf{x}' \cdot \mathbf{T} \cdot \mathbf{a}_n\right).
\end{aligned}
$$

But

$$
\begin{aligned}
\mathbf{x}' \cdot \mathbf{T} \cdot \mathbf{a}_k &= (\alpha_1\mathbf{a}_1 + \alpha_2\mathbf{a}_2 + \cdots + \alpha_n\mathbf{a}_n) \cdot \mathbf{T} \cdot \mathbf{a}_k \\
&= \alpha_1 (\mathbf{a}_1 \cdot \mathbf{T} \cdot \mathbf{a}_k) + \alpha_2 (\mathbf{a}_2 \cdot \mathbf{T} \cdot \mathbf{a}_k) + \cdots + \alpha_n (\mathbf{a}_n \cdot \mathbf{T} \cdot \mathbf{a}_k) \\
&= \alpha_k
\end{aligned}
$$

on using the orthonormality relations. Hence

$$
\mathbf{x}' \cdot \mathbf{V} \cdot \mathbf{x} = \alpha_1^2\,\omega_1^2 + \alpha_2^2\,\omega_2^2 + \cdots + \alpha_n^2\,\omega_n^2
$$

and, by a similar argument,

$$
\mathbf{x}' \cdot \mathbf{T} \cdot \mathbf{x} = \alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2.
$$

Hence

$$
\begin{aligned}
F(\mathbf{x}) &= \frac{\alpha_1^2\,\omega_1^2 + \alpha_2^2\,\omega_2^2 + \cdots + \alpha_n^2\,\omega_n^2}{\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2} \ge \frac{\alpha_1^2\,\omega_1^2 + \alpha_2^2\,\omega_1^2 + \cdots + \alpha_n^2\,\omega_1^2}{\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2} \\
&= \omega_1^2
\end{aligned}
$$

which is the required result. It is also evident that equality can only occur when $\alpha_2 = \alpha_3 = \cdots = \alpha_n = 0$, that is, when $\mathbf{x} = \alpha_1\mathbf{a}_1$. ■

This result means that $F(\mathbf{x})$ is an **upper bound** for ω_1^2, for *any* choice of the column vector $\mathbf{x}$; this upper bound has been obtained *without solving the oscillation problem*. Moreover, if we could substitute every value of $\mathbf{x}$ into the function $F(\mathbf{x})$, then the vectors that yield the least value of F must be multiples of the amplitude vector $\mathbf{a}_1$.

---

* The **fundamental frequency** is the lowest of the normal frequencies and the corresponding normal mode is the **fundamental mode**.

† This follows because the column vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ satisfy the orthogonality relations (15.22). A set of mutually orthogonal vectors must be linearly independent, and, since there are n of them, they form a basis for the space of column vectors of dimension n.

In normal mode theory, this result is of little consequence since the normal frequencies are simply the roots of a polynomial equation which can always be solved numerically. However, Rayleigh's principle has extensions to many areas of applied mathematics and physics such as continuum mechanics and quantum mechanics. In these subjects, the oscillation problems often cannot be solved, even numerically, and Rayleigh's principle is one of the few ways in which information can be gained about the fundamental mode. For example, in **quantum mechanics** Rayleigh's minimum principle takes the form:

Suppose $\mathcal{S}$ is a quantum mechanical system with Hamiltonian H *and ground state energy E_1. Then*

$$\frac{\langle\, \mathsf{x} \,|\, \mathsf{H} \,|\, \mathsf{x} \,\rangle}{\langle\, \mathsf{x} \,|\, \mathsf{x} \rangle} \;\geq\; E_1,$$

for any choice of the quantum state x.

15.7 GENERAL SMALL OSCILLATIONS

Normal modes are special small motions but, from them, we can generate the **general solution** of the small oscillation equations. The result is as follows:

General solution of small oscillation equations

The general solution of the small oscillation equations can be expressed as a linear combination of normal modes in the form

$$\boldsymbol{q}(t) = C_1 \boldsymbol{a}_1 \cos(\omega_1 t - \gamma_1) + C_2 \boldsymbol{a}_2 \cos(\omega_2 t - \gamma_2) + \cdots + C_n \boldsymbol{a}_n \cos(\omega_n t - \gamma_n)$$

where the amplitude factors $\{C_j\}$ and phase factors $\{\gamma_j\}$ are arbitrary constants.

Proof. Suppose that $\boldsymbol{q}(t)$ is *any* solution of the small oscillation equations (15.11). We will now show that we can construct a linear combination of normal modes that satisfies the small oscillation equations and also satisfies the same initial conditions as $\boldsymbol{q}(t)$. To do this, take a general linear combination of normal modes in the form

$$\begin{aligned}\boldsymbol{q}^*(t) &= C_1 \boldsymbol{a}_1 \cos(\omega_1 t - \gamma_1) + C_2 \boldsymbol{a}_2 \cos(\omega_2 t - \gamma_2) + \cdots + C_n \boldsymbol{a}_n \cos(\omega_n t - \gamma_n)\\ &= \boldsymbol{a}_1 \left(A_1 \cos \omega_1 t + B_1 \sin \omega_2 t\right) + \boldsymbol{a}_2 \left(A_2 \cos \omega_2 t + B_2 \sin \omega_2 t\right) + \\ &\qquad \cdots + \boldsymbol{a}_n \left(A_n \cos \omega_n t + B_n \sin \omega_n t\right),\end{aligned}$$

on writing $C_j \cos(\omega_j t - \gamma_j) = A_j \cos \omega_j t + B_j \sin \omega_j t$. Since the small oscillation equations are linear and homogeneous, $\boldsymbol{q}^*(t)$ is a solution for all choices of the coefficients $\{A_j\}$, $\{B_j\}$. We now need to choose the coefficients $\{A_j\}$, $\{B_j\}$ so that $\boldsymbol{q}^*(0) = \boldsymbol{q}(0)$ and $\dot{\boldsymbol{q}}^*(0) = \dot{\boldsymbol{q}}(0)$. This requires that the coefficients $\{A_j\}$ be chosen so that

$$A_1 \boldsymbol{a}_1 + A_2 \boldsymbol{a}_2 + \cdots + A_n \boldsymbol{a}_n = \boldsymbol{q}(0),$$

and that the coefficients $\{B_j\}$ be chosen so that

$$(B_1 \omega_1)\, \boldsymbol{a}_1 + (B_2 \omega_2)\, \boldsymbol{a}_2 + \cdots + (B_n \omega_n)\, \boldsymbol{a}_n = \dot{\boldsymbol{q}}(0).$$

This is *always possible* because the n amplitude vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ form a basis for the space of vectors of dimension n. The vectors $\boldsymbol{q}(t)$ and $\dot{\boldsymbol{q}}(t)$ *can* therefore be expanded in the required forms.

We have thus constructed a solution $\boldsymbol{q}^*(t)$ of the small oscillation equations that satisfies the same initial conditions as the solution $\boldsymbol{q}(t)$. But ODE theory tells us that there can be only *one* such solution and so $\boldsymbol{q} = \boldsymbol{q}^*$. Since $\boldsymbol{q}$ can be any solution and $\boldsymbol{q}^*$ is a linear combination of normal modes, it follows that *any solution of the small oscillation equations can be expressed as a linear combination of normal modes.* ■

Example 15.7 *General small motion of the double pendulum*

Find the general solution of the small oscillation equations for the double pendulum problem.

Solution

The normal modes for the double pendulum problem have been found to be

$$\left.\begin{aligned}\theta &= \epsilon_1 \cos(\sqrt{2/3}\,nt - \gamma_1)\\ \phi &= 2\epsilon_1 \cos(\sqrt{2/3}\,nt - \gamma_1)\end{aligned}\right\}, \qquad \left.\begin{aligned}\theta &= \epsilon_2 \cos(\sqrt{2}\,nt - \gamma_2)\\ \phi &= -2\epsilon_2 \cos(\sqrt{2}\,nt - \gamma_2)\end{aligned}\right\}.$$

The general small motion is therefore

$$\begin{aligned}\theta &= \epsilon_1 \cos(\sqrt{2/3}\,nt - \gamma_1) + \epsilon_2 \cos(\sqrt{2}\,nt - \gamma_2),\\ \phi &= 2\epsilon_1 \cos(\sqrt{2/3}\,nt - \gamma_1) - 2\epsilon_2 \cos(\sqrt{2}\,nt - \gamma_2),\end{aligned}$$

where $\epsilon_1, \epsilon_2, \gamma_1, \gamma_2$ are arbitrary constants. ■

General small motion not usually periodic

The general small motion is a sum of periodic motions, but it is *not* usually periodic itself. Periodicity will occur only if there is some time interval τ that is an integer multiple of each of the periods $\tau_1, \tau_2, \ldots, \tau_n$ of the normal modes. This only happens when the *ratios* of the normal mode periods are all *rational numbers*. In the double pendulum example, $\tau_1/\tau_2 = \omega_2/\omega_1 = \sqrt{3}$, which is irrational. The general small motion is therefore *not* periodic.

15.8 NORMAL COORDINATES

The preceding theory applies for any choice of the generalised coordinates $\{q_j\}$. Changing the generalised coordinates will change the V- and T-matrices, but the normal frequencies and the physical forms of the normal modes will be the same. This suggests that it might be possible to make a clever choice of coordinates so that the V- and T-matrices have a simple form leading to a much simplified theory. In particular, it would be very advantageous if $\mathbf{T}$ and $\mathbf{V}$ had **diagonal** form.

Definition 15.7 ***Normal coordinates*** *A set of generalised coordinates in terms of which the T- and V-matrices have diagonal form are called* ***normal coordinates****.*

Actually, every oscillating system has normal coordinates, as we will now show. Let $\boldsymbol{q}$ be the original choice of coordinates with corresponding matrices $\mathbf{V}$ and $\mathbf{T}$. Then

$$T^{\text{app}} = \dot{\mathbf{q}}' \cdot \mathbf{T} \cdot \dot{\mathbf{q}}, \qquad V^{\text{app}} = \mathbf{q}' \cdot \mathbf{V} \cdot \mathbf{q}. \tag{15.27}$$

Now consider a change of coordinates from $\boldsymbol{q}$ to $\boldsymbol{\eta}$ defined by the linear transformation*

$$\mathsf{q} = \mathsf{P} \cdot \boldsymbol{\eta} \quad \Longleftrightarrow \quad \boldsymbol{\eta} = \mathsf{P}^{-1} \cdot \mathsf{q} \tag{15.28}$$

where P can be any non-singular matrix. On substituting the transformation (15.28) into the expressions (15.27), we obtain

$$T^{\text{app}} = (\mathsf{P} \cdot \dot{\boldsymbol{\eta}})' \cdot \mathsf{T} \cdot (\mathsf{P} \cdot \dot{\boldsymbol{\eta}}) = \dot{\boldsymbol{\eta}}' \cdot (\mathsf{P}' \cdot \mathsf{T} \cdot \mathsf{P}) \cdot \dot{\boldsymbol{\eta}},$$
$$V^{\text{app}} = (\mathsf{P} \cdot \boldsymbol{\eta})' \cdot \mathsf{V} \cdot (\mathsf{P} \cdot \boldsymbol{\eta}) = \boldsymbol{\eta}' \cdot (\mathsf{P}' \cdot \mathsf{V} \cdot \mathsf{P}) \cdot \boldsymbol{\eta},$$

from which we see that this transformation of coordinates causes V and T to be transformed as

$$\mathsf{T} \to \mathsf{P}' \cdot \mathsf{T} \cdot \mathsf{P}, \qquad\qquad \mathsf{V} \to \mathsf{P}' \cdot \mathsf{V} \cdot \mathsf{P}. \tag{15.29}$$

Can we now choose the transformation matrix P so that the new T- and V-matrices are diagonal?

Let $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ be the amplitude vectors of the normal modes when they are expressed in terms of the coordinates $\boldsymbol{q}$ and let $\omega_1, \omega_2, \ldots, \omega_n$ be the corresponding normal frequencies. We will suppose that these amplitude vectors have been chosen so that they satisfy the **orthonormality relations** (15.24), that is

$$\mathsf{a}'_j \cdot \mathsf{T} \cdot \mathsf{a}_k = \begin{cases} 0 & (j \neq k), \\ 1 & (j = k). \end{cases} \tag{15.30}$$

Now consider the matrix P whose columns are the amplitude vectors $\{\mathsf{a}_j\}$, that is,

$$\mathsf{P} = (\mathsf{a}_1 | \mathsf{a}_2 | \cdots | \mathsf{a}_n)\,. \tag{15.31}$$

Since the amplitude vectors are known to be linearly independent, P has linearly independent columns and is therefore a non-singular matrix. Let us now try this P as the transformation matrix. Then

$$\mathsf{P}' \cdot \mathsf{T} \cdot \mathsf{P} = \begin{pmatrix} \mathsf{a}'_1 \\ \hline \mathsf{a}'_2 \\ \hline \vdots \\ \hline \mathsf{a}'_n \end{pmatrix} \cdot \mathsf{T} \cdot (\mathsf{a}_1 | \mathsf{a}_2 | \cdots | \mathsf{a}_n)\,.$$

---

* For example, in the case of two degrees of freedom, this transformation has the form

$$q_1 = p_{11}\,\eta_1 + p_{12}\,\eta_2,$$
$$q_2 = p_{21}\,\eta_1 + p_{22}\,\eta_2.$$

The jk-th element of this matrix is given by

$$\mathbf{a}_j' \cdot \mathbf{T} \cdot \mathbf{a}_k = \begin{cases} 0 & (j \neq k) \\ 1 & (j = k) \end{cases}$$

by the orthonormality relations. Hence, with this choice of $\mathbf{P}$,

$$\mathbf{P}' \cdot \mathbf{T} \cdot \mathbf{P} = \mathbf{1},$$

where $\mathbf{1}$ is the **identity matrix**. In the same way,

$$\mathbf{P}' \cdot \mathbf{V} \cdot \mathbf{P} = \begin{pmatrix} \mathbf{a}_1' \\ \hline \mathbf{a}_2' \\ \hline \vdots \\ \hline \mathbf{a}_n' \end{pmatrix} \cdot \mathbf{V} \cdot (\mathbf{a}_1 | \mathbf{a}_2 | \cdots | \mathbf{a}_n) \,.$$

The jk-th element of this matrix is given by

$$\begin{aligned} \mathbf{a}_j' \cdot \mathbf{V} \cdot \mathbf{a}_k &= \mathbf{a}_j' \cdot (\mathbf{V} \cdot \mathbf{a}_k) = \mathbf{a}_j' \cdot \left(\omega_k^2 \mathbf{T} \cdot \mathbf{a}_k\right) = \omega_k^2 \left(\mathbf{a}_j' \cdot \mathbf{T} \cdot \mathbf{a}_k\right) \\ &= \begin{cases} 0 & (j \neq k), \\ \omega_j^2 & (j = k). \end{cases} \end{aligned}$$

Hence, with this choice of $\mathbf{P}$,

$$\mathbf{P}' \cdot \mathbf{V} \cdot \mathbf{P} = \mathbf{\Omega}^2,$$

where $\mathbf{\Omega}$ is the diagonal matrix whose diagonal elements are the normal frequencies, that is,

$$\mathbf{\Omega} = \begin{pmatrix} \omega_1 & 0 & \cdots & 0 \\ 0 & \omega_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \omega_n \end{pmatrix}.$$

We have thus succeeded in reducing both $\mathbf{V}$ and $\mathbf{T}$ to diagonal form. Hence the coordinates $\{\eta_j\}$ defined by (15.28) with $\mathbf{P} = (\mathbf{a}_1 | \mathbf{a}_2 | \cdots | \mathbf{a}_n)$ are a set of **normal coordinates**. They are given explicitly by

$$\boldsymbol{\eta} = \mathbf{P}^{-1} \cdot \mathbf{q} = \left(\mathbf{P}' \cdot \mathbf{T}\right) \cdot \mathbf{q},$$

on using the formula $\mathbf{P}' \cdot \mathbf{T} \cdot \mathbf{P} = \mathbf{1}$. This can also be written in the semi-expanded form

$$\eta_j = \left(\mathbf{a}_j' \cdot \mathbf{T}\right) \cdot \mathbf{q} \qquad (1 \leq j \leq n). \tag{15.32}$$

From this last formula, we can see that, if the amplitude vectors $\{\mathbf{a}_j\}$ are *not* normalised, then the coordinates $\{\eta_j\}$ are simply multiplied by constants. They are therefore *still normal coordinates*. The corresponding V- and T-matrices are still diagonal, but $\mathbf{T}$ is no longer reduced to the identity.

Our results are summarised as follows:

Finding normal coordinates

Let $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ be the amplitude vectors of the normal modes when expressed in terms of the coordinates $\{q_j\}$. Then the coordinates $\{\eta_j\}$ defined by

$$\eta_j = \left(\mathbf{a}_j' \cdot \mathbf{T}\right) \cdot \mathbf{q} \qquad (1 \le j \le n)$$

are a set of **normal coordinates**, as are any constant multiples of them. (The amplitude vectors only need to be normalised if it is required to reduce the matrix $\mathbf{T}$ to the identity.)

When expressed in terms of normal coordinates, the **small oscillation equations** become

$$\ddot{\boldsymbol{\eta}} + \boldsymbol{\Omega}^2 \cdot \boldsymbol{\eta} = \mathbf{0}.$$

In expanded form, this is

$$\ddot{\eta}_j + \omega_j^2 \eta_j = 0 \qquad (1 \le j \le n),$$

a system of n *uncoupled* SHM equations. The solution $\eta_1 = C_1 \cos(\omega_1 t - \gamma_1)$, $\eta_2 = \eta_3 = \cdots = \eta_n = 0$ is the first normal mode, the solution $\eta_2 = C_2 \cos(\omega_2 t - \gamma_2)$, $\eta_1 = \eta_3 = \cdots = \eta_n = 0$ is the second normal mode, and so on.

Note. Using normal coordinates is *not* a practical way of solving normal mode problems. Indeed the problem has to be solved before the normal coordinates can be found! Normal coordinates are important because they simplify further developments of the general theory.

Example 15.8 *Finding normal coordinates*

Find a set of normal coordinates for the double pendulum problem.

Solution

For the double pendulum problem, we have already found that

$$\mathbf{T} = \tfrac{1}{2} m b^2 \begin{pmatrix} 4 & 1 \\ 1 & 1 \end{pmatrix}, \qquad \mathbf{a}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \qquad \mathbf{a}_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

Hence, on dropping the inessential constant factor $\frac{1}{2}mb^2$, a set of normal coordinates is given by

$$\eta_1 = (\,1\;\;2\,)\cdot\begin{pmatrix}4 & 1\\ 1 & 1\end{pmatrix}\cdot\begin{pmatrix}\theta\\ \phi\end{pmatrix} = 6\theta + 3\phi$$

and

$$\eta_2 = (\,1\;\;-2\,)\cdot\begin{pmatrix}4 & 1\\ 1 & 1\end{pmatrix}\cdot\begin{pmatrix}\theta\\ \phi\end{pmatrix} = 2\theta - \phi.$$

Since normal coordinates may always be scaled, we can equally well take

$$\eta_1 = 2\theta + \phi,$$
$$\eta_2 = 2\theta - \phi,$$

as our **normal coordinates**. ■

Problems on Chapter 15

Answers and comments are at the end of the book.

Harder problems carry a star (∗).

Two degrees of freedom

15.1 A particle P of mass $3m$ is connected to a particle Q of mass $8m$ by a light elastic spring of natural length a and strength α. Two similar springs are used to connect P and Q to the fixed points A and B respectively, which are a distance $3a$ apart on a smooth horizontal table. The particles can perform longitudinal oscillations along the straight line AB. Find the normal frequencies and the forms of the normal modes.

The system is in equlilibrium when the particle P receives a blow that gives it a speed u in the direction $\overrightarrow{AB}$. Find the displacement of each particle at time t in the subsequent motion.

15.2 A particle A of mass $3m$ is suspended from a fixed point O by a spring of strength α and a second particle B of mass $2m$ is suspended from A by a second identical spring. The system performs small oscillations in the vertical straight line through O. Find the normal frequencies, the forms of the normal modes, and a set of normal coordinates.

15.3 ***Rod pendulum*** A uniform rod of length $2a$ is suspended from a fixed point O by a light inextensible string of length b attached to one of its ends. The system moves in a vertical plane through O. Take as coordinates the angles θ, ϕ between the string and the rod respectively and the downward vertical. Show that the equations governing small oscillations of the system about $\theta = \phi = 0$ are

$$b\ddot{\theta} + a\ddot{\phi} = -g\theta,$$

$$b\ddot{\theta} + \tfrac{4}{3}a\ddot{\phi} = -g\phi.$$

For the special case in which $b = 4a/5$, find the normal frequencies and the forms of the normal modes. Is the general motion periodic?

Three or more degrees of freedom

15.4 ***Triple pendulum*** A triple pendulum has three strings of equal length a and the three particles (starting from the top) have masses $6m$, $2m$, m respectively. The pendulum performs small oscillations in a vertical plane. Show that the normal frequencies satisfy the equation

$$12\mu^3 - 60\mu^2 + 81\mu - 27 = 0,$$

where $\mu = a\omega^2/g$. Find the normal frequencies, the forms of the normal modes, and a set of normal coordinates. [$\mu = 3$ is a root of the equation.]

15.5 A light *elastic* string is stretched to tension T_0 between two fixed points A and B a distance $3a$ apart, and two particles of mass m are attached to the string at equally spaced intervals. The strength of *each* of the three sections of the string is α. The system performs small oscillations in a plane through AB. Without making any prior assumptions, prove that the particles oscillate longitudinally in two of the normal modes and transversely in the other two. Find the four normal frequencies.

15.6 A rod of mass M and length L is suspended from two fixed points at the same horizontal level and a distance L apart by two equal strings of length b attached to its ends. From each end of the rod a particle of mass m is suspended by a string of length a. The system of the rod and two particles performs small oscillations in a vertical plane. Find $\mathbf{V}$ and $\mathbf{T}$ for this system. For the special case in which $b = 3a/2$ and $M = 6m/5$, find the normal frequencies. Show that the general small motion is periodic and find the period.

15.7 A uniform rod is suspended in a horizontal position by *unequal* vertical strings of lengths b, c attached to its ends. Show that the frequency of the in-plane swinging mode is $((b + c)g/2bc)^{1/2}$, and that the frequencies of the other modes satisfy the equation

$$bc\mu^2 - 2a(b + c)\mu + 3a^2 = 0,$$

where $\mu = a\omega^2/g$. Find the normal frequencies for the particular case in which $b = 3a$ and $c = 8a$.

15.8* A uniform rod BC has mass M and length $2a$. The end B of the rod is connected to a fixed point A on a smooth horizontal table by an elastic string of strength α_1, and the end C is connected to a second fixed point D on the table by a second elastic string of strength α_2. In equilibrium, the rod lies along the line AD with the strings having tension T_0 and lengths b, c respectively. Show that the frequency of the longitudinal mode is $((\alpha_1 + \alpha_2)/M)^{1/2}$ and that the frequencies of the transverse modes satisfy the equation

$$b^2c^2\mu^2 - 2bc(2ab + 3bc + 2ac)\mu + 6abc(2a + b + c) = 0,$$

where $\mu = Ma\omega^2/T_0$. [The calculation of V^{app} is very tricky.]

Find the frequencies of the transverse modes for the particular case in which $a = 3c$ and $b = 5c$.

15.9* A light *elastic* string is stretched between two fixed points A and B a distance $(n+1)a$ apart, and n particles of mass m are attached to the string at equally spaced intervals. The strength of *each* of the $n+1$ sections of the string is α. The system performs small *longitudinal* oscillations along the line AB. Show that the normal frequencies satisfy the determinantal equation

$$\Delta_n \equiv \begin{vmatrix} 2\cos\theta & -1 & 0 & \cdots & 0 & 0 \\ -1 & 2\cos\theta & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2\cos\theta & -1 \\ 0 & 0 & 0 & \cdots & -1 & 2\cos\theta \end{vmatrix} = 0,$$

where $\cos\theta = 1 - (m\omega^2/2\alpha)$.

By expanding the determinant by the top row, show that Δ_n satisfies the recurrence relation

$$\Delta_n = 2\cos\theta\,\Delta_{n-1} - \Delta_{n-2},$$

for $n \geq 3$. Hence, show by induction that

$$\Delta_n = \sin(n+1)\theta/\sin\theta.$$

Deduce the normal frequencies of the system.

15.10 A light string is stretched to a tension T_0 between two fixed points A and B a distance $(n+1)a$ apart, and n particles of mass m are attached to the string at equally spaced intervals. The system performs small plane *transverse* oscillations. Show that the normal frequencies satisfy the same determinantal equation as in the previous question, except that now $\cos\theta = 1 - (ma\omega^2/2T_0)$. Find the normal frequencies of the system.

Vibrating molecules

15.11 ***Unsymmetrical linear molecule*** A general linear triatomic molecule has atoms A_1, A_2, A_3 with masses m_1, m_2, m_3. The chemical bond between A_1 and A_2 is represented by a spring of strength α_{12} and the bond between A_2 and A_3 is represented by a spring of strength α_{23}. Show that the vibrational frequences of the molecule satisfy the equation

$$m_1 m_2 m_3\,\omega^4 - [\alpha_{12} m_3(m_1+m_2) + \alpha_{23} m_1(m_2+m_3)]\,\omega^2$$

$$+\,\alpha_{12}\alpha_{23}(m_1+m_2+m_3) = 0.$$

Find the vibrational frequencies for the special case in which $m_1 = 3m$, $m_2 = m$, $m_3 = 2m$ and $\alpha_{12} = 3\alpha$, $\alpha_{23} = 2\alpha$.

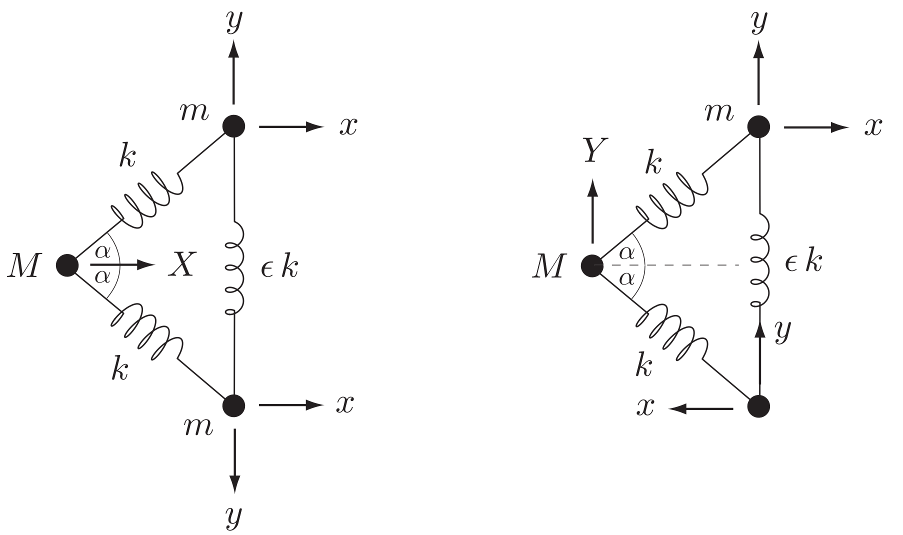


FIGURE 15.7 Vibrations of a symmetric V-shaped molecule. **Left:** a symmetric motion, **Right:** an antisymmetric motion.

The molecule O–C–S (carbon oxysulphide) is known to be linear. Use the λ_1^{-1} values given in Table 2 to estimate the ratio of its vibrational frequencies. [The experimentally measured value is 2.49.]

15.12* ***Symmetric V-shaped molecule*** Figure 15.7 shows the symmetric V-shaped triatomic molecule XY_2; the X–Y bonds are represented by springs of strength k, while the Y–Y bond is represented by a spring of strength $\epsilon\, k$. Common examples of such molecules include water, hydrogen sulphide, sulphur dioxide and nitrogen dioxide; the apex angle 2α is typically between 90° and 120°. In planar motion, the molecule has six degrees of freedom of which three are rigid body motions; there are therefore *three* vibrational modes. It is best to exploit the reflective symmetry of the molecule and solve separately for the symmetric and antisymmetric modes. Figure 15.7 (left) shows a symmetric motion while (right) shows an antisymmetric motion; the displacements X, Y, x, y are measured *from the equilibrium position*. Show that there is one antisymmetric mode whose frequency ω_3 is given by

$$\omega_3^2 = \frac{k}{mM}(M + 2m\sin^2\alpha),$$

and show that the frequencies of the symmetric modes satisfy the equation

$$\mu^2 - \Big(1 + 2\gamma\cos^2\alpha + 2\epsilon\Big)\,\mu + 2\epsilon\cos^2\alpha(1 + 2\gamma) = 0,$$

where $\mu = m\omega^2/k$ and $\gamma = m/M$.

Find the three vibrational frequencies for the special case in which $M = 2m$, $\alpha = 60°$ and $\epsilon = 1/2$.

15.13 ***Plane triangular molecule*** The molecule BCl_3 (boron trichloride) is plane and symmetrical. In equlibrium, the Cl atoms are at the vertices of an equilateral triangle with the

B atom at the centroid. Show that the molecule has six vibrational modes of which five are in the plane of the molecule; show also that the out-of-plane mode and one of the in-plane modes have axial symmetry; and show finally that the remaining four in-plane modes are in doubly degenerate pairs. Deduce that the BCl_3 molecule has a total of four distinct vibrational frequencies.

Computer assisted problem

15.14 ***Sulphur dioxide molecule*** Use computer assistance to obtain an equation satisfied by the squares of the frequencies of the symmetric modes of a V-shaped molecule. For the special case in which $M = 2m$ and $\alpha = 60°$, show that the frequencies of the symmetrical modes satisfy the equation

$$4\mu^2 - (5 + 8\epsilon)\mu + 4\epsilon = 0,$$

where $\mu = m\omega^2/k$.

The sulphur dioxide molecule O–S–O has mass ratio $M/m = 2$ and an apex angle very close to 120°. Its infrared absorbtion wave numbers are found to be $\lambda_1^{-1} = 1151 \text{ cm}^{-1}$, $\lambda_2^{-1} = 525 \text{ cm}^{-1}$, $\lambda_3^{-1} = 1336 \text{ cm}^{-1}$. Show that there is no value of ϵ that fits this data with reasonable accuracy. This is a deficiency of our simple (central force) model of interatomic forces, which gives poor results for V-shaped molecules (see Herzberg [13]).

Chapter Sixteen

Vector angular velocity and rigid body kinematics

KEY FEATURES

The key features in this chapter are **vector angular velocity** and the kinematics of **rigid bodies in general motion**.

This chapter is concerned with the kinematics of **rigid bodies in general motion**. In Chapter 2 we considered only those rigid body motions that were essentially two-dimensional, and angular velocity appeared there as a *scalar* quantity. In general three-dimensional rigid body motion, this approach is no longer adequate and angular velocity must be introduced in its proper rôle as a **vector** quantity. The principal result of the Chapter is that any motion of a rigid body can be represented as a sum of **translational** and **rotational** contributions.

16.1 ROTATION ABOUT A FIXED AXIS

In this chapter we adopt a more rigorous approach to rigid body rotation than we did in Chapter 2. We begin with a proper definition of rigidity.

Definition 16.1 ***Rigidity*** *A body $\mathcal{B}$ is said to be a* ***rigid body*** *if the distance between any pair of its particles remains constant. That is, if P_i and P_j are typical particles of $\mathcal{B}$ with position vectors $\boldsymbol{r}_i(t)$ and $\boldsymbol{r}_j(t)$ at time t, then*

$$|\boldsymbol{r}_i(t) - \boldsymbol{r}_j(t)| = c_{ij}, \tag{16.1}$$

where the c_{ij} are constants.

Suppose a rigid body $\mathcal{B}$ is rotating about a fixed axis with angular speed ω. This motion certainly satisfies the rigidity conditions (16.1). Let $\widehat{\boldsymbol{n}}$ be a *unit* vector parallel to the rotation axis. Then the **vector angular velocity** of $\mathcal{B}$ is defined as follows:

Definition 16.2 ***Vector angular velocity*** *The* ***angular velocity vector*** *of the body $\mathcal{B}$ is defined to be*

$$\boldsymbol{\omega} = \pm\,\omega\,\widehat{\boldsymbol{n}}, \tag{16.2}$$

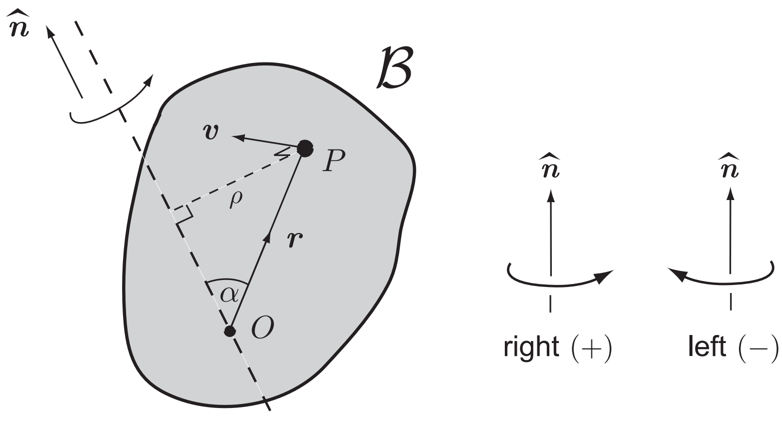


FIGURE 16.1 The rigid body $\mathcal{B}$ rotates with angular speed ω about a fixed axis parallel to the unit vector $\widehat{\boldsymbol{n}}$. Its angular velocity vector is defined by $\boldsymbol{\omega} = \pm\,\omega\,\widehat{\boldsymbol{n}}$, where the sign is determined by the sense of the rotation.

where the sign is taken to be plus or minus depending on whether the sense of the rotation (relative to the vector $\widehat{\boldsymbol{n}}$*) is right- or left-handed. These senses are shown in Figure 16.1.*

Note. From the vector $\boldsymbol{\omega}$ we can deduce the angular speed, the axis direction, and the rotation sense about the axis. It tells us nothing about the position of the axis either in space or in the body.

Example 16.1 *Calculation of* $\boldsymbol{\omega}$

A rigid body $\mathcal{B}$ is rotating with angular speed 7 radians per second about a fixed axis through the points $A(2, 3, -1)$, $B(-4, 0, 1)$. The rotation is in the left-handed sense relative to $\overrightarrow{AB}$. Find the angular velocity of $\mathcal{B}$.

Solution

The position vectors of the points A and B are $\boldsymbol{a} = 2\boldsymbol{i} + 3\,\boldsymbol{j} - \boldsymbol{k}$ and $\boldsymbol{b} = -4\boldsymbol{i} + \boldsymbol{k}$ so that

$$\overrightarrow{AB} = \boldsymbol{b} - \boldsymbol{a} = -6\boldsymbol{i} - 3\,\boldsymbol{j} + 2\boldsymbol{k}.$$

The vector $\widehat{\boldsymbol{n}}$ is then

$$\widehat{\boldsymbol{n}} = \frac{-6\boldsymbol{i} - 3\,\boldsymbol{j} + 2\boldsymbol{k}}{|-6\boldsymbol{i} - 3\,\boldsymbol{j} + 2\boldsymbol{k}|} = \frac{-6\boldsymbol{i} - 3\,\boldsymbol{j} + 2\boldsymbol{k}}{7}.$$

The rotation sense is left-handed relative to the direction of $\boldsymbol{n}$ so that the angular velocity of $\mathcal{B}$ is

$$\begin{aligned}\boldsymbol{\omega} &= -7\,\widehat{\boldsymbol{n}}\\ &= -7\left(\frac{-6\boldsymbol{i} - 3\,\boldsymbol{j} + 2\boldsymbol{k}}{7}\right)\\ &= 6\boldsymbol{i} + 3\,\boldsymbol{j} - 2\boldsymbol{k} \quad \text{radians per second.}\ \blacksquare\end{aligned}$$

Particle velocities and accelerations

The **particle velocities** can be conveniently calculated in terms of the vector $\boldsymbol{\omega}$. Let P be a particle of $\mathcal{B}$ with position vector $\boldsymbol{r}$ relative to an origin O located **on the rotation axis** (see Figure 16.1). Then the velocity of P has the same direction as the vector $\boldsymbol{\omega}\times\boldsymbol{r}$. Hence $\boldsymbol{v}$ can be written in the form

$$\boldsymbol{v} = \lambda\, \boldsymbol{\omega}\times\boldsymbol{r},$$

where λ is a positive scalar. To determine λ, consider the magnitude of each side. The magnitude of $\boldsymbol{v}$ is the circumferential speed $\omega\rho$ (see Figure 16.5) and so

$$\begin{aligned}\omega\rho &= \lambda|\boldsymbol{\omega}\times\boldsymbol{r}| = \lambda|\boldsymbol{\omega}||\boldsymbol{r}|\sin\alpha = \lambda\omega\,(OP\sin\alpha)\\ &= \lambda\omega\rho.\end{aligned}$$

Hence $\lambda = 1$ and the formula for $\boldsymbol{v}$ is

$$\boldsymbol{v} = \boldsymbol{\omega}\times\boldsymbol{r}. \tag{16.3}$$

This formula applies only when the origin of vectors lies on the rotation axis, but a more general result is easy to obtain. Let B be *any* fixed point on the rotation axis. Then the velocity formula (16.3) still holds if the position vector $\boldsymbol{r}$ is replaced by $\overrightarrow{BP} = \boldsymbol{r} - \boldsymbol{b}$, where $\boldsymbol{b}$ is the position vector of the point B. Hence the general formula for the velocity of P is given by

$$\boxed{\boldsymbol{v} = \boldsymbol{\omega}\times(\boldsymbol{r} - \boldsymbol{b})} \tag{16.4}$$

where B is any point on the rotation axis.

Example 16.2 *Finding particle velocities and accelerations*

A rigid body is rotating with constant angular speed 7 radians per second about a fixed axis through the points $A(2, 3, -1)$, $B(-4, 0, 1)$, distances being measured in centimetres. The rotation is in the left-handed sense relative to $\overrightarrow{AB}$. Find the instantaneous velocity, speed, and acceleration of the particle P of the body at the point $(-3, 3, 5)$.

Solution

The angular velocity of this body has been determined in the last example to be

$$\boldsymbol{\omega} = 6\boldsymbol{i} + 3\boldsymbol{j} - 2\boldsymbol{k} \quad \text{radians per second.}$$

The velocity of P can now be found using (16.4) with

$$\boldsymbol{r} = -3\boldsymbol{i} + 3\boldsymbol{j} + 5\boldsymbol{k} \quad \text{and} \quad \boldsymbol{b} = -4\boldsymbol{i} + \boldsymbol{k}.$$

This gives

$$\begin{aligned} \boldsymbol{v} &= (6\boldsymbol{i} + 3\boldsymbol{j} - 2\boldsymbol{k}) \times (\boldsymbol{i} + 3\boldsymbol{j} + 4\boldsymbol{k}) \\ &= \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 6 & 3 & -2 \\ 1 & 3 & 4 \end{vmatrix} \\ &= 18\boldsymbol{i} - 26\boldsymbol{j} + 15\boldsymbol{k} \quad \text{cm s}^{-1}. \end{aligned}$$

This is the instantaneous **velocity** of P. The **speed** is therefore $|\boldsymbol{v}| = \left(18^2 + (-26)^2 + 15^2\right)^{1/2} = 35\ \text{cm s}^{-1}$.

The **acceleration** of P can be found by differentiating the formula (16.4) with repect to t. This gives

$$\boldsymbol{a} = \dot{\boldsymbol{\omega}} \times (\boldsymbol{r} - \boldsymbol{b}) + \boldsymbol{\omega} \times (\dot{\boldsymbol{r}} - \dot{\boldsymbol{b}}).$$

But $\boldsymbol{\omega}$ is known to have constant direction and magnitude and so $\dot{\boldsymbol{\omega}} = \boldsymbol{0}$. Also $\dot{\boldsymbol{b}} = \boldsymbol{0}$ since B is a fixed particle. This leaves

$$\begin{aligned} \boldsymbol{a} &= \boldsymbol{\omega} \times \dot{\boldsymbol{r}} \\ &= \boldsymbol{\omega} \times \boldsymbol{v} \\ &= (6\boldsymbol{i} + 3\boldsymbol{j} - 2\boldsymbol{k}) \times (18\boldsymbol{i} - 26\boldsymbol{j} + 15\boldsymbol{k}) \\ &= \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 6 & 3 & -2 \\ 18 & -26 & 15 \end{vmatrix} \\ &= -7\boldsymbol{i} - 126\boldsymbol{j} - 210\boldsymbol{k} \quad \text{cm s}^{-2}. \blacksquare \end{aligned}$$

16.2 GENERAL RIGID BODY KINEMATICS

We now move on to the more general case in which the rigid body does not have a fixed rotation axis. We first consider a rigid body that has *one particle* O that does not move, and we take O to be the origin of position vectors. Now the rigidity conditions (16.1) are equivalent to

$$(\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{r}_i - \boldsymbol{r}_j) = c_{ij}^2 \tag{16.5}$$

and because O is a particle $\mathcal{B}$ it follows in particular that

$$\boldsymbol{r}_i \cdot \boldsymbol{r}_i = d_i, \tag{16.6}$$

where the d_i are constants. On expanding the dot product in (16.5) and using (16.6) we obtain

$$\boldsymbol{r}_i \cdot \boldsymbol{r}_j = e_{ij}, \tag{16.7}$$

where the e_{ij} are constants. If we now differentiate (16.7) with respect to t we obtain

$$\dot{\boldsymbol{r}}_i \cdot \boldsymbol{r}_j + \boldsymbol{r}_i \cdot \dot{\boldsymbol{r}}_j = 0 \qquad \text{for all} \quad i, j, \tag{16.8}$$

which is our preferred form of the **rigidity conditions**.

We now prove the **fundamental theorems** of rigid body kinematics. The details of the proofs are mainly of interest to mathematics students.

Theorem 16.1 ***Existence of angular velocity I*** *Let a rigid body $\mathcal{B}$ be in motion with one of its particles O fixed. Then there exists a unique vector $\boldsymbol{\omega}(t)$ such that the velocity of any particle P of $\mathcal{B}$ is given by the formula*

$$\boldsymbol{v} = \boldsymbol{\omega} \times \boldsymbol{r}, \tag{16.9}$$

where $\boldsymbol{r}$ is the position vector of P relative to O.

This result means that, at each instant, $\mathcal{B}$ is rotating about an *instantaneous* axis through O. This axis is not fixed in space or in the body.

Proof. Suppose that there exist particles E_1, E_2, E_3 of $\mathcal{B}$ such that their position vectors $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ relative to O form a standard basis set. Then if there does exist an $\boldsymbol{\omega}$ satisfying (16.9), it must in particular satisfy

$$\dot{\boldsymbol{e}}_k = \boldsymbol{\omega} \times \boldsymbol{e}_k \tag{16.10}$$

for $k = 1, 2, 3$. Taking the cross product of this equation with $\boldsymbol{e}_k$ gives

$$\begin{aligned} \boldsymbol{e}_k \times \dot{\boldsymbol{e}}_k &= \boldsymbol{e}_k \times (\boldsymbol{\omega} \times \boldsymbol{e}_k) = (\boldsymbol{e}_k \cdot \boldsymbol{e}_k)\,\boldsymbol{\omega} - (\boldsymbol{\omega} \cdot \boldsymbol{e}_k)\,\boldsymbol{e}_k \\ &= \boldsymbol{\omega} - (\boldsymbol{\omega} \cdot \boldsymbol{e}_k)\,\boldsymbol{e}_k \end{aligned}$$

since $\boldsymbol{e}_k$ is a unit vector. Summing these equations over $1 \le k \le 3$ gives

$$\sum_k \boldsymbol{e}_k \times \dot{\boldsymbol{e}}_k = 3\boldsymbol{\omega} - \sum_k (\boldsymbol{\omega} \cdot \boldsymbol{e}_k)\,\boldsymbol{e}_k = 3\boldsymbol{\omega} - \boldsymbol{\omega}$$

since the sum on the right is just the expansion of $\boldsymbol{\omega}$ with respect to the basis set $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$. Hence

$$\boldsymbol{\omega} = \frac{1}{2} \sum_k \boldsymbol{e}_k \times \dot{\boldsymbol{e}}_k. \tag{16.11}$$

Thus if $\boldsymbol{\omega}$ does exist, it must be given by the formula (16.11), which shows that $\boldsymbol{\omega}$ is *unique*. We must now show that this $\boldsymbol{\omega}$ satisfies (16.9) for *all* the particles of the body. It is a simple exercise to verify that this is true for the particles E_1, E_2, E_3 by substituting (16.11) into (16.10) and using the rigidity conditions (16.8). Now let P be any other particle of the body and expand its position vector $\boldsymbol{r}$ with respect to the basis set $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ in the form

$$\boldsymbol{r} = \sum_k (\boldsymbol{r} \cdot \boldsymbol{e}_k)\,\boldsymbol{e}_k.$$

The velocity of P is then given by

$$\boldsymbol{v} = \dot{\boldsymbol{r}} = \sum_k \left(\dot{\boldsymbol{r}} \cdot \boldsymbol{e}_k + \boldsymbol{r} \cdot \dot{\boldsymbol{e}}_k\right) \boldsymbol{e}_k \; + \sum_k (\boldsymbol{r} \cdot \boldsymbol{e}_k)\,\dot{\boldsymbol{e}}_k.$$

Now $\dot{\boldsymbol{r}} \cdot \boldsymbol{e}_k + \boldsymbol{r} \cdot \dot{\boldsymbol{e}}_k = 0$ by the rigidity conditions (16.8), and we have directly verified that $\dot{\boldsymbol{e}}_k = \boldsymbol{\omega} \times \boldsymbol{e}_k$. Hence

$$\begin{aligned}\boldsymbol{v} &= \sum_k (\boldsymbol{r} \cdot \boldsymbol{e}_k)\,(\boldsymbol{\omega} \times \boldsymbol{e}_k) = \boldsymbol{\omega} \times \sum_k (\boldsymbol{r} \cdot \boldsymbol{e}_k)\,\boldsymbol{e}_k \\ &= \boldsymbol{\omega} \times \boldsymbol{r}\end{aligned}$$

as required. ■

The above proof cannot even begin if any of the particles E_1, E_2, E_3 are not actually present (for example the body could be a lamina). In such a case, suppose that the body has at least *two* particles A, B in addition to O, and that O, A, B are not collinear. Then define the standard basis set $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ by

$$\boldsymbol{e}_1 = \frac{\boldsymbol{a}}{|\boldsymbol{a}|}, \qquad \boldsymbol{e}_2 = \frac{\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\,\boldsymbol{a}}{|\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\,\boldsymbol{a}|}, \qquad \boldsymbol{e}_3 = \boldsymbol{e}_1 \times \boldsymbol{e}_2.$$

It can be shown that the points of space E_1, E_2, E_3 that have the position vectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ satisfy the *same rigidity conditions* as the real particles of the body. They can therefore be regarded as real particles and the proof given above then holds.

We now extend the result in Theorem 16.1 to the case of *completely general* motion in which no particle of the body is fixed.

Theorem 16.2 ***Existence of angular velocity II*** *Suppose a rigid body is in completely general motion and let B be any one of its particles. Then there exists a unique vector $\boldsymbol{\omega}(t)$ such that the velocity of any particle P of the body is given by the formula*

$$\boldsymbol{v} = \boldsymbol{v}^B + \boldsymbol{\omega} \times (\boldsymbol{r} - \boldsymbol{b}), \tag{16.12}$$

where $\boldsymbol{r}$ and $\boldsymbol{b}$ are the position vectors of P and B, and $\boldsymbol{v}^B$ is the velocity of B.

Proof. We view the motion of $\mathcal{B}$ from a reference frame $\mathcal{F}'$ with origin at B and moving without rotation relative to the original frame $\mathcal{F}$. Then (see section 1.4) the position vector $\boldsymbol{r}'$ and velocity $\boldsymbol{v}'$ of a particle P relative to $\mathcal{F}'$ are related to the original $\boldsymbol{r}$ and $\boldsymbol{v}$ by

$$\boldsymbol{r} = \boldsymbol{r}' + \boldsymbol{b}, \qquad \boldsymbol{v} = \boldsymbol{v}' + \boldsymbol{v}^B. \tag{16.13}$$

It follows that if P_i and P_j are particles of $\mathcal{B}$

$$|\boldsymbol{r}'_i - \boldsymbol{r}'_j| = |\boldsymbol{r}_i - \boldsymbol{r}_j| = c_{ij},$$

where the c_{ij} are constants, so that the rigidity conditions are also satisfied in $\mathcal{F}'$. But in $\mathcal{F}'$ the particle B is fixed (at the origin) and so Theorem 16.1 applies. Hence there exists a unique vector $\boldsymbol{\omega}(t)$ such that, for any particle of $\mathcal{B}$,

$$\boldsymbol{v}' = \boldsymbol{\omega} \times \boldsymbol{r}'. \tag{16.14}$$

On using (16.13) into (16.14) we obtain

$$\boldsymbol{v} - \boldsymbol{v}^B = \boldsymbol{\omega} \times (\boldsymbol{r} - \boldsymbol{b}),$$

that is

$$\boldsymbol{v} = \boldsymbol{v}^B + \boldsymbol{\omega} \times (\boldsymbol{r} - \boldsymbol{b}), \tag{16.15}$$

as required. ■

What the last theorem shows is that **any rigid body motion** can be resolved into a **translation** with velocity $\boldsymbol{v}^B$ and a **rotation** with some angular velocity $\boldsymbol{\omega}$ about an axis

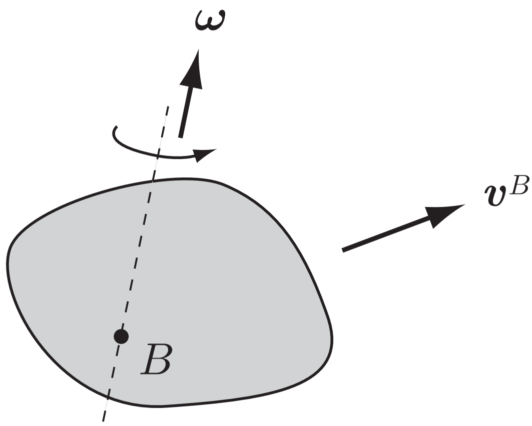


FIGURE 16.2 A rigid body in general motion. The velocities of its particles are the sum of those due to a **translation** with velocity $\boldsymbol{v}^B$ and a **rotation** with angular velocity $\boldsymbol{\omega}$ about an axis through B.

through B, where the particle B can be any particle of the body. This is shown in Figure 16.2.

Now suppose we choose a new reference particle C. Then the translational velocity would become $\boldsymbol{v}^C$, but what happens to the angular velocity $\boldsymbol{\omega}$? The answer is nothing; the angular velocity $\boldsymbol{\omega}$ is *independent* of the choice of reference particle. This means that we can refer to *the* angular velocity of a rigid body without specifying the reference particle. The proof of this is as follows:

Proof. Suppose that, with reference particles B, C, the body has angular velocities $\boldsymbol{\omega}^B, \boldsymbol{\omega}^C$ respectively. Then the velocity $\boldsymbol{v}$ of any particle P of the body is given by either of the two formulae

$$\boldsymbol{v} = \boldsymbol{v}^B + \boldsymbol{\omega}^B \times (\boldsymbol{r} - \boldsymbol{b}),$$
$$\boldsymbol{v} = \boldsymbol{v}^C + \boldsymbol{\omega}^C \times (\boldsymbol{r} - \boldsymbol{c}),$$

where $\boldsymbol{r}$ is the position vector of P. It follows that

$$\boldsymbol{v}^B + \boldsymbol{\omega}^B \times (\boldsymbol{r} - \boldsymbol{b}) = \boldsymbol{v}^C + \boldsymbol{\omega}^C \times (\boldsymbol{r} - \boldsymbol{c}) \tag{16.16}$$

for any $\boldsymbol{r}$ that is the position vector of a particle of the body. In particular, since B and C are particles of the body, it follows that

$$\boldsymbol{v}^B = \boldsymbol{v}^C + \boldsymbol{\omega}^C \times (\boldsymbol{b} - \boldsymbol{c}),$$

$$\boldsymbol{v}^B + \boldsymbol{\omega}^B \times (\boldsymbol{c} - \boldsymbol{b}) = \boldsymbol{v}^C,$$

and if we now subtract each of these formulae from the equality (16.16), we obtain

$$\boldsymbol{x} \times (\boldsymbol{r} - \boldsymbol{b}) = \boldsymbol{0},$$
$$\boldsymbol{x} \times (\boldsymbol{r} - \boldsymbol{c}) = \boldsymbol{0},$$

where $\boldsymbol{x} = \boldsymbol{\omega}^B - \boldsymbol{\omega}^C$. Now let P be any particle of the body not collinear with B and C and suppose that $\boldsymbol{x} \neq \boldsymbol{0}$. Then $\boldsymbol{x}$ must be parallel to both of the vectors $\boldsymbol{r} - \boldsymbol{b}$ and $\boldsymbol{r} - \boldsymbol{c}$, which are *not* parallel to each other. This is impossible and so $\boldsymbol{x} = \boldsymbol{0}$, which means that $\boldsymbol{\omega}^B = \boldsymbol{\omega}^C$. Hence the angular velocity of the body is the same, irrespective of the choice of reference particle. ■

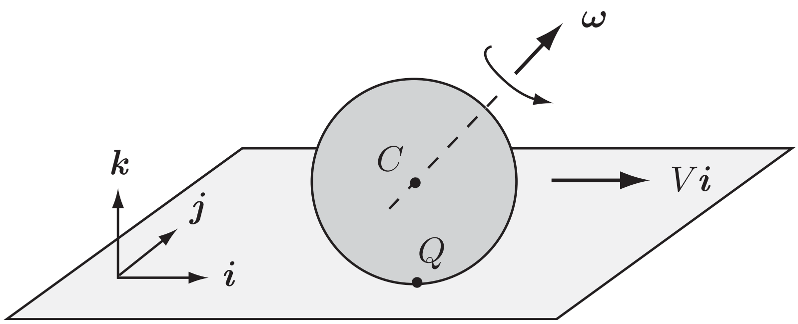


FIGURE 16.3 The ball rolls with velocity $V\boldsymbol{i}$ and has angular velocity $\boldsymbol{\omega}$.

Our results are summarised below:

Particle velocities in a rigid body

Suppose a rigid body is in general motion and that B is one of its particles. Then the velocity of any particle P of the body is given by the formula

$$\boldsymbol{v} = \boldsymbol{v}^B + \boldsymbol{\omega}\times(\boldsymbol{r} - \boldsymbol{b}), \tag{16.17}$$

where $\boldsymbol{v}^B$ is the velocity of the reference particle B and the angular velocity $\boldsymbol{\omega}$ is independent of the choice of reference particle.

Example 16.3 *Rolling snooker ball*

A rigid ball of radius b rolls without slipping on a horizontal table. Find the most general form of $\boldsymbol{\omega}$ consistent with the rolling condition.

Solution

Suppose the ball is rolling with velocity $V\boldsymbol{i}$ (where $\boldsymbol{i}$ is a horizontal unit vector), and has an unknown angular velocity $\boldsymbol{\omega}$. Then, on taking the centre C of the ball as the reference particle, the velocity of any particle P is given by (16.15) to be

$$\boldsymbol{v} = V\boldsymbol{i} + \boldsymbol{\omega}\times(\boldsymbol{r} - \boldsymbol{c}).$$

In particular, the velocity of the contact particle Q is given by

$$\boldsymbol{v}^Q = V\boldsymbol{i} + \boldsymbol{\omega}\times(-b\boldsymbol{k})\,,$$

where the unit vector $\boldsymbol{k}$ points vertically upwards. Since the rolling condition requires that $\boldsymbol{v}^Q = \boldsymbol{0}$, it follows that $\boldsymbol{\omega}$ must satisfy the condition

$$V\boldsymbol{i} + b\,\boldsymbol{k}\times\boldsymbol{\omega} = \boldsymbol{0}. \tag{16.18}$$

On taking the cross product of this equation with $\boldsymbol{k}$, we obtain

$$V\boldsymbol{k}\times\boldsymbol{i} + b\,\boldsymbol{k}\times(\boldsymbol{k}\times\boldsymbol{\omega}) = \boldsymbol{0},$$

that is

$$V\boldsymbol{j} + b\big((\boldsymbol{\omega}\cdot\boldsymbol{k})\boldsymbol{k} - (\boldsymbol{k}\cdot\boldsymbol{k})\boldsymbol{\omega}\big) = \boldsymbol{0}.$$

Since $\boldsymbol{k}$ is a unit vector, $\boldsymbol{k}\cdot\boldsymbol{k} = 1$ and we obtain

$$\boldsymbol{\omega} = \frac{V}{b}\,\boldsymbol{j} + (\boldsymbol{\omega}\cdot\boldsymbol{k})\boldsymbol{k}.$$

It follows that any $\boldsymbol{\omega}$ consistent with the rolling condition must have the form

$$\boldsymbol{\omega} = \frac{V}{b}\,\boldsymbol{j} + \lambda\,\boldsymbol{k}, \tag{16.19}$$

where λ is a scalar function of the time. Conversely, it is easy to verify that the formula (16.19) for $\boldsymbol{\omega}$ satisfies the rolling condition (16.18) for *any* choice of the scalar λ. This is therefore the most **general form** of $\boldsymbol{\omega}$ consistent with rolling.

This result is surprising at first. If the motion were planar, the value of ω would be V/b, the corresponding $\boldsymbol{\omega}$ being $(V/b)\,\boldsymbol{j}$. This is the special case $\lambda = 0$. But in general three dimensional rolling, $\lambda \neq 0$ and the rotation axis is not horizontal. This effect is well known to pool and snooker players and is achieved by striking the ball to the right (or left) of centre, thereby giving λ a positive (or negative) value. Players call this putting 'side' on the ball. It makes no difference to the rolling but affects the bounce when the ball strikes a cushion. ■

Time to relax Find a pool table and experiment by striking a ball slowly but firmly well to the right of centre. The marking on the ball should enable you to 'see' the rotation axis (a ball with spots is best). Check that giving the ball 'right hand side' produces a positive value of $\boldsymbol{\omega}\cdot\boldsymbol{k}$.

Example 16.4 *Wheel rolling around a circular path*

A circular wheel of radius b has its plane vertical and rolls with constant speed V around a circular path of radius R marked on a horizontal floor. Find the angular velocity of the wheel and the acceleration of the contact particle.

Solution

Suppose we view the motion of the wheel from the *rotating* reference frame $\{O;\ \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ shown in Figure 16.4. This frame rotates about the axis $\{O, \boldsymbol{k}\}$ with angular velocity $\boldsymbol{\Omega}$ given by

$$\boldsymbol{\Omega} = \dot{\theta}\boldsymbol{k} = \frac{V}{R}\,\boldsymbol{k}. \tag{16.20}$$

Viewed from the rotating frame, the wheel is rotating about a *fixed* axis parallel to the vector $\boldsymbol{i}$. On applying the rolling condition, the angular velocity $\boldsymbol{\omega}'$ of the wheel, viewed from the rotating frame, is given by

$$\boldsymbol{\omega}' = -\frac{V}{b}\boldsymbol{i}. \tag{16.21}$$

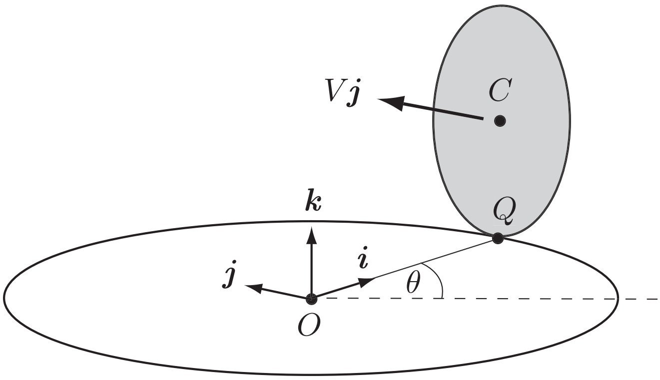


FIGURE 16.4 A wheel of radius b rolls around a circle of radius R. The unit vectors $\boldsymbol{i}(t)$ and $\boldsymbol{j}(t)$ follow the wheel as it moves around the circle; $\boldsymbol{k}$ is a constant unit vector.

The true angular velocity $\boldsymbol{\omega}$ of the wheel is then given by the sum

$$\boldsymbol{\omega} = \boldsymbol{\omega}' + \boldsymbol{\Omega}. \tag{16.22}$$

Here we are using a result from Chapter 17, the *angular velocity addition theorem*; it is the rotational counterpart of the addition theorem for linear velocities that we obtained in Chapter 2 . Although we have yet to prove this result, it is clear enough what it means and we will use it anyway! On substituting 16.20) and (16.21) into (16.22) we find the **angular velocity** of the wheel to be

$$\boldsymbol{\omega} = -\frac{V}{b}\boldsymbol{i} + \frac{V}{R}\boldsymbol{k}. \tag{16.23}$$

Now for the particle accelerations. These can be found by differentiating the velocity formula

$$\boldsymbol{v} = V\boldsymbol{j} + \boldsymbol{\omega}\times(\boldsymbol{r} - \boldsymbol{c}), \tag{16.24}$$

with respect to t, where we have taken C, the centre of the wheel, as the reference particle. This gives

$$\begin{aligned}\boldsymbol{a} &= V\frac{d\boldsymbol{j}}{dt} + \dot{\boldsymbol{\omega}}\times(\boldsymbol{r} - \boldsymbol{c}) + \boldsymbol{\omega}\times(\dot{\boldsymbol{r}} - \dot{\boldsymbol{c}})\\ &= V\frac{d\boldsymbol{j}}{dt} + \dot{\boldsymbol{\omega}}\times(\boldsymbol{r} - \boldsymbol{c}) + \boldsymbol{\omega}\times(\boldsymbol{v} - V\boldsymbol{j}),\end{aligned}$$

since $\dot{\boldsymbol{r}} = \boldsymbol{v}$ and $\dot{\boldsymbol{c}} = V\boldsymbol{j}$. In particular, since $\boldsymbol{r}^Q = \boldsymbol{c} - b\boldsymbol{k}$ and $\boldsymbol{v}^Q = \boldsymbol{0}$, the acceleration of the contact particle Q is given by

$$\begin{aligned}\boldsymbol{a}^Q &= V\frac{d\boldsymbol{j}}{dt} + \dot{\boldsymbol{\omega}}\times(-b\boldsymbol{k}) + \boldsymbol{\omega}\times(-V\boldsymbol{j})\\ &= V\frac{d\boldsymbol{j}}{dt} + V\frac{d\boldsymbol{i}}{dt}\times\boldsymbol{k} + \frac{V^2}{b}\boldsymbol{k} + \frac{V^2}{R}\boldsymbol{i},\end{aligned} \tag{16.25}$$

on using the formula (16.23) to replace $\boldsymbol{\omega}$ and $\dot{\boldsymbol{\omega}}$.

The only unknown quantities left are $d\boldsymbol{i}/dt$ and $d\boldsymbol{j}/dt$. However, the vectors $\boldsymbol{i}$, $\boldsymbol{j}$ correspond precisely to the polar unit vectors $\widehat{\boldsymbol{r}}$, $\widehat{\boldsymbol{\theta}}$ treated in section 2.3, from which we deduce that

$$\frac{d\boldsymbol{i}}{dt} = \dot{\theta}\,\boldsymbol{j} = \frac{V}{R}\,\boldsymbol{j} \qquad \text{and} \qquad \frac{d\boldsymbol{j}}{dt} = -\dot{\theta}\,\boldsymbol{i} = -\frac{V}{R}\,\boldsymbol{i}.$$

On substituting these formulae into equation (16.25) we obtain

$$\boldsymbol{a}^Q = \frac{V^2}{R}\,\boldsymbol{i} + \frac{V^2}{b}\,\boldsymbol{k}$$

as the **acceleration** of the contact particle Q. ■

Problems on Chapter 16

Answers and comments are at the end of the book.

Harder problems carry a star (∗).

16.1 A rigid body is rotating in the right-handed sense about the axis Oz with a constant angular speed of 2 radians per second. Write down the angular velocity vector of the body, and find the instantaneous velocity, speed and acceleration of the particle of the body at the point (4, −3, 7), where distances are measured in metres.

16.2 A rigid body is rotating with constant angular speed 3 radians per second about a fixed axis through the points $A(4, 1, 1)$, $B(2, -1, 0)$, distances being measured in centimetres. The rotation is in the left-handed sense relative to the direction $\overrightarrow{AB}$. Find the instantaneous velocity and acceleration of the particle P of the body at the point (4, 4, 4).

16.3 A spinning top (a rigid body of revolution) is in general motion with its vertex (a particle on the axis of symmetry) fixed at the origin O. Let $\boldsymbol{a}(t)$ be the unit vector pointing along the axis of symmetry and let $\boldsymbol{\omega}(t)$ be the angular velocity of the top. (In general, $\boldsymbol{\omega}$ does *not* point along the axis of symmetry.) By considering the velocities of particles of the top that lie on the axis of symmetry, show that $\boldsymbol{a}$ satisfies the equation

$$\dot{\boldsymbol{a}} = \boldsymbol{\omega} \times \boldsymbol{a}.$$

Deduce that the most general form $\boldsymbol{\omega}$ can have is

$$\boldsymbol{\omega} = \boldsymbol{a} \times \dot{\boldsymbol{a}} + \lambda\,\boldsymbol{a},$$

where λ is a scalar function of the time. [This formula is needed in the theory of the spinning top.]

16.4 A penny of radius a rolls without slipping on a rough horizontal table. The penny rolls in such a way that its centre G remains fixed (see Figure 16.5). The plane of the penny makes a constant angle α with the table and the point of contact C traces out a circle with centre O and

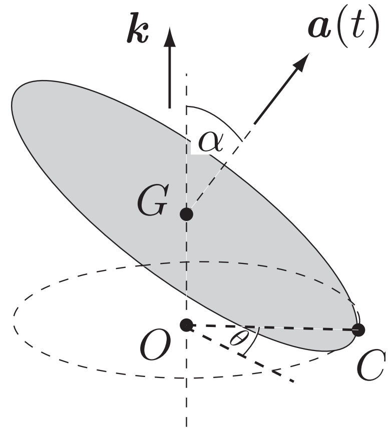


FIGURE 16.5 A penny of radius b rolls on a rough horizontal table in such a way that its centre G remains fixed.

radius $a \cos\alpha$, as shown. At time t, the angle between the radius OC and some fixed radius is θ. Find the angular velocity vector of the penny in terms of the unit vectors $\boldsymbol{a}(t)$, $\boldsymbol{k}$ shown.

Find the velocity of the highest particle of the penny.

16.5 A rigid circular cone with altitude h and semi-angle α rolls without slipping on a rough horizontal table. Explain why the vertex O of the cone never moves. Let $\theta(t)$ be the angle between OC, the line of the cone that is in contact with the table, and some fixed horizontal reference line OA. Show that the angular velocity $\boldsymbol{\omega}$ of the cone is given by

$$\boldsymbol{\omega} = -\left(\dot{\theta}\cot\alpha\right)\boldsymbol{i},$$

where $\boldsymbol{i}(t)$ is the unit vector pointing in the direction $\overrightarrow{OC}$. [First identify the *direction* of $\boldsymbol{\omega}$, and then consider the velocities of those particles of the cone that lie on the axis of symmetry.]

Identify the particle of the cone that has the maximum speed and find this speed.

16.6* Two rigid plastic panels lie in the planes $z = -b$ and $z = b$ respectively. A rigid ball of radius b can move in the space between the panels and is gripped by them so that it does not slip. The panels are made to rotate with angular velocities $\omega_1\boldsymbol{k}$, $\omega_2\boldsymbol{k}$ about fixed vertical axes that are a distance $2c$ apart. Show that, *with a suitable choice of origin*, the position vector $\boldsymbol{R}$ of the centre of the ball satisfies the equation

$$\dot{\boldsymbol{R}} = \boldsymbol{\Omega}\times\boldsymbol{R},$$

where $\boldsymbol{\Omega} = \frac{1}{2}(\boldsymbol{\omega}_1 + \boldsymbol{\omega}_2)$. Deduce that the ball must move in a circle and find the position of the centre of this circle. [This arrangement is sometimes seen as a shop window display. The panels are transparent and the ball seems to be executing a circle in mid-air.]

16.7 Two hollow spheres have radii a and b ($b > a$), and their common centre O is fixed. A rigid ball of radius $\frac{1}{2}(b-a)$ can move in the annular space between the spheres and is gripped by them so that it does not slip. The spheres are made to rotate with constant angular velocities $\boldsymbol{\omega}_1$, $\boldsymbol{\omega}_2$ respectively. Show that the ball must move in a circle whose plane is perpendicular to the vector $a\boldsymbol{\omega}_1 + b\boldsymbol{\omega}_2$.

Chapter Seventeen

Rotating reference frames

KEY FEATURES

The key features of this chapter are the **transformation** of **velocity** and **acceleration** between frames in general relative motion, and the dynamical **effects of the Earth's rotation**.

So far we have viewed the motion of mechanical systems from an inertial reference frame. The reason for this is simple; the Second Law, in its standard form, applies only in inertial frames. However, circumstances arise in which it is convenient to view the motion from a **non-inertial frame**. The most important instance of this occurs when the motion takes place near the surface of the Earth. Previously we have argued that the dynamical effects of the Earth's rotation are small enough to be neglected. While this is usually true, there are circumstances in which it has a significant effect. In long range artillery, the Earth's rotation gives rise to an important correction, and, in the hydrodynamics of the atmosphere and oceans, the Earth's rotation can have a dominant effect. If we wish to calculate such effects (as seen by an observer on the Earth), we must take our reference frame fixed to the Earth, thus making it a *non-inertial* frame. The downside of this choice is that the Second Law does not hold and must be replaced by a considerably more complicated equation.

In addition to applications involving the Earth's rotation, there are instances where the motion of a system looks much simpler when viewed from a suitably chosen rotating frame. The **Larmor precession** of a charged particle moving in a uniform magnetic field is one example; a second is the motion of a rigid body relative to its own principal axes of inertia, which leads to **Euler's equations**.

17.1 TRANSFORMATION FORMULAE

In this section we derive the transformation formulae that link the velocity and acceleration of a particle measured in a **moving frame**, with the same quantities measured in a **fixed frame**. For the purposes of *kinematics*, the labels 'fixed' and 'moving' are arbitrary. Each frame is moving relative to the other and the labels could be reversed; we use them purely for convenience. In *dynamics* however the distinction between the fixed and moving frames is real. The 'fixed' frame is an inertial frame, in which Newton's laws

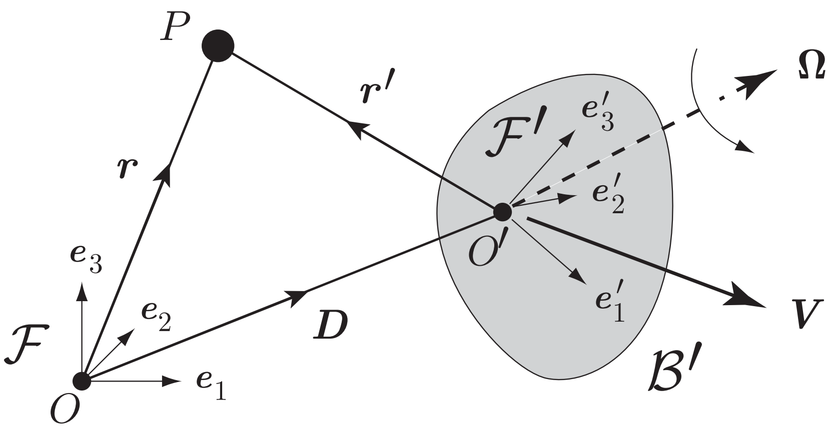


FIGURE 17.1 The moving frame $\mathcal{F}' \equiv \{O'; \boldsymbol{e}'_1, \boldsymbol{e}'_2, \boldsymbol{e}'_3\}$ has translational velocity $\boldsymbol{V}$ and angular velocity $\boldsymbol{\Omega}$ relative to the fixed frame $\mathcal{F} \equiv \{O; \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$.

apply, and the moving frame is a non-inertial frame, in which Newton's laws do not apply, at least in their standard form.

Let $\mathcal{F} \equiv \{O; \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ be the fixed frame and $\mathcal{F}' \equiv \{O'; \boldsymbol{e}'_1, \boldsymbol{e}'_2, \boldsymbol{e}'_3\}$ be the moving frame, as shown in Figure 17.1. At time t, the frame $\mathcal{F}'$ has translational velocity* $\boldsymbol{V}$ and angular velocity $\boldsymbol{\Omega}$ relative to the frame $\mathcal{F}$. It is convenient to regard the moving reference frame $\mathcal{F}'$ as being embedded in a rigid body $\mathcal{B}'$ with reference particle O'. Then $\boldsymbol{V}$ and $\boldsymbol{\Omega}$ are the translational and angular velocities of the body $\mathcal{B}'$. The most important feature of reference frames in relative motion is this:

The rate of change of a **vector** *quantity measured in the frame $\mathcal{F}$ is generally not the same as the rate of change of the same quantity measured in the frame $\mathcal{F}'$.*

Question ***Why are there different rates of change?***

Suppose $\boldsymbol{u}(t)$ is a vector quantity. Why should it have different rates of change when measured in the frames $\mathcal{F}$ and $\mathcal{F}'$?

Answer

Suppose the expression for $\boldsymbol{u}$ in terms of the basis set $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ is

$$\boldsymbol{u} = u_1\,\boldsymbol{e}_1 + u_2\,\boldsymbol{e}_2 + u_3\,\boldsymbol{e}_3, \tag{17.1}$$

and in terms of the basis set $\{\boldsymbol{e}'_1, \boldsymbol{e}'_2, \boldsymbol{e}'_3\}$ is

$$\boldsymbol{u} = u'_1\,\boldsymbol{e}'_1 + u'_2\,\boldsymbol{e}'_2 + u'_3\,\boldsymbol{e}'_3. \tag{17.2}$$

In general, the components $\{u_1, u_2, u_3\}$ and $\{u'_1, u'_2, u'_3\}$ will be functions of the time t.

* This means that the origin O' has velocity $\boldsymbol{V}$ relative to $\mathcal{F}$. The angular velocity $\boldsymbol{\Omega}$ is independent of the choice of O'.

In the rate of change of $\boldsymbol{u}$ measured in $\mathcal{F}$, the **basis set** $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ is, by definition, **constant** so that

$$\left(\frac{d\boldsymbol{u}}{dt}\right)_{\mathcal{F}} = \dot{u}_1\,\boldsymbol{e}_1 + \dot{u}_2\,\boldsymbol{e}_2 + \dot{u}_1\,\boldsymbol{e}_3. \tag{17.3}$$

In contrast, in the rate of change of $\boldsymbol{u}$ measured in $\mathcal{F}'$, the **basis set** $\{\boldsymbol{e}'_1, \boldsymbol{e}'_2, \boldsymbol{e}'_3\}$ is, by definition, **constant** so that

$$\left(\frac{d\boldsymbol{u}}{dt}\right)_{\mathcal{F}'} = \dot{u}'_1\,\boldsymbol{e}'_1 + \dot{u}'_2\,\boldsymbol{e}'_2 + \dot{u}'_3\,\boldsymbol{e}'_3. \tag{17.4}$$

Note that the components $\{u_k\}$ and $\{u'_k\}$ are *scalar* functions of the time and so their rates of change are *independent of the reference frame*. This is why we do not need to label them as being observed in $\mathcal{F}$ or $\mathcal{F}'$. There is no reason why the two expressions (17.3) and (17.4) should be equal and, in general, they are *not* equal. Consider, for example, the case in which $\boldsymbol{u}$ is constant in $\mathcal{F}'$ so that $(d\boldsymbol{u}/dt)_{\mathcal{F}'} = \boldsymbol{0}$. However, the motion of $\mathcal{F}'$ relative to $\mathcal{F}$ means that $\boldsymbol{u}$ will not be constant in $\mathcal{F}$ and $(d\boldsymbol{u}/dt)_{\mathcal{F}}$ will not be zero. ■

True and apparent values

In order to simplify the writing, we will, from now on, refer to the value of a quantity measured in the fixed frame $\mathcal{F}$ as its **true** value, and the value measured in the moving frame $\mathcal{F}'$ as its **apparent** value. For example, $(d\boldsymbol{u}/dt)_{\mathcal{F}}$ will be referred to the *true* value of $d\boldsymbol{u}/dt$, while $(d\boldsymbol{u}/dt)_{\mathcal{F}'}$ will be referred to as its *apparent* value.

Rates of change of the basis vectors $\{\boldsymbol{e}'_1, \boldsymbol{e}'_2, \boldsymbol{e}'_3\}$

Our first step is to find the rates of change of the fundamental basis vectors $\{\boldsymbol{e}'_1, \boldsymbol{e}'_2, \boldsymbol{e}'_3\}$ belonging to the frame $\mathcal{F}'$. Since these vectors are, by definition, constant in $\mathcal{F}'$, their *apparent* rates of change are zero. What we need to find are their *true* rates of change.

Let E be any particle of the body $\mathcal{B}'$ in which the frame $\mathcal{F}'$ is embedded, and let $\boldsymbol{e}$ and $\boldsymbol{e}'$ be the position vectors of E relative to O and O' respectively. Then, by the triangle law,

$$\boldsymbol{e} = \boldsymbol{D} + \boldsymbol{e}',$$

where $\boldsymbol{D}$ is the position vector of O' relative to O. Then

$$\left(\frac{d\boldsymbol{e}}{dt}\right)_{\mathcal{F}} = \left(\frac{d\boldsymbol{D}}{dt}\right)_{\mathcal{F}} + \left(\frac{d\boldsymbol{e}'}{dt}\right)_{\mathcal{F}} \tag{17.5}$$

$$= \boldsymbol{V} + \left(\frac{d\boldsymbol{e}'}{dt}\right)_{\mathcal{F}}, \tag{17.6}$$

where $\boldsymbol{V}$ is the true velocity of O' relative to O. But $(d\boldsymbol{e}/dt)_{\mathcal{F}}$ is, by definition, $\boldsymbol{v}^E$, the true velocity of the particle E, and, since E is a particle of the rigid body $\mathcal{B}'$, $\boldsymbol{v}^E$ is given by

$$\boldsymbol{v}^E = \boldsymbol{V} + \boldsymbol{\Omega}\times\boldsymbol{e}', \tag{17.7}$$

on using the kinematical formula (16.17). On comparing the equations (17.6) and (17.7), we see that

$$\left(\frac{d\boldsymbol{e}'}{dt}\right)_{\mathcal{F}} = \boldsymbol{\Omega}\times\boldsymbol{e}'.$$

This result applies to *any* vector $\boldsymbol{e}'$ that is the position vector (relative to O') of a particle of the rigid body $\mathcal{B}'$. In particular, since the basis vectors $\{\boldsymbol{e}'_1, \boldsymbol{e}'_2, \boldsymbol{e}'_3\}$ can be regarded as the position vectors of particles of $\mathcal{B}'$, we obtain the **fundamental relations**

$$\left(\frac{d\boldsymbol{e}'_j}{dt}\right)_{\mathcal{F}} = \boldsymbol{\Omega}\times\boldsymbol{e}'_j \qquad (1 \le j \le 3).$$

The notation $(d\boldsymbol{e}'_j/dt)_{\mathcal{F}}$ for the true rate of change of the vector $\boldsymbol{e}'_j$ is accurate but cumbersome. Since the apparent rate of change of these vectors is zero, there is little chance of confusion if, from now on, we replace $(d\boldsymbol{e}'_j/dt)_{\mathcal{F}}$ by the simple notation $\dot{\boldsymbol{e}}'_j$. Our result can then be expressed in the form:

True rates of change of the basis vectors $\{\boldsymbol{e}'_1, \boldsymbol{e}'_2, \boldsymbol{e}'_3\}$

$$\dot{\boldsymbol{e}}'_1 = \boldsymbol{\Omega}\times\boldsymbol{e}'_1 \qquad \dot{\boldsymbol{e}}'_2 = \boldsymbol{\Omega}\times\boldsymbol{e}'_2 \qquad \dot{\boldsymbol{e}}'_3 = \boldsymbol{\Omega}\times\boldsymbol{e}'_3 \tag{17.8}$$

where $\boldsymbol{\Omega}$ is the angular velocity of the frame $\mathcal{F}'$ relative to the frame $\mathcal{F}$.

Relation between the true and apparent values of $d\boldsymbol{u}/dt$

Our next step is to find the relationship between the true and apparent values of $d\boldsymbol{u}/dt$, where $\boldsymbol{u}$ is any vector function of the time. To do this we differentiate the representation (17.2) with respect to t *while keeping the basis set* $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ *constant*. This gives

$$\begin{aligned}
\left(\frac{d\boldsymbol{u}}{dt}\right)_{\mathcal{F}} &= \left(\frac{d(u'_1\boldsymbol{e}'_1)}{dt}\right)_{\mathcal{F}} + \left(\frac{d(u'_2\boldsymbol{e}'_2)}{dt}\right)_{\mathcal{F}} + \left(\frac{d(u'_3\boldsymbol{e}'_3)}{dt}\right)_{\mathcal{F}}\\
&= \left(\dot{u}'_1\,\boldsymbol{e}'_1 + \dot{u}'_2\,\boldsymbol{e}'_2 + \dot{u}'_3\,\boldsymbol{e}'_3\right) + u'_1\,\dot{\boldsymbol{e}}'_1 + u'_2\,\dot{\boldsymbol{e}}'_2 + u'_3\,\dot{\boldsymbol{e}}'_3\\
&= \left(\frac{d\boldsymbol{u}}{dt}\right)_{\mathcal{F}'} + u'_1\left(\boldsymbol{\Omega}\times\boldsymbol{e}'_1\right) + u'_2\left(\boldsymbol{\Omega}\times\boldsymbol{e}'_2\right) + u'_3\left(\boldsymbol{\Omega}\times\boldsymbol{e}'_3\right)\\
&= \left(\frac{d\boldsymbol{u}}{dt}\right)_{\mathcal{F}'} + \boldsymbol{\Omega}\times\left(u'_1\,\boldsymbol{e}'_1 + u'_2\,\boldsymbol{e}'_2 + u'_3\,\boldsymbol{e}'_3\right)\\
&= \left(\frac{d\boldsymbol{u}}{dt}\right)_{\mathcal{F}'} + \boldsymbol{\Omega}\times\boldsymbol{u}.
\end{aligned}$$

The true and apparent values of $d\boldsymbol{u}/dt$ are therefore related by the formula:

Transformation formula for $d\boldsymbol{u}/dt$

$$\left(\frac{d\boldsymbol{u}}{dt}\right)_{\mathcal{F}} = \left(\frac{d\boldsymbol{u}}{dt}\right)_{\mathcal{F}'} + \boldsymbol{\Omega}\times\boldsymbol{u} \tag{17.9}$$

where $\boldsymbol{\Omega}$ is the angular velocity of the frame $\mathcal{F}'$ relative to the frame $\mathcal{F}$.

Velocity transformation formula

It is now easy to find how particle velocities transform between the two frames. Suppose a particle P has position vector $\boldsymbol{r}$ relative to $\mathcal{F}$ and position vector $\boldsymbol{r}'$ relative to $\mathcal{F}'$. It follows from the triangle law that

$$\boldsymbol{r} = \boldsymbol{D} + \boldsymbol{r}', \tag{17.10}$$

where $\boldsymbol{D}$ is the position vector of O' relative to O. Then

$$\left(\frac{d\boldsymbol{r}}{dt}\right)_{\mathcal{F}} = \left(\frac{d\boldsymbol{D}}{dt}\right)_{\mathcal{F}} + \left(\frac{d\boldsymbol{r}'}{dt}\right)_{\mathcal{F}},$$

that is,

$$\boldsymbol{v} = \boldsymbol{V} + \left(\frac{d\boldsymbol{r}'}{dt}\right)_{\mathcal{F}},$$

where $\boldsymbol{v}$ is the true velocity of P. From the transformation formula (17.9), it follows that

$$\begin{aligned}\left(\frac{d\boldsymbol{r}'}{dt}\right)_{\mathcal{F}} &= \left(\frac{d\boldsymbol{r}'}{dt}\right)_{\mathcal{F}'} + \boldsymbol{\Omega}\times\boldsymbol{r}' \\ &= \boldsymbol{v}' + \boldsymbol{\Omega}\times\boldsymbol{r}',\end{aligned} \tag{17.11}$$

where $\boldsymbol{v}'$ is the apparent velocity of P. On combining these results, we obtain:

Velocity transformation formula

$$\boldsymbol{v} = \boldsymbol{V} + \boldsymbol{\Omega}\times\boldsymbol{r}' + \boldsymbol{v}' \tag{17.12}$$

This is the **velocity transformation formula** connecting the true and apparent velocities of the particle P.

This formula has a simple interpretation. The true velocity $\boldsymbol{v}$ is the sum of (i) the apparent velocity $\boldsymbol{v}'$, and (ii) the velocity $\boldsymbol{V} + \boldsymbol{\Omega}\times\boldsymbol{r}'$ that is given to the particle by the motion of the frame $\mathcal{F}'$.

An immediate consequence of the velocity transformation formula is the corresponding transformation rule for **angular velocities**.

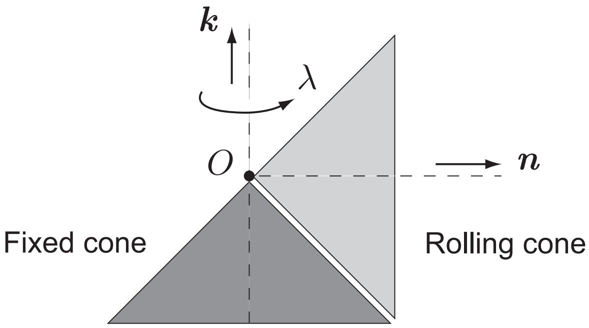


FIGURE 17.2 The dark shaded cone is fixed and the light shaded cone rolls around it. The vertex of each cone is at rest at O.

Theorem 17.1 ***Addition theorem for angular velocities*** *Suppose a rigid body in general motion has true angular velocity* $\boldsymbol{\omega}$ *and apparent angular velocity* $\boldsymbol{\omega}'$. *Then* $\boldsymbol{\omega}$ *and* $\boldsymbol{\omega}'$ *are related by the addition formula*

$$\boldsymbol{\omega} = \boldsymbol{\Omega} + \boldsymbol{\omega}', \tag{17.13}$$

where $\boldsymbol{\Omega}$ *is the angular velocity of the frame* $\mathcal{F}'$ *relative to* $\mathcal{F}$. [This result is the counterpart of the velocity transformation formula $\boldsymbol{v} = \boldsymbol{V} + \boldsymbol{v}'$ for *translating* frames.]

The proof of this theorem is the subject of Problem 17.2. We have already used this result in Chapter 16, Example 16.4, and we give a second example now.

Example 17.1 *Cone rolling on another cone*

A circular cone with semi-angle 45° is fixed with its axis of symmetry vertical and its vertex O upwards. An identical cone also has its vertex fixed at O and rolls on the first cone so that its axis of symmetry precesses around the upward vertical with angular speed λ as shown. Find the angular speed of the rolling cone.

Solution

Let $\mathcal{F}$ be a fixed frame and $\mathcal{F}'$ be a frame with origin at O and precessing with the rolling cone. Then $\mathcal{F}'$ has angular velocity $\lambda\,\boldsymbol{k}$ relative to $\mathcal{F}$, where the unit vector $\boldsymbol{k}$ points vertically upwards. In the frame $\mathcal{F}'$ the second cone has its axis of symmetry fixed and the first cone has angular velocity $-\lambda\,\boldsymbol{k}$. The rolling contact between the two cones means that the second cone must be rotating about its axis of symmetry $\{O, \boldsymbol{n}\}$ with angular speed λ in the negative sense. Its angular velocity is therefore $-\lambda\,\boldsymbol{n}$. In the notation of Theorem 17.1 we therefore have $\boldsymbol{\Omega} = \lambda\,\boldsymbol{k}$ and $\boldsymbol{\omega}' = -\lambda\,\boldsymbol{n}$. It follows that the **angular velocity** of the rolling cone is

$$\boldsymbol{\omega} = \boldsymbol{\Omega} + \boldsymbol{\omega}' = \lambda\,\boldsymbol{k} - \lambda\,\boldsymbol{n}.$$

The **angular speed** of the rolling cone is therefore $\sqrt{2}\,\lambda$. ■

Acceleration transformation formula

From the velocity transformation formula (17.12), it follows that

$$\left(\frac{d\boldsymbol{v}}{dt}\right)_{\mathcal{F}} = \left(\frac{d\boldsymbol{V}}{dt}\right)_{\mathcal{F}} + \left(\frac{d\boldsymbol{\Omega}}{dt}\right)_{\mathcal{F}} \times \boldsymbol{r}' + \boldsymbol{\Omega} \times \left(\frac{d\boldsymbol{r}'}{dt}\right)_{\mathcal{F}} + \left(\frac{d\boldsymbol{v}'}{dt}\right)_{\mathcal{F}},$$

that is,

$$\boldsymbol{a} = \boldsymbol{A} + \dot{\boldsymbol{\Omega}} \times \boldsymbol{r}' + \boldsymbol{\Omega} \times \left(\frac{d\boldsymbol{r}'}{dt}\right)_{\mathcal{F}} + \left(\frac{d\boldsymbol{v}'}{dt}\right)_{\mathcal{F}}, \qquad (17.14)$$

where $\boldsymbol{a}$ is the true acceleration of the particle P, and $\boldsymbol{A}$, $\boldsymbol{\Omega}$ and $\dot{\boldsymbol{\Omega}}$ are the translational acceleration, the angular velocity, and the angular acceleration of the frame $\mathcal{F}'$, all relative to the frame $\mathcal{F}$.

The derivative $\left(d\boldsymbol{r}'/dt\right)_{\mathcal{F}}$ has already been evaluated in (17.11), and the derivative $\left(d\boldsymbol{v}'/dt\right)_{\mathcal{F}}$ is given by the transformation formula (17.9) to be

$$\begin{aligned}\left(\frac{d\boldsymbol{v}'}{dt}\right)_{\mathcal{F}} &= \left(\frac{d\boldsymbol{v}'}{dt}\right)_{\mathcal{F}'} + \boldsymbol{\Omega} \times \boldsymbol{v}', \\ &= \boldsymbol{a}' + \boldsymbol{\Omega} \times \boldsymbol{v}',\end{aligned} \qquad (17.15)$$

where $\boldsymbol{a}'$ is the apparent acceleration of P. On combining these results, we obtain

$$\boldsymbol{a} = \boldsymbol{A} + \dot{\boldsymbol{\Omega}} \times \boldsymbol{r}' + \boldsymbol{\Omega} \times \left(\boldsymbol{v}' + \boldsymbol{\Omega} \times \boldsymbol{r}'\right) + \left(\boldsymbol{a}' + \boldsymbol{\Omega} \times \boldsymbol{v}'\right)$$

which, on simplification, gives the **acceleration transformation formula**:

Acceleration transformation formula

$$\boldsymbol{a} = \boldsymbol{A} + \dot{\boldsymbol{\Omega}} \times \boldsymbol{r}' + 2\,\boldsymbol{\Omega} \times \boldsymbol{v}' + \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \boldsymbol{r}') + \boldsymbol{a}' \qquad (17.16)$$

where $\boldsymbol{A}$ is the translational acceleration, $\boldsymbol{\Omega}$ is the angular velocity and $\dot{\boldsymbol{\Omega}}$ is the angular acceleration of the frame $\mathcal{F}'$, all relative to the frame $\mathcal{F}$.

The acceleration formula does not have a simple interpretation. The apparent acceleration is $\boldsymbol{a}'$ and the acceleration given to the particle by the motion of the frame $\mathcal{F}'$ is $\boldsymbol{A} + \dot{\boldsymbol{\Omega}} \times \boldsymbol{r}' + \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \boldsymbol{r}')$. However, the true acceleration is *not* the sum of these two contributions! There is the additional term $2\,\boldsymbol{\Omega} \times \boldsymbol{v}'$ that appears *only when the particle is moving within the moving frame*. This **Coriolis*** term, as it is called, is hard to explain physically.

* After Gaspard-Gustave de Coriolis (1792–1843) whose paper on the subject was published in 1835. He also introduced the notion of 'work' with its present scientific meaning.

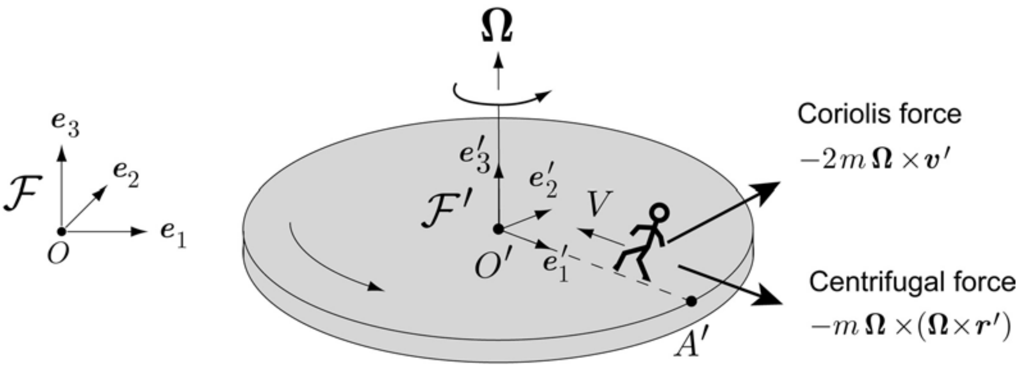


FIGURE 17.3 Man walking on a rotating roundabout.

17.2 PARTICLE DYNAMICS IN A NON-INERTIAL FRAME

Suppose now that $\mathcal{F}$ is an **inertial frame**. Then the equation of motion of a particle P of mass m moving under the force $\boldsymbol{F}$ is simply given by the Second Law

$$m\boldsymbol{a} = \boldsymbol{F}, \tag{17.17}$$

where $\boldsymbol{a}$ is the true acceleration of P. It follows from the acceleration transformation formula (17.16), that, when the *same motion* is observed from the frame $\mathcal{F}'$, equation (17.17) is replaced by the **transformed equation of motion**

Second Law in a general non-inertial frame

$$m\left[\boldsymbol{A} + \dot{\boldsymbol{\Omega}}\times\boldsymbol{r}' + 2\,\boldsymbol{\Omega}\times\boldsymbol{v}' + \boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\boldsymbol{r}') + \boldsymbol{a}'\right] = \boldsymbol{F}. \tag{17.18}$$

Here $\boldsymbol{r}'$, $\boldsymbol{v}'$, $\boldsymbol{a}'$ are the apparent position, velocity and acceleration of P, and $\boldsymbol{A}$, $\boldsymbol{\Omega}$ and $\dot{\boldsymbol{\Omega}}$ are the translational acceleration, angular velocity and angular acceleration of the frame $\mathcal{F}'$ relative to the frame $\mathcal{F}$.

This is the form taken by the Second Law in a general **non-inertial frame**. Note that there is *no new physics* in this equation. It is simply the Second Law written in terms of the apparent quantities observed in the non-inertial frame $\mathcal{F}'$.

Example 17.2 *Man walking on a rotating roundabout*

A fairground roundabout is made to rotate with constant angular speed Ω about its axis of symmetry which is fixed in a vertical position, the sense of the rotation being as shown in Figure 17.3. A man is on the roundabout and is walking grimly towards the centre O' along the moving radius $A'O'$ with constant speed V. What is the force that the roundabout exerts on the man?

Solution

Let the frame $\mathcal{F} \equiv \{O; \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ be an inertial frame of reference attached to the ground and the frame $\mathcal{F}' \equiv \{O'; \boldsymbol{e}'_1, \boldsymbol{e}'_2, \boldsymbol{e}'_3\}$ be attached to the roundabout, as shown in Figure 17.3. Then $\mathcal{F}'$ has zero translational velocity and constant angular velocity $\boldsymbol{\Omega}\ (= \Omega\, \boldsymbol{e}_3)$ relative to $\mathcal{F}$. Hence $\boldsymbol{A} = \boldsymbol{0}$ and $\dot{\boldsymbol{\Omega}} = \boldsymbol{0}$ in this problem.

In the frame $\mathcal{F}'$, the motion of the man is uniform and rectilinear so that $\boldsymbol{r}' = x'_1\, \boldsymbol{e}'_1$, $\boldsymbol{v}' = -V\, \boldsymbol{e}'_1$ and $\boldsymbol{a}' = \boldsymbol{0}$. The transformed equation of motion (17.18) therefore reduces to

$$m\left[2\,(\Omega\,\boldsymbol{e}_3)\times(-V\,\boldsymbol{e}'_1) + (\Omega\,\boldsymbol{e}_3)\times\big((\Omega\,\boldsymbol{e}_3)\times(x'_1\,\boldsymbol{e}'_1)\big)\,\right] = -mg\,\boldsymbol{e}_3 + \boldsymbol{X},$$

where $\boldsymbol{X}$ is the force that the roundabout exerts on the man. Hence, $\boldsymbol{X}$ is given by

$$\boldsymbol{X} = (mg)\,\boldsymbol{e}_3 - (2m\Omega V)\,\boldsymbol{e}'_2 - (m\Omega^2 x'_1)\,\boldsymbol{e}'_1.$$

Because of the additional terms* in this expression, the man must lean forwards and to his left; otherwise he will fall over. ■

Fictitious forces

Equation (17.18) can be made to resemble the *standard* form of the Second Law by the simple device of transferring the four new terms on the left to the right side of the equation and regarding them as **fictitious forces** that act on P in addition to the real force $\boldsymbol{F}$. This gives

$$m\boldsymbol{a}' = \boldsymbol{F} + (-m\boldsymbol{A}) + (-m\,\dot{\boldsymbol{\Omega}}\times\boldsymbol{r}') + \underbrace{(-2m\,\boldsymbol{\Omega}\times\boldsymbol{v}')}_{\text{Coriolis force}} + \underbrace{(-m\,\boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\boldsymbol{r}'))}_{\text{centrifugal force}}, \tag{17.19}$$

which we will call the **fictitious force equation**.

In other words, the fact that the frame $\mathcal{F}'$ is non-inertial may be ignored provided that the four fictitious forces $-m\boldsymbol{A}$, $-m\dot{\boldsymbol{\Omega}}\times\boldsymbol{r}'$, $-2m\,\boldsymbol{\Omega}\times\boldsymbol{v}'$ and $-m\boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\boldsymbol{r}')$ are added to the real force $\boldsymbol{F}$. Two of these 'forces' have names. The 'force' $-2m\,\boldsymbol{\Omega}\times\boldsymbol{v}'$ is called the **Coriolis force** and the 'force' $-m\,\boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\boldsymbol{r}')$ is called the **centrifugal force**.† In the problem of the man on the rotating roundabout, only the Coriolis and centrifugal 'forces' are non-zero and their directions are shown in Figure 17.3.

The transformed equation of motion and the fictitious force equation are obviously equivalent and there is no logical need to introduce the notion of fictitious forces at all. However, the fictitious force approach is often useful when looking at a problem *qualitatively*. For example, in the problem of the man walking on the roundabout, the fictitious force description would be as follows:

* If the roundabout were at rest, then $\boldsymbol{X}$ would be simply $mg\,\boldsymbol{e}_3$.

† Although the name 'centrifugal force' for the expression $-m\,\boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\boldsymbol{r}')$ is standard in the literature, it is not really appropriate unless $O' = O$, that is, the two frames of reference have a common origin.

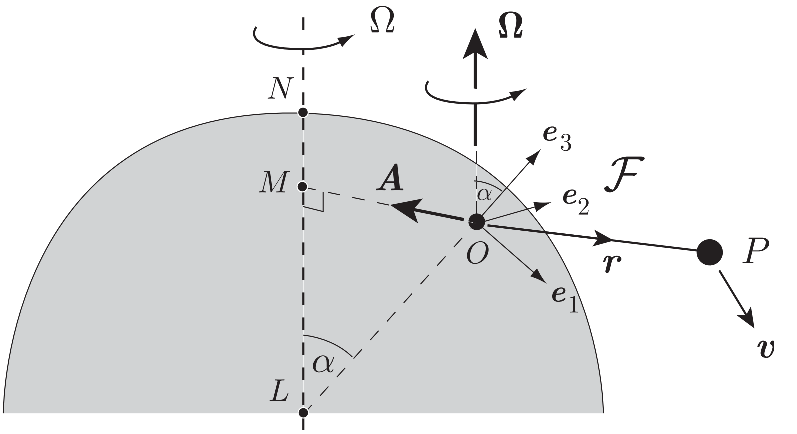


FIGURE 17.4 The inertial frame $\mathcal{F} \equiv \{O; \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ is attached to the surface of the Earth.

'Forget that the roundabout is rotating, and introduce the Coriolis and centrifugal forces as shown in Figure 17.3. The man will then feel the Coriolis force pushing him to his right and the centrifugal force pushing him backwards. This is why he must lean forwards and to his left'.

Such arguments are fine at the qualitative level, where they provide a picture of what is going on. However, it is asking for trouble to have two parallel approaches available when solving problems. A choice must be made and, in this book, *we will always solve problems in rotating frame dynamics by using the transformed equation of motion* (17.18); *ficticious forces will only be mentioned in passing*.

17.3 MOTION RELATIVE TO THE EARTH

The most common non-inertial reference frame is the Earth. The effects of the Earth's motion are generally small and we have so far completely neglected them when solving problems in dynamics. However, there are circumstances in which the Earth's motion has a significant effect. In long range artillery, the Earth's rotation gives rise to an important correction,* and, in the hydrodynamics of the atmosphere and oceans, the Earth's rotation can have a dominant effect.

The Earth moves in its orbit around the Sun and also rotates on its axis. The effect of its orbital motion is *very* much smaller than the effect of its rotation and we will neglect

* Artillery shells are deflected to the *right* in the northern hemisphere and to the *left* in the southern. The following anecdote is due to the great English mathematician J.E. Littlewood. '*I heard an account of the battle of the Falkland Islands (early in the 1914 war) from an officer who was there. The German ships were destroyed at extreme range, but it took a long time and salvos were continually falling 100 yards to the left. The effect of the rotation of the Earth was incorporated into the gun-sights. But this involved the tacit assumption that Naval battles take place round about latitude* 50°*N, whereas the Falkland Islands are at about latitude* 50°*S. At extreme range, this double difference is of the order of 100 yards!*

the orbital motion of the Earth altogether. We will regard the Earth as an *axially symmetric rigid body** whose axis of symmetry is at rest in an inertial frame $\mathcal{I}$. Previously we have denoted the moving frame by $\mathcal{F}'$, but it would be insufferable to attach a dash to every quantity from now until the end of the chapter. We will therefore drop the dashes and, from now on, we will let $\mathcal{F} \equiv \{O; \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ be a *non-inertial* frame attached to the surface of the Earth, as shown in Figure 17.4. The orientations of the basis vectors $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ will be described later. Because of the Earth's rotation, the frame $\mathcal{F}$ has an acceleration $\boldsymbol{A}$ and an angular velocity $\boldsymbol{\Omega}$ relative to the inertial frame $\mathcal{I}$. The acceleration $\boldsymbol{A}$ is caused by the circular motion of the origin O around the Earth's rotation axis and acts towards the axis. The angular velocity $\boldsymbol{\Omega}$ is the same as the angular velocity of the Earth itself; it is therefore parallel to the rotation axis and *constant*. If $\boldsymbol{n}$ is the unit vector in the direction $\overrightarrow{SN}$, where S and N are the north and south poles, then $\boldsymbol{\Omega} = +\Omega\boldsymbol{n}$, where Ω is 2π radians per day,† or about 0.000073 radians per second.

Let P be a particle of mass m and position vector $\boldsymbol{r}$ relative to O. Then the transformed equation of motion for P is

$$m\left[\boldsymbol{A} + 2\,\boldsymbol{\Omega}\times\boldsymbol{v} + \boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\boldsymbol{r}) + \boldsymbol{a}\right] = \boldsymbol{F}^E. \tag{17.20}$$

Here $\boldsymbol{r}$, $\boldsymbol{v}$ and $\boldsymbol{a}$ are the apparent position, velocity and acceleration of P, and $\boldsymbol{F}^E$ is the gravitational force exerted on P by the Earth. There may be other forces acting on P, but, for the moment, we suppose that the only force is the Earth's gravity.

Suppose first that P is *released from rest* (in $\mathcal{F}$) at the point with position vector $\boldsymbol{r}$. Then its initial acceleration $\boldsymbol{a}_0$ is given by equation (17.20) to be

$$\boldsymbol{a}_0 = m^{-1}\boldsymbol{F}^E - \boldsymbol{A} - \boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\boldsymbol{r}). \tag{17.21}$$

If the Earth were not rotating, this acceleration would be the acceleration due to the Earth's gravity at the point $\boldsymbol{r}$. With rotation present, it is still the observed acceleration of a body released from rest, but it is not caused by gravity alone. We call this the **apparent gravitational acceleration** (at the point $\boldsymbol{r}$). The direction opposite to $\boldsymbol{a}_0$ is the **apparent vertical** direction. Let $\boldsymbol{k}\,(=\boldsymbol{k}(\boldsymbol{r}))$ be the unit vector in the apparent vertical direction. Then $\boldsymbol{a}^I$ can be written in the form

$$\boldsymbol{a}_0 = -g(\boldsymbol{r})\,\boldsymbol{k}(\boldsymbol{r}), \tag{17.22}$$

where g is the magnitude of the **apparent gravitational acceleration**. These values of g and $\boldsymbol{k}$ differ slightly (by less than 1%) from the corresponding quantities due to Earth's gravity alone. In terms of the quantities g and $\boldsymbol{k}$, the transformed equation of motion for P simplifies to give

$$m\left[\frac{d\boldsymbol{v}}{dt} + 2\,\boldsymbol{\Omega}\times\boldsymbol{v}\right] = -m\,g(\boldsymbol{r})\,\boldsymbol{k}(\boldsymbol{r}).$$

* We are not restricted to regard the Earth as spherically symmetric.

† Actually, the Earth revolves not once every day, but every 23 h 56 m. Why?

As in the non-rotational theory, we will usually suppose that the extent of the motion of P is small compared to the Earth's radius so that the spatial variations of $g(\boldsymbol{r})$ and $\boldsymbol{k}(\boldsymbol{r})$ can be neglected. In this approximation, the transformed equation of motion for P becomes

Projectile equation on a rotating Earth

$$m\left[\frac{d\boldsymbol{v}}{dt} + 2\boldsymbol{\Omega}\times\boldsymbol{v}\right] = -mg\,\boldsymbol{k} \tag{17.23}$$

where g and $\boldsymbol{k}$ are now constants. This is the **projectile equation** for a **rotating Earth**. If any other forces act on P, they should be included on the right side of this equation. Note that, by introducing apparent gravity, this equation differs from the standard projectile equation only by the appearance of the Coliolis term.

We will now assign directions to the basis vectors $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ of our frame $\mathcal{F}$. The basis vector $\boldsymbol{k}$ is taken to be the same as the apparent upward vertical vector $\boldsymbol{k}$ (now assumed constant); the vector $\boldsymbol{i}$ is taken so that the Earth's rotation axis lies in the plane Oxz (see Figure 17.4); the vector $\boldsymbol{j}$ is then determined. In geographical language, $\boldsymbol{i}$ points south and $\boldsymbol{j}$ points east. With these choices, the Earth's angular velocity is then

$$\boldsymbol{\Omega} = \Omega(-\sin\beta\,\boldsymbol{i} + \cos\beta\,\boldsymbol{k}), \tag{17.24}$$

where β is the angle between the apparent vertical at O and the Earth's rotation axis. β is approximately $90° - \theta$, where θ is the geographical latitude of O. [Why only approximately?] Despite this small discrepancy, we will call the angle β the **co-latitude** of O.

Example 17.3 *Tower of Pisa experiment*

A lead ball is dropped from the top of the tower of Pisa. Show that its path deviates from that of a plumbline and find where it lands. [Neglect air resistance but include the effect of the Earth's rotation.]

Solution

Consider first the **plumbline**. The bob is subject to the apparent gravity force $-mg\boldsymbol{k}$ and the string tension force $\boldsymbol{T}$. When the plumbline is in equilibrium it follows from equation (17.23) that $\boldsymbol{0} = -mg\boldsymbol{k} + \boldsymbol{T}$, that is, $\boldsymbol{T} = mg\boldsymbol{k}$. Hence the tension in the plumbline is mg and the *string is parallel to the apparent vertical*.

Now consider the **ball**. Take the origin O of coordinates at the point where the ball is released and suppose that the apparent vertical through O meets the ground at the point $(0, 0, -h)$. The equation of motion for the ball is the **projectile equation** (17.23), together with the initial conditions $\boldsymbol{v} = \boldsymbol{0}$ and $\boldsymbol{r} = \boldsymbol{0}$ when $t = 0$. Since $\boldsymbol{\Omega}$ is constant, equation (17.23) can easily be integrated once with respect to t to give

$$\frac{d\boldsymbol{r}}{dt} + 2\,\boldsymbol{\Omega}\times\boldsymbol{r} = -gt\,\boldsymbol{k} + \boldsymbol{C},$$

where $\boldsymbol{C}$ is the constant of integration. The initial conditions then imply that $\boldsymbol{C} = \boldsymbol{0}$ so that the displacement of the ball satisfies the first order equation

$$\frac{d\boldsymbol{r}}{dt} + 2\,\boldsymbol{\Omega}\times\boldsymbol{r} = -gt\,\boldsymbol{k} \tag{17.25}$$

with the initial condition $\boldsymbol{r} = \boldsymbol{0}$ when $t = 0$. Since we expect the effect of the Earth's rotation to be a small correction, we will solve this equation approximately by an iterative method. On integrating equation (17.25) with respect to t and using the initial condition, we find that the unknown displacement $\boldsymbol{r}(t)$ satisfies the **integral equation***

$$\boldsymbol{r}(t) = -\tfrac{1}{2}gt^2\,\boldsymbol{k} - 2\,\boldsymbol{\Omega}\times\int_0^t \boldsymbol{r}(t')\,dt'. \tag{17.26}$$

(The variable of integration has been changed to the dummy variable t' because t appears as a limit of integration.) We now solve this equation approximately by iteration. The zeroth order approximation $\boldsymbol{r}^{(0)}$ corresponds to the case $\Omega = 0$, that is, the case of the non-rotating Earth. Thus the **zeroth order** approximation is

$$\boldsymbol{r}^{(0)} = -\tfrac{1}{2}gt^2\,\boldsymbol{k}, \tag{17.27}$$

which is just the elementary solution for vertical motion under uniform gravity. The first order approximation $\boldsymbol{r}^{(1)}$ is now obtained by substituting the zeroth order approximation into the integral term in equation (17.26). This gives

$$\begin{aligned}
\boldsymbol{r}^{(1)} &= -\tfrac{1}{2}gt^2\,\boldsymbol{k} - 2\boldsymbol{\Omega}\times\int_0^t \boldsymbol{r}^{(0)}(t')\,dt' \\
&= -\tfrac{1}{2}gt^2\,\boldsymbol{k} - 2\boldsymbol{\Omega}\times\int_0^t \left(-\tfrac{1}{2}gt'^2\,\boldsymbol{k}\right)\,dt' \\
&= -\tfrac{1}{2}gt^2\,\boldsymbol{k} + \tfrac{1}{3}gt^3(\boldsymbol{\Omega}\times\boldsymbol{k}).
\end{aligned}$$

Now the angular velocity $\boldsymbol{\Omega}$ is given by equation (17.24), where β is the co-latitude of the point O. It follows that $\boldsymbol{\Omega}\times\boldsymbol{k} = (\Omega\sin\beta)\,\boldsymbol{j}$ and so the **first order approximation** for the displacement $\boldsymbol{r}(t)$ is given by

$$\boldsymbol{r}^{(1)} = -\tfrac{1}{2}gt^2\,\boldsymbol{k} + \tfrac{1}{3}gt^3(\Omega\sin\beta)\,\boldsymbol{j}. \tag{17.28}$$

Higher approximations can be obtained in the same way, but they are progressively less important† (and more complicated). The first order appoximation predicts that

* Any equation in which an unknown function appears under an integral sign is called an *integral equation*. There is an extensive theory of such equations. Integral equations like (17.26) have the impressive title of *Volterra integral equations of the second kind*. Their most important feature is that they can always be solved by the iteration process described.

† The iterative approximation scheme described above will converge rapidly when the dimensionless product $\Omega\tau$ is small, where τ is the total travel time. ($\Omega\tau$ is the angle through which the Earth rotates during the motion.) In the Tower of Pisa problem, the value of $\Omega\tau$ is about 0.0002, which explains why the correction is so small.

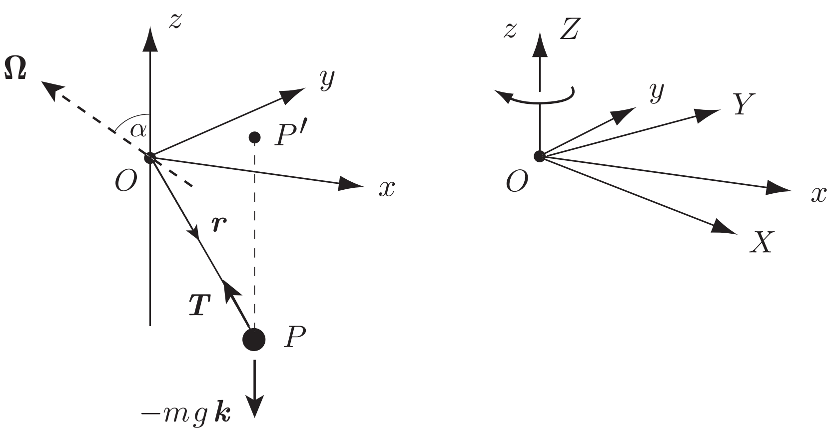


FIGURE 17.5 **Left**: The Foucault pendulum is an ordinary pendulum suspended from a point O fixed to the Earth. **Right**: The precessing frame $OXYZ$ rotates around the axis Oz with angular velocity $-(\Omega \cos \beta)\boldsymbol{k}$.

the ball will not travel down the apparent vertical, but will be displaced in the positive $\boldsymbol{j}$-direction, that is, to the east. It remains to determine this easterly displacement in terms of h, the height of the tower. For this we need the total travel time τ, but it is consistent with the first order approximation to determine τ from the *zero order* theory which gives $\tau = (2h/g)^{1/2}$. Hence, because of the Earth's rotation, the ball does not travel down the apparent vertical exactly but **drifts to the east** a distance

$$\left(\frac{8h^3}{9g}\right)^{1/2} \Omega \sin \beta.$$

For a ball dropped from the top of the **Tower of Pisa**, which is about 50 m high and is at latitude 44°N, this easterly drift is about 5.6 mm. ■

Example 17.4 *The Foucault pendulum*

How does the Earth's rotation affect the small oscillations of an ordinary pendulum?

Solution

Let the origin O be the suspension point of the pendulum as shown in Figure 17.5. The **transformed equation of motion** for the pendulum bob P is

$$m\left[\frac{d\boldsymbol{v}}{dt} + 2\boldsymbol{\Omega} \times \boldsymbol{v}\right] = -mg\,\boldsymbol{k} + \boldsymbol{T},$$

where $\boldsymbol{T}$ is the tension force exerted by the string. The magnitude T of the tension force is unknown but its direction is opposite to the position vector $\boldsymbol{r}$ so that

$$\boldsymbol{T} = -\left(\frac{T}{a}\right)\boldsymbol{r},$$

where a is the length of the string. The equation of motion for P can therefore be written

$$m\left[\frac{d\boldsymbol{v}}{dt}+2\,\boldsymbol{\Omega}\times\boldsymbol{v}\right]=-mg\,\boldsymbol{k}-\left(\frac{T}{a}\right)\boldsymbol{r}. \tag{17.29}$$

This non-linear equation, which holds for large oscillations of the pendulum, cannot be solved in closed form. We will therefore restrict attention to the case of **small oscillations** which allows us to replace (17.29) by a **linear approximation**. The zero order approximation corresponds to the 'motion' $\boldsymbol{r}=-a\boldsymbol{k}$, in which the pendulum hangs in equilibrium. Thus the zero order approximation to the tension is $T=mg$.

Furthermore, if we write $\boldsymbol{r}=x\boldsymbol{i}+y\boldsymbol{j}+z\boldsymbol{k}$, then, in any motion of the pendulum, $x^2+y^2+z^2=a^2$, from which it follows that

$$\begin{aligned} z&=\left(a^2-(x^2+y^2)\right)^{1/2}\\ &=a\left(1-\frac{x^2+y^2}{2a^2}+\cdots\right), \end{aligned}$$

which shows that the *change* in z is second order in the small quantities x and y. It can similarly be shown that v_z and dv_z/dt are also second order quantities. On using these approximations in equation (17.29), **linearised equation** for the **horizontal motion** of P is found to be

$$\frac{d\boldsymbol{v}^H}{dt}+2\Omega\cos\beta(\boldsymbol{k}\times\boldsymbol{v}^H)+n^2\boldsymbol{r}^H=\boldsymbol{0}, \tag{17.30}$$

where $\boldsymbol{r}^H=x\boldsymbol{i}+y\boldsymbol{j}$, $\boldsymbol{v}^H=v_x\boldsymbol{i}+v_y\boldsymbol{j}$ and $n^2=g/a$. Note that n is the angular frequency the pendulum pendulum would have if the Earth were not rotating. Equation (17.30) is the linearised equation for the motion of the point P' shown in Figure 17.5.

The two-dimensional vector equation (17.30) is equivalent to a pair of homogeneous second order linear equations with constant coefficients and can therefore be solved by standard methods. However, this is a messy job and there is a much simpler way of generating the solutions. The term $2\Omega\cos\beta(\boldsymbol{k}\times\boldsymbol{v}^H)$ can be eliminated by viewing the motion from a frame* $OXYZ$ rotating with angular velocity $-(\Omega\cos\beta)\boldsymbol{k}$ relative to the frame $\mathcal{F}$ (see Figure 17.5). We will call this new frame the **precessing frame** and the horizontal displacement, velocity and acceleration of P in this frame will be denoted by $\boldsymbol{R}^H$, $\boldsymbol{V}^H$ and $\boldsymbol{A}^H$. The standard transformation formulae then give

$$\boldsymbol{v}^H=\boldsymbol{V}^H+(-\Omega\cos\beta\boldsymbol{k})\times\boldsymbol{R},\qquad \boldsymbol{a}^H=\boldsymbol{A}^H+(-\Omega\cos\beta\boldsymbol{k})\times\boldsymbol{V} \tag{17.31}$$

and if these formulae are substituted into equation (17.30), we find that the linearised equation of motion for P in the precessing frame can be written in the form

$$\frac{d^2\boldsymbol{R}^H}{dt^2}+\left(n^2-\Omega^2\cos^2\beta\right)\boldsymbol{R}^H=\boldsymbol{0}, \tag{17.32}$$

* Note that this operation does *not* simply return us to the underlying inertial frame.

which is the **two-dimensional SHM equation**. Its solutions are known to be ellipses with centre O and angular frequency $(n^2 - \Omega^2 \cos^2 \beta)^{1/2}$. The difference between this frequency and n is completely negligible.

Conclusion *In small oscillations, the period of the pendulum is unchanged by the Earth's rotation, but its motion now precesses* around the apparent vertical with angular velocity* $-(\Omega \cos \beta)\boldsymbol{k}$. *In particular, an oscillation that would take place in a fixed vertical plane through* O *when* $\Omega = 0$ *will now take place in a vertical plane through* O *that precesses slowly around the axis* Oz *with angular velocity* $-(\Omega \cos \beta)\boldsymbol{k}$. *This precession is in the negative sense in the northern hemisphere, and in the positive sense in the southern hemisphere.* ■

This practical demonstration of the Earth's rotation is called **Foucault's pendulum**† and is a favourite exhibit in science museums the world over. [But not in Singapore. Why?] From the duration of your museum visit, and the angle turned by the plane of motion of the pendulum while you were inside, you can work out your latitude!

Winds, ocean currents and bathwater

Atmospheric winds and ocean currents are often strongly influenced by the Coriolis force induced by the Earth's rotation. We may analyse these fluid motions by analogy with particle dynamics.

By analogy with equation (17.23) for the motion of a particle over the Earth, the equation of motion for the steady flow of the atmosphere or the ocean is

$$\rho\, \boldsymbol{a} + 2\rho\, \boldsymbol{\Omega} \times \boldsymbol{v} = -\rho g\, \boldsymbol{k} - \operatorname{grad} p. \tag{17.33}$$

Here, $\boldsymbol{v}(\boldsymbol{r})$ and $\boldsymbol{a}(\boldsymbol{r})$ are the velocity and acceleration of the *fluid particle* with position vector $\boldsymbol{r}$, ρ is the fluid density and p is the pressure field in the fluid. The term $-\operatorname{grad} p$ is called the **pressure gradient** and represents the force exerted on a fluid particle by the fluid around it. The acceleration $\boldsymbol{a}$ can be written in terms of $\boldsymbol{v}$. If we follow the progress in time of a fluid particle, it follows from the chain rule that

$$\begin{aligned} \boldsymbol{a} &= \frac{\partial \boldsymbol{v}}{\partial x} \times \frac{dx}{dt} + \frac{\partial \boldsymbol{v}}{\partial y} \times \frac{dy}{dt} + \frac{\partial \boldsymbol{v}}{\partial z} \times \frac{dz}{dt} \\ &= v_x \frac{\partial \boldsymbol{v}}{\partial x} + v_y \frac{\partial \boldsymbol{v}}{\partial y} + v_z \frac{\partial \boldsymbol{v}}{\partial z}. \end{aligned} \tag{17.34}$$

Let us now compare the sizes of the two acceleration terms on the left of equation (17.33). Suppose that the flow has characteristic length scale L and velocity scale V. Then the Coriolis acceleration $2\,\boldsymbol{\Omega} \times \boldsymbol{v}$ is of order ΩV and, from equation (17.34), the acceleration $\boldsymbol{a}$ is of order V^2/L. The dimensionless ratio of these two magnitudes, given by

$$Ro = \frac{V}{\Omega L}, \tag{17.35}$$

is called the **Rossby number**‡ of the flow. *When the Rossby number is small, the Coriolis acceleration dominates over the real acceleration.* When this condition is satisfied, the equation of motion (17.33) can be approximated by

$$2\rho\, \boldsymbol{\Omega} \times \boldsymbol{v} = -\rho g\, \boldsymbol{k} - \operatorname{grad} p. \tag{17.36}$$

* A **precession** is a slow rotation superimposed on a fast motion.

† After Jean Bernard Leon Foucault (1819–1868), who first exhibited it at the 1851 Paris Exhibition. This was the first time the Earth's rotation had been demonstrated by an entirely terrestrial method.

‡ After the Swedish-American meteorologist Carl-Gustav Rossby (1898–1957).

which is called a **geostrophic flow**. If the flow velocity $\boldsymbol{v}$ is everywhere horizontal, then, on taking the scalar product of this equation with $\boldsymbol{v}$, we obtain

$$\boldsymbol{v} \cdot \operatorname{grad} p = 0. \tag{17.37}$$

Hence, in a horizontal geostrophic flow, the fluid velocity is perpendicular to the horizontal pressure gradient. In other words, the **fluid flows along the isobars**.

Weather maps often show winds circulating around areas of low pressure. In a large low pressure system on the Earth, $V \approx 10\,\mathrm{m\,s^{-1}}$, $L \approx 1000$ km and $\Omega \approx 10^{-4}$ radians per second. The corresponding Rossby number is $Ro \approx 0.1$. Such a flow is therefore quite accurately geostrophic. We would therefore expect the wind to follow the isobars in the sense predicted by equation (17.36) (anti-clockwise around a 'low' in the northern hemisphere) and qualitatively this is what is observed. On the other hand, in the wind flow near the centre of a hurricane, $V \approx 50\,\mathrm{m\,s^{-1}}$ and $L \approx 50$ km so that the Rossby number $Ro \approx 10$. Hence, contrary to common belief, the Coriolis force has an insignificant* effect on the wind in this part of a hurricane. Even less significant is the effect of the Earth's rotation on the swirling motion of water leaving a bath or a toilet bowl. The direction of the water in such flows is strongly influenced by plumbing but not by Coriolis force!

17.4 MULTI-PARTICLE SYSTEM IN A NON-INERTIAL FRAME

Some of the principles of **multi-particle mechanics** can be extended to non-inertial reference frames.

Linear momentum principle

The **linear momentum principle** (see Chapter 10) can easily be extended to the case in which the system is observed from a **non-inertial frame**.

In the notation for systems of particles used in Part II, the transformed equation of motion for the particle P_i of mass m_i is

$$m_i\left[\boldsymbol{A} + \dot{\boldsymbol{\Omega}}\times\boldsymbol{r}_i + 2\,\boldsymbol{\Omega}\times\boldsymbol{v}_i + \boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\boldsymbol{r}_i) + \frac{d\boldsymbol{v}_i}{dt}\right] = \boldsymbol{F}_i + \sum_{j=1}^{N}\boldsymbol{G}_{ij}$$

$(1 \le i \le N)$. Here, all quantities are observed in the moving frame $\mathcal{F}'$, the dashes being understood throughout. On summing these equations over all the particles we obtain

$$\left(\sum_{i=1}^{N} m_i\right)\boldsymbol{A} + \dot{\boldsymbol{\Omega}}\times\left(\sum_{i=1}^{N} m_i\boldsymbol{r}_i\right) + 2\,\boldsymbol{\Omega}\times\left(\sum_{i=1}^{N} m_i\boldsymbol{v}_i\right) + \boldsymbol{\Omega}\times\left[\boldsymbol{\Omega}\times\left(\sum_{i=1}^{N} m_i\boldsymbol{r}_i\right)\right] + \sum_{i=1}^{N} m_i\frac{d\boldsymbol{v}_i}{dt}$$

$$= \sum_{i=1}^{N}\boldsymbol{F}_i + \sum_{i=1}^{N}\left(\sum_{j=1}^{N}\boldsymbol{G}_{ij}\right).$$

* And yet the direction of such winds is always that predicted by the geostrophic theory! However, the flow *was* geostrophic in the early stages of the hurricane's formation, and this presumably determines the direction of flow thereafter.

As in the standard linear momentum principle, the double sum of the $\{\boldsymbol{G}_{ij}\}$ is zero. In terms of centre of mass variables, this equation can be written

$$M\left[\boldsymbol{A}+\dot{\boldsymbol{\Omega}}\times\boldsymbol{R}+2\,\boldsymbol{\Omega}\times\boldsymbol{V}+\boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\boldsymbol{R})+\frac{d\boldsymbol{V}}{dt}\right]=\boldsymbol{F},\tag{17.38}$$

where $\boldsymbol{R}$ and $\boldsymbol{V}$ are the position and velocity of the centre of mass, M is the total mass, and $\boldsymbol{F}$ is the total external force. This is the centre of mass form of the **linear momentum principle** in a **non-inertial frame**.

We have therefore proved that:

Linear momentum principle in a non-inertial frame

In *any* reference frame, the centre of mass of a system moves as if it were a particle of mass the total mass, and all the external forces acted upon it.

Hence, even in a non-inertial frame, the motion of the centre of mass of a *system* can still be calculated by *particle* mechanics.

Energy principle

There is also a form of **energy conservation** that applies in frames with a *fixed origin* and *constant angular velocity*. The proof is as follows:

In this case the acceleration $\boldsymbol{A}$ and the angular acceleration $\dot{\boldsymbol{\Omega}}$ are zero and the transformed equation of motion for the particle P_i becomes

$$m_i\left[2\,\boldsymbol{\Omega}\times\boldsymbol{v}_i+\boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\boldsymbol{r}_i)+\frac{d\boldsymbol{v}_i}{dt}\right]=\boldsymbol{F}_i+\sum_{j=1}^{N}\boldsymbol{G}_{ij},$$

in the standard notation. If we now take the scalar poduct of this equation with $\boldsymbol{v}_i$, the Coriolis term vanishes and we obtain

$$m_i\left[\boldsymbol{v}_i\cdot\left[\boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\boldsymbol{r}_i)\right]+\boldsymbol{v}_i\cdot\frac{d\boldsymbol{v}_i}{dt}\right]=\boldsymbol{F}_i\cdot\boldsymbol{v}_i+\sum_{j=1}^{N}\boldsymbol{G}_{ij}\cdot\boldsymbol{v}_i\tag{17.39}$$

$(1\le i\le N)$. This is the same equation as for an inertial frame except for the centrifugal term. This term can be written

$$\begin{aligned}\boldsymbol{v}_i\cdot\left[\boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\boldsymbol{r}_i)\right]&=\boldsymbol{v}_i\cdot\left[(\boldsymbol{\Omega}\cdot\boldsymbol{r}_i)\boldsymbol{\Omega}-\Omega^2\boldsymbol{r}_i\right]\\&=(\boldsymbol{\Omega}\cdot\boldsymbol{v}_i)(\boldsymbol{\Omega}\cdot\boldsymbol{r}_i)-\Omega^2\boldsymbol{v}_i\cdot\boldsymbol{r}_i\\&=\frac{d}{dt}\left[\tfrac{1}{2}(\boldsymbol{\Omega}\cdot\boldsymbol{r}_i)^2-\tfrac{1}{2}\Omega^2\boldsymbol{r}_i\cdot\boldsymbol{r}_i\right]\\&=\frac{d}{dt}\left[-\tfrac{1}{2}\Omega^2p_i^2\right],\end{aligned}$$

where p_i is the perpendicular distance of P from the rotation axis. On substituting this formula into equation (17.39) and summing over all the particles we obtain

$$\frac{d}{dt}\left[\sum_{i=1}^{N}\tfrac{1}{2}m_i\boldsymbol{v}_i\cdot\boldsymbol{v}_i-\tfrac{1}{2}\Omega^2\sum_{i=1}^{N}m_ip_i^2\right]=\sum_{i=1}^{N}\boldsymbol{F}_i\cdot\boldsymbol{v}_i+\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{G}_{ij}\cdot\boldsymbol{v}_i,$$

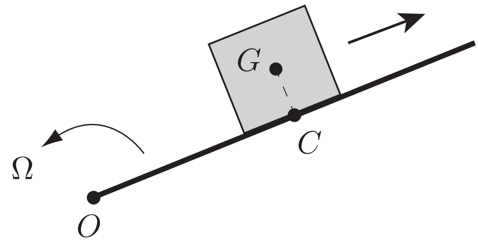


FIGURE 17.6 The straight rod rotates with constant angular speed Ω around O and the square lamina slides along the rod.

that is

$$\frac{d}{dt}\left(T - \tfrac{1}{2}I_{\mathcal{A}}\Omega^2\right) = \sum_{i=1}^{N} \boldsymbol{F}_i \cdot \boldsymbol{v}_i + \sum_{i=1}^{N}\sum_{j=1}^{N} \boldsymbol{G}_{ij} \cdot \boldsymbol{v}_i, \qquad (17.40)$$

where T is the apparent total kinetic energy and $I_{\mathcal{A}}$ is the moment of inertia of the system about the rotation axis $\mathcal{A}$. The right side of equation (17.40) is the apparent total rate of working of all the forces. If the apparent total work done can be represented by the potential energy V, we will say that the system is **apparently conservative**. In this case, we have **conservation of energy** in the form

$$T + V - \tfrac{1}{2}I_{\mathcal{A}}\Omega^2 = E,$$

where E is a constant. The term $-\frac{1}{2}I_{\mathcal{A}}\Omega^2$ is sometimes called the **centrifugal potential energy** and we will call this equation the **transformed energy equation**. Our result is summarised as follows:

Energy conservation in a uniformly rotating frame

If a system is apparently conservative when viewed from a rotating reference frame with a fixed origin and constant angular velocity, **energy conservation** holds in the modified form

$$T + V - \tfrac{1}{2}I_{\mathcal{A}}\Omega^2 = E, \qquad (17.41)$$

where T and V are the apparent kinetic and potential energies, Ω is the angular speed of the reference frame, and $I_{\mathcal{A}}$ is the moment of inertia of the system about the rotation axis.

For systems with one degree of freedom, this energy conservation principle is sufficient to determine the motion, as in the following example.

Example 17.5

A long straight smooth rod lies on a smooth horizontal table. One of its ends is fixed at point O on the table and the rod is made to rotate around O with constant angular speed Ω. A uniform square lamina of side $2a$ can slide freely on the table. One of its sides is in contact with the rod at all times and the lamina slides freely along the rod. Initially, the lamina is at rest (relative to the rod) with one corner at O. Find the subsequent displacement of the lamina as a function of the time.

Solution

We will view the motion of the lamina from a reference frame with origin O in which the rod is at rest. This frame has a *fixed* origin and *constant* angular velocity.

First we find the apparent kinetic energy. Let r be the distance OC shown in Figure 17.6. In the rotating frame, the motion of the lamina is rectilinear and its apparent **kinetic energy** is therefore given by $T = \frac{1}{2}M\dot{r}^2$.

Next we find the apparent potential energy. There are no specified external forces, but there is the constraint force exerted on the lamina by the rod, which, since the rod is smooth, acts *perpendicular* to the rod. In an inertial frame, this constraint force does work, but, in the frame rotating with the rod, the lamina simply moves *along* the stationary rod. The apparent rate of working of the constraint force is therefore zero. The internal forces within the lamina enforce rigidity and do no apparent work in total. The system is therefore **apparently conservative** with potential energy $V = 0$.

The transformed energy equation is therefore

$$\tfrac{1}{2}M\dot{r}^2 - \tfrac{1}{2}I_{\mathcal{A}}\Omega^2 = E,$$

where $I_{\mathcal{A}}$ is the moment of inertia of the lamina about the vertical axis through O, and Ω is the angular speed of the rod. By the 'parallel axes' theorem (see the Appendix at the end of the book), $I_{\mathcal{A}}$ is given by

$$I_{\mathcal{A}} = I_G + \left(a^2 + r^2\right),$$

so that the equation becomes

$$\tfrac{1}{2}M\dot{r}^2 - \tfrac{1}{2}\left(I_G + a^2 + r^2\right)\Omega^2 = E.$$

The **initial conditions** $r = a$ and $\dot{r} = 0$ when $t = 0$ give

$$E = -\tfrac{1}{2}\left(I_G + 2a^2\right)$$

so that the **transformed energy equation** for the lamina is finally given by

$$\dot{r}^2 = \Omega^2\left(r^2 - a^2\right).$$

This equation is sufficient to determine the motion. On taking square roots and integrating, we find that the **displacement** of the lamina at time t is

$$r = a\cosh\Omega t,$$

which is the same answer as for a particle sliding along a rotating wire. ■

Problems on Chapter 17

Answers and comments are at the end of the book.

Harder problems carry a star (∗).

Kinematics in rotating frames

17.1 Use the velocity and acceleration transformation formulae to derive the standard expressions for the velocity and acceleration of a particle in plane polar coordinates.

17.2 ***Addition of angular velocities*** Prove the 'addition of angular velocities' theorem, Theorem 17.13. [*Hint*. For a general particle of the rigid body, find $\boldsymbol{v}$ and $\boldsymbol{v}'$ in terms of $\boldsymbol{\omega}$ and $\boldsymbol{\omega}'$ respectively. Then relate $\boldsymbol{v}$ and $\boldsymbol{v}'$ by the velocity transformation formula.]

17.3 A circular cone with semi-angle α is fixed with its axis of symmetry vertical and its vertex O upwards. A second circular cone has semi-vertical angle $(\pi/2)-\alpha$ and has its vertex fixed at O. The second cone rolls on the first cone so that its axis of symmetry precesses around the upward vertical with angular speed λ. Find the angular speed of the rolling cone.

Dynamics in rotating frames

17.4 A particle P of mass m can slide along a smooth rigid straight wire. The wire has one of its points fixed at the origin O, and is made to rotate in a plane through O with constant angular speed Ω. Show that r, the distance of P from O, satisfies the equation

$$\ddot{r} - \Omega^2 r = 0.$$

Initially, P is at rest (relative to the wire) at a distance a from O. Find r as a function of t in the subsequent motion.

17.5 ***Larmor precession*** A particle of mass m and charge e moves in the force field $\boldsymbol{F}(\boldsymbol{r})$ and the uniform magnetic field $B\boldsymbol{k}$, where $\boldsymbol{k}$ is a constant unit vector. Its equation of motion is then

$$m\frac{d\boldsymbol{v}}{dt} = \left(\frac{eB}{c}\right)\boldsymbol{v}\times\boldsymbol{k} + \boldsymbol{F}(\boldsymbol{r})$$

in cgs Gaussian units. Show that the term $(eB/c)\boldsymbol{v}\times\boldsymbol{k}$ can be removed from the equation by viewing the motion from an appropriate rotating frame.

For the special case in which $\boldsymbol{F}(\boldsymbol{r}) = -m\omega_0^2\boldsymbol{r}$, show that circular motions with two different frequencies are possible.

17.6 A bullet is fired vertically upwards with speed u from a point on the Earth with co-latitude β. Show that it returns to the ground west of the firing point by a distance $4\Omega u^3 \sin\beta/3g^2$.

17.7 An artillery shell is fired from a point on the Earth with co-latitude β. The direction of firing is due **south**, the muzzle speed of the shell is u and the angle of elevation of the barrel is

α. show that the effect of the Earth's rotation is to deflect the shell to the east by a distance

$$\frac{4\Omega u^3}{3g^2}\sin^2\alpha\,(3\cos\alpha\cos\beta+\sin\alpha\sin\beta)\,.$$

17.8 An artillery shell is fired from a point on the Earth with co-latitude β. The direction of firing is due **east**, the muzzle speed of the shell is u and the angle of elevation of the barrel is α. Show that the effect of the Earth's rotation is to deflect the shell to the south by a distance

$$\frac{4\Omega u^3}{3g^2}\sin^2\alpha\cos\alpha\cos\beta.$$

∗ Show also that the easterly range is increased by

$$\frac{4\Omega u^3}{3g^2}\sin\alpha\sin\beta\left(3-4\sin^2\alpha\right).$$

[*Hint*. The second part requires a corrected value for the flight time.]

Energy conservation in rotating frames

17.9 Consider Problem 17.4 again. This time find the motion of the particle by using the transformed energy equation.

17.10 One end of a straight rod is fixed at a point O on a smooth horizontal table and the rod is made to rotate around O with constant angular speed Ω. A uniform circular disk of radius a lies flat on the table and can slide freely upon it. The disc remains in contact with the rod at all times and is constrained to *roll* along the rod. Initially, the disk is at rest (relative to the rod) with its point of contact at a distance a from O. Find the displacement of the disk as a function of the time.

17.11 A horizontal turntable is made to rotate about a fixed vertical axis with constant angular speed Ω. A *hollow* uniform circular cylinder of mass M and radius a can *roll* on the turntable. Initially the cylinder is at rest (relative to the turntable), with its centre of mass on the rotation axis, when it is slightly disturbed. Find the speed of the cylinder when it has rolled a distance x on the turntable.

∗ Find also an expression (in terms of x) for the force that the turntable exerts on the cylinder.

Hydrostatics in rotating frames

When a fluid is at rest in an **inertial frame**, the equation of hydrostatics is

$$\mathbf{0}=\boldsymbol{F}-\operatorname{grad}p,$$

where p is the pressure field in the fluid, and $\boldsymbol{F}$ is the body force (force per unit volume) acting on the fluid. This equation means that the body force is balanced by the pressure gradient. In the case of uniform gravity, $\boldsymbol{F}=-\rho g\boldsymbol{k}$ and, if ρ is constant, the equation integrates to give the well known 'hydrostatic pressure formula' $p=p_0-\rho gz$, where Oz points vertically upwards.

If a fluid is at rest in a **rotating frame** with a fixed origin and constant angular velocity $\boldsymbol{\Omega}$, then the equation of 'hydrostatics' in this frame is

$$\rho\,\boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\boldsymbol{r})=\boldsymbol{F}-\operatorname{grad}p,$$

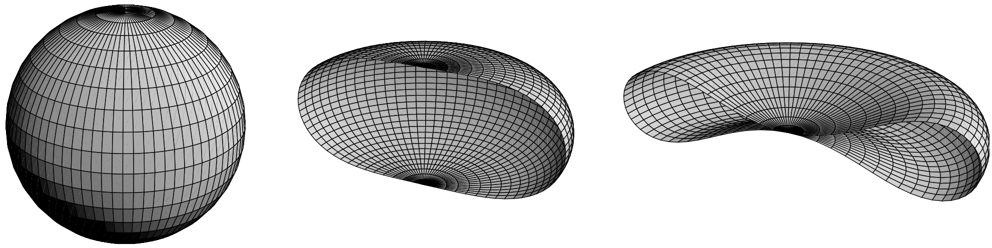


FIGURE 17.7 A mass of liquid in equilibrium (left) and rotating like a rigid body (centre and right, cut in half to show the cross sections). The right figure shows the mass rotating with the critical angular speed.

where ρ is the fluid density. (This equation means that the body force, the pressure gradient and the centrifugal 'force' must balance.) This is all you need to know to answer the following questions.

17.12 ***Newton's bucket*** A bucket half full of water is made to rotate with angular speed Ω about its axis of symmetry, which is vertical. Find, to within a constant, the pressure field in the fluid. By considering the isobars (surfaces of constant pressure) of this pressure field, find the shape of the free surface of the water.

What would the shape of the free surface be if the bucket were replaced by a cubical box?

17.13 A sealed circular can of radius a is three-quarters full of water of density ρ, the remainder being air at pressure p_0. The can is taken into gravity free space and then rotated about its axis of symmetry with constant angular speed Ω. Where will the water be when it comes to rest relative to the can? Find the water pressure at the wall of the can.

Computer assisted problem

The following problem is suitable for a **supervised project**.

17.14** If a mass of liquid in gravity free space is in equilibrium, then it has the form of a sphere, stabilised by its own surface tension (see Figure 17.7 (left)). Suppose that the mass now rotates like a rigid body with constant angular velocity. What will its equilibrium shape be now? Show that, as the angular speed increases, the mass becomes more oblate until its top and bottom surfaces become quite flat (Figure 17.7 (centre)). Show further that, at higher angular speeds, the top and bottom surfaces approach each other until they meet (Figure 17.7 (right)) and the mass breaks up. Find this critical angular speed for a sphere of mercury of radius 5 cm.

Chapter Eighteen

Tensor algebra and the inertia tensor

KEY FEATURES

The key features of this chapter are the **transformation formulae** for the components of tensors and **tensor algebra**; the **inertia tensor** and the calculation of the angular momentum and kinetic energy of a rigid body; and the **principal axes** and **principal moments of inertia** of a rigid body.

We have previously regarded a vector as a quantity that has magnitude and direction. Picturing a vector as a line segment with an arrow on it has helped us to understand essentially difficult concepts, such as the acceleration of a particle and the angular momentum of a rigid body. Neither of these quantities has any direct connection with line segments, but the picture is a valuable aid nonetheless. Useful though this notion of vectors is, it is not capable of generalisation and this is the main reason why we will now look at vectors from a different perspective. In reality, what we actually observe are the three *components* of a vector, and the values of these components will depend on the coordinate system in which they are measured. However, a general triple of real numbers $\{v_1, v_2, v_3\}$, defined in each coordinate system, does not neccessarily constitute a vector. The reason is that the components of a vector in different coordinate systems are related to each other (in a way that we shall determine) and, unless the quantities $\{v_1, v_2, v_3\}$ satisfy this **transformation formula**, they are *not* the components of a vector. This leads us to a new definition of a vector as *a quantity that has three components in each coordinate system, whose values in different coordinate systems are related by the vector transformation formula*.

Defined in the above way, vectors can be regarded as part of a hierarchy of entities called **tensors** with a tensor of order n having 3^n components. A tensor of order zero, which has one component, is a **scalar**, and a tensor of order one, which has three components, is a **vector**. Tensors of higher order, which are entirely new objects, are defined to be quantities with 3^n components that obey the appropriate transformation formula. In this chapter we give a gentle introduction to the principles of **tensor algebra**, which should be useful, not only for its present application to dynamics, but also in other areas such as relativity. Any account of tensor algebra is mathematics, not mechanics, and some readers may find this indigestible. However, there are very few basic principles involved and, once these are mastered, the subject seems (and actually is) rather easy.

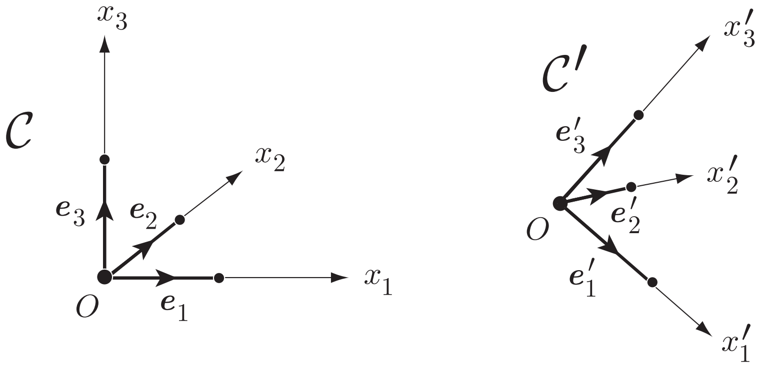


FIGURE 18.1 The Cartesian coordinate systems $\mathcal{C} \equiv Ox_1x_2x_3$ and $\mathcal{C}' \equiv Ox'_1x'_2x'_3$ have the associated unit vectors $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ and $\{\boldsymbol{e}'_1, \boldsymbol{e}'_2, \boldsymbol{e}'_3\}$ respectively. The two systems actually have a common origin but are drawn separately for clarity.

We then introduce the **inertia tensor**, a second order tensor that is used to calculate the angular momentum $\boldsymbol{L}$ and kinetic energy T of a generally rotating rigid body. It transpires that, by taking a special choice of axes, known as **principal axes**, the formulae for $\boldsymbol{L}$ and T are greatly simplified and reduce to non-tensorial forms. This makes it possible to solve most problems in rigid body dynamics without any knowledge of tensors at all.* The few results that one *really* needs in rigid body dynamics are given at the end of Section 18.6.

18.1 ORTHOGONAL TRANSFORMATIONS

Let $\mathcal{C} \equiv Ox_1x_2x_3$ and $\mathcal{C}' \equiv Ox'_1x'_2x'_3$ be two Cartesian coordinate systems having the common origin O and associated unit vectors $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ and $\{\boldsymbol{e}'_1, \boldsymbol{e}'_2, \boldsymbol{e}'_3\}$ respectively (see Figure 18.1). Then any vector $\boldsymbol{v}$ (understood in the elementary sense of Chapter 1) can be expanded in the form

$$\boldsymbol{v} = v_1\,\boldsymbol{e}_1 + v_2\,\boldsymbol{e}_2 + v_3\,\boldsymbol{e}_3, \tag{18.1}$$

where $\{v_1, v_2, v_3\}$ are the components of $\boldsymbol{v}$ in the coordinate system $\mathcal{C}$. Similarly $\boldsymbol{v}$ can be expanded in the form

$$\boldsymbol{v} = v'_1\,\boldsymbol{e}'_1 + v'_2\,\boldsymbol{e}'_2 + v'_3\,\boldsymbol{e}'_3, \tag{18.2}$$

where $\{v'_1, v'_2, v'_3\}$ are the components of $\boldsymbol{v}$ in the coordinate system $\mathcal{C}'$. These two sets of components are related to each other. On taking the scalar product of the equality

$$v'_1\,\boldsymbol{e}'_1 + v'_2\,\boldsymbol{e}'_2 + v'_3\,\boldsymbol{e}'_3 = v_1\,\boldsymbol{e}_1 + v_2\,\boldsymbol{e}_2 + v_3\,\boldsymbol{e}_3 \tag{18.3}$$

* Don't spread this information around!

successively with $\boldsymbol{e}_1'$, $\boldsymbol{e}_2'$ and $\boldsymbol{e}_3'$ we find that

$$
\begin{aligned}
v_1' &= \left(\boldsymbol{e}_1' \cdot \boldsymbol{e}_1\right) v_1 + \left(\boldsymbol{e}_1' \cdot \boldsymbol{e}_2\right) v_2 + \left(\boldsymbol{e}_1' \cdot \boldsymbol{e}_3\right) v_3,\\
v_2' &= \left(\boldsymbol{e}_2' \cdot \boldsymbol{e}_1\right) v_1 + \left(\boldsymbol{e}_2' \cdot \boldsymbol{e}_2\right) v_2 + \left(\boldsymbol{e}_2' \cdot \boldsymbol{e}_3\right) v_3,\\
v_3' &= \left(\boldsymbol{e}_3' \cdot \boldsymbol{e}_1\right) v_1 + \left(\boldsymbol{e}_3' \cdot \boldsymbol{e}_2\right) v_2 + \left(\boldsymbol{e}_3' \cdot \boldsymbol{e}_3\right) v_3,
\end{aligned}
$$

which can be written in the matrix form

$$
\boxed{\mathsf{v}' = \mathsf{A} \cdot \mathsf{v}} \tag{18.4}
$$

where

$$
\mathsf{v}' = \begin{pmatrix} v_1' \\ v_2' \\ v_3' \end{pmatrix}, \qquad \mathsf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \qquad \mathsf{A} = \begin{pmatrix} (\boldsymbol{e}_1' \cdot \boldsymbol{e}_1) & (\boldsymbol{e}_1' \cdot \boldsymbol{e}_2) & (\boldsymbol{e}_1' \cdot \boldsymbol{e}_3) \\ (\boldsymbol{e}_2' \cdot \boldsymbol{e}_1) & (\boldsymbol{e}_2' \cdot \boldsymbol{e}_2) & (\boldsymbol{e}_2' \cdot \boldsymbol{e}_3) \\ (\boldsymbol{e}_3' \cdot \boldsymbol{e}_1) & (\boldsymbol{e}_3' \cdot \boldsymbol{e}_2) & (\boldsymbol{e}_3' \cdot \boldsymbol{e}_3) \end{pmatrix}, \tag{18.5}
$$

and the '$\cdot$' means the *matrix product*. This is the required relationship between the components of $\boldsymbol{v}$ in the two coordinate systems. The elements of the **transformation matrix A** depend only on the *orientation* of the system $\mathcal{C}'$ relative to the system $\mathcal{C}$.

The **inverse relationship** to (18.4) can be found by taking the scalar product of equation (18.3) successively with $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$. The result is that

$$
\begin{aligned}
v_1' &= \left(\boldsymbol{e}_1 \cdot \boldsymbol{e}_1'\right) v_1 + \left(\boldsymbol{e}_1 \cdot \boldsymbol{e}_2'\right) v_2 + \left(\boldsymbol{e}_1 \cdot \boldsymbol{e}_3'\right) v_3,\\
v_2' &= \left(\boldsymbol{e}_2 \cdot \boldsymbol{e}_1'\right) v_1 + \left(\boldsymbol{e}_2 \cdot \boldsymbol{e}_2'\right) v_2 + \left(\boldsymbol{e}_2 \cdot \boldsymbol{e}_3'\right) v_3,\\
v_3' &= \left(\boldsymbol{e}_3 \cdot \boldsymbol{e}_1'\right) v_1 + \left(\boldsymbol{e}_3 \cdot \boldsymbol{e}_2'\right) v_2 + \left(\boldsymbol{e}_3 \cdot \boldsymbol{e}_3'\right) v_3,
\end{aligned}
$$

which can be written in the matrix form

$$
\boxed{\mathsf{v} = \mathsf{A}^T \cdot \mathsf{v}'} \tag{18.6}
$$

where A^T is the **transpose** of A. If we now substitute equation (18.4) into equation (18.6) we obtain

$$
\mathsf{v} = \mathsf{A}^T \cdot (\mathsf{A} \cdot \mathsf{v}) = \left(\mathsf{A}^T \cdot \mathsf{A}\right) \cdot \mathsf{v}.
$$

Since this formula holds for every choice of v, it follows that $\mathsf{A}^T \cdot \mathsf{A} = \mathbf{1}$, where $\mathbf{1}$ is the identity matrix. This means that the transformation matrix A has the special property that $\mathsf{A}^T = \mathsf{A}^{-1}$. Such matrices are called **orthogonal**.

Definition 18.1 ***Orthogonal matrix*** *A matrix having the property that its transpose is also its inverse, that is,*

$$
\mathsf{A}^T \cdot \mathsf{A} = \mathsf{A} \cdot \mathsf{A}^T = \mathbf{1}, \tag{18.7}
$$

*is called an **orthogonal** matrix.*

Thus the transformation matrix **A**, defined by (18.5), is an orthogonal matrix. For this reason, the transformation formula (18.4) is called an **orthogonal transformation**.

Example 18.1 ***Prove a matrix is orthogonal***

Prove that the matrix

$$\mathsf{B} = \frac{1}{9}\begin{pmatrix} 4 & 7 & -4 \\ 1 & 4 & 8 \\ 8 & -4 & 1 \end{pmatrix}$$

is orthogonal.

Solution

$$\mathsf{B}^T \cdot \mathsf{B} = \frac{1}{81}\begin{pmatrix} 4 & 1 & 8 \\ 7 & 4 & -4 \\ -4 & 8 & 1 \end{pmatrix}\begin{pmatrix} 4 & 7 & -4 \\ 1 & 4 & 8 \\ 8 & -4 & 1 \end{pmatrix} = \frac{1}{81}\begin{pmatrix} 81 & 0 & 0 \\ 0 & 81 & 0 \\ 0 & 0 & 81 \end{pmatrix} = \mathbf{1}.$$

There is no need to check that $\mathsf{B} \cdot \mathsf{B}^T = \mathbf{1}$ as this now follows automatically. ■

18.2 ROTATED AND REFLECTED COORDINATE SYSTEMS

In this section we work out the transformation matrices that correspond to specific **rotations** and/or **reflections** of the coordinate system $\mathcal{C}$. Important though this is, it is not needed for the general understanding of tensors.

Suppose now that the coordinate system $\mathcal{C}'$ is *defined* by a specified rotation of $\mathcal{C}$ about an axis through O. This notion needs a little thought. It should be remembered that the coordinate system $\mathcal{C}$ does not actually move; the ‘rotation’ of $\mathcal{C}$ is hypothetical and is merely a way of describing the orientation of $\mathcal{C}'$ relative to $\mathcal{C}$. What we are saying is that the orientation of $\mathcal{C}'$ is the same as $\mathcal{C}$ *would* have *if* it were rotated in the prescribed manner.

Consider first the **special case** in which $\mathcal{C}'$ is obtained by rotating $\mathcal{C}$ through an angle ψ about the axis Ox_3, as shown in Figure 18.2 (Left). In this case,

$$\begin{array}{lll} \boldsymbol{e}'_1 \cdot \boldsymbol{e}_1 = \cos\psi & \boldsymbol{e}'_1 \cdot \boldsymbol{e}_2 = \sin\psi & \boldsymbol{e}'_1 \cdot \boldsymbol{e}_3 = 0 \\ \boldsymbol{e}'_2 \cdot \boldsymbol{e}_1 = -\sin\psi & \boldsymbol{e}'_2 \cdot \boldsymbol{e}_2 = \cos\psi & \boldsymbol{e}'_2 \cdot \boldsymbol{e}_3 = 0 \\ \boldsymbol{e}'_3 \cdot \boldsymbol{e}_1 = 0 & \boldsymbol{e}'_3 \cdot \boldsymbol{e}_2 = 0 & \boldsymbol{e}'_3 \cdot \boldsymbol{e}_3 = 1 \end{array}$$

and the **transformation matrix** between $\mathcal{C}$ and $\mathcal{C}'$ is therefore

$$\mathsf{A} = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{18.8}$$

Since this example is a special case of the preceding theory, the matrix **A** in (18.8) must be orthogonal. [You should check this directly.] A rotation of $\mathcal{C}$ about the Ox_1 or Ox_2 axis is treated in the same way.

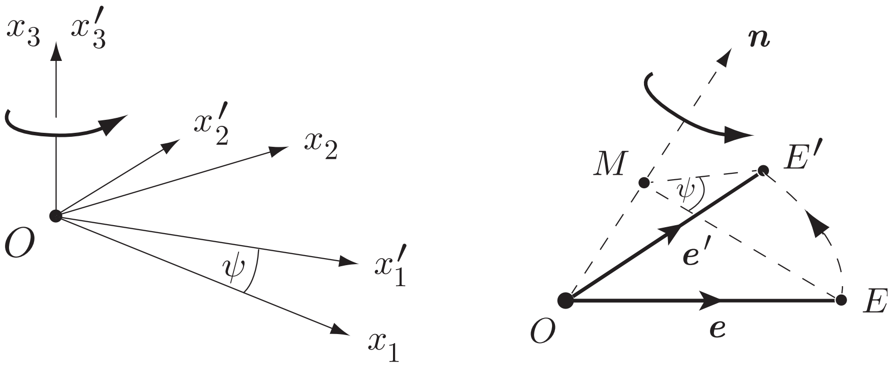


FIGURE 18.2 **Left**: The coordinate system $\mathcal{C}' \equiv Ox_1'x_2'x_3'$ is obtained by rotating the coordinate system $\mathcal{C} \equiv Ox_1x_2x_3$ through an angle ψ about the axis Ox_3. **Right**: The unit vector $\boldsymbol{e}'$ is obtained by rotating the unit vector $\boldsymbol{e}$ through an angle ψ about the general axis $\{O, \boldsymbol{n}\}$.

Example 18.2 *Rotation of $\mathcal{C}$ about a coordinate axis*

The coordinate system $\mathcal{C}'$ is obtained by rotating the coordinate system $\mathcal{C}$ through an angle of 30° about the axis Ox_1. Find the transformation matrix **A** between $\mathcal{C}$ and $\mathcal{C}'$.

A vector $\boldsymbol{v}$ has components $\{3, -1, 2\}$ in $\mathcal{C}$. What are its components in $\mathcal{C}'$?

Solution

The formula corresponding to (18.8) when $\mathcal{C}$ is rotated through an angle ψ about the axis Ox_1 is

$$\mathsf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\psi & \sin\psi \\ 0 & -\sin\psi & \cos\psi \end{pmatrix}.$$

Hence, when $\psi = 30°$, the **transformation matrix** is

$$\mathsf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sqrt{3}/2 & 1/2 \\ 0 & -1/2 & \sqrt{3}/2 \end{pmatrix}.$$

The components of the vector $\boldsymbol{v}$ in $\mathcal{C}'$ are then given by the elements of the column vector v', where

$$\mathsf{v}' = \mathsf{A} \cdot \begin{pmatrix} 3 \\ -4 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sqrt{3}/2 & 1/2 \\ 0 & -1/2 & \sqrt{3}/2 \end{pmatrix} \begin{pmatrix} 3 \\ -4 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 - 2\sqrt{3} \\ 2 + \sqrt{3} \end{pmatrix}.$$

The **components** of the vector $\boldsymbol{v}$ in $\mathcal{C}'$ are therefore $\{3,\ 1 - 2\sqrt{3},\ 2 + \sqrt{3}\}$. ■

In the **general case**, the rotation of $\mathcal{C}$ can take place about an axis $\{O, \boldsymbol{n}\}$, where $\boldsymbol{n}$ is *any* unit vector. The argument is messier but we can still find an explicit formula for **A**.

Figure 18.2 (Right) shows how a typical unit vector $\boldsymbol{e}'$ of the coordinate system $\mathcal{C}'$ is obtained from the corresponding vector $\boldsymbol{e}$ belonging to $\mathcal{C}$. Let $\boldsymbol{e}$ be resolved into parts parallel and perpendicular to the axis $\{O, \boldsymbol{n}\}$ in the form

$$\boldsymbol{e} = \boldsymbol{e}^{\parallel} + \boldsymbol{e}^{\perp},$$

where

$$\boldsymbol{e}^{\parallel} = \overrightarrow{OM} = (\boldsymbol{e} \cdot \boldsymbol{n})\,\boldsymbol{n}, \qquad \boldsymbol{e}^{\perp} = \overrightarrow{ME} = \boldsymbol{e} - (\boldsymbol{e} \cdot \boldsymbol{n})\,\boldsymbol{n}.$$

Then, if $\boldsymbol{e}'$ is similarly resolved as

$$\boldsymbol{e}' = \boldsymbol{e}'^{\parallel} + \boldsymbol{e}'^{\perp},$$

it follows that $\boldsymbol{e}'^{\parallel} = \boldsymbol{e}^{\parallel}$ and

$$\boldsymbol{e}'^{\perp} = \cos\psi\, \boldsymbol{e}^{\perp} + \sin\psi \left(\boldsymbol{n} \times \boldsymbol{e}^{\perp}\right) = \cos\psi\, (\boldsymbol{e} - (\boldsymbol{e} \cdot \boldsymbol{n})\,\boldsymbol{n}) + \sin\psi\, (\boldsymbol{n} \times \boldsymbol{e}).$$

Hence $\boldsymbol{e}'$ is related to $\boldsymbol{e}$ by the formula

$$\boldsymbol{e}' = \cos\psi\, \boldsymbol{e} + (1 - \cos\psi)\, (\boldsymbol{e} \cdot \boldsymbol{n})\,\boldsymbol{n} + \sin\psi\, (\boldsymbol{n} \times \boldsymbol{e}). \tag{18.9}$$

If we now take $\boldsymbol{e} = \boldsymbol{e}_1$, $\boldsymbol{e}' = \boldsymbol{e}_1'$ and take the scalar product of equation (18.9) successively with $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$, we obtain

$$\begin{aligned}
\boldsymbol{e}_1' \cdot \boldsymbol{e}_1 &= \cos\psi (\boldsymbol{e}_1 \cdot \boldsymbol{e}_1) + (1 - \cos\psi)(\boldsymbol{n} \cdot \boldsymbol{e}_1)(\boldsymbol{n} \cdot \boldsymbol{e}_1) + \sin\psi (\boldsymbol{n} \times \boldsymbol{e}_1) \cdot \boldsymbol{e}_1 \\
&= \cos\psi + (1 - \cos\psi) n_1^2 + 0, \\
\boldsymbol{e}_1' \cdot \boldsymbol{e}_2 &= \cos\psi (\boldsymbol{e}_1 \cdot \boldsymbol{e}_2) + (1 - \cos\psi)(\boldsymbol{n} \cdot \boldsymbol{e}_1)(\boldsymbol{n} \cdot \boldsymbol{e}_2) + \sin\psi (\boldsymbol{n} \times \boldsymbol{e}_1) \cdot \boldsymbol{e}_2 \\
&= 0 + (1 - \cos\psi) n_1 n_2 + n_3 \sin\psi, \\
\boldsymbol{e}_1' \cdot \boldsymbol{e}_3 &= \cos\psi (\boldsymbol{e}_1 \cdot \boldsymbol{e}_3) + (1 - \cos\psi)(\boldsymbol{n} \cdot \boldsymbol{e}_1)(\boldsymbol{n} \cdot \boldsymbol{e}_3) + \sin\psi (\boldsymbol{n} \times \boldsymbol{e}_1) \cdot \boldsymbol{e}_3 \\
&= 0 + (1 - \cos\psi) n_1 n_3 - n_2 \sin\psi,
\end{aligned}$$

where n_1, n_2, n_3 are the components of the vector $\boldsymbol{n}$ in the coordinate system $\mathcal{C}$. This gives the elements of the first row of the transformation matrix $\mathbf{A}$, and the elements of the second and third rows can be found in a similar manner. The full transformation matrix $\mathbf{A}$ is given by

$$\begin{pmatrix}
\cos\psi + (1 - \cos\psi) n_1^2 & (1 - \cos\psi) n_1 n_2 + n_3 \sin\psi & (1 - \cos\psi) n_1 n_3 - n_2 \sin\psi \\
(1 - \cos\psi) n_2 n_1 - n_3 \sin\psi & \cos\psi + (1 - \cos\psi) n_2^2 & (1 - \cos\psi) n_2 n_3 + n_1 \sin\psi \\
(1 - \cos\psi) n_3 n_1 + n_2 \sin\psi & (1 - \cos\psi) n_3 n_2 - n_1 \sin\psi & \cos\psi + (1 - \cos\psi) n_3^2
\end{pmatrix} \tag{18.10}$$

This is the transformation matrix $\mathbf{A}$ when $\mathcal{C}'$ is obtained by rotating $\mathcal{C}$ through an angle ψ about the axis $\{O, \boldsymbol{n}\}$; n_1, n_2, n_3 are components of the unit vector $\boldsymbol{n}$ in the coordinate system $\mathcal{C}$.

Example 18.3 *Rotation of $\mathcal{C}$ about a general axis*

The coordinate system $\mathcal{C}'$ is obtained by rotating the coordinate system $\mathcal{C}$ through an angle of 60° about the axis $\overrightarrow{OP}$, where P is the point with coordinates $(1, 1, -1)$ in $\mathcal{C}$. Find the transformation matrix $\mathbf{A}$ between $\mathcal{C}$ and $\mathcal{C}'$.

A vector $\boldsymbol{v}$ has components $\{3, -6, 9\}$ in $\mathcal{C}'$. What are its components in $\mathcal{C}$?

Solution

The unit vector $\boldsymbol{n}$ in the direction $\overrightarrow{OP}$ has components $\{1/\sqrt{3}, 1/\sqrt{3}, -1/\sqrt{3}\}$ in $\mathcal{C}$, and $\psi = 60°$. It follows from the formula (18.10) that the tranformation matrix from

$\mathcal{C}$ to $\mathcal{C}'$ is

$$\mathsf{A} = \frac{1}{3}\begin{pmatrix} 2 & -1 & -2 \\ 2 & 2 & 1 \\ 1 & -2 & 2 \end{pmatrix}.$$

The components of the vector $\boldsymbol{v}$ in $\mathcal{C}$ are then given by the elements of the column vector v', where

$$\mathsf{v}' = \mathsf{A}^T \cdot \begin{pmatrix} 3 \\ -6 \\ 9 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 2 & 2 & 1 \\ -1 & 2 & -2 \\ -2 & 1 & 2 \end{pmatrix}\begin{pmatrix} 3 \\ -6 \\ 9 \end{pmatrix} = \begin{pmatrix} 1 \\ -11 \\ 2 \end{pmatrix}.$$

The components of the vector $\boldsymbol{v}$ in $\mathcal{C}$ are therefore $\{1, -11, 2\}$. ■

Reflections

The coordinate system $\mathcal{C}'$ may also be defined by **reflecting** $\mathcal{C}$ in a plane through O. For example, if $\mathcal{C}$ is reflected in the coordinate plane Ox_1x_2, the transformation matrix is easily found to be

$$\mathsf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

In the general case, if $\mathcal{C}$ is reflected in the plane through O with unit normal vector $\boldsymbol{n}$, then the transformation matrix is

$$\mathsf{A} = \begin{pmatrix} 1-2n_1^2 & -2n_1n_2 & -2n_1n_3 \\ -2n_2n_1 & 1-2n_2^2 & -2n_2n_3 \\ -2n_3n_1 & -2n_3n_2 & 1-2n_3^2 \end{pmatrix}, \qquad (18.11)$$

where n_1, n_2, n_3 are the components of the vector $\boldsymbol{n}$ in $\mathcal{C}$. [The method of proof is similar to that for the general rotation.]

Example 18.4 *Reflection of axes*

The coordinate system $\mathcal{C}'$ is obtained by reflecting the coordinate system $\mathcal{C}$ in the plane $x_3 = 2x_1 + 2x_2$. Find the transformation matrix A between $\mathcal{C}$ and $\mathcal{C}'$.

Find also the transformation matrix when $\mathcal{C}'$ is obtained from $\mathcal{C}$ by performing the above reflection followed by a rotation of 90° about the *new* x_3-axis.

Solution

The equation $x_3 = 2x_1 + 2x_2$ can be written in the form

$$(x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + x_3\boldsymbol{e}_3) \cdot (2\boldsymbol{e}_1 + 2\boldsymbol{e}_2 - \boldsymbol{e}_3)$$

and, by comparison with the standard vector equation $\boldsymbol{r} \cdot \boldsymbol{n} = p$ of a plane, we see that the unit vector

$$\boldsymbol{n} = \frac{2}{3}\boldsymbol{e}_1 + \frac{2}{3}\boldsymbol{e}_2 - \frac{1}{3}\boldsymbol{e}_3$$

is normal to the reflection plane. Hence $n_1 = 2/3$, $n_2 = 2/3$ and $n_3 = -1/3$. It follows from the formula (18.11) that the transformation matrix between $\mathcal{C}$ and $\mathcal{C}'$ is

$$\mathbf{A} = \frac{1}{9}\begin{pmatrix} 1 & -8 & 4 \\ -8 & 1 & 4 \\ 4 & 4 & 7 \end{pmatrix}.$$

If this transformation is followed by a rotation whose transformation matrix is

$$\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

then the overall transformation has matrix

$$\frac{1}{9}\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -8 & 4 \\ -8 & 1 & 4 \\ 4 & 4 & 7 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} -8 & 1 & 4 \\ -1 & 8 & -4 \\ 4 & 4 & 7 \end{pmatrix}. \blacksquare$$

It can be proved that the most **general orthogonal transformation** corresponds either to (i) a rotation of coordinates, or (ii) a reflection followed by a rotation. Which of these classes a particular orthogonal transformation belongs to can be decided by examining the *sign* of the determinant det **A**.

If we take determinants of equation (18.7) we obtain

$$\det \mathbf{1} = \det\left(\mathbf{A}^T \cdot \mathbf{A}\right) = \det \mathbf{A}^T \times \det \mathbf{A} = \det \mathbf{A} \times \det \mathbf{A} = (\det \mathbf{A})^2,$$

from which it follows that det **A** can only take the values ± 1. Surprisingly, we can work out the determinant of the general **A** defined by equation (18.5) quite easily. The first row of **A** contains the three components of the vector $\boldsymbol{e}'_1$ in the coordinate system $\mathcal{C}$. Likewise, the second and third rows contain the components of the vectors $\boldsymbol{e}'_2$ and $\boldsymbol{e}'_3$. It follows from the vector formula (1.7) that, provided $\mathcal{C}$ is a **right-handed** coordinate system,

$$\det \mathbf{A} = \left(\boldsymbol{e}'_1 \times \boldsymbol{e}'_2\right) \cdot \boldsymbol{e}'_3.$$

We can now see how the different signs for det **A** can occur. If $\mathcal{C}'$ is also a **right-handed** coordinate system, then the triple scalar product $\left(\boldsymbol{e}'_1 \times \boldsymbol{e}'_2\right) \cdot \boldsymbol{e}'_3 = +1$, whereas if $\mathcal{C}'$ is a **left-handed** system then $\left(\boldsymbol{e}'_1 \times \boldsymbol{e}'_2\right) \cdot \boldsymbol{e}'_3 = -1$. If two coordinate systems have the same handedness, then one can be made coincident with the other by a suitable **rotation** about an axis through O. If they have opposite handedness then a **reflection** in a plane through O is also needed. For this reason, orthogonal matrices with determinant $+1$ are sometimes called **rotation matrices**.

18.3 SCALARS, VECTORS AND TENSORS

Our view of a vector as a quantity that has magnitude and direction has served us well. By picturing a vector as a line segment with an arrow on it, we have been able to understand essentially difficult concepts, such as acceleration and angular momentum, far more easily. Neither of these quantities has any direct connection with line segments, but the picture is a valuable aid nonetheless. However, this notion of vectors cannot be generalised to tensors. This is the main reason we will now look at vectors (and even scalars) in a different light.

Scalars

We begin by giving the true definition of a scalar.

Definition 18.2 ***Scalar*** *Let ϕ be a real number defined in each coordinate system.** *Then ϕ is said to be a* **scalar** *if it has the same value in every coordinate system. That is, for any pair of coordinate systems $\mathcal{C}$ and $\mathcal{C}'$,*

Definition of a scalar

$$\phi' = \phi. \tag{18.12}$$

This can be alternatively expressed by saying that a scalar is ***invariant*** *under change of coordinate system.*

It is this invariance under change of coordinate system that is the essential feature of a scalar. Thus the **mass** of a particle and the **length** of a line are scalars. However, the sum of the coordinates of a given point of space (which *is* a single real quantity defined in each coordinate system) is not a scalar since it is *not* invariant under change of coordinate system.

Vectors

We turn now to the definition of a vector in terms of its components. The individual components of a vector are certainly not invariants. They transform according to the transformation formula (18.4). It is this transformation formula that becomes our **new definition** of a vector.

Definition 18.3 ***Vector*** *Let $\{v_1, v_2, v_3\}$ be a set of three real numbers defined in each coordinate system. Then $\{v_1, v_2, v_3\}$ are said to be the components of a* **vector** *if their values in any pair of coordinate systems $\mathcal{C}$ and $\mathcal{C}'$ are related by the transformation formula*

Definition of a vector (matrix form)

$$\mathbf{v}' = \mathbf{A}\cdot\mathbf{v} \tag{18.13}$$

where $\mathbf{A}$ *is the transformation matrix between $\mathcal{C}$ and $\mathcal{C}'$.*

* These are rectangular, Cartesian coordinate systems with common origin O.

This definition of a vector is clearly consistent with our previous point of view. Note that a set of three real numbers defined in each coordinate system does *not* in general constitute a vector; *the components must be related to each other by the transformation formula* (18.13). For example, the Cartesian coordinates (b_1, b_2, b_3) of a particle B are a set of three real numbers defined in each coordinate system and these numbers *are* related by the transformation formula (18.13). Thus $\{b_1, b_2, b_3\}$ is a **vector**. Suppose now that B is a moving particle with time dependent coordinates $(b_1(t), b_2(t), b_3(t))$. Then, at each time t, $\{\dot{b}_1(t), \dot{b}_2(t), \dot{b}_3(t)\}$ is also a set of three of real numbers defined in each coordinate system. The transformation formula (18.13) is again satisfied since the matrix $\mathbf{A}$ does not depend upon the time. Hence, at each time t, $\{\dot{b}_1(t), \dot{b}_2(t), \dot{b}_3(t)\}$ is a **vector**. This is, of course, the **velocity vector** of the particle B at time t, and a second differentiation yields the **acceleration vector** of B. In contrast, $\{2b_1, b_2, b_3\}$ is a set of three real numbers defined in each coordinate system, but the transformation formula is not satisfied. Therefore $\{2b_1, b_2, b_3\}$ is not a vector.

Tensors

In order to generalise our definition of a vector, it is necessary to write the vector transformation formula (18.13) in the following suffix form:

Definition of a vector (suffix form)

$$v'_i = \sum_{j=1}^{3} a_{ij}\, v_j \qquad (18.14)$$

$(1 \leq i \leq 3)$, where a_{pq} is the element in the p-th row and q-th column of the matrix $\mathbf{A}$. The matrix product in equation (18.13) is equivalent to the summation in equation (18.14).

Defined in the above way, vectors can be regarded as part of a hierarchy of entities called **tensors** with a tensor of order n having 3^n components. A tensor of order one, which has three components, is a vector; a tensor of order zero, which has one component, is a scalar. The general definition of tensors looks extremely daunting when first encountered. Below we give the definitions of second and third order tensors, from which the general case can be inferred.

Definition 18.4 ***Second order tensor*** *Let* $\{t_{ij}\}$ $(1 \leq i, j \leq 3)$ *be a set of* **nine** *real numbers defined in each coordinate system. Then the* $\{t_{ij}\}$ *are said to be the components of a* ***second order tensor*** *if their values in any pair of coordinate systems* $\mathcal{C}$ *and* $\mathcal{C}'$ *are related by the* ***transformation formula***

Definition of a second order tensor (suffix form)

$$t'_{ij} = \sum_{k=1}^{3}\sum_{l=1}^{3} a_{ik}\, a_{jl}\, t_{kl} \qquad (18.15)$$

$(1 \le i, j \le 3)$, *where the* $\{a_{pq}\}$ *are the elements of the transformation matrix between* $\mathcal{C}$ *and* $\mathcal{C}'$.

Definition 18.5 ***Third order tensor*** *Let* $\{t_{ijk}\}$ $(1 \le i, j, k \le 3)$ *be a set of* ***twenty seven*** *real numbers defined in each coordinate system. Then the* $\{t_{ijk}\}$ *are said to be the components of a* ***third order tensor*** *if their values in any pair of coordinate systems* $\mathcal{C}$ *and* $\mathcal{C}'$ *are related by the* ***transformation formula***

Definition of a third order tensor (suffix form)

$$t'_{ijk} = \sum_{l=1}^{3}\sum_{m=1}^{3}\sum_{n=1}^{3} a_{il}\, a_{jm}\, a_{kn}\, t_{lmn} \qquad (18.16)$$

$(1 \le i, j, k \le 3)$, *where the* $\{a_{pq}\}$ *are the elements of the transformation matrix between* $\mathcal{C}$ *and* $\mathcal{C}'$.

Notes on the tensor transformation formulae

The transformation formulae (18.14), (18.15), and (18.16) may seem incomprehensible but they do follow a pattern. In the definition of a vector there is only one summation and one appearance of a_{pq}; in the definition of a tensor of the second order, there are two summations and two appearances of a_{pq}, and so on. *Correct positioning of the suffices is vital*. The suffices of the tensor on the left (in order) must be the same as the first suffix of each of the a_{pq} (in order); and the suffices of the tensor on the right (in order) must be the same as the second suffix of each of the a_{pq} (in order).*

Power users of tensors write formulae such as those above *without* the summation signs. They adopt the **summation convention**† that any repeated suffix is deemed to be summed over its range. Although the summation convention is widely used in most applications of tensors, we do *not* use it here. This is for two reasons: (i) it is a good thing to see the summation signs written in when tensors are first encountered, and (ii) we have no need of the heavy tensor algebra for which the summation convention was designed.

* Be a hero. Write out the transformation formula for a fourth order tensor.

† Invented by Einstein to simplify the writing of the theory of general relativity.

Second order tensors in matrix form

In the general case it is not possible to give any simpler form for the tensor transformation formula. However, for the special case of tensors of the **second order** (which is what we are mainly concerned with), the transformation formula can be written in a more user-friendly form in terms of **matrix products**.

In each coordinate system, a second order tensor $\{t_{ij}\}$ has nine components, indexed by the integers i, j $(1 \le i, j \le 3)$. It is natural then to display these components as a 3×3 array, that is, as the elements of a 3×3 matrix. The tensor $\{t_{ij}\}$ can then be regarded as a 3×3 matrix

$$\mathbf{T} = \begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{pmatrix}.$$

defined in each coordinate system. Also, the tensor transformation rule (18.15) can be wrtten in the form

$$\begin{aligned} t'_{ij} &= \sum_{k=1}^{3}\sum_{l=1}^{3} a_{ik}\, a_{jl}\, t_{kl} = \sum_{l=1}^{3}\left(\sum_{k=1}^{3} a_{ik}\, t_{kl}\right) a_{jl} \\ &= \sum_{l=1}^{3}\left(\sum_{k=1}^{3} a_{ik}\, t_{kl}\right) a^{T}_{lj}, \end{aligned}$$

where the $\{a^{T}_{pq}\}$ are the elements of the transposed matrix $\mathbf{A}^{T}$. The summations in this last expression are equivalent to matrix products and so the transformation rule for a second order tensor can be expressed in the form

$$\mathbf{T}' = \mathbf{A} \cdot \mathbf{T} \cdot \mathbf{A}^{T}.$$

This gives the following **alternative definition**:

Definition 18.6 ***Second order tensor (matrix form)*** *Let* $\mathbf{T}$ *be a* 3×3 *matrix defined in each coordinate system. Then* $\mathbf{T}$ *is said to represent a* ***second order tensor*** *if its values in any pair of coordinate systems* $\mathcal{C}$ *and* $\mathcal{C}'$ *are related by the* ***transformation formula***

Definition of a second order tensor (matrix form)

$$\mathbf{T}' = \mathbf{A} \cdot \mathbf{T} \cdot \mathbf{A}^{T} \qquad (18.17)$$

where the $\mathbf{A}$ *is the transformation matrix between* $\mathcal{C}$ *and* $\mathcal{C}'$.

This definition has the virtue of being expressed in terms of the familiar operation of matrix mutiplication. It only applies to second order tensors but that is exactly what we need. The **inertia tensor** is second order, as are many other important tensors of physics.

Example 18.5 *Transforming a second order tensor*

In the coordinate system $\mathcal{C}$, a second order tensor has the components $t_{11} = 1$, $t_{12} = 2$, $t_{13} = 3$, $t_{21} = 4$, $t_{22} = 5$, $t_{23} = 6$, $t_{31} = 7$, $t_{32} = 8$, $t_{33} = 9$. Find the components of the tensor in $\mathcal{C}'$ when

(i) $\mathcal{C}'$ is obtained from $\mathcal{C}$ by a rotation of 90° about the axis Ox_3, and

(ii) when $\mathcal{C}'$ is obtained from $\mathcal{C}$ by a reflection in the plane $x_3 = 0$.

Solution

In the coordinate system $\mathcal{C}$, the matrix of components is given to be

$$\mathbf{T} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$

Part (i). From the formula (18.8), the transformation matrix of the required rotation of coordinates is

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The transformation formula (18.17) then implies that, in the coordinate systen $\mathcal{C}'$, the tensor is represented by the matrix

$$\mathbf{T}' = \mathbf{A} \cdot \mathbf{T} \cdot \mathbf{A}^T = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} 5 & -4 & 6 \\ -2 & 1 & -3 \\ 8 & -7 & 9 \end{pmatrix}.$$

Hence the **components** of the tensor in $\mathcal{C}'$ are $t'_{11} = 5$, $t'_{12} = -4$, $t'_{13} = 6$, $t'_{21} = -2$, $t'_{22} = 1$, $t'_{23} = -3$, $t'_{31} = 8$, $t'_{32} = -7$, $t'_{33} = 9$.

Part (ii). The transformation matrix of the required **reflection** of coordinates is

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

The transformation formula (18.17) then implies that

$$\mathbf{T}' = \mathbf{A} \cdot \mathbf{T} \cdot \mathbf{A}^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$
$$= \begin{pmatrix} 1 & 2 & -3 \\ 4 & 5 & -6 \\ -7 & -8 & 9 \end{pmatrix}.$$

Hence the **components** of the tensor in $\mathcal{C}'$ are $t'_{11} = 1$, $t'_{12} = 2$, $t'_{13} = -3$, $t'_{21} = 4$, $t'_{22} = 5$, $t'_{23} = -6$, $t'_{31} = -7$, $t'_{32} = -8$, $t'_{33} = 9$. ■

The identity tensor

The simplest second order tensor is the **identity tensor**. This tensor has the components $\{\delta_{ij}\}$ in *every* coordinate system, where δ_{ij} is the Krönecker delta, defined by

$$\delta_{ij} = \begin{cases} 1 & (i = j) \\ 0 & (i \neq j) \end{cases}$$

In each coordinate system, the identity tensor is represented by the **identity matrix** $\mathbf{1}$. The transformation formula (18.17) is certainly satisfied since

$$\mathbf{A} \cdot \mathbf{1} \cdot \mathbf{A}^T = \mathbf{A} \cdot \mathbf{A}^T = \mathbf{1},$$

thus confirming that $\{\delta_{ij}\}$ is a tensor.

The identity tensor one of a very restricted class of tensors that have the *same components in all coordinate systems*. Such tensors are called **isotropic** and they have an important physical rôle. All **scalars** are, by definition, isotropic; there are no (non-trival) isotropic **vectors**; and the only isotropic second order **tensors** are scalar multiples of the identity tensor. It becomes increasingly difficult to identify the isotropic tensors of higher orders!

18.4 TENSOR ALGEBRA

From the linearity of the transformation formulae, it follows that a **scalar multiple** of a tensor is a tensor, and the **sum** of two tensors (of the same order) is a tensor. We will now look at two other ways in which new tensors can be created. This gives an indication how tensors can arise naturally.

The outer product of two tensors

Suppose, for example, that $\{u_{ij}\}$ is a **second order** tensor and $\{v_{ijk}\}$ a **third order** tensor. Then $\{t_{ijklm}\}$ defined* by

$$t_{ijklm} = u_{ij} v_{klm}$$

is a 243 component quantity, defined in each coordinate system. Although we will not prove this here, $\{t_{ijklm}\}$ is a **fifth order tensor** by virtue of satisfying the appropriate transformation formula. This is an example of the **outer product** of two tensors. In general, the *outer product of a tensor of order m and a tensor of order n is a new tensor of order $m + n$*.

---

* To see what this means, write out the definitions of a few elements of $\{t_{ijklm}\}$; for example, $t_{13231} = u_{13} v_{231}$. Note that, in an outer product, all the suffix *names* must be different.

The outer product is one way in which higher order tensors can be constructed from lower order ones. For example, suppose $\{u_i\}$ and $\{v_i\}$ are vectors (tensors of order one). Then the nine component quantity $\{u_i v_j\}$ is a **second order tensor** whose matrix form is

$$\begin{pmatrix} u_1v_1 & u_1v_2 & u_1v_3 \\ u_2v_1 & u_2v_2 & u_2v_3 \\ u_3v_1 & u_3v_2 & u_3v_3 \end{pmatrix}.$$

In the same way, if $\{u_i\}$, $\{v_i\}$ and $\{w_i\}$ are vectors, then the twenty seven component quantity $\{u_i v_j w_k\}$ is a **third order tensor**. Not all higher order tensors can be so constructed, but it is a common procedure. The inertia tensor makes use of this construction.

Contraction of a tensor

Suppose, for example, that $\{t_{ijkl}\}$ is a fourth order tensor. Now select **two** of its suffices (k and l say), set them **equal** (to m say)* and **sum** over the suffix m. The result is the nine component quantity

$$w_{ij} = \sum_{m=1}^{3} t_{ijmm},$$

defined in each coordinate system. Although we will not prove this here, $\{w_{ij}\}$ is a **second order tensor** by virtue of satisfying the appropriate transformation formula. The tensor $\{w_{ij}\}$ is called the **contraction** of $\{t_{ijkl}\}$ with respect to the suffix pair k, l. [There are six different contractions of $\{t_{ijkl}\}$. What are they?] In the general case, *contraction of a tensor produces a new tensor whose order is two less than that of the original tensor*.

Contraction of a second order tensor is the simplest case. If $\{t_{ij}\}$ is a second order tensor, then its contraction with respect to the suffix pair i, j, namely,

$$\sum_{i=1}^{3} t_{ii} = t_{11} + t_{22} + t_{33},$$

must be a tensor of order zero, that is, a **scalar invariant**. If $\{t_{ij}\}$ is represented by the matrix $\mathbf{T}$, then the contraction $t_{11} + t_{22} + t_{33}$ is the sum of the diagonal elements of $\mathbf{T}$. In linear algebra this is called the **trace** of $\mathbf{T}$. The fact that the *trace of a second order tensor is invariant*† is an important result. For example, suppose $\{u_i\}$ and $\{v_i\}$ are vectors. Then the outer product $\{u_i v_j\}$ is a second order tensor. It now follows that the trace of this tensor, namely $u_1v_1 + u_2v_2 + u_3v_3$ is an invariant. This is actually no surprise, since $u_1v_1 + u_2v_2 + u_3v_3$ is the scalar product of the vectors $\{u_i\}$ and $\{v_i\}$, which is known to be independent of the coordinate system in which the components are measured.

* The name of this repeated suffix can be any name not already in use. However, it is permissible to re-use either of the old suffix names that were set equal (k or l in this example).

† This result is by no means obvious. After all, the individual components of $\{t_{ij}\}$ are not invariants.

Another important application of contraction is as follows. Suppose that $\{t_{ij}\}$ is a second order tensor and that $\{u_i\}$ is a vector. Then the outer product $\{t_{ij}u_k\}$ is a third order tensor. If we now contract this third order tensor with respect to the suffix pair j, k, we obtain the vector $\{w_i\}$ given by

$$w_i = \sum_{j=1}^{3} t_{ij}u_j.$$

This formula looks nice in the matrix formulation. If $\{t_{ij}\}$ is represented by the matrix T and $\{u_i\}$ by the column vector u, then $\{v_i\}$ is represented by the column vector v given by

$$\mathsf{v} = \mathsf{T} \cdot \mathsf{u}.$$

In other words, *if the matrix* T *represents a second order tensor and the column vector* u *represents a vector, then the product* $\mathsf{T} \cdot \mathsf{u}$ *represents a vector*. Thus pre-multiplicaton by T transforms one vector into another vector. This is a common way in which tensors act in physics. For example, in crystalline materials, the electric vectors $\boldsymbol{E}$ and $\boldsymbol{D}$ are not parallel. In each coordinate system, they are related by the formula

$$\mathsf{D} = \mathsf{K} \cdot \mathsf{E},$$

where the 3×3 matrix K represents the **dielectric tensor**, the crystalline equivalent of the dielectric constant. This tensor relationship between $\boldsymbol{D}$ and $\boldsymbol{E}$ is the cause of double refraction in crystalline materials. The inertia tensor has a similar rôle, transforming the angular velocity vector into the angular momentum vector.

The angular velocity vector

We have already shown that the **velocity** of a particle is a vector in the sense of the definition (18.13), but we have so far said nothing about the transformation properties of the **angular velocity** of a rigid body. This is made more awkward to decide by the fact that $\boldsymbol{\omega}$ is defined *indirectly*. Suppose that one of the particles of a rigid body is fixed at the origin O. Then $\boldsymbol{\omega}$ is essentially defined to be that 'vector' that gives the velocities of the particles of the body by the formula $\boldsymbol{v} = \boldsymbol{\omega} \times \boldsymbol{r}$. In each coordinate system this formula can also be expressed by the *matrix* product

$$\mathsf{v} = \boldsymbol{\Omega} \cdot \mathsf{x},$$

where

$$\mathsf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \qquad \boldsymbol{\Omega} = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}, \qquad \mathsf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

[Check this.] Now $\{x_i\}$ and $\{v_i\}$ are known to transform as vectors, but nothing is known about the transformation properties of the matrix $\boldsymbol{\Omega}$. However, it can be proved that if a matrix maps vectors into vectors by matrix multiplication, then that matrix represents a second order tensor. Hence $\mathbf{\Omega}$ transforms as a second order tensor. This must, in turn, imply a rule for transforming the $\{\omega_i\}$. It looks as if this rule will be something horrendous, but, amazingly, it turns out that the rule is

$$\begin{pmatrix} \omega_1' \\ \omega_2' \\ \omega_3' \end{pmatrix} = (\det \mathbf{A})\, \mathbf{A} \cdot \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix}, \tag{18.18}$$

which is *not quite* the transformation formula for a vector because of the factor $\det \mathbf{A}$. If the transformation $\mathbf{A}$ acts between systems with the *same* handedness, then $\det \mathbf{A} = +1$ and (18.18) is the vector transformation formula. However, if the transformation acts between systems with opposite handedness, then $\det \mathbf{A} = -1$ and the quantity produced by (18.18) has the wrong sign. Three-component quantities that have this strange behaviour are known as **pseudovectors**. Pseudovectors are reasonably common. The cross product of any two (genuine) vectors is a pseudovector, not a vector. In particular, the moment of a force and the angular momentum of a body are pseudovectors. We do not wish to make an issue of this distinction. Instead, *we will restrict all our coordinate systems to be right-handed*. With this restriction, $\det \mathbf{A}$ is always $+1$ and $\{\omega_i\}$ (and all other pseudovectors) transform as vectors.* We will therefore have no need to distinguish between vectors and pseudovectors and we will call them all 'vectors'.

18.5 THE INERTIA TENSOR

Suppose the **rigid body** $\mathcal{B}$ has one of its particles held fixed at the origin O, but is in otherwise general motion. Then the angular momentum of $\mathcal{B}$ about O is defined by

$$\boldsymbol{L}_O = \sum_{i=1}^{N} \boldsymbol{r}_i \times (m_i \boldsymbol{v}_i) \tag{18.19}$$

in the standard notation, where the position vectors $\{\boldsymbol{r}_i\}$ are measured from the origin O. In Chapter 11 we found the formula for the angular momentum of a rigid body *about its own rotation axis*; this angular momentum is a scalar quantity. Now we will find the formula for the full vector value of $\boldsymbol{L}_O$.

In order that we may introduce component suffices without confusion, it is convenient to omit the suffix i in formula (18.19) and simply write

$$\boldsymbol{L}_O = \sum \boldsymbol{r} \times (m\boldsymbol{v}), \tag{18.20}$$

* It is an interesting fact that angular velocity exists, as a *vector*, only in a space of three dimensions. Had the universe been created with (say) four spatial dimensions, then $\mathbf{x}$ and $\mathbf{v}$ would be (four-dimensional) vectors and $\mathbf{v}$ would be given by the formula $\mathbf{v} = \mathbf{\Omega} \cdot \mathbf{x}$, where $\mathbf{\Omega}$ is a 4×4 anti-symmetric matrix representing a second order tensor. This matrix has *six* independent elements which cannot be fitted into a column of length *four*! The same applies to any spatial dimension other than three.

where the sum is deemed to be taken over all the particles of $\mathcal{B}$.

Since $\mathcal{B}$ is a rigid body and O is fixed, the velocity $\boldsymbol{v}$ is given by $\boldsymbol{v} = \boldsymbol{\omega}\times\boldsymbol{r}$, where $\boldsymbol{\omega}$ is the angular velocity of the body. Then

$$\begin{aligned}\boldsymbol{r}\times\boldsymbol{v} &= \boldsymbol{r}\times(\boldsymbol{\omega}\times\boldsymbol{r})\\ &= (\boldsymbol{r}\cdot\boldsymbol{r})\,\boldsymbol{\omega} - (\boldsymbol{r}\cdot\boldsymbol{\omega})\,\boldsymbol{r}.\end{aligned}$$

If the typical particle P has coordinates (x_1, x_2, x_3) and $\boldsymbol{\omega}$ has the components $\{\omega_1, \omega_2, \omega_3\}$, then the i-th component of $\boldsymbol{r}\times\boldsymbol{v}$ can be written

$$\begin{aligned}(\boldsymbol{r}\times\boldsymbol{v})_i &= \left(x_1^2 + x_2^2 + x_3^2\right)\omega_i - (x_1\omega_1 + x_2\omega_2 + x_3\omega_3)\,x_i\\ &= \left(\sum_{k=1}^{3} x_k x_k\right)\omega_i - \left(\sum_{j=1}^{3} x_j\,\omega_j\right)x_i.\end{aligned}$$

What we want to do now is to make ω_j a factor of this expression, which means that we would like ω_i to be replaced by ω_j. There is a standard trick for doing this, that is, to write

$$\omega_i = \sum_{j=1}^{3} \delta_{ij}\,\omega_j,$$

where δ_{ij} is the Krönecker delta. [This is equivalent to pre-multiplication of the column vector $\boldsymbol{\omega}$ by the identity matrix.] We then obtain

$$\begin{aligned}(\boldsymbol{r}\times\boldsymbol{v})_i &= \left(\sum_{k=1}^{3} x_k x_k\right)\left(\sum_{j=1}^{3} \delta_{ij}\,\omega_j\right) - \left(\sum_{j=1}^{3} x_j\,\omega_j\right)x_i\\ &= \sum_{j=1}^{3}\left(\left(\sum_{k=1}^{3} x_k x_k\right)\delta_{ij} - x_i\,x_j\right)\omega_j\end{aligned}$$

It follows that if $\{L_1, L_2, L_3\}$ are the components of the angular momentum $\boldsymbol{L}_O$, then

$$L_i = \sum_{j=1}^{3} I_{ij}\,\omega_j, \tag{18.21}$$

where the nine-component quantity $\{I_{ij}\}$ is defined by

$$I_{ij} = \sum m\left(\left(\sum_{k=1}^{3} x_k x_k\right)\delta_{ij} - x_i\,x_j\right).$$

It is easy to show directly that $\{I_{ij}\}$ is a second order **tensor**. For each particle of the body, the Cartesian coordinates $\{x_1, x_2, x_2\}$ transform as a vector. It follows that the outer product $\{x_i x_j\}$ is a tensor and its contraction $\sum_{k=1}^{3} x_k x_k$ is a scalar. It then follows by linearity that $\left(\sum_{k=1}^{3} x_k x_k\right)\delta_{ij} - x_i x_j$ is a tensor. Since the particle masses are certainly invariants, it once again follows by linearity that $\{I_{ij}\}$ must be a tensor.

Definition 18.7 ***Inertia tensor*** *The second order tensor $\{I_{ij}\}$ defined by*

Definition of the inertia tensor

$$I_{ij} = \sum m \left(\left(\sum_{k=1}^{3} x_k x_k \right) \delta_{ij} - x_i x_j \right) \tag{18.22}$$

is called the ***inertia tensor*** *of the body $\mathcal{B}$ at the point O. Note that the inertia tensor is* ***symmetric****, that is, $I_{ji} = I_{ij}$.*

We have therefore proved that, when a rigid body is rotating with angular velocity $\boldsymbol{\omega}$ about an axis through origin O, its angular momentum $\boldsymbol{L}_O$ about O is given by

$$\begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix} = \begin{pmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix}. \tag{18.23}$$

This can be written in matrix form as

Angular momentum formula

$$\mathsf{L}_O = \mathbf{I} \cdot \boldsymbol{\omega} \tag{18.24}$$

where the column vectors L_O and $\boldsymbol{\omega}$ are formed from the components of $\boldsymbol{L}_O$ and $\boldsymbol{\omega}$ respectively, and $\mathbf{I}$ is the matrix form of the inertia tensor $\{I_{ij}\}$. Because L_O is obtained from $\boldsymbol{\omega}$ by a *matrix* multiplication, it follows that $\boldsymbol{L}_O$ and $\boldsymbol{\omega}$ will generally lie in *different directions*. This fact (which does not show itself in planar mechanics) is what gives three-dimensional rigid body motion its special character.

The inertia tensor also appears in the corresponding expression for the **kinetic energy** of $\mathcal{B}$. By following a similar procedure to that used for angular momentum, it is found that T is given by

$$T = \tfrac{1}{2} \begin{pmatrix} \omega_1 & \omega_2 & \omega_3 \end{pmatrix} \begin{pmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \tag{18.25}$$

which can be written in the matrix form

Kinetic energy formula

$$T = \tfrac{1}{2} \boldsymbol{\omega}^T \cdot \mathbf{I} \cdot \boldsymbol{\omega} \tag{18.26}$$

The elements of the inertia tensor

The **diagonal** elements of $\mathbf{I}$ are actually familiar quantities. For example

$$\begin{aligned} I_{11} &= \sum m\left(\left(x_1^2+x_2^2+x_3^2\right)\delta_{11}-x_1x_1\right)\\ &= \sum m\left(x_2^2+x_3^2\right)\\ &= \sum mp_1^2, \end{aligned}$$

where p_1 is the perpendicular distance of the typical particle P from the axis Ox_1. Hence I_{11} is just the *ordinary moment of inertia* of the body about the axis Ox_1. Similarly, I_{22} and I_{33} are the moments of inertia of the body about the axes Ox_2 and Ox_3 respectively.

The **off-diagonal** elements of $\mathbf{I}$ are given by

$$I_{12} = I_{21} = -\sum mx_1x_2, \tag{18.27}$$

$$I_{23} = I_{32} = -\sum mx_2x_3, \tag{18.28}$$

$$I_{31} = I_{13} = -\sum mx_3x_1, \tag{18.29}$$

where, as usual, the sum is taken over the particles of the body. The quantities $\sum mx_1x_2$, $\sum mx_2x_3$, $\sum mx_3x_1$ are known as **products of inertia**. Hence the off-diagonal elements of $\mathbf{I}$ are the *negatives* of their corresponding products of inertia.* Our results are summarised as follows:

Elements of the inertia tensor

- The **diagonal elements** of the inertia tensor, namely I_{11}, I_{22}, I_{33}, are the ordinary moments of inertia of the body about the coordinate axes Ox_1, Ox_2, Ox_3 respectively.
- The **off-diagonal elements** of the inertia tensor are the *negatives* of their corresponding products of inertia.

Products of inertia are new quantities and the reader will probably expect that we should now give several examples on their calculation. In fact, we will give just one easy one, which is a part of the next example. The reason is that, in almost all problems, the products of inertia can be inferred from symmetry and known results. Only a body with virtually no symmetry would *require* products of inertia to be calculated from first principles.

* Some authors include the negative sign in their definition of product of inertia, but this seems unnatural.

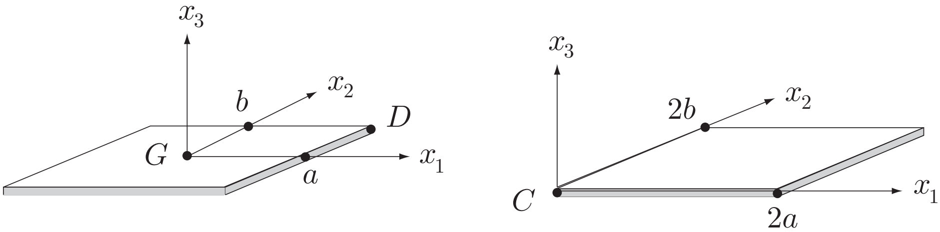


FIGURE 18.3 The uniform rectangular plate has mass M and sides $2a$ and $2b$

Example 18.6 *Calculating the inertia tensor*

A uniform rectangular plate has mass M and sides $2a$ and $2b$ as shown in Figure 18.3. Find the inertia tensor at the points G and C in the coordinate systems shown.

Solution

Consider first the **centre of mass** point G. Then I_{11} is the moment of inertia of the plate about the axis Gx_1 which is given in the Appendix (at the end of the book) to be $\frac{1}{3}Mb^2$. Similarly, $I_{22} = \frac{1}{3}Ma^2$ and, by the perpendicular axes theorem, $I_{33} = I_{11}+I_{22} = \frac{1}{3}M(a^2+b^2)$. The products of inertia are all zero. The products $\sum mx_2x_3$ and $\sum mx_3x_1$ are zero because all the mass lies in the plane $x_3 = 0$. The product $\sum mx_1x_2$ is zero because of symmetry; the plate is symmetrical about the axis $x_1 = 0$ but the terms of the sum are *odd* functions of x_1. The contributions to the sum from the right and left halves of the plate therefore cancel. Hence $I_{23} = I_{31} = I_{12} = 0$. Thus, in the given coordinate system, the **inertia tensor** at G is represented by the matrix

$$\mathbf{I}_G = \frac{1}{3}M \begin{pmatrix} b^2 & 0 & 0 \\ 0 & a^2 & 0 \\ 0 & 0 & a^2+b^2 \end{pmatrix}.$$

Now consider the **corner point** C. Then I_{11} is the moment of inertia of the plate about the axis Cx_1 which is given by the parallel axes theorem to be $\frac{1}{3}Mb^2 + Mb^2 = \frac{4}{3}Mb^2$. Similarly, $I_{22} = \frac{4}{3}Ma^2$ and, by the perpendicular axes theorem, $I_{33} = I_{11} + I_{22} = \frac{4}{3}M(a^2 + b^2)$. The elements I_{23} and I_{31} are zero for the same reason as before. However, the product $\sum mx_1x_2$ is not zero now because the plate is no longer symmetrically placed relative to the coordinate system based at C. We will therefore evaluate this product of inertia from first principles. Let the plate have uniform mass per unit area σ. Then

$$\sum mx_1x_2 = \int_{x_1=0}^{2a}\int_{x_2=0}^{2b} \sigma x_1x_2\,dx_1dx_2 = 4\sigma a^2b^2 = Mab,$$

since $M = 4\sigma ab$. Hence $I_{12} = Mab$. Thus, in the given coordinate system, the **inertia tensor** of the plate at C is represented by the matrix

$$\mathbf{I}_C = \frac{1}{3}M\begin{pmatrix} 4b^2 & -3ab & 0 \\ -3ab & 4a^2 & 0 \\ 0 & 0 & 4(a^2+b^2) \end{pmatrix}. \blacksquare$$

Example 18.7 *Transforming the inertia tensor*

Find $\mathbf{I}_G$ in the set of axes obtained by rotating the axes $Gx_1x_2x_3$ about Gx_3 so that the new axis Gx_1' lies along the diagonal GD.

Solution

In order to achieve the required position, the axes $Gx_1x_2x_3$ must be rotated through the acute angle $\alpha = \tan^{-1}(b/a)$. The corresponding transformation matrix is given by the formula (18.8) to be

$$\mathbf{A} = \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The matrix $\mathbf{I}_G'$ representing the inertia tensor at G in the rotated coordinates is given by the transformation formula (18.17) to be

$$\begin{aligned}
\mathbf{I}_G' &= \mathbf{A}\cdot\mathbf{I}_G\cdot\mathbf{A}^T \\
&= \frac{1}{3}M\begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} b^2 & 0 & 0 \\ 0 & a^2 & 0 \\ 0 & 0 & a^2+b^2 \end{pmatrix}\begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \frac{1}{3}M\begin{pmatrix} a^2\sin^2\alpha + b^2\cos^2\alpha & (a^2-b^2)\sin\alpha\cos\alpha & 0 \\ (a^2-b^2)\sin\alpha\cos\alpha & a^2\cos^2\alpha + b^2\sin^2\alpha & 0 \\ 0 & 0 & a^2+b^2 \end{pmatrix} \\
&= \frac{M}{3(a^2+b^2)}\begin{pmatrix} 2a^2b^2 & ab(a^2-b^2) & 0 \\ ab(a^2-b^2) & a^4+b^4 & 0 \\ 0 & 0 & (a^2+b^2)^2 \end{pmatrix},
\end{aligned}$$

on inserting the values $\sin\alpha = b/(a^2+b^2)^{1/2}$ and $\cos\alpha = a/(a^2+b^2)^{1/2}$. This is the inertia tensor of the plate at G in the **rotated coordinate system**. ■

Example 18.8 *Finding kinetic energy using the inertia tensor*

Suppose that the plate in Figure 18.3 is made to rotate about one of its diameters with angular speed λ. Find its kinetic energy.

Solution

The kinetic energy can be evaluated by using either the original or the rotated coordinate system.

In the **original** coordinate system,

$$\begin{aligned}
T &= \tfrac{1}{2}\boldsymbol{\omega}^T \cdot \mathbf{I}_G \cdot \boldsymbol{\omega} \\
&= \tfrac{1}{6}M \begin{pmatrix} \lambda\cos\alpha & \lambda\sin\alpha & 0 \end{pmatrix} \begin{pmatrix} b^2 & 0 & 0 \\ 0 & a^2 & 0 \\ 0 & 0 & a^2+b^2 \end{pmatrix} \begin{pmatrix} \lambda\cos\alpha \\ \lambda\sin\alpha \\ 0 \end{pmatrix} \\
&= \tfrac{1}{6}M\lambda^2 \left(b^2\cos^2\alpha + a^2\sin^2\alpha \right) \\
&= \frac{Ma^2b^2\lambda^2}{3(a^2+b^2)}.
\end{aligned}$$

In the **rotated** coordinate system,

$$\begin{aligned}
T &= \tfrac{1}{2}\boldsymbol{\omega}^T \cdot \mathbf{I}_G \cdot \boldsymbol{\omega} \\
&= \frac{M}{6(a^2+b^2)} \begin{pmatrix} \lambda & 0 & 0 \end{pmatrix} \begin{pmatrix} 2a^2b^2 & ab(a^2-b^2) & 0 \\ ab(a^2-b^2) & a^4+b^4 & 0 \\ 0 & 0 & (a^2+b^2)^2 \end{pmatrix} \begin{pmatrix} \lambda \\ 0 \\ 0 \end{pmatrix} \\
&= \frac{Ma^2b^2\lambda^2}{3(a^2+b^2)}.
\end{aligned}$$

Starting from scratch, the first calculation is quicker. ■

18.6 PRINCIPAL AXES OF A SYMMETRIC TENSOR

The nine components of a second order tensor depend on the coordinate system, their values in different coordinate systems being related by the transformation formula (18.17). The question naturally arises as to whether we can simplify the representation of a tensor by a clever choice of coordinate system. If a tensor is represented by the matrix $\mathbf{T}$ in one coordinate system $\mathcal{C}$, then, in another coordinate system $\mathcal{C}'$, it is represented by the matrix $\mathbf{T}'$ given by

$$\mathbf{T}' = \mathbf{A} \cdot \mathbf{T} \cdot \mathbf{A}^T, \tag{18.30}$$

where $\mathbf{A}$ is the orthogonal transformation matrix between $\mathcal{C}$ and $\mathcal{C}'$. We therefore wish to find an orthogonal matrix $\mathbf{A}$ (with determinant $+1$) that makes the product $\mathbf{A} \cdot \mathbf{T} \cdot \mathbf{A}^T$ as 'simple' as possible, preferably a **diagonal** matrix. The question is therefore equivalent to a problem in linear algebra. There are a number of theorems available concerning the diagonalisation of matrices by transformations resembling (18.30) (see Anton [7]) and one of these results is exactly what we need. It is the **orthogonal diagonalisation** theorem:

Theorem 18.1 ***Orthogonal diagonalisation*** *Given any real* ***symmetric*** *matrix* $\mathbf{T}$, *there is an orthogonal matrix* $\mathbf{A}$ *such that the product* $\mathbf{A} \cdot \mathbf{T} \cdot \mathbf{A}^T$ *is a* ***diagonal*** *matrix.*

Since the **inertia tensor** is symmetric, the above theorem applies. Thus, at any point O of the body, one can find a set of axes in which $\mathbf{I}$ is a diagonal matrix. These are called **principal axes** of the body at O. The diagonal elements of $\mathbf{I}$ are simply the moments of inertia of the body about the three principal axes. They are called the **principal moments of inertia** at O and will be denoted by symbols A, B and C. This is summarised as follows:

Principal axes and principal moments of inertia

Relative to a set of **principal axes** $Ox_1x_2x_3$ at O, the inertia tensor $\mathbf{I}$ has the diagonal form

$$\mathbf{I} = \begin{pmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & C \end{pmatrix}, \qquad (18.31)$$

where the **principal moments of inertia** A, B, C are the moments of inertia of the body about the axes Ox_1, Ox_2, Ox_3 respectively.

Finding principal axes

There is a standard procedure for finding the orthogonal matrix **A** that reduces any symmetric matrix **T** to diagonal form in the above manner. The diagonal elements of the reduced matrix are the **eigenvalues** of the matrix **T**, and the rows of **A** contain the components of the corresponding normalised **eigenvectors** of **T**. However, for the purpose of solving problems in mechanics, we (almost) never need to carry through this procedure. The reason is that, in almost all problems, the orientation of the principal axes can be deduced by symmetry, as we will show in Section 18.7.

Expressions for L and T in principal axes

When principal axes are used, the expressions (18.24) for the angular momentum and (18.26) for the kinetic energy of a rigid body are much simplified. They reduce to the non-tensorial expressions

Expressions for $\boldsymbol{L}_O$ and T in principal axes at O

$$\boldsymbol{L}_O = (A\,\omega_1)\,\boldsymbol{e}_1 + (B\,\omega_2)\,\boldsymbol{e}_2 + (C\,\omega_3)\,\boldsymbol{e}_3$$

$$T = \tfrac{1}{2}\left(A\,\omega_1^2 + B\,\omega_2^2 + C\,\omega_3^2\right) \qquad (18.32)$$

where $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ are the unit vectors of the principal axes, and $\{\omega_1, \omega_2, \omega_3\}$ are the components of $\boldsymbol{\omega}$ in the principal axes, that is, $\boldsymbol{\omega} = \omega_1\boldsymbol{e}_1 + \omega_2\boldsymbol{e}_2 + \omega_3\boldsymbol{e}_3$.

These results apply to the case in which the body has a permanently fixed particle O (the vertex of a spinning top, for example). However, similar formulae apply to the case of completely **general motion**. In this case we have

> **Expressions for L_G and T in principal axes at G**
>
> $$\boldsymbol{L}_G = (A\,\omega_1)\,\boldsymbol{e}_1 + (B\,\omega_2)\,\boldsymbol{e}_2 + (C\,\omega_3)\,\boldsymbol{e}_3 \tag{18.33}$$
>
> $$T^G = \tfrac{1}{2}\left(A\,\omega_1^2 + B\,\omega_2^2 + C\,\omega_3^2\right)$$

where G is the centre of mass of the body. Here $\boldsymbol{L}_G$ is the angular momentum of the body about G, and T^G is the *rotational* kinetic energy of the body about G.

These are the results that you *really* need in order to solve problems in rigid body dynamics. Almost always, principal axes are used and only the equations (18.32) or (18.33) are required. Thus, it is possible to solve most problems in rigid body dynamics without any knowledge of tensors at all!

18.7 DYNAMICAL SYMMETRY

Finding principal axes from geometrical symmetries of the body

The directions of the principal axes of the inertia tensor of a body at a point O can often be inferred from the **geometrical symmetry** of the body about O. The following rules (which are consequences of the tensor transformation formula) are useful.

Rule 1: *If the body has* ***reflective symmetry*** *in a plane through O, then the line through O perpendicular to this plane is a principal axis.*

Rule 2: *If the body has any* ***rotational symmetry*** *about an axis through O, then this axis is a principal axis.*

Either of these rules is enough to show that, for a uniform **rectangular plate**, the axes shown in Figure 18.3 (left) are a set of principal axes at G. The plate has reflective symmetry in *each* of the coordinate planes, and also has rotational symmetry of order two about *each* of the coordinate axes. The set of parallel axes at C is not a principal set. For a **uniform cone** on a pentagonal base (see Figure 18.4 (left)), the axis of symmetry is a principal axis at G and there are *five* other principal axes through G. [What are they?] For a **spinning top** (see Figure 18.4 (centre)), the axis of symmetry is a principal axis at O and so is *any* axis through O that is perpendicular to it! We seem to be finding too many principal axes, but this will soon be explained.

Dynamical symmetry

Consider a rigid body $\mathcal{B}$ pivoted at the origin O. Then the motion of $\mathcal{B}$ under known forces is determined by the **angular momentum principle**, which, in view of equation (18.24), can be written

$$\frac{d}{dt}\left(\mathbf{I}_O\cdot\boldsymbol{\omega}\right) = \mathbf{K}_O$$

where $\boldsymbol{K}_O$ is the total moment about O of the external forces acting on $\mathcal{B}$. In principal axes, $\mathbf{I}_O$ contains only the three principal moments of inertia A, B and C and so the shape

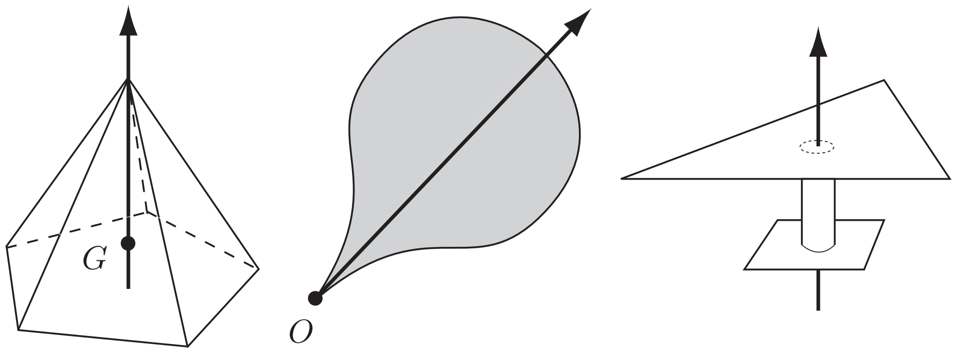


FIGURE 18.4 Three bodies with **dynamical axial symmetry**.

of the body enters into the equations of motion only by the values of these three numbers. Under the same set of forces and initial conditions, *different bodies that have the same principal moments of inertia have the same motion*.

The form of the motion, and the chance of solving problems, is very much influenced by the presence or absence of **dynamical symmetry**. There are three cases:

A, B, C all different

When $\{A, B, C\}$ are all different, the body is said to be **dynamically unsymmetric**, and, as might be expected, this is the most difficult case to treat analytically. For the rectangular plate shown in Figure 18.3, the principal moments of inertia at G are $A = \frac{1}{3}Mb^2$, $B = \frac{1}{3}Ma^2$ and $C = \frac{1}{3}M(a^2+b^2)$, which are all different. Thus, despite the apparently simple shape of the plate, it is no easier to treat than a lump of scrap metal. In the unsymmetrical case, the principal axes at O are *essentially unique*.*

A and B equal, C different

Suppose now that two of the principal moments of inertia (A and B say) are equal, the third, C, being different. The body is then said to be **dynamically axially symmetric** about the axis Ox_3. This time, the axis of symmetry is a principal axis, and *the other two axes can be chosen arbitrarily*, provided that they form a right-handed orthogonal set. One might expect this to occur only when the body is a body of revolution about an axis through O, as in the case of a traditional spinning top. However, the following remarkable result shows that this is not neccessary.

Rule 3: *If a body has a rotational symmetry about an axis through O, and the order*† *of this symmetry is three or more, then the body has* ***dynamical axial symmetry*** *about this axis.*

* The principal axes *are* unique in the sense that the three orthogonal unlabelled lines in space are unique. There are then 24 ways of labelling these lines to create a right-handed system of coordinates.

† If the mass distribution of a rigid body is left unchanged when the body is rotated through an angle $2\pi/n$, then the body is said to have **rotational symmetry of order** n about the rotation axis. Thus a uniform lamina having the shape of a regular pentagon has rotational symmetry of order five about the axis through G perpendicular to its plane.

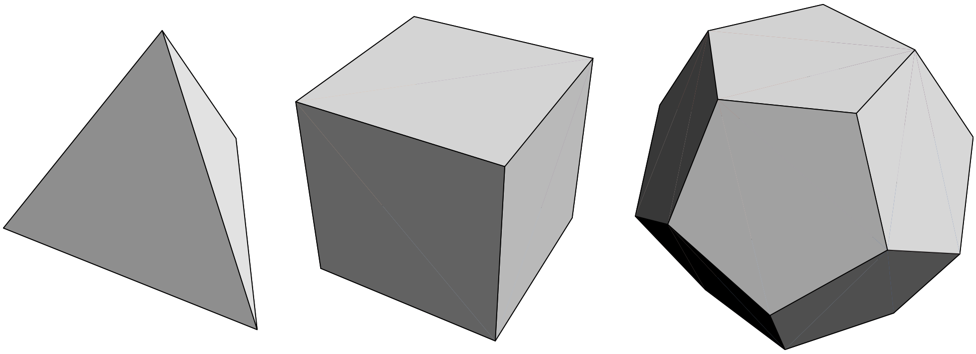


FIGURE 18.5 Three bodies with **dynamical spherical symmetry**.

The pentagonal cone in Figure 18.4 has rotational symmetry of order five and so Rule 3 applies. In particular, it follows that the cone has the same moment of inertia about *any* axis through G pependicular to the symmetry axis!

Interestingly, it is possible for a body to have dynamical axial symmetry and yet have no rotational symmetry at all. The body in Figure 18.4 (right) has no rotational symmetry, but, since it is made up of three bodies that do have such symmetry, it must be dynamically axially symmetric!

A, B and C all equal

Suppose now that all three of the principal moments of inertia at O are equal. The body is then said to be **dynamically spherically symmetric** about O. This time, the inertia tensor is a multiple of the identity and *any* axis through O is a principal axis. The principal axes can therefore be chosen arbitrarily, provided that they form a right-handed orthogonal set. A uniform ball is obviously dynamically spherically symmetric, but there are many other examples. This follows from our last rule, Rule 4.

Rule 4: *If a body has two different axes of dynamical axial symmetry through O, then it must be* ***dynamically spherically symmetric*** *about O.*

Figure 18.5 shows three bodies that have **dynamical spherical symmetry** about their centres of mass. None of these bodies is a ball, but they move as if they were. For example, the uniform tetrahedron has *four* different axes of dynamical *axial* symmetry through G and so, by Rule 4, it must be dynamically *spherically* symmetric about G. As a consequence, it must have the same moment of inertia about *any* axis through G. [Don't try proving this any other way!]

Problems on Chapter 18

Answers and comments are at the end of the book.

Harder problems carry a star (∗).

Orthogonal transformations

18.1 Show that the matrix

$$\mathbf{A} = \frac{1}{7}\begin{pmatrix} 3 & 2 & 6 \\ -6 & 3 & 2 \\ 2 & 6 & -3 \end{pmatrix}$$

is orthogonal. If $\mathbf{A}$ is the transformation matrix between the coordinate systems $\mathcal{C}$ and $\mathcal{C}'$, do $\mathcal{C}$ and $\mathcal{C}'$ have the same, or opposite, handedness?

18.2 Find the transformation matrix between the coordinate systems $\mathcal{C}$ and $\mathcal{C}'$ when $\mathcal{C}'$ is obtained

(i) by rotating $\mathcal{C}$ through an angle of 45° about the axis Ox_2,
(ii) by reflecting $\mathcal{C}$ in the plane $x_2 = 0$,
(iii) by rotating $\mathcal{C}$ through a right angle about the axis $\overrightarrow{OB}$, where B is the point with coordinates (2, 2, 1),
(iv) by reflecting $\mathcal{C}$ in the plane $2x_1 - x_2 + 2x_3 = 0$.

In each case, find the new coordinates of the point D whose coordinates in $\mathcal{C}$ are (3, −3, 0).

18.3 Show that the matrix

$$\mathbf{A} = \frac{1}{3}\begin{pmatrix} 2 & -1 & -2 \\ 2 & 2 & 1 \\ 1 & -2 & 2 \end{pmatrix}$$

is orthogonal and has determinant +1. Find the column vectors $\mathbf{v}$ that satisfy the equation $\mathbf{A} \cdot \mathbf{v} = \mathbf{v}$. If $\mathbf{A}$ is the transformation matrix between the coordinate systems $\mathcal{C}$ and $\mathcal{C}'$, show that $\mathbf{A}$ represents a rotation of $\mathcal{C}$ about the axis $\overrightarrow{OE}$ where E is the point with coordinates (1, 1, −1) in $\mathcal{C}$.

∗ Find the rotation angle.

Tensor algebra

18.4 Write out the transformation formula for a fifth order tensor. [The main difficulty is finding enough suffix names!]

18.5 In the coordinate system $\mathcal{C}$, a certain second order tensor is represented by the matrix

$$\mathbf{T} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Find the matrix representing the tensor in the coordinate system $\mathcal{C}'$, where $\mathcal{C}'$ is obtained

(i) by rotating $\mathcal{C}$ through an angle of 45° about the axis Ox_1,
(ii) by reflecting $\mathcal{C}$ in the plane $x_3 = 0$.

18.6 The quantities t_{ijk} and u_{ijkl} are third and fourth order tensors respectively. Decide if each of the following quantities is a tensor and, if it is, state its order:

(i) $t_{ijk}u_{lmnp}$ (ii) $t_{ijk}t_{lmn}$ (iii) $\sum_{j=1}^{3} t_{ijj}$

(iv) $\sum_{j=1}^{3} t_{jij}$ (v) $\sum_{i=1}^{3} t_{iii}$ (vi) $\sum_{k=1}^{3} t_{ijk}u_{klmn}$

(vii) $\sum_{i=1}^{3}\sum_{j=1}^{3} u_{iijj}$ (viii) $\sum_{k=1}^{3} u_{klmn}$ (ix) $\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3} t_{ijk}t_{ijk}$.

18.7 Show that the sum of the squares of the elements of a tensor is an invariant. [First and second order tensors will suffice.]

18.8 If the matrix T represents a second order tensor, show that $\det \mathsf{T}$ is an invariant. [We have now found three invariant functions of a second order tensor: the sum of the diagonal elements, the sum of the squares of all the elements, and the determinant.]

18.9 In crystalline materials, the ordinary elastic moduli are replaced by c_{ijkl}, a fourth order tensor with eighty one elements. It appears that the most general material has eighty one elastic moduli, but this number is reduced because c_{ijkl} has the following symmetries:

(i) $c_{jikl} = c_{ijkl}$ (ii) $c_{ijlk} = c_{ijkl}$ (iii) $c_{klij} = c_{ijkl}$

How many elastic moduli does the most general material actually have?

Inertia tensor and principal axes

The following problems do *not* require moments or products of inertia to be evaluated by integration.

18.10 Show that $I_{\{O,\boldsymbol{n}\}}$, the moment of inertia of a body about an axis through O parallel to the unit vector $\boldsymbol{n}$, is given by

$$I_{\{O,\boldsymbol{n}\}} = \mathsf{n}^T \cdot \mathbf{I}_O \cdot \mathsf{n},$$

where $\mathbf{I}_O$ is the matrix representing the inertia tensor of the body at O (in some coordinate system), and n is the column vector that contains the components of $\boldsymbol{n}$ (in the same coordinate system).

Find the moment of inertia of a uniform rectangular plate with sides $2a$ and $2b$ about a diagonal.

18.11 Find the principal moments of inertia of a uniform circular disk of mass M and radius a (i) at its centre of mass, and (ii) at a point on the edge of the disk.

18.12 A uniform circular disk has mass M and radius a. A spinning top is made by fitting the disk with a light spindle AB which passes through the disk and is fixed along its axis of symmetry. The distance of the end A from the disk is equal to the disk radius a. Find the principal moments of inertia of the top at the end A of the spindle.

18.13 A uniform hemisphere has mass M and radius a. A spinning top is made by fitting the hemisphere with a light spindle AB which passes through the hemisphere and is fixed along its axis of symmetry with the curved surface of the hemisphere facing away from the end A. The distance of A from the point where the spindle enters the flat surface is equal to the radius a of the hemisphere. Find the principal moments of inertia of the top at the end A of the spindle.

18.14 Find the principal moments of inertia of a uniform cube of mass M and side $2a$ (i) at its centre of mass, (ii) at the centre of a face, and (iii) at a corner point.

Find the moment of inertia of the cube (i) about a space diagonal, (ii) about a face diagonal, and (iii) about an edge.

18.15 A uniform rectangular block has mass M and sides $2a$, $2b$ and $2c$. Find the principal moments of inertia of the block (i) at its centre of mass, (ii) at the centre of a face of area $4ab$. Find the moment of inertia of the block (i) about a space diagonal, (ii) about a diagonal of a face of area $4ab$.

18.16 Find the principal moments of inertia of a uniform cylinder of mass M, radius a and length $2b$ at its centre of mass G. Is it possible for the cylinder to have dynamical *spherical* symmetry about G?

18.17 Determine the dynamical symmetry (if any) of each the following bodies about their centres of mass:

(i) a frisbee,
(ii) a piece of window glass having the shape of an isosceles triangle,
(iii) a two bladed aircraft propellor,
(iv) a three-bladed ship propellor,
(v) an Allen screw (ignore the thread),
(vi) eight particles of equal mass forming a rigid cubical structure,
(vii) a cross-handled wheel nut wrench,
(viii) the great pyramid of Giza,
(ix) a molecule of carbon tetrachloride.

18.18* A uniform rectangular plate has mass M and sides $2a$ and $4a$. Find the principal axes and principal moments of inertia at a *corner* point of the plate. [Make use of the formula for $\mathbf{I}_C$ obtained in Example 18.6, with $b = 2a$.]

If you know how, do this question by finding the eigenvalues and eigenvectors of $\mathbf{I}_C$. If not, try the following homespun method: Starting from $\mathbf{I}_C$ in the coordinates used in Example 18.6, find $\mathbf{I}'_C$ in the coordinate system obtained by rotating through an angle α around the axis Cx_3. Then choose α to eliminate the off-diagonal elements.

Chapter Nineteen

Problems in rigid body dynamics

KEY FEATURES

The key features of this chapter are the use of the linear and angular momentum principles to generate the **equations of rigid body motion**; the importance of **body symmetry** in simplifying the problem; and the choice of solution method, **vectorial, Lagrangian** or **Eulerian**.

Readers who have reached this point have come a long way and deserve to be congratulated. With the linear and angular momentum principles and the inertia tensor behind us, we are now able to solve some of the most interesting and puzzling problems in mechanics. Rigid body motion is the pinnacle of achievment* in the 'problem solving' kind of mechanics but, given the background we now have, many problems are surprisingly easy. In this chapter, we classify problems by the dynamical symmetry of the body; first we consider **spherically symmetric** bodies, then **axially symmetric** bodies and finally **unsymmetric** bodies. Most problems can be done by either vectorial or analytical methods and we choose whatever is appropriate in each case.

19.1 EQUATIONS OF RIGID BODY DYNAMICS

Rigid body in general motion

Translational motion of G:
$$M\,\frac{d\boldsymbol{V}}{dt} = \boldsymbol{F} \qquad (19.1)$$

Rotational motion about G:
$$\frac{d\boldsymbol{L}_G}{dt} = \boldsymbol{K}_G \qquad (19.2)$$

* More advanced topics exist, but they are more concerned with 'grand general principles' than problem solving!

The governing equations of rigid body dynamics are the **linear** and **angular momentum principles** in their centre of mass form, as shown above; the notation is that used in Chapters 10 and 11. The linear momentum principle (19.1) determines the *translational* motion of the centre of mass G, and the angular momentum principle (19.2) determines the *rotational* motion of the body relative to G.

The above equations could be applied in all cases, but, when the body is **pivoted*** at a fixed point O, it is more convenient to drop the first equation and apply the angular momentum principle about the pivot point O instead of G. With this choice, the unknown reaction at the pivot is eliminated. We then have:

Rigid body pivoted at a fixed point O

Rotational motion about O :
$$\frac{d\boldsymbol{L}_O}{dt} = \boldsymbol{K}_O \tag{19.3}$$

Calculating $\boldsymbol{L}_G$ and $\boldsymbol{L}_O$ in principal axes

In order to use the equation (19.2) (or (19.3)) we need to express $\boldsymbol{L}_G$ (or $\boldsymbol{L}_O$) in terms of $\boldsymbol{\omega}$, the angular velocity of the body. In general, this is a tensor relation, but, in principal axes, it reduces to the vector form obtained in Section 18.6. The results are as follows:

Suppose $Gx_1x_2x_3$ are principal axes at G with associated unit vectors $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$. Then

$$\boldsymbol{L}_G = A\omega_1\,\boldsymbol{e}_1 + B\omega_2\,\boldsymbol{e}_2 + C\omega_3\,\boldsymbol{e}_3 \tag{19.4}$$

where A, B, C are the principal moments of inertia of the body about the axes Gx_1, Gx_2, Gx_3 respectively. Similarly, the angular momentum $\boldsymbol{L}_O$ can be expressed in terms of the angular velocity $\boldsymbol{\omega}$, relative to principal axes of the body at O, by the formula

$$\boldsymbol{L}_O = A\omega_1\,\boldsymbol{e}_1 + B\omega_2\,\boldsymbol{e}_2 + C\omega_3\,\boldsymbol{e}_3 \tag{19.5}$$

where A, B, C are now the principal moments of inertia of the body about the axes Ox_1, Ox_2, Ox_3 respectively. In each case, $\{\omega_1, \omega_2, \omega_3\}$ are the components of $\boldsymbol{\omega}$ in the principal axes, that is,

$$\boldsymbol{\omega} = \omega_1\,\boldsymbol{e}_1 + \omega_2\,\boldsymbol{e}_2 + \omega_3\,\boldsymbol{e}_3.$$

Dynamical symmetry of the body

The most important general feature of a moving rigid body is its **dynamical symmetry** (see Section 18.7). The more symmetry the body has, the simpler is its motion and the

* The spinning top with its vertex held at a fixed point is a typical example.

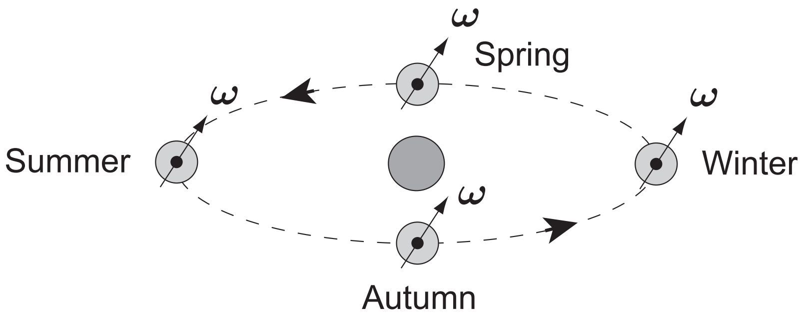


FIGURE 19.1 In the motion of the Earth, its angular velocity $\boldsymbol{\omega}$ is preserved and is *not* normal to the plane of the orbit. This gives rise to the seasons (shown here for the northern hemisphere).

easier it is to calculate it. We begin with the simplest case, which is when the body has dynamical **spherical symmetry**. Remember that, in order to have *dynamical* spherical symmetry, the body need not actually be spherical; indeed, one can find bodies that have no geometrical symmetry at all that are dynamically spherical.

19.2 MOTION OF 'SPHERES'

Suppose the body has dynamical **spherical symmetry** about its centre of mass G. The body could be a uniform sphere, but it does not have to be. All we require is that $A = B = C$ at G, in which case the body has the same moment of inertia A about every axis through G. Equation (19.4) then simplifies to give

$$\boldsymbol{L}_G = A\omega_1 \boldsymbol{e}_1 + A\omega_2 \boldsymbol{e}_2 + A\omega_3 \boldsymbol{e}_3 = A\,(\omega_1 \boldsymbol{e}_1 + \omega_2 \boldsymbol{e}_2 + \omega_3 \boldsymbol{e}_3) = A\,\boldsymbol{\omega},$$

so that $\boldsymbol{L}_G$ is simply proportional to $\boldsymbol{\omega}$. Equation (19.2) then becomes

$$A\dot{\boldsymbol{\omega}} = \boldsymbol{K}_G.$$

If the body is moving under **no forces**, or under **uniform gravity**, then the total moment $\boldsymbol{K}_G = \boldsymbol{0}$ and we obtain

$$\boldsymbol{\omega} = \boldsymbol{\omega}_0,$$

a constant. Thus, in either of these cases, the *angular velocity of the sphere is preserved* in the motion. Meanwhile, the motion of G, as determined by equation (19.1), is either straight line or parabolic motion respectively. Hence, when one throws a ball (or a cube!), G traces out the usual parabola and $\boldsymbol{\omega}$ retains the value that it was given initially. Note that $\boldsymbol{\omega}$ retains its *direction* as well as its magnitude, so that even though the direction of $\boldsymbol{V}$ changes, the rotation axis maintains a constant direction in space.

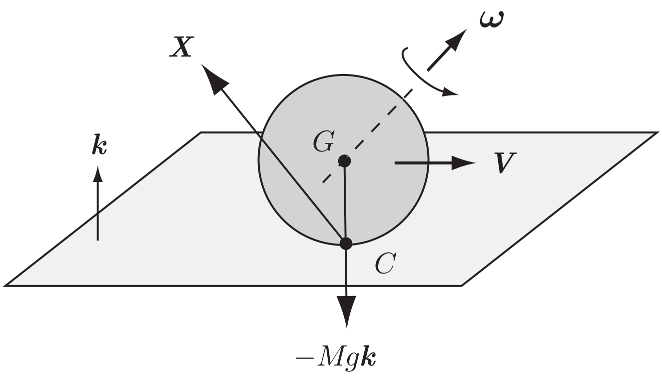


FIGURE 19.2 The snooker ball is a uniform sphere of mass M and radius b that moves in contact with a horizontal table.

The most important instance of this is the motion of the Earth in its orbit. If we assume that the Sun and the Earth are spherically symmetric bodies,* then the total moment of forces $\boldsymbol{K}_G$ exerted on the Earth by the Sun is zero and $\boldsymbol{\omega}$ is preserved. Since $\boldsymbol{\omega}$ is *not* normal to the plane of the orbit, the motion looks like Figure 19.1 (not to scale!), and this gives rise to the seasons.

19.3 THE SNOOKER BALL

Consider a snooker or pool ball, with mass M and radius b, that moves in contact with a horizontal table (see Figure 19.2). First we will consider **general motion** of the ball. In particular, we will allow skidding so that there is no special relation between $\boldsymbol{V}$, the velocity of G, and the angular velocity $\boldsymbol{\omega}$. The table is rough† so that the reaction $\boldsymbol{X}$ exerted by the table on the ball is not necessarily vertical. The governing equations (19.1) and (19.2) become

$$\begin{aligned} M\dot{\boldsymbol{V}} &= \boldsymbol{X} - Mg\,\boldsymbol{k} \\ \dot{\boldsymbol{L}}_G &= (-b\,\boldsymbol{k})\times\boldsymbol{X}, \end{aligned}$$

where $\boldsymbol{L}_G = A\boldsymbol{\omega}$. On eliminating the reaction $\boldsymbol{X}$ we find that

$$\begin{aligned} A\dot{\boldsymbol{\omega}} &= b\left(M\dot{\boldsymbol{V}} + Mg\,\boldsymbol{k}\right)\times\boldsymbol{k} \\ &= Mb\,\dot{\boldsymbol{V}}\times\boldsymbol{k}. \end{aligned}$$

* This is a very good approximation but (like everything else in astronomy) it is not exact. The Earth is slightly spheroidal and the gravitational fields of the Sun and Moon give rise to a small resultant moment about G. As a consequence, $\boldsymbol{\omega}$ actually precesses very slowly around the normal to the plane of the orbit (once every 25,730 years).

† But not so rough that the ball is *compelled* to roll.

On integrating with respect to t, we obtain the non-standard **conservation principle***

$$A\boldsymbol{\omega} + Mb\,\boldsymbol{k}\times\boldsymbol{V} = \boldsymbol{C}, \tag{19.6}$$

where $\boldsymbol{C}$ is a constant vector. This is true in *any* motion of the ball, whether skidding or rolling.

In particular, if we take the scalar product of this equation with $\boldsymbol{k}$, we obtain the scalar conservation principle

$$\boldsymbol{\omega}\cdot\boldsymbol{k} = n, \tag{19.7}$$

where n is a constant. Thus the *vertical component of* $\boldsymbol{\omega}$ *is conserved in* ***any*** *motion of the ball.*

We now examine the special case of **rolling**. In this case the contact particle C has zero velocity, so that the rolling condition is

$$\boldsymbol{V} + \boldsymbol{\omega}\times(-b\,\boldsymbol{k}) = \boldsymbol{0},$$

that is,

$$\boldsymbol{V} + b\,\boldsymbol{k}\times\boldsymbol{\omega} = \boldsymbol{0}. \tag{19.8}$$

Now, from the conservation principle (19.6), it follows that

$$\begin{aligned} A\,\boldsymbol{k}\times\boldsymbol{\omega} &= \boldsymbol{k}\times\boldsymbol{C} - Mb\,\boldsymbol{k}\times(\boldsymbol{k}\times\boldsymbol{V}) \\ &= \boldsymbol{k}\times\boldsymbol{C} - Mb\left((\boldsymbol{k}\cdot\boldsymbol{V})\,\boldsymbol{k} - (\boldsymbol{k}\cdot\boldsymbol{k})\,\boldsymbol{V}\right) \\ &= \boldsymbol{k}\times\boldsymbol{C} + Mb\,\boldsymbol{V}, \end{aligned}$$

since $\boldsymbol{k}$ is a unit vector and is perpendicular to $\boldsymbol{V}$. On substituting this result into the rolling condition (19.8) we obtain

$$\boldsymbol{V} + \left(\frac{Mb^2}{A}\right)\boldsymbol{V} = \left(\frac{b}{A}\right)\boldsymbol{C}\times\boldsymbol{k},$$

which shows that $\boldsymbol{V}$ must be constant in any rolling motion.

The corresponding value of $\boldsymbol{\omega}$ is

$$\boldsymbol{\omega} = \frac{1}{b}\,\boldsymbol{k}\times\boldsymbol{V} + n\,\boldsymbol{k}, \tag{19.9}$$

which is also constant. Hence:

---

* It is actually the angular momentum principle applied about the non-standard point C.

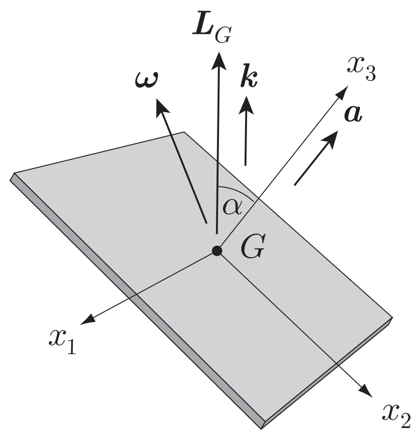


FIGURE 19.3 The free motion of a body with dynamic axial symmetry, depicted here as a square plate. The principal axis Gx_3 is the axis of symmetry and $\boldsymbol{a}$ $(= \boldsymbol{a}(t))$ is the axial unit vector.

The only rolling motions that are possible are straight line motions with constant $\mathbf{V}$ *and constant* $\boldsymbol{\omega}$*; this is despite the presence of the vertical angular velocity component* n.

This fact makes the games of pool and snooker possible. In practice, it is difficult to strike a ball and not give it some 'side' (a non-zero value of n). If rolling took place along curved lines when the ball had 'side', the game would be impossibly difficult. As it is, players often give the ball plenty of 'side' deliberately. The ball still rolls in a straight line, but the 'side' affects the bounce of the ball when it hits a cushion; good players use this fact to control the ball. As a corollary, it follows that, if a player actually wants to *swerve* a ball, then he must make the ball *skid* on the table.

19.4 FREE MOTION OF BODIES WITH AXIAL SYMMETRY

Bodies with dynamical **axial symmetry** are the most important objects whose motion we study. This is for three reasons: (i) They are much more common than spherically symmetric objects, (ii) they perform more interesting motions than spherically symmetric bodies, (iii) their motions can often be calculated in closed form.

Consider a rigid body with dynamical axial symmetry. It could be a baseball bat, a pencil, a top, a spacecraft, a freely-pivoted gyroscope, or the square plate shown in Figure 19.3. The centre of mass G of such a body must lie on the axis of symmetry and we will suppose that the principal moments of inertia of the body at G are $\{A, A, C\}$, with C being the moment of inertia about the *symmetry axis*. C may be greater or less than A. For example, if the body is a short fat cylinder, then $C > A$, but, if the the body is a long thin cylinder, then $C < A$.

Consider such a body moving under either (i) **no forces**, or (ii) **uniform gravity**. Then the motion of G is given by particle mechanics and it only remains to calculate the rotational motion of the body relative to G. Since $\boldsymbol{K}_G = \mathbf{0}$ in each of these cases, the equation of rotational motion is $\dot{\boldsymbol{L}}_G = \mathbf{0}$, where, in the principal axes $Gx_1x_2x_3$,

$$\boldsymbol{L}_G = A\omega_1\,\boldsymbol{e}_1 + A\omega_2\,\boldsymbol{e}_2 + C\omega_3\,\boldsymbol{e}_3. \tag{19.10}$$

It should be remembered that the principal axes $Gx_1x_2x_3$ move ***with the body*** *and that the unit vectors $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ are therefore functions of the time.*

In what follows we will denote the **axial unit vector** $\boldsymbol{e}_3$ by $\boldsymbol{a}$ $(= \boldsymbol{a}(t))$; this is simply to improve readability. We first obtain an expression for the angular velocity $\boldsymbol{\omega}$ in terms of $\boldsymbol{a}$. Consider the point of space that has position vector $\boldsymbol{a}$ relative to G. Since the axial vector $\boldsymbol{a}$ moves with the body, this is the position vector of the particle A of the body that is at unit distance from O along the axis Gx_3. (If this point lies outside the body, then it can simply be included as an extra particle of zero mass.) It follows that $\dot{\boldsymbol{a}}$, the velocity of A, is given by

$$\dot{\boldsymbol{a}} = \boldsymbol{\omega} \times \boldsymbol{a}.$$

If we now take the vector product of this equation with $\boldsymbol{a}$, we obtain

$$\begin{aligned}\boldsymbol{a} \times \dot{\boldsymbol{a}} &= \boldsymbol{a} \times (\boldsymbol{\omega} \times \boldsymbol{a})\\ &= (\boldsymbol{a} \cdot \boldsymbol{a})\,\boldsymbol{\omega} - (\boldsymbol{a} \cdot \boldsymbol{\omega})\,\boldsymbol{a}\\ &= \boldsymbol{\omega} - (\boldsymbol{a} \cdot \boldsymbol{\omega})\,\boldsymbol{a}.\end{aligned}$$

Hence $\boldsymbol{\omega}$ must have the form

$$\boldsymbol{\omega} = \boldsymbol{a} \times \dot{\boldsymbol{a}} + \lambda\,\boldsymbol{a}, \tag{19.11}$$

where λ $(= \boldsymbol{\omega} \cdot \boldsymbol{a})$ is a scalar function of the time.

The second step is to calculate the corresponding angular momentum $\boldsymbol{L}_G$. In the expression (19.11), the term $\lambda\boldsymbol{a}$ lies in the direction of the principal axis Gx_3 and the term $\boldsymbol{a} \times \dot{\boldsymbol{a}}$ is perpendicular to $\boldsymbol{a}$ and is therefore also in a principal direction. It follows from the formula (19.10) that the corresponding angular momentum $\boldsymbol{L}_G$ is given by

$$\boldsymbol{L}_G = A\,\boldsymbol{a} \times \dot{\boldsymbol{a}} + C\lambda\,\boldsymbol{a}. \tag{19.12}$$

The equation for the rotational motion of the body about G is therefore

$$\frac{d}{dt}\left(A\,\boldsymbol{a} \times \dot{\boldsymbol{a}} + C\lambda\,\boldsymbol{a}\right) = \boldsymbol{0}, \tag{19.13}$$

that is,

$$A\,\boldsymbol{a} \times \ddot{\boldsymbol{a}} + C\,(\dot{\lambda}\,\boldsymbol{a} + \lambda\,\dot{\boldsymbol{a}}) = \boldsymbol{0}.$$

If we now take the scalar product of this equation with $\boldsymbol{a}$, we obtain

$$A\,\boldsymbol{a} \cdot (\boldsymbol{a} \times \ddot{\boldsymbol{a}}) + C\,(\dot{\lambda}(\boldsymbol{a} \cdot \boldsymbol{a}) + \lambda(\boldsymbol{a} \cdot \dot{\boldsymbol{a}})) = 0.$$

Now the triple scalar product $\boldsymbol{a} \cdot (\boldsymbol{a} \times \ddot{\boldsymbol{a}})$ has two elements the same and is therefore zero; also, since $\boldsymbol{a}$ is a *unit* vector, $\boldsymbol{a} \cdot \boldsymbol{a} = 1$ and $\boldsymbol{a} \cdot \dot{\boldsymbol{a}} = 0$. It follows that $\dot{\lambda} = 0$, that is $\lambda = n$, a constant. Hence *the axial component of* $\boldsymbol{\omega}$ *is a constant*. We will call this axial component of $\boldsymbol{\omega}$ the **spin** of the body.

The **equation of motion** for the body is therefore

Equation of motion for a free axisymmetric body

$$A\,\boldsymbol{a}\times\dot{\boldsymbol{a}} + C n\,\boldsymbol{a} = \boldsymbol{L}_G \tag{19.14}$$

where the spin n $(= \boldsymbol{\omega}\cdot\boldsymbol{a})$ and the angular momentum $\boldsymbol{L}_G$ are constants determined by the initial conditions.

Surprisingly, this equation has a simple exact solution. First, let us write $\boldsymbol{L}_G = L\,\boldsymbol{k}$, where L is the magnitude of $\boldsymbol{L}_G$ and $\boldsymbol{k}$ is the unit vector in the same direction as $\boldsymbol{L}_G$, as shown in Figure 19.3. Then, on taking the *scalar* product of equation (19.14) with $\boldsymbol{a}$, we obtain

$$A\,\boldsymbol{a}\cdot(\boldsymbol{a}\times\dot{\boldsymbol{a}}) + C n\,\boldsymbol{a}\cdot\boldsymbol{a} = L\,(\boldsymbol{a}\cdot\boldsymbol{k}),$$

which simplifies to give

$$C n = L(\boldsymbol{a}\cdot\boldsymbol{k}) = L\cos\alpha,$$

where α is the angle between $\boldsymbol{a}$ and $\boldsymbol{k}$.* It follows that α is constant and that n, the constant axial component of $\boldsymbol{\omega}$, is given by

$$n = \frac{L\cos\alpha}{C}. \tag{19.15}$$

Thus the *axis of symmetry of the body makes a constant angle with* $\boldsymbol{k}$ *and so sweeps out a cone around the axis* $\{G, \boldsymbol{k}\}$.

The progress of the axis of symmetry in time can be found by taking the *vector* product of equation (19.14) with $\boldsymbol{a}$. This gives

$$A\,\boldsymbol{a}\times(\boldsymbol{a}\times\dot{\boldsymbol{a}}) + C n\,\boldsymbol{a}\times\boldsymbol{a} = L\,(\boldsymbol{a}\times\boldsymbol{k}),$$

that is,

$$A\left((\boldsymbol{a}\cdot\dot{\boldsymbol{a}})\,\boldsymbol{a} - (\boldsymbol{a}\cdot\boldsymbol{a})\,\dot{\boldsymbol{a}}\right) + \boldsymbol{0} = L\,(\boldsymbol{a}\times\boldsymbol{k}).$$

Since $\boldsymbol{a}$ is a *unit* vector, $\boldsymbol{a}\cdot\boldsymbol{a} = 1$ and $\boldsymbol{a}\cdot\dot{\boldsymbol{a}} = 0$ and we obtain

$$\dot{\boldsymbol{a}} = \left(\frac{L}{A}\,\boldsymbol{k}\right)\times\boldsymbol{a}. \tag{19.16}$$

This equation shows that the *axis of symmetry of the body precesses around the axis* $\{G, \boldsymbol{k}\}$ *with constant angular speed* L/A.

* We will suppose that n is positive so that α is an acute angle.

Motion viewed from the precessing frame

It is instructive to view this motion from a rotating frame in which the axis of symmetry is at rest. The *true* angular velocity of the body is given by

$$\begin{aligned}\boldsymbol{\omega} &= \boldsymbol{a}\times\dot{\boldsymbol{a}} + n\,\boldsymbol{a}\\ &= \boldsymbol{a}\times\left(\frac{L}{A}(\boldsymbol{k}\times\boldsymbol{a})\right) + \left(\frac{L\cos\alpha}{C}\right)\boldsymbol{a}\\ &= \frac{L}{A}\,\boldsymbol{k} + L\cos\alpha\left(\frac{A-C}{AC}\right)\boldsymbol{a}.\end{aligned} \tag{19.17}$$

Suppose we now view the motion from a frame with origin at G that is rotating about the axis $\{G, \boldsymbol{k}\}$ with angular speed $(L/A)\boldsymbol{k}$. In this **precessing frame**, the axial vector $\boldsymbol{a}$ is at rest. Also, Theorem 17.1 (on the addition of angular velocities) tells us that, in the precessing frame, the *apparent* angular velocity of the body is

$$\boldsymbol{\omega}' = L\cos\alpha\left(\frac{A-C}{AC}\right)\boldsymbol{a},$$

that is, the body apparently rotates about its fixed axis of symmetry with angular velocity $L\cos\alpha(A-C)/AC$. The *true* motion is thus composed of (i) this constant axial rotation, and (ii) a precession around the axis $\{G, \boldsymbol{k}\}$ with constant angular speed L/A.

Our results are summarised as follows:

Free motion of an axisymmetric body

- The axis of symmetry of the body makes a constant angle α with the angular momentum vector $L\,\boldsymbol{k}$ and precesses around the axis $\{G, \boldsymbol{k}\}$ (in the positive sense) with constant angular speed L/A.
- In this motion, the axial spin n of the body has the constant value $L\cos\alpha/C$.
- In the precessing frame, the body apparently rotates about its fixed axis of symmetry with constant angular velocity $L\cos\alpha(A-C)/AC$.

The motion described above is very familiar, at least qualitatively. The precession of the symmetry axis looks like a 'wobble' superimposed on the spinning motion of the body about the symmetry axis.

Example 19.1 *Wobble on a frisbee*

A frisbee is spinning with constant angular speed Ω about its axis of symmetry when its motion is slightly disturbed. What is the angular frequency of the resulting wobble?

Solution

Suppose that the new angular momentum of the frisbee is $L\boldsymbol{k}$ with the axis of the frisbee making a small angle α with the fixed axis $\{G, \boldsymbol{k}\}$. Then the axis of the frisbee

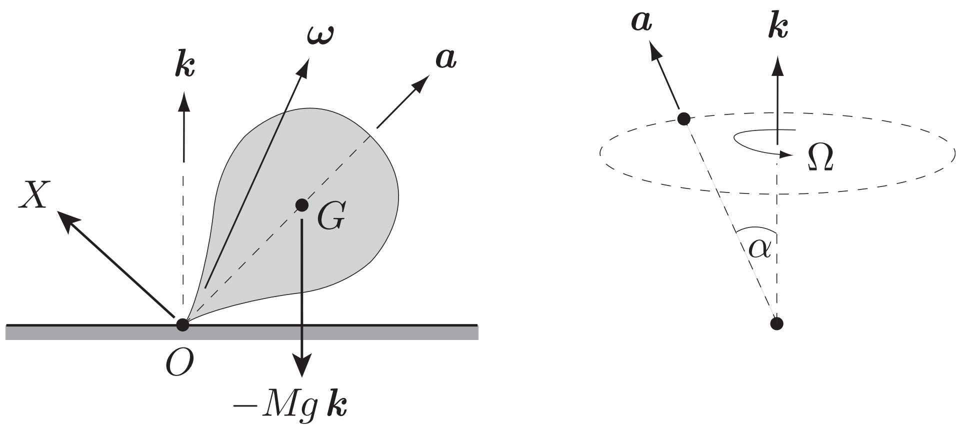


FIGURE 19.4 A symmetrical top with its vertex fixed at O. The top moves under uniform gravity and the reaction force $\boldsymbol{X}$ of the table.

precesses around the axis $\{G, \boldsymbol{k}\}$ with angular speed L/A. *It is this precession that is observed as a wobble*. The angular momentum magnitude L is not known exactly, but it differs only slightly from that in the undisturbed state, namely $C\Omega$. Hence the precession rate is $C\Omega/A$ approximately, and this is the angular frequency of the wobble. If we regard the frisbee as a circular disk, then $C = 2A$ and the **angular frequency of the wobble** is 2Ω, approximately twice the spin of the frisbee about its axis. ■

The above results can also be obtained by using Lagrangian mechanics. The method is described in Section 19.6.

19.5 THE SPINNING TOP

Consider the symmetrical spinning top* shown in Figure 19.4 (left), which has its vertex fixed at the origin O. One may imagine that the vertex of the top is lodged in a small smooth pit in the table from which it cannot escape. The top moves under uniform gravity and the unknown reaction force $\boldsymbol{X}$ exerted by the table. Since the top has a particle fixed at O, we take the equation of motion in the form

$$\dot{\boldsymbol{L}}_O = \boldsymbol{K}_O,$$

where $\boldsymbol{K}_O$ is the moment of the external forces about O. In the present case,

$$\begin{aligned}\boldsymbol{K}_O &= \boldsymbol{0}\times\boldsymbol{X} + (h\boldsymbol{a})\times(-Mg\,\boldsymbol{k})\\ &= -Mgh\,\boldsymbol{a}\times\boldsymbol{k},\end{aligned}$$

* A 'top' should be understood to mean *any* rigid body with dynamical axial symmetry that is pivoted at a point on its symmetry axis, and is moving under uniform gravity.

where M is the mass of the top and h is the distance OG.

The angular momentum $\boldsymbol{L}_O$ is calculated exactly in the same way as in the last section. The angular velocity $\boldsymbol{\omega}$ can be expressed in the form

$$\boldsymbol{\omega} = \boldsymbol{a}\times\dot{\boldsymbol{a}} + \lambda\,\boldsymbol{a},$$

and $\boldsymbol{L}_O$ is then given by

$$\boldsymbol{L}_O = A\,\boldsymbol{a}\times\dot{\boldsymbol{a}} + C\lambda\,\boldsymbol{a},$$

where $\{A, A, C\}$ are the principal moments of inertia of the top *at the vertex* O, with C being the moment of inertia of the top about its axis of symmetry. The equation of motion for the top is therefore

$$\frac{d}{dt}\left(A\,\boldsymbol{a}\times\dot{\boldsymbol{a}} + C\lambda\,\boldsymbol{a}\right) = -Mgh\,\boldsymbol{a}\times\boldsymbol{k},$$

that is

$$A\,\boldsymbol{a}\times\ddot{\boldsymbol{a}} + C(\dot{\lambda}\boldsymbol{a} + \lambda\dot{\boldsymbol{a}}) = -Mgh\,\boldsymbol{a}\times\boldsymbol{k}.$$

If we now take the scalar product of this equation with $\boldsymbol{a}$, we obtain

$$A\,\boldsymbol{a}\cdot(\boldsymbol{a}\times\ddot{\boldsymbol{a}}) + C(\dot{\lambda}(\boldsymbol{a}\cdot\boldsymbol{a}) + \lambda(\boldsymbol{a}\cdot\dot{\boldsymbol{a}})) = -Mgh\,\boldsymbol{a}\cdot(\boldsymbol{a}\times\boldsymbol{k}).$$

The triple scalar products are both zero since each has two elements the same. Also, since $\boldsymbol{a}$ is a *unit* vector, $\boldsymbol{a}\cdot\boldsymbol{a} = 1$ and $\boldsymbol{a}\cdot\dot{\boldsymbol{a}} = 0$. It follows that $\dot{\lambda} = 0$, so that $\lambda = n$, a constant. Hence *the axial component of* $\boldsymbol{\omega}$ *is a constant*. We call this axial component of $\boldsymbol{\omega}$ the **spin** of the top. The **equation of motion** for the top can therefore be written

Equation of motion for the top

$$\frac{d}{dt}\left(A\,\boldsymbol{a}\times\dot{\boldsymbol{a}} + Cn\,\boldsymbol{a}\right) = -Mgh\,\boldsymbol{a}\times\boldsymbol{k} \qquad (19.18)$$

where the spin n $(= \boldsymbol{\omega}\cdot\boldsymbol{a})$ is a constant determined by the initial conditions.

Steady precession

We will calculate the *general* motion of the top in the next section by using the Lagrangian method. Here we will investigate only the special, but important, motion called **steady precession**. In this motion, the axial vector $\boldsymbol{a}$ maintains a constant angle α with the vertical and precesses round the axis $\{O, \boldsymbol{k}\}$ with constant angular speed Ω, as shown in Figure 19.4 (right). In this steady precession the rate of change of $\boldsymbol{a}$ is given by

$$\dot{\boldsymbol{a}} = (\Omega\boldsymbol{k})\times\boldsymbol{a}. \qquad (19.19)$$

We will now investigate whether steady precession of the top can actually occur. From equation (19.16), it follows that, in steady precession,

$$\begin{aligned} \boldsymbol{a}\times\dot{\boldsymbol{a}} &= \boldsymbol{a}\times(\Omega\,\boldsymbol{k}\times\boldsymbol{a}) \\ &= \Omega(\boldsymbol{a}\cdot\boldsymbol{a})\,\boldsymbol{k} - \Omega(\boldsymbol{a}\cdot\boldsymbol{k})\,\boldsymbol{a} \\ &= \Omega\,(\boldsymbol{k} - \cos\alpha\,\boldsymbol{a})\,. \end{aligned}$$

Hence

$$\begin{aligned} \frac{d}{dt}(\boldsymbol{a}\times\dot{\boldsymbol{a}}) &= \Omega\,(\boldsymbol{0} - \cos\alpha\,\dot{\boldsymbol{a}}) \\ &= -\Omega\cos\alpha\,(\Omega\boldsymbol{k})\times\boldsymbol{a} \\ &= \Omega^2\cos\alpha\,(\boldsymbol{a}\times\boldsymbol{k}). \end{aligned} \tag{19.20}$$

On substituting equations (19.20) and (19.19) into the equation of motion (19.18), we obtain

$$\left(A\cos\alpha\,\Omega^2 - Cn\,\Omega + Mgh\right)(\boldsymbol{a}\times\boldsymbol{k}) = \boldsymbol{0}.$$

This is the condition for steady precssion of the top. Since $\boldsymbol{a}\times\boldsymbol{k} \neq \boldsymbol{0}$, steady precession can only occur if there are values of α and Ω that satisfy the equation

$$A\cos\alpha\,\Omega^2 - Cn\,\Omega + Mgh = 0. \tag{19.21}$$

This quadratic equation will have *real roots* for the angular rate Ω if

$$C^2n^2 \geq 4AMgh\cos\alpha. \tag{19.22}$$

If the angle α is obtuse, then this condition always holds, even when $n = 0$. However, the case in which the top executes a conical type of motion *below* the pivot O is not interesting. The interesting case is that in which α is an *acute* angle and the top precesses in an 'upright' position. Our result is as follows:

The top can undergo steady precession at an acute angle α to the upward vertical if n (the axial component of $\boldsymbol{\omega}$) is large enough to satisfy the condition (19.22). If this condition is satisfied there will generally be two different values of the precession rate for each choice of the angle α.

Fast and slow precession

The two solutions of equation (19.21) for the precession rate Ω are

$$\begin{aligned} \Omega^{F,S} &= \frac{Cn \pm \left(C^2n^2 - 4AMgh\cos\alpha\right)^{1/2}}{2A\cos\alpha} \\ &= \frac{Cn}{2A\cos\alpha}\left[1 \pm \left(1 - \frac{4AMgh\cos\alpha}{C^2n^2}\right)^{1/2}\right], \end{aligned} \tag{19.23}$$

where the fast (F) precession rate corresponds to the 'plus' choice and the slow precession rate (S) corresponds to the 'minus' choice. In practical circumstances the value of the spin n is often such that the dimensionless ratio

$$\frac{4AMgh}{C^2n^2}$$

is *small compared to unity*. In this case, Ω^F and Ω^S are given approximately by

$$\Omega^F \approx \frac{Cn}{A\cos\alpha}, \qquad \Omega^S \approx \frac{Mgh}{Cn}. \tag{19.24}$$

The **fast precession** rate is approximately *directly proportional* to n and is independent of the gravitational acceleration! In fact, it approximates the *force-free precession* found in the last section. This motion is almost impossible to observe in a real top. It has a precession rate of similar magnitude to n which would make the vertex of the top extremely difficult to secure. The fast precession may however explain the trembling motion sometimes seen when a top is spinning slowly with its axis almost vertical.

In contrast, the **slow precession** rate is approximately *inversely proportional to n and independent of the angle* α. This is the motion commonly observed; the faster the spin of the top, the slower the rate of precession.

Example 19.2 *A simple top*

A top is made by sticking a light pin of length 3 cm through the centre of a uniform circular disk of mass M and radius 8 cm. Find the rates of slow and fast precession of the top when it is given a spin of 10 revolutions per second and α is small. [Take $g = 10\ \mathrm{m\,s^{-2}}$.]

Solution

For a uniform disk of mass M and radius a, the principal moments of inertia at its centre of mass are $\{\frac{1}{4}Ma^2, \frac{1}{4}Ma^2, \frac{1}{2}Ma^2\}$ and so , by the theorem of parallel axes, the principal moments at the point of the pin are $A = B = \frac{1}{4}M(a^2 + 4h^2)$, $C = \frac{1}{2}Ma^2$, where h is the length of the pin. Hence, on putting in the given dimensions, $A = B = 25M/10000$ and $C = 32M/10000$, where M is the mass of the disk. We are also given that $n = 20\pi$ radians per second and that α is small. The quadratic equation for the precession rate Ω is therefore

$$\frac{25M}{10000}\Omega^2 - \frac{32M}{10000}(40\pi)\Omega + \frac{30M}{100} = 0.$$

The roots of this equation are $\Omega^S = 1.52$ and $\Omega^F = 78.9$ radians per second. The approximate formulae in equation (19.24) give 1.49 and 80.4 respectively. [In this case the dimensionless ratio $4AMgh/(C^2n^2) \approx 0.074$ so a reasonable approximation could be expected.] ■

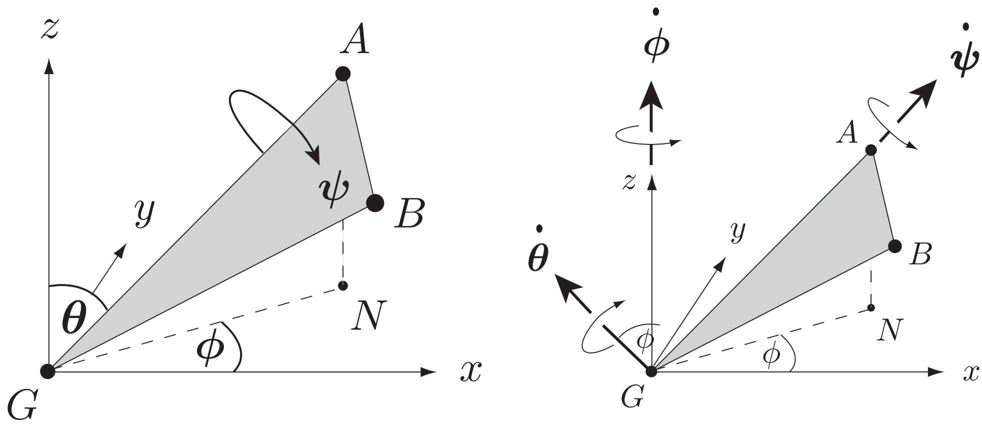


FIGURE 19.5 **Left:** The Euler angles θ, ϕ and ψ. **Right:** The corresponding 'velocity' diagram.

19.6 LAGRANGIAN DYNAMICS OF THE TOP

The Euler angles

In order to solve problems in rigid body dynamics by Lagrangian mechanics, we need a set of **generalised coordinates** for a generally moving rigid body. The position of the centre of mass G can be specified by its three Cartesian coordinates, but the position of the body *relative* to G still needs to be specified. The best practical coordinates for this purpose are the **Euler angles*** θ, ϕ and ψ shown in Figure 19.5 (left). In order to fix the position of the body, it is sufficient to specify the positions of *two more* of its particles, A and B say, chosen so that G, A, B do not lie in a straight line. The position of A is fixed by the 'polar angles' θ and ϕ, measured relative to the fixed Cartesian coordinate system $Gxyz$. This does not yet fix the position of B since the rigid triangle GAB can still rotate around GA. The position of B becomes becomes fixed when we specifiy the angle ψ through which the triangle GAB has been rotated from some reference position.† Since the angles θ, ϕ, ψ are clearly independent variables, it follows that *the **Euler angles** θ, ϕ, ψ are a set of generalised coordinates for the **rotational** motion of a rigid body.*

The Euler angles are a particularly appropriate set of coordinates when the body has **axial symmetry**. In this case it is usual to take the A to be a particle *lying on the symmetry axis*. Then θ and ϕ determine the orientation of the symmetry axis and ψ is the rotation angle of the body around the symmetry axis.

We may now construct the **velocity diagram** corresponding to this choice of coordinates. When the coordinates θ, ϕ, ψ increase individually, each gives the body an *angular* velocity as shown in Figure 19.5 (right). The **angular velocity** of the body in general motion is then obtained by adding these three angular velocities vectorially.

* After Leonhard Euler (1707–1783), the greatest Swiss mathematician, and the most prolific of all time.

† The triangle GAB is in its standard reference position when it lies in the plane AGz with B 'below' the line GA. The configuraton shown in Figure 19.5 thus has a ψ value of nearly 2π.

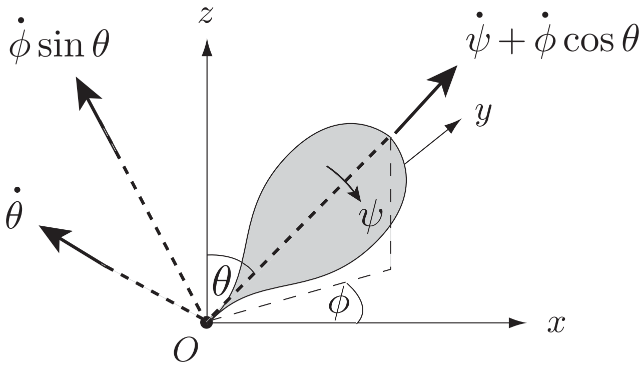


FIGURE 19.6 The angular velocity of the top resolved along three perpendicular principal directions.

Lagrangian for the top

For a top with its vertex fixed at the origin O, we take a fixed system of Cartesian coordinates $Oxyz$ with Oz vertically upwards and define the Euler angles as shown in Figure 19.5 (left) with G replaced by O. Since the top has axial symmetry we also take A to lie on the symmetry axis, which is a principal axis of the top at O. The top is thus a holonomic system and can be solved by Lagrangian method. To find the kinetic energy of the top in terms of the Euler angles, the angular velocities in Figure 19.5 need to be resolved along three perpendicular principal axes, as shown in Figure 19.6.* The **kinetic energy** of the top is therefore given by

$$T = \tfrac{1}{2}A\dot{\theta}^2 + \tfrac{1}{2}A\left(\dot{\phi}\sin\theta\right)^2 + \tfrac{1}{2}C\left(\dot{\psi}+\dot{\phi}\cos\theta\right)^2, \tag{19.25}$$

where $\{A, A, C\}$ are the principal moments of inertia of the top at O. The gravitational **potential energy** of the top is simply

$$V = MgZ = Mgh\cos\theta,$$

where h is the distance OG. Hence the **Lagrangian** of the top, with the Euler angles as coordinates, is:

Lagrangian for the top

$$L = \tfrac{1}{2}A\dot{\theta}^2 + \tfrac{1}{2}A\left(\dot{\phi}\sin\theta\right)^2 + \tfrac{1}{2}C\left(\dot{\psi}+\dot{\phi}\cos\theta\right)^2 - Mgh\cos\theta \tag{19.26}$$

Conserved quantities

Since ϕ and ψ appear in the Lagrangian only as $\dot{\phi}$ and $\dot{\psi}$, they are both **cyclic coordinates**. It follows that the generalised momenta p_ϕ and p_ψ are *constants of the motion*. On

* These principal axes are not 'embedded' in the top since they do not rotate as ψ increases.

evaluating p_ϕ ($= \partial L/\partial\dot\phi$) and p_ψ ($= \partial L/\partial\dot\psi$), these conservation principles become

$$A\dot\phi \sin^2\theta + C(\dot\psi + \dot\phi\cos\theta)\cos\theta = L_z, \tag{19.27}$$

$$C(\dot\psi + \dot\phi\cos\theta) = Cn, \tag{19.28}$$

where n and L_z are constants, determined by the initial conditions. In fact, *L_z is the z-component of the* ***angular momentum*** $\boldsymbol{L}_O$*, and n is the axial component of the angular velocity* $\boldsymbol{\omega}$, the quantity known as the **spin** of the top.* Without losing generality, we will assume that n is *positive*.

These two conservation principles, together with the **energy conservation** equation

$$\tfrac{1}{2}A\dot\theta^2 + \tfrac{1}{2}A(\dot\phi\sin\theta)^2 + \tfrac{1}{2}C(\dot\psi + \dot\phi\cos\theta)^2 + Mgh\cos\theta = E, \tag{19.29}$$

are sufficient to determine the motion of the top. The energy conservation equation is preferable to the Lagrange equation for the coordinate θ because it contains only *first* derivatives of the Euler angles.

The equation for the inclination angle θ

By making use of the conservation equations (19.27) and (19.28) in the energy equation (19.29), we can eliminate $\dot\phi$ and $\dot\psi$ to give

$$A\dot\theta^2 = 2E - Cn^2 - \frac{(L_z - Cn\cos\theta)^2}{A\sin^2\theta} - 2Mgh\cos\theta, \tag{19.30}$$

where the constants n, L_z, E are determined by the initial conditions. This ODE for the unknown function $\theta(t)$ has the form

$$\dot\theta^2 = \frac{F(\cos\theta)}{A^2\sin^2\theta},$$

where the function $F(u)$ is defined by

$$F(u) = A(2E - Cn^2)(1-u^2) - (L_z - Cnu)^2 - 2AMghu(1-u^2).$$

This equation is similar to the corresponding equation in the theory of the spherical pendulum (see Chapter 11) and we can make similar deductions about the behaviour of the coordinate θ. Because the left side of (19.30) is positive, it follows that the motion is restricted to those values of θ that make the function $F(\cos\theta)$ positive. Moreover, the maximium and minimum values of θ can only occur when $F(\theta) = 0$.

The equation $F(u) = 0$ is a cubic in the variable u and the positions of its roots can be deduced by considering the sign of F. We have the following information:

* Note that $n = \dot\psi + \dot\phi\cos\theta$, not $\dot\psi$.

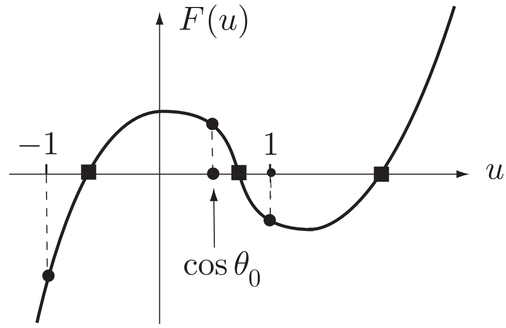


FIGURE 19.7 The form of the cubic function $F(u)$. The black squares are at its roots.

(i) As u tends to positive (or negative) infinity, $F(u)$ tends to positive (or negative) infinity.

(ii) $F(1) = -(L_z - Cn)^2$ and $F(-1) = -(L_z + Cn)^2$ which are both *negative*.†

(iii) Given that θ_0, the initial value of θ, lies in the range $0 < \theta_0 < \pi$ and that $\dot{\theta}$ is not zero initially, then $F(\cos\theta_0)$ must be *positive*.

The graph of $F(u)$ must therefore have the form shown in Figure 19.7. It follows that the equation $F(u) = 0$ must have a root $u_{\min}$ lying in $-1 < u < \cos\theta_0$, a root $u_{\max}$ lying in $\cos\theta_0 < u < 1$, and a third root lying in $u > 1$, as indicated by the black squares in Figure 19.7; this accounts for all three roots. The root greater than unity is not physically admissible since it lies outside the range of u when $u = \cos\theta$. It follows that, in the motion, u must oscillate in the range $u_{\min} \le u \le u_{\max}$. Hence, as in the case of the spherical pendulum, the inclination angle θ must perform **periodic oscillations** between two extreme values α and β. The difference here is that it is possible for both α and β to be *acute* angles (measured from the upward vertical); in other words, the top can stand 'upright'. This oscillation of the inclination angle θ is called **nutation** of the top.

Example 19.3 *Finding the range of the inclination angle*

Suppose that the top is released with its axis at rest and making an angle of $\pi/3$ with the upward vertical. Find the function $F(u)$. For the case in which $C^2n^2 = 4AMgh$, find the range of values of θ that occur in the subsequent motion.

Solution

On using the initial conditions $\theta = \pi/3$, $\dot{\theta} = 0$, $\dot{\phi} = 0$ when $t = 0$, we find that $L_z = \frac{1}{2}Cn$ and $E = \frac{1}{2}Cn^2 + \frac{1}{2}Mgh$. The function $F(u)$ therefore reduces to

$$F(u) = \tfrac{1}{4}(1-2u)\Big[4AMgh(1-u^2) - C^2n^2(1-2u)\Big],$$

and when $C^2n^2 = 4AMgh$, this becomes

$$F(u) = AMgh\,u(1-2u)(2-u).$$

† $F(\pm 1)$ is zero when $L_z = \pm Cn$ respectively. This corresponds to a special (and not very important) class of motions in which the axis of the top passes through the upward or downward vertical. We exclude these cases from consideration.

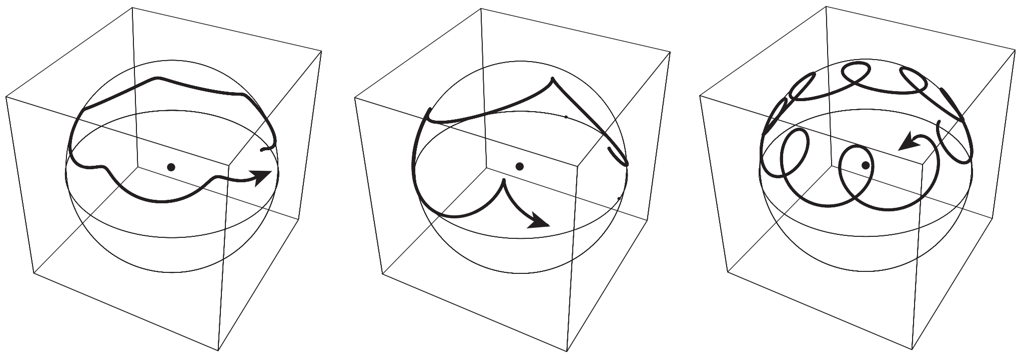


FIGURE 19.8 Three different forms for the motion of the axis of symmetry of the top. The paths show the movement of the point in which the axis of symmetry intersects a fixed sphere with centre O.

The physical roots of the equation $F(u) = 0$ are $u = 0$ and $u = \frac{1}{2}$ and so θ oscillates in the range $\pi/3 \le \theta \le \pi/2$. This is the motion shown in Figure 19.8 (centre). ■

Question ***Finding the period of*** θ

Find the period of the angle θ in the same motion.

Answer

With the same set of initial conditions (and with $C^2n^2 = 4AMgh$), the ODE for θ reduces to

$$\dot{\theta}^2 = \frac{Mgh\cos\theta(1 - 2\cos\theta)(2 - \cos\theta)}{A\sin^2\theta}.$$

On taking square roots and integrating, the period τ of the function $\theta(t)$ is found to be

$$\tau = 2\left(\frac{A}{Mgh}\right)^{1/2} \int_{\pi/3}^{\pi/2} \frac{\sin\theta\, d\theta}{[\cos\theta(1 - 2\cos\theta)(2 - \cos\theta)]^{1/2}} \approx 3.37\left(\frac{A}{Mgh}\right)^{1/2}. ■$$

Precession of the axis

Once the function $\theta(t)$ has been determined (at least in principle), the precession angle $\phi(t)$ can be be found from the angular momentum conservation equation

$$A\dot{\phi}\sin^2\theta + Cn\cos\theta = L_z. \tag{19.31}$$

Note that the precession rate $\dot{\phi}$ is *not constant* when the top is in general motion.

With the initial conditions $\theta = \alpha$, $\dot{\theta} = 0$, $\phi = 0$, $\dot{\phi} = \Omega$, the value of the constant L_z is $A\Omega\sin^2\alpha + Cn\cos\alpha$, and equation (19.31) becomes

$$A\dot{\phi}\sin^2\theta = A\Omega\sin^2\alpha + Cn(\cos\alpha - \cos\theta). \tag{19.32}$$

Suppose that $C^2n^2 \geq 4AMgh$ so that steady precession can exist at any inclination. Then if the initial precession rate Ω is set equal to Ω^S (the rate of slow steady precession at inclination α), the resulting motion of the top is (unsurprisingly) slow steady precession. However, if $\Omega < \Omega^S$, it can be shown (by differentiating equation (19.30) with respect to t) that the initial value of $\ddot{\theta}$ is positive so that θ initially increases and α is the *minimum* value taken by θ. (In other words, the axis of the top *falls* initially.) In this case, we can see from equation (19.32) that

$$A\Omega \sin^2 \alpha \leq A\dot{\phi} \sin^2 \theta \leq A\Omega \sin^2 \alpha + Cn(\cos \alpha - \cos \beta),$$

where β is the *maximum* value taken by θ. It follows that

(i) If $\Omega > 0$, then the precession rate $\dot{\phi}$ is always *positive*. This is the case shown in Figure 19.8 (left). The critical case in which $\Omega = 0$ (when the axis is released from rest) is shown in Figure 19.8 (centre).

(ii) There is no reason why the top axis cannot be projected the 'wrong' way. If $\Omega < 0$ (but is not so negative that $A\Omega \sin^2 \alpha + Cn(\cos \alpha - \cos \beta) < 0$), then $\dot{\phi}$ is *sometimes positive and sometimes negative*! In this case, the path of the axis crosses itself, as shown in Figure 19.8 (right).

(iii) If the value of Ω is so negative that $A\Omega \sin^2 \alpha + Cn(\cos \alpha - \cos \beta) < 0$, then $\dot{\phi}$ is always *negative*. However, this requires a 'fast' precession rate that is unlikely to be observed.

In a similar way, one can show that when $\Omega^S < \Omega < \Omega^F$, the axis of the top *rises* initially and that $\dot{\phi}$ is always positive. The motion of the axis resembles that shown in Figure 19.8 (left).

Example 19.4 *The precession during a period of* θ

In the last example, find the angle through which the top precesses during one period of the inclination angle θ.

Solution

In the last example, the axis of the top is released from rest with $\theta = \pi/3$, and with $C^2n^2 = 4AMgh$. With $\alpha = \pi/3$ and $\Omega = 0$, the equation for the precession rate $\dot{\phi}$ becomes

$$\dot{\phi} = \frac{Cn(1 - 2\cos\theta)}{2A \sin^2 \theta}.$$

It follows that Φ, the increase in ϕ during one period of the coordinate θ is given by

$$\begin{aligned}\Phi = \phi(\tau) - \phi(0) &= \int_0^\tau \dot{\phi}\, dt \\ &= \frac{Cn}{2A} \int_0^\tau \frac{1 - 2\cos\theta}{\sin^2\theta}\, dt = \left(\frac{Mgh}{A}\right)^{1/2} \times 2 \int_{\pi/3}^{\pi/2} \left(\frac{1 - 2\cos\theta}{\sin^2\theta}\right) \frac{d\theta}{\dot{\theta}}.\end{aligned}$$

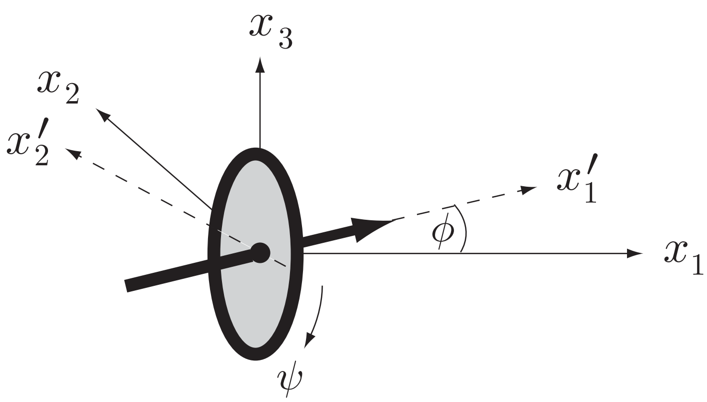


FIGURE 19.9 The gyrocompass is a gyroscope whose symmetry axis is constrained to remain in a horizontal plane.

On using the expression for $\dot{\theta}$ from the last example, we obtain

$$\Phi = 2\int_{\pi/3}^{\pi/2} \left[\frac{1-2\cos\theta}{\cos\theta(2-\cos\theta)}\right]^{1/2} \frac{d\theta}{\sin\theta} \approx 1.69,$$

which is about 97°. This motion is shown in Figure 19.8 (centre). ■

19.7 THE GYROCOMPASS

The gyrocompass is the most important navigational instrument used in the world today. Invented by Elmer Sperry* in 1908, it consists essentially of a gyroscope which is smoothly pivoted so that its centre of mass G remains fixed and its *axis of symmetry remains in the* ***horizontal*** *plane through* G; the gyroscope is however free to rotate about its axis of symmetry and about the vertical through G (see Figure 19.9). The constraint that the axis of the gyro should remain horizontal is vital to its operation.

Non-rotating Earth

Consider first the case in which the Earth has **no rotation**. Let $Gx_1x_2x_3$ be a frame fixed to the Earth, with the axis Gx_3 pointing vertically upwards and the axis of the gyro moving in the horizontal (x_1, x_2)-plane. Let $Gx_1'x_2'x_3$ be a second set of axes with Gx_1' pointing along the axis of symmetry of the gyro; these axes therefore *rotate with the gyro*. Since the Euler angle θ is constrained to be $\pi/2$ in this problem, the system has two degrees of freedom and the Euler angles ϕ and ψ (see Figure 19.9) form a set of generalised coordinates. In terms of these angles, the **angular velocity** of the gyroscope is

$$\boldsymbol{\omega} = \dot{\psi}\,\boldsymbol{e}_1' + \dot{\phi}\,\boldsymbol{e}_3, \tag{19.33}$$

* Elmer Ambrose Sperry (1860–1930), American inventor and entrepreneur.

where $\{\boldsymbol{e}_1', \boldsymbol{e}_2', \boldsymbol{e}_3\}$ are the unit vectors of the coordinate system $Gx_1'x_2'x_3$; note that $\boldsymbol{e}_1'$ and $\boldsymbol{e}_2'$ are functions of the time. The corresponding **angular momentum** is therefore

$$\boldsymbol{L}_G = C\dot{\psi}\,\boldsymbol{e}_1' + A\dot{\phi}\,\boldsymbol{e}_3,$$

where $\{C, A, A\}$ are principal moments of inertia of the gyro at G. Now consider the **applied moment** $\boldsymbol{K}_G$. Since the gyro is free to rotate about the axes Gx_1' and Gx_3, it follows that the applied moment about these axes is zero. However there is a moment* applied about the axis Gx_2' which enforces the constraint that the axis of the gyro should remain in the (x_1, x_2)-plane. Hence the **equation of motion** of the gyro is

$$\frac{d}{dt}\left(C\dot{\psi}\,\boldsymbol{e}_1' + A\dot{\phi}\,\boldsymbol{e}_3\right) = K\,\boldsymbol{e}_2', \qquad (19.34)$$

where K is the unknown 'moment of constraint'. Now

$$\frac{d}{dt}\left(C\dot{\psi}\,\boldsymbol{e}_1' + A\dot{\phi}\,\boldsymbol{e}_3\right) = C\ddot{\psi}\,\boldsymbol{e}_1' + C\dot{\psi}\,\frac{d\boldsymbol{e}_1'}{dt} + A\ddot{\phi}\,\boldsymbol{e}_3,$$

and

$$\frac{d\boldsymbol{e}_1'}{dt} = \frac{d\boldsymbol{e}_1'}{d\phi} \times \frac{d\phi}{dt} = \dot{\phi}\,\boldsymbol{e}_2'.$$

The last step follows since the unit vectors $\boldsymbol{e}_1'$, $\boldsymbol{e}_2'$ are analogous to the polar unit vectors $\widehat{\boldsymbol{r}}, \widehat{\boldsymbol{\theta}}$ and hence satisfy the relations

$$\frac{d\boldsymbol{e}_1'}{d\phi} = \boldsymbol{e}_2', \qquad \frac{d\boldsymbol{e}_2'}{d\phi} = -\boldsymbol{e}_1'. \qquad (19.35)$$

On combining these results, we obtain the three component equations

$$\ddot{\psi} = 0, \qquad C\dot{\phi}\dot{\psi} = K, \qquad \ddot{\phi} = 0.$$

Hence the **motion** of the gyro has the form

$$\dot{\psi} = n, \qquad \dot{\phi} = \Omega,$$

where n and Ω are constants determined by the initial conditions. Thus the *gyro has constant spin n and precesses at a constant rate* Ω. The 'moment of constraint' that keeps the gyro axis horizontal is $K = Cn\Omega$. This is the complete solution for the gyro on a non-rotating Earth.

Rotating Earth

A gyro that simply precesses at a constant rate does not seem much like a direction finding device, but everything changes when we introduce the effect of the **Earth's rotation**. Suppose now that the axes $Gx_1x_2x_3$ are fixed at some location on the Earth with Gx_3 in the direction of the *local* vertical; the axes are oriented so that Gx_1 points north and Gx_2 points west. Then $Gx_1x_2x_3$ is no longer an inertial frame and our equations of motion must be modified.

* This moment is applied by the gimbal ring that holds the gyro.

The angular velocity of the gyro *relative to the frame* $Gx_1x_2x_3$ is still given by equation (19.33) but now the frame $Gx_1x_2x_3$ is not inertial but rotates with the Earth's angular velocity $\boldsymbol{\Omega}^E$. This can be written in the form

$$\begin{aligned}\boldsymbol{\Omega}^E &= \Omega^E \sin\alpha\, \boldsymbol{e}_1 + \Omega^E \cos\alpha\, \boldsymbol{e}_3\\ &= \Omega^E \sin\alpha \cos\phi\, \boldsymbol{e}_1' - \Omega^E \sin\alpha \sin\phi\, \boldsymbol{e}_2' + \Omega^E \cos\alpha\, \boldsymbol{e}_3 \end{aligned} \tag{19.36}$$

where the angle α is the co-latitude of the location where the gyro is situated. Hence, by the addition theorem for angular velocities, the true **angular velocity** of the gyro (relative to an *inertial* frame) is

$$\begin{aligned}\boldsymbol{\omega} &= (\dot{\psi}\, \boldsymbol{e}_1' + \dot{\phi}\, \boldsymbol{e}_3) + \boldsymbol{\Omega}^E\\ &= \left(\dot{\psi} + \Omega^E \sin\alpha\cos\phi\right) \boldsymbol{e}_1' - \left(\Omega^E \sin\alpha \sin\phi\right) \boldsymbol{e}_2' + \left(\dot{\phi} + \Omega^E \cos\alpha\right) \boldsymbol{e}_3.\end{aligned}$$

It follows that the true **angular momentum** of the gyro is given by

$$\boldsymbol{L}_G = C\left(\dot{\psi} + \Omega^E \sin\alpha\cos\phi\right) \boldsymbol{e}_1' - A\left(\Omega^E \sin\alpha \sin\phi\right) \boldsymbol{e}_2' + A\left(\dot{\phi} + \Omega^E \cos\alpha\right) \boldsymbol{e}_3. \tag{19.37}$$

For the same reasons as before, the **applied moment** has the form $\boldsymbol{K}_G = K\,\boldsymbol{e}_2'$, where K is the unknown moment of constraint. The **equation of motion** of the gyro is therefore

$$\frac{d\boldsymbol{L}_G}{dt} = K\,\boldsymbol{e}_2', \tag{19.38}$$

where $\boldsymbol{L}_G$ is given by equation (19.37). Now

$$\begin{aligned}\frac{d\boldsymbol{L}_G}{dt} &= C\frac{d}{dt}\left(\dot{\psi} + \Omega^E \sin\alpha\cos\phi\right) \boldsymbol{e}_1' \;+\; C\left(\dot{\psi} + \Omega^E \sin\alpha\cos\phi\right) \frac{d\boldsymbol{e}_1'}{dt}\\ &- A\frac{d}{dt}\left(\Omega^E \sin\alpha \sin\phi\right) \boldsymbol{e}_2' \;-\; A\left(\Omega^E \sin\alpha \sin\phi\right) \frac{d\boldsymbol{e}_2'}{dt} \;+\; A\ddot{\phi}\,\boldsymbol{e}_3 \;+\; A\left(\dot{\phi} + \Omega^E \cos\alpha\right) \frac{d\boldsymbol{e}_3}{dt}.\end{aligned}$$

The vectors $\{\boldsymbol{e}_1', \boldsymbol{e}_2', \boldsymbol{e}_3\}$ are the unit vectors of the frame $Gx_1'x_2'x_3$ and so their (true) rates of change are given by

$$\frac{d\boldsymbol{e}_1'}{dt} = \boldsymbol{\omega}^* \times \boldsymbol{e}_1', \qquad \frac{d\boldsymbol{e}_2'}{dt} = \boldsymbol{\omega}^* \times \boldsymbol{e}_2', \qquad \frac{d\boldsymbol{e}_3}{dt} = \boldsymbol{\omega}^* \times \boldsymbol{e}_3,$$

where $\boldsymbol{\omega}^*$ is the angular velocity of the frame $Gx_1'x_2'x_3$ relative to an inertial frame. By the addition theorem for angular velocities, this is given by

$$\begin{aligned}\boldsymbol{\omega}^* &= \dot{\phi}\,\boldsymbol{e}_3 + \boldsymbol{\Omega}^E\\ &= \left(\Omega^E \sin\alpha\cos\phi\right) \boldsymbol{e}_1' - \left(\Omega^E \sin\alpha \sin\phi\right) \boldsymbol{e}_2' + \left(\dot{\phi} + \Omega^E \cos\alpha\right) \boldsymbol{e}_3.\end{aligned}$$

On combining these results and simplifying, we obtain the three component equations of motion for the gyro. As before, the second of these equations serves to determine K and the first and third are

$$\frac{d}{dt}\left(\dot{\psi} + \Omega^E \sin\alpha\cos\phi\right) = 0,$$

and

$$A\ddot{\phi} + \Omega^E \sin\alpha \left[C\left(\dot{\psi} + \Omega^E \sin\alpha\cos\phi\right) - A\Omega^E \sin\alpha\cos\alpha \right]\sin\phi = 0.$$

It follows that

$$\dot{\psi} + \Omega^E \sin\alpha\cos\phi = n,$$

where n is a constant. Hence the **spin** $\boldsymbol{\omega}\cdot\boldsymbol{e}_1'$ of the gyro is a constant of the motion. The equation for the **precession angle** ϕ then becomes

$$A\ddot{\phi} + \Omega^E \sin\alpha \left[Cn - A\Omega^E \sin\alpha\cos\alpha \right]\sin\phi = 0. \qquad (19.39)$$

In practice, the ratio $A\Omega^E/Cn$ is *very* small so that $Cn - A\Omega^E \sin\alpha\cos\alpha$ can be closely approximated by Cn.

Equation (19.39) is the equation for the (large) oscillations of a pendulum about the direction $\phi = 0$, that is, about the axis Gx_1. Since this axis points north, it follows that *the axis of the gyro performs periodic oscillations about the northerly direction.* If these oscillations are damped (a feature not included in our model), then the axis of the gyro will eventually settle pointing north. The gyro does therefore 'home' to north and (despite statements in the literature to the contrary) does not have to be initially set pointing north.

Example 19.5 *The oscillaton period of the gyrocompass*

Estimate the period of small oscillations of the gyro in a typical case.

Solution

In a typical case the ratio $\Omega^E/n < 10^{-6}$ so that the equation (19.39) is *very* accurately approximated by

$$A\ddot{\phi} + \left(Cn\Omega^E \sin\alpha\right)\sin\phi = 0,$$

which, for **small oscillations** is approximated by

$$A\ddot{\phi} + \left(Cn\Omega^E \sin\alpha\right)\phi = 0.$$

The period of these simple harmonic oscillations is

$$\tau = 2\pi\left(\frac{A}{Cn\Omega^E \sin\alpha}\right)^{1/2}.$$

With $\Omega^E = 0.000073$, $n = 50$ Hz, $C = 2A$ and $\alpha = 45°$ this gives a **period** of about 35 seconds. ■

19.8 EULER'S EQUATIONS

The simplest set of equations for the rotational motion of a rigid body are **Euler's equations**. They apply to any body, symmetrical or not, but they come with some deficiencies.

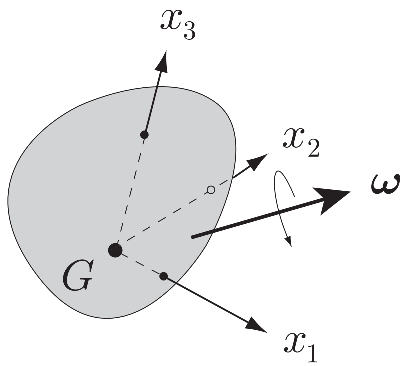


FIGURE 19.10 The principal axes $Gx_1x_2x_3$ are embedded in the body and have the same angular velocity as the body.

Consider a general rigid body with centre of mass G and principal axes $Gx_1x_2x_3$, as shown in Figure 19.10. If the body has an axis of dynamical symmetry through G, then there are infinitely many sets of principal axes at G and it is possible to find principal axes that are not rigidly attached to the body; indeed, this has been an essential feature when calculating the angular momentum $\boldsymbol{L}_G$ for a symmetrical body. However, we now wish to take a set of principal axes that is *rigidly fixed* within the body, even when this is not necessary. To emphasise the point, we will call these a set of **embedded axes** at G. The most important property of a reference frame embedded in a body is that its *angular velocity is the same as the angular velocity of the body in which it is embedded.*

Let $\boldsymbol{\omega}$, the **angular velocity** of the body, be expanded in the form

$$\boldsymbol{\omega} = \omega_1\, \boldsymbol{e}_1 + \omega_2\, \boldsymbol{e}_2 + \omega_3\, \boldsymbol{e}_3, \tag{19.40}$$

where $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ are the unit vectors of the coordinate system $Gx_1x_2x_3$. Then the corresponding **angular momentum** is

$$\boldsymbol{L}_G = A\omega_1\, \boldsymbol{e}_1 + B\omega_2\, \boldsymbol{e}_2 + C\omega_3\, \boldsymbol{e}_3, \tag{19.41}$$

where $\{A, B, C\}$ are the principal moments of inertia of the body. If the body is acted upon by the external moment $\boldsymbol{K}_G$, then the equation of rotational motion about G is

$$\frac{d}{dt}\big(A\omega_1\, \boldsymbol{e}_1 + B\omega_2\, \boldsymbol{e}_2 + C\omega_3\, \boldsymbol{e}_3\big) = \boldsymbol{K}_G. \tag{19.42}$$

Now

$$\frac{d}{dt}\,(A\omega_1\, \boldsymbol{e}_1 + B\omega_2\, \boldsymbol{e}_2 + C\omega_3\, \boldsymbol{e}_3) = A\dot{\omega}_1\, \boldsymbol{e}_1 + B\dot{\omega}_2\, \boldsymbol{e}_2 + C\dot{\omega}_3\, \boldsymbol{e}_3 + A\omega_1 \frac{d\boldsymbol{e}_1}{dt} + B\omega_2 \frac{d\boldsymbol{e}_2}{dt} + C\omega_3 \frac{d\boldsymbol{e}_3}{dt},$$

and, since the unit vectors $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ are embedded in the body, it follows that

$$\frac{d\boldsymbol{e}_1}{dt} = \boldsymbol{\omega}\times\boldsymbol{e}_1, \qquad \frac{d\boldsymbol{e}_2}{dt} = \boldsymbol{\omega}\times\boldsymbol{e}_2, \qquad \frac{d\boldsymbol{e}_3}{dt} = \boldsymbol{\omega}\times\boldsymbol{e}_3,$$

where $\boldsymbol{\omega}$ is given by equation (19.40). On combining these results and simplifying, we obtain the three component equations

Euler's equations

$$\begin{aligned} A\,\dot{\omega}_1 - (B - C)\,\omega_2\,\omega_3 &= K_1 \\ B\,\dot{\omega}_2 - (C - A)\,\omega_3\,\omega_1 &= K_2 \\ C\,\dot{\omega}_3 - (A - B)\,\omega_1\,\omega_2 &= K_3 \end{aligned} \qquad (19.43)$$

where $\{\omega_1, \omega_2, \omega_3\}$ and $\{K_1, K_2, K_3\}$ are the components of the angular velocity $\boldsymbol{\omega}$ and the applied moment $\boldsymbol{K}_G$ *in the embedded coordinate system* $Gx_1x_2x_3$. These are **Euler's equations of motion**.

As an example, we re-consider a problem solved earlier: the free motion of a body with axial symmetry.

Example 19.6 *Body with axial symmetry*

Solve Euler's equations for the free motion of a body with axial symmetry.

Solution

Let Gx_3 be the axis of symmetry so that $B = A$. In free motion, $K_1 = K_2 = K_3 = 0$ and the Euler equations reduce to

$$\begin{aligned} A\,\dot{\omega}_1 - (A - C)\,\omega_2\,\omega_3 &= 0, \\ A\,\dot{\omega}_2 - (C - A)\,\omega_3\,\omega_1 &= 0, \\ C\,\dot{\omega}_3 &= 0. \end{aligned}$$

The third equation gives

$$\omega_3 = n,$$

where n is a constant that we recognise as the spin $\boldsymbol{\omega} \cdot \boldsymbol{e}_3$. The first two equations then reduce to

$$\dot{\omega}_1 - \Omega\,\omega_2 = 0, \qquad \dot{\omega}_2 + \Omega\,\omega_1 = 0,$$

where $\Omega = (A - C)n/A$. On eliminating ω_2, we find that ω_1 satisfies the SHM equation

$$\ddot{\omega}_1 + \Omega^2\omega_1 = 0,$$

which has general solution

$$\omega_1 = P\,\sin(\Omega t + \gamma),$$

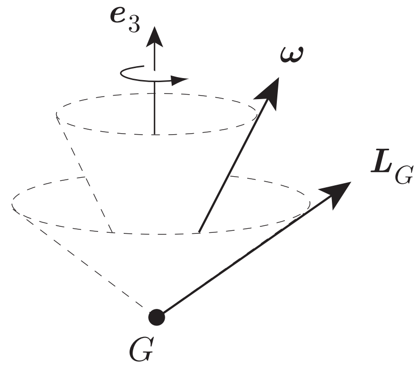


FIGURE 19.11 Relative to the embedded frame, $\boldsymbol{\omega}$ and $\boldsymbol{L}_G$ precess around the symmetry axis.

where P and γ are constants of integration; the corresponding solution for ω_2 is

$$\omega_2 = P\cos(\Omega t + \gamma).$$

The **general solution** of the Euler equations in this case is therefore

$$\begin{aligned}\omega_1 &= P\sin(\Omega t + \gamma),\\ \omega_2 &= P\cos(\Omega t + \gamma),\\ \omega_3 &= n,\end{aligned}$$

where P and γ are constants. The corresponding components of the **angular momentum** vector $\boldsymbol{L}_G$ are

$$\begin{aligned}L_1 &= AP\sin(\Omega t + \gamma),\\ L_2 &= AP\cos(\Omega t + \gamma),\\ L_3 &= Cn.\end{aligned}$$

The geometrical interpretation of these results is that *the vectors $\boldsymbol{e}_3$, $\boldsymbol{\omega}$ and $\boldsymbol{L}_G$ all lie in the same plane, where they make consant angles wth each other, and that this plane rotates around the embedded axis Gx_3 at the constant rate* $(A - C)n/A$. This situation is shown in Figure 19.11 for the case in which $A > C$.

Motion of the body

What we have found is the time variation of $\boldsymbol{\omega}$ and $\boldsymbol{L}$ as seen from the *principal embedded frame* $Gx_1x_2x_3$. The true picture (relative to an inertial frame) is quite different. In this frame, $\boldsymbol{L}_G$ is constant and the time variation of $\boldsymbol{\omega}$ and the motion of the body have yet to be determined. In general, deducing the motion of the body is *more* difficult than solving the Euler equations. In the present case though, this can can be done fairly easily. From the above expressions for $\boldsymbol{\omega}$ and $\boldsymbol{L}_G$, it follows that $\boldsymbol{\omega}$ can be written in the form

$$\begin{aligned}\boldsymbol{\omega} &= \frac{1}{A}\boldsymbol{L}_G + \left(1 - \frac{C}{A}\right) n\,\boldsymbol{e}_3\\ &= \frac{1}{A}\boldsymbol{L}_G + \left(\frac{A-C}{AC}\right) L_3\,\boldsymbol{e}_3\\ &= \frac{1}{A}\boldsymbol{L}_G + \left(\frac{A-C}{AC}\right) (|\,\boldsymbol{L}_G\,|\cos\alpha)\,\boldsymbol{e}_3,\end{aligned}$$

where α is the constant angle between $\boldsymbol{L}_G$ and the axis Gx_3. If we now write $L = |\boldsymbol{L}_G|$ and $\boldsymbol{k} = \boldsymbol{L}_G/L$, then

$$\boldsymbol{\omega} = \frac{L}{A}\,\boldsymbol{k} + L\cos\alpha\left(\frac{A-C}{AC}\right)\boldsymbol{e}_3,$$

which is precisely the solution found in section 19.4. ■

Deficiencies of Euler's equations

The above example illustrates the **first deficiency** in Euler's equations:

The solutions of Euler's equations yield the time variation of $\boldsymbol{\omega}$ *as seen from the* ***embedded reference frame****. The position of the body is still unknown.*

Thus Euler's equations do not tell us *where the body is* at time t, which is the most important unknown in the problem. In the example above, we were able to *deduce* the true motion of the body fairly simply. However, for an unsymmetrical body (which is the problem of principal interest that remains), there is no simple method and a further substantial calculation is required.

A **second** (but related) **deficiency** appears when the applied moment $\boldsymbol{K}_G$ is not zero:

When the applied moment $\boldsymbol{K}_G$ *is not zero, the components* K_1, K_2, K_3 *will not generally be known since the orientation of the body is unknown. In such a case, one cannot even begin the solution.*

For example, in the problem of the spining top, the moment of the gravity force is known, but the *components* of this moment in the embedded reference frame are not known. The right sides of Euler's equations for the spinning top are therefore unknowns.

Dynamical balancing

The difficulties mentioned above disappear when the motion of the body is given, and we wish to find the moments that are being applied to the body to permit that motion. In such a case, Euler's equations give the answer immediately. Consider the case of an axially symmetric body which is smoothly constrained to rotate about a fixed axis through G but is inaccurately mounted so that the rotation axis is not quite coincident with the symmetry axis, as shown in Figure 19.12. What effect does this mis-alignment have on the pivots?

Let the symmetry axis be Gx_3 and take the embedded axis Gx_2 to be perpendicular to the *rotation axis* PQ; the direction of the embedded axis Gx_1 is then determined (see Figure 19.12). If the body is rotating about PQ at a constant rate λ, then the angular

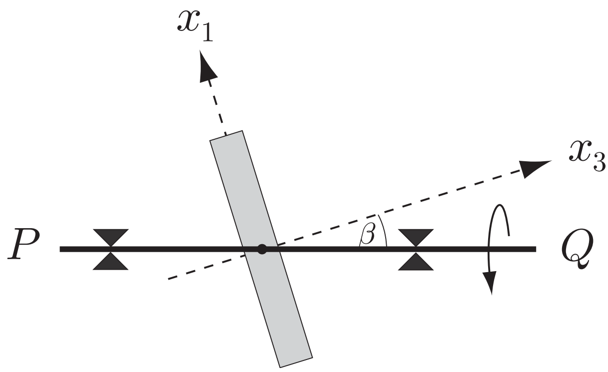


FIGURE 19.12 The axially symmetric body is constrained to rotate about the fixed axis PQ which makes an angle β with the axis of symmetry.

velocity components in the principal embedded frame $Gx_1x_2x_3$ are

$$\omega_1 = -\lambda \sin\beta$$
$$\omega_2 = 0$$
$$\omega_3 = \lambda \cos\beta.$$

To find the moment exerted by the pivots, we simply substitute these expressions into the left sides of Euler's equations. This gives

$$K_1 = 0, \qquad K_2 = (A - C)\lambda^2 \sin\beta \cos\beta, \qquad K_3 = 0.$$

Thus the applied moment is constant in the embedded frame. However, it is *not constant in an inertial frame*. The axis Gx_2 is perpendicular to PQ and rotates around it at a constant rate λ. *The pivots are therefore subjected to oscillating forces with angular frequency λ and magnitude proportional to λ^2*. In such a configuration, the body is said to be **dynamically unbalanced**. Since G lies on the rotation axis, the body can rest in equilibrium in any position, but, when rotating, it exerts oscillating forces on the pivots that may result in vibration of the mounting. This is why the wheels of motor cars must be re-balanced* when a new tyre is fitted. For a turbine with a mass of many tonnes rotating at 50 Hz say, extremely accurate alignment is needed, for otherwise the forces needed to restrain the turbine would be so large that it would break loose its mountings!

19.9 FREE MOTION OF AN UNSYMMETRICAL BODY

The free motion of unsymmetrical bodies is an important topic in space technology because of its application to the tumbling motion of spacecraft and space satellites.

* This is done by adding lead weights at determined positions around the rim of the wheel.

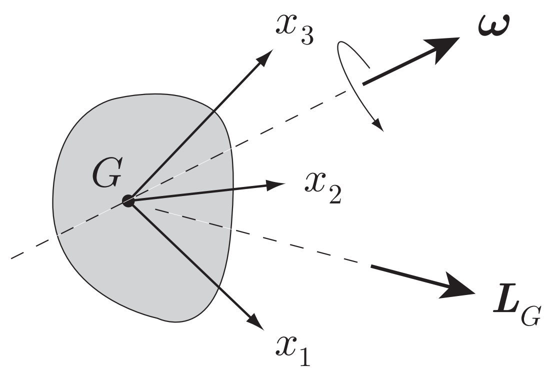


FIGURE 19.13 The unsymmetrical body moves under no forces with G at rest.

Such motions are considerably more difficult to analyse than for bodies with axial symmetry. The reader may be surprised then to learn that there is an exact **analytical solution** to the problem of an unsymmetrical body moving under no forces (see Landau & Lifshitz [6], Chapter VI). However, since this solution is expressed in terms of an impressive array of elliptic functions and theta functions, it gives one almost no idea of what the motion is actually like! Instead, we will use **Euler's equations** combined with **geometrical arguments**. This enables us to make interesting predictions about the motion without too much analysis.

Rotation of a free body about a fixed axis

Suppose a **free unsymmetrical body** $\mathcal{B}$ (an asteroid, for example) is observed from an inertial frame in which its centre of mass G is at rest. Is it possible for the body to rotate about a **fixed*** **axis**?

This is one of the few questions about unsymmetrical bodies that can be answered quite easily. If such an axis does exist, then it must pass through the fixed point G. Let $Gx_1x_2x_3$ be a set of principal axes of $\mathcal{B}$ at G. Since $\mathcal{B}$ is unsymmetrical, these principal axes are unique (apart from labelling) and are therefore necessarily *embedded* in $\mathcal{B}$; Euler's equations therefore apply.

The kinetic energy of $\mathcal{B}$ is given by $T = \frac{1}{2}I|\boldsymbol{\omega}|^2$, where I is the moment of inertia of $\mathcal{B}$ about the rotation axis and $\boldsymbol{\omega}$ is its angular velocity. Since $\mathcal{B}$ is rotating about a fixed axis, I is a constant. Also, since the motion takes place under no forces, the kinetic energy T is constant. It follows that the angular speed $|\boldsymbol{\omega}|$ must be constant and that the components $\{\omega_1, \omega_2, \omega_3\}$ of $\boldsymbol{\omega}$ in the embedded principal frame $Gx_1x_2x_3$ must be constants. Hence $\dot{\omega}_1 = \dot{\omega}_2 = \dot{\omega}_3 = 0$ and **Euler's equations** then imply that

$$(B - C)\omega_2\omega_3 = (C - A)\omega_3\omega_1 = (A - B)\omega_1\omega_2 = 0,$$

* This means an axis fixed in space and fixed in the body.

where $\{A, B, C\}$ are the principal moments of inertia about the axes Gx_1, Gx_2, Gx_3. Since $\mathcal{B}$ is an *unsymmetrical* body, A, B and C are all different and so

$$\omega_2\omega_3 = \omega_3\omega_1 = \omega_1\omega_2 = 0.$$

The only possible motions are therefore

$$\text{(i)}\quad \omega_1 = \text{constant}, \qquad \omega_2 = \omega_3 = 0,$$

$$\text{(ii)}\quad \omega_2 = \text{constant}, \qquad \omega_3 = \omega_1 = 0,$$

$$\text{(iii)}\quad \omega_3 = \text{constant}, \qquad \omega_1 = \omega_2 = 0.$$

The first solution corresponds to the body rotating with constant angular speed about the principal axis Gx_1. The other solutions correspond to rotations about the principal axes Gx_2 and Gx_3. Hence:

Free unsymmetrical body rotating about a fixed axis

A free unsymmetrical body can only rotate about a fixed axis if this axis is one of the three **principal axes** of the body at G. The rotation will then have constant angular speed.

For example, if it is required that a spacecraft should rotate about a fixed axis, this axis *must* be one of the three principal axes through its centre of mass. Later we will examine the **stability** of each of these steady motions.

General motion of a free unsymmetrical body

Once again we will suppose that the motion is viewed from an inertial frame in which G is at rest. It will be convenient to express the Euler equations in terms of the components of $\boldsymbol{L}$ in the frame $Gx_1x_2x_3$ (instead of the components of $\boldsymbol{\omega}$). Since $L_1 = A\,\omega_1$, $L_2 = B\,\omega_2$ and $L_3 = C\,\omega_3$, this gives the equivalent system of equations

$$\begin{aligned} BC\,\dot{L}_1 &= (B - C)\,L_2\,L_3, \\ AC\,\dot{L}_2 &= (C - A)\,L_3\,L_1, \\ AB\,\dot{L}_3 &= (A - B)\,L_1\,L_2. \end{aligned} \tag{19.44}$$

If we multiply the first equation by AL_1, the second by BL_2, the third by CL_3 and add, we obtain

$$ABC\left(L_1\dot{L}_1 + L_2\dot{L}_2 + L_3\dot{L}_3\right) = 0,$$

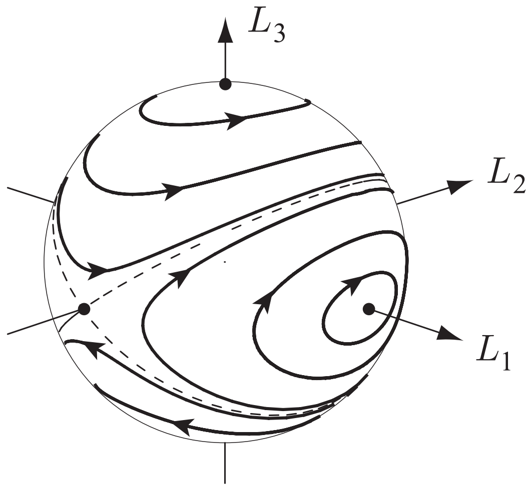


FIGURE 19.14 The L-sphere and the curves in which it meets different T-ellipsoids for the case $A < B < C$.

which integrates to give

$$L_1^2 + L_2^2 + L_3^2 = L^2, \tag{19.45}$$

where L is a constant. Similarly, if we multiply the first equation by L_1, the second by L_2, the third by L_3 and add, we obtain

$$BC\, L_1\dot{L}_1 + AC\, L_2\dot{L}_2 + AB\, L_3\dot{L}_3 = 0,$$

which integrates to give

$$\frac{L_1^2}{A} + \frac{L_2^2}{B} + \frac{L_3^2}{C} = 2T, \tag{19.46}$$

where T is a constant. The first of these integrals means that the **magnitude** of $\boldsymbol{L}_G$ is conserved* and the second means that the **kinetic energy** of the body is conserved.

The equations (19.45), (19.46) place geometric restrictions on the time variation of $\boldsymbol{L}_G$, *as seen from the embedded frame*. Equation (19.45) means that, if $\boldsymbol{L}_G$ is considered to be the position vector of a 'point' $\boldsymbol{L}$ in (L_1, L_2, L_3)-space, then that point must move on the surface of a sphere with centre O and radius L, which we will call L**-sphere**. In the same way, equation (19.46) means that the $\boldsymbol{L}$-point must also move on the surface of the ellipsoid† in (L_1, L_2, L_3)-space defined by equation (19.46). We will call this ellipsoid,

* Note that, in the embedded frame, the *individual components* L_1, L_2, L_3 *are not conserved*, but the magnitude $\left(L_1^2 + L_2^2 + L_3^2\right)^{1/2}$ is.

† An ellipsoid is the three-dimensional counterpart of an ellipse. The standard equation of an ellipsoid with semi-axes a, b, c is

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} + \frac{x_3^2}{c^2} = 1.$$

which has the semi-axes $(2AT)^{1/2}$, $(2BT)^{1/2}$, $(2CT)^{1/2}$, the T**-ellipsoid**. It follows that *the path of the* $\boldsymbol{L}$*-point must be a curve of intersection of the* L*-sphere and the* T*-ellipsoid*. Figure 19.14 shows the L-sphere for a fixed value of the parameter L. The paths on its surface are examples of the curves in which it is cut by the T-ellipsoid for different values of the parameter T.

Paths of the $\boldsymbol{L}$-point

Suppose that $A < B < C$ and that L is fixed. For each value of L, values of T that are too small or too large are not admissible. Indeed we can see from equation (19.46) that

$$T_{\min} = \frac{L^2}{2C} \le T \le \frac{L^2}{2A} = T_{\max}.$$

If T lies outside this range, then the L-sphere and the T-ellipsoid do not intersect.

(i) When $T = T_{\min}$ the intersection consists of the two isolated points $(0, 0, \pm L)$. The constant solution $\boldsymbol{L} = (0, 0, L)$ corresponds to steady rotation in the positive sense about the principal axis Gx_3. The solution $\boldsymbol{L} = (0, 0, -L)$ corresponds to a steady rotation in the opposite sense.

(ii) Suppose now that T is slightly increased. The intersection point $(0, 0, L)$ then becomes a small closed curve* on the surface of the L-sphere that encircles the point $(0, 0, L)$ as shown in Figure 19.14. The $\boldsymbol{L}$-point must therefore move along this curve. The direction of motion of the $\boldsymbol{L}$-point can be deduced by examining the signs of $\dot{L}_1$ and $\dot{L}_2$ given by the first and second equations of (19.44). As T increases, the intersection curve becomes larger (and changes its shape) but it still encircles the point $(0, 0, L)$. This pattern continues until $T = L^2/2B$.

(iii) When $T = L^2/2B$, the L-sphere and the T-ellipsoid touch at the points $(0, \pm L, 0)$ and the intersection curves are the dashed lines in Figure 19.14. This is a transitional case between two different regimes of curves. It includes the case of steady rotation about the principal axis Gx_2.

(iv) When $T > L^2/2B$, the intersection curves are completely different. Now they encircle the point $(L, 0, 0)$ (or $(-L, 0, 0)$) and, as T approaches $T_{\max}$ they shrink to zero around these points. The case $T = T_{\max}$ corresponds to steady rotation about the principal axis Gx_1.

Stabilty of steady rotation about a principal axis

We can make some interesting deductions from the form of the above paths.

(i) In the general motion, the $\boldsymbol{L}$-point moves around a closed curve. It follows that the **time variation** of $\boldsymbol{L}$, as seen from the embedded frame, is **periodic**, a property not apparent from equations (19.44). This does not however mean that the motion of the *body* is periodic.

(ii) If the body is in steady rotation in the positive sense about the axis Gx_1 and is slightly disturbed, then the path of $\boldsymbol{L}$ remains close to the point $(L, 0, 0)$. This implies that,

* The projection of this curve on to the (x_1, x_2)-plane is a small ellipse with centre the origin.

in an inertial frame where $\boldsymbol{L}$ is constant, the axis Gx_1 remains close to $\boldsymbol{L}$. In other words, the original steady rotation about the axis Gx_1 is **stable** to small disturbances. The same applies to a steady rotation about the axis Gx_3. However, if a steady rotation about the axis Gx_2 is slightly disturbed, the paths of the $\boldsymbol{L}$-point lead far away from the original constant positions at $(0, \pm L, 0)$. This implies that a steady rotation about the axis Gx_2 is **unstable** to small disturbances. Hence:

The steady rotation of an unsymmetrical body about a principal axis is stable for the axes with the greatest and least moments of inertia, but unstable for the other axis.

Thus, it is useless to try to stabilise a satellite in steady rotation about the principal axis with the 'middle' moment of inertia.

(iii) The stability argument above is modified if the body has a means of **energy dissipation**. This cannot happen with rigid bodies, but real bodies, such as a satellite, can flex and slowly dissipate mechanical energy as heat. In this case, steady rotation of the body about the axis with the *least* moment of inertia (which corresponds to $T = T_{\max}$) is no longer stable. The 'radius' of the path of $\boldsymbol{L}$ around the axis Gx_1 will slowly increase until $T = L^2/(2B)$ when it will switch to a large path encircling one of the points $(\pm 0, 0, \pm L)$. This path will then gradually shrink to zero. Thus, *in the presence of energy dissipation, the body will end up rotating about the axis with the largest moment of inertia*. In the early days of space exploration* this fact was learned by hard experience!

Motion of the body

Suppose that Euler's equations have been solved for the unknowns ω_1, ω_2, ω_3. What is the corresponding motion of the body? Let $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ be the unit vectors of the principal embedded frame $Gx_1x_2x_3$. Then the time variation of these vectors, as seen from an inertial frame, is given by

$$\dot{\boldsymbol{e}}_1 = \boldsymbol{\omega} \times \boldsymbol{e}_1, \qquad \dot{\boldsymbol{e}}_2 = \boldsymbol{\omega} \times \boldsymbol{e}_, \qquad \dot{\boldsymbol{e}}_3 = \boldsymbol{\omega} \times \boldsymbol{e}_3.$$

Unfortunately, these equations are not really uncoupled since Euler's equations give us $\boldsymbol{\omega}$ in the form

$$\boldsymbol{\omega} = \omega_1 \boldsymbol{e}_1 + \omega_2 \boldsymbol{e}_2 + \omega_3 \boldsymbol{e}_3,$$

which also involves the unknown vectors $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$. On combining these equations, we get the set of *coupled* ODEs

$$\begin{aligned} \dot{\boldsymbol{e}}_1 &= \omega_3 \boldsymbol{e}_2 - \omega_2 \boldsymbol{e}_3, \\ \dot{\boldsymbol{e}}_2 &= \omega_1 \boldsymbol{e}_3 - \omega_3 \boldsymbol{e}_1, \\ \dot{\boldsymbol{e}}_3 &= \omega_2 \boldsymbol{e}_1 - \omega_1 \boldsymbol{e}_2, \end{aligned} \tag{19.47}$$

* Kaplan [5] recounts the following incident: The first U.S. satellite, Explorer I, was set into orbit rotating about its longitudinal axis, which was the principal axis with the *least* moment of inertia. After only a few hours, radio signals indicated that a tumbling motion had developed and was increasing in amplitude in an unstable manner. It was concluded that the satellite's flexible antennae were dissipating energy and causing a transfer of body spin from the axis of minimum inertia to a transverse axis of maximum inertia.

for the unknown vectors $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$. These equations cannot usually be solved explicitly, but they are suitable for numerical integration (see Problem 19.17) and for approximate solution by perturbation analysis. This is illustrated by the following example.

Example 19.7 *Motion of a principal axis in space*

An unsymmetrical body is in steady rotation about the principal axis with the *largest* moment of inertia. Find an approximation to the wobble of this principal axis if the body is slightly disturbed.

Solution

Suppose that $A < B < C$ and that, in the *initial motion*, $\omega_1 = \omega_2 = 0$ and $\omega_3 = \Lambda$, where Λ is a constant. In the *disturbed motion* we suppose that ω_1, ω_2 (and their time derivatives) remain small. If we now apply **Euler's equations** and then **linearise**, we obtain

$$\begin{aligned} A\,\dot{\omega}_1 - (B - C)\Lambda\,\omega_2 &= 0,\\ B\,\dot{\omega}_2 - (C - A)\Lambda\,\omega_1 &= 0,\\ C\,\dot{\omega}_3 &= 0, \end{aligned}$$

Hence, in the linear approximation to the disturbed motion, the coupled equations for ω_1, ω_2 are easily solved to give

$$\begin{aligned} \omega_1 &= \epsilon\Lambda\,((C - B)/A)^{1/2}\cos(\Omega t + \gamma),\\ \omega_2 &= \epsilon\Lambda\,((C - A)/B)^{1/2}\sin(\Omega t + \gamma), \end{aligned} \tag{19.48}$$

where ϵ, γ are dimensionless constants and

$$\Omega = \Lambda\left(\frac{(C - A)(C - B)}{AB}\right)^{1/2}.$$

The corresponding expressions for L_1, L_2 are

$$\begin{aligned} L_1 &= \epsilon\Lambda\,(A(C - B))^{1/2}\cos(\Omega t + \gamma),\\ L_2 &= \epsilon\Lambda\,(B(C - A))^{1/2}\sin(\Omega t + \gamma). \end{aligned} \tag{19.49}$$

Hence the projection of the path of the $\boldsymbol{L}$-point on to the (x_1, x_2)-plane is an ellipse executed in the anti-clockwise sense. This is entirely consistent with the paths shown in Figure 19.14 and our discussion of stability.

Now to find the actual time variaton of the unit vector $\boldsymbol{e}_3$. In the initial motion we have

$$\begin{aligned} \boldsymbol{e}_1 &= \cos\Lambda t\,\boldsymbol{i} + \sin\Lambda t\,\boldsymbol{j},\\ \boldsymbol{e}_2 &= -\sin\Lambda t\,\boldsymbol{i} + \cos\Lambda t\,\boldsymbol{j},\\ \boldsymbol{e}_3 &= \boldsymbol{k}, \end{aligned}$$

where $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ is a *fixed* orthonormal set. Since ω_1, ω_2 are small in the disturbed motion, the vectors $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ that appear in the third equation of (19.47) can be

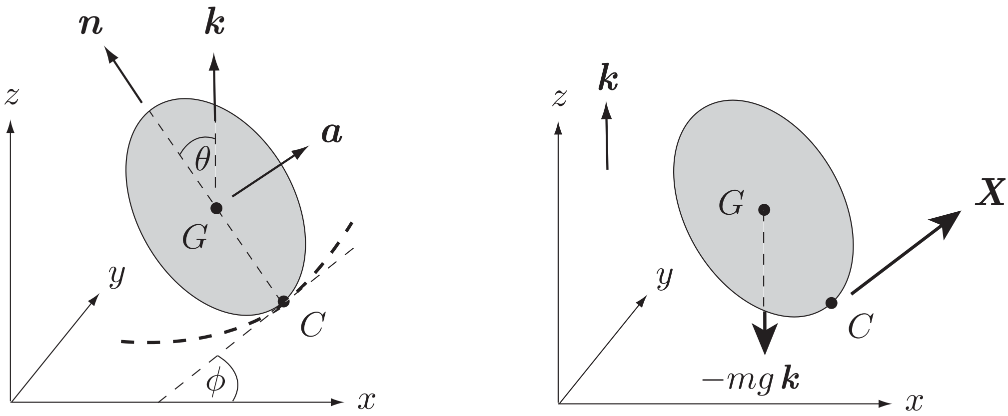


FIGURE 19.15 The wheel rolls on a rough horizonal plane. The dashed curve is the path of the point of contact C.

replacd by their steady (zero order) approximations to give

$$\dot{\boldsymbol{e}}_3 = \omega_2\,(\cos\Lambda t\,\boldsymbol{i} + \sin\Lambda t\,\boldsymbol{j}) - \omega_1\,(-\sin\Lambda t\,\boldsymbol{i} + \cos\Lambda t\,\boldsymbol{j})$$

where ω_1 and ω_2 are given by the expressions (19.48). This equation can now be integrated explicitly to give the first order approximation to the wobble of the axis vector $\boldsymbol{e}_3$.

We will not actually do this integration because the result is messy and the details of the formula are less interesting than its structure. By examining the form of the right side of the above equation one can see that $\boldsymbol{e}_3$ must have the form $\boldsymbol{e}_3 = \boldsymbol{k} +$ (small terms in the $\boldsymbol{i}$ and $\boldsymbol{j}$ directions), where these small terms contain oscillating functions with *two different frequencies* $\Omega + \Lambda$ and $\Omega - \Lambda$. In general the ratio of these frequencies will not be a rational number and so the **motion of the axis vector $\boldsymbol{e}_3$ is not periodic**. In fact, the projection of $\boldsymbol{e}_3$ onto the fixed plane through G perpendicular to $\boldsymbol{k}$ describes a Lissajous type path bounded by two ellipses. (See however Problem 19.14.) ■

19.10 THE ROLLING WHEEL

Suppose a wheel is free to *roll* on a rough horizontal floor. How does it move? This, our final problem, typifies problem solving in classical mechanics. Nothing could be easier to describe, but few could just sit down and write out the solution to this classic problem.

Since the rolling condition is a non-holonomic constraint, the rolling wheel cannot be solved by the Lagrangian methods presented in this book.* We therefore use vectorial mechanics. Figure 19.15 (left) shows a thin circular disk† of mass M and radius b rolling

* For an extension of the Lagrangian method to non-holonomic systems, see Goldstein [4], chapter 2.

† The disk need not be uniform, but it must have axial symmetry about its centre. For example, it could be a hoop.

on the (x, y)-plane. The thick dashed curve is the path of the point of contact C. The vector $\boldsymbol{a}$ $(= \boldsymbol{a}(t))$ is the axial unit vector and the vector $\boldsymbol{n}$ $(= \boldsymbol{n}(t))$ is the unit vector in the direction $\overrightarrow{CG}$. Since the disk has axisymmetry, the angular velocity of the disk can be expressed in the form

$$\boldsymbol{\omega} = \boldsymbol{a}\times\dot{\boldsymbol{a}} + \lambda\,\boldsymbol{a}, \tag{19.50}$$

where the spin λ $(= \boldsymbol{\omega}\cdot\boldsymbol{a})$ is an unknown scalar function of the time. The corresponding angular momentum $\boldsymbol{L}_G$ is then given by

$$\boldsymbol{L}_G = A\,\boldsymbol{a}\times\dot{\boldsymbol{a}} + C\lambda\,\boldsymbol{a}, \tag{19.51}$$

where $\{A, A, C\}$ are the principal moments of inertia of the disk at G.

The disk moves under the forces shown in Figure 19.15 (right). It follows that the equation of **translational motion** of the disk is

$$M\frac{d\boldsymbol{V}}{dt} = \boldsymbol{X} - Mg\,\boldsymbol{k}, \tag{19.52}$$

where $\boldsymbol{V}$ is the velocity of G, and that the equation of **rotational motion** is

$$\frac{d}{dt}\left(A\,\boldsymbol{a}\times\dot{\boldsymbol{a}} + C\lambda\,\boldsymbol{a}\right) = (-b\boldsymbol{n})\times\boldsymbol{X}. \tag{19.53}$$

In addition, we have the **rolling condition** that the particle of the disk in contact with the floor has zero velocity, that is

$$\boldsymbol{V} + \boldsymbol{\omega}\times(-b\boldsymbol{n}) = \boldsymbol{0}. \tag{19.54}$$

Equations (19.52)–(19.54) are the **governing equations** of the motion. On eliminating $\boldsymbol{X}$ and $\boldsymbol{V}$, we obtain the equation

$$\frac{d}{dt}\left(A\,\boldsymbol{a}\times\dot{\boldsymbol{a}} + C\lambda\,\boldsymbol{a}\right) + Mb^2\left[\dot{\boldsymbol{\omega}} - (\dot{\boldsymbol{\omega}}\cdot\boldsymbol{n})\,\boldsymbol{n} - (\boldsymbol{\omega}\cdot\boldsymbol{n})\,\dot{\boldsymbol{n}}\right] + Mgb\,\boldsymbol{n}\times\boldsymbol{k} = \boldsymbol{0}. \tag{19.55}$$

where $\boldsymbol{\omega}$ is given by equation (19.50). The vectors $\boldsymbol{a}$ and $\boldsymbol{n}$ are not independent. In fact $\boldsymbol{n}$ can be expressed in terms of $\boldsymbol{a}$ and $\boldsymbol{k}$ in the form

$$\boldsymbol{n} = \frac{\boldsymbol{k} - (\boldsymbol{k}\cdot\boldsymbol{a})\,\boldsymbol{a}}{|\,\boldsymbol{k} - (\boldsymbol{k}\cdot\boldsymbol{a})\,\boldsymbol{a}\,|}\,. \tag{19.56}$$

Thus it is possible in principle to eliminate $\boldsymbol{n}$ from equation (19.55) and obtain a vector equation solely in terms of the unknown **axial vector** $\boldsymbol{a}$ and the **spin** λ. This equation is sufficient to determine $\boldsymbol{a}$ and λ.

All this begins to look easier when we take components of the vector equation (19.55) in three *well chosen* directions, namely, the (instantaneous) principal directions $\boldsymbol{a}$, $\boldsymbol{n}$ and

$\boldsymbol{a}\times\boldsymbol{n}$. On taking the scalar product of equation (19.55) with each of these vectors, we obtain the scalar equations

$$\begin{aligned}(C+Mb^2)\dot\lambda - Mb^2\left[\boldsymbol{a},\dot{\boldsymbol{a}},\boldsymbol{n}\right]\left(\dot{\boldsymbol{n}}\cdot\boldsymbol{a}\right) &= 0,\\ A\left[\boldsymbol{a},\ddot{\boldsymbol{a}},\boldsymbol{n}\right]+C\lambda\left(\dot{\boldsymbol{a}}\cdot\boldsymbol{n}\right) &= 0,\\ (A+Mb^2)\left(\ddot{\boldsymbol{a}}\cdot\boldsymbol{n}\right)-(C+Mb^2)\left[\boldsymbol{a},\dot{\boldsymbol{a}},\boldsymbol{n}\right]- &\\ Mb^2\left[\boldsymbol{a},\dot{\boldsymbol{a}},\boldsymbol{n}\right]\left[\boldsymbol{a},\boldsymbol{n},\dot{\boldsymbol{n}}\right]-Mgb\,(\boldsymbol{a}\cdot\boldsymbol{k}) &= 0.\end{aligned} \tag{19.57}$$

Here $[\,\boldsymbol{p},\boldsymbol{q},\boldsymbol{r}\,]$ means the triple scalar product of the vectors $\boldsymbol{p}$, $\boldsymbol{q}$ and $\boldsymbol{r}$. These equations look better still when expressed in terms of scalar variables. Let θ be the angle between the plane of the disk and the vertical, and let ϕ be the angle between the tangent line to the disk at C and the axis Ox, as shown in Figure 19.15 (left). Then $\boldsymbol{a}$ and $\boldsymbol{n}$ take the form

$$\boldsymbol{a} = (\cos\theta\sin\phi)\,\boldsymbol{i}-(\cos\theta\cos\phi)\,\boldsymbol{j}+(\sin\theta)\,\boldsymbol{k}, \tag{19.58}$$

$$\boldsymbol{n} = -(\sin\theta\sin\phi)\,\boldsymbol{i}+(\sin\theta\cos\phi)\,\boldsymbol{j}+(\cos\theta)\,\boldsymbol{k}, \tag{19.59}$$

where $\{\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}\}$ are the constant unit vectors of the inertial frame $Oxyz$. In terms of θ and ϕ, the equations (19.57) become

$$\begin{aligned}(C+Mb^2)\,\dot\lambda+Mb^2\cos\theta\,\dot\theta\,\dot\phi &= 0,\\ A\left(\cos\theta\,\ddot\phi-2\sin\theta\,\dot\theta\dot\phi\right)+C\lambda\,\dot\theta &= 0,\\ (A+Mb^2)\,\ddot\theta+A\cos\theta\sin\theta\,\dot\phi^2-(C+Mb^2)\lambda\cos\theta\,\dot\phi-Mgb\sin\theta &= 0.\end{aligned} \tag{19.60}$$

This is our final form of the **governing equations** for the rolling disk. Note that, unlike other problems involving an axially symmetric body, the *spin λ is not a constant of the motion*, but must be solved for, along with the angular coordinates θ, ϕ. The complexity of these coupled non-linear equations is such that explicit solutions are rare. However, they are suitable for numerical integration and (if they can be linearised) for finding approximate solutions. This is illustrated by the next example.

Example 19.8 *Stability of straight line rolling*

How fast must a wheel roll in a straight line so that it is stable to small disturbances?

Solution

In the **unperturbed motion** (along the x-axis say), we have $\theta=0$, $\phi=0$ and $\lambda=\Lambda$, where Λ is a positive constant. In the **perturbed motion**, we suppose that θ, ϕ and $\lambda-\Lambda$ (and their time derivatives) remain small so that we may **linearise** equations (19.60). This gives

$$\begin{aligned}\dot\lambda &= 0,\\ A\ddot\phi+C\Lambda\dot\theta &= 0,\\ A(A+Mb^2)\ddot\theta-(C+Mb^2)\Lambda\dot\phi-Mgb\theta &= 0.\end{aligned}$$

Hence, in this linear approximation, $\lambda = \Lambda$, a constant. The second equation gives

$$A\dot{\phi} + C\Lambda\theta = \text{constant},$$

and, in order to be satisfied by the unperturbed state, the constant must be zero. On substituting this into the third equation, we obtain

$$A(A + Mb^2)\ddot{\theta} + \left[\Lambda^2 C(C + Mb^2) - AMgb\right]\theta = 0. \qquad (19.61)$$

If the expression in the square brackets is *positive*, this is the SHM equation and θ will perform small oscillations about $\theta = 0$; the same is then true of ϕ. Hence, the **condition for stability** of straight line rolling is

$$\Lambda^2 C(C + Mb^2) - AMgb > 0,$$

that is

$$V^2 > \frac{AMgb^3}{C(C + Mb^2)},$$

where V is the speed of the disk. For a **hoop** with $A = \frac{1}{2}Mb^2$ and $C = Mb^2$, this condition becomes $V^2 > gb/4$. ■

Path of the point of contact

If $\boldsymbol{r}$ is the position vector of C, the point of contact, then

$$\boldsymbol{r} = \boldsymbol{R} - b\,\boldsymbol{n},$$

where $\boldsymbol{R}$ is the position vector of G. It follows that

$$\frac{d\boldsymbol{r}}{dt} = \boldsymbol{V} - b\,\dot{\boldsymbol{n}},$$

where $\boldsymbol{V}$ is the velocity of G. On using the rolling condition (19.54) and the expression (19.50) for $\boldsymbol{\omega}$, the differential equation for $\boldsymbol{r}$ becomes

$$\frac{d\boldsymbol{r}}{dt} = b\left[(\boldsymbol{a}\times\dot{\boldsymbol{a}})\times\boldsymbol{n} + \lambda\,\boldsymbol{a}\times\boldsymbol{n} - \dot{\boldsymbol{n}}\right]. \qquad (19.62)$$

This is the **path equation** for the point of contact. Once the equations (19.60) have been solved for λ, θ and ϕ, the right side of (19.62) is known and the path of C can then be found by integration. Numerical integration shows that some possible paths can be quite exotic, as shown in Figure 19.16.

Stability of the bicycle

Our analysis of the rolling wheel gives some idea of the complexity of the mechanics of the bicycle. Indeed, the question of stability is not completely settled and rival sets of

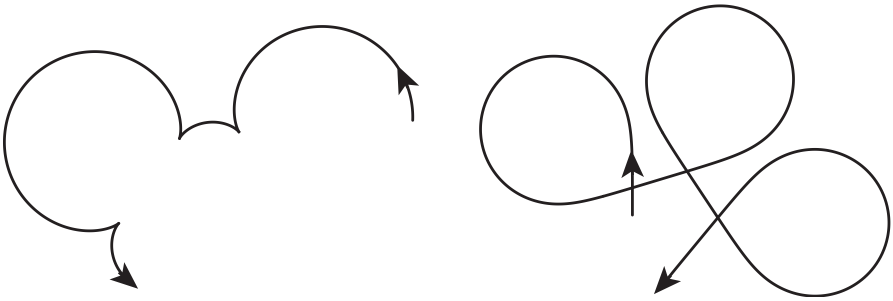


FIGURE 19.16 Two not-so-typical paths for the point of contact of a rolling wheel.

equations appear in the literature. Probably the most important paper on bicycle physics is that by Jones.* Jones constructed a non-standard bicycle in which the angular momentum of the wheels (the 'gyroscopic effect') was cancelled out by fitting extra wheels that spun in the opposite direction. It turned out that this bicycle was just as stable as a normal one and could be ridden 'no hands'. Despite Jones's findings, the myth that a bicycle is stabilised by the the gyroscopic effect of the spinning wheels still persists. However, the concensus of informed opinion is that the stability of a bicycle depends crucially on the geometry of the forks holding the front wheel. Jones constructed another non-standard bicycle in which the front wheel was four inches ahead of its normal position; it was almost unridable! The website `http://ruina.tam.cornell.edu/research/index.htm` has references to further interesting work on bicycle physics.

Problems on Chapter 19

Answers and comments are at the end of the book.

Harder problems carry a star ($*$).

Rolling balls

19.1 ***Ball rolling on a slope*** A uniform ball can roll or skid on a rough plane inclined at an angle β to the horizontal. Show that, in *any* motion of the ball, the component of $\boldsymbol{\omega}$ perpendicular to the plane is conserved. If the ball *rolls* on the plane, show that the path of the ball must be a parabola.

19.2 ***Ball rolling on a rotating turntable*** ✱ A rough horizontal turntable is made to rotate about a fixed vertical axis through its centre O with *constant* angular velocity $\Omega\boldsymbol{k}$, where the unit vector $\boldsymbol{k}$ points vertically upwards. A uniform ball of radius a can roll or skid on the turntable. Show that, in *any* motion of the ball, the vertical spin $\boldsymbol{\omega}\cdot\boldsymbol{k}$ is conserved. If the ball *rolls* on the turntable, show that

$$\dot{\boldsymbol{V}} = \tfrac{2}{7}\Omega\,\boldsymbol{k}\times\boldsymbol{V},$$

* Jones, D.E.H., The stability of the bicycle, *Physics Today* **23** (1970) pp. 34–40.

where $\boldsymbol{V}$ is the velocity of the centre of the ball viewed from a *fixed* reference frame. Deduce the amazing result that the path of the rolling ball must be a circle.

Suppose the ball is held at rest (relative to the turntable), with its centre a distance b from the axis $\{O, \boldsymbol{k}\}$, and is then released. Given that the ball rolls, find the radius and the centre of the circular path on which it moves.

19.3 ***Ball rolling on a fixed sphere*** ∗ A uniform ball with radius a and centre C *rolls* on the rough outer surface of a fixed sphere of radius b and centre O. Show that the radial spin $\boldsymbol{\omega} \cdot \boldsymbol{c}$ is conserved, where $\boldsymbol{c}$ $(= \boldsymbol{c}(t))$ is the *unit* vector in the radial direction $\overrightarrow{OC}$. [Take care!] Show also that $\boldsymbol{c}$ satisfies the equation

$$7(a+b)\,\boldsymbol{c}\times\ddot{\boldsymbol{c}} + 2an\,\dot{\boldsymbol{c}} + 5g\,\boldsymbol{c}\times\boldsymbol{k} = \boldsymbol{0},$$

where n is the constant value of $\boldsymbol{\omega} \cdot \boldsymbol{c}$ and $\boldsymbol{k}$ is the unit vector pointing vertically upwards.

By comparing this equation with that for the spinning top, deduce the amazing result that the ball can roll on the spherical surface without ever falling off. Find the minimum value of n such that the ball is stable at the highest point of the sphere.

Axisymmetric bodies

19.4 Investigate the steady precession of a top for the case in which the axis of the top moves in the horizontal plane through O. Show that for any $n \neq 0$ there is just *one* rate of steady precession and find its value.

19.5 ***The sleeping top*** By performing a perturbation analysis, show that a top will be stable in the vertically upright position if

$$C^2n^2 > 4AMgh,$$

in the standard notation. [You can do this by either the vectorial or Lagrangian method.]

19.6 Estimate how large the spin n of a pencil would have to be for it to be stable in the vertically upright position, spinning on its point. [Take the pencil to be a uniform cylinder 15 cm long and 7 mm in diameter.]

19.7 A juggler is balancing a spinning ball of diameter 20 cm on the end of his finger. Estimate the spin required for stability (i) for a uniform solid ball, (ii) for a uniform thin hollow ball. Which do you suppose the juggler uses?

19.8 Solve the problem of the free motion of an axisymmetric body by the Lagrangian method. Compare your results with those in Section 19.4. [Surprisingly awkward. Take the axis Gz of coordinates $Gxyz$ to point in the direction of $\boldsymbol{L}_G$ and make use of the Lagrange equations for the Euler angles ϕ and ψ.]

19.9 ***Frisbee with resistance*** A (wobbling) frisbee moving through air is subject to a frictional couple equal to $K\,\boldsymbol{\omega}$. Find the time variation of the axial spin λ $(= \boldsymbol{\omega} \cdot \boldsymbol{a})$, where $\boldsymbol{a}$ is the axial unit vector. Show also that $\boldsymbol{a}$ satisfies the equation

$$A\,\boldsymbol{a}\times\ddot{\boldsymbol{a}} + K\,\boldsymbol{a}\times\dot{\boldsymbol{a}} + C\lambda\,\dot{\boldsymbol{a}} = \boldsymbol{0}.$$

∗ By taking the cross product of this equation with $\dot{\boldsymbol{a}}$, find the time variation of $|\dot{\boldsymbol{a}}|$. Deduce that the angle between $\boldsymbol{\omega}$ and $\boldsymbol{a}$ decreases with time if $C > A$ (which it is for a normal frisbee). Thus, in the presence of linear resistance, the wobble dies away.

19.10 ***Spinning hoop on a smooth floor*** A uniform circular hoop of radius a rolls and slides on a *perfectly smooth* horizontal floor. Find its Lagrangian in terms of the Euler angles, and determine which of the generalised momenta are conserved. [Suppose that G has no *horizontal* motion.]

Investigate the existence of motions in which the angle between the hoop and the floor is a constant α. Show that Ω, the rate of steady precession, must satisfy the equation

$$\cos\alpha\,\Omega^2 - 2n\,\Omega - 2\frac{g}{a}\cot\alpha = 0,$$

where n is the constant axial spin. Deduce that, for $n \neq 0$, there are two possible rates of precession, a faster one going the 'same way' as n, and a slower one in the opposite direction. [These are interesting motions but one would need a *very* smooth floor to observe them.]

Euler's equations

19.11 ***Bicycle wheel*** A bicycle wheel (a hoop) of mass M and radius a is fitted with a smooth spindle lying along its symmetry axis. The wheel is spun with the spindle horizontal, and the spindle is then made to turn with angular speed Ω about a fixed vertical axis through the centre of the wheel. Show that n, the axial spin of the wheel, remains constant and find the moment that must be applied to the spindle to produce this motion.

19.12 ***Stability of steady rotation*** An unsymmetrical body is in steady rotation about a principal axis through G. By performing a perturbation analysis, investigate the stability of this motion for each of the three principal axes.

19.13 ***Frisbee with resistance*** Re-solve the problem of the frisbee with resistance (Problem 19.9) by using Euler's equations. [Find the time dependencies of ω_3 and $\omega_1^2 + \omega_2^2$.]

19.14 ***Wobble on spinning lamina*** An unsymmetrical lamina is in steady rotation about the axis through G perpendicular to its plane. Find an approximation to the wobble of this axis if the body is slightly disturbed. [This is a repeat of Example 19.7 for the special case in which the body is an unsymmetrical *lamina*; in this case $C = A+B$ and there is much simplification.]

19.15 ∗ ***Euler theory for the unsymmetrical lamina*** An unsymmetrical lamina has principal axes $Gx_1x_2x_3$ at G with the corresponding moments of inertia $\{A, B, A+B\}$. Initially the lamina is rotating with angular velocity Ω about an axis through G that lies in the (x_1, x_2)-plane and makes an acute angle α with Gx_1. By using Euler's equations, show that, in the subsequent motion,

$$\begin{aligned} \omega_1^2 + \omega_2^2 &= \Omega^2, \\ (B-A)\,\omega_2^2 + (B+A)\,\omega_3^2 &= (B-A)\,\Omega^2\sin^2\alpha. \end{aligned}$$

Interpret these results in terms of the motion of the '$\boldsymbol{\omega}$-point' moving in $(\omega_1, \omega_2, \omega_3)$-space and deduce that $\boldsymbol{\omega}$ is periodic when viewed from the embedded frame.

Find an ODE satisfied by ω_2 alone and deduce that the lamina will once again be rotating about the same axis after a time

$$\frac{4}{\Omega}\left(\frac{B-A}{B+A}\right)^{1/2}\int_0^{\pi/2}\frac{d\theta}{(1-\sin^2\alpha\,\sin^2\theta)^{1/2}}.$$

Computer assisted problems

19.16 Solve Lagrange's equations for the top numerically and obtain the motions of the axis shown in Figure 19.8. Use Euler's angles as coordinates and use computer assistance to obtain the equations as well as solve them. [It seems easier not to use the conservation relations for p_ϕ and p_ψ in their integrated form, since this requires the initial conditons to be incorporated into the equations of motion.]

19.17 Obtain the paths of the $\boldsymbol{L}$-point for an unsymmetrical body, as shown in Figure 19.14. These were obtained by solving Euler's equations, together with the equations

$$\dot{\boldsymbol{e}}_1 = \omega_3\boldsymbol{e}_2 - \omega_2\boldsymbol{e}_3, \qquad \dot{\boldsymbol{e}}_2 = \omega_1\boldsymbol{e}_3 - \omega_3\boldsymbol{e}_1, \qquad \dot{\boldsymbol{e}}_3 = \omega_2\boldsymbol{e}_1 - \omega_1\boldsymbol{e}_2,$$

numerically. [*Mathematica* handled the solution of these twelve simultaneous equations with ease!]

Appendix

Centres of mass and moments of inertia

A.1 CENTRE OF MASS

Suppose we have a system $\mathcal{S}$ of particles $P_1, P_2, \ldots, P_N$ with masses $m_1, m_2, \ldots, m_N$, and position vectors $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N$ respectively. Then the centre of mass of $\mathcal{S}$ is defined as follows:

Definition A.1 ***Centre of mass*** *The* ***centre of mass*** *of $\mathcal{S}$ is the* ***point of space*** *whose position vector $\boldsymbol{R}$ is defined by*

$$\boxed{\boldsymbol{R} = \frac{\sum_{i=1}^{N} m_i \boldsymbol{r}_i}{\sum_{i=1}^{N} m_i} = \frac{\sum_{i=1}^{N} m_i \boldsymbol{r}_i}{M},} \tag{A.1}$$

where M is the total mass of $\mathcal{S}$.

In Chapter 2, we gave some simple examples of centre of mass, but here we will suppose the system is a **rigid body** $\mathcal{B}$. Then G, the centre of mass of $\mathcal{B}$, moves as if it were a particle of the body and its position can be calculated once and for all.

Finding centres of mass by symmetry

For most bodies, the position of G can be found by symmetry arguments so that no summations (or integrations) are needed. Instead, one simply applies the following rules which follow fron the definition (A.1) :

Rule 1: *If the body has* ***reflective symmetry*** *in a plane, then the centre of mass must lie in this plane.*

Rule 2: *If the body has any* ***rotational symmetry**** *about an axis, then the centre of mass must lie on this axis.*

* For a body to have a rotational symmetry, the mass distribution need not be preserved for *all* rotation angles about the axis. Just *one* angle (less than 2π) is enough.

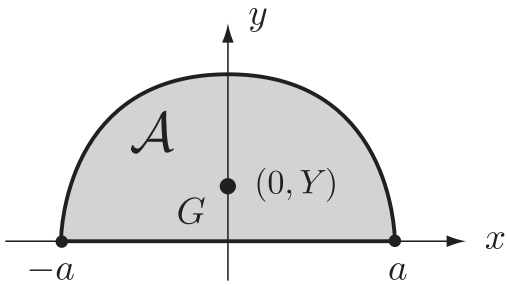


FIGURE A.1 A semi-circular lamina of radius a occupies the region $\mathcal{A}$ shown.

Suppose for example that the body $\mathcal{B}$ is a uniform **circular cylinder** occupying the region $x^2 + y^2 \le a^2$, $-b \le z \le b$. Then $\mathcal{B}$ has *reflective symmetry* in the plane $z = 0$ and so G lies in this plane. $\mathcal{B}$ also has full *rotational symmetry* about the z-axis and so G must lie on this axis. It follows that G must be at the origin. Similarly, if $\mathcal{B}$ is a uniform lamina in the shape of an **equilateral triangle**, $\mathcal{B}$ has *rotational symmetry* about each of its three medians. G must therefore lie at the intersection of the medians.

Not all bodies have enough symmetry to find G completely however. If $\mathcal{B}$ is a uniform **circular cone**, then it has full rotational symmetry about the axis connecting its vertex to the centre of its base. G must therefore lie somewhere on this axis but there is no other symmetry to determine exactly where. In such cases, a summation (or integration) becomes neccessary. This process is illustrated by the following two examples.

Example A.1 *Centre of mass of a semi-circular lamina*

Find the position of the centre of mass of a uniform semi-circular lamina.

Solution

Suppose that the lamina has radius a and occupies the region $\mathcal{A}$ shown in Figure A.1. The lamina has *rotational symmetry* about the y-axis so that G must lie on the y-axis, as shown. It remains to find its y-coordinate Y which, from the definition (A.1), is given by

$$MY = \sum_{i=1}^{N} m_i y_i, \tag{A.2}$$

where M is the mass of the lamina. However, since the lamina is a *continuous distribution* of mass, this sum must now be interpreted as an integral.

Consider an element of area dA of the lamina; this has mass $M dA/\left(\frac{1}{2}\pi a^2\right)$. The contribution to the sum in equation (A.2) from this element is therefore

$$\left(M dA/\left(\tfrac{1}{2}\pi a^2\right)\right) y,$$

where y is the y-coordinate of the element. On 'summing' these contributions (and cancelling by M) we find that

$$Y = \frac{2}{\pi a^2} \int_{\mathcal{A}} y \, dA, \tag{A.3}$$

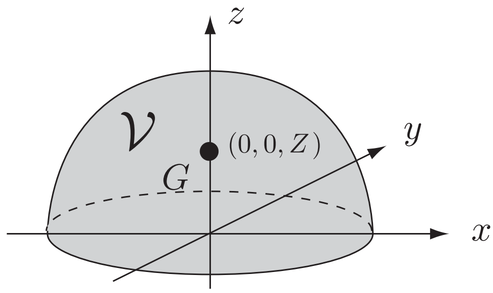


FIGURE A.2 A solid hemisphere of radius a occupies the region $\mathcal{V}$ shown.

where the integral is taken over the region $\mathcal{A}$ occupied by the lamina.

This double integral is most easily evaluated using standard polar coordinates. In these coordinates* $dA = (dr)(r\,d\theta) = r\,dr\,d\theta$, and $y = r\sin\theta$. The ranges of integration for r, θ are $0 \le r \le a$ and $0 \le \theta \le \pi$. We therefore obtain

$$\begin{aligned} Y &= \frac{2}{\pi a^2}\int_{r=0}^{r=a}\int_{\theta=0}^{\theta=\pi} r^2 \sin\theta\,dr\,d\theta \\ &= \frac{2}{\pi a^2}\left(\int_{r=0}^{r=a} r^2\,dr\right)\left(\int_{\theta=0}^{\theta=\pi}\sin\theta\,d\theta\right) \\ &= \frac{2}{\pi a^2}\left(\tfrac{1}{3}a^3\right)(2) = \frac{4a}{3\pi} \approx 0.424\,a. \end{aligned}$$

Hence the **centre of mass** of the semi-circular lamina is on the y-axis a distance $4a/3\pi$ from the origin. ■

Example A.2 *Centre of mass of a solid hemisphere*

Find the position of the centre of mass of a uniform solid hemisphere.

Solution

Suppose that the lamina has radius a and occupies the region $\mathcal{V}$ shown in Figure A.2. The hemisphere has full *rotational symmetry* about the z-axis so that G must lie on the z-axis, as shown. It remains to find its z-coordinate Z which, from the definition (A.1), is given by

$$MZ = \sum_{i=1}^{N} m_i z_i, \tag{A.4}$$

where M is the mass of the hemisphere. However, since the hemisphere is a *continuous distribution* of mass, this sum must now be interpreted as an integral.

Consider a volume element dv of the hemisphere; this has mass $M dv/\left(\frac{2}{3}\pi a^2\right)$. The contribution to the sum in equation (A.4) from this element is therefore

$$\left(M dv/\left(\tfrac{2}{3}\pi a^2\right)\right)z,$$

* See Figure 3.6 (right) for a diagram of the element of area in plane polars.

where z is the z-coordinate of the volume element. On 'summing' these contributions (and cancelling by M) we find that

$$Z = \frac{3}{2\pi a^2} \int_{\mathcal{V}} z\, dv, \tag{A.5}$$

where the integral is taken over the region $\mathcal{V}$ occupied by the hemisphere.

This volume integral is most easily evaluated using standard spherical polar coordinates r, θ, ϕ. In these coordinates* $dv = (dr)(r\, d\theta)(r \sin\theta\, d\phi) = r^2 \sin\theta\, dr\, d\theta\, d\phi$, and $z = r\cos\theta$. The ranges of integration for r, θ and ϕ are $0 \le r \le a$, $0 \le \theta \le \pi/2$ and $0 \le \phi \le 2\pi$. We therefore obtain

$$\begin{aligned} Z &= \frac{3}{2\pi a^2} \int_{r=0}^{r=a} \int_{\theta=0}^{\theta=\pi/2} \int_{\phi=0}^{\phi=2\pi} r^3 \sin\theta \cos\theta\, dr\, d\theta\, d\phi \\ &= \frac{3}{2\pi a^2} \left(\int_{r=0}^{r=a} r^3\, dr \right) \left(\int_{\theta=0}^{\theta=\pi/2} \sin\theta \cos\theta\, d\theta \right) \left(\int_{\phi=0}^{\phi=2\pi} d\phi \right) \\ &= \frac{3}{2\pi a^2} \left(\tfrac{1}{4} a^4 \right) \left(\tfrac{1}{2} \right) (2\pi) = \frac{3a}{8}. \end{aligned}$$

Hence the **centre of mass** of the hemisphere is on the z-axis a distance $3a/8$ from the origin. ■

A.2 MOMENT OF INERTIA

Suppose that we have the same system $\mathcal{S}$ as before and that CD is a straight line. Then the moment of inertia of $\mathcal{S}$ about the axis† CD is defined as follows:

Definition A.2 ***Moment of inertia*** *The* ***moment of inertia*** *of the system $\mathcal{S}$ about the axis CD is defined by*

$$\boxed{I_{CD} = \sum_{i=1}^{N} m_i\, p_i^{\,2}} \tag{A.6}$$

where p_i is the perpendicular distance of the mass m_i from the axis CD.

Any system has a moment of inertia about any axis, but here we will suppose that the system is a **rigid body** $\mathcal{B}$. Then the value of I_{CD} about an *embedded* axis is a constant

* See Figure 3.7 (right) for a diagram of the volume element in spherical polars.

† It is traditional to call CD an *axis* whether or not $\mathcal{S}$ is a rigid body rotating about CD.

and can be calculated once and for all. Calculation of moments of inertia required us to evaluate the sum (or integral) appearing in the definition (A.6). This process is illustrated by the following examples.

Example A.3 *Uniform rod (about a perpendicular axis through G)*

Find the moment of inertia of a uniform rod of mass M and length $2a$ about an axis through its centre and perpendicular to its length.

Solution

Consider an element of length dx of the rod; this has mass $M dx/2a$. The contribution to the sum in (A.6) from this element is therefore $p^2(M dx/2a)$, where p is the distance of the element dx from the specified axis. On 'summing' these contributions over all the elements of length of the disk, we find that

$$I_{CD} = \frac{M}{2a}\int_{-a}^{a} p^2\,dx = \frac{M}{2a}\int_{-a}^{a} x^2\,dx,$$

since $p = |x|$ for the specified axis. Hence

$$I_{CD} = \frac{M}{2a}\left(\tfrac{2}{3}a^3\right) = \tfrac{1}{3}Ma^2.$$

Hence the **moment of inertia** of the rod about the specified axis is $\frac{1}{3}Ma^2$. ■

Example A.4 *Uniform circular disk (about its axis of symmetry)*

Find the moment of inertia of a uniform circular disk of mass M and radius a about its axis of symmetry.

Solution

Consider an element of area dA of the disk; this has mass $M dA/\pi a^2$. The contribution to the sum in (A.6) from this element is therefore

$$\left(M dA/\pi a^2\right)p^2,$$

where p is the distance of the element from the symmetry axis. On 'summing' these contributions over all the elements of area of the disk we find that

$$I_{CD} = \frac{M}{\pi a^2}\int_{\mathcal{A}} p^2\,dA, \tag{A.7}$$

where the integral is taken over the region $\mathcal{A}$ occupied by the disk.

This double integral is most easily evaluated using standard plane polar coordinates r, θ. In these coordinates* $dA = (dr)(r\,d\theta) = r\,dr\,d\theta$, and $p = r$. The ranges

* See Figure 3.6 (right) for a diagram of the element of area in plane polars.

of integration for r, θ are $0 \le r \le a$ and $0 \le \theta \le 2\pi$. We therefore obtain

$$\begin{aligned} I_{CD} &= \frac{M}{\pi a^2} \int_{r=0}^{r=a} \int_{\theta=0}^{\theta=2\pi} r^3 \, dr \, d\theta \\ &= \frac{M}{\pi a^2} \left(\int_{r=0}^{r=a} r^3 \, dr \right) \left(\int_{\theta=0}^{\theta=2\pi} d\theta \right) \\ &= \frac{M}{\pi a^2} \left(\tfrac{1}{4} a^4 \right) (2\pi) = \tfrac{1}{2} M a^2. \end{aligned}$$

Hence the **moment of inertia** of the disk about its symmetry axis is $\frac{1}{2}Ma^2$. ■

Example A.5 *Uniform solid sphere (about any axis through G)*

Find the moment of inertia of a uniform solid sphere of mass M and radius a about an axis through its centre.

Solution

Consider a volume element dv of the sphere; this has mass $M dv / \left(\frac{4}{3}\pi a^2\right)$. The contribution to the sum in equation (A.6) from this element is therefore

$$\left(M dv / \left(\tfrac{4}{3} \pi a^3 \right) \right) p^2,$$

where p is the distance of the volume element dv from the axis CD. On 'summing' these contributions over all the volume elements of the sphere, we find that

$$I_{CD} = \frac{3M}{4\pi a^3} \int_{\mathcal{V}} p^2 \, dv, \tag{A.8}$$

where the integral is taken over the region $\mathcal{V}$ occupied by the sphere.

This volume integral is most easily evaluated using spherical polar coordinates r, θ, ϕ centred at the centre of the sphere, with the line $\theta = 0$ lying along the axis CD. In these coordinates* $dv = (dr)(r\, d\theta)(r \sin\theta \, d\phi) = r^2 \sin\theta \, dr \, d\theta \, d\phi$, and $p = r \sin\theta$. The ranges of integration for r, θ and ϕ are $0 \le r \le a$, $0 \le \theta \le \pi$ and $0 \le \phi \le 2\pi$. We therefore obtain

$$\begin{aligned} I_{CD} &= \frac{3M}{4\pi a^3} \int_{r=0}^{r=a} \int_{\theta=0}^{\theta=\pi} \int_{\phi=0}^{\phi=2\pi} r^4 \sin^3\theta \, dr \, d\theta \, d\phi \\ &= \frac{3M}{4\pi a^3} \left(\int_{r=0}^{r=a} r^4 \, dr \right) \left(\int_{\theta=0}^{\theta=\pi} \sin^3\theta \, d\theta \right) \left(\int_{\phi=0}^{\phi=2\pi} d\phi \right) \\ &= \frac{3M}{4\pi a^3} \left(\tfrac{1}{5} a^5 \right) \left(\tfrac{4}{3} \right) (2\pi) = \tfrac{2}{5} M a^2. \end{aligned}$$

Hence the **moment of inertia** of the sphere about an axis through its centre is $\frac{2}{5}Ma^2$.
■

* See Figure 3.7 (right) for a diagram of the volume element in spherical polars.

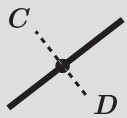
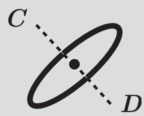
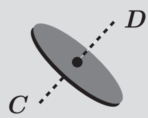
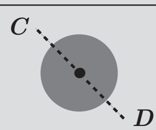
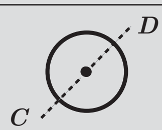
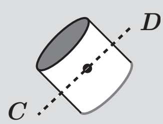
TABLE OF MOMENTS OF INERTIA		
Body	**Axis**	**Moment of inertia**
Thin rod mass M length $2a$		$I_{CD} = \frac{1}{3}Ma^2$
Circular hoop mass M radius a		$I_{CD} = Ma^2$
Circular disk mass M radius a		$I_{CD} = \frac{1}{2}Ma^2$
Solid sphere mass M radius a		$I_{CD} = \frac{2}{5}Ma^2$
Spherical shell mass M radius a		$I_{CD} = \frac{2}{3}Ma^2$
Circular cylinder mass M radius a length $2b$		$I_{CD} = \frac{1}{4}Ma^2 + \frac{1}{3}Mb^2$

Table 3 Some useful moments of inertia. All the bodies are uniform and in each case the axis CD passes through the centre of mass.

Table of moments of inertia

We can now build up a table of the most useful moments of inertia. Table 3 consists of the three results derived earlier together with one that is obvious (the circular hoop), and two that are new* (the spherical shell and the circular cylinder).

It may seem that there are some obvious omissions. There is no mention of the circular cylinder about its axis of symmetry, nor the rectangular plate about *any* axis. However, the table has been deliberately kept short to make it more memorable. Any other moments of inertia that are likely to be required can be quickly deduced from those in the table by the methods given below. We begin with a couple of use ful tricks.

* These two are less common than the others. They can be obtained by integration.

Two useful tricks

The **first trick** is based on the simple observation that the moment of inertia of a body is unaltered if its masses are moved in any manner *parallel* to the specified axis. In particular, the moment of inertia of a body is unaltered if it is squashed into a *lamina* perpendicular to the specified axis.

Suppose for example that the body is a uniform circular cylinder of radius a and length $2b$. Then its moment of inertia about its axis of symmetry is the same as that of a *uniform circular disk* of the same mass and radius about its own axis of symmetry. But this moment of inertia is known from Table 3 to be $\frac{1}{2}Ma^2$. Therefore the moment of inertia of the **circular cylinder** about its axis of symmetry is also $\frac{1}{2}Ma^2$. As a second example, suppose that the body is a uniform rectangular plate occupying the region $-a \le x \le a$, $-b \le y \le b$ of the (x, y)-plane. Then the moment of inertia of the plate about the x-axis is the same as that of a uniform rod of mass M and length $2b$ lying along the interval $-b \le y \le b$ of the y-axis. But this moment of inertia is known from the table to be $\frac{1}{3}Mb^2$. Therefore the moment of inertia of the **rectangular plate** about the x-axis is also $\frac{1}{3}Mb^2$. Similarly, the moment of inertia of the plate about the y-axis is $\frac{1}{3}Ma^2$. The moment of inertia of the plate about the z-axis will be deduced in the next section by using the perpendicular axes theorem.

To illustrate the **second trick**, suppose a uniform solid sphere of mass M and radius a is cut into two hemispheres. Then each of these hemispheres has mass $\frac{1}{2}M$ and (from the table) has moment of inertia $\frac{1}{2}\left(\frac{2}{5}Ma^2\right) = \frac{1}{5}Ma^2$ about its axis of symmetry, where M is the mass of the *whole* sphere. Hence, the moment of inertia of a **hemisphere** of mass M and radius a about its axis of symmetry must be $\frac{2}{5}Ma^2$, which is the same formula* as that for the complete sphere! *A similar argument applies whenever a body can be cut into a number of parts of equal mass that make equal contributions to the moment of inertia. The formula for the moment of inertia of each part is the same as that for the whole body.* Thus the moment of inertia of a segment of a *Terry's* chocolate orange about its straight edge is $\frac{2}{5}Ma^2$, where M is the mass of the segment and a is the radius of the orange.

A.3 PARALLEL AND PERPENDICULAR AXES

In this final section, we prove two important theorems that enable many more moments of inertia to be deduced. These are the **parallel axes** theorem and the **perpendicular axes** theorem.

Theorem A.1 ***Parallel axes theorem*** *Let I_G be the moment of inertia of a body about some axis through its centre of mass G, and let I be the moment of inertia of the body about a* ***parallel*** *axis. Then*

* It is the same formula, but remember that M is now the mass of the *hemisphere*.

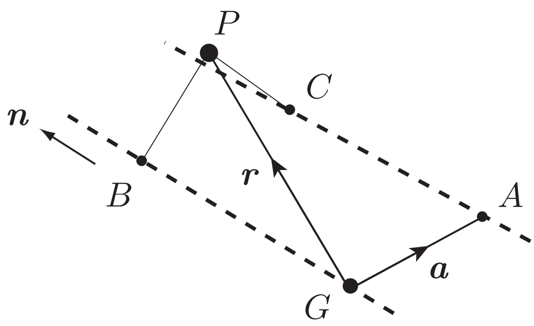


FIGURE A.3 The parallel axes theorem. The bold dashed lines are the two parallel axes.

$$\boxed{I = I_G + Ma^2} \tag{A.9}$$

where M is the mass of the body and a is the distance between the two parallel axes.

This is a powerful result. It means that, if you know the moment of inertia of a body about some axis, then you can easily calculate its moment of inertia about *any parallel axis*. The proof is as follows:

Proof. The configuration is shown in Figure A.3. The axis* $\{G, \boldsymbol{n}\}$ passes through G and $\{A, \boldsymbol{n}\}$ is the parallel axis. Here $\boldsymbol{n}$ is a unit vector parallel to the two axes, and A is chosen so that GA is perpendicular to the two axes. Then if $\overrightarrow{GA}$ represents the vector $\boldsymbol{a}$, the magnitude $|\boldsymbol{a}| = a$, the distance between the two axes.

Let P be a typical particle of the body with mass m and position vector $\boldsymbol{r}$ *relative to* G. Construct the perpendiculars PB, PC to the two axes, as shown. Then, by applications of Pythagoras,

$$PB^2 = GP^2 - GB^2 = |\boldsymbol{r}|^2 - (\boldsymbol{r} \cdot \boldsymbol{n})^2$$

and

$$\begin{aligned} PC^2 = AP^2 - AC^2 &= |\boldsymbol{r} - \boldsymbol{a}|^2 - \big((\boldsymbol{r} - \boldsymbol{a}) \cdot \boldsymbol{n}\big)^2 \\ &= (\boldsymbol{r} - \boldsymbol{a}) \cdot (\boldsymbol{r} - \boldsymbol{a}) - \big(\boldsymbol{r} \cdot \boldsymbol{n} - \boldsymbol{a} \cdot \boldsymbol{n}\big)^2 \\ &= |\boldsymbol{r}|^2 + |\boldsymbol{a}|^2 - 2\boldsymbol{r} \cdot \boldsymbol{a} - (\boldsymbol{r} \cdot \boldsymbol{n})^2 \\ &= PB^2 + a^2 - 2\boldsymbol{r} \cdot \boldsymbol{a}. \end{aligned}$$

If we now multiply this equality by m and sum over all the particles of the body, we obtain

$$\begin{aligned} I &= I_G + \left(\sum_{i=1}^{N} m_i\right) a^2 - 2\sum_{i=1}^{N} m_i\,\boldsymbol{r}_i \cdot \boldsymbol{a} = I_G + Ma^2 - 2\boldsymbol{a} \cdot \sum_{i=1}^{N} m_i\,\boldsymbol{r}_i \\ &= I_G + Ma^2 - 2\boldsymbol{a} \cdot (M\boldsymbol{R}), \end{aligned}$$

* The notation $\{G, \boldsymbol{n}\}$ means '*the axis through G parallel to the vector* $\boldsymbol{n}$'.

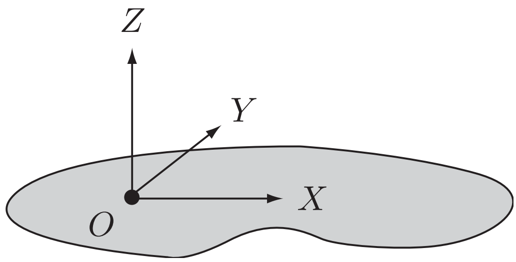


FIGURE A.4 The perpendicular axes theorem. The axes OX, OY lie in the plane of the lamina and OZ is perpendicular.

where $\boldsymbol{R}$ is the position vector of G. But since G is itself the origin of position vectors, $\boldsymbol{R} = \boldsymbol{0}$ and we obtain

$$I = I_G + Ma^2,$$

which is the required result. ■

Example A.6 *Rectangular plate (about an edge)*

A uniform rectangular plate has mass M and sides $2a$, $2b$. Find its moment of inertia about a side of length $2a$.

Solution

Suppose the plate occupies the region $-a \le x \le a$, $-b \le y \le b$ of the (x, y)-plane. It is already known that the moment of inertia of the plate about the x-axis is $\frac{1}{3}Mb^2$. The specified axis is parallel to the x-axis, the distance between the two axes being b. Hence the required **moment of inertia** is given by the parallel axes theorem to be $\frac{1}{3}Mb^2 + Mb^2 = \frac{4}{3}Mb^2$. ■

We will do more examples later, but first we will prove the perpendicular axes theorem.

Theorem A.2 ***Perpendicular axes theorem*** *Suppose the body is a* ***lamina*** *and that O is some point in its plane. Let $OXYZ$ be a set of Cartesian axes with OX, OY in the plane of the lamina and OZ perpendicular (see Figure A.4). Then the moments of inertia of the lamina about these* ***three perpendicular axes*** *are related by*

$$\boxed{I_{OZ} = I_{OX} + I_{OY}} \qquad \text{(A.10)}$$

that is, the moment of inertia of the lamina about the axis perpendicular to its plane is the ***sum*** *of the moments of inertia about the two in-plane axes.*

Proof. Let P be a typical particle of the lamina with mass m and coordinates $(x, y, 0)$ in the system $OXYZ$. Then p_X, p_Y, p_Z, the perpendicular distances of P from the axes OX, OY OZ are given by

$$p_X = |y|, \qquad p_Y = |x|, \qquad p_Z = \left(x^2 + y^2\right)^{1/2},$$

from which it follows that

$$p_Z^2 = p_X^2 + p_Y^2.$$

If we multiply this equality by m and sum over all the particles of the body, we obtain

$$I_{OZ} = I_{OX} + I_{OY},$$

which is the required result. ■

Example A.7 *Rectangular plate (about the perpendicular axis through G)*

A uniform rectangular plate has mass M and sides $2a$, $2b$. Find its moment of inertia about the axis through its centre perpendicular to its plane.

Solution

Let the axes $OXYZ$ be placed so that O is at the centre of the plate with OZ perpendicular to its plane. Then, since the plate is a lamina, it follows from the perpendicular axes theorem that

$$I_{OZ} = I_{OX} + I_{OY}.$$

But it is already known that $I_{OX} = \frac{1}{3}Mb^2$ and $I_{OY} = \frac{1}{3}Ma^2$ so that $I_{OZ} = \frac{1}{3}Mb^2 + \frac{1}{3}Ma^2 = \frac{1}{3}M(a^2 + b^2)$. Hence the required **moment of inertia** is $\frac{1}{3}M(a^2 + b^2)$. ■

Example A.8 *Circular disk (about a diameter)*

Find the moment of inertia of a uniform circular disk of mass M and radius a about a diameter.

Solution

Let the axes $OXYZ$ be placed so that O is at the centre of the disk with OZ perpendicular to its plane. Then, since the disk is a lamina, it follows from the perpendicular axes theorem that

$$I_{OZ} = I_{OX} + I_{OY}.$$

But I_{OZ} is known to be $\frac{1}{2}Ma^2$ and, by the rotational symmetry of the disk, $I_{OX} = I_{OY}$. It follows that $I_{OX} = I_{OY} = \frac{1}{4}Ma^2$. Hence the **moment of inertia** of the disk about any of its diameters is $\frac{1}{4}Ma^2$. ■

Example A.9 *Rectangular block (about any of its edges)*

A uniform rectangular block of mass M has edges of lengths $2a$, $2b$, $2c$. Find its moment of inertia about any edge.

Solution

Suppose the block occupies the region $-a \le x \le a$, $-b \le y \le b$, $-c \le z \le c$. We will first find its moment of inertia about each of the coordinate axes. Suppose the block is squashed parallel to the x-axis into a lamina of mass M occupying the rectangular region $-b \leq y \leq b$, $-c \leq z \leq c$ of the (y, z)-plane. The moment of inertia of this 'plate' about the x-axis is already known to be $\frac{1}{3}M(b^2 + c^2)$ and the block must have the same moment of inertia about the x-axis. Hence the **moments of inertia** of the block about the x- y- and z-axes are $\frac{1}{3}M(b^2 + c^2)$, $\frac{1}{3}M(a^2 + c^2)$ and $\frac{1}{3}M(a^2 + b^2)$ respectively.

Consider now an axis lying along an edge of length $2a$. This axis is parallel to the x-axis, the distance between the rwo axes being $(b^2 + c^2)^{1/2}$. Hence the moment of inertia of the block about this edge is given by the theorem of parallel axes to be $\frac{1}{3}M(b^2 + c^2) + M(b^2 + c^2) = \frac{4}{3}M(b^2 + c^2)$. Hence the **moments of inertia** of the block about the edges of lengths $2a$, $2b$, $2c$ are $\frac{4}{3}M(b^2 + c^2)$, $\frac{4}{3}M(a^2 + c^2)$, $\frac{4}{3}M(a^2 + b^2)$ respectively. ■

Answers to the problems

Chapter 1 Vectors (page 21)

1.1 (i) $18\,\boldsymbol{i} + 17\,\boldsymbol{j} - 26\,\boldsymbol{k},\ 6.$ (ii) $3,\ 5,\ \cos^{-1}(14/15).$ (iii) $1/3,\ -9/5.$
(iv) $4\,\boldsymbol{i} + 2\,\boldsymbol{j} + 3\,\boldsymbol{k},\ -20\,\boldsymbol{i} - 13\,\boldsymbol{j} - 15\,\boldsymbol{k},\ 9\,\boldsymbol{i} - 12\,\boldsymbol{k}.$
(v) +3. The set is *right*-handed. (vi) $-11\,\boldsymbol{i} + 70\,\boldsymbol{j} - 46\,\boldsymbol{k}.$

1.2 $\cos^{-1}(1/3)$

1.3 $\boldsymbol{c} = \boldsymbol{b} - \boldsymbol{a},\ \ \boldsymbol{d} = -\boldsymbol{a},\ \ \boldsymbol{e} = -\boldsymbol{b},\ \ \boldsymbol{f} = \boldsymbol{a} - \boldsymbol{b}.$

1.14 $6t\,\boldsymbol{i} + 3t^2\,\boldsymbol{j} + \boldsymbol{k},\ 6\,\boldsymbol{i} + 6t\,\boldsymbol{j},\ -6t(t+3)\,\boldsymbol{i} + 6(t+3)\,\boldsymbol{j} + 12t(t^2-2)\,\boldsymbol{k}.$

1.15 $2\,\boldsymbol{v}\cdot\dot{\boldsymbol{v}},\ \left(\dot{\boldsymbol{v}}\cdot\boldsymbol{k}\right)\boldsymbol{v} + (\boldsymbol{v}\cdot\boldsymbol{k})\,\dot{\boldsymbol{v}},\ \left[\boldsymbol{v},\,\ddot{\boldsymbol{v}},\,\boldsymbol{k}\right].$

1.16 $\boldsymbol{t} = -\sin\theta\,\boldsymbol{i} + \cos\theta\,\boldsymbol{j},\ \boldsymbol{n} = -\cos\theta\,\boldsymbol{i} - \sin\theta\,\boldsymbol{j},\ \kappa^{-1} = a.$

1.17 $\boldsymbol{t} = (-a\sin\theta\,\boldsymbol{i} + a\cos\theta\,\boldsymbol{j} + b\,\boldsymbol{k})\,/(a^2+b^2)^{1/2}$, $\boldsymbol{n} = (-\cos\theta\,\boldsymbol{i} - \sin\theta\,\boldsymbol{j}),\ \kappa^{-1} = (a^2 + b^2)/a.$

1.18 $\boldsymbol{t} = (p\,\boldsymbol{i} + \boldsymbol{j})\,/(1+p^2)^{1/2},\ \boldsymbol{n} = (\boldsymbol{i} - p\,\boldsymbol{j})\,/(1+p^2)^{1/2},\ \kappa^{-1} = 2a(1+p^2)^{3/2}.$

Chapter 2 Velocity, acceleration and scalar angular velocity (page 43)

2.1 $v = 12t - 3t^2$ m s^{-1}, $a = 12 - 6t$ m s^{-2}. P comes to rest when $t = 0$ and $t = 4$ s. At these times its displacement is 1 m and 33 m respectively.

2.2 $v = 3t^2 - 4t - 15$ m s^{-1}, $x = t^3 - 2t^2 - 15t + 20$ m. P comes to rest after 3 s. (The time $t = -5/3$ s is *before* the motion was set up.) At this time P has displacement -16 m.

2.3 If the acceleration were constant, then, from $v = u + at$, $a = 14.9$ ft s^{-2}. (Use feet and seconds as units.) However, from $s = ut + \frac{1}{2}at^2$, $a = 20.3$ ft s^{-2}. This indicates that the acceleration was *not* constant.

2.4 $|\boldsymbol{v}| = (\Omega^2 b^2 + c^2)^{1/2},\ \ |\boldsymbol{a}| = \Omega^2 b.$

2.5 0.034 m s^{-2} directed towards the centre of the Earth, 0.0060 m s^{-2} directed towards the Sun.

2.6 $\boldsymbol{v} = \left(\Omega b\, e^{\Omega t}\right)\widehat{\boldsymbol{r}} + \left(\Omega b\, e^{\Omega t}\right)\widehat{\boldsymbol{\theta}},\ \ \boldsymbol{a} = \left(2\Omega^2 b\, e^{\Omega t}\right)\widehat{\boldsymbol{\theta}}.$

2.7 $\cos^{-1}\left(b/(b^2 + \alpha^2 t^4)^{1/2}\right).$

2.8 $|\boldsymbol{a}| = v^2/(b\cos\alpha).$

2.9 $\boldsymbol{v} = 2b\tau^{-2}(\tau - t)\,\widehat{\boldsymbol{r}} + b\tau^{-3}t(2\tau - t)\,\widehat{\boldsymbol{\theta}}$. Minimum $|\boldsymbol{v}|$ occurs when $t = \tau$. When $t = \tau$, $\boldsymbol{a} = -3b\tau^{-2}\,\widehat{\boldsymbol{r}}.$

2.13 Maximum speed relative to ground is 120 mph. $|\boldsymbol{a}| = 121g.$

2.14 $|\boldsymbol{v}| = \Omega b(\cosh 2\Omega t)^{1/2},\ \ \boldsymbol{a} = (2\Omega^2 b\sinh\Omega t)\,\widehat{\boldsymbol{\theta}}.$

2.15 Maximum speed is ωe, maximum acceleration is $\omega^2 e$.

2.16 $\omega^2 b\,(1 - (b/c)).$

2.17 Speed of C is $\frac{1}{2}(\omega_1 b_1 + \omega_2 b_2)$. Angular velocity of gear is $(\omega_2 b_2 - \omega_1 b_1)/(b_2 - b_1)$. Angular velocity of arm is $(\omega_1 b_1 + \omega_2 b_2)/(b_1 + b_2)$.

2.18 The angular velocity $\omega = -\Omega b \cos\Omega t \left(a^2 - b^2 \sin^2 \omega t\right)^{-1/2}$, and the speed of C is $\Omega a b |\cos\Omega t| \left(a^2 - b^2 \sin^2 \omega t\right)^{-1/2}$.

2.19 Bearing is 18.6° west of north, time taken is 3 h 10 m.

2.20 $aU \sin\alpha \left(U^2 + u^2 + 2Uu\cos\alpha\right)^{-1/2}$.

2.21 Maximum range is $R_0(v^2 - u^2)^{1/2}/v$, achieved when $\theta = \pm\pi/2$.

Chapter 3 Newton's laws and gravitation (page 71)

3.1 $\sqrt{6}m^2G/a^2$, $\sqrt{6}m^2G/4a^2$.

3.2 $\left[\sqrt{3} + \sqrt{3/2} + 1/3\right] m^2G/a^2$, $4\left[1 + 1/\sqrt{2} + 1/(3\sqrt{3})\right] m^2G/a^2$.

3.3 $mMG/(x'^2 - a^2)$, $(M^2G/4a^2)\ln\left(b^2/(b^2 - 4a^2)\right)$.

3.4 $(2MM'G/a^2b)\left[a + b - (a^2 + b^2)^{1/2}\right]$.

3.6 $(mMG/a^3)r$.

3.7 $2\pi M\sigma G$.

3.8 $3M^2G/4a^2$.

Chapter 4 Problems in particle dynamics (page 98)

4.1 $\frac{1}{2}T_0$. $3T_0/2M$, $2T_0$.

4.2 $(m(b - z)/b + M)\,g$, $3\,(m(b - z)/b + M)\,g$.

4.4 $(\sin\alpha - \mu\cos\alpha)g$. If $\mu > \tan\alpha$, the motion will come to rest and remain at rest.

4.5 $20\,\mathrm{m\,s^{-1}}$.

4.6 1445 m.

4.7 Terminal velocity is $(U^2 + V^2)^{1/2}$, inclined at an angle $\tan^{-1}(U/V)$ to the downward vertical.

4.8 $\tan^{-1}(ebE_0/mu^2)$.

4.9 $(2h - g\tau^2)^2/8g\tau^2$.

4.10 $r = \left[a^{3/2} + 3(MG/2)^{1/2}\,t\right]^{2/3}$, which tends to infinity as t tends to infinity.

4.11 $\sqrt{2}a^2/\gamma^{1/2}$.

4.12 65 days.

4.14 The acceleratons of the three masses are $g/9$ *downwards*, $7g/9$ upwards and $5g/9$ downwards respectively.

4.15 $r = a\cosh\Omega t$, $N = 2ma\Omega^2\sinh\Omega t$.

4.17 Maximum height is $(V^2/2g)\ln\left[1 + (u^2/V^2)\right]$, time taken is $(V/g)\tan^{-1}(u/V)$.

4.18 Calculated descent times are 2.32 s and 2.27 s, which could hardly have been distinguished.

4.19 Maximum height is $(V^2/2g)\tan^{-1}(u^2/V^2)$.

4.20 Charge is $6\pi a\mu(v_1 + v_2)/E_0$, where the droplet radius $a = 3(\mu v_1/2\rho g)^{1/2}$.

4.21 Ranges are 55 m and 57 m respectively.

4.22 Projection angle is 60°.

4.23 Man is not safe. At 60 m away he would need to be at least 25 m high.

4.24 Least projection speed to clear hill is $\left(g(a^2+h^2)^{1/2}-gh\right)^{1/2}$. Least muzzle speed needed to cover the hill is 20 m s^{-1}, so this gun *can* do the job.

4.25 Landing point has displacement $(u^2/g)\,(\sin(2\alpha+\beta)-\sin\beta)\,/\cos^2\beta$ up the plane. $R^U = (u^2/g)(1-\sin\beta)/\cos^2\beta$, $\;R^D = (u^2/g)(1+\sin\beta)/\cos^2\beta$.

4.27 Path is an inclined straight line.

4.28 Estimated mass of Earth is 6.02×10^{24} kg. Calculation supposes the Earth fixed and so ignores the motion of the Earth induced by the Moon. Radius of geostationary orbit is approximately 42,000 km.

4.30 Tension is $(mg/\sqrt{2})+(mu^2/a)$, normal reaction is $(mg/\sqrt{2})-(mu^2/a)$, maximum u is $(ag/\sqrt{2})^{1/2}$. If u exceeds the maximum, the particle loses contact with the sphere.

4.31 Angle turned by string is u/Kb. Tension is $(mu^2/b)e^{-2Kt}$.

4.32 Particle of mass m meets y-axis again at the point $y=-2mV/eB_0$.

Chapter 5 Linear oscillations and normal modes (page 126)

5.1 Amplitude is 2, time taken is $\pi/6$.

5.2 In parallel, $\omega^2=\Omega_1^2+\Omega_2^2$. In series, $\omega^2=\Omega_1^2\Omega_2^2/(\Omega_1^2+\Omega_2^2)$.

5.3 Distance travelled is $a\left(1+e^{-\pi k/n}\right)$.

5.4 Greater than 8.

5.5 $\alpha=16.3$, $\beta=4.4$.

5.8 Lower block leaves floor after time $(2\pi/3)(a/g)^{1/2}$.

5.9 Angular frequencies less than 20 radians s^{-1} and greater than 40 radians s^{-1} are safe.

5.11 Amplitude is
$$\left(\frac{\Omega^2+4K^2p^2}{(\Omega^2-p^2)^2+4K^2p^2}\right)^{1/2} h_0.$$

5.12 Response is
$$x = 2F_0\sum_{n=1}^{\infty}\left(\frac{(-1)^{n+1}}{n}\right)\left(\frac{(\Omega^2-n^2)\sin nt+2Kn\cos nt}{(\Omega^2-n^2)^2+4K^2n^2}\right).$$

5.13 Period is $2(\pi+4)/\Omega$.

5.14 Body comes to rest at $x=+F_0/(2\Omega^2)$. Time taken is $2\pi/\Omega$.

5.15 Period is $\pi(\Omega+\Omega_D)/(\Omega\Omega_D)$, where $\Omega_D=(\Omega^2-K^2)^{1/2}$. Ratio of sucessive maxima is $e^{-K\pi/\Omega_D}$.

5.16 Slow mode: $\omega_1=n/\sqrt{6}$, $A/B=2/3$, where $n=(\alpha/m)^{1/2}$. Fast mode: $\omega_2=n$, $A/B=-1$. General motion is
$$x = 2\delta_1\cos(\omega_1 t-\gamma_1)+\delta_2\cos(\omega_2 t-\gamma_2),$$
$$y = 3\delta_1\cos(\omega_1 t-\gamma_1)-\delta_2\cos(\omega_2 t-\gamma_2),$$
where $\delta_1,\delta_2,\gamma_1,\gamma_2$ are arbitrary constants.

5.17 Slow mode: $\omega_1 = n$, $A/B = 1$, where $n = (T/ma)^{1/2}$. Fast mode: $\omega_2 = \sqrt{3}\,n$, $A/B = -1$. General motion is

$$x = \delta_1 \cos(\omega_1 t - \gamma_1) + \delta_2 \cos(\omega_2 t - \gamma_2),$$
$$y = \delta_1 \cos(\omega_1 t - \gamma_1) - \delta_2 \cos(\omega_2 t - \gamma_2),$$

where δ_1, δ_2, γ_1, γ_2 are arbitrary constants. General motion is not periodic.

5.18 Slow mode: $\omega_1 = \sqrt{2/3}\,n$, $A/B = 1/2$, where $n = (g/a)^{1/2}$. Fast mode: $\omega_2 = \sqrt{2}\,n$, $A/B = -1/2$.

Chapter 6 Energy conservation (page 151)

6.1 Both are 16 J which illustrates the energy priciple.

6.2 Man works at 500 W, which is quite a feat.

6.4 3.0×10^9 J.

6.5 Motion takes place in the interval $3 \le x \le 6$ and the period is $9\pi/\sqrt{2}$.

6.7 Particle penetrates a distance $\left(3mu^2/2K\right)^{1/3}$.

6.8 $V = -2mMG\left(a^2 + x^2\right)^{-1/2}$ and $v_{\max} = (4MG/5a)^{1/2}$.

6.9 $V = (2mMG/a^2)\left[z - (a^2 + z^2)^{1/2}\right]$ and speed on impact is $(8MG/3a)^{1/2}$.

6.10 Projection speed is $(\alpha a^2/4m)^{1/2}$.

6.11 Spring is compressed by $a/3$.

6.13 Minimum speed needed is $(g(a + 2h))^{1/2}$. When $h = \frac{1}{2}a$ the mortar should be placed a distance $a\left(\sqrt{3} - 1\right)/2$ from the wall of the building and inclined at 60° to the horizontal.

6.14 Escape speed is $\left(ae^2/m(b^2 - a^2)\right)^{1/2}$.

6.15 Escape speed is $\left(a^3e^2/mb^2(b^2 - a^2)\right)^{1/2}$.

6.16 (i) is stable while (ii) is unstable.

6.17 Arrival speed is $(4\pi bg)^{1/2}$, time taken is $\left(4\pi(a^2 + b^2)/gb\right)^{1/2}$.

6.18 When a/b is small, period is approximately

$$2\pi\left(\frac{b}{g}\right)^{1/2}\left[1 + \frac{a^2}{4b^2} + \cdots\right]$$

6.20 Time taken to hit post is $[(8/9) + (2/5)\ln 3]\,a/u$.

6.21 String makes 60° with *upward* vertical when it first becomes slack.

6.22 Car was parked about 5 m from edge of building.

Chapter 7 Orbits in a central field (page 188)

7.1 Distance of closest approach is $(p^2V^2 + \gamma)^{1/2}/V$.

7.2 $a = c$, $b = c\sin\alpha$.

7.4 $a = \sqrt{2}c\,\cos(\frac{1}{2}\alpha)$, $b = \sqrt{2}c\,\sin(\frac{1}{2}\alpha)$.

7.5 120°.

7.6 Deflection angle is $2\tan^{-1}(M_{\odot}G/pV^2)$ and distance of closest approach is $(p^2V^2 + M_{\odot}G)^{1/2}/V$.

7.7 Path is $r = (4p/15)/\sin(4\theta/15)$. Distance of closest approach is $4p/15$.

7.9 Path is $r = a\cos(\theta/3)$. Time taken to reach centre is $\pi a^2/(\sqrt{2}\gamma)$.

7.10 Apsidal angle is $\pi/(1-\epsilon)^{1/2}$.

7.11 Advance of perihelion is $2\pi\epsilon$ per 'year'.

7.12 Advance of perihelion is $-4\pi^2\epsilon$ per 'year'.

7.13 Advance of perihelion is $6\pi MG/(ac^2)$ per 'year'.

7.14 Differential scattering cross section is $a^2/4$, a constant.

7.15 Differential scattering cross section is

$$\sigma(\theta) = \frac{\pi^2\gamma(\pi-\theta)}{V^2\theta^2(2\pi-\theta)^2\sin\theta}.$$

7.16 Period was 89.6 minutes. Maximum speed was 7.84 km per second.

7.17 Apogee is 3910 km above the Earth. Apogee speed is 5.50 km per second. Period is 128 minutes.

7.18 Spacecraft will escape if $k \ge \sqrt{2}$. Eccentricity of escape orbit is $k^2 - 1$.

7.19 Spacecraft will go into orbit if

$$k^2 < \frac{2mG}{cu^2 + 2MG}.$$

7.20 Time average of kinetic energy is $\gamma/(2a)$.

7.22 $\Delta v^E = +3.13$ km per second, $\Delta v^M = +0.83$ km per second. Travel time = 119 hours.

7.23 The thrust should be applied in the direction of motion when the spacecraft is at its apogee.

7.24 The most fuel efficient one-impulse strategy is to use the first impulse of the Hohmann orbit. $\Delta v = -2.50$ km per second for Venus, and 2.95 km per second for Mars. Hence the Venus flyby uses less fuel. The travel time is 146 days to Venus, and 259 days to Mars.

7.25 The most fuel efficient one-impulse strategy is to apply the impulse parallel to the direction of motion, Δv being chosen to give an orbit with the correct period. For a synchronous orbit, $\Delta v = 2.94$ km per second. The apogee is 121,000 km from the Earth's surface.

7.26 In the given approximation, $r = (L_0^2/\gamma)e^{-2Kt}$ and $\dot{\theta} = (\gamma^2/L_0)e^{+3Kt}$. Hence $r\dot{\theta} = (\gamma/L_0)e^{+Kt}$, which *increases* as t increases.

7.27 In the given approximation, $r = (L_0/\Omega)^{1/2}e^{-Kt/2}$ and $\dot{\theta} = \Omega$. Hence $r\dot{\theta} = (\Omega L_0)^{1/2}e^{-Kt/2}$, which *decreases* as t increases.

Chapter 8 Non-linear oscillations and phase space (page 214)

8.1 When the oscillation has unit amplitude, $\omega = 1 - \frac{3}{8}\epsilon + O\left(\epsilon^2\right)$ and $x(t) = \cos s + \frac{1}{32}(\cos s - \cos 3s)\,\epsilon + O\left(\epsilon^2\right)$ where $s = \omega t$.

8.2 When the oscillation has unit amplitude, $\omega = 1 + \frac{5}{16}\epsilon + O\left(\epsilon^2\right)$.

8.3 When the maximum value achieved by $x(t)$ is unity, $x(t) = \cos s + \frac{1}{6}(-3 + 2\cos s + \cos 2s)\,\epsilon + O\left(\epsilon^2\right)$, where $s = \left(1 + O(\epsilon^2)\right)t$. The minimum value achieved by $x(t)$ is $-1 - \frac{2}{3}\epsilon + O\left(\epsilon^2\right)$.

8.4 At zero order, the limit cycle is a circle centre the origin and radius 2. Correct to order ϵ^2, $\omega = 1 - \frac{1}{16}\epsilon^2 + O\left(\epsilon^3\right)$, and correct to order ϵ, $x(t) = 2\cos s + \frac{1}{12}\left(-3\sin s + \sin 3s\right) + O\left(\epsilon^2\right)$.

8.9 The period of the limit cycle is 2π.

8.13 Correct to order ϵ, the driven response is

$$x(t) = -\frac{\cos pt}{p^2-1} + \left(-\frac{3\cos pt}{4(p^2-1)^4} - \frac{\cos 3pt}{4(p^2-1)^3(9p^2-1)}\right)\epsilon + O\left(\epsilon^2\right)$$

which is valid when $p \neq 1, 1/3, 1/5, \ldots$.

Chapter 9 The energy principle (page 241)

9.1 Equilibrium position is $\theta = \tan^{-1}(m/M)$.

9.2 The equilibrium positions are (i) vertically downwards (unstable), and (ii) inclined at 60° to the upward vertical (stable).

9.3 Angular frequency is $\left(4V_0/mb^2\right)^{1/2}$.

9.5 Acceleration of mass m up the plane, and the string tension, are

$$\left(\frac{M\sin\alpha - m\sin\beta}{M+m}\right)g \quad \text{and} \quad \frac{Mmg}{M+m}\left(\sin\alpha + \sin\beta\right).$$

9.6 Normal reactions on P and Q are

$$\tfrac{2}{3}mg\left(7\cos\theta + 2\sin\theta - 3\sqrt{2}\right) \quad \text{and} \quad \tfrac{1}{3}mg\left(4\cos\theta + 5\sin\theta - 3\sqrt{2}\right).$$

9.7 Speed of rope on leaving peg is $(ag)^{1/2}$

9.8 The velocity and acceleration of the free end are given by

$$v^2 = \left(\frac{x(2a-x)}{a-x}\right)g, \qquad \frac{dv}{dt} = \left(\frac{2a^2 - 2ax + x^2}{(a-x)^2}\right)g,$$

and the free end has acceleration $5g$ when $x = \frac{2}{3}a$.

9.9 Amplitude of oscillations is a, period is $4\pi(a/g)^{1/2}$.

9.10 Final speed of hoop is $\left(v^2 + gh\right)^{1/2}$.

9.11 Acceleration of ball is $\frac{5}{7}g\sin\alpha$.

9.12 Acceleration of yo-yo is $\frac{2}{3}g$.

9.13 Speed of roll when its radius is b is

$$\left(\frac{a^2V^2}{b^2} - \frac{4g(b^3-a^3)}{3b^2}\right)^{1/2}.$$

9.14 Period of small oscillations is

$$2\pi\left(\frac{a^2+3b^2}{3gb}\right)^{1/2}.$$

9.15 The speed of G is $\left(\frac{10}{7}g(a+b)(1-\cos\theta)\right)^{1/2}$.

9.16 Period of small oscillations is $2\pi\,(7(b-a)/5g)^{1/2}$.

9.17 Period of small oscillations is $2\pi\,\left(b^2/3ga\right)^{1/2}$.

Chapter 10 The linear momentum principle (page 279)

10.1 The time average of the apparent weight is Mg.

10.2 The time average of the total force applied by the juggler is $10Mg$. Juggling the balls as he crosses the bridge is worse than carrying them. The average force his feet apply to the bridge is $10Mg$ and, since this is certainly not a constant, there will be times when the force *exceeds* $10Mg$.

10.4 The reaction exerted by the support is

$$R = \frac{MG}{4a}\left(\frac{2a^2+2ax-3x^2}{a-x}\right).$$

The support will give way when $x = \frac{2}{3}a$.

10.5 The speed of the vertical part of the chain when it has been pulled up a distance x is given by $v^2 = g\left(a - \frac{2}{3}x\right)$.

10.6 The speed of G is given by $v^2 = \frac{10}{7}g(a+b)(1-\cos\theta)$.

10.8 The speed and height at burnout are $2100\,\mathrm{m\,s^{-1}}$ and $100\,\mathrm{km}$.

10.9 The maximum speed achieved is

$$\frac{u}{\epsilon}\left(1-\gamma^{-\epsilon}\right) \sim u\ln\gamma\left[1-\tfrac{1}{2}(\ln\gamma)\,\epsilon + O\left(\epsilon^2\right)\right]$$

when ϵ is small; $\gamma = M/m$.

10.13 Velocity of composite particle is $(m_1\boldsymbol{v}_1 + m_2\boldsymbol{v}_2)/(m_1+m_2)$.

10.15 The proton lost 14% of its energy and the recoil angle of the heium numcleus was 55°.

10.16 The mystery nucleus has atomic mass 16 and is therefore oxygen.

10.18 Neutrons of energy $\frac{1}{4}E$ will be found at angles of 30° and 60° to the incident beam. Neutrons of energy $\frac{3}{4}E$ will also be found at these angles.

10.20 The dark companion in Cygnus X-1 has a mass of about $16M_{\odot}$.

10.21 The emerging particles are scattered alpha particles and recoiling helium nucleii, which are identical. The angular distribution of emerging particles is

$$\frac{q^4}{E^2}\left(\frac{\cos\theta}{\sin^4\theta} + \frac{1}{\cos^3\theta}\right),$$

where θ is measured from the direction of the incident beam.

10.22 The scattering cross section for the incident neutrons is

$$\sigma^{TB} = \frac{A}{\pi}\cos\theta_1,$$

and the angular distribution of the recoiling protons is the same, that is, $(A/\pi)\cos\theta_2$.

10.24 The rod will hit the table after a time

$$\left(\frac{a}{3g}\right)^{1/2}\int_{\pi/3}^{\pi/2}\left(\frac{4-3\cos^2\theta}{1-2\cos\theta}\right)^{1/2}d\theta.$$

Chapter 11 The angular momentum principle (page 317)

11.2 Angular speed of target after impact is

$$\left(\frac{4mb}{Ma^2+4mb^2}\right)u.$$

11.3 Angular speed of cylinder after impact is

$$\left(\frac{2mb}{a^2(M+2m)}\right)u.$$

11.4 Angular speed of disk is $\frac{4}{5}(a^2/b^2)\Omega$ and increase in kinetic energy is

$$\frac{Ma^2}{25b^2}(4a^2-5b^2)\Omega^2.$$

11.5 Angle of the new conical motion is 84° approximately.

11.6 In the elastic case, the speed of the ball after impact is

$$\left(\frac{1-\beta}{1+\beta}\right)u, \quad \text{where} \quad \beta=\frac{m}{M}\left(1+\frac{b^2}{k^2}\right).$$

11.8 In Case A,

$$r=\left(a^2+u^2t^2\right)^{1/2}, \qquad \phi=\frac{\tan^{-1}(ut/a)}{\sin\alpha}.$$

In Case B, the required value of u is $\left(\frac{4}{3}ag\right)^{1/2}$.

11.9 After one lap by the bug, the hoop has rotated through the angle

$$\left(\frac{2m}{M+2m}\right)\pi.$$

11.10 Period of small oscillations is

$$2\pi\left(\frac{a^2+3b^2}{3gb}\right)^{1/2}.$$

11.11 Speed of ball after the onset of rolling is $\frac{5}{7}V$ and $\frac{2}{7}$ of the kinetic energy is lost in the process.

11.12 Ball will slide if $\tan\alpha > \frac{7}{2}\mu$ and will roll if $\tan\alpha < \frac{7}{2}\mu$. If ball slides, the acceleration is $g(\sin\alpha - \mu\cos\alpha)$. If ball rolls, the acceleration is $\frac{5}{7}g\sin\alpha$.

11.13 Bug will reach top of disk if

$$u^2 > \frac{8mag}{M+2m}.$$

11.14 Acceleration of yo-yo is $\frac{2}{3}g - \frac{1}{3}\ddot{Z}$ downwards.

11.15 Forward velocity of cylinder is

$$\left(\frac{k^2}{a^2+k^2}\right)V.$$

11.16

$$S = \tfrac{1}{4}\gamma(3\gamma-2)Mg\sin\theta, \qquad K = \tfrac{1}{2}\gamma^2(1-\gamma)Mga\sin\theta.$$

11.17 Reactions at B and C are $\left(\frac{1}{6}mg\right)\boldsymbol{i}$ and $\left(\frac{1}{6}mg\right)\boldsymbol{j}$. Reaction at the floor is $\frac{1}{6}mg\,(-\boldsymbol{i}-\boldsymbol{j}+\boldsymbol{k})$.

Chapter 12 Lagrange's equations and conservation principles (page 361)

12.1 Closed chain has N degrees of freedom. There are *four* conserved quantities (two horizontal components of linear momentum, the vertical component of angular momentum about G, and kinetic energy). Motion can be determined from conservation principles if $N \le 4$.

12.2 Upward acceleration of mass m is $2g/5$.

12.3 The accelerations of the three masses are (in order) $g/9$ *downwards*, $7g/9$ upwards and $5g/9$ downwards.

12.4 Period of small oscillations is $2\pi\,(4a/(3g\sin\alpha))^{1/2}$.

12.5 Normal reaction is $mg(7\cos\theta - 4)/3$. It is not realistic to assume that rolling persists until the normal reaction becomes zero. This would require an infinite coefficient of friction!

12.6 $p_x = (3/2)M\dot{x} + m(\dot{x} + b\,\dot{\theta}\cos\theta)$, which is *not* the horizontal linear momentum. Period of small oscillations is $2\pi(3Mb/(3M+2m)g)^{1/2}$.

12.7 Accelerations are (i) $2g\sin\alpha\cos\alpha/(7-2\cos^2\alpha)$, (ii) $5g\sin\alpha/(7-2\cos^2\alpha)$.

12.8 Reaction of the floor is $mg\left(4 - 6\cos\alpha\cos\theta + 3\cos^2\theta\right)/\left(1+3\sin^2\theta\right)^2$.

12.9 $p_\theta = mR^2\dot{\theta}$, which is the vertical component of angular momentum about O. Kinetic energy is not conserved because the force pulling the string down does work. Tension in the string is $m\left((L^2/R^3) - \ddot{R}\right)$.

12.10 Solution is $r = a\cosh\Omega t$. Energy function $h = \frac{1}{2}ma^2\left(\dot{r}^2 - \Omega^2 r^2\right) = -\frac{1}{2}ma^2\Omega^2$, which is constant.

12.11 Downward acceleration of yo-yo is $(2g - \ddot{Z})/3$, so yo-yo can remain at same height (or move with constant velocity) if $\ddot{Z} = 2g$. The total energy at time t is $M\left(\dot{Z}^2 + 2gZ\right)/6$.

12.12 Neither of E and h is conserved in general. (Consider, for example, the case in which $\dot{Z}$ is constant.)

12.13 Angle turned by hoop is $2\pi m/(M+2m)$.

12.14 $U = -f(t)\,x$.

12.15 The conserved momenta are $p_x = m\dot{x} + etz$ and $p_y = m\dot{y}$.

12.17 Conserved momentum $p_\theta = mr^2\dot{\theta}$, which is the component of angular momentum about O perpendicular to the plane of motion.

12.18 Conserved momentum $p_\phi = mr^2 \sin^2\alpha\,\dot{\phi}$, which is the angular momentum about the axis of the cone.

12.19 Conserved momenta are $p_\theta = mr^2\dot{\theta}$ and $p_z = m\dot{z} - (e\mu_0 I/2\pi)\ln r$.

12.20 In terms of obvious coordinates, the conserved quantities are (i) the vector $\boldsymbol{L}$ (ii) P_x, P_y and L_z (iii) L_z (iv) L_z (v) P_z and L_z.

12.21 Conserved quantity is $L_z + cP_z$.

Chapter 13 Calculus of variations and Hamilton's principle (page 388)

13.1 $x = t^4 + 2$.

13.2 $x = \sin t$.

13.4

$$z = \frac{1}{k}\ln\left(\frac{\cos kx}{\cos ka}\right).$$

13.5

$$\rho = \frac{a\cos\left(\frac{1}{2}\pi\,\sin\alpha\right)}{\cos\left(\theta\sin\alpha\right)}.$$

13.6 $x = t\,(2t + X - 8)\,/4$.

Chapter 14 Hamilton's equations and phase space (page 413)

14.1 $G = -v_1^2 + 3v_1v_2 - 2v_2^2 + 6wv_1 - 9wv_2 - 9w^2$.

14.2 The Hamiltonian is $H = p_\theta^2/(2m(a^2+b^2)) + mgb\,\theta$ and the equations are $\dot{\theta} = p_\theta/(m(a^2+b^2)),\ \ \dot{p}_\theta = -mgb$.

14.3 The Hamiltonian is $H = (p_x^2 + p_z^2)/2m + mgz$ and the equations are $\dot{x} = p_x/m,\ \ \dot{z} = p_z/m,\ \ \dot{p}_x = 0,\ \ \dot{p}_z = -mg$. The coordinate x is cyclic.

14.4 The Hamiltonian is

$$H = \frac{p_\theta^2}{2ma^2} + \frac{p_\phi^2}{2ma^2\sin^2\theta} - mga\cos\theta$$

and the equations are

$$\dot{\theta} = \frac{p_\theta}{ma^2},\quad \dot{\phi} = \frac{p_\phi}{ma^2\sin^2\theta},\quad \dot{p}_\theta = \frac{p_\phi^2\cos\theta}{ma^2\sin^3\theta} - mga\sin\theta,\quad \dot{p}_\phi = 0.$$

14.5 The Hamiltonian is

$$H = \frac{a^2p_x^2 - 2a\cos\theta p_x p_\theta + 2p_\theta^2}{2ma^2\left(2-\cos^2\theta\right)} - mga\cos\theta.$$

14.6 The Hamiltonian is

$$H = \frac{p_\theta^2}{2m(a-Z)^2} - \tfrac{1}{2}m\dot{Z}^2 - mg(a-Z)\cos\theta$$

and the equations are $\dot{\theta} = p_\theta/m(a-Z)^2$, $\dot{p}_\theta = -mg(a-Z)\sin\theta$. Since H has an explicit time dependence through $Z(t)$, it is not conserved.

14.7 The Hamiltonian H is conserved when the fields $\{\boldsymbol{E}, \boldsymbol{B}\}$ are static.

Chapter 15 General theory of small oscillations (page 452)

15.1 $\omega_1^2 = \alpha/6m$, $\omega_2^2 = 3\alpha/4m$, $\boldsymbol{a}_1 = (2, 3)$, $\boldsymbol{a}_2 = (4, -1)$. In the subsequent motion,

$$x = \frac{2u}{14\omega_1\omega_2}\left(\omega_2 \sin\omega_1 t + 6\omega_1 \sin\omega_2 t\right), \qquad y = \frac{3u}{14\omega_1\omega_2}\left(\omega_2 \sin\omega_1 t - \omega_1 \sin\omega_2 t\right).$$

15.2 $\omega_1^2 = \alpha/6m$, $\omega_2^2 = \alpha/m$, $\boldsymbol{a}_1 = (2, 3)$, $\boldsymbol{a}_2 = (1, -1)$. The variables $\eta_1 = x + y$, $\eta_2 = 3x - 2y$ are a set of normal coordinates.

15.3 $\omega_1^2 = g/2a$, $\omega_2^2 = 15g/2a$, $\boldsymbol{a}_1 = (5, 6)$, $\boldsymbol{a}_2 = (3, -2)$. The ratio $\tau_1/\tau_2 = \sqrt{15}$, which is irrational; so general small motion is *not* periodic.

15.4 $\omega_1^2 = g/2a$, $\omega_2^2 = 3g/2a$, $\omega_3^2 = 3g/a$, $\boldsymbol{a}_1 = (1, 2, 3)$, $\boldsymbol{a}_2 = (1, 0, -3)$, $\boldsymbol{a}_3 = (1, -3, 3)$. The variables $\eta_1 = 3\theta + 2\phi + \psi$, $\eta_2 = 3\theta - \psi$, $\eta_3 = 3\theta - 3\phi + \psi$ are a set of normal coordinates.

15.5 The longitudinal modes have frequencies $\omega_1^2 = \alpha/m$ and $\omega_2^2 = 3\alpha/m$, and the transverse modes have frequencies $\omega_3^2 = T_0/ma$ and $\omega_4^2 = 3T_0/ma$.

15.6 $\omega_1^2 = 4g/9a$, $\omega_2^2 = g/a$, $\omega_3^2 = 4g/a$. Then $\tau_2/\tau_1 = 3/2$ and $\tau_2/\tau_1 = 3$ which are both rational. The general small motion is periodic with period $6\pi(a/g)^{1/2}$.

15.7 The longitudinal mode has frequency $\omega_1^2 = 11g/48a$, and the transverse modes have frequencies $\omega_2^2 = g/6a$ and $\omega_3^2 = 3g/4a$.

15.8 The transverse modes have frequencies $\omega_2^2 = 12T_0/5Ma$ and $\omega_3^2 = 18T_0/Ma$.

15.9 The j-th normal frequency is

$$\omega_j = 2\left(\frac{\alpha}{m}\right)^{1/2} \sin\left(\frac{j\pi}{2(n+1)}\right) \qquad (1 \le j \le n).$$

15.10 The j-th normal frequency is

$$\omega_j = 2\left(\frac{T_0}{ma}\right)^{1/2} \sin\left(\frac{j\pi}{2(n+1)}\right) \qquad (1 \le j \le n).$$

15.11 Vibrational frequencies are $\omega_1^2 = \alpha/m$, $\omega_2^2 = 6\alpha/m$. Estimated frequency ratio is 2.54, remarkably close to the measured value of 2.49.

15.12 The antisymmetric mode has frequency $\omega_3^2 = 7k/4m$, and the symmetric modes have frequencies $\omega_1^2 = k/4m$ and $\omega_2^2 = 2k/m$.

Chapter 16 Vector angular velocity and rigid body kinematics (page 467)

16.1 Angular velocity is $2\boldsymbol{k}$ radians per second, particle velocity is $6\boldsymbol{i} + 8\boldsymbol{j}\,\mathrm{m\,s^{-1}}$, speed is $10\,\mathrm{m\,s^{-1}}$, and acceleration is $-16\boldsymbol{i} + 12\boldsymbol{j}\,\mathrm{m\,s^{-2}}$.

16.2 Angular velocity is $2\boldsymbol{i}+2\boldsymbol{j}+\boldsymbol{k}$ radians per second, particle velocity is $3\boldsymbol{i}-6\boldsymbol{j}+6\boldsymbol{k}\,\mathrm{cm\,s^{-1}}$, speed is $9\,\mathrm{cm\,s^{-1}}$, and acceleration is $18\boldsymbol{i} - 9\boldsymbol{j} - 18\boldsymbol{k}\,\mathrm{cm\,s^{-2}}$.

16.4 Angular velocity of penny is $\dot\theta\,(\boldsymbol{k} - \cos\alpha\,\boldsymbol{a})$. Velocity of highest particle is zero.

16.5 Highest particle has the greatest speed which is $2h\cos^2\alpha\,|\dot\theta\,|$.

Chapter 17 Rotating reference frames (page 489)

17.3 Angular speed is $\lambda\sec\alpha$.

17.4 $r = a\cosh\Omega t$.

17.5 Take the angular velocity of the rotating frame to be $\boldsymbol{\Omega} = -(eB/2mc)\boldsymbol{k}$. The two frequencies are $\omega_0 \pm (eB/2mc)$.

17.10 $r = a\cosh\left(\sqrt{\tfrac{2}{3}}\,\Omega\,t\right)$.

17.11 Speed of cylinder is $\Omega x/\sqrt{2}$. Turntable exerts force $-\frac{1}{2}M\Omega^2 x\,\boldsymbol{i} + \sqrt{2}M\Omega^2 x\,\boldsymbol{j} + Mg\,\boldsymbol{k}$ on the cylinder, where $\boldsymbol{i}$ points in the direction of motion and $\boldsymbol{k}$ points vertically upwards.

17.12 In cylindrical coordinates with Oz along the rotation axis of the bucket, the isobars are the family of surfaces $z = (\Omega^2/2g)r^2 + c$, where c is a constant. The free surface of the water must be one of these, whatever the shape of the container.

17.13 In cylindrical coordinates, the water occupies the region $\frac{1}{2}a \le r \le a$. The pressure at the wall is $p_0 + \frac{3}{8}\rho\Omega^2 a^2$.

Chapter 18 Tensor algebra and the inertia tensor (page 519)

18.1 $\mathcal{C}$ and $\mathcal{C}'$ have opposite handedness.

18.2 The transformation matrices are

$$\text{(i)}\ \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 0 & -1\\ 0 & \sqrt{2} & 0\\ 1 & 0 & 1\end{pmatrix}\quad \text{(ii)}\ \begin{pmatrix}1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 1\end{pmatrix}\quad \text{(iii)}\ \frac{1}{9}\begin{pmatrix}4 & 7 & -4\\ 1 & 4 & 8\\ 8 & -4 & 1\end{pmatrix}\quad \text{(iv)}\ \frac{1}{9}\begin{pmatrix}1 & 4 & -8\\ 4 & 7 & 4\\ -8 & 4 & 1\end{pmatrix}$$

The new coordinates of the point D are $(3/\sqrt{2},\ -3,\ 3/\sqrt{2})$, $(3,\ 3,\ 0)$, $(-1,-1,\ 4)$, $(-1,-1,-4)$ respectively.

18.3 Solution for v is $(1, 1, 0)'$ or any multiple of it. The rotation angle about OE is $\pi/3$.

18.4 The transformation formula for a fifth order tensor is

$$t'_{ijklm} = \sum_{p=1}^{3}\sum_{q=1}^{3}\sum_{r=1}^{3}\sum_{s=1}^{3}\sum_{t=1}^{3} a_{ip}\,a_{jq}\,a_{kr}\,a_{ls}\,a_{mt}\,t_{pqrst}$$

18.5 The matrices representing the tensor are

$$\text{(i)}\ \frac{1}{\sqrt{2}}\begin{pmatrix}\sqrt{2} & 1 & 1\\ 1 & \sqrt{2} & 0\\ 1 & 0 & \sqrt{2}\end{pmatrix}\qquad \text{(ii)}\ \begin{pmatrix}1 & 0 & -1\\ 0 & 1 & 0\\ -1 & 0 & 1\end{pmatrix}.$$

18.6 (i) Yes, order 7, (ii) Yes, order 6, (iii) Yes, order 1 (a vector), (iv) Yes, order 1 (a vector), (v) No, (vi) Yes, order 5, (vii) Yes, order zero (a scalar), (viii) No, (ix) Yes, order zero (a scalar).

18.9 The most general material has 21 elastic moduli.

18.10 Moment of inertia of the plate about a diagonal is $\frac{2}{3}Ma^2b^2/(a^2+b^2)$.

18.11 Principal moments of inertia of the disk are (i) $\frac{1}{4}Ma^2$, $\frac{1}{4}Ma^2$, $\frac{1}{2}Ma^2$, (ii) $\frac{5}{4}Ma^2$, $\frac{1}{4}Ma^2$, $\frac{3}{2}Ma^2$.

18.12 Principal moments of inertia of the top at A are $\frac{5}{4}Ma^2$, $\frac{5}{4}Ma^2$, $\frac{1}{2}Ma^2$.

18.13 Principal moments of inertia of the top at A are $\frac{43}{20}Ma^2$, $\frac{43}{20}Ma^2$, $\frac{2}{5}Ma^2$.

18.14 Principal moments of inertia of the cube are (i) $\frac{2}{3}Ma^2$, $\frac{2}{3}Ma^2$, $\frac{2}{3}Ma^2$ (ii) $\frac{5}{3}Ma^2$, $\frac{5}{3}Ma^2$, $\frac{2}{3}Ma^2$ (iii) $\frac{11}{3}Ma^2$, $\frac{11}{3}Ma^2$, $\frac{2}{3}Ma^2$. The required moments of inertia are (i) $\frac{2}{3}Ma^2$, (ii) $\frac{5}{3}Ma^2$, (iii) $\frac{8}{3}Ma^2$.

18.15 Principal moments of inertia are (i) $\frac{1}{3}M(b^2+c^2)$, $\frac{1}{3}M(a^2+c^2)$, $\frac{1}{3}M(a^2+b^2)$ (ii) $\frac{1}{3}M(b^2+4c^2)$, $\frac{1}{3}M(a^2+4c^2)$, $\frac{1}{3}M(a^2+b^2)$. The required moments of inertia are

$$\text{(i)} \quad \frac{2M(a^2b^2+a^2c^2+b^2c^2)}{3(a^2+b^2+c^2)} \qquad \text{(ii)} \quad \frac{2M(a^2b^2+2a^2c^2+2b^2c^2)}{3(a^2+b^2)}$$

18.16 Principal moments of inertia at G are $\frac{1}{4}Ma^2+\frac{1}{3}Mb^2$, $\frac{1}{4}Ma^2+\frac{1}{3}Mb^2$, $\frac{1}{2}Ma^2$. Cylinder is dynamically spherical when $b=\sqrt{3}\,a/2$.

18.17 Symmetries are (i) axial, (ii) none, (iii) none, (iv) axial, (v) axial, (vi) spherical, (vii) axial, (viii) axial, (ix) spherical.

18.18 Principal moments of inertia at a corner point of the plate are $\frac{2}{3}(5+3\sqrt{2})Ma^2$, $\frac{2}{3}(5-3\sqrt{2})Ma^2$, $\frac{20}{3}Ma^2$. A set of principal axes is obtained by rotating the axes $Cx_1x_2x_3$ shown in Figure 18.3 through an angle of $-\pi/8$ about the axis Cx_3.

Chapter 19 Problems in rigid body dynamics (page 560)

19.2 Radius of circle is $\frac{7}{2}b$ and centre is a distance $\frac{5}{2}b$ from O.

19.3 Ball is stable at highest point if $n^2 > 35(a+b)g/a^2$.

19.4 Precession rate is Mgh/Cn.

19.6 For upright stability, the spin would have to be at least 3860 revolutions per second!

19.7 Spin nedded for stability is (i) 9.3 and (ii) 6.2 revolutions per second. The juggler should use a hollow ball.

19.11 The required moment is $\frac{1}{2}Ma^2\Omega n$ about the horizontal axis *perpendicular* to the symmetry axis of the wheel.

19.14 The motion of the axis is approximately periodic with period π/Λ.

Bibliography

This bibliography is not a survey of the mechanics literature. The books that appear below are included simply because they are referred to at some point in the text. Thus, no inference should be drawn regarding books that do not appear.

Classical mechanics

[1] Acheson, D., *From Calculus to Chaos: An Introduction to Dynamics*, Oxford University Press, 1997.
Not a mechanics textbook as such, but a selection of topics leading, as the title says, from calculus to chaos. Lots of interesting problems and photographs, including one of the Indian Rope Trick!

[2] Arnold, V.I., *Mathematical Methods of Classical Mechanics*, Springer-Verlag, 1978.
An elegant exposition of classical mechanics expressed in the language of differential topology. A classic, but you need to know classical mechanics *and* differential topology before reading it.

[3] French, A.P., *Newtonian Mechanics*, W.W. Norton & Co., 1971.
A leisurely introduction to classical mechanics from the Newtonian standpoint. Very good on the foundations and the background physics.

[4] Goldstein, H., Poole, C. & Safko, J., *Classical Mechanics*, third edition, Addison-Wesley, 2002.
Goldstein's classic graduate-level text, skilfully revised and extended by Poole & Safko. This is the first place to look for material not included in this book.

[5] Kaplan, M.H., *Modern Spacecraft Dynamics and Control*, John Wiley & Sons, 1976.
A good source for applications of classical mechanics to spacecraft orbits and manoeuvers.

[6] Landau, L.D. & Lifshitz, E.M., *Mechanics (Course of Theoretical Physics, Volume I)*, third edition, Butterworth-Heinemann, 1995.
Volume I of this famous *Course of Theoretical Physics* is a terse account of analytical mechanics. A classic, but not easy to follow.

Mathematics

[7] Anton, H., *Elementary Linear Algebra*, John Wiley & Sons, 2005.
An enduring introductory text on linear algebra. It has a good chapter on eigenvalues and eigenvectors, but does not *quite* prove the orthogonal diagonalisation theorem for symmetric matrices. For a full proof see Lang [10].

[8] Boyce, W. E. & DiPrima, R. C., *Elementary Differential Equations and Boundary Value Problems*, eighth edition, John Wiley & Sons, 2004.
One of several good textbooks on the *solution* of ODEs. It contains all the solution methods for ODEs that are used in this book.

[9] Coddington, E.A. & Levinson, N. *Theory of Ordinary Differential Equations*, McGraw-Hill, 1984.
The ultimate authority on the *theory* of ODEs. In particular, it has a correct proof of the Poincaré-Bendixson theorem in all its cases. The authors do everything as it should be done, but they take no prisoners.

[10] Lang, S., *Linear Algebra*, Springer-Verlag, 1993.
Contains a proof of the important result that *any* real symmetrtic matrix has a full set of eigenvectors (the orthogonal diagonalisation theorem).

[11] Schey, H.M., *Div, Grad, Curl and All That: An Informal Text on Vector Calculus*, W.W. Norton & Co., 1996.
A very readable book on vector calculus.

[12] Walters, P. *An Introduction to Ergodic Theory*, Springer-Verlag, 2000.
Despite its title, this is a graduate text on ergodic theory. In particular, it develops the theory of Poincaré recurrence further than we were able to do in this book.

Physics and chemistry

[13] Herzberg, G., *Molecular Spectra and Molecular Structure, Volume II: Infrared and Raman Spectra of Polyatomic Molecules*, Krieger Publishing Co., 1990.
This is the second part of Herzberg's epic trilogy on molecular spectra. It contains the classical theory of molecular vibrations, including group theoretic methods, and also has detailed analyses of the vibrational spectra of many common polyatomic molecules.

[14] Kaye, E.W.C., & Laby, T.H., *Tables of Physical and Chemical Constants*, Longman, 1986.
The title says it all.

Index

A page number printed in **bold** indicates that a definition or an important statement about the index entry appears on that page.