N. M. J. Woodhouse

SPRINGER UNDERGRADUATE
MATHEMATICS SERIES

Introduction to Analytical Dynamics

New Edition

Springer

Springer Undergraduate Mathematics Series

Advisory Board

For other titles published in this series, go to
www.springer.com/series/3423

N.M.J. Woodhouse

Introduction to Analytical Dynamics

New Edition

Springer

Nicholas M.J. Woodhouse
Mathematical Institute
University of Oxford
24-29 St Giles'
Oxford, OX1 3LB
United Kingdom
nwoodh@maths.ox.ac.uk

Springer Undergraduate Mathematics Series ISSN 1615-2085
ISBN 978-1-84882-815-5 e-ISBN 978-1-84882-816-2
DOI 10.1007/978-1-84882-816-2
Springer London Dordrecht Heidelberg New York

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Control Number: 2009941530

Mathematics Subject Classification (2000): 70HXX, 03,05,15,25, 99

First Edition published by Oxford University Press, 1987

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface to the New Edition

This is a revised edition of a text on classical mechanics that was originally published 20 years ago by Oxford University Press. I have taken the opportunity to simplify some of the presentation, while keeping to the original intention of confronting rather than evading the various notational and pedagogical difficulties that one encounters in the journey from Newton to Lagrange and Hamilton. I have also responded to comments over the years from colleagues and, more recently, from the new publisher's referees.

There are two major changes. I have gathered together the material and examples on systems with one degree of freedom into a separate chapter. The intention here is to give a first introduction to the core ideas of the Lagrangian theory in a context in which they make strong contact with familiar elementary techniques from the treatment of ordinary differential equations, without the distraction of indices and the summation convention. Second, I have added a chapter on differential geometric methods.

I am grateful to the many students and colleagues who have commented on the first edition, and pointed out mistakes.

N.M.J. Woodhouse
Oxford, February 2009

Preface to the First Edition

It may seem odd that Newtonian mechanics should still hold a central place in the university mathematics curriculum. But there are good reasons.

– It is one of the most accurate physical theories ever devised. Three hundred years after the publication of Newton's *Philisophiae naturalis principia mathematica* (1687), we should be surprised not that some of his ideas have been superseded by relativity and quantum theory, but that it is still necessary to exercise great subtlety and scientific ingenuity to detect any error at all in the three laws of motion. Even in the prediction of the orbit of the planet Mercury, for example, which was a crucial failure of the classical theory, the discrepancy[1] is only one part in 10^7.

  Newton's theory is the prime example of what Wigner calls the 'unreasonable effectiveness of mathematics' as a tool for understanding the physical world – an aspect of the truth of mathematics that can easily be lost in a course overburdened with abstraction [14].

– Quantum theory and relativity have overthrown the classical view of physics, but the mathematical formalism of classical mechanics still plays an essential part. It provides both a framework for interpretation and a first introduction to key ideas and techniques (frames of reference, general coordinate transformations, the connection between observables and symmetries,...). It is an essential prerequisite for any advanced course on applications of mathematics in modern theoretical physics.

– It develops geometric intuition and gives invaluable practice in problem

[1] The radius vector from the Sun to Mercury sweeps out a total angle of $150,000^\circ$ per century. The prediction of the Newtonian theory is $43''$ less than the observed angle. For the other planets, it is much less.

solving and mathematical modelling. It is easy to poke fun at the seemingly endless supply of light rods, inextensible strings, and smooth hemispheres. But all undergraduate exercises are necessarily artificial, however cleverly they are dressed. The strength of mechanics is the vast range of its examples, something that their familiarity can make us overlook, and the diversity of different mathematical ideas that they illustrate.

– The problems of classical mechanics and, in particular, the centuries of work on planetary motion, stimulated the development of much of modern mathematics. It is no coincidence that the great names of mechanics, Newton, Euler, Poisson, Lagrange, Hamilton,..., also occur over and over again throughout many branches of pure mathematics. It is essential to study classical mechanics to understand the roots of mathematics.

– The influence of classical mechanics is still present in modern pure mathematical research. The study of Hamilton's equations, for example, led to the development of symplectic geometry, which in turn has found recent applications in the analysis of partial differential equations and in the representation theory of Lie groups.

A glance through the pages that follow will not reveal anything strikingly unfamiliar. The range of topics is central and traditional, partly because I want the book to be short and (OUP willing) cheap, and partly because I intend it to be no more than an introduction. I hope that it will be read in conjunction with the classics and that it will encourage further exploration (in, for example, Arnol'd's *Mathematical methods of classical mechanics* [1]).

The book is written for second year mathematics undergraduates and assumes familiarity with elementary linear algebra, the chain rule for partial derivatives, and vector mechanics in three dimensions (the last is not absolutely essential). The main intention is, first, to give a confident understanding of the chain of argument that leads from Newton's laws through Lagrange's equations and Hamilton's principle to Hamilton's equations and canonical transformations; and, second, to give practice in problem solving. Most of the exercises and examples are taken from recent Oxford examination papers.

I have concentrated on trying to clarify the points that come up most frequently in tutorials and that I myself found confusing when I first met these ideas. For example: why are you allowed to say that q and $\dot{q}$ are independent? and: why can I not deduce from $\partial L/\partial t = -\partial h/\partial t$ that $h + L$ is independent of t?

It is true, of course, that the most satisfactory way to come to terms with the mathematics of classical mechanics would be to approach the subject from modern differential geometry. But that would mean reducing analytical mechanics to a minority option at the end of the undergraduate course or in the

first year of graduate work, which would be a great loss. Instead, I have tried to make use of lessons that I have learnt from differential geometry, but without ever going outside the framework of local, coordinate-based arguments.

I am particularly grateful to Paul Tod and Tom Cooper for many comments on an earlier version of this book; and to Rob Baston, Andy Clark, Mike Dobson, Steve Lloyd, Diana Mountain, Charles Sanderson, and Steve Thorsett for working through the final version.

Oxford 1986
N.M.J. Woodhouse

Note. Examples and exercises marked with a dagger (†) are adapted from examination questions set at the University of Oxford.

Contents

1
Frames of Reference

1.1 Introduction

The solution to a mechanical problem begins with the kinematic analysis, the analysis of how a system can move, as opposed to how it actually does move under the influence of a particular set of forces. In this first stage, the essential step is the introduction of coordinates to label the configurations of the system. These might be Cartesian coordinates for the position of a particle, or angular coordinates for the orientation of a rigid body, or some complicated combination of distances and angles. The only conditions are that each physically possible configuration should correspond to a particular set of values of the coordinates; and that, conversely, the coordinates should be *independent*, which can be understood informally to mean that each set of values of the coordinates should determine a unique configuration. The number of coordinates is called the *number of degrees of freedom* of the system.

Example 1.1

A particle moving in space has three degrees of freedom. Its position is determined by three Cartesian coordinates or by three spherical polar coordinates.

Example 1.2

A particle moving on the surface of a sphere has two degrees of freedom, labelled by the two polar angles θ and φ (Figure 1.1)

N.M.J. Woodhouse, *Introduction to Analytical Dynamics: New Edition*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-1-84882-816-2_1,

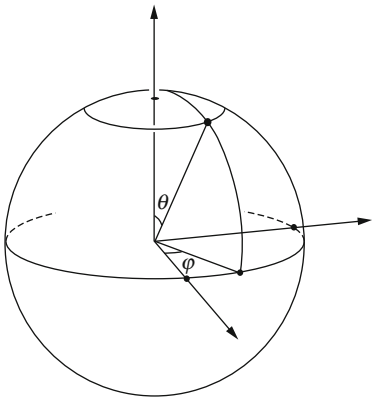

Figure 1.1

Example 1.3

Two particles connected by a rigid rod have five degrees of freedom: if the position of one particle is given (three coordinates), then the other can be anywhere on a sphere with its centre at the position of the first particle (two further coordinates).

Example 1.4

A rigid body has six degrees of freedom: three for the position of the centre of mass; two for the direction of some axis fixed in the body; and one for rotations about this axis.

The second stage, the dynamical part of the problem, is to use Newton's second law to determine the actual motion: to find out how the coordinates evolve as functions of time when the system is subjected to given forces. In the next few chapters we look at a number of techniques for finding and solving dynamical equations in general coordinate systems. These make it possible to simplify the second stage of a variety of mechanical problems, particularly problems involving constraints, by choosing well adapted coordinates in the first stage.

First, however, let us get our bearings by considering a very simple problem, the motion of a single particle moving in space under the influence of a given force. The kinematic analysis is easy. We can describe the motion of the particle by introducing a frame of reference R, which defines a standard of rest, but we need to think about what freedom is available in the choice of frame, what other coordinate systems could be used, and about how they are related.

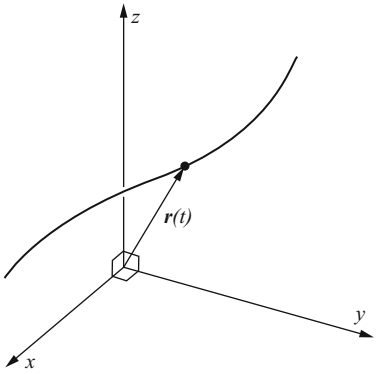

Figure 1.2

Definition 1.5

A *frame of reference* is an origin together with a set of right-handed Cartesian coordinate axes.

Let $\boldsymbol{r}$ denote the particle's position vector from the origin of R. Then the components of $\boldsymbol{r}$ along the axes, x, y, and z, are its Cartesian coordinates. A motion of the particle is represented by a curve $\boldsymbol{r} = \boldsymbol{r}(t)$ in space, along which x, y, and z are functions of time (Figure 1.2).

Definition 1.6

The *velocity* and *acceleration* of the particle relative to R are the vectors $\boldsymbol{v}$ and $\boldsymbol{a}$ with respective components $(\dot{x}, \dot{y}, \dot{z})$ and $(\ddot{x}, \ddot{y}, \ddot{z})$, where the dot denotes the derivative with respect to time.

It is important to remember that 'velocity' and 'acceleration' do not make sense unless one adds, either explicitly or by implication, 'with respect to such-and-such a frame of reference'.

Turning to the dynamics, it is an axiom of Newtonian mechanics that there exist special frames of reference, called *inertial frames*, in which Newton's second law holds. If R is such a frame, then

$$m\boldsymbol{a} = \boldsymbol{F} \tag{1.1}$$

where m is the mass of the particle and $\boldsymbol{F} = \boldsymbol{F}(\boldsymbol{r}, \boldsymbol{v}, t)$ is the force acting on it, which we allow to depend on the position and velocity of the particle, and on

the time t. When written out in components, (1.1) becomes a system of three simultaneous second-order differential equations,

$$\begin{aligned} \ddot{x} &= F_1(x, y, z, \dot{x}, \dot{y}, \dot{z}, t) \\ \ddot{y} &= F_2(x, y, z, \dot{x}, \dot{y}, \dot{z}, t) \\ \ddot{z} &= F_3(x, y, z, \dot{x}, \dot{y}, \dot{z}, t) \end{aligned}$$

which determine the three functions $x(t)$, $y(t)$, $z(t)$ in terms of the initial position and velocity.

There is considerable freedom in the choice of the inertial frame. We could pick a different origin or we could choose new directions for the axes. We could also replace R by a second frame $\tilde{R}$ moving relative to R without rotation and at constant velocity. Then the particle would have the same acceleration relative to $\tilde{R}$ and (1.1) would still hold.

If we ignore the effects of the Earth's rotation and acceleration, then a set of axes fixed on the Earth's surface is an inertial frame. But, again ignoring the effects of rotation and acceleration, Newton's laws are equally valid on the Moon, which is moving relative to the Earth at about one kilometre per second, or on the sun (about 30 km s^{-1}), or in the Andromeda galaxy (about 270 km s^{-1}). Only an extreme geochauvinist would insist on giving special status to terrestrial frames. As far as mechanical problems are concerned, all non-accelerating, non-rotating frames must be treated equally. To develop this idea in detail, we need to understand the relationship between coordinates and vector components measured in different frames of reference.

Exercise 1.1

Count the number of degrees of freedom in each of the following systems.

(a) A small bead sliding on a wire.

(b) A lamina moving in its own plane.

(c) A double pendulum confined to a vertical plane. This consists of a point mass A suspended from a fixed point by a thin rod; and a second point mass B suspended from A by a second thin rod. The rods are hinged at A.

(d) A double pendulum which is not confined to a vertical plane.

1.2 Frames of Reference

A general frame of reference consists of an origin and a set of right-handed Cartesian coordinate axes. These should not be regarded as fixed: the origin can be a moving point and the axes can be rotating.

Such a frame is represented by a pair $R = (O, \mathcal{T})$, where O is the origin and $\mathcal{T} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ is the triad of unit vectors along the coordinate axes. The $\boldsymbol{e}_i$s satisfy three conditions. First,

$$\boldsymbol{e}_1 \cdot \boldsymbol{e}_1 = \boldsymbol{e}_2 \cdot \boldsymbol{e}_2 = \boldsymbol{e}_3 \cdot \boldsymbol{e}_3 = 1$$

at all times because they are unit vectors. Second,

$$\boldsymbol{e}_1 \cdot \boldsymbol{e}_2 = \boldsymbol{e}_2 \cdot \boldsymbol{e}_3 = \boldsymbol{e}_3 \cdot \boldsymbol{e}_1 = 0$$

because the Cartesian axes are orthogonal, and, third,

$$\boldsymbol{e}_1 \cdot (\boldsymbol{e}_2 \wedge \boldsymbol{e}_3) = 1 \tag{1.2}$$

because the axes are right-handed. The first two conditions can be combined in the more compact form

$$\boldsymbol{e}_i \cdot \boldsymbol{e}_j = \delta_{ij}, \tag{1.3}$$

for $i, j = 1, 2, 3$. Here δ_{ij} is the *Kronecker delta symbol*, defined by

$$\begin{aligned} &\delta_{11} = \delta_{22} = \delta_{33} = 1 \\ &\delta_{12} = \delta_{21} = \delta_{23} = \delta_{32} = \delta_{31} = \delta_{13} = 0. \end{aligned}$$

Three vectors satisfying (1.2) and (1.3) at all times are said to make up a right-handed *orthonormal triad*, which we shorten to 'orthonormal triad', taking 'right-handed' as understood.

Any vector $\boldsymbol{x}$ can be expressed as a linear combination of the triad vectors in the form

$$\boldsymbol{x} = x_1 \boldsymbol{e}_i + x_2 \boldsymbol{e}_2 + x_3 \boldsymbol{e}_3, \tag{1.4}$$

where the x_is are functions of time, called the $\mathcal{T}$*-components* of $\boldsymbol{x}$. By taking the dot product with each triad vector in turn,

$$x_1 = \boldsymbol{x} \cdot \boldsymbol{e}_1, \quad x_2 = \boldsymbol{x} \cdot \boldsymbol{e}_2, \quad x_3 = \boldsymbol{x} \cdot \boldsymbol{e}_3.$$

A point P can be specified by its O*-position vector*, which is the vector from O to P. The components of this vector are the Cartesian coordinates of P in the frame.

In thinking about the relationship between vector components and coordinates in different frames of reference, one should keep in mind that a vector is not localized at a point. It is simply a quantity with magnitude and direction. If the distance from A to B is the same as the distance from $\tilde{A}$ to $\tilde{B}$, and if AB is parallel to $\tilde{A}\tilde{B}$, then the vector from A to B is the same as the vector from $\tilde{A}$ to $\tilde{B}$.

1.3 Transition Matrices

Suppose that $\mathcal{T} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ and $\tilde{\mathcal{T}} = (\tilde{\boldsymbol{e}}_1, \tilde{\boldsymbol{e}}_2, \tilde{\boldsymbol{e}}_3)$ are two orthonormal triads, which may be rotating relative to each other. Then, for $i, j = 1, 2, 3$,

$$\boldsymbol{e}_i \cdot \boldsymbol{e}_j = \delta_{ij} = \tilde{\boldsymbol{e}}_i \cdot \tilde{\boldsymbol{e}}_j$$

at all times. Put $H_{ij} = \boldsymbol{e}_i \cdot \tilde{\boldsymbol{e}}_j$ and put

$$H = \begin{pmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{pmatrix}. \tag{1.5}$$

Definition 1.7

The matrix H is the *transition matrix* from $\tilde{\mathcal{T}}$ to $\mathcal{T}$.

Note the transition matrix from $\mathcal{T}$ to $\tilde{\mathcal{T}}$ is the transposed matrix H^{t}. The H_{ij}s are nine functions of time, labelled by i and j. They determine the relative orientation of the two triads. With i fixed, H_{i1}, H_{i2}, H_{i3} are the $\tilde{\mathcal{T}}$-components of $\boldsymbol{e}_i$. With j fixed, H_{1j}, H_{2j}, H_{3j} are the $\mathcal{T}$-components of $\tilde{\boldsymbol{e}}_j$. Thus

$$\begin{aligned} \boldsymbol{e}_i &= H_{i1}\tilde{\boldsymbol{e}}_1 + H_{i2}\tilde{\boldsymbol{e}}_2 + H_{i3}\tilde{\boldsymbol{e}}_3 \\ \tilde{\boldsymbol{e}}_i &= H_{1i}\boldsymbol{e}_1 + H_{2i}\boldsymbol{e}_2 + H_{3i}\boldsymbol{e}_3, \end{aligned} \tag{1.6}$$

for $i = 1, 2, 3$.

Example 1.8

Suppose that

$$\begin{aligned} \tilde{\boldsymbol{e}}_1 &= \cos\theta\, \boldsymbol{e}_1 - \sin\theta\, \boldsymbol{e}_2 \\ \tilde{\boldsymbol{e}}_2 &= \sin\theta\, \boldsymbol{e}_1 + \cos\theta\, \boldsymbol{e}_2 \\ \tilde{\boldsymbol{e}}_3 &= \boldsymbol{e}_3 \end{aligned}$$

Then

$$H = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and

$$\begin{aligned} \boldsymbol{e}_1 &= \cos\theta\, \tilde{\boldsymbol{e}}_1 + \sin\theta\, \tilde{\boldsymbol{e}}_2 \\ \boldsymbol{e}_2 &= -\sin\theta\, \tilde{\boldsymbol{e}}_1 + \cos\theta\, \tilde{\boldsymbol{e}}_2 \\ \boldsymbol{e}_3 &= \tilde{\boldsymbol{e}}_3. \end{aligned}$$

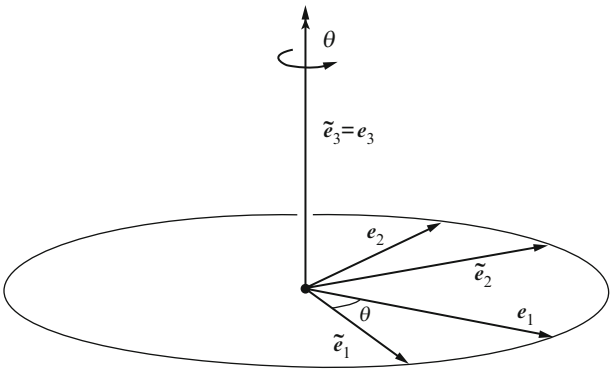

Figure 1.3

The triad $\mathcal{T}$ is obtained from $\tilde{\mathcal{T}}$ by rotation through θ about an axis parallel to $\tilde{\boldsymbol{e}}_3$ (Figure 1.3). In the same way, rotations through θ about the axes parallel to $\tilde{\boldsymbol{e}}_1$ and $\tilde{\boldsymbol{e}}_2$ are given respectively by the transition matrices

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}, \qquad \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix}. \tag{1.7}$$

Now let $\boldsymbol{x}$ be a general vector and put $x_i = \boldsymbol{x} \cdot \boldsymbol{e}_i$, and $\tilde{x}_i = \boldsymbol{x} \cdot \tilde{\boldsymbol{e}}_i$. The x_is are the $\mathcal{T}$-components of $\boldsymbol{x}$ and the $\tilde{x}_i$s are the $\tilde{\mathcal{T}}$-components of $\boldsymbol{x}$. By substituting from (1.6), we obtain

$$x_i = \boldsymbol{x} \cdot \boldsymbol{e}_i = \sum_{j=1}^{3} H_{ij} \boldsymbol{x} \cdot \tilde{\boldsymbol{e}}_j = \sum_{j=1}^{3} H_{ij} \tilde{x}_j$$

and

$$\tilde{x}_i = \boldsymbol{x} \cdot \tilde{\boldsymbol{e}}_i = \sum_{j=1}^{3} H_{ji} \boldsymbol{x} \cdot \boldsymbol{e}_j = \sum_{j=1}^{3} H_{ji} x_j$$

or, in matrix notation,

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = H \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \tilde{x}_3 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \tilde{x}_3 \end{pmatrix} = H^{\mathrm{t}} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \tag{1.8}$$

where the 't' denotes transposition. Because this holds for any $\boldsymbol{x}$, H^{t} must also be the inverse of H. Hence

$$H^{\mathrm{t}} H = I = H H^{\mathrm{t}},$$

where I is the 3×3 identity matrix. In other words, H is an orthogonal matrix. Exercise (1.7) contains an outline of a demonstration that the right-handedness of the two triads implies that H is also *proper* in the sense that $\det(H) = 1$.

EXERCISES

1.2. Suppose that the matrix H in (1.5) is given by

$$H = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{2} & \sqrt{2} & \sqrt{2} \\ -2 & 1 & 1 \\ 0 & -\sqrt{3} & \sqrt{3} \end{pmatrix}$$

Check that $H^{\mathrm{t}}H = HH^{\mathrm{t}} = I$. Write down the components of $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, and $\boldsymbol{e}_3$ in $\tilde{\mathcal{T}}$, and the components of $\tilde{\boldsymbol{e}}_1$, $\tilde{\boldsymbol{e}}_2$, and $\tilde{\boldsymbol{e}}_3$ in $\mathcal{T}$.

1.3. What are the transition matrices for rotations through $\pm 2\pi/3$ about an axis aligned with the vector with $\mathcal{T}$-components $(1, 1, 1)$?

1.4. Show that if H is a proper orthogonal matrix such that $H_{33} = 1$, then there is a unique angle $\alpha \in [0, 2\pi)$ such that

$$H = \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Show that if $H_{33} = -1$, then there is a unique angle $\alpha \in [0, 2\pi)$ such that

$$H = \begin{pmatrix} -\cos\alpha & -\sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Sketch a diagram of two orthonormal triads with this transition matrix, showing the angle α.

1.5. Show that if H is an orthogonal matrix, then $H^{\mathrm{t}}(H - I) = (I - H)^{\mathrm{t}}$. Deduce that if H is also proper, then $\det(I - H) = 0$. Hence show that if $\mathcal{T}$ and $\tilde{\mathcal{T}}$ are two (right-handed) orthonormal triads, then there exists a nonzero vector that has the same components in both triads.

1.6. Show that if $\tilde{\boldsymbol{e}}_i = \sum_j H_{ji}\boldsymbol{e}_j$ where $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ is an orthonormal triad, then

$$\tilde{\boldsymbol{e}}_1 \cdot (\tilde{\boldsymbol{e}}_2 \wedge \tilde{\boldsymbol{e}}_3) = \det(H)\, \boldsymbol{e}_1 \cdot (\boldsymbol{e}_2 \wedge \boldsymbol{e}_3).$$

Deduce that if $\tilde{\mathcal{T}} = (\tilde{\boldsymbol{e}}_1, \tilde{\boldsymbol{e}}_2, \tilde{\boldsymbol{e}}_3)$ and $\mathcal{T} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ are right-handed orthonormal triads, then the transition matrix from $\tilde{\mathcal{T}}$ to $\mathcal{T}$ is a *proper* orthogonal matrix.

1.4 Euler Angles

Let $\mathcal{T}$ a right-handed orthonormal triad and let H denote the transition matrix from a second right-handed orthonormal triad $\tilde{\mathcal{T}}$ to $\mathcal{T}$. The nine entries in H determine the relative orientation of the two triads, but we cannot specify them independently of each other because H must satisfy the orthogonality condition

$$HH^{\mathrm{t}} = I.$$

There are six independent equations here, for example, three for the diagonal entries and three for the entries above the diagonal, so we should be able to express H in terms of three independent parameters. The following proposition shows that this is indeed the case.

Proposition 1.9

Let H be a 3×3 proper orthogonal matrix. Then there exist angles $\theta \in [0, \pi]$, $\varphi, \psi \in [0, 2\pi)$ such that

$$H = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Moreover ψ, θ, and φ are uniquely determined by H provided that $|H_{33}| \neq 1$.

Proof

First we deal with uniqueness. If ψ, θ, and φ exist, then

$$\begin{aligned} H_{33} &= \cos\theta, \\ H_{31} &= \sin\theta\cos\varphi, \qquad & H_{32} &= \sin\theta\sin\varphi, \\ H_{13} &= -\sin\theta\cos\psi, \qquad & H_{23} &= \sin\theta\sin\psi. \end{aligned} \tag{1.9}$$

The first equation fixes the value of θ uniquely in the interval $[0, \pi]$. The second pair then determine φ uniquely in $[0, 2\pi)$, provided that $\sin\theta \neq O$, that is, provided that $|H_{33}| \neq 1$. Finally, under the same condition, the last pair determine ψ uniquely in $[0, 2\pi)$.

To establish existence, we consider first the case $|H_{33}| \neq 1$. Choose φ and ψ so that

$$H_{31}\sin\varphi - H_{32}\cos\varphi = 0, \qquad H_{13}\sin\psi + H_{23}\cos\psi = 0.$$

These determine φ and ψ up to the addition of integral multiples of π. Consider the matrix K defined by

$$K = \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} H \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.10}$$

Note that $K_{33} = H_{33}$, and that

$$K_{32} = -\sin\varphi\, H_{31} + \cos\varphi\, H_{32} = 0, \quad K_{23} = \sin\psi\, H_{13} + \cos\psi\, H_{23} = 0.$$

Because K is an orthogonal matrix, its second and third columns are orthogonal, so we also have

$$K_{12}K_{13} = 0.$$

Similarly, its second and third rows are orthogonal, so

$$K_{21}K_{31} = 0.$$

Also its third row and its third column are unit vectors, so

$$K_{31}^2 + K_{33}^2 = K_{13}^2 + K_{33}^2 = 1.$$

But $K_{33}^2 = H_{33}^2 \neq 1$, so K_{13} and K_{31} are both nonzero. It follows that $K_{21} = K_{12} = 0$. Also the middle column and the middle row of K are unit vectors. So finally $K_{22} = \pm 1$.

If we add an odd multiple of π to ψ, then these conclusions still hold, but the entries in the first two rows of K are multiplied by -1. Similarly if we add an odd multiple of π to φ, then the first two columns of K are multiplied by -1. By using this freedom, we can fix φ and ψ uniquely in the interval $[0, 2\pi)$ by imposing the requirements that $K_{22} = 1$ and $K_{31} > 0$. We then have that K is of the form

$$K = \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix}$$

where $\theta \in (0, \pi)$. The displayed equation in the proposition now follows from (1.10).

When $H_{33} = \pm 1$, we get the displayed equation directly by taking $\psi = 0$ and either $\theta = 0$ or $\theta = \pi$. □

Definition 1.10

Let $\mathcal{T}$ and $\tilde{\mathcal{T}}$ be two right-handed orthonormal triads and let H be the transition matrix from $\tilde{\mathcal{T}}$ to $\mathcal{T}$. The angles θ, φ, and ψ are called the *Euler angles* of $\mathcal{T}$ relative to $\tilde{\mathcal{T}}$.

Note from (1.9) that $\boldsymbol{e}_3$ has $\tilde{\mathcal{T}}$-components

$$(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta) \tag{1.11}$$

so that θ and φ are the spherical polar angles of $\boldsymbol{e}_3$ in $\tilde{\mathcal{T}}$; and that $\tilde{\boldsymbol{e}}_3$, has $\mathcal{T}$-components

$$(-\sin\theta\cos\psi, \sin\theta\sin\psi, \cos\theta) \tag{1.12}$$

so that θ and $\pi - \psi$ are the spherical polar angles of $\tilde{\boldsymbol{e}}_3$ in $\mathcal{T}$.

A key to understanding the geometry of the Euler angles is to introduce the unit vector

$$\boldsymbol{j} = \sin\psi\, \boldsymbol{e}_1 + \cos\psi\, \boldsymbol{e}_2. \tag{1.13}$$

This is orthogonal to $\boldsymbol{e}_3$ and also, from (1.12), to $\tilde{\boldsymbol{e}}_3$. It is the unit vector in the direction of $\tilde{\boldsymbol{e}}_3 \wedge \boldsymbol{e}_3$ and its direction determines the so-called *line of nodes*. From (1.11), it can also be written in the form

$$\boldsymbol{j} = -\sin\varphi\, \tilde{\boldsymbol{e}}_1 + \cos\varphi\, \tilde{\boldsymbol{e}}_2. \tag{1.14}$$

The proposition splits the transformation from $\tilde{\mathcal{T}}$ to $\mathcal{T}$ into three steps: a rotation about $\tilde{\boldsymbol{e}}_3$ through φ, which brings $\tilde{\boldsymbol{e}}_2$ into coincidence with $\boldsymbol{j}$; a rotation about $\boldsymbol{j}$ through θ, which brings $\tilde{\boldsymbol{e}}_3$ into coincidence with $\boldsymbol{e}_3$; and, finally, a rotation about $\boldsymbol{e}_3$ through ψ, which brings $\boldsymbol{j}$ into coincidence with $\boldsymbol{e}_2$ (Figure 1.4).

EXERCISES

1.7. Find the Euler angles of

$$H = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

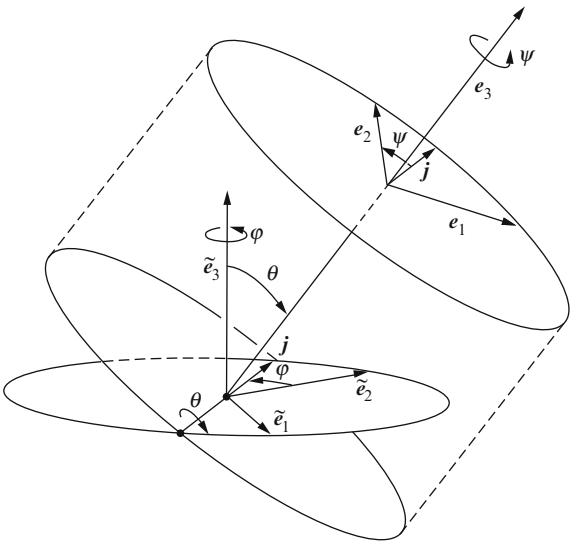

Figure 1.4

1.8. Show that the trace of any 3×3 proper orthogonal matrix H is given in terms of Euler angles by

$$\mathrm{tr}H = \big(\cos\theta + 1\big)\big(\cos(\theta+\psi) + 1\big) - 1.$$

1.5 Angular Velocity

We now consider how to describe the *rate* of rotation of $\mathcal{T}$ relative to a second orthonormal triad $\tilde{\mathcal{T}}$. By differentiating the orthogonality condition $HH^{\mathrm{t}} = I$ with respect to time, we obtain

$$\dot{H}H^{\mathrm{t}} + H\dot{H}^{\mathrm{t}} = 0,$$

where $\dot{H}$ denotes the matrix whose entries are the time derivatives of the entries in H. Because $H\dot{H}^{\mathrm{t}}$ is the transpose of $\dot{H}H^{\mathrm{t}}$, it follows that the matrix

$$\Omega = H\dot{H}^{\mathrm{t}} = -\dot{H}H^{\mathrm{t}}$$

is skew-symmetric. That is,

$$\Omega = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix} \tag{1.15}$$

for some functions of time ω_1, ω_2, and ω_3.

Definition 1.11

The *angular velocity* of $\mathcal{T}$ relative to $\tilde{\mathcal{T}}$ is the vector $\boldsymbol{\omega} = \sum_{i=1}^{3} \omega_i \boldsymbol{e}_i$.

We also use this to define the angular velocity of one frame $R = (O, \mathcal{T})$ relative to another frame $\tilde{R} = (\tilde{O}, \tilde{\mathcal{T}})$ as the angular velocity of $\mathcal{T}$ relative to $\tilde{\mathcal{T}}$.

Angular velocity is not localized at a particular point or on a particular axis. It is the 'angular velocity of $\mathcal{T}$ relative to $\tilde{\mathcal{T}}$', not 'the angular velocity of $\mathcal{T}$ relative to $\tilde{\mathcal{T}}$ about such-and-such an axis'. It is also related to Ω through the following lemma.

Lemma 1.12

Suppose that $\boldsymbol{x}$ has $\mathcal{T}$-components x_1, x_2, x_3. Then

$$\Omega \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \omega_2 x_3 - \omega_3 x_1 \\ \omega_3 x_1 - \omega_1 x_3 \\ \omega_1 x_2 - \omega_2 x_1 \end{pmatrix}.$$

The entries in the column vector on the right are the components of $\boldsymbol{\omega} \wedge \boldsymbol{x}$.

Proof

The lemma follows by evaluating the matrix product on the left. □

Example 1.13

Suppose that H is the matrix in Example (1.8), where θ is now a function of time. Then

$$\dot{H} = \dot{\theta} \begin{pmatrix} -\sin\theta & \cos\theta & 0 \\ -\cos\theta & -\sin\theta & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{1.16}$$

and so

$$\Omega = -\dot{H}H^{\mathrm{t}} = \begin{pmatrix} 0 & -\dot{\theta} & 0 \\ \dot{\theta} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{1.17}$$

Therefore the angular velocity of $\mathcal{T}$ relative to $\tilde{\mathcal{T}}$ is $\omega = \dot{\theta}\boldsymbol{e}_3 = \dot{\theta}\check{\boldsymbol{e}}_3$.

EXERCISES

1.9. Show that for the two rotations in (1.7), Ω is given by, respectively,

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\dot{\theta} \\ 0 & \dot{\theta} & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 0 & \dot{\theta} \\ 0 & 0 & 0 \\ -\dot{\theta} & 0 & 0 \end{pmatrix}.$$

Show that the corresponding angular velocities are $\dot{\theta}\boldsymbol{e}_1$ and $\dot{\theta}\boldsymbol{e}_2$.

1.10. Show that if $\boldsymbol{\omega}$, and Ω are related by (1.15), then

$$\boldsymbol{\omega} \cdot \boldsymbol{\omega} = -\tfrac{1}{2}\mathrm{tr}(\Omega^2)$$

and also

$$\Omega \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} = 0.$$

Show that if there are constants x_i such

$$H \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0.$$

for all t, then $\boldsymbol{\omega}$ is parallel to $\boldsymbol{x} = \sum_i x_i \boldsymbol{e}_i$.

1.6 The Coriolis Theorem

In any frame of reference, one can define the time derivative of a vector by differentiating its components in the frame. This seems natural if one thinks of the frame as fixed. However if one applies this definition in two different frames which are rotating relative to each other, then the resulting vector depends on the choice of frame. In fact the angular velocity of one frame relative to the other has an alternative characterization in terms of the relationship between the corresponding two notions of 'vector time derivative'.

Definition 1.14

The *time derivative* of the time-dependent vector $\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + x_3\boldsymbol{e}_3$ relative to an orthonormal triad $\mathcal{T} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ is the vector

$$\mathrm{D}\boldsymbol{x} = \dot{x}_1\boldsymbol{e}_1 + \dot{x}_2\boldsymbol{e}_2 + \dot{x}_3\boldsymbol{e}_3.$$

That is, $\mathrm{D}\boldsymbol{x}$ is obtained from $\boldsymbol{x}$ by taking the components of $\boldsymbol{x}$ in $\mathcal{T}$ and by differentiating them with respect to time, or alternatively, by differentiating the right-hand side of (1.4), treating the $\boldsymbol{e}_i$s as constants. For example, if $\boldsymbol{r}$ is the position vector of a particle from the origin of a frame $R = (O, \mathcal{T})$, then $\mathrm{D}\boldsymbol{r}$ and $\mathrm{D}^2\boldsymbol{r}$ are the velocity and acceleration of the particle relative to R.

The derivative of a scalar quantity with respect to time is defined without reference to a particular choice of axes. Because one can compare the values of a scalar at two different times without ambiguity, it makes sense to talk about the 'rate of change of a scalar' without adding a qualification.

The derivative of a vector, on the other hand, is a more subtle concept. We know how to compare two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ at different points of space at a fixed time: they are the same if they have the same magnitude and the same direction. So there is no difficulty in differentiating vectors with respect to parameters such as arc length along a curve, with the time held fixed. But how do we compare $\boldsymbol{a}$ at time t_1 with $\boldsymbol{b}$ at time t_2? Comparing their lengths is easy, but to compare their directions, we must know how to interpret the statement '$\boldsymbol{a}(t_1)$ and $\boldsymbol{b}(t_2)$ point in the same direction', and that depends on who is doing the comparing. For example someone on Earth might say that $\boldsymbol{a}(t_1)$ and $\boldsymbol{b}(t_2)$ pointed in the same direction if they both happened to point vertically upwards. But then they would not appear to have the same directions to an astronaut standing on the Moon because the Earth and the Moon rotate relative to each other between t_1, and t_2.

Proposition 1.15 (The Coriolis Theorem)

The time derivatives $\mathrm{D}\boldsymbol{x}$ and $\tilde{\mathrm{D}}\boldsymbol{x}$ of $\boldsymbol{x}$ relative to $\mathcal{T}$ and $\tilde{\mathcal{T}}$ are related by

$$\tilde{\mathrm{D}}\boldsymbol{x} = \mathrm{D}\boldsymbol{x} + \boldsymbol{\omega} \wedge \boldsymbol{x} \tag{1.18}$$

where $\boldsymbol{\omega}$ is the angular velocity of $\mathcal{T}$ relative to $\tilde{\mathcal{T}}$.

Proof

The components of $\boldsymbol{x}$ in two triads are related by

$$\begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \tilde{x}_3 \end{pmatrix} = H^{\mathrm{t}} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

On differentiating the matrix entries on each side, we get

$$\begin{pmatrix} \dot{\tilde{x}}_1 \\ \dot{\tilde{x}}_2 \\ \dot{\tilde{x}}_3 \end{pmatrix} = H^{\mathrm{t}} \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} + \dot{H}^{\mathrm{t}} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

and therefore, by multiplying on the left by H,

$$H \begin{pmatrix} \dot{\tilde{x}}_1 \\ \dot{\tilde{x}}_2 \\ \dot{\tilde{x}}_3 \end{pmatrix} = \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} + \Omega \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

The $\tilde{\mathcal{T}}$-components of $\tilde{\mathrm{D}}\boldsymbol{x}$ are $\dot{\tilde{\boldsymbol{x}}}_1$, $\dot{\tilde{\boldsymbol{x}}}_2$, $\dot{\tilde{\boldsymbol{x}}}_3$, so the entries in the column vector on left-hand side are the $\mathcal{T}$-components of $\tilde{\mathrm{D}}\boldsymbol{x}$. By using Lemma 1.12, we conclude that

$$\tilde{\mathrm{D}}\boldsymbol{x} = \mathrm{D}\boldsymbol{x} + \boldsymbol{\omega} \wedge \boldsymbol{x}.$$

□

The Coriolis theorem is the key to the interpretation of angular velocity. If $\boldsymbol{x}$ is a vector with fixed components in a triad $\mathcal{T}$, so that $\mathrm{D}\boldsymbol{x} = 0$, then its time derivative with respect to a second triad $\tilde{\mathcal{T}}$ is

$$\tilde{\mathrm{D}}\boldsymbol{x} = \boldsymbol{\omega} \wedge \boldsymbol{x}, \tag{1.19}$$

where $\boldsymbol{\omega}$ is the angular velocity of $\mathcal{T}$ relative $\tilde{\mathcal{T}}$. Equation (1.19) determines the way in which $\boldsymbol{x}$ changes relative to frame of reference $(\tilde{O}, \tilde{\mathcal{T}})$. If we take $\tilde{R}$ as the standard of rest and ignore terms of order δt^2, then $\boldsymbol{x}$ appears to change between times t and $t + \delta t$ by the addition of $\delta\boldsymbol{x} = \boldsymbol{\omega} \wedge \boldsymbol{x}\, \delta t$. This is a vector orthogonal to $\boldsymbol{\omega}$ and to $\boldsymbol{x}$ and of modulus $|\boldsymbol{\omega}|\, |\boldsymbol{x}| \sin\theta\, \delta t$, where θ is the angle between $\boldsymbol{\omega}$ and $\boldsymbol{x}$.

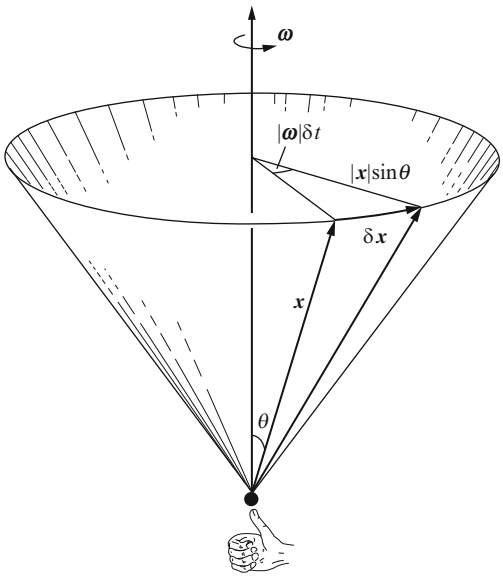

Figure 1.5

Measured in $\tilde{R}$, therefore, the change in $\boldsymbol{x}$ between t and $t+\delta t$ is produced by rotating $\boldsymbol{x}$ through an angle $|\boldsymbol{\omega}|\delta t$ about an axis parallel to $\boldsymbol{\omega}$, in the right-handed sense. That is, when the thumb of the right hand points along $\boldsymbol{\omega}$, the fingers curl in the direction of the rotation; see Figure 1.5.

Equation 1.18 can be taken as the definition of $\boldsymbol{\omega}$. It determines $\boldsymbol{\omega}$ uniquely because if $\boldsymbol{\omega}'$ is another vector with the property

$$\tilde{\mathrm{D}}\boldsymbol{x} = \mathrm{D}\boldsymbol{x} + \boldsymbol{\omega}' \wedge \boldsymbol{x} \tag{1.20}$$

for all $\boldsymbol{x}$, then $(\boldsymbol{\omega}-\boldsymbol{\omega}')\wedge\boldsymbol{x}=0$ for all $\boldsymbol{x}$, and so $\boldsymbol{\omega}=\boldsymbol{\omega}'$. However, some care is needed. Without going into the construction of the matrix Ω, it is not obvious that there exists a vector such that (1.18) holds for all $\boldsymbol{x}$.

Immediate corollaries of the Coriolis theorem are the following.

Proposition 1.16

If the angular velocity of $\mathcal{T}$ relative to $\tilde{\mathcal{T}}$ is $\boldsymbol{\omega}$, then the angular velocity of $\tilde{\mathcal{T}}$ relative to $\mathcal{T}$ is $-\boldsymbol{\omega}$.

Proposition 1.17

If $\mathcal{T}$ has angular velocity $\boldsymbol{\omega}$ relative to $\tilde{\mathcal{T}}$ and $\tilde{\mathcal{T}}$ has angular velocity $\tilde{\boldsymbol{\omega}}$ relative to $\mathcal{T}'$, then $\mathcal{T}$ has angular velocity $\boldsymbol{\omega}+\boldsymbol{\omega}'$ relative to $\mathcal{T}'$.

Proof

Because $\tilde{\mathrm{D}}\boldsymbol{x} = \mathrm{D}\boldsymbol{x} + \boldsymbol{\omega} \wedge \boldsymbol{x}$,

$$\mathrm{D}\boldsymbol{x} = \tilde{\mathrm{D}}\boldsymbol{x} - \boldsymbol{\omega} \wedge \boldsymbol{x} \tag{1.21}$$

for all $\boldsymbol{x}$. Hence the angular velocity of $\tilde{\mathcal{T}}$ relative to $\mathcal{T}$ is $-\boldsymbol{\omega}$. Similarly, if D' denotes the time derivative relative to $\mathcal{T}'$, then, for all $\boldsymbol{x}$,

$$\begin{aligned} \tilde{\mathrm{D}}\boldsymbol{x} &= \mathrm{D}\boldsymbol{x} + \boldsymbol{\omega} \wedge \boldsymbol{x} \\ \mathrm{D}'\boldsymbol{x} &= \tilde{\mathrm{D}}\boldsymbol{x} + \tilde{\boldsymbol{\omega}} \wedge \boldsymbol{x}. \end{aligned} \tag{1.22}$$

Therefore

$$\mathrm{D}'\boldsymbol{x} = \mathrm{D}\boldsymbol{x} + (\boldsymbol{\omega} + \tilde{\boldsymbol{\omega}}) \wedge \boldsymbol{x} \tag{1.23}$$

for all $\boldsymbol{x}$ and hence the angular velocity of $\mathcal{T}$ relative to $\mathcal{T}'$ is $\boldsymbol{\omega} + \tilde{\boldsymbol{\omega}}$. □

We can specify the changing orientation of an orthonormal triad $\mathcal{T}$ relative to a second orthonormal $\tilde{\mathcal{T}}$ by giving the Euler angles θ, φ, and ψ as functions of time. So we can express the angular velocity of $\mathcal{T}$ relative to $\tilde{\mathcal{T}}$ in terms of the derivatives of the Euler angles, by making use of Proposition 1.9. The calculation is greatly simplified by appealing to Proposition 1.17.

In the notation of Section 1.4, we introduce two intermediate triads

$$\mathcal{T}' = (\boldsymbol{j} \wedge \boldsymbol{e}_3, \boldsymbol{j}, \boldsymbol{e}_3) = (\cos\psi\, \boldsymbol{e}_1 - \sin\psi\, \boldsymbol{e}_2, \sin\psi\, \boldsymbol{e}_1 + \cos\psi\, \boldsymbol{e}_2, \boldsymbol{e}_3)$$

and

$$\mathcal{T}'' = (\boldsymbol{j} \wedge \tilde{\boldsymbol{e}}_3, \boldsymbol{j}, \tilde{\boldsymbol{e}}_3) = (\cos\varphi\, \tilde{\boldsymbol{e}}_1 + \sin\varphi\, \tilde{\boldsymbol{e}}_2, -\sin\varphi\, \tilde{\boldsymbol{e}}_1 + \cos\varphi\, \tilde{\boldsymbol{e}}_2, \tilde{\boldsymbol{e}}_3)$$

The respective transition matrices from $\tilde{\mathcal{T}}$ to $\mathcal{T}''$, from $\mathcal{T}''$ to $\mathcal{T}'$, and from $\mathcal{T}'$ to $\mathcal{T}$ are

$$\begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix}, \quad \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

So the angular velocity of $\mathcal{T}$ relative to $\mathcal{T}'$ is $\dot{\psi}\boldsymbol{e}_3$, the angular velocity of $\mathcal{T}'$ relative to $\mathcal{T}''$ is $\dot{\theta}\boldsymbol{j}$, and the angular velocity of $\mathcal{T}''$ relative to $\tilde{\mathcal{T}}$ is $\dot{\varphi}\tilde{\boldsymbol{e}}_3$ (see Example 1.13 and Exercise 1.9). Therefore the angular velocity of $\mathcal{T}$ relative to $\tilde{\mathcal{T}}$ is

$$\begin{aligned} \boldsymbol{\omega} &= \dot{\psi}\, \boldsymbol{e}_3 + \dot{\theta}\, \boldsymbol{j} + \dot{\varphi}\, \tilde{\boldsymbol{e}}_3 \\ &= \dot{\psi}\, \boldsymbol{e}_3 + \dot{\theta}(\sin\psi\, \boldsymbol{e}_1 + \cos\psi\, \boldsymbol{e}_2) \\ &\quad + \dot{\varphi}(-\sin\theta\cos\psi\, \boldsymbol{e}_1 + \sin\theta\sin\psi\, \boldsymbol{e}_2 + \cos\theta\, \boldsymbol{e}_3), \end{aligned}$$

where we have substituted for $\boldsymbol{j}$ and $\tilde{\boldsymbol{e}}_3$ from (1.13) and (1.12). Hence the $\mathcal{T}$-components of $\boldsymbol{\omega}$ are

$$\begin{aligned} \omega_1 &= \dot\theta \sin\psi - \dot\varphi \sin\theta \cos\psi \\ \omega_2 &= \dot\theta \cos\psi + \dot\varphi \sin\theta \sin\psi \\ \omega_3 &= \dot\psi + \dot\varphi \cos\theta. \end{aligned} \tag{1.24}$$

Similarly, the $\tilde{\mathcal{T}}$-components are

$$\begin{aligned} \tilde\omega_1 &= -\dot\theta \sin\varphi + \dot\psi \sin\theta \cos\varphi \\ \tilde\omega_2 &= \dot\theta \cos\varphi + \dot\psi \sin\theta \sin\varphi \\ \tilde\omega_3 &= \dot\varphi + \dot\psi \cos\theta. \end{aligned} \tag{1.25}$$

Put in less formal terms, $\boldsymbol{\omega}$ is the sum of three vectors. The first, $\dot\psi \boldsymbol{e}_3$ is the angular velocity of $\mathcal{T}$ relative to $\tilde{\mathcal{T}}$ when θ and φ are held constant. The second, $\dot\theta \boldsymbol{j}$ is the angular velocity when ψ and φ are held constant, and the third, $\dot\varphi \boldsymbol{e}_3$, is the angular velocity when θ and ψ are held constant. This is the simplest way to obtain the angular velocity in practice.

EXERCISES

1.11. Show that

(i) $\mathrm{D}(\boldsymbol{a} \wedge \boldsymbol{b}) = (\mathrm{D}\boldsymbol{a}) \wedge \boldsymbol{b} + \boldsymbol{a} \wedge \mathrm{D}\boldsymbol{b}$,

(ii) $\dfrac{\mathrm{d}}{\mathrm{d}t}\big(\boldsymbol{a} \cdot \boldsymbol{b}\big) = (\mathrm{D}\boldsymbol{a}) \cdot \boldsymbol{b} + \boldsymbol{a} \cdot (\mathrm{D}\boldsymbol{b})$.

1.12. Derive the addition law for angular velocities from the composition rule $H = H'H''$ for transition matrices.

1.7 Rotating, Accelerating, and Inertial Frames

Let us compare the motions of a particle relative to the two frames of reference, $R = (O, \mathcal{T})$ and $\tilde R = (\tilde O, \tilde{\mathcal{T}})$. Let $\boldsymbol{r}$ denote the O-position vector of the particle, let $\tilde{\boldsymbol{r}}$ denote its $\tilde O$-position vector, and let $\boldsymbol{x}$ denote the $\tilde O$-position vector of O, that is the vector from $\tilde O$ to O. Keep in mind that $\boldsymbol{r}$, $\tilde{\boldsymbol{r}}$, and $\boldsymbol{x}$ all depend on time.

The position vectors of the particle in the two frames are related by

$$\tilde{\boldsymbol{r}} = \boldsymbol{r} + \boldsymbol{x}. \tag{1.26}$$

From this, we can determine the relationship between $\boldsymbol{a} = \mathrm{D}^2\boldsymbol{r}$, which is the acceleration relative to R, and $\tilde{\boldsymbol{a}} = \tilde{\mathrm{D}}^2\tilde{\boldsymbol{r}}$, which is the acceleration relative to $\tilde{R}$. By taking the second time derivative of (1.26) relative to $\tilde{\mathcal{T}}$ we get

$$\begin{aligned} \tilde{\boldsymbol{a}} &= \tilde{\mathrm{D}}^2(\boldsymbol{r} + \boldsymbol{x}) \\ &= \tilde{\mathrm{D}}(\mathrm{D}\boldsymbol{r} + \boldsymbol{\omega} \wedge \boldsymbol{r}) + \boldsymbol{A} \\ &= \boldsymbol{a} + (\mathrm{D}\boldsymbol{\omega}) \wedge \boldsymbol{r} + 2\boldsymbol{\omega} \wedge \mathrm{D}\boldsymbol{r} + \boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge \boldsymbol{r}) + \boldsymbol{A} \end{aligned} \tag{1.27}$$

where $\boldsymbol{A} = \tilde{\mathrm{D}}^2\boldsymbol{x}$ is the acceleration of O relative to $\tilde{R}$ (see Exercise 1.11).

When $\boldsymbol{\omega} = 0$ and $\boldsymbol{A} = 0$, we have $\tilde{\boldsymbol{a}} = \boldsymbol{a}$. This is the case in which R is not rotating relative to $\tilde{R}$, so the transition matrix from $\tilde{\mathcal{T}}$ to $\mathcal{T}$ is constant, and the point O is moving relative to $\tilde{R}$ in a straight line with constant velocity. Under these conditions, the particle has the same acceleration relative to the two frames, whatever its motion.

Definition 1.18

The two frames R and $\tilde{R}$ are *equivalent* if $\boldsymbol{\omega} = 0$ and O is moving relative to $\tilde{R}$ in a straight line with constant velocity.

It is easy to see that equivalence of frames is symmetric and transitive. If R is equivalent to $\tilde{R}$, then $\tilde{R}$ is equivalent to R, and if R is equivalent to $\tilde{R}$ and $\tilde{R}$ is equivalent to R', then R is equivalent to R'.

The inertial frames are an equivalence class of frames. Newton's second law holds in every inertial frame, and every frame in which it holds for every choice of $\boldsymbol{F}$ in (1.1) is inertial. The axioms of classical mechanics assert the existence of inertial frames, but they do not single out a particular frame. In classical mechanics, there is no absolute standard of rest.

There is, however, an absolute standard of acceleration. We can say that a particle has acceleration $\boldsymbol{a}$ if it has acceleration a relative to some, and hence to every, inertial frame. Although it is not possible to test whether or not a frame is 'at rest', it is possible to contrive mechanical experiments that test whether or not a frame is rotating or accelerating. For example, Foucault's pendulum experiment (1851) demonstrated that axes fixed on the Earth's surface are not exactly inertial, although the delicacy of the experiment indicated that for most practical purposes they can be treated as such.

Because the axes in two equivalent frames of reference are related by a constant orthogonal matrix, the angular velocity of a frame relative to an equivalent frame is zero. The following is therefore an immediate consequence of Proposition 1.17.

Proposition 1.19

If R has angular velocity $\boldsymbol{\omega}$ relative to $\tilde{R}$, then every frame equivalent to R has angular velocity $\boldsymbol{\omega}$ relative to every frame equivalent to $\tilde{R}$.

It follows that the angular velocity of a frame relative to an inertial frame is the same, whichever inertial frame is chosen.

Definition 1.20

The *angular velocity* of a frame is its angular velocity relative to an inertial frame.

The idea that frames in uniform relative motion are equivalent for describing mechanical phenomena predates Newton's formulation of the laws of classical mechanics. In Galileo's *Dialogue concerning the two chief world systems* [5], Salviati (who speaks for Galileo) argues, by considering the behaviour of butterflies and fish, that 'in the main cabin below the decks on some large ship' it is impossible to detect the motion of the ship in its effect on physical phenomena so long as 'the motion be uniform and not fluctuating this way and that'. Newton himself stressed the idea. Corollary V to his laws of motion states: 'The motions of bodies included in a given space are the same among themselves, whether that space is at rest or moves uniformly forward in a right line without circular motion'. However, he also asserted the existence of 'absolute space', which remains 'always similar and immovable', and distinguished between absolute motion, that is, motion relative to absolute space, and relative motion [9].

In the late nineteenth century, the electromagnetic ether was thought to provide an absolute standard of rest, although attempts to detect the effects of motion relative to the ether on electromagnetic processes, such as the experiment of Michelson and Morley (1887), ended in failure. Einstein finally abolished the ether and 'absolute space' in 1905. He asserted the equivalence of inertial frames for describing all physical phenomena as a fundamental principle, the *principle of relativity*.

In an inertial frame, the dynamical equations governing the motion of a particle are particularly straightforward. Nevertheless sometimes it is simpler to refer the motion to a non-inertial frame, even though it is then necessary to introduce correction terms to compensate for the frame's acceleration and rotation.

Suppose that $\tilde{R} = (\tilde{O}, \tilde{\mathcal{T}})$ is the inertial frame and that $R = (O, \mathcal{T})$ is some other frame. Then,

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

where $\boldsymbol{F}$ is the force acting on the particle, and so

$$m\mathrm{D}^2\boldsymbol{r} + m(D\boldsymbol{\omega}) \wedge \boldsymbol{r} + 2m\boldsymbol{\omega} \wedge \mathrm{D}\boldsymbol{r} + m\boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge \boldsymbol{r}) + m\boldsymbol{A} = \boldsymbol{F}$$

by substituting for $\tilde{\boldsymbol{a}}$ from (1.27). Here $\boldsymbol{\omega}$ is the angular velocity of R and $\boldsymbol{A}$ is the acceleration of O relative to $\tilde{R}$. On transferring terms from one side to the other, this becomes

$$m\boldsymbol{a} = \boldsymbol{F} - m(D\boldsymbol{\omega}) \wedge \boldsymbol{r} - 2m\boldsymbol{\omega} \wedge \mathrm{D}\boldsymbol{r} - m\boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge \boldsymbol{r}) - m\boldsymbol{A}, \tag{1.28}$$

which is the same as the equation of motion that we would write down if R were an inertial frame and there were additional forces

$$\begin{aligned} \boldsymbol{F}_1 &= -m(\mathrm{D}\boldsymbol{\omega}) \wedge \boldsymbol{r}, & \boldsymbol{F}_3 &= -m\boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge \boldsymbol{r}) \\ \boldsymbol{F}_2 &= -2m\boldsymbol{\omega} \wedge \mathrm{D}\boldsymbol{r}, & \boldsymbol{F}_4 &= -m\boldsymbol{A}, \end{aligned}$$

acting on the particle. To think of them as real forces, however, is to court disaster. They are nothing more than correction terms that compensate for the acceleration and rotation of R.

The first, $\boldsymbol{F}_1$ arises from the angular acceleration of R relative to $\tilde{R}$. The second, $\boldsymbol{F}_2$, is the 'Coriolis force'. It is orthogonal to the angular velocity $\boldsymbol{\omega}$ and to the velocity of the particle relative to R. It is responsible, for example, for the circulation of air around an area of low pressure, anti-clockwise in the northern hemisphere. The third, $\boldsymbol{F}_3$, is the 'centrifugal force', much exploited in fairgrounds. The fourth, $\boldsymbol{F}_4$, has the form of a uniform gravitational field. It cancels the Earth's gravitational field in a free-falling lift or aircraft.

The notation in (1.28) is not very convenient for calculation. It is simpler to use a dot for the time derivative D relative to $\mathcal{T}$ and to rewrite the equation of motion as

$$m\ddot{\boldsymbol{r}} + m\dot{\boldsymbol{\omega}} \wedge \boldsymbol{r} + 2m\boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge \boldsymbol{r}) + m\boldsymbol{A} = \boldsymbol{F}. \tag{1.29}$$

So long as we work entirely in the rotating frame R, and make no further reference to the auxiliary inertial frame $\tilde{R}$, then there is no danger of ambiguity.

Example 1.21

†A small bead can slide on a smooth wire in the shape of a circle of radius a. The wire is forced to rotate with constant angular speed ω about a vertical axis through the centre of the circle. This axis makes a constant angle $\alpha \in (0, \pi/2)$ with the normal to the plane of the circle. The problem is to find the positions at which the bead can remain at rest relative to the wire.

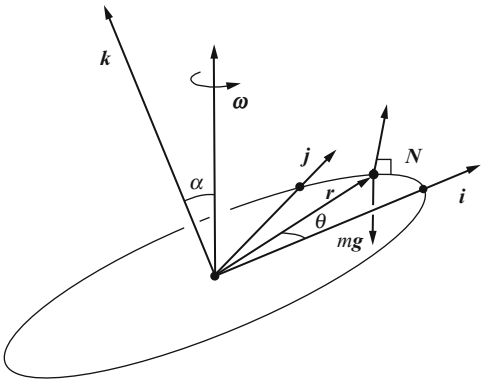

Figure 1.6

Choose the rotating frame R so that O is at the centre of the circle and $\mathcal{T} = (\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k})$, where $\boldsymbol{i}$, $\boldsymbol{j}$, and $\boldsymbol{k}$ are fixed relative to the wire. Choose their directions so that $\boldsymbol{k}$ is normal to the circle, $\boldsymbol{j}$ is horizontal, and the axis lies in the plane spanned by $\boldsymbol{i}$ and $\boldsymbol{k}$ (see Figure 1.6). Then

$$\boldsymbol{\omega} = \omega \sin\alpha\, \boldsymbol{i} + \omega \cos\,\alpha \boldsymbol{k}$$

and the position vector of the bead is

$$\boldsymbol{r} = a\cos\theta \boldsymbol{i} + a\sin\theta \boldsymbol{j}.$$

The forces on the bead are gravity

$$m\boldsymbol{g} = -mg\sin\alpha\, \boldsymbol{i} - mg\cos\alpha\, \boldsymbol{k}$$

and the normal reaction $\boldsymbol{N}$ of the wire, which is orthogonal to the wire. Thus the equation of motion is

$$m\ddot{\boldsymbol{r}} + 2m\boldsymbol{\omega} \wedge \dot{\boldsymbol{r}} + m\boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge \boldsymbol{r}) = m\boldsymbol{g} + \boldsymbol{N},$$

where the dot is the time derivative relative to R. By taking the scalar product with

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

which is orthogonal to $\boldsymbol{N}$ and parallel to $\boldsymbol{r}$, we obtain

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\theta} \cdot (\ddot{\boldsymbol{r}} + a\omega\cos\theta\sin\alpha\, \boldsymbol{\omega}) = ga\sin\alpha\sin\theta.$$

This reduces to

$$a^2\ddot{\theta} - a^2\omega^2\cos\theta\sin\theta\sin^2\alpha = ga\sin\alpha\sin\theta \tag{1.30}$$

on substituting

$$\ddot{\boldsymbol{r}} = -a\ddot{\theta}\sin\theta\boldsymbol{i} + a\ddot{\theta}\cos\boldsymbol{j} - a\dot{\theta}^2\cos\theta\boldsymbol{i} - a\dot{\theta}^2\sin\theta\boldsymbol{j}.$$

The positions of equilibrium relative to the wire are given by the constant solutions to (1.30). These are $\theta = \beta$, where

$$\sin\beta(g + a\omega^2\cos\beta\sin\alpha) = 0.$$

If $\omega^2\sin\alpha < g/a$, then there are just two positions at $\beta = 0$ and $\beta = \pi$. If $\omega^2\sin\alpha > g/a$, then there are two additional values of β for which

$$\cos\beta = -\frac{g}{a\omega^2\sin\alpha}.$$

If $\theta = \pi + \varepsilon$, then, to the first order in ε,

$$\ddot{\varepsilon} - \varepsilon\omega^2\sin^2\alpha + \frac{g}{a}\varepsilon\sin\alpha = 0,$$

Hence the lowest position is stable when $\omega^2\sin\alpha < g/a$ and unstable when $\omega^2\sin\alpha > g/a$.

Similarly, if $\theta = \varepsilon$, then, to the first order in ε,

$$\ddot{\varepsilon} - \varepsilon\omega^2\sin^2\alpha - \frac{g}{a}\varepsilon\sin\alpha = 0.$$

Hence the highest position is always unstable.

EXERCISES

1.13. Two frames R and $\tilde{R}$ are moving relative to each other without rotation. A particle has kinetic energy T measured in R and $\tilde{T}$ measured in $\tilde{R}$. Show that

$$T = \tilde{T} + \boldsymbol{u}\cdot\boldsymbol{p} + c,$$

where $\boldsymbol{p} = m\boldsymbol{v}$ is the momentum of the particle measure in R, $\boldsymbol{u}$ is the velocity of $\tilde{R}$ relative to R, and c is a constant.

1.14. †A small bead is threaded on a smooth wire in the shape of a curve given parametrically by $\boldsymbol{r} = \boldsymbol{r}(q)$. The wire rotates with constant angular velocity about a vertical axis. Show that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\tfrac{1}{2}\dot{q}^2\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}q}\cdot\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}q}\right] = \dot{q}\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}q}\cdot\ddot{\boldsymbol{r}},$$

where the dot is the time derivative relative to a frame rotating with the wire, and $\boldsymbol{r}$ is measured from an origin on the axis.

Deduce that

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}q} \cdot \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}q}\,\dot{q}^2 - (\boldsymbol{\omega} \wedge \boldsymbol{r}) \cdot (\boldsymbol{\omega} \wedge \boldsymbol{r}) - 2\boldsymbol{g} \cdot \boldsymbol{r} = \text{constant}.$$

Does this result still hold when the axis of rotation is not vertical?

1.15. A plastic ball is held at the bottom of a bucket of water and then released. As it is released, the bucket is dropped over the edge of a cliff. What happens?

1.16. In Example (1.21), investigate the stability of the two additional equilibrium points in the case $\omega^2 \sin\alpha > g/a$.

1.17. †A pendulum consists of a light rod and a heavy bob. Initially it is at rest in vertical stable equilibrium. The upper end is then made to move down a straight line of slope α (with the horizontal) with constant acceleration. Show that in the subsequent motion, the pendulum just becomes horizontal if $g = f(\cos\alpha + \sin\alpha)$.

1.8 The Rotation of the Earth

Suppose that the frame $R = (O, \mathcal{T})$ is fixed relative to the Earth, with the origin O on the Earth's surface, and that the inertial frame $\tilde{R} = (\tilde{O}, \tilde{\mathcal{T}})$ has its origin at the centre of the Earth (Figure 1.7). Then $\boldsymbol{x}$, the vector from $\tilde{O}$ to O, is constant relative to R because both O and $\tilde{O}$ are at rest relative to the Earth. Therefore $\ddot{\boldsymbol{x}} = 0 = \dot{\boldsymbol{x}}$, where the dot denotes the time derivative relative to $\mathcal{T}$, and so

$$\boldsymbol{A} = \boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge \boldsymbol{x}).$$

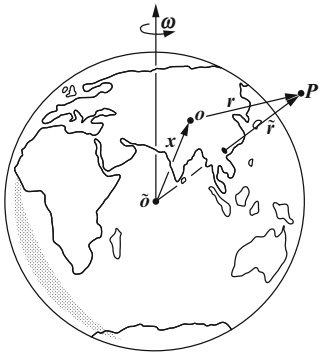

Figure 1.7

We assume that $\boldsymbol{\omega}$, which is the angular velocity of the Earth, is constant relative to the inertial frame. Then (1.29) reduces to

$$m\ddot{\boldsymbol{r}} + 2m\boldsymbol{\omega} \wedge \dot{\boldsymbol{r}} + m\boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge \boldsymbol{r}) + m\boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge \boldsymbol{x}) = m\boldsymbol{g} + \boldsymbol{T}, \tag{1.31}$$

where $\boldsymbol{g}$ is the gravitational acceleration and $\boldsymbol{T}$ is the resultant of any other forces acting on the particle.

If the Earth were not rotating, then we could measure the gravitational acceleration $\boldsymbol{g}$ by measuring the force $\boldsymbol{T}$ needed to keep the particle at rest relative to the Earth, for example by hanging the particle from a spring and calculating $\boldsymbol{T}$ from the extension and direction of the direction of the spring. But when we take account of the effects of the Earth's rotation, $\boldsymbol{T}$ is no longer equal to $-m\boldsymbol{g}$. If the particle is at rest relative to R at the point P with O-position vector $\boldsymbol{r}$, then $\dot{\boldsymbol{r}} = \ddot{\boldsymbol{r}} = 0$ and $\boldsymbol{T} = -m\boldsymbol{g}'$, where

$$\boldsymbol{g}' = \boldsymbol{g} - \boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge \boldsymbol{x}) - \boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge \boldsymbol{r}). \tag{1.32}$$

This vector is called the *apparent gravity*.

At the equator, where the difference is at its maximum, the magnitude of $\boldsymbol{g} - \boldsymbol{g}'$ is $3.4 \times 10^{-2}\,\mathrm{ms}^{-2}$, which is less than half of one per cent of the true gravitational acceleration.

By substituting (1.32) into (1.31), we obtain the final form of the equation of motion

$$m(\ddot{\boldsymbol{r}} + 2\boldsymbol{\omega} \wedge \dot{\boldsymbol{r}}) = \boldsymbol{T} + m\boldsymbol{g}'. \tag{1.33}$$

The apparent gravity depends on $\boldsymbol{r}$ both through the last term in its definition and through the dependence of $\boldsymbol{g}$ on P. For motion near O, however, we can ignore the variation, and treat $\boldsymbol{g}'$ as constant, relative to R.

At the Earth's surface, the variations in $\boldsymbol{g}$ over 10km are of the same order as the difference between real and apparent gravity. This puts a limit on the distance over which it is sensible to take account of the difference between $\boldsymbol{g}$ and $\boldsymbol{g}'$, but to ignore the variations in $\boldsymbol{g}'$.

Example 1.22 (Foucault's Pendulum)

The motion of Foucault's pendulum provides a concrete application of this form of the equation of motion. The pendulum is named after Léon Foucault's demonstration of the Earth's rotation at the Paris Exhibition in 1851. It consists simply of a bob of mass m suspended from a fixed point by a long wire of length a. It is so constructed that, when set in motion, it will swing for long enough for the Coriolis term $2m\boldsymbol{\omega}\wedge\dot{\boldsymbol{r}}$ in (1.33) to have an appreciable cumulative effect. In practice, additional energy is usually provided to keep the pendulum swinging.

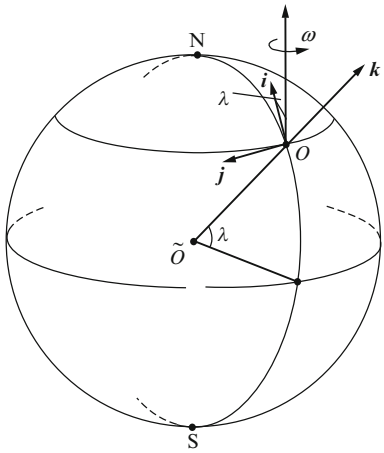

Figure 1.8

Let the origin O of R be the point of suspension and denote the unit vectors along the axes of R as $\boldsymbol{i}$, $\boldsymbol{j}$, and $\boldsymbol{k}$. We suppose that these have been chosen so that $\boldsymbol{k}$ is along the apparent vertical and $\boldsymbol{i}$ is a horizontal vector pointing due north. Then,

$$\begin{aligned} \boldsymbol{g}' &= -g'\boldsymbol{k} \\ \boldsymbol{\omega} &= \Omega(\cos\lambda\,\boldsymbol{i} + \sin\lambda\,\boldsymbol{k}) \\ \boldsymbol{r} &= r\cos\theta\,\boldsymbol{i} + r\sin\theta\,\boldsymbol{j} + z\boldsymbol{k}, \end{aligned}$$

where Ω is the angular speed of the Earth ($2\pi/24$ radians per hour), λ is the apparent latitude, which is defined to be the angle between $\boldsymbol{g}'$ and the equatorial plane, and r, θ, and z are the cylindrical polar coordinates of the bob. See Figures 1.8 and 1.9. Note that $z < 0$ in the configuration shown.

The equation of motion is

$$m(\ddot{\boldsymbol{r}} + 2\boldsymbol{\omega} \wedge \dot{\boldsymbol{r}}) = -mg'\boldsymbol{k} + \boldsymbol{T}, \tag{1.34}$$

where $\boldsymbol{T}$ is the tension in the wire and $\boldsymbol{r}$ is subject to the constraint $\boldsymbol{r} \cdot \boldsymbol{r} = a^2$, which fixes the length of the wire. Note that the constraint implies that $\boldsymbol{r} \cdot \dot{\boldsymbol{r}} = 0$, so that $\dot{\boldsymbol{r}}$ is orthogonal to $\boldsymbol{r}$ and hence also to $\boldsymbol{T}$, which is along the wire. By taking the scalar product of (1.34) with $\dot{\boldsymbol{r}}$, and by cancelling m, we obtain

$$\dot{\boldsymbol{r}} \cdot \ddot{\boldsymbol{r}} = -g'\boldsymbol{k} \cdot \dot{\boldsymbol{r}} = -g'\dot{z}$$

which, on integration with respect to time, yields the *energy conservation equation* relative to R,

$$\tfrac{1}{2}\dot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}} + g'z = E \tag{1.35}$$

in which E is a constant.

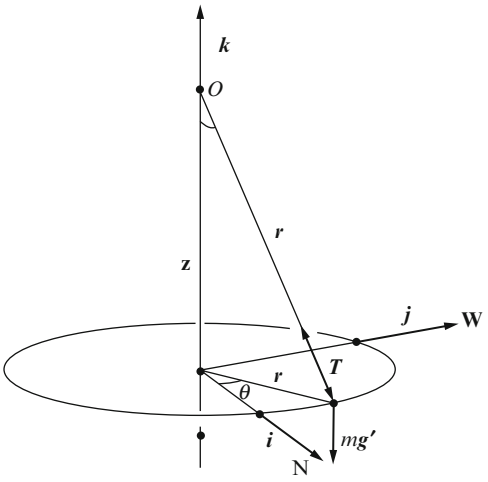

Figure 1.9

Similarly, by taking the vector product with $\boldsymbol{r}$ and then the scalar product with $\boldsymbol{k}$, we obtain the *angular momentum equation*

$$\dot{h} - 2(\boldsymbol{k} \cdot \dot{\boldsymbol{r}})(\boldsymbol{r} \cdot \boldsymbol{\omega}) = 0 \tag{1.36}$$

in which $h = \boldsymbol{k} \cdot (\boldsymbol{r} \wedge \dot{\boldsymbol{r}})$. Now

$$\dot{\boldsymbol{r}} = \dot{r}(\cos\theta\, \boldsymbol{i} + \sin\theta\, \boldsymbol{j}) + \dot{\theta}(-r\sin\theta\, \boldsymbol{i} + r\cos\theta\, \boldsymbol{j}) + \dot{z}\boldsymbol{k},$$

and by substituting this into (1.36) and (1.35), we obtain

$$\dot{h} = 2\dot{z}\Omega(r\cos\theta\cos\lambda + z\sin\lambda) \tag{1.37}$$

$$\tfrac{1}{2}\left(\dot{r}^2 + \dot{z}^2 + \frac{h^2}{r^2}\right) + g'z = E \tag{1.38}$$

together with $h = r^2\dot{\theta}$. The two equations (1.37) and (1.38), together with the constraint $z^2 + r^2 = a^2$, determine r, θ, and z in terms of the initial position and velocity. The second equation (1.38) does not involve the latitude explicitly.

When $\lambda = \pi/2$, so that the pendulum is over the North Pole, the first equation (1.37) reduces to

$$\dot{h} = 2z\dot{z}\Omega. \tag{1.39}$$

In this case, we know what happens without solving the equations: O is also at rest relative to the inertial frame and the Earth simply turns under the swinging pendulum. The rotation has no other effect on the motion. So, for example, if the bob is pulled to one side and released, then the pendulum swings in a

vertical plane, but this plane rotates relative to the Earth through a complete revolution once every 24 h, in a clockwise direction.

For a general value of λ, (1.37) is

$$\dot{h} = 2z\dot{z}\Omega' + 2\dot{z}r\Omega\cos\theta\cos\lambda$$

where $\Omega' = \Omega\sin\lambda$. Apart from the replacement of Ω by Ω', this differs from (1.39) only by the term

$$2\dot{z}r\Omega\cos\theta\cos\lambda.$$

For small oscillations, for which $r/a \ll 1$, this is less significant than the first term on the right-hand side because $|r/z|$ is then much less that 1. If we neglect this term, then we obtain an equation identical to (1.39), but with Ω replaced by Ω'. At a general latitude, therefore, the motion is the same as at the North Pole, except that the rate of rotation is $\Omega\sin\lambda$, rather than Ω. For oscillations in a vertical plane at 52°N, for example, the plane of oscillation turns through a complete revolution once every 30.5 h.

Although we have not considered the motion in any detail, this general discussion has been sufficient to derive the result on which Foucault based his demonstration. A more thorough analysis, such as is given in Synge and Griffith's *Principles of mechanics* [10], reveals, however, a serious difficulty: there is another effect which causes a rotation of the plane of oscillation and which can completely mask the effect of the Coriolis term. Without solving the full equations, we can see the origin of the second effect by using phase-plane techniques. We come back to this in Example 2.4.

1.9 The Kinematics of Rigid Bodies

A collection of particles make up a *rigid body* if there exists a frame R relative to which all the particles are at rest at all times. Such a frame is called a *rest frame* of the body. In general, rest frames are *not* inertial.

Equivalently, a rigid body is a collection of particles separated by fixed distances. By taking a limit as the number of particles goes to infinity, we can also think of a rigid body as a continuous distribution of matter, with the separation between different elements remaining constant as the body moves. The exact form of the limiting process is not easy to tie down, but in any case both the 'finite collection of particles' and the 'rigid distribution of matter' models are of limited validity.

Provided that the particles do not all lie along a line, the rest frame is determined by the particles up to a rotation of the axes by a constant orthogonal transformation and a translation of the origin by a constant vector, that is, a

vector constant in R. In particular, all the rest frames of a general rigid body are equivalent. So a rigid body has a well-defined angular velocity relative to some other frame $\tilde{R}$, namely the angular velocity of any rest frame relative to $\tilde{R}$. If $\tilde{R}$ is inertial, then we speak simply of the 'angular velocity of the body'.

Proposition 1.23

Let P and Q be two points fixed a rigid body and let $\boldsymbol{r}$ be the vector from Q to P. Then the velocities $\boldsymbol{v}_P$ and $\boldsymbol{v}_Q$ of P and Q relative to a frame $\tilde{R}$ are related by

$$\boldsymbol{v}_P = \boldsymbol{v}_Q + \boldsymbol{\omega} \wedge \boldsymbol{r} \tag{1.40}$$

where $\boldsymbol{\omega}$ is the angular velocity of the body relative to $\tilde{R}$.

Proof

Let $\boldsymbol{r}_P$, $\boldsymbol{r}_Q$ denote the respective position vectors of P and Q from the origin of $\tilde{R}$. Then $\boldsymbol{r}_P = \boldsymbol{r}_Q + \boldsymbol{r}$ and

$$\boldsymbol{v}_P = \tilde{\mathrm{D}}\boldsymbol{r}_P, \qquad \boldsymbol{v}_Q = \tilde{\mathrm{D}}\boldsymbol{r}_Q$$

Therefore

$$\boldsymbol{v}_P = \boldsymbol{v}_Q + \tilde{\mathrm{D}}\boldsymbol{r}.$$

However $\boldsymbol{r}$ is constant in R, so $\tilde{\mathrm{D}}\boldsymbol{r} = \boldsymbol{\omega} \wedge \boldsymbol{r}$ by the Coriolis theorem. □

At any instant, therefore, the motion of the body relative to $\tilde{R}$ is completely specified by two vectors: the angular velocity $\boldsymbol{\omega}$ and the velocity of any one particle of the body. We use (1.40) to define $\boldsymbol{v}_P$ at points outside the body, in an 'imaginary extension' of the body.

Definition 1.24

A line L is an *instantaneous axis of rotation* at time t if $\boldsymbol{v}_P(t) = 0$ for every $P \in L$.

If one can find an instantaneous axis L, either in the body or in its imaginary extension, then it is easy to picture the motion: the points on L are at rest relative to $\tilde{R}$ at time t and the body is 'rotating about L'. In general, however, every point of the body is moving relative to $\tilde{R}$ and there is no instantaneous axis, as the following argument shows. From (1.40),

$$\boldsymbol{v}_P \cdot \boldsymbol{\omega} = \boldsymbol{v}_Q \cdot \boldsymbol{\omega}$$

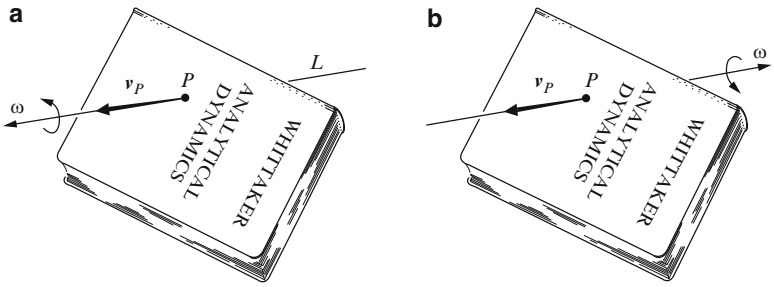

Figure 1.10 (a) Right-handed motion, (b) left-handed motion.

Therefore $v_P \cdot \boldsymbol{\omega}$ takes the same value at every point. If the body is set in motion with $\boldsymbol{v}_Q$ not orthogonal to $\boldsymbol{\omega}$, then $\boldsymbol{v}_P$ cannot vanish for any P.

The converse is also true. If $\boldsymbol{v}_Q \cdot \boldsymbol{\omega} = 0$ and $\boldsymbol{\omega} \neq Q$, then the line L given in terms of a parameter s by

$$\boldsymbol{r} = \frac{1}{\boldsymbol{\omega} \cdot \boldsymbol{\omega}} \boldsymbol{\omega} \wedge \boldsymbol{v}_Q + s\boldsymbol{\omega} \tag{1.41}$$

is an instantaneous axis.

The line L given by (1.41) is still significant even when $\boldsymbol{v}_Q \cdot \boldsymbol{\omega} \neq 0$. It is characterized by the property that for any P on L, $\boldsymbol{v}_P$ is proportional to $\boldsymbol{\omega}$. In general, therefore, one can picture the body as having a 'corkscrew' motion, moving along L as it rotates about it. If $\boldsymbol{v}_Q \cdot \boldsymbol{\omega} > 0$, then $\boldsymbol{v}_P$ is in the same direction as $\boldsymbol{\omega}$ on L and the corkscrew is right-handed. That is, when the thumb of the right hand points in the direction of $\boldsymbol{v}_P$, the fingers curl in the sense of the rotation. If $\boldsymbol{v}_Q \cdot \boldsymbol{\omega} < 0$, then $\boldsymbol{v}_P$ is anti-parallel to $\boldsymbol{\omega}$ and the corkscrew is left-handed (Figure 1.10).

Definition 1.25

A motion of a rigid body is *right-handed* (*left-handed*) relative to $\tilde{R}$ if $\boldsymbol{v} \cdot \boldsymbol{\omega} > 0$ ($\boldsymbol{v} \cdot \boldsymbol{\omega} < 0$), where $\boldsymbol{v}$ is the velocity of any point of the body.

Example 1.26 (Rolling Conditions)

If two rigid bodies are in contact at a point P and if there is no slipping, then the particles of the two bodies at P must have the same velocities. For example, consider a sphere of radius a rolling without slipping on a rough plane. Let $\boldsymbol{v}$ denote the velocity of the centre and let $\boldsymbol{\omega}$ denote the angular velocity of the sphere, both relative to a frame fixed on the plane. Then the velocity of the

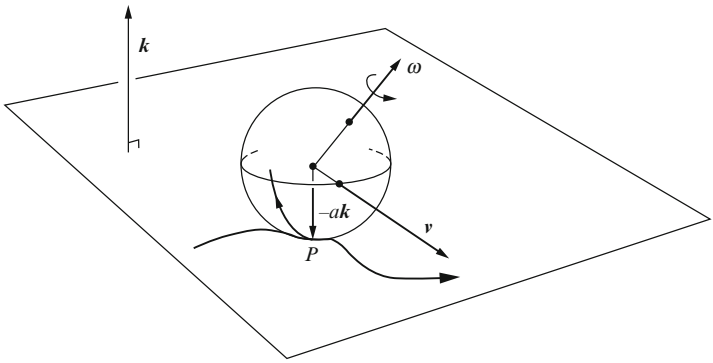

Figure 1.11 Rolling: the curves are the loci of the point of contact on the plane and on the sphere.

particle of the sphere in contact with the plane is $\boldsymbol{v} + \boldsymbol{\omega} \wedge (-a\boldsymbol{k})$, where $\boldsymbol{k}$ is the unit normal to the plane. The corresponding particle of the plane is at rest. Hence we have the rolling condition

$$\boldsymbol{v} - a\boldsymbol{\omega} \wedge \boldsymbol{k} = 0. \tag{1.42}$$

There is a potential confusion here. Equation (1.42) is the condition that the particle of the sphere in contact with the plane should be instantaneously at rest. However, as the sphere rolls, different particles come into contact and the *point* of contact moves both relative to the sphere and relative to the plane (Figure 1.11).

Example 1.27 (A Rolling Problem)

†A rough hollow sphere S of radius $3a$ is free to rotate about its centre O. Inside S are five rough solid spheres S_0, S_1, S_2, S_3, S_4, each of radius a (Figure 1.12). Initially the centre of S_0 is at O and the centres of S_1, S_2, S_3, and S_4 are at the vertices of a regular tetrahedron; S_1, S_2, S_3, and S_4 are in contact with the inner surface of S and with the outer surface of S_0. No slipping occurs.

The problem is to show that if S is forced to rotate about O with variable angular velocity, then at all times during the subsequent motion the centres of S_1, S_2, S_3, and S_4 form the vertices of a regular tetrahedron.

The first step is to establish the following: suppose that n moving points have position vectors $\boldsymbol{r}_1(t)$, $\boldsymbol{r}_2(t)$,..., $\boldsymbol{r}_n(t)$ from the origin of some frame $\tilde{R}$. If there exists a *single* time-dependent vector $\boldsymbol{\theta}(t)$ such that

$$\dot{\boldsymbol{r}}_i = \boldsymbol{\theta} \wedge \boldsymbol{r}_i, \quad i = 1, 2, \ldots, n, \tag{1.43}$$

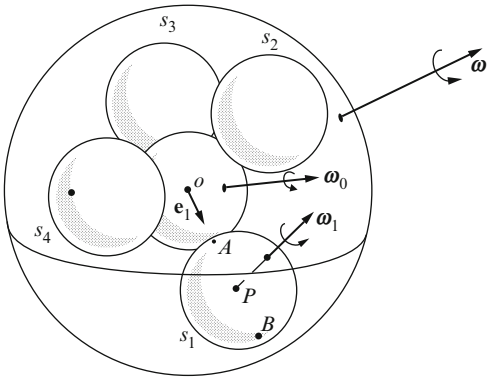

Figure 1.12

where the dot is the time derivative relative to $\tilde{R}$, then the distances between the points are constant. This follows from

$$\frac{\mathrm{d}}{\mathrm{d}t}\big((\boldsymbol{r}_i-\boldsymbol{r}_j)\cdot(\boldsymbol{r}_i-\boldsymbol{r}_j)\big)=2\big(\boldsymbol{\theta}\wedge(\boldsymbol{r}_i-\boldsymbol{r}_j)\big)\cdot(\boldsymbol{r}_i-\boldsymbol{r}_j)=0.$$

Now consider S_1. Denote its angular velocity by $\boldsymbol{\omega}_1$, and denote the respective angular velocities of S_0 and S by $\boldsymbol{\omega}_0$ and $\boldsymbol{\omega}$. Let $\boldsymbol{e}_1$ be the unit vector from O towards the centre P of S_1, so that the O-position vector of P is $\boldsymbol{r}_1=2a\boldsymbol{e}_1$.

The rolling condition at the point A of contact between S_1 and S_0 is

$$\boldsymbol{\omega}_0\wedge(a\boldsymbol{e}_1)=\dot{\boldsymbol{r}}_1+\boldsymbol{\omega}_1\wedge(-a\boldsymbol{e}_1) \tag{1.44}$$

because the left-hand side is the velocity of the particle of S_0 at A and the right-hand side is the velocity of the particle of S_1 at A. Similarly, the rolling condition at the point B of contact between S_1 and S is

$$\boldsymbol{\omega}\wedge(3a\boldsymbol{e}_1)=\dot{\boldsymbol{r}}_1+\boldsymbol{\omega}_1\wedge(a\boldsymbol{e}_1).$$

By adding,

$$\dot{\boldsymbol{r}}_1=\frac{a}{2}(3\boldsymbol{\omega}+\boldsymbol{\omega}_0)=(\tfrac{3}{4}\boldsymbol{\omega}+\tfrac{1}{4}\boldsymbol{\omega}_0)\wedge\boldsymbol{r}_1.$$

By the same argument applied to S_2, S_3, and S_4,

$$\boldsymbol{r}_i=\boldsymbol{\theta}\wedge\boldsymbol{r}_i,\qquad i=1,2,3,4, \tag{1.45}$$

where $\boldsymbol{\theta}=\frac{3}{4}\boldsymbol{\omega}+\frac{1}{4}\boldsymbol{\omega}_0$ and the $\boldsymbol{r}_i$s are the O-position vectors of the centres of S_1, S_2, S_3, and S_4. Therefore the distances between the centres remain constant and the result follows.

EXERCISES

1.18. Establish the properties of the line L defined by (1.41) for the general motion of a rigid body. That is, show that if P is a point of L, then $\boldsymbol{v}_P$ is proportional to $\boldsymbol{\omega}$; and show conversely that if $\boldsymbol{v}_P$ is proportional to $\boldsymbol{\omega}$, then P lies on L. (Assume that $\boldsymbol{\omega} \neq 0$.)

1.19. A rigid body has a right-handed motion relative to a frame R. If the body is observed in a mirror, does the motion appear to be right or left-handed? If the motion is recorded and the recording is then run backwards, does the motion appear to be right- or left-handed?

1.20. †A rigid body has angular velocity $\boldsymbol{\omega}$ and has one point O fixed relative to a frame $\tilde{R}$. Show that if $\boldsymbol{\omega} \wedge \tilde{\mathrm{D}}\boldsymbol{\omega} \neq 0$, then O is the only point with zero acceleration relative to $\tilde{R}$.

1.21. Show that the motion of a rigid body is determined at any instant by the velocities of three non-collinear points.

Three particles A, B, and C have velocities $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{w}$ respectively relative to a frame R. Show that they can belong to a rigid body if and only if

$$(\boldsymbol{a} - \boldsymbol{b}) \cdot (\boldsymbol{u} - \boldsymbol{v}) = (\boldsymbol{b} - \boldsymbol{c}) \cdot (\boldsymbol{v} - \boldsymbol{w}) = (\boldsymbol{c} - \boldsymbol{a}) \cdot (\boldsymbol{w} - \boldsymbol{u}) = 0$$

where $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ are the position vectors of A, B, and C from the origin of R.

1.22. Two identical coins lie on a table top in contact. If the second coin is rolled around the first, through what angle does the radius from the centre of the first coin to the point of contact turn before the second coin has rotated through an angle π? Check your answer experimentally.

1.23. A sphere of radius a is rolling without slipping on a rough horizontal plane in such a way that its centre traces out a horizontal circle, radius b and centre O, with constant angular speed Ω. Let $(\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k})$ be an orthonormal triad with $\boldsymbol{k}$ vertical and $\boldsymbol{i}$ in the direction from O to the centre of the sphere. Show that the angular velocity of the sphere relative to the plane satisfies

$$\boldsymbol{\omega} = n\boldsymbol{k} - \frac{b}{a}\Omega\boldsymbol{i}$$

where $n = \boldsymbol{\omega} \cdot \boldsymbol{k}$. Show that if n is constant, then the locus of the point of contact on the sphere is a circle. What is its radius?

2
One Degree of Freedom

2.1 The Equation of Motion

A system with one degree of freedom is one in which just one coordinate is needed to determine the configuration. The motion is specified by expressing the coordinate as a function of time. Typically this involves solving a second-order differential equation, called the *equation of motion*.

An obvious example is a particle of mass m moving along the x-axis under the influence of a force F, where the equation of motion is

$$m\ddot{x} = F.$$

The force could depend on the position of the particle, as in the case of tension in a spring, or its velocity, as in the case of air resistance, or on time, or on all three. In the latter case, we are confronted with the problem of solving an ordinary differential equation of the form

$$m\ddot{x} = F(x, \dot{x}, t). \tag{2.1}$$

Other examples are a pendulum swinging in a vertical plane, where the coordinate is the angle with the vertical, or a system of interlocking cog-wheels in a machine, such as old-fashioned watch, where there are a large number of moving parts, but the rotation of one part determines the configuration of the whole system, so there is only a single degree of freedom.

In coming to understand the dynamical behaviour of such systems, we are faced with a number of problems.

N.M.J. Woodhouse, *Introduction to Analytical Dynamics: New Edition*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-1-84882-816-2_2,

1. How to determine the equation of motion from the physical description of the system.
2. How to transform the equation into a simpler form by changing the coordinate.
3. How to find an exact solution of the equation of motion, when this is possible.
4. How to find approximate solutions when it is not.
5. How to understand the general nature of the motion when an exact solution is not possible.

The third problem (3) opens the door to a vast area of theory that we do not explore in any depth, although we assume familiarity with the simpler techniques for solving ordinary differential equations. In looking at the other problems in the list, we concentrate in this chapter on ideas that extend in a straightforward way to systems with many degrees of freedom, and on ideas that illuminate the general behaviour of mechanical systems and shed light on the underlying physical processes. We pay particular heed to aspects of classical mechanics which connect with quantum theory and relativity.

2.2 The Phase Plane

A key idea that extends in a natural way to the more general setting is to analyse the motion in the phase plane. Here we reinterpret (2.1) as a system of first order differential equations by treating the velocity as a new dependent variable. That is, we put $v = \dot{x}$ and rewrite (2.1) in the form

$$m\dot{v} = F(x, v, t), \qquad \dot{x} = v.$$

There is a unique solution for each choice of initial values $x(0) = x_0$, $v(0) = v_0$, so for each choice of initial point (x_0, v_0), the motion traces out a curve

$$x = x(t), \qquad v = v(t) \tag{2.2}$$

in the x, v-plane.

In this context, the x, v-plane is called the *phase plane*, and is denoted by P. The dynamical evolution of the system is pictured as a flow in the phase plane, analogous to a fluid flow, with a point (x, v) flowing along the solution curve (2.2) from the initial point (x_0, v_0) as t increases. The solution curves are the *dynamical trajectories* of the system.

Sometimes it is possible to determine broad features of the dynamical behaviour by considering qualitative aspects of the flow in the phase plane without actually solving the equation of motion. A simple case is that of a system which is *autonomous* and *conservative*. To say that it is autonomous means that F depends only on x and v, but not on t. It then follows that the system behaves in the same way for given initial values of x and v whenever it is set in motion. To say that it is also conservative means that F is independent of v and is given by

$$F = -\frac{\mathrm{d}U}{\mathrm{d}x}$$

for some *potential function* $U = U(x)$. So if it is both conservative and autonomous, we can write the equation of motion in the form

$$m\ddot{x} + \frac{\mathrm{d}U}{\mathrm{d}x} = 0,$$

In this case, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\tfrac{1}{2}m\dot{x}^2 + U\right) = m\dot{x}\ddot{x} + \dot{x}\frac{\mathrm{d}U}{\mathrm{d}x} = 0.$$

So

$$E = \tfrac{1}{2}mv^2 + U(x)$$

is constant along the dynamical trajectories in the phase plane. It is the total energy. By sketching the level curves of E, it is possible to deduce the general behaviour of the system without solving the equation of motion explicitly.

Example 2.1 (The Harmonic Oscillator)

A particle of mass m moves along the x-axis under the influence of the force $F = -mx$. The equation of motion is

$$m\ddot{x} = -mx \tag{2.3}$$

The system is conservative and autonomous, with potential $U = \frac{1}{2}mx^2$. Its equations of motion in the phase plane are

$$m\dot{v} = -mx, \qquad \dot{x} = v. \tag{2.4}$$

Because these imply that $E = \frac{1}{2}m(x^2 + v^2)$ is constant, the trajectories are circles.

Figure 2.1 illustrates the 'phase portrait' of the system. The arrows point in the direction of increasing x when $v = \dot{x} > 0$, and in the direction of decreasing x when $v = \dot{x} < 0$. On each trajectory, x oscillates between two equal and opposite extremes, namely the intersection points of the circle with the x-axis.

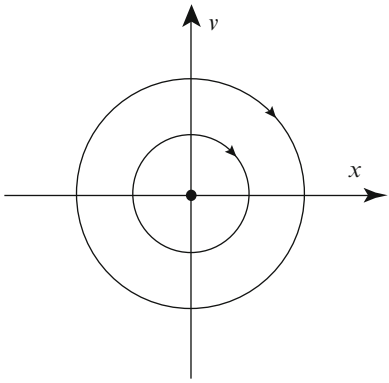

Figure 2.1

The only point at which the particle can remain at rest is the origin $x = 0$. This is a position of *stable* equilibrium because if the particle is displaced slightly, then its trajectory in P is a small circle centre $(0, 0)$, on which x and v remain close to zero.

Example 2.2

If instead $F = mx$, then the trajectories are the hyperbolas

$$v^2 - x^2 = \text{constant}$$

and the origin is a position of unstable equilibrium. There are several different types of trajectory, illustrated in Figure 2.2.

(A) $v^2 - x^2 < 0$, $x > 0$: the particle approaches the origin from $x = \infty$; x comes to a minimum (the intersection point with the x-axis) and then increases again without ever reaching $x = 0$.

(B) $v^2 - x^2 = 0$, $x > 0$, $v > 0$: here $\dot{x} = x$ and the particle moves away from the equilibrium position with x increasing exponentially with time.

(C) $v^2 - x^2 = 0$, $x > 0$, $v < 0$: here $\dot{x} = -x$. The particle approaches the origin, with x decreasing like e^{-t}, but never reaches it.

(D) $v^2 - x^2 > 0$, $v > 0$: the particle approaches the origin from $x = -\infty$, slows down, passes through $x = 0$, and then accelerates again towards $x = \infty$.

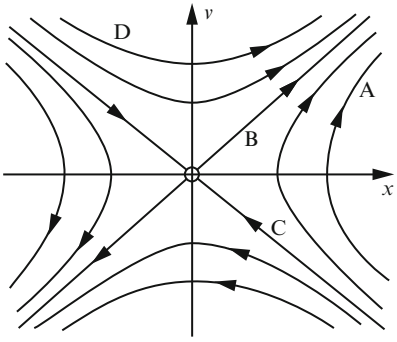

Figure 2.2

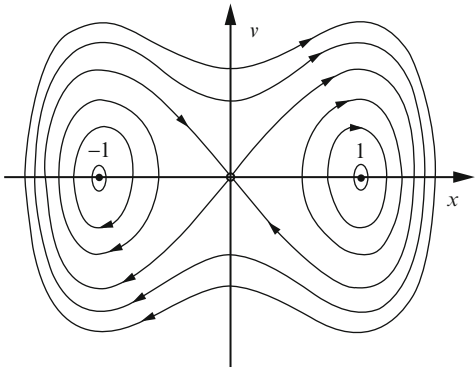

Figure 2.3

Example 2.3

A more interesting case is $F = 2mx(1 - x^2)$, for which the trajectories are

$$(x-1)^2(x+1)^2 + v^2 = \text{constant} \tag{2.5}$$

(Figure 2.3). There are two positions of stable equilibrium at $x = -1$ and $x = 1$, near which the system behaves like the harmonic oscillator; and one position of unstable equilibrium at $x = 0$, near which it behaves like the previous example, although the behaviour for larger values of $|x|$ is very different.

In general, if $F = F(x)$, then the equilibrium points are the values of x for which $F(x) = 0$. If $F'(x) < 0$ at equilibrium, then $F(x + \varepsilon)$ is negative for small positive ε and positive for small negative ε, so the force nearby is towards x, making the equilibrium stable. If $F'(x) > 0$, then the force near x is directed away from x and the equilibrium is unstable. We consider motion near equilibrium more systematically in Chapter 7.

'Phase plane analysis' is a powerful tool for tackling nonlinear ordinary differential equations in a general context [11]. As an illustration, we return to Foucault's pendulum.

Example 2.4 (The Area Effect)

This is a continuation of Example 1.22. By dividing (1.38) by $\frac{1}{2}\dot{\theta}^2$, and substituting $h = r^2\dot{\theta}$ and $r^2 = a^2 - z^2$ (which implies that $r\dot{r} = -z\dot{z}$), we obtain

$$v^2 = \frac{(a^2 - z^2)^2}{a^2 h^2}(F(z) - h^2) \tag{2.6}$$

where $v = \mathrm{d}z/d\theta$ and

$$F(z) = 2(E - g'z)(a^2 - z^2). \tag{2.7}$$

See Figure 2.4. Apart from the fact that z rather than x is the dependent variable and θ rather than t is the independent variable, this is the same sort of equation as those that arose in the previous examples.

As the pendulum oscillates, both v and z vary as functions of θ and so the different possible motions give rise to a family of curves in the z, v-plane, the 'phase plane' for this problem. See Figure 2.5.

For the moment, put $\Omega = 0$, so the rotation of the Earth is ignored. Then $g' = g$, $\dot{h} = 0$, and the phase plane curves, which are given by (2.6), are labelled by the various values of the constants E and h. We must exclude the singular case $h = 0$ in which the pendulum oscillates in a vertical plane and θ is not a good parameter.

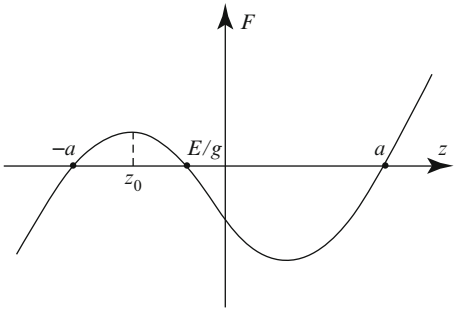

Figure 2.4 The graph of $F(z)$.

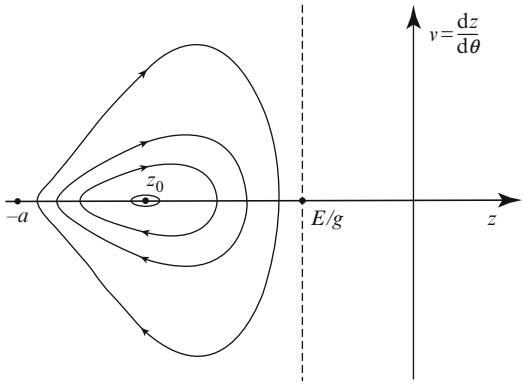

Figure 2.5 All the trajectories in the z, v-plane lie to the left of $z = E/g$.

Fix a value of $E > -ag$. Then there is a critical value h_0 of h for which the cubic $F(z) - h^2$ has a repeated root $z = z_0$. Because z_0 is also a root of $F'(z)$, we have

$$g(a^2 - z_0^2) + 2z_0(E - gz_0) = 0. \tag{2.8}$$

Hence

$$z_0 = \frac{1}{3g}(E - \sqrt{E^2 + 3a^2g^2}), \tag{2.9}$$

the sign of the square root being chosen by considering the graph of $F(z)$. The phase plane trajectory labelled by $h = h_0$ is the single point $(z_0, 0)$. In this case the path of the bob is a horizontal circle in the plane $z = z_0$.

We now look at what happens when the motion is close to this horizontal circle. The argument involves two approximations. We assume that the radius of the circle is small compared with a, that is that $a - |z_0| \ll a$. We further assume that the departure of the path of the bob from the circle is small compared with the radius of the circle, that is, that $|z - z_0| \ll a - |z_0|$.

When h is near the critical value h_0, so that $0 < |h_0 - h| \ll |h_0|$, we have $z = z_0 + y$ where $|y| \ll a - |z_0|$. By ignoring terms of order y^3, we find from (2.6) that the dependence of y on θ is determined by

$$\frac{1}{2}\left(\frac{dy}{d\theta}\right)^2 = \frac{(a^2 - z_0^2)^2}{2a^2h_0^2}\left(h_0^2 - h^2 + \tfrac{1}{2}y^2F''(z_0)\right).$$

This requires a little thought: the crucial point is that $h_0^2 - h^2$ is of the same order as y^2. On differentiating with respect to y,

$$\frac{d^2y}{d\theta^2} + \lambda^2 y = 0$$

where

$$\lambda^2 = -\frac{(a^2 - z_0^2)^2 F''(z_0)}{2a^2 h_0^2} = \frac{2z_0(3gz_0 - E)}{a^2 g}. \tag{2.10}$$

Here we have used (2.8) together with $h_0^2 = F(z_0)$.

Therefore y, and hence also z, oscillate as functions of θ with period $2\pi/\lambda$. The bob moves close to a horizontal circle, but it rises and falls slightly as it rotates about the vertical axis.

If λ were an integer, then we would have $z(0) = z(2\pi)$ and the bob would return to its initial position after each complete circuit. This is what happens in the limit $E \to -ag$, $z_0 \to -a$, when the circular path collapses to a point. In that case, (2.10) gives $\lambda^2 = 4$.

But if E is slightly above the limiting value, so that $E = -ag + \varepsilon g$ where $|y| \ll \varepsilon \ll a$, then

$$z_0 = -a + \tfrac{1}{2}\varepsilon + O(\varepsilon^2/a^2), \tag{2.11}$$

from (2.8), and

$$\lambda^2 = 4(1 - 3\varepsilon/4a) + O(\varepsilon^2/a^2), \tag{2.12}$$

from (2.10). Thus the period of z as a function of θ is

$$\pi\left(1 + \frac{3\varepsilon}{8a}\right) = \pi + \frac{3A}{8a^2} \tag{2.13}$$

where A is the area of the circular trajectory. On each complete revolution, the point at which z reaches its maximum advances through an angle $3A/4a^2$, in the same sense as the bob's motion. This is the *area effect*.

The derivation is based on the assumption that $|y| \ll \varepsilon \ll a$. A more careful analysis shows that the area effect is given by the same expression when y is comparable with ε. In that case, the projection of the path of the bob into the horizontal r, θ-plane is approximately an ellipse with its centre at the origin, but the axes of the ellipse rotate through an angle $3A/4a^2$ in the forward sense on each revolution, where A is the area of the ellipse.

When $\Omega \neq 0$, the area effect is superimposed on the Coriolis rotation and dominates it unless the pendulum is set in motion in such a way that $A^2 g \ll \Omega^2 a^5$.

EXERCISES

2.1. Sketch the phase portrait and describe the dynamical trajectories for a particle of mass m moving under the following forces

(1) $F = m(\frac{3}{2}x^2 - x)$

(2) $F = 2mx(x^2 - 1)$

(3) $F = -m \sin x$

(4) $F = -mx\dot{x}/(1 + \dot{x})$

In the fourth case, what at happens as the trajectories approach the line $v = -1$ in the phase plane?

2.3 Phase Space and Extended Phase Space

In some problems, the coordinate labelling the configuration of the system is an angle rather than a distance. An example is a simple pendulum, consisting of a particle of mass m attached to one end of a light rod. The other end is smoothly pivoted at a fixed point O, and the rod can swing freely under gravity is a vertical plane. The equation of motion is

$$\ddot{\theta} = -g \sin \theta,$$

where θ is the angle between the rod and the downward vertical. We think of the system as consisting of a single particle constrained to move on a circle in a vertical plane.

Because θ and $\theta + 2\pi$ label the same configuration of the pendulum, we should represent the space of configurations as a circle rather than a line. Instead of the phase plane, we consider the cylindrical *phase space*. The angular coordinate θ increases around the circular cross-sections and the velocity coordinate $v = \dot{\theta}$ increases along the generators of the cylinder. The trajectories in the phase space are illustrated in Figure 2.6. There is a position of stable equilibrium at $\theta = 0$, shown at the front of the cylinder, and a position of unstable equilibrium at $\theta = \pi$, shown at the back.

We denote the phase space by P. In problems with one degree of freedom, P is either the phase plane or a cylinder, as in the example above. In general problems, with more degrees of freedom, both the configuration space and the phase space can have more elaborate topologies.

In an autonomous system, there is a unique trajectory in P for given initial values of x and v, independently of the time when the motion is initiated. That is, if $x = x(t)$, $v = v(t)$ is the solution with

$$x(0) = x_0, \qquad v(0) = v_0,$$

then $\tilde{x} = x(t - \tau)$, $\tilde{v} = v(t - \tau)$ is the solution with initial conditions

$$\tilde{x}(\tau) = x_0, \qquad \tilde{v}(\tau) = v_0.$$

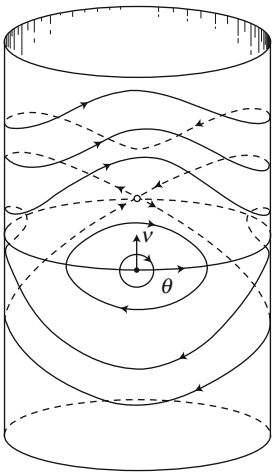

Figure 2.6 The phase space of the simple pendulum.

The path traced out in P is the same in both cases; only the parametrization by time has changed. The system behaves in the same way whether it is set in motion at time 0 or at time τ.

In non-autonomous systems, where F in (2.1) depends on t as well as on x and v, this is no long true because the motion may change when the system is set in motion at the later time, but with the same initial position and velocity. So the trajectory in P will generally be different in the two cases.

It is helpful in such cases to represent the dynamical behaviour in terms of a flow not in P, but in the *extended phase space* $P \times \mathbb{R}$, with coordinates x, v, t (or θ, v, t in the example of the pendulum). In both autonomous and non-autonomous systems, there is a unique trajectory through each point of $P \times \mathbb{R}$. The trajectory through (x_0, v_0, t_0) has parametric form

$$t \mapsto \big(x(t), v(t), t\big)$$

where $v(t) = \dot{x}(t)$, and $x(t)$ is the unique solution to the equation of motion with

$$x(t_0) = x_0, \qquad \dot{x}(t_0) = v_0.$$

For the harmonic oscillator, for example, the trajectories in extended phase space are the helices

$$t \mapsto \big(x_0 \cos(t - t_0) + v_0 \sin(t - t_0), v_0 \cos(t - t_0) - x_0 \sin(t - t_0), t\big).$$

(See Figure 2.7.)

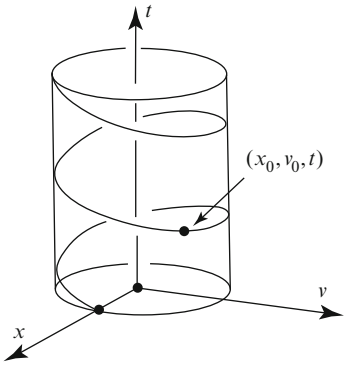

Figure 2.7

To allow for the possibility that it might be either a distance or an angle or some more general quantity, it is conventional to denote the position coordinate by q rather that by x or θ. The coordinates on P are denoted q, v, with $v = \dot{q}$ along the trajectory. Those on the extended phase space are q, v, t. In the case of the harmonic oscillator, $q = x$. In the case of the pendulum, $q = \theta$.

2.4 Constrained Systems

We now consider an example of the first problem in the list on p. 36, that of determining the equation of motion from the physical description of a mechanical problem with one degree of freedom. The example is important because it illustrates some of the central ideas of the Lagrangian theory. It concerns the motion of a particle in three dimensions, but in a case in which there is only one degree of freedom because the particle is constrained to move on a particular curve in space. The problem is to determine the equation of motion for the single residual degree of freedom in the simplest and most direct possible way. To be specific, the particle might be a bead sliding on a smooth wire in the shape of the curve.

A curve in space can be represented in two ways. It can be given *implicitly* in terms of two *constraint equations*

$$f_1(x, y, z) = 0, \qquad f_2(x, y, z) = 0.$$

For example, the unit circle in the x, y-plane is determined by the equations

$$x^2 + y^2 - 1 = 0, \qquad z = 0.$$

In this representation, the curve is seen as the intersection of two surfaces, the surfaces on which f_1 and f_2 individually vanish. Alternatively, it can be specified *parametrically* by giving x, y, z as functions of a parameter q. For example, the unit circle in the x, y-plane is determined by

$$x = \cos q, \quad y = \sin q, \quad z = 0.$$

Here the parameter is the angle q between the radius and the x-axis. Either way, there are technical conditions that must be satisfied to ensure that the equations really do determine a smooth curve with no kinks: in the first description, the gradient vectors of f_1 and f_2 must be linearly independent at points of the curve. In the second, the derivative $\mathrm{d}\boldsymbol{r}/\mathrm{d}q$ must be non-vanishing. This vector is tangent to the curve.

In a dynamical context, we may want to allow for the possibility that the curve itself may be changing its shape or moving in some predetermined way. For example our wire might be rotating about a fixed axis as the bead slides along it. The word 'predetermined' is important here. We must know in advance how the curve is changing with time. If we do not, then determining its motion becomes part of the dynamical problem, and we have gone beyond 'one degree of freedom'.

So we may have two *moving constraints*, given by equations of the form

$$f_1(x, y, z, t) = 0, \qquad f_2(x, y, z, t) = 0 \tag{2.14}$$

or else by a parametric representation of the constraint curve of the form

$$\boldsymbol{r} = \boldsymbol{r}(q, t).$$

In the first case, the gradients of f_1 and f_2 must be linearly independent on the curve at each fixed t . In the second, we must have $\partial \boldsymbol{r}/\partial q \neq 0$ on the curve.

Let us focus now on the example of a bead sliding smoothly on a moving wire. The forces on the bead are the normal reaction $\boldsymbol{N}$ of the wire on the bead, which is orthogonal to the wire, but otherwise unknown, and some other known force $\boldsymbol{F}$, such as gravity. By Newton's second law,

$$m\ddot{x} = N_1 + F_1, \quad m\ddot{y} = N_2 + F_2, \quad m\ddot{z} = N_3 + F_3, \tag{2.15}$$

In (2.15), we have three equations in the six unknowns x, y, z, N_1, N_2, and N_3. To these we can add the two constraint equations (2.14) and the condition that $\boldsymbol{N}$ should be orthogonal to the wire,

$$\boldsymbol{N} \cdot \frac{\partial \boldsymbol{r}}{\partial q} = N_1 \frac{\partial x}{\partial q} + N_2 \frac{\partial y}{\partial q} + N_3 \frac{\partial z}{\partial q} = 0. \tag{2.16}$$

In all, six equations in six unknowns. In principle, therefore, the motion is determined, but in practice a frontal attack by differentiation and elimination might not lead very directly to the solution.

In vector notation, (2.15) is

$$\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = \boldsymbol{N} + \boldsymbol{F},$$

where the time derivative is with respect to some fixed inertial frame of reference and

$$\boldsymbol{p} = m\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}$$

is the momentum. As is illustrated in Example 2.6 below, a single equation of motion in which $\boldsymbol{N}$ does not appear follows on taking the scalar product with $\partial\boldsymbol{r}/\partial q$, which is tangent to the wire and therefore orthogonal to $\boldsymbol{N}$. Rather than go directly to this equation, however, we do what amounts to the same manipulation, but in a slightly different way. At first reading, the steps may seem a little obscure; but it is worth following them carefully because they contain the key step to the Lagrangian theory.

Put

$$p = \boldsymbol{p}\cdot\frac{\partial\boldsymbol{r}}{\partial q}, \quad F = \boldsymbol{F}\cdot\frac{\partial\boldsymbol{r}}{\partial q}, \quad v = \dot{q}. \tag{2.17}$$

We call p, F, and v the *q-components* of the momentum $\boldsymbol{p}$, of the force $\boldsymbol{F}$, and of the velocity $\mathrm{d}\boldsymbol{r}/\mathrm{d}t$. The variable v is also called the *generalized velocity* corresponding to q.

Here, and below, we distinguish between $\mathrm{d}/\mathrm{d}t$, which is the derivative of a quantity that varies with time during the motion, and $\partial/\partial t$, which is the partial derivative with respect to t of a quantity expressed as a function of q and t or of q, v and t. So if $f = f(q, v, t)$ is a function of q, v, t, then

1. The *derivative along the motion* $\mathrm{d}f/\mathrm{d}t$ is obtained by substituting $q = q(t)$, $v = \dot{q}(t)$ into $f = f(q, v, t)$ and then differentiating with respect to t.
2. The *partial derivative* $\partial f/\partial t$ is obtained by differentiating f with respect to t, holding the values of q and v fixed.

The difference is illustrated in Figure 2.8: when calculating $\mathrm{d}f/\mathrm{d}t$, one compares the values of f at two nearby points on the curve $q = q(t)$, $v = v(t)$; for $\partial f/\partial t$, on the other hand, one compares the values of f at two different times, but for the same values of q and v, that is, for the same position and state of motion.

Example 2.5

Let $f = qvt$ and suppose that the motion is $q = t^2$. Then $v = 2t$ and

$$\begin{aligned}\frac{\partial f}{\partial t} &= qv = 2t^3\\ \frac{\mathrm{d}f}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t}(2t^4) = 8t^3.\end{aligned} \tag{2.18}$$

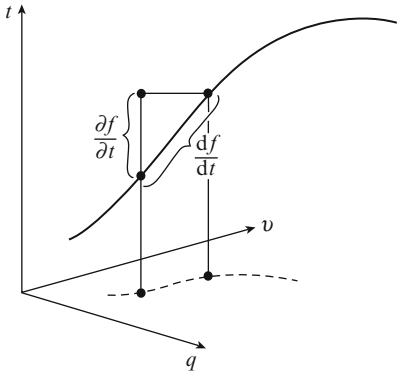

Figure 2.8

For our bead sliding on a moving wire,

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \frac{\partial \boldsymbol{r}}{\partial q}\dot{q} + \frac{\partial \boldsymbol{r}}{\partial t}$$

and therefore

$$\boldsymbol{p} = m\left(\frac{\partial \boldsymbol{r}}{\partial q}\dot{q} + \frac{\partial \boldsymbol{r}}{\partial t}\right).$$

By taking the time derivative of the first of Equations (2.17),

$$\dot{p} = \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t}\cdot\frac{\partial \boldsymbol{r}}{\partial q} + \boldsymbol{p}\cdot\left(\frac{\partial^2 \boldsymbol{r}}{\partial q^2}\dot{q} + \frac{\partial^2 \boldsymbol{r}}{\partial q \partial t}\right) = F + \boldsymbol{p}\cdot\left(\frac{\mathrm{d}^2\boldsymbol{r}}{\mathrm{d}q^2}v + \frac{\partial^2 \boldsymbol{r}}{\partial q \partial t}\right), \tag{2.19}$$

because $\boldsymbol{N}$ is orthogonal to $\partial \boldsymbol{r}/\partial q$. But for the 'correction term'

$$\boldsymbol{p}\cdot\left(\frac{\partial^2 \boldsymbol{r}}{\partial q^2}v + \frac{\partial^2 \boldsymbol{r}}{\partial q \partial t}\right), \tag{2.20}$$

(2.19) gives a simple and appealing equality between the time derivative of the q-component of momentum and the q-component of $\boldsymbol{F}$.

The key fact that underlies the Lagrangian approach is that both p and the correction term can be obtained very simply from the expression for the kinetic energy, not just in this system, but, as we see below, in a wide class of systems with an arbitrary number of degrees of freedom.

The kinetic energy in this case is

$$T = \tfrac{1}{2}m\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\cdot\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \tfrac{1}{2}m\left(\frac{\partial \boldsymbol{r}}{\partial q}v + \frac{\partial \boldsymbol{r}}{\partial t}\right)\cdot\left(\frac{\partial \boldsymbol{r}}{\partial q}v + \frac{\partial \boldsymbol{r}}{\partial t}\right). \tag{2.21}$$

It depends on the variables q and v, which together determine the position and state of motion of the particle, and the time t. By taking the partial derivatives

of T, first with respect to v with q and t held fixed and then with respect to q with v and t held fixed, we find that

$$\begin{aligned}\frac{\partial T}{\partial v} &= m\frac{\partial \boldsymbol{r}}{\partial q}\cdot\left(\frac{\partial \boldsymbol{r}}{\partial q}v+\frac{\partial \boldsymbol{r}}{\partial t}\right)=\frac{\partial \boldsymbol{r}}{\partial q}\cdot\boldsymbol{p}=p\\ \frac{\partial T}{\partial q} &= m\left(\frac{\partial \boldsymbol{r}}{\partial q}v+\frac{\partial \boldsymbol{r}}{\partial t}\right)\cdot\left(\frac{\partial^2 \boldsymbol{r}}{\partial q^2}v+\frac{\partial^2 \boldsymbol{r}}{\partial q\partial t}\right)=\boldsymbol{p}\cdot\left(\frac{\partial^2 \boldsymbol{r}}{\partial q^2}v+\frac{\partial^2 \boldsymbol{r}}{\partial q\partial t}\right).\end{aligned} \tag{2.22}$$

In other words, p is the partial derivative $\partial T/\partial v$ and the correction term is $\partial T/\partial q$. Therefore the equation of motion is

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial v}\right)-\frac{\partial T}{\partial q}=F\,. \tag{2.23}$$

Thus the work of finding the equation of motion is all but done once the kinetic energy has been expressed in terms of q and v. The gain that represents will soon be obvious.

There is considerable scope for confusion in the interpretation of the various derivatives in (2.23), and some attention will be given to their precise meaning when we come to derive the general version. For the moment, we must be content with the following recipe for decoding the left-hand side.

1. Express T in terms of q, v and t.
2. Take the partial derivatives with respect to q and v by treating q and v as if they were independent variables; in other words, forget that $v=\dot{q}$.
3. After finding the partial derivatives, substitute $\dot{q}$ for v and then take the time derivative of $\partial T/\partial v$, regarding q as a function of time.

The following examples illustrate the way in which the manipulations work in a concrete setting. In the first, the wire is fixed, so the constraint equations do not involve t.

Example 2.6

A bead of mass m is slides under gravity along a smooth wire in the shape of the elliptical helix

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1,\qquad x=a\cos\left(\frac{z}{c}\right) \tag{2.24}$$

where x, y, and z are Cartesian coordinates in an inertial frame (Figure 2.9).

If we set about solving this problem by using vector methods, then we start by writing the equation of the curve in parametric form $\boldsymbol{r}=\boldsymbol{r}(q)$, by putting

$$x=a\cos q,\quad y=b\sin q,\quad z=cq. \tag{2.25}$$

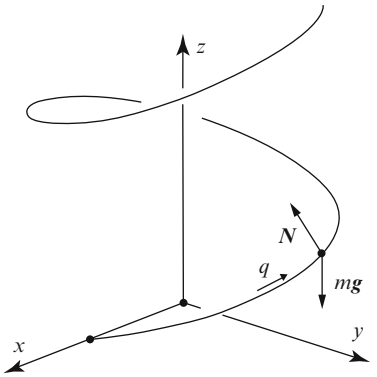

Figure 2.9

So the components of the acceleration relative to R are

$$\ddot{x} = -a\dot{q}^2 \cos q - a\ddot{q} \sin q,$$
$$\ddot{y} = -b\dot{q}^2 \sin q + b\ddot{q} \cos q, \tag{2.26}$$
$$\ddot{z} = c\ddot{q}. \tag{2.27}$$

By combining (2.16) and (2.15) to get

$$m\dot{x}\ddot{x} + m\dot{y}\ddot{y} + m\dot{z}(\ddot{z} + g) = 0 \tag{2.28}$$

and then by substituting from (2.26), we obtain

$$\ddot{q}(a^2 \sin^2 q + b^2 \cos^2 q + c^2) + \dot{q}^2(a^2 - b^2) \cos q \sin q + cg = 0, \tag{2.29}$$

after some cancellation and rearrangement. This is a single second-order differential equation for the single unknown q. The irrelevant normal reaction has been eliminated. But, by approaching the problem in this way, we have had to undertake the unnecessary task of calculating the components of the acceleration in terms of the first two time derivatives of q. It was easy enough in this particular case, but it can be a source of complication in more general problems.

If instead we follow the steps above, then we start with the expressions

$$T = \tfrac{1}{2}m(a^2 \sin^2 q + b^2 \cos^2 q + c^2)v^2.$$

for the kinetic energy and

$$F = \boldsymbol{F} \cdot \frac{\partial \boldsymbol{r}}{\partial q} = -mg\frac{\partial z}{\partial q} = -mgc$$

for the q component of the gravitational force. The second step gives

$$\begin{aligned}\frac{\partial T}{\partial q} &= m(a^2-b^2)v^2\sin q\cos q\,,\\ \frac{\partial T}{\partial v} &= \frac{\partial T}{\partial v} = m(a^2\sin^2 q + b^2\cos^2 q + c^2)v.\end{aligned} \tag{2.30}$$

Hence the equation of motion is

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(m(a^2\sin^2 q + b^2\cos^2 q + c^2)\dot{q}\Big) - m(a^2-b^2)\dot{q}^2\sin q\cos q = -mcg\,, \tag{2.31}$$

which immediately reduces to (2.29).

Example 2.7

In Example 1.21, the velocity of the bead relative to the inertial frame is

$$\dot{\boldsymbol{r}} + \boldsymbol{\omega}\wedge\boldsymbol{r} = a\dot{\theta}(-\sin\theta\boldsymbol{i} + \cos\theta\boldsymbol{j}) + a\omega(\sin\alpha\boldsymbol{i} + \cos\alpha\boldsymbol{k})\wedge(\cos\theta\boldsymbol{i} + \sin\theta\boldsymbol{j}).$$

Therefore

$$\begin{aligned}T &= \tfrac{1}{2}m(\dot{\boldsymbol{r}} + \boldsymbol{\omega}\wedge\boldsymbol{r})\cdot(\dot{\boldsymbol{r}} + \boldsymbol{\omega}\wedge\boldsymbol{r})\\ &= \tfrac{1}{2}ma^2\big(v^2 + 2\omega v\cos\alpha + \omega^2(\cos^2\alpha + \sin^2\alpha\sin^2 q)\big),\end{aligned}$$

with $q=\theta$ and $v=\dot{\theta}$. Also we have

$$F = m\boldsymbol{g}\cdot\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\theta} = mga\sin\alpha\sin q.$$

Therefore the equation of motion is

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial v}\right) - \frac{\partial T}{\partial q} &= \frac{\mathrm{d}}{\mathrm{d}t}\big(ma^2(v+\omega\cos\alpha)\big) - ma^2\omega^2\sin^2\alpha\sin q\cos q\\ &= ma^2\ddot{q} - ma^2\omega^2\sin^2\alpha\sin q\cos q\\ &= mga\sin\alpha\sin q,\end{aligned}$$

in agreement with (1.30).

EXERCISES

2.2. The motion of a particle of mass m in space is subject to the two constraint equations

$$x^2 + y^2 + z^2 = 1 \quad\text{and}\quad x\sin(\omega t) - y\cos(\omega t) = 0,$$

where ω is constant. Show that the particle is moving on a circle that is rotating with constant angular speed ω about a vertical diameter. Show that its position can be specified parametrically by

$$x = \sin q \cos(\omega t), \quad y = \sin q \sin(\omega t), \quad z = \cos q,$$

and express the kinetic energy as a function of q, $\dot{q}$, and t.

2.5 Lagrange's Equation

We can further streamline the process when the known force $\boldsymbol{F}$ can be derived from a *potential function* $U(x, y, z, t)$. That is, $\boldsymbol{F} = -\operatorname{grad} U$, or in components,

$$F_1 = -\frac{\partial U}{\partial x}, \quad F_2 = -\frac{\partial U}{\partial y}, \quad F_3 = -\frac{\partial U}{\partial z}.$$

In this case, we put $U(q, t) = U(\boldsymbol{r}(q), t)$ and we find that

$$F = -\operatorname{grad} U \cdot \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}q} = -\frac{\partial U}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}q} - \frac{\partial U}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}q} - \frac{\partial U}{\partial z}\frac{\mathrm{d}z}{\mathrm{d}q} = -\frac{\partial U}{\partial q}$$

by the chain rule. Thus the equation of motion can be written

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial v}\right) - \frac{\partial T}{\partial q} + \frac{\partial U}{\partial q} = 0,$$

or in a simpler form,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v}\right) - \frac{\partial L}{\partial q} = 0, \tag{2.32}$$

where $L(q, v, t) = T(q, v, t) - U(q, t)$. The function L is called the *Lagrangian*. It determines the motion through (2.32), which is *Lagrange's equation*.

Example 2.8

In Example 2.6, the kinetic energy is

$$T = \tfrac{1}{2}m(a^2 \sin^2 q + b^2 \cos^2 q + c^2)v^2.$$

and the potential is $U = mgz = mgcq$. So

$$L = \tfrac{1}{2}m(a^2 \sin^2 q + b^2 \cos^2 q + c^2)v^2 - mgcq.$$

Example 2.9

In Example 2.7, $U = mga \sin\alpha \cos\theta$ and so the Lagrangian is

$$L = \tfrac{1}{2}mv^2 + m\omega v \cos\alpha + \tfrac{1}{2}m\omega^2(\cos^2\alpha + \sin^2\alpha \sin^2 q) - mga \sin\alpha \cos\theta. \tag{2.33}$$

2.6 Conservation of Energy

The partial derivative $\partial L/\partial t$ is the derivative of L with respect to t with q and v treated as constant, while $\mathrm{d}L/\mathrm{d}t$ is the derivative taken *after* making the substitution

$$q = q(t), \qquad v = \dot{q}(t).$$

By applying the chain rule,

$$\frac{\mathrm{d}L}{\mathrm{d}t} = \frac{\partial L}{\partial q}\dot{q} + \frac{\partial L}{\partial v}\ddot{q} + \frac{\partial L}{\partial t}.$$

So if $q = q(t)$ is a solution of Lagrange's equation, then

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{q}\frac{\partial L}{\partial v}\right) = \ddot{q}\frac{\partial L}{\partial v} + \dot{q}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v}\right) = \ddot{q}\frac{\partial L}{\partial v} + \dot{q}\frac{\partial L}{\partial q} = \frac{\mathrm{d}L}{\mathrm{d}t} - \frac{\partial L}{\partial t}. \tag{2.34}$$

It follows that if

$$\frac{\partial L}{\partial t} = 0,$$

which is the case when L is given by an expression in q and v with no explicit dependence on t, then

$$h = v\frac{\partial L}{\partial v} - L$$

is constant during the motion. This quantity h is called the *Hamiltonian*.

When $L = T - U$, where the potential U is a function of q alone and the kinetic energy T is an expression of the form

$$T = \tfrac{1}{2}f(q)v^2,$$

the Hamiltonian is

$$h = v\frac{\partial T}{\partial v} - T + U = \tfrac{1}{2}f(q)v^2 + U = T + U.$$

So it is the *total energy*, the sum of the kinetic and potential energies, that is conserved in this case.

A system with the property that $\partial L/\partial t = 0$ behaves in the same way at whatever time it is set in motion because if L does not depend on t then Lagrange's equation is unchanged when t is replaced by $t+k$ for any constant k. Thus in Lagrangian mechanics, conservation of energy is intimately connected with symmetry under *time translation*.

A word of warning: the Hamiltonian does not always coincide with the total energy. There are examples in which the Hamiltonian is constant during the motion, but the total energy is not.

Example 2.10

In Example 2.8,

$$L = \tfrac{1}{2}m(a^2 \sin^2 q + b^2 \cos^2 q + c^2)v^2 - mgcq.$$

Here $\partial L/\partial t = 0$ and $f = m(a^2 \sin^2 q + b^2 \cos^2 q + c^2)$. So the Hamiltonian is

$$h = \tfrac{1}{2}m(a^2 \sin^2 q + b^2 \cos^2 q + c^2)v^2 + mgcq.$$

It is constant during the motion, and is equal to the total energy.

Example 2.11

In Example 2.9, the Lagrangian is given by (2.33) and so the Hamiltonian is

$$h = \tfrac{1}{2}mv^2 - \tfrac{1}{2}m\omega^2(\cos^2 \alpha + \sin^2 \alpha \sin^2 q) + mga \sin \alpha \cos \theta.$$

It is conserved because $\partial L/\partial t = 0$, but is not equal to the total energy $T + U$. The total energy is not conserved in this case because the force maintaining the rotation of the wire does work during the motion of the bead.

2.7 The Calculus of Variations

Lagrange's equation also arises in another context, the *calculus of variations*. This is a method for determining which function within a certain class maximizes or minimizes an integral expression involving the function.

Example 2.12 (Minimal Surfaces of Revolution)

Amongst all continuously differentiable functions $y = y(x)$ with given values at $x = a$ and $x = b$, find the one that minimizes the integral

$$A = 2\pi \int_a^b y\sqrt{1 + y'^2}\, \mathrm{d}x.$$

This quantity is the area between a and b of the surface obtained by rotating the graph of the function about the x axis. If we are given that $y(a) = c$ and $y(b) = d$, then we can think of the problem in physical terms as that of finding the shape of a soap film between two circular hoops of wire of radius c and d, respectively. The circles are in planes orthogonal to the x-axis, with their respective centres at $x = a$ and $x = b$.

Example 2.13 (Brachistochrone Problem)

Consider a wire in the shape of a curve $y = y(x)$, where the y-axis is vertical and the x-axis is horizontal. The ends of the wire are at the points (a, c) and (b, d), with $c > d > 0$.

A bead slides smoothly down the wire from (a, c), where it starts at rest, to (b, d). Its velocity is

$$(\dot{x}, \dot{y}) = (1, y')\dot{x},$$

where $y' = \mathrm{d}y/\mathrm{d}x$. Because energy is conserved, and because $\dot{x} = 0$ at the initial point $x = a$, we have

$$\tfrac{1}{2}\dot{x}^2(1 + y'^2) + g(y - c) = 0.$$

Therefore

$$\dot{x} = \sqrt{\frac{2g(c - y)}{1 + y'^2}}$$

and so the time taken to reach the lower end is

$$T = \int_a^b \frac{\mathrm{d}x}{\dot{x}} = \int_a^b \sqrt{\frac{1 + y'^2}{2g(c - y)}}\,\mathrm{d}x.$$

The *brachistochrone problem* is to find, with a, b, c, d fixed, the shape of the wire that minimizes the time for the descent. That is, to find the function $y = y(x)$ with $y(a) = c$ and $y(b) = d$ that minimizes T.

Example 2.14

A problem that links mechanics and the calculus of variations in a way that we revisit many times below is that of finding how a particle should move along the x-axis between the point $x = c$ at time a and $x = d$ at time b if it is to minimize its average (over time) kinetic energy. This is the problem of finding a function $x(t)$ with $x(a) = c$ and $x(b) = d$ which minimizes the integral

$$\int_a^b \tfrac{1}{2}m\dot{x}^2\,\mathrm{d}t.$$

We see below that the answer is that the motion should be at constant speed. To put it the other way round, the force-free motion between c and d is singled out amongst all possible motions from c at time a to d at time b by the condition that the average kinetic energy should be minimal.

In the calculus of variations, one tackles such problems in the same way that one tackles the problem of finding the maxima and minima of a function $f(x)$ of one variable. If f achieves one of its extreme values at x, then x must be a *critical point*. That is, a point at which the derivative $f'(x)$ is zero. Setting aside analytical niceties, we can characterize the critical points by the condition that the change δf in f should vanish when x is changed to $x + \delta x$ for some small value of δx. This is equivalent to the vanishing of the derivative because

$$f(t + \delta x) - f(x) = \delta x f'(x)$$

to the first order in δx. Not all critical points are maxima or minima, and in any case f need not achieve its extreme values at finite values of x. For example the function

$$f(x) = \tan^{-1}(x)$$

has no critical points and approaches the extreme values of $\pm\pi/2$ only in the limit $x \to \pm\infty$.

In problems like those in the examples above, we are given a function $L = L(q, v, t)$ of three variables, and we seek to find the extrema of

$$J = \int_a^b L\big(q(t), \dot{q}(t), t\big)\, \mathrm{d}t \tag{2.35}$$

over all smooth functions $q = q(t)$ such that $q(a) = c$, $q(b) = d$, with a, b, c, d given and fixed. In order to avoid analytical complexity, we take 'smooth' to mean 'having a continuous second derivative'. We also assume that L has continuous second partial derivatives. It is worth noting, however, that the problem still makes sense if we allow $q(t)$ to vary over a wider class of functions with only continuous first derivatives. With a little more work, one can show that the functions $q(t)$ that solve this problem necessarily have continuous second derivatives. This is the *du Bois-Reymond theorem* [3].

As with the problem of finding the extrema of a function of a single variable, one looks first for 'critical functions' $q(t)$. That is, for functions $q = q(t)$ such that $\delta J = 0$ when $q(t)$ is replaced by any *variation* $q(t) + \delta q(t)$, where $\delta q(t)$ is a small change in $q(t)$, subject to the constraints

$$\delta q(a) = 0, \qquad \delta q(b) = 0.$$

Second-order terms in δq and its derivatives are ignored.

Proposition 2.15 (Euler's Theorem)

The critical functions are the solutions of Lagrange's equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v}\right) - \frac{\partial L}{\partial q} = 0, \qquad v = \dot{q}. \tag{2.36}$$

Proof

Make a variation in $q(t)$ by putting

$$\delta q(t) = su(t).$$

where $u(a) = u(b) = 0$ and s is small. Then

$$\begin{aligned} J + \delta J &= \int_a^b L(q + su, \dot{q} + s\dot{u}, t)\,\mathrm{d}t + O(s^2) \\ &= J + s\int_a^b \left(\frac{\partial L}{\partial q}u + \frac{\partial L}{\partial v}\dot{u}\right)\,\mathrm{d}t + O(s^2). \end{aligned} \tag{2.37}$$

Hence, by integrating by parts and by dropping second-order terms in s,

$$\begin{aligned} \delta J &= s\int_a^b \left(\frac{\partial L}{\partial q}u + \frac{\partial L}{\partial v}\dot{u}\right)\,\mathrm{d}t \\ &= s\int_a^b \left[\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v}\right)\right] u\,\mathrm{d}t + \left[\frac{\partial L}{\partial v}u\right]_a^b \\ &= s\int_a^b G(t)u(t)\,\mathrm{d}t, \end{aligned} \tag{2.38}$$

because $u(a) = 0 = u(b)$, where

$$G(t) = \frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v}\right).$$

The definition of G is interpreted in the same way as before: first calculate the partial derivatives of L with respect to q and v, then substitute $q = q(t)$ and $v = \dot{q}(t)$, and take the time derivative $\mathrm{d}/\mathrm{d}t$.

Suppose now that $q(t)$ is a critical function. Then

$$\int_a^b G(t)u(t)\,\mathrm{d}t = 0 \tag{2.39}$$

for every such variation. We deduce that $G(t) = 0$ by assuming that it is not true and arriving at a contradiction. If $G(\tau) > 0$ for some $\tau \in (a, b)$, then, by continuity, G is also positive in a neighbourhood of τ in (a, b). So there exists $\sigma > 0$ such that

$$(\tau - \sigma, \tau + \sigma) \subset (a, b)$$

and such that $G(t)$ is positive for all $t \in (\tau - \sigma, \tau + \sigma)$. Put

$$u(t) \;=\; \begin{cases} (\tau + \sigma - t)^3(t - \tau + \sigma)^3 & t \in (\tau - \sigma, \tau + \sigma) \\ 0 & t \notin (\tau - \sigma, \tau + \sigma) \end{cases} \tag{2.40}$$

Then

$$\int_a^b Gu\,\mathrm{d}t = \int_{\tau-\sigma}^{\tau+\sigma} Gu\,\mathrm{d}t > 0. \tag{2.41}$$

But u vanishes at a and b and is twice continuously differentiable, with the third power ensuring the continuity of the second derivative at $\tau \pm \sigma$. Hence (2.41) contradicts (2.39). A similar argument disposes of $G(\tau) < 0$. Therefore $G(t) = 0$ for all $t \in (a, b)$, and by continuity, also at the endpoints of the interval. □

Example 2.16

In Example 2.12, y plays the role of q and x the role of t. We have

$$L(y, v, x) = y\sqrt{1+v^2},$$

where $v = y'$. Lagrange's equations are

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{yv}{\sqrt{1+v^2}}\right) - \sqrt{1+v^2} = 0, \qquad v = \frac{\mathrm{d}y}{\mathrm{d}x}.$$

On substituting $v = y' = \mathrm{d}y/\mathrm{d}x$ and simplifying, we find

$$yy'' = 1 + y'^2.$$

It can be verified that the solution is $y = A\cosh(A^{-1}x + B)$ for constant A and B. The graph of such a function is called a *catenary*, because it is the curve formed by a heavy chain hanging between two fixed points. Our surface is therefore a *catenoid*, which is the surface formed by rotating a catenary about the x-axis.

Example 2.17

In the brachistochrone problem (Example 2.13), we have

$$L = \sqrt{\frac{1+v^2}{2g(c-y)}},$$

with $v = y'$, again with x playing the role of t and y playing the role of q. Because L is independent of x, the quantity

$$y'\frac{\partial L}{\partial v} - L = \frac{y'^2 F}{1+y'^2} - F = -\frac{1}{\sqrt{1+y'^2}\sqrt{2g(c-y)}}$$

is constant. It is the 'Hamiltonian' for this problem (see Section 2.6).

We deduce that

$$(1+y'^2)(c-y) = 2k$$

where k is constant. This is solved in terms of a parameter θ by

$$x = k(\theta - \sin\theta), \qquad c - y = k(1 - \cos\theta),$$

which are the parametric equations of a cycloid.

EXERCISES

2.3. Find the critical functions in the case $L = \frac{1}{2}v^2 - gq$, with g constant. Solve Lagrange's equation with boundary condition $q = 0$ at $t = 0$ and at $t = 1$. Evaluate J for the family of functions $q = st(t-1)$ labelled by s and show that the solution of Lagrange's equation minimizes J as a function of s.

2.8 One-Parameter Variations

The 'δ' notation used in the definition of critical curves and in the proof of Euler's theorem is convenient and familiar, but perhaps too informal to permit complete confidence in the proof. One can instead phrase the definition and the statement of the proposition in terms of one-parameter variations.

Denote by Γ the set of twice continuously differentiable functions

$$q : [a, b] \to \mathbb{R} \quad \text{with} \quad q(a) = c, \quad q(b) = d.$$

A one-parameter variation of $q \in \Gamma$ is a curve in Γ through q. The formal definition is as follows.

Definition 2.18

Let $q \in \Gamma$. A *one-parameter variation* of q is a twice continuously differentiable map

$$w : [a, b] \times [-\varepsilon, \varepsilon] \to \mathbb{R}$$

such that

(1) $w(a, s) = c$ and $w(b, s) = d$ for all $s \in [-\varepsilon, \varepsilon]$ and

(2) $w(t, 0) = q(t)$ for all $t \in [a, b]$.

In other words, if we define $q_s(t) = w(t, s)$, then q_s is a family of functions in Γ depending smoothly on s, with $q_0 = q$. The integral (2.35) can be understood to define a map

$$J : \Gamma \to \mathbb{R}.$$

By evaluating J at each q_s, we obtain a function of s, given by

$$J(s) = \int_a^b L\big(w(t,s), w_t(t,s), t\big)\,\mathrm{d}t,$$

where w_t denotes the partial derivative $\partial w/\partial t$.

With this definition, one avoids the delta notation by defining a critical curve q to be one for which $J'(0) = 0$ for every one-parameter variation of q. Euler's theorem remains true with this definition, and the proof is the same except that the first step is to calculate $J'(0)$ by differentiation under the integral sign to obtain

$$J'(0) = \int_a^b \left(\frac{\partial L}{\partial q}u + \frac{\partial L}{\partial v}\dot{u}\right)\,\mathrm{d}t$$

where now $u = w_s(t,0)$. If u is defined by (2.40), then $w(t,s) = q(t) + su(t)$ is a one-parameter variation of q.

EXERCISES

2.4. Suppose that $L = v^2$, $a = c = 0$, and $b = d = 1$, with $n = 1$. Show that the critical function is $q = t$. Show that $w(s,t) = t + st(t-1)$ is a one-parameter variation. Show that $J(s) = \frac{1}{3}s^2 + 1$ and so check that $J(s)$ is minimal when $s = 0$.

2.9 Hamilton's Principle

Consider a system with one degree of freedom and with Lagrangian equation of motion

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v}\right) - \frac{\partial L}{\partial q} = 0,$$

where $L = T - U$. Suppose that we are given that $q(a) = c$ and $q(b) = d$, so that we know the initial and final position. How do we determine the position $q(t)$ when t is between a and b? The simple answer is 'by solving the equation of motion and imposing the given boundary conditions'.

Hamilton's principle is an alternative way to characterize the solution $q(t)$ amongst all smooth functions with $q(a) = c$ and $q(b) = d$. It follows from the observation that the equation of motion determines the critical functions of

$$J_L(q) = \int_a^b L(q, \dot{q}, t)\mathrm{d}t.$$

Thus we have the following corollary of Euler's theorem.

Proposition 2.19 (Hamilton's Principle)

The dynamical trajectories of a system with Lagrangian L are the critical functions of J_L.

The integral J_L is called the *action* and Hamilton's principle is sometimes called the *principle of least action*. However this is misleading. The trajectories are critical functions of the action, but they are not necessarily minima.

Example 2.20 (The Harmonic Oscillator)

The force $F = -mq$ is independent of time, so the system is conservative and autonomous, with potential $U = \frac{1}{2}mq^2$. The equation of motion

$$m\ddot{q} = -mq$$

is generated by the Lagrangian $L(q, v, t) = \frac{1}{2}m(v^2 - q^2)$, with $v = \dot{q}$.

If we impose the boundary conditions $q(0) = 1$, $q(2\pi) = 1$, then the solution is $q = \cos t$. When we evaluate the action with this choice of $q(t)$, we obtain

$$J_L(q) = \int_0^{2\pi} \tfrac{1}{2}m(\sin^2 t - \cos^2 t)\,\mathrm{d}t = 0.$$

However the boundary conditions, but not of course the equation of motion, are also satisfied by the two functions $q_1(t) = 1$ and $q_2(t) = \cos 2t$. For the first of these, the action is

$$J_L(q_1) = -\int_0^{2\pi} \tfrac{1}{2}m\,\mathrm{d}t = -\pi m.$$

For the second, $L = 2m\sin^2 2t - \frac{1}{2}m\cos^2 2t$ and

$$J_L(q_2) = \int_0^{2\pi} \tfrac{1}{2}m(4\sin^2 2t - \tfrac{1}{2}\cos^2 2t)\,\mathrm{d}t = 7\pi m/4.$$

So $J_L(q_1) < J_L(q) < J_L(q_2)$. Although the solution is a critical function for the action, it is neither a maximum nor a minimum.

One must be wary in all this of the *first fundamental confusion of calculus*. We have used q is two senses: when we write $L(q, v, t)$, q is one of the three independent real variables on which L depends. When we write $q \in \Gamma$ or $J_L(q)$, we are thinking of q as a function of t, with given values at a and b. One does something similar in elementary calculus when one writes $y = y(x)$, with the y on the right denoting a function and that on the left denoting its value at a particular value of x. The first fundamental confusion can result, particularly

when functions with standard names, such as cos and sin are involved. It is possible, and from some points of view it might be better, to draw a notational distinction between the two meanings of 'q', but the result can be clumsy and inelegant. As with the conventional misuse of notation in writing $y = y(x)$, it is better simply to be aware of the shifting meaning and to be resolute in agreeing not to be confused by it. So we name the confusion in order to tame it.

2.10 Change of Coordinates

One technique for simplifying an equation of motion of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial v}\right) - \frac{\partial T}{\partial q} = F\,. \tag{2.42}$$

is to replace the dependent variable q by some function of q and t. For example, in Example 2.6, one could change the parameter along the curve. It is usually a simple matter to make the corresponding substitution to write the equation of motion in terms of the new variable, although the need to express the second derivatives of the new variable in terms of the derivatives of the old can be a source of complication. In systems with more degrees of freedom, such a head-on approach can be very inefficient.

A more direct method is to write the kinetic energy as a function of q, $v = \dot{q}$, and t, and then to express the kinetic energy in terms of the new position variable and its derivative. If the new position variable is $\tilde{q} = \tilde{q}(q,t)$, then along any trajectory

$$\dot{\tilde{q}} = \frac{\partial \tilde{q}}{\partial q}\dot{q} + \frac{\partial \tilde{q}}{\partial t}.$$

So the new velocity variable $\tilde{v}$ is expressed in terms of q, v and t by

$$\tilde{v} = \frac{\partial \tilde{q}}{\partial q}v + \frac{\partial \tilde{q}}{\partial t}.$$

In other words, we are looking at a change of coordinates on the extended phase space of the form

$$\tilde{q} = \tilde{q}(q,t), \qquad \tilde{v} = \frac{\partial \tilde{q}}{\partial q}v + \frac{\partial \tilde{q}}{\partial t}, \qquad \tilde{t} = t, \tag{2.43}$$

with the last equation expressing the fact that the time coordinates in the original system q, v, t and the transformed system $\tilde{q}, \tilde{v}, \tilde{t}$ are the same. The kinetic energy can be written either as a function of q, v, t or as a function of the new variables $\tilde{q}, \tilde{v}, \tilde{t}$.

The following proposition enables us to obtain the equation of motion in terms of the new variables without the need to make a direct determination of the second derivatives.

Proposition 2.21

Let T be a twice continuously differentiable function of the three variable q, v, t and suppose that q is given a function of t. Then under the change of variables (2.43),

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial v}\right) - \frac{\partial T}{\partial q} = \frac{\partial \tilde{q}}{\partial q}\left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \tilde{v}}\right) - \frac{\partial T}{\partial \tilde{q}}\right]. \tag{2.44}$$

Because this proposition and its generalization for systems with many degrees of freedom is so central to the Lagrangian theory, we look at two proofs. The first exploits the fact that

$$\int_a^b T(q, \dot{q}, t)\,\mathrm{d}t = \int_a^b T(\tilde{q}, \dot{\tilde{q}}, t)\,\mathrm{d}t. \tag{2.45}$$

The second is a direct application of the chain rule.

Before turning to the proofs, we note that despite the fact that $t = \tilde{t}$, the partial derivative $\partial/\partial t$ is not necessarily the same as $\partial/\partial\tilde{t}$, although the derivative $\mathrm{d}/\mathrm{d}t$ along the motion is the same in both coordinate systems. An illustration is provided by the behaviour of the derivatives of a function $g(q, t)$ under the coordinate transformation $\tilde{q} = q+t$ and $\tilde{t} = t$. By using the chain rule,

$$\frac{\partial g}{\partial t} = \frac{\partial g}{\partial \tilde{q}}\frac{\partial \tilde{q}}{\partial t} + \frac{\partial g}{\partial \tilde{t}}\frac{\partial \tilde{t}}{\partial t} = \frac{\partial g}{\partial \tilde{q}} + \frac{\partial g}{\partial \tilde{t}} \tag{2.46}$$

so that $\partial g/\partial t \neq \partial g/\partial\tilde{t}$, even though $t = \tilde{t}$. The difficulty is that there is a potential ambiguity in the notation $\partial g/\partial t$. It does not make sense unless one knows which other variables are to be held fixed. In (2.46), we held q fixed when we took the partial derivative with respect to t, and $\tilde{q}$ fixed when we differentiated with respect to $\tilde{t}$.

Here and below we avoid this notational pitfall by adopting the following device: when taking one of the partial derivatives $\partial/\partial\tilde{q}$, $\partial/\partial\tilde{v}$, or $\partial/\partial\tilde{t}$, we hold fixed all the other variables with tildes, but when taking $\partial/\partial q$, $\partial/\partial v$, or $\partial/\partial t$, we hold fixed the other variables without tildes. This allows us to distinguish between $\partial/\partial t$ and $\partial/\partial\tilde{t}$ in spite of the fact that $t = \tilde{t}$.

Proof (First Proof)

Fix a and b and consider the integral

$$J_T = \int_a^b T(q, \dot{q}, t)\,\mathrm{d}t.$$

Now make a variation, replacing $q(t)$ by $q(t) + \delta q(t)$, with $\delta q(a) = \delta q(b) = 0$. Then to the first order

$$\delta J_T = \int_a^b \left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial v}\right) - \frac{\partial T}{\partial q}\right] \delta q(t)\,\mathrm{d}t,$$

as in the proof of Proposition 2.15. However, from (2.45), we also have

$$\delta J_T = \int_a^b \left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \tilde{v}}\right) - \frac{\partial T}{\partial \tilde{q}}\right] \delta \tilde{q}(t)\,\mathrm{d}t,$$

where

$$\delta \tilde{q} = \frac{\partial \tilde{q}}{\partial q}\,\delta q.$$

Therefore any $\delta q(t)$ vanishing at a and b, we have

$$\int_a^b \left\{\left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial v}\right) - \frac{\partial T}{\partial q}\right] - \left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \tilde{v}}\right) - \frac{\partial T}{\partial \tilde{q}}\right]\frac{\partial \tilde{q}}{\partial q}\right\} \delta q\,\mathrm{d}t = 0.$$

So the proposition follows by the same argument as in the proof of Proposition 2.15. □

Proof (Second Proof)

By the chain rule,

$$\begin{aligned} \tfrac{\partial T}{\partial q} &= \frac{\partial T}{\partial \tilde{q}}\frac{\partial \tilde{q}}{\partial q} + \frac{\partial T}{\partial \tilde{v}}\frac{\partial \tilde{v}}{\partial q} + \frac{\partial T}{\partial \tilde{t}}\frac{\partial \tilde{t}}{\partial q} \\ &= \frac{\partial T}{\partial \tilde{q}}\frac{\partial \tilde{q}}{\partial q} + \frac{\partial T}{\partial \tilde{v}}\left[\frac{\partial^2 \tilde{q}}{\partial q^2}v + \frac{\partial^2 \tilde{q}}{\partial q \partial t}\right] \end{aligned} \tag{2.47}$$

because $\partial \tilde{t}/\partial q = 0$ and $\partial \tilde{v}/\partial q$ is equal to the expression in square brackets, by differentiating the second equation in (2.43). Similarly

$$\begin{aligned} \tfrac{\partial T}{\partial v} &= \frac{\partial T}{\partial \tilde{q}}\frac{\partial \tilde{q}}{\partial v} + \frac{\partial T}{\partial \tilde{v}}\frac{\partial \tilde{v}}{\partial v} + \frac{\partial T}{\partial \tilde{t}}\frac{\partial \tilde{t}}{\partial v} \\ &= \frac{\partial T}{\partial \tilde{v}}\frac{\partial \tilde{q}}{\partial q} \end{aligned}$$

because $\partial\tilde{q}/\partial v = \partial\tilde{t}/\partial v = 0$ and $\partial\tilde{v}/\partial v = \partial\tilde{q}/\partial q$, from (2.43). Now for any function $k = k(q,t)$ which depends only on q and t,

$$\frac{\mathrm{d}k}{\mathrm{d}t} = \frac{\partial k}{\partial q}\frac{\mathrm{d}q}{\mathrm{d}t} + \frac{\partial k}{\partial t} = \frac{\partial k}{\partial q}v + \frac{\partial k}{\partial t}\,.$$

Therefore

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\tilde{q}}{\partial q}\right) = \frac{\partial^2\tilde{q}}{\partial q^2}v + \frac{\partial^2\tilde{q}}{\partial q\partial t}$$

and hence

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial v}\right) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial\tilde{v}}\right)\frac{\partial\tilde{q}}{\partial q} + \frac{\partial T}{\partial\tilde{v}}\left[\frac{\partial^2\tilde{q}}{\partial q^2}v + \frac{\partial^2\tilde{q}}{\partial q\partial t}\right]. \tag{2.48}$$

The proposition now follows on subtracting (2.47) from (2.48). □

The proposition establishes that the combination of derivatives on the left-hand side of (2.42) transforms in a simple way. The transformation rule is that of a *covector*, a term that is explained in Chapter 8. If we transform the right-hand side in the same way, then we arrive very simply at the transformed equation of motion. In fact, it follows that if the equation of motion is (2.42) in the original coordinates, then in the new coordinates it is

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial\tilde{v}}\right) - \frac{\partial T}{\partial\tilde{q}} = \tilde{F} \quad \text{where} \quad \tilde{F} = \frac{\partial q}{\partial\tilde{q}}\,F.$$

In particular, if the equation of motion is in Lagrange's form (2.32) in the original coordinates, then it takes the same form in the new coordinates, with the Lagrangian written in terms of the new coordinates. This follows from the fact that if $F = -\partial U/\partial q$, then

$$\tilde{F} = -\frac{\partial q}{\partial\tilde{q}}\frac{\partial U}{\partial q} = -\frac{\partial U}{\partial\tilde{q}}.$$

by the chain rule.

EXERCISES

2.5. The Lagrangian $L = \frac{1}{2}\dot{q}^2 - \frac{1}{2}q^2$ generates the equation of simple harmonic motion

$$\ddot{q} + q = 0.$$

Show directly that under change of variable to $\tilde{q} = q^2$, the equation of motion is transformed to

$$2\tilde{q}\ddot{\tilde{q}} - \dot{\tilde{q}}^2 + 4\tilde{q}^2 = 0.$$

Derive the same result by showing that the transformed Lagrangian is

$$L = \frac{\dot{\tilde{q}}^2 - 4\tilde{q}^2}{8\tilde{q}}$$

and by writing down the corresponding Lagrange equation.

3
Lagrangian Mechanics

3.1 Several Degrees of Freedom

In this chapter, we extend the techniques developed in Chapter 2 to systems with several degrees of freedom. These are systems in which more than one coordinate is needed to specify the configuration at any time. Our aim is to determine and then to analyse the differential equations that determine the evolution of the configuration by applying Newton's laws of motion to the component parts of the system. The Lagrangian method greatly simplifies the task. Indeed, in many of the examples that we consider, any more direct approach would be intractable.

The set of all possible configurations of a mechanical system is called its *configuration space* and is denoted by C. The coordinates label the points of C. Some simple examples illustrate this idea.

– For a particle moving in a plane, C is the plane itself. Here we need two coordinates, for example the x and y coordinates of the particle, or its polar coordinates r, θ.

– For a particle moving in space, C is Euclidean space $\mathbb{E}$, and we can use Cartesian coordinates x, y and z coordinates relative to some choice of axes.

– For two particles moving in space, $C = \mathbb{E} \times \mathbb{E}$, and we need six coordinates, three for the position of the first particle and three for the position of the second.

– For a spherical pendulum made up of a particle attached to a rigid rod of

N.M.J. Woodhouse, *Introduction to Analytical Dynamics: New Edition*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-1-84882-816-2_3,

length a, with the other end freely pivoted at a fixed point, C is the sphere of radius a centred on the fixed point. If we take the fixed point as origin, then the Cartesian coordinates of the particle satisfy the constraint

$$x^2 + y^2 + z^2 = a^2.$$

We can label the points of C by the two polar coordinates θ and φ, by writing

$$x = a\sin\theta\cos\varphi, \qquad y = a\sin\theta\sin\varphi, \qquad z = \cos\theta.$$

– For two particles moving in space connected by a rigid rod of length a, we need five coordinates, three for the position of one particle, and then two more, as in the spherical pendulum, to determine the position of the second when that of the first is given.

– For a rigid body, we need six coordinates: three for the position of the centre of mass, two to determine the orientation in space of some axis fixed in the body, and one more for the rotations about this axis.

In systems involving many particles and constraints, the number of coordinates can become large and the topology of the configuration space can be very complicated.

3.2 Configuration Space and Phase Space

Our model for a general mechanical system is a collection of particles subject to constraints, as in the example of two particles connected by a light rigid rod. The 'constraint' in that case is that the distance between the particles is constant. Because the rod is 'light', we ignore its effect on the dynamics, beyond assuming that it exerts the forces needed to maintain the constraint.

A more general example is a rigid body, which we think of a collection of a large number of particles, subject to the constraints that the distances between them should be constant. This is not implausible if one thinks of the particles as atoms, with the constraints maintained by inter-atomic forces. Despite its artificiality and clear physical limitations, it gives good results in situations in which 'rigidity' is maintained.

Suppose that our general system is made up of N particles with masses

$$m_1, m_2, \ldots, m_N$$

and position vectors

$$\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N$$

relative to some inertial frame. If for each $\alpha = 1, 2, \ldots, N$ the particle with position vector $\boldsymbol{r}_\alpha$ is subject to the force $\boldsymbol{F}_\alpha$, then the equations of motion are

$$m_\alpha \ddot{\boldsymbol{r}}_\alpha = \boldsymbol{F}_\alpha.$$

As the system evolves, each particle traces out a curve in space, so the motion of the system can be represented by a set of N curves. The first step in the Lagrangian theory is to make a shift in viewpoint and think of these N curves as a single curve in a $3N$-dimensional space, called the *configuration space* of the particles. We denote this space by C. Its points are labelled by the $3N$ coordinates

$$r_{11}, r_{12}, r_{13}, r_{21}, r_{22}, \ldots, r_{N3}$$

Each point of C corresponds to a particular configuration of the system. As the particles move under the influence of the forces, their successive configurations trace out the curve in C, which is the *dynamical trajectory* of the system.

The configuration space is the Cartesian product of N copies of three-dimensional Euclidean space, but to emphasize that it is to be thought of as a single space rather than as a product, we relabel the coordinates, by writing

$$q_1 = r_{11},\ q_2 = r_{12},\ q_3 = r_{13},\ q_4 = r_{21},\ q_5 = r_{22},\ \ldots,\ q_n = r_{N3}$$

where $n = 3N$; and we relabel the masses, introducing constants

$$\mu_1 = m_1,\ \mu_2 = m_1,\ \mu_3 = m_1,\ \mu_4 = m_2,\ \mu_5 = m_2,\ \ldots,\ \mu_n = m_N \tag{3.1}$$

Then the equations of motion become

$$\mu_1 \ddot{q}_1 = Q_1, \quad \mu_2 \ddot{q}_2 = Q_2, \ldots, \mu_n \ddot{q}_n = Q_n \tag{3.2}$$

where Q_1, Q_2, Q_3 are the components of $\boldsymbol{F}_1$, Q_4, Q_5, Q_6 are the components of $\boldsymbol{F}_2$, and so on.

If we know the forces as functions of the positions and velocities of the particles and of time, then the equations of motion form a system of differential equations that determine the evolution from the initial values of q_a and $\dot{q}_a$ $(a = 1, 2, \ldots, n)$. There are many dynamical trajectories through a given point of C, depending on when and in what state of motion the system passes through the corresponding configuration.

To get a cleaner picture, we double the number of variables, by introducing the *phase space* P in which

$$q_1, q_2, \ldots, q_n, v_1, v_2, \ldots,\ v_n \tag{3.3}$$

are coordinates and then write the equations of motion in the first-order form

$$\begin{gathered} \mu_1 \dot{v}_1 = Q_1, \ldots,\ \mu_n \dot{v}_n = Q_n \\ \dot{q}_1 = v_1, \ldots,\ \dot{q}_n = v_n. \end{gathered} \tag{3.4}$$

This is a common trick in the theory of differential equations. We have replaced the n second-order differential equations (3.2) by $2n$ first-order equations (3.4). Each point of P corresponds to a particular state of motion: a particular configuration together with a particular set of velocities for the particles.

In an *autonomous* system, the forces $\boldsymbol{F}_\alpha$ can depend on the positions and velocities of the particles, but not on t. In this case, the dynamical behaviour of the system does not depend on the time at which it is set in motion. So there is just one dynamical trajectory through each point of phase space.

If, however, the forces depend on time as well as on configuration and velocity, then this is no longer true. In this case the equations of motion are of the form

$$\mu\dot{v} = Q(q, v, t), \qquad \dot{q} = v. \tag{3.5}$$

If the system is set in motion at some time t_0 with given initial values of the coordinates q_a and v_a, then the subsequent behaviour depends on t_0. In such *non-autonomous* systems, it is convenient to picture the motion in the *extended configuration space* $C \times \mathbb{R}$, which has coordinates $q_1, q_2, \ldots q_n, t$, or in the *extended phase space* $P \times \mathbb{R}$, which has coordinates $q_1, \ldots, q_n, v_1, \ldots, v_n, t$. There is always just one dynamical trajectory through each point of $P \times \mathbb{R}$, whether or not the system is autonomous.

3.3 Coordinate Transformations

It is not always convenient to use their individual Cartesian coordinates to label the configuration of a system of particles. This is particularly true when the system is subject to constraints, which may look simpler when expressed, for example, in terms of the separations of the particles rather than of the positions relative to the frame of reference.

The techniques developed in the last chapter extend to a general method for determining the equations of motion of a system of particles in a general coordinate system. They enable us to transform the equations of motion from the original Cartesian system q_a to a new coordinate system $\tilde{q}_a$ which is related to the original by expressions of the form

$$\tilde{q}_a = \tilde{q}_a(q_1, q_2, \ldots, q_n, t). \tag{3.6}$$

The coordinates $\tilde{q}_a$ in the new system are called *generalized coordinates*.

By allowing the $\tilde{q}_a$s to depend on t as well as on the q_as, we are allowing for the possibility that the new coordinate system is moving relative to the old. For a single particle, for example, the $\tilde{q}_a$s might be Cartesian coordinates in a rotating frame, or spherical polar coordinates. For two particles, they

might be the three coordinates of the centre of mass together with three other coordinates that determine the position of one of the particles relative to the centre of mass, and so on.

More formally, we are looking at a coordinate transformation in the extended configuration space $C \times \mathbb{R}$. The original coordinates on $C \times \mathbb{R}$ are $q_1, \ldots, q_n, t$. The new ones are the tilded coordinates, given in terms of the old by expressions of the form

$$\tilde{q}_a = \tilde{q}_a(q_1, q_2, \ldots, q_n, t), \qquad \tilde{t} = t + c,$$

where c is constant. By permitting c to be nonzero, we allow a change in the origin of the time coordinate at the same time as a change in the configuration coordinates. Even when $c = 0$, we maintain the distinction between t, the time coordinate in the original system, and $\tilde{t}$, the time coordinate in the new system, because the partial derivatives with respect to time in the two systems are not the same: $\partial/\partial t$ denotes the derivative with respect to time with $q_1, \ldots, q_n$ held fixed, while $\partial/\partial \tilde{t}$ denotes the derivative with respect to time with $\tilde{q}_1, \ldots, \tilde{q}_n$ held fixed.

The first step in determining the equations of motion in the new coordinates is to extend the transformation to the coordinates on $P \times \mathbb{R}$. During any kinematically possible motion of the system, not necessarily the one determined by the equations of motion, the configuration of the system changes traces out a curve γ in $C \times \mathbb{R}$, given by

$$q_a = q_a(t),$$

and its state of motion traces out a curve in $P \times \mathbb{R}$ given by

$$q_a = q_a(t), \qquad v_a = \dot{q}_a(t). \tag{3.7}$$

Such curves are called *kinematic trajectories*.

Any curve in $C \times \mathbb{R}$ which can be parametrized by t is a possible kinematic trajectory. However, not every such curve in $P \times \mathbb{R}$ is a possible kinematic trajectory because the second of Equations (3.7) implies a special relationship between the velocity and configuration coordinates along the curve, which does not generally hold when the q_as and v_as are arbitrary functions of time.

Example 3.1

For a system with one degree of freedom, the kinematic trajectory given by $q = \cos t$ has corresponding phase space trajectory given by

$$t \mapsto (q, v, t) = (\cos t, -\sin t, t),$$

which is a helix in $\mathbb{R}^3$, as in Figure 2.7, p. 45. However the curve in $P \times \mathbb{R}$ given by $q(t) = \cos t$, $v(t) = -2 \sin t$ is not a kinematic trajectory because $v(t) \neq \dot{q}(t)$.

The coordinate transformation in $P \times \mathbb{R}$ is determined by the condition that for any kinematic trajectory in $C \times \mathbb{R}$, the corresponding trajectory in $P \times \mathbb{R}$ should be the same in the new coordinates as in the old. In other words, when we transform (3.7) from the original coordinate system q_a, v_a, t to the new system $\tilde{q}_a, \tilde{v}_a, t$, we have the same relationship

$$\tilde{v}_a = \dot{\tilde{q}}_a$$

as before for any motion of the system. By applying the chain rule,

$$\dot{\tilde{q}}_a(t) = \frac{\partial \tilde{q}_a}{\partial q_1}\dot{q}_1 + \frac{\partial \tilde{q}_a}{\partial q_2}\dot{q}_2 + \cdots + \frac{\partial \tilde{q}_a}{\partial q_n}\dot{q}_n + \frac{\partial \tilde{q}_a}{\partial t}$$

for $a = 1, 2, \ldots, n$. Therefore the coordinate transformation in $P \times \mathbb{R}$ is

$$\begin{aligned} \tilde{q}_a &= \tilde{q}_a(q,t) \\ \tilde{v}_a &= \tilde{v}_a(q,v,t) = \sum_{b=1}^{n} \frac{\partial \tilde{q}_a}{\partial q_b} v_b + \frac{\partial \tilde{q}_a}{\partial t} \\ \tilde{t} &= t + c. \end{aligned} \tag{3.8}$$

It is necessary to add the condition that the coordinate transformation in $P \times \mathbb{R}$ should be non-singular. We should be able to invert (3.8) and express the original coordinates q_a, v_a, t as smooth functions of $\tilde{q}_a$, $\tilde{v}_a$, and $\tilde{t}$. The condition for this to be possible, at least locally, is given by the inverse function theorem. To invert (3.6) and so obtain q_a as a function of $\tilde{q}_1, ..., \tilde{q}_n$ and $\tilde{t} = t$, we require that the Jacobian matrix with entries $\partial \tilde{q}_a / \partial q_b$ should have nowhere vanishing determinant. When this condition holds, we can write $q_a = q_a(\tilde{q}, \tilde{t})$ and invert (3.8) to obtain

$$v_a = \sum_{b=1}^{n} \frac{\partial q_a}{\partial \tilde{q}_b}\tilde{v}_b + \frac{\partial q_a}{\partial \tilde{t}}$$

The summation signs that appear in such coordinate transformations can make the expressions involved look unnaturally complicated and hard to read. It is therefore useful to introduce some new notation. First, we use $\tilde{q}_a = \tilde{q}_a(q,t)$ as a convenient shorthand for $\tilde{q}_a = \tilde{q}_a(q_1, \ldots, q_n, t)$, and so on. Second we apply the *Einstein conventions* to the indices $a, b, c, \ldots = 1, 2, \ldots, n$. For an index in the first part of the alphabet, a repetition of the index in a term is taken to indicate a sum over $1, 2, ..., n$, and any equation is understood to hold for all possible values of its free (unsummed) indices. With these conventions, the second of Equations (3.8) is understood to hold for $a = 1, 2, \ldots, n$, and the sum over $b = 1, 2, \ldots, n$ in the expression $(\partial \tilde{q}_a / \partial q_b) v_b$ is implied by the repetition of b, so the summation sign is redundant.

Example 3.2

With these conventions, we have

$$\tilde{v}_a(q,v,t) = \frac{\partial \tilde{q}_a}{\partial q_b} v_b + \frac{\partial \tilde{q}_a}{\partial t}, \qquad v_a(\tilde{q},\tilde{v},\tilde{t}) = \frac{\partial q_a}{\partial \tilde{q}_b}\tilde{v}_b + \frac{\partial q_a}{\partial \tilde{t}}.$$

Example 3.3

The *Kronecker delta* is defined by $\delta_{ab} = 1$ when $a = b$ and $\delta_{ab} = 0$ when $a \neq b$. For any $x_1, x_2, \ldots, x_n$, we have $\delta_{ab} x_b = x_a$.

Example 3.4

The expression $x_a x_a$ has the same value as $x_b x_b$. One can change *dummy indices*, that is, indices which are summed over, in the same way that one can change dummy variables in an integral.

The equation $x_a = y_a$ can equally well be written $x_b = y_b$. Either way , it is equivalent to

$$x_1 = y_1, \quad x_2 = y_2, \ldots, x_n = y_n.$$

But it is not equivalent to $x_a = y_b$ because the conventions do not tell us to take as understood the qualification "for $a = b = 1, 2, \ldots, n$" that would be necessary to get equations equivalent to $x_a = y_a$. For this reason, for an equation to make sense according to the Einstein conventions, the free indices on each must match up.

Example 3.5

Let L, M, N be $n \times n$ matrices. In terms of the entries in the matrices, the matrix equation $L = MN$ is

$$L_{ab} = \sum_{c=1}^{n} M_{ac} N_{cb}, \qquad a, b = 1, \ldots, n.$$

With the Einstein conventions, this is

$$L_{ab} = M_{ac} N_{cb},$$

with the summation over c and the range of the free indices a, b understood. The entries in the identity matrix are given by the Kronecker delta, so for example the condition for N to be the inverse of M is $M_{ab} N_{bc} = \delta_{ac}$.

Example 3.6

By the chain rule,

$$\frac{\partial \tilde{q}_b}{\partial q_a}\frac{\partial q_a}{\partial \tilde{q}_c} = \frac{\partial \tilde{q}_b}{\partial \tilde{q}_c} = \delta_{bc}. \tag{3.9}$$

3.4 Transformation of the Equations of Motion

Our task is to rewrite the equations of motion in terms of the new coordinates. The direct substitution of (3.8) into (3.4) can lead to cumbersome and unmanageable expressions. A more elegant approach is to make use of the following fundamental fact, which is a direct extension of Proposition 2.21 to several degrees of freedom.

Proposition 3.7

Let γ be a kinematic trajectory in $C \times \mathbb{R}$ given by $q_a = q_a(t)$ and suppose that T is a twice continuously differentiable function of the $2n+1$ variables q_a, v_a, and t. Then

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial v_a}\right) - \frac{\partial T}{\partial q_a} = \frac{\partial \tilde{q}_b}{\partial q_a}\left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \tilde{v}_b}\right) - \frac{\partial T}{\partial \tilde{q}_b}\right]$$

where $\mathrm{d}/\mathrm{d}t$ denotes differentiation with respect to time along the corresponding phase space trajectory.

Remember that on the right-hand side, there is a sum over $b = 1, 2, ..., n$. Just as in the case of one degree of freedom, there is a distinction between $\mathrm{d}/\mathrm{d}t$ and $\partial/\partial t$. If $f = f(q, v, t)$ is a function on $P \times \mathbb{R}$, then

(i) $\mathrm{d}f/\mathrm{d}t$ is obtained by substituting $q_a = q_a(t)$, $v_a = \dot{q}_a(t)$ into $f = f(q, v, t)$ and then differentiating with respect to t;

(ii) $\partial f/\partial t$ is obtained by differentiating f with respect to t, holding the values of q_a and v_a fixed.

We follow the same convention as before in interpreting partial derivatives. The partial derivative with respect to a variable without a tilde is taken with the other variables without tildes held constant, and the partial derivative with respect to a variable with a tilde is taken with the other variables with tildes held constant. Thus $\partial/\partial\tilde{t}$ is the partial derivate with respect to time with the values of $\tilde{q}_a$ and $\tilde{v}_a$ held fixed. In particular, $\partial/\partial t$ is not the same derivative operator as $\partial/\partial\tilde{t}$, despite the fact that we may have $t = \tilde{t}$.

Again there are two proofs of the proposition. We give here the longer direct proof based on the chain rule and later give the shorter proof based on the calculus of variations. As this first proof does no more that add suffices to the variables in the first proof of Proposition 2.21 and make use of the Einstein conventions to avoid covering the page in summation signs, it is a useful exercise to try to write it out before reading the following.

Proof

The only difference from the case of a single degree of freedom is that when calculating partial derivatives by the chain rule, we must sum over the subscripts on the q_a's and v_a's. Thus

$$\begin{aligned}\frac{\partial T}{\partial q_a} &= \frac{\partial T}{\partial \tilde{q}_b}\frac{\partial \tilde{q}_b}{\partial q_a} + \frac{\partial T}{\partial \tilde{v}_b}\frac{\partial \tilde{v}_b}{\partial q_a} + \frac{\partial T}{\partial \tilde{t}}\frac{\partial \tilde{t}}{\partial q_a} \\ &= \frac{\partial T}{\partial \tilde{q}_b}\frac{\partial \tilde{q}_b}{\partial q_a} + \frac{\partial T}{\partial \tilde{v}_b}\left[\frac{\partial^2 \tilde{q}_b}{\partial q_a \partial q_c}v_c + \frac{\partial^2 \tilde{q}_b}{\partial q_a \partial t}\right] \\ \frac{\partial T}{\partial v_a} &= \frac{\partial T}{\partial \tilde{q}_b}\frac{\partial \tilde{q}_b}{\partial v_a} + \frac{\partial T}{\partial \tilde{v}_b}\frac{\partial \tilde{v}_b}{\partial v_a} + \frac{\partial T}{\partial \tilde{t}}\frac{\partial \tilde{t}}{\partial v_a} \\ &= \frac{\partial T}{\partial \tilde{v}_b}\frac{\partial \tilde{q}_b}{\partial q_a}.\end{aligned} \tag{3.10}$$

Hence

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial v_a}\right) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \tilde{v}_b}\right)\frac{\partial \tilde{q}_b}{\partial q_a} + \frac{\partial T}{\partial \tilde{v}_b}\left[\frac{\partial^2 \tilde{q}_b}{\partial q_a \partial q_c}v_c + \frac{\partial^2 \tilde{q}_b}{\partial q_a \partial t}\right] \tag{3.11}$$

because

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \tilde{q}_b}{\partial q_a}\right) = \frac{\partial^2 \tilde{q}_b}{\partial q_a \partial q_c}v_c + \frac{\partial^2 \tilde{q}_b}{\partial q_a \partial t}.$$

Remember that the repetition of an index in a term implies a sum over $1, 2, \ldots, n$. As before, the proposition follows on subtracting (3.10) from (3.11). □

The proposition shows that the combination of derivatives

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial v_a}\right) - \frac{\partial T}{\partial q_a}$$

has a very straightforward transformation rule under a change of coordinates, although, as is clear from the proof, the individual terms do not behave in a particularly simple way. The rule is characteristic of a *covector*, a term that is explained in Chapter 8. As in the case of a single degree of freedom, if the

forces on the right-hand side are transformed in the same way, then we arrive very simply at the equations of motion in the new system.

The first step is to introduce the total kinetic energy

$$T = \tfrac{1}{2}(\mu_1 v_1^2 + \mu_2 v_2^2 + \cdots + \mu_n v_n^2),$$

where the μ_a's are defined by (3.1). Because $\partial T/\partial q_a = 0$, the equations of motion (3.2) can be written

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial v_a}\right) - \frac{\partial T}{\partial q_a} = Q_a, \qquad \frac{\mathrm{d}q_a}{\mathrm{d}t} = v_a$$

The point of this apparent complication is that we can now use the transformation rule to write down the equations in the new coordinates. Apply the proposition to the first equation to get

$$\frac{\partial \tilde{q}_b}{\partial q_a}\left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \tilde{v}_b}\right) - \frac{\partial T}{\partial \tilde{q}_b}\right] = Q_a$$

Multiply both sides by $\partial q_a/\partial \tilde{q}_c$, sum over a, and make use of (3.9). The result is

$$\delta_{bc}\left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \tilde{v}_b}\right) - \frac{\partial T}{\partial \tilde{q}_b}\right] = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \tilde{v}_c}\right) - \frac{\partial T}{\partial \tilde{q}_c} = \tilde{Q}_c,$$

where the $\tilde{Q}_c$s are defined by

$$\tilde{Q}_c = \frac{\partial q_a}{\partial \tilde{q}_c} Q_a \quad \text{or equivalently by} \quad Q_a = \frac{\partial \tilde{q}_c}{\partial q_a}\tilde{Q}_c. \tag{3.12}$$

Hence the equations of motion in the new coordinates are

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \tilde{v}_a}\right) - \frac{\partial T}{\partial \tilde{q}_a} = \tilde{Q}_a, \quad \frac{\mathrm{d}\tilde{q}_a}{\mathrm{d}t} = \tilde{v}_a. \tag{3.13}$$

The first equation in (3.12) is again the covector transformation rule. It applies not only between the original Cartesian coordinates and the new coordinates but also between two systems of generalized coordinates. If we have a second generalized coordinate system $\hat{q}_a$, then the corresponding quantities $\hat{Q}_a$ are given by

$$\hat{Q}_a = \frac{\partial q_b}{\partial \hat{q}_a} Q_b = \frac{\partial q_b}{\partial \hat{q}_a}\frac{\partial \tilde{q}_c}{\partial q_b}\tilde{Q}_c$$

by substituting from the second equation in (3.12). However

$$\frac{\partial q_b}{\partial \hat{q}_a}\frac{\partial \tilde{q}_c}{\partial q_b} = \frac{\partial \tilde{q}_c}{\partial \hat{q}_a}$$

by the chain rule. Therefore, when we change from one generalized coordinate system to another, we have the same transformation rule

$$\hat{Q}_a = \frac{\partial \tilde{q}_c}{\partial \hat{q}_a}\tilde{Q}_c. \tag{3.14}$$

As always, remember that there is a sum here over c.

3.5 Generalized Forces

In order to make use of the transformed equations of motion (3.13), we need a rather more practical way to find the quantities $\tilde{Q}_a$ than working directly from their definition in (3.12). This is provided by an analysis of the work done by the forces acting on the particles under small displacements.

If we know the values of the $\tilde{q}_a$s and $\tilde{t}$, then we also know the q_as and hence the positions of all the particles. Therefore $\boldsymbol{r}_\alpha = \boldsymbol{r}_\alpha(\tilde{q}, \tilde{t})$ and we can rewrite the definition in the form

$$\tilde{Q}_a = \sum_{\alpha=1}^{N} \boldsymbol{F}_\alpha \cdot \frac{\partial \boldsymbol{r}_\alpha}{\partial \tilde{q}_a}.$$

Suppose now that we make a small displacement in the configuration of the system by sending $\tilde{q}_a$ to $\tilde{q}_a + \delta\tilde{q}_a$, keeping t fixed (Figure 3.1). Then particle α is moved from the point with position vector $\boldsymbol{r}_\alpha$ by translation through

$$\delta \boldsymbol{r}_\alpha = \frac{\partial \boldsymbol{r}_\alpha}{\partial \tilde{q}_a} \delta\tilde{q}_a$$

to the first order. Therefore

$$\tilde{Q}_a \delta\tilde{q}_a = \sum_{\alpha} \boldsymbol{F}_\alpha \cdot \delta \boldsymbol{r}_\alpha.$$

The quantity on the right-hand side is the work done by the forces $\boldsymbol{F}_\alpha$ during the displacement.

This observation gives a convenient method for calculating the $\tilde{Q}_a$'s. To find, for example, $\tilde{Q}_1$, consider the displacement of the system given by $\tilde{q}_1 \mapsto \tilde{q}_1 + \delta\tilde{q}_1$, with $\tilde{q}_2, \ldots, \tilde{q}_n$, and $\tilde{t}$ held fixed. Calculate, to the first order in $\delta\tilde{q}_1$, the total work done by all the forces during this displacement and equate the result to $\delta\tilde{q}_1 \tilde{Q}_1$.

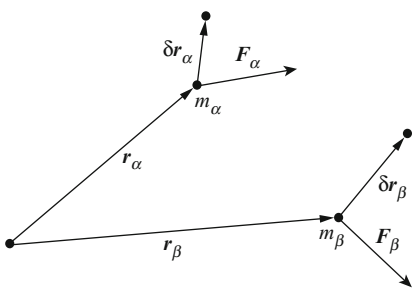

Figure 3.1

Many of the forces we come across in examples are *conservative* in the sense that they can be derived from a *potential* function U. For these, there is a simpler way of the finding the $\tilde{Q}_a$s. For suppose that there exists a function U of the positions of the particles such that

$$\boldsymbol{F}_\alpha = -\boldsymbol{\nabla}_\alpha U,$$

where $\boldsymbol{\nabla}_\alpha$ is the gradient operator acting on U as a function of the position particle α, with all the other particles held fixed. Then, in this case,

$$\tilde{Q}_a = -\sum_{\alpha=1}^{N} \boldsymbol{\nabla}_\alpha U \cdot \frac{\partial \boldsymbol{r}_\alpha}{\partial \tilde{q}_a} = -\frac{\partial U}{\partial \tilde{q}_a},$$

by the chain rule. So the $\tilde{Q}_a$s can be found simply by expressing U in terms of the transformed coordinates and differentiating. This is still true if the forces are time-dependent and U depends on t as well as the positions of the particles. Most authors would not describe such forces as *conservative* because energy is generally not conserved during motion under their influence. However, they *are* conservative in the sense that they do no net work under imaginary displacements around a closed path at a fixed value of t. For this reason, we still use the term, with caution, for time-dependent forces that can be derived from a potential. More is said about this below.

Example 3.8

A particle of mass m moves under the influence of a force $\boldsymbol{F}$. The original coordinates are

$$q_1 = x, \quad q_2 = y, \quad q_3 = z,$$

where x, y, and z are Cartesian coordinates in some inertial frame. Take the $\tilde{q}_a$ to be spherical polar coordinates: $\tilde{q}_1 = r$, $\tilde{q}_2 = \theta$, $\tilde{q}_3 = \varphi$ where

$$\begin{aligned} x &= r\sin\theta\cos\varphi \\ y &= r\sin\theta\sin\varphi \\ z &= r\cos\theta. \end{aligned}$$

In terms of these

$$\begin{aligned} T &= \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\dot{\varphi}^2\sin^2\theta) \\ &= \tfrac{1}{2}m(\tilde{v}_1^2 + \tilde{q}_1^2\tilde{v}_2^2 + \tilde{q}_1^2\tilde{v}_3^2\sin^2\tilde{q}_2). \end{aligned}$$

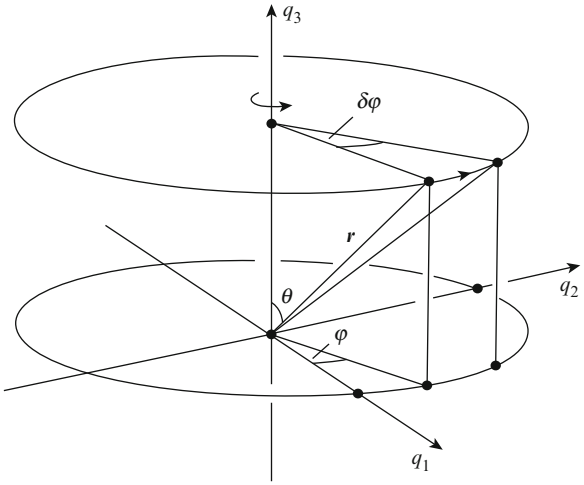

Figure 3.2

Thus the equations of motion are

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \tilde{v}_1}\right) - \frac{\partial T}{\partial \tilde{q}_1} &= m\ddot{r} - mr(\dot{\theta}^2 + \dot{\varphi}^2 \sin^2\theta) = \tilde{Q}_1 \\
\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \tilde{v}_2}\right) - \frac{\partial T}{\partial \tilde{q}_2} &= mr^2\ddot{\theta} + 2mr\dot{r}\dot{\theta} - mr^2\dot{\varphi}^2 \sin\theta\cos\theta = \tilde{Q}_2 \\
\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \tilde{v}_3}\right) - \frac{\partial T}{\partial \tilde{q}_3} &= \frac{\mathrm{d}}{\mathrm{d}t}(mr^2\dot{\varphi}\sin^2\theta) = \tilde{Q}_3
\end{aligned}$$

To find the $\tilde{Q}_1$, $\tilde{Q}_2$, and $\tilde{Q}_3$, we follow the instructions. For example, for $\tilde{Q}_3$, we consider the displacement $\varphi \mapsto \varphi + \delta\varphi$, with r, θ, and t fixed (Figure 3.2). This changes the position of the particle by

$$\delta x = -\delta\varphi\, r\sin\theta\sin\varphi, \quad \delta y = \delta\varphi\, r\sin\theta\cos\varphi, \quad \delta z = 0$$

to the first order in $\delta\varphi$. Therefore, by cancelling $\delta\varphi$,

$$\tilde{Q}_3 = -Q_1 r\sin\theta\sin\varphi + Q_2 r\sin\theta\cos\varphi,$$

where Q_1, Q_2, and Q_3 are the x, y, and z components of $\boldsymbol{F}$.

Now that they have served their purpose, we drop the tildes. Henceforth the q_as and v_as are arbitrary generalized coordinates and velocities. We have shown that the equations of motion for a system of particles are

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial v_a}\right) - \frac{\partial T}{\partial q_a} = Q_a, \quad \dot{q}_a = v_a,$$

where the Q_as are found by considering the work done by the forces acting on the particles under small displacements at constant time.

Definition 3.9

The Q_as are *generalized forces* and the v_as are *generalized velocities*. The Q_as are also called the *q-components* of the $\boldsymbol{F}_\alpha$s.

We often omit the adjective 'generalized'. At times we also use a common, but potentially dangerous shorthand notation, in which v_a is replaced by $\dot{q}_a$, for example in writing the equations of motion in the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_a}\right) - \frac{\partial T}{\partial q_a} = Q_a. \tag{3.15}$$

The danger is that of confusing two meanings of '$\dot{q}_a$': at one stage it is a generalized velocity coordinate on $P \times \mathbb{R}$, treated as independent of q_a and t; at another, it is the time derivative $\mathrm{d}q_a/\mathrm{d}t$. The q_a, v_a notation, in which the two meanings are clearly distinguished, is safer, but more cumbersome.

We see later that a small displacement in the configuration of the system at constant time is represented by a *tangent vector* to C. The generalized forces are the components of a *covector*, and the work done under the displacement is the quantity given by the natural pairing of vectors and covectors. Proposition 3.7 establishes that quantities

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial v_a}\right) - \frac{\partial T}{\partial q_a}$$

are also the components of a covector. So equations of motion equate two covectors in a way that is independent of the choice of generalized coordinates.

EXERCISES

3.1. In Example 3.8, find $\tilde{Q}_1$ and $\tilde{Q}_2$ in terms of the components of $\boldsymbol{F}$ by considering the work done under suitable small displacements. Check that the same expressions for the $\tilde{Q}_a$s follow from the chain rule in the case that $\boldsymbol{F} = -\boldsymbol{\nabla} U$.

3.6 Lagrange's Equations

There is a notable simplification in the case that the forces acting on the system can be derived from a potential $U = U(q, t)$, so that

$$Q_a = -\frac{\partial U}{\partial q_a}. \tag{3.16}$$

If we change the system of generalized coordinates from q_a to $\tilde{q}_a$, then we have

$$\frac{\partial U}{\partial \tilde{q}_a} = \frac{\partial q_b}{\partial \tilde{q}_a}\frac{\partial U}{\partial q_b}$$

while the generalized forces transform according to (3.14). Thus if (3.16) holds, then we also have $\tilde{Q}_a = -\partial U/\tilde{q}_a$. So the relationship between the generalized forces and the potential does not depend on the choice of generalized coordinates. Moreover, in this case the equations of motion take the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v_a}\right) - \frac{\partial L}{\partial q_a} = 0, \quad \frac{\mathrm{d}q_a}{\mathrm{d}t} = v_a.$$

where $L = T - U$, in any system of generalized coordinates. These are *Lagrange's equations*. Written out in full, they are

$$\frac{\partial^2 L}{\partial v_a \partial v_b}\ddot{q}_b + \frac{\partial^2 L}{\partial v_a \partial q_b}\dot{q}_b + \frac{\partial^2 L}{\partial v_a \partial t} - \frac{\partial L}{\partial q_a} = 0, \quad \dot{q}_a = v_a.$$

Definition 3.10

The function $L = T - U$ on the extended phase space is called the *Lagrangian*.

It is important to be able to recognize forces that can be derived from a potential function. The first point is that such forces can depend only on $q_1, \ldots, q_n$ and t, but not, for example, on $v_1, \ldots, v_n$. It must be possible to determine the forces given the configuration and the time, but not the state of motion. There are generalizations of the notion for velocity-dependent forces, but we do not consider them.

The second point is concerned with the origin of the term 'conservative'. Consider, as before, imaginary displacements of the system at a fixed time t. Suppose that we move the particles through a continuous sequence of configurations represented by a curve $q_a = q_a(s)$ in C, where $s \in [s_1, s_2]$ is some parameter (Figure 3.3). Then $\boldsymbol{r}_\alpha = \boldsymbol{r}_\alpha(s)$ and the total work done by the forces during this displacement is

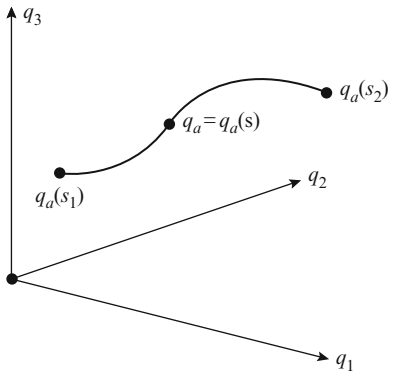

Figure 3.3

$$W = \sum_{\alpha} \int_{s_1}^{s_2} \boldsymbol{F}_{\alpha} \cdot \frac{\mathrm{d}\boldsymbol{r}_{\alpha}}{\mathrm{d}s}\,\mathrm{d}s = \int_{s_1}^{s_2} Q_a(q(s),t)\frac{\mathrm{d}q_a}{\mathrm{d}s}\,\mathrm{d}s.$$

If $Q_a = -\partial U/\partial q_a$ for some potential function $U(q,t)$, then

$$\begin{aligned} W \;&=\; -\int_{s_1}^{s_2} \frac{\partial U}{\partial q_a}\frac{\mathrm{d}q_a}{\mathrm{d}s}\,\mathrm{d}s \;=\; -\int_{s_1}^{s_2} \frac{\mathrm{d}}{\mathrm{d}s}\Big(U(q(s),t)\Big)\,\mathrm{d}s \\ &=\; U\big(q(s_1),t\big) - U\big(q(s_2),t\big)\,, \end{aligned} \tag{3.17}$$

and so the work done depends only on $q_a(s_1)$ and $q_a(s_2)$. It depends only on the initial and final configurations and not on the intermediate configurations.

The converse is also true. If for any curve in C, the work done in displacing the system along the curve at fixed t depends only on the endpoints of the curve, then the forces can be derived from a potential. Equivalently, the forces are conservative if the work done in displacing the system around any closed curve at fixed t vanishes (a curve is closed if its two endpoints coincide). Hence the term 'conservative'. For such curves, there is no net expenditure of energy, but take care to note the qualification 'at fixed t'. The displacement is not a feasible motion, but an imaginary one undertaken with the values of the forces frozen at the chosen value of t. If the forces are time-dependent, then energy will generally not be conserved during actual motions.

One can see from this that gravitational and elastic forces are conservative. For example, no net work is done by gravity in lifting a particle and then returning it to its initial position, and no net work is done by elastic forces in stretching and then relaxing a perfectly elastic string. However, friction is not conservative.

Equation (3.17) also gives an interpretation of U. The difference in the values of U at two different configurations at the same time t is the energy required to change the first configuration into the second. Thus in problems involving gravitational or elastic forces, U is the potential energy, to within the addition of an arbitrary constant.

At least locally, there exists a potential $U(q,t)$ for forces $Q_a(q,t)$ if and only if

$$\frac{\partial Q_a}{\partial q_b} = \frac{\partial Q_b}{\partial q_a}.$$

This integrability condition is sometimes useful. The qualification 'at least locally' is needed to avoid problems that arise when C is not simply-connected. In general the statement holds only in some neighbourhood in C.

Example 3.11

If the z-axis is vertical and the force $\boldsymbol{F}$ in Example 3.10 is gravity, then $U = mgz = mgr\cos\theta$. From this, we can read off that $Q_1 = Q_2 = 0$, $Q_3 = -mg$ and that

$$\begin{aligned}\tilde{Q}_1 &= -\frac{\partial U}{\partial r} = -mg\cos\theta\\ \tilde{Q}_2 &= -\frac{\partial U}{\partial \theta} = mgr\sin\theta\\ \tilde{Q}_3 &= -\frac{\partial U}{\partial \varphi} = 0.\end{aligned}$$

Example 3.12

Suppose that q_1, q_2, q_3 are the Cartesian coordinates of a particle of mass m in an inertial frame R. Put

$$\tilde{q}_1 = q_1 - tu, \quad \tilde{q}_2 = q_2, \quad \tilde{q}_3 = q_3, \quad \tilde{t} = t.$$

Then the $\tilde{q}_a$s are Cartesian coordinates of a new inertial frame $\tilde{R}$, which is moving relative to the original frame R at constant speed u in the q_1-direction.

In the new coordinates, $\tilde{v}_1 = v_1 - u$ and the kinetic energy relative to R is

$$T = \tfrac{1}{2}m(v_1^2 + v_2^2 + v_3^2) = \tfrac{1}{2}m(\tilde{v}_1^2 + \tilde{v}_2^2 + \tilde{v}_3^2 + 2u\tilde{v}_1 + u^2),$$

This is slightly alarming. If we had begun by referring the motion to $\tilde{R}$ instead of R, then we would have started by writing

$$\tilde{T} = \tfrac{1}{2}m(\tilde{v}_1^2 + \tilde{v}_2^2 + \tilde{v}_3^2),$$

which is the kinetic energy relative to $\tilde{R}$. However, the discrepancy is not serious because

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial v_a}\right) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \tilde{T}}{\partial \tilde{v}_a}\right), \qquad \frac{\partial \tilde{T}}{\partial \tilde{q}_a} = \frac{\partial T}{\partial q_a} = 0,$$

so the equations of motion are the same, whichever route is followed.

Example 3.13

We can transform from the Cartesian coordinates q_1, q_2, q_3 of an inertial frame $R = (O, \mathcal{T})$ to the Cartesian coordinates $\tilde{q}_1$, $\tilde{q}_2$, $\tilde{q}_3$ of a uniformly rotating frame $\tilde{R} = (O, \tilde{\mathcal{T}})$ by writing

$$\begin{aligned}
\tilde{q}_1 &= q_1 \cos\omega t + q_2 \sin\omega t \\
\tilde{q}_2 &= -q_1 \sin\omega t + q_2 \cos\omega t \\
\tilde{q}_3 &= q_3
\end{aligned}$$

where ω is constant. The $\tilde{q}$-axes are rotating relative to the q-axes with angular speed ω about the q_3-axis (Figure 3.4).

The corresponding velocity coordinates are related by

$$\begin{aligned}
\tilde{v}_1 &= v_1 \cos\omega t + v_2 \sin\omega t - \omega q_1 \sin\omega t + \omega q_2 \cos\omega t, \\
\tilde{v}_2 &= -v_1 \sin\omega t + v_2 \cos\omega t - \omega q_1 \cos\omega t - \omega q_2 \sin\omega t, \\
\tilde{v}_3 &= v_3.
\end{aligned}$$

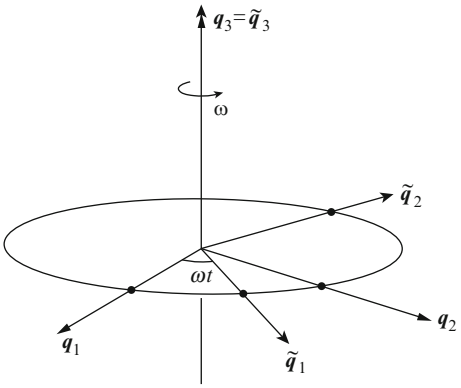

Figure 3.4

For a single particle, the kinetic energy relative to R is

$$\begin{aligned} T &= \tfrac{1}{2}m(v_1^2 + v_2^2 + v_3^2) \\ &= \tfrac{1}{2}m(\tilde{v}_1^2 + \tilde{v}_2^2 + \tilde{v}_3^2 - 2\omega\tilde{q}_2\tilde{v}_1 + 2\omega\tilde{q}_1\tilde{v}_2 + \omega^2\tilde{q}_1^2 + \omega^2\tilde{q}_2^2). \end{aligned}$$

Hence the equations of motion in the absence of forces are

$$\begin{aligned} m\frac{\mathrm{d}}{\mathrm{d}t}(\tilde{v}_1 - \omega\tilde{q}_2) - m\omega\tilde{v}_2 - m\omega^2\tilde{q}_1 &= 0, & \frac{\mathrm{d}\tilde{q}_1}{\mathrm{d}t} &= \tilde{v}_1 \\ m\frac{\mathrm{d}}{\mathrm{d}t}(\tilde{v}_2 + \omega\tilde{q}_1) + m\omega\tilde{v}_1 - m\omega^2\tilde{q}_2 &= 0, & \frac{\mathrm{d}\tilde{q}_2}{\mathrm{d}t} &= \tilde{v}_2 \\ m\frac{\mathrm{d}}{\mathrm{d}t}(\tilde{v}_3) &= 0, & \frac{\mathrm{d}\tilde{q}_3}{\mathrm{d}t} &= \tilde{v}_3. \end{aligned}$$

It is easy to see that these are equivalent to

$$m\ddot{\boldsymbol{r}} + 2m\boldsymbol{\omega} \wedge \dot{\boldsymbol{r}} + m\boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge \boldsymbol{r}) = 0,$$

where $\boldsymbol{r}$ is the position vector from O, $\boldsymbol{\omega} = \omega\boldsymbol{e}_3$, and the dot denotes the time derivative with respect to the non-inertial frame.

EXERCISES

3.2. A particle of mass m is subject to a force $\boldsymbol{F}$. Obtain the equations of motion in cylindrical polar coordinates.

3.3. A particle of unit mass is subject to an inverse-square-law central force

$$\boldsymbol{F} = -\frac{\boldsymbol{r}}{r^3}$$

where $r = |\boldsymbol{r}|$ and $\boldsymbol{r}$ is the position vector from the origin of an inertial frame. Show that the motion is governed by the Lagrangian

$$L = \tfrac{1}{2}\dot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}} + \frac{1}{r}.$$

Write down the equations of motion in spherical polar coordinates and show that there are solutions with $\theta = \pi/2$ throughout the motion.

3.4. Two particles, each of mass m, are moving under their mutual gravitational attraction, which is given by the potential $U = -\gamma m/2r$, where $2r$ is their separation and γ is a constant. Find the equations of motion in terms of the coordinates X, Y, Z, r, θ, φ, where X, Y,

and Z are the Cartesian coordinates of the centre of mass and r, θ, and φ are the polar coordinates of one particle relative to the centre of mass.

3.5. The dynamics of a system with n degrees of freedom are governed by a Lagrangian $L(q, v, t)$. Show that if $f(q, t)$ is any function on CT, then

$$L' = L + \frac{\partial f}{\partial q_a} v_a + \frac{\partial f}{\partial t}$$

generates the same dynamics.

3.6. A system has Lagrangian $L = \frac{1}{2} T_{ab} v_a v_b$ where the T_{ab}s are functions of the q_as alone. Show that

$$\frac{\mathrm{d}L}{\mathrm{d}t} = 0.$$

Show that if f is any function of one variable, then $L' = f(L)$ generates the same dynamics.

3.7. †Two particles P_1 and P_2 have respective masses m_1 and m_2 and are attracted to each other by a force with time-independent potential $U(\boldsymbol{r})$, where $\boldsymbol{r}$ is the vector from P_1 to P_2.

(1) Show that the motion of the centre of mass is the same as in the case $U = 0$.

(2) Show that the motion of P_2 relative to P_1 is the same as in the case that P_1 is fixed and P_2 is attracted to P_1 by a force with potential

$$V = \frac{m_1 + m_2}{m_1} U\,.$$

3.8. A particle of unit mass is moving in the x, y plane under the influence of the potential

$$U = -\frac{1}{r_1} - \frac{1}{r_2}\,,$$

where r_1 and r_2 are the respective distances from the points $(1, 0)$ and $(-1, 0)$. Show that if new coordinates are introduced by putting

$$x = \cosh\varphi \cos\theta, \qquad y = \sinh\varphi \sin\theta\,,$$

then the Lagrangian governing the motion becomes

$$L = \tfrac{1}{2}\Omega(\dot\theta^2 + \dot\varphi^2) + 2\Omega^{-1}\cosh\varphi,$$

where $\Omega = \cosh^2\varphi - \cos^2\theta$.

Write down the equations of motion and show that if the particle is set in motion with $T + U = 0$, then

$$\tfrac{1}{2}(\dot\theta^2 + \dot\varphi^2) - \frac{2\cosh\varphi}{\Omega^2} = 0\,.$$

Deduce that

$$\frac{\mathrm{d}^2\theta}{\mathrm{d}\tau^2} = 0, \qquad \frac{\mathrm{d}^2\varphi}{\mathrm{d}\tau^2} = 2\sinh\varphi$$

where $\mathrm{d}t/\mathrm{d}\tau = \Omega$. Hence find the path of the particle.

3.7 Holonomic Constraints

In a wide class of mechanical problems we know that the motion of the system satisfies certain constraints, but we are not interested in the forces that maintain the constraints. All that we want to understand is the behaviour of the remaining unconstrained degrees of freedom. The Lagrangian formalism is particularly useful in this type of problem because it provides a simple general method for eliminating unknown and unwanted constraint forces.

Before turning to the details, let us consider some examples.

Example 3.14

A particle confined to the surface of a smooth sphere of radius a. Here there is one constraint equation

$$x^2 + y^2 + z^2 = a^2$$

and two remaining degrees of freedom, corresponding to the two polar angles θ and φ. The force that maintains the constraint is the normal reaction $\boldsymbol{N}$. To find the motion of the particle, we need two differential equations for θ and φ which do not involve $\boldsymbol{N}$.

Example 3.15

A bead sliding along a smooth rotating wire. Here there are two constraint equations and one degree of freedom, the position of the bead on the wire.

Example 3.16

In an idealized, and somewhat unrealistic model, a rigid body is a large collection of N particles with position vectors are subject to the constraints

$$|\boldsymbol{r}_\alpha - \boldsymbol{r}_\beta| = \text{constant}. \tag{3.18}$$

These reduce the number of degrees of freedom from 3N to six, three for the position of the centre of mass and three for rotations about the centre of mass.

Example 3.17

A rigid sphere rolling on a rough plane. Here, in addition to (3.18), we have the extra condition

$$\dot{\boldsymbol{r}} + \boldsymbol{\omega} \wedge (-a\boldsymbol{k}) = 0$$

where $\boldsymbol{r}$ is the position vector of the centre, $\boldsymbol{\omega}$ is the angular velocity of the sphere, a is the radius, and $\boldsymbol{k}$ is the unit vector normal to the plane. This is the condition that the particle of the sphere in contact with the plane should be instantaneously at rest.

In the first and third examples, the constraint equations are of the form

$$f(q) = 0. \tag{3.19}$$

In the second example, we have a moving constraint. The condition that the bead remains on the wire is given by two equations of the form

$$f(q, t) = 0. \tag{3.20}$$

In the final example, the constraint equations also involve the velocities of the particles making up the sphere. Equation (2.4.3) gives rise to two additional equations of the form

$$f(q, v, t) = 0. \tag{3.21}$$

Constraints like (3.19) and (3.20) that do not involve the velocities are said to be *holonomic*. Holonomic constraints are further subdivided into *fixed* or *scleronomic* constraints that do not involve t, as in (3.19), and *moving* or *rheonomic* constraints that do involve t, as in (3.20). For the moment we deal only with holonomic constraints.

Suppose that we have a system of N particles, as in sections 3.2, subject to $n - n'$ constraints of the form

$$f_r(q, t) = 0; \qquad r = 1, 2, ..., n - n' \tag{3.22}$$

where the q_as are generalized coordinates. We think of these as defining an $(n' + 1)$-dimensional *constraint surface* in $C \times \mathbb{R}$ (Figure 3.5).

Provided that the constraints are independent, we should be able to use (3.22) to determine the configuration of the system from a knowledge of t and n' of the coordinates q_a. The constraints should reduce the number of degrees of freedom from $n = 3N$ to n'. This is made precise by the following.

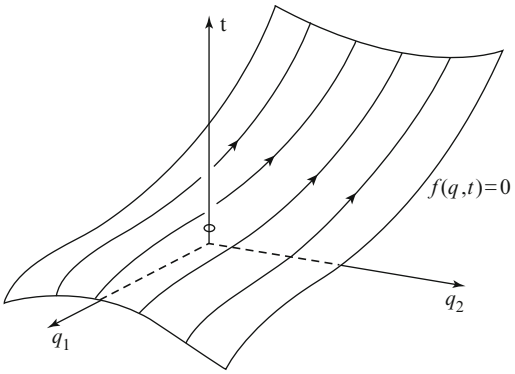

Figure 3.5 The constraint surface ($n = 2$, $n' = 1$).

Definition 3.18

The constraints $f_r = 0$ are *independent* if the $(n - n') \times n$ matrix with entries $\partial f_r / \partial q_a$ has maximal rank at each point.

Proposition 3.19

If the constraints $f_r = 0$ are independent, then there exists a system of generalized coordinates $\tilde{q}_a$ in which the constraints are

$$\tilde{q}_{n'+1} = 0, \quad \tilde{q}_{n'+2} = 0, \ldots, \tilde{q}_n = 0. \tag{3.23}$$

Independence precludes meaningless constraints (such as $t = 0$) and prevents double counting. It fails, for example, if $f_r = f_s$ for some $s \neq r$. We shall not go through the proof of the proposition, but it would be a simple exercise in any course on advanced calculus. All that is needed is to find n' functions on $C \times \mathbb{R}$ which, together with t and the f_r's themselves, make up a coordinate system on $C \times \mathbb{R}$.

Definition 3.20

If (3.23) holds, then the generalized coordinates $\tilde{q}_a$ are said to be *adapted* to the constraints.

Example 3.21

For the particle moving on the surface of a sphere, the coordinates $q_1 = \theta$, $q_2 = \varphi$, $q_3 = r - a$ are adapted to the constraint $x^2 + y^2 + z^2 - a^2 = 0$.

3.8 External and Constraint Forces

Suppose that we can write the force $\boldsymbol{F}_\alpha$ on each particle as the sum of two terms

$$\boldsymbol{F}_\alpha = \boldsymbol{E}_\alpha + \boldsymbol{K}_\alpha$$

where the $\boldsymbol{E}_\alpha$ are known external forces such as gravity and the $\boldsymbol{K}_\alpha$ are constraint forces responsible for maintaining the constraints. In the first example (3.14), the normal reaction of the sphere on the particle is a constraint force, as are the normal reaction of the wire on the bead in the second example, and the forces between the particles in the rigid body in the third example.

Then we can also write

$$Q_a = E_a + K_a$$

where the E_as are the q-components of the $\boldsymbol{E}_\alpha$s and the K_as are q-components of the $\boldsymbol{K}_\alpha$s.

The only assumption that we make about the constraint forces K_a is that they are *workless* in the sense that *they do no work during any instantaneous displacement of the system consistent with the constraints*. The precise meaning of this assumption must be clearly understood. Fix a time t, and imagine that at this time, we take a photograph of the system with a camera that records not only the positions of the individual particles, but also the forces acting on them (as in Figure 3.6). Now imagine making a small change in the recorded configuration in which $q_a \mapsto q_a + \delta q_a$ with t held fixed. The particle α is moved from the point with position vector $\boldsymbol{r}_\alpha$ to the point with position vector $\boldsymbol{r}_\alpha + \delta\boldsymbol{r}_\alpha$ where

$$\delta\boldsymbol{r}_\alpha = \frac{\partial \boldsymbol{r}_\alpha}{\partial q_a}\delta q_a,$$

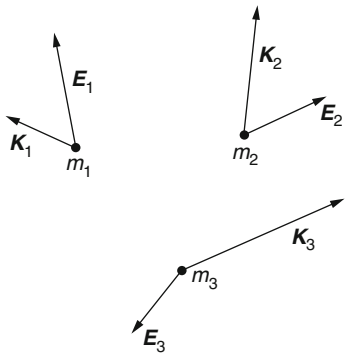

Figure 3.6

ignoring second-order terms in δq_a. The work done by the constraint forces is

$$W = \sum_\alpha \boldsymbol{K}_\alpha \cdot \delta \boldsymbol{r}_\alpha = K_a \delta q_a,$$

to the first order.

The assumption is that $W = 0$ whenever the displacement is consistent with the constraints. This means the displaced configuration must also be one that satisfies the constraint equations at time t. That is,

$$0 = f_r(q + \delta q, t) = \delta q_a \frac{\partial f_r}{\partial q_a} \tag{3.24}$$

to the first order for each r, because $f_r(q, t) = 0$ when the constraints are satisfied. Thus our assumption is that

$$K_a \delta q_a = 0 \quad \text{whenever} \quad \delta q_a \frac{\partial f_r}{\partial q_a} = 0 \quad \forall r \tag{3.25}$$

to the first order in the δq_a.

We can express this in a less clumsy way. Put $\delta q_a = \varepsilon u_a$, where $u_a \in \mathbb{R}$ and ε is a small parameter. Then, by cancelling ε, (3.25) becomes

$$K_a u_a = 0 \quad \text{whenever} \quad u_a \frac{\partial f_r}{\partial q_a} = 0 \quad \forall r,$$

which is now required to hold exactly, without the qualification 'to the first order'.

When the generalized coordinates q_a are adapted to the constraints, the second of Equations (3.25) is equivalent to

$$\delta q_{n'+1} = \delta q_{n'+2} = \cdots = \delta q_n = 0.$$

In this case the condition that the constraint forces should be workless reduces to

$$K_1 = K_2 = \cdots = K_{n'} = 0$$

and the equations of motion for $a = 1, 2, \ldots, n'$ become

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial v_a}\right) - \frac{\partial T}{\partial q_a} = E_a, \qquad \frac{\mathrm{d}q_a}{\mathrm{d}t} = v_a. \tag{3.26}$$

These combine to give n' second-order equations, the number needed, in which the constraint forces do not appear.

It is sometimes convenient to work in coordinates that are not adapted to the constraints. We can then still write

$$u_a \left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial v_a}\right) - \frac{\partial T}{\partial q_a} - E_a\right] = 0, \quad \text{whenever} \quad u_a \frac{\partial f_r}{\partial q_a} = 0 \quad \forall r.$$

This is a form of *d'Alembert's principle*.

It is important to understand two points. First, in the case of moving constraints, when the f_rs depend on time, a displacement satisfying (3.24) need not represent a possible motion of the system, whatever the external forces. If in time δt, the configuration changes from q_a to $q_a + \delta q_a$ during the actual motion, then the later configuration will satisfy $f_r(q + \delta q, t + \delta t) = 0$; that is

$$0 = \frac{\partial f_r}{\partial q_a}\delta q_a + \frac{\partial f_r}{\partial t}\delta t$$

to the first order in δq_a and δt. In general this is incompatible with (3.24).

Second, one can give no general justification for the assumption about the constraint forces. It is simply that in a large number of examples, and, in particular in the first three cited above, one can split the forces $\boldsymbol{F}_\alpha$ into a sum of known external forces $\boldsymbol{E}_\alpha$ and unknown constraint forces $\boldsymbol{K}_\alpha$ that can easily be seen to satisfy (3.25).

Example 3.22

For the bead sliding on a rotating wire, the constraint force is the normal reaction $\boldsymbol{N}$. If we freeze the motion of the wire and slide the bead a small distance along the wire, then $\boldsymbol{N}$ does no work because it is orthogonal to the wire. Hence the moving constraints in this system are workless. This is despite of the fact that $\boldsymbol{N}$ does do work during the actual motion, because the rotation contributes to the velocity of the bead a component normal to the wire.

It is not actually necessary to find all the coordinates q_a to make use of (3.26), and, for example, in the case of a rigid body made up of 10^{25} particles, it would not be practical in any case. With the coordinates adapted to the constraints, we know that $q_a = v_a = 0$ for $a = n' + 1, \ldots, n$ throughout the actual motion. To write down (3.26), therefore, we need only know the expression for T when $q_a = v_a = 0$ for $a = n' + 1, \ldots, n$. So we can obtain the equations of motion for the remaining coordinates $q_1, \ldots, q_{n'}$ by doing the following.

1. Choose coordinates $q_1, q_2, \ldots, q_{n'}$ that label the configurations satisfying the constraints.
2. Express T in terms of $q_1, \ldots, q_{n'}, v_1, \ldots, v_{n'}$, and t on the assumption that the constraints are satisfied at all times.
3. Find the q-components of the external forces by considering, for example, a small displacement consistent with the constraints of the form

$$q_1 \mapsto q_1 + \delta q_1, \qquad q_2 \mapsto q_2, \ldots, q_{n'} \mapsto q_{n'},$$

with t held fixed, and equating $E_1\delta q_1$ to the work done by the external forces.

When the external forces are conservative, the third step is much simpler. In this case there is a potential U for the $\boldsymbol{E}_\alpha$ and we can write

$$E_a = -\frac{\partial U}{\partial q_a} \qquad a = 1, 2, \ldots, n'.$$

The first of Equations (3.26) then reduces to the Lagrangian form

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v_a}\right) - \frac{\partial L}{\partial q_a} = 0$$

where $L = T - U$ and a runs over $1, 2, \ldots, n'$.

Definition 3.23

The integer n' is the number of *residual degrees of freedom*.

Unless otherwise stated, the q_a in problems with constraints will always be adapted coordinates; and the Einstein conventions will apply for $a, b, \ldots = 1, 2, \ldots, n'$, where n' is the number of residual degrees of freedom.

The q_as for $a = 1, 2, \ldots, n'$ are coordinates in the *constrained configuration space*; q_a and v_a $(a = 1, \ldots, n')$ are coordinates in the *constrained phase space*, and so on. When there is no danger of confusion, the prime on n' and the qualifications 'constrained' and 'residual' will be dropped. The notation C, $C \times \mathbb{R}$, P, and $P \times \mathbb{R}$ will be used for the constrained spaces associated with the residual degrees of freedom.

Example 3.24

Consider a particle of mass m moving on a smooth sphere of radius a, with gravity as the only external force. The constraint force is the normal reaction, which is workless, because it does no work during small displacements of the particle along the surface.

Here we can take (Figure 3.7)

$$q_1 = \theta, \qquad q_2 = \varphi$$

in terms of which the kinetic energy and the gravitational potential are given by

$$T = \tfrac{1}{2}ma^2(\dot{\theta}^2 + \dot{\varphi}^2 \sin^2\theta)$$
$$U = mga\cos\theta.$$

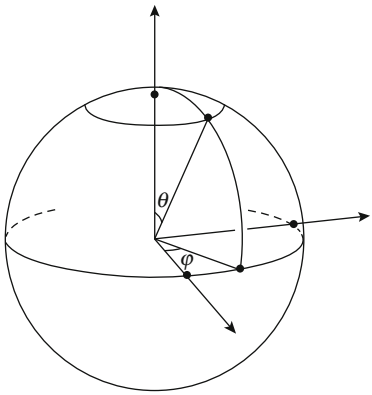

Figure 3.7

Thus the equations of motion are

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}(ma^2\dot{\theta}) - ma^2\dot{\varphi}^2\cos\theta\sin\theta - mga\sin\theta &= 0 \\ \frac{\mathrm{d}}{\mathrm{d}t}(ma^2\dot{\varphi}\sin^2\theta) &= 0. \end{aligned}$$

Example 3.25

In Example 1.21, the bead has one residual degree of freedom corresponding to the coordinate $q = \theta$. The velocity of the bead relative to the inertial frame is $\dot{\boldsymbol{r}} + \boldsymbol{\omega} \wedge \boldsymbol{r}$. That is

$$a\dot{\theta}(-\sin\theta\,\boldsymbol{i} + \cos\theta\,\boldsymbol{j}) + \omega a(\sin\alpha\sin\theta\,\boldsymbol{k} + \cos\alpha\cos\theta\,\boldsymbol{j} - \cos\alpha\sin\theta\,\boldsymbol{i}).$$

Hence

$$T = \tfrac{1}{2}ma^2\dot{\theta}^2 + ma^2\omega\dot{\theta}\cos\alpha + \tfrac{1}{2}ma^2\omega^2(\sin^2\alpha\sin^2\theta + \cos^2\alpha)$$

and

$$U = mga\sin\alpha\cos\theta.$$

Therefore the equation of motion is

$$ma^2\frac{\mathrm{d}}{\mathrm{d}t}(\dot{\theta} + \omega\cos\alpha) - ma^2\omega^2\sin^2\alpha\sin\theta\cos\theta - mga\sin\alpha\sin\theta = 0$$

in agreement with the result derived by vector methods.

The constrained configuration and phase spaces can have complicated topological properties and it may not be possible to use only a single coordinate system. For example, for the particle moving on a sphere, the polar coordinates θ and φ are singular on the polar axis at $\theta = 0$. Other coordinate systems must be used to describe motion near the north and south poles.

Example 3.26

Take k rods of various lengths which are smoothly jointed together at their ends to form a chain. Fix the two ends of the chain at two points separated by a distance slightly less than the total length of the rods. Then the configuration space is the hypersphere S^{2k-1} in $2k$-dimensional Euclidean space. The proof is by Morse theory, as was explained to me by Graeme Segal.

EXERCISES

3.9. A particle of mass m is free to move in a horizontal plane. It is attached to a fixed point O by a light elastic string, of natural length a and modulus λ. Show that the tension in the string is conservative and show that the Lagrangian for the motion when $r > a$ is

$$L = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - \frac{\lambda}{2a}(r-a)^2$$

where r and θ are plane polar coordinates with origin O.

3.10. The double pendulum (Figure 3.8) is free to move in a vertical plane. Show that the motion is governed by the Lagrangian

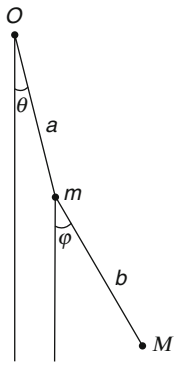

Figure 3.8

$$L = \tfrac{1}{2}ma^2\dot{\theta}^2 + \tfrac{1}{2}M(a^2\dot{\theta}^2 + b^2\dot{\varphi}^2 + 2ab\dot{\theta}\dot{\varphi}\cos(\theta - \varphi)) + mga\cos\theta + Mg(b\cos\varphi + a\cos\theta).$$

3.11. A particle of mass m is constrained to move under gravity on the surface of a smooth right circular cone of semi-vertical angle $\pi/4$. The axis of the cone is vertical, with the vertex downwards. Find the equations of motion in terms of z (the height above the vertex) and θ (the angular coordinate around the circular cross-sections). Show that

$$\dot{z}^2 + \frac{h^2}{2z^2} + gz = E$$

where E and h are constant. Sketch and interpret the trajectories in the z, $\dot{z}$-plane for a fixed value of h.

3.12. A particle P of mass m is attached to two light inextensible strings, each of length a. The strings pass over two smooth pegs A and B, which are at the same height and distance $2b$ apart. At the other ends of the strings hang two particles of mass m, which can move up and down the vertical lines through A and B. The particle P can move in the vertical plane containing A and B.

Show that if $2b\cosh\varphi = PA + PB$ and $2b\cos\theta = PA - PB$, then the kinetic energy of P is

$$T = \tfrac{1}{2}mb^2(\cosh^2\varphi - \cos^2\theta)(\dot{\theta}^2 + \dot{\varphi}^2).$$

Hence find the Lagrangian of the system in terms of θ and φ.

3.9 Stationary Action

As in the case of systems with one degree of freedom, Lagrange's equations also determine the critical curves of a variational problem. In fact we can extend the theory developed in Chapter 2 to several degrees of freedom with little further work.

Let $C \times \mathbb{R}$ and $P \times \mathbb{R}$ denote the extended configuration and phase spaces of some system with Lagrangian L, which we assume to be a twice continuously differentiable function of the coordinates $q_1, \ldots, q_n, v_1, \ldots, v_n, t$. Fix two points in $C \times \mathbb{R}$ with respective coordinates

$$(q_a, t) = (x_a, t_1), \qquad (q_a, t) = (y_a, t_2),$$

and let $\gamma \subset C \times \mathbb{R}$ be a path from the first point to the second, given by $q_a = q_a(t)$, where the $q_a(t)$s are twice continuously differentiable functions of t. So γ represents a kinematically possible trajectory of the system from the configuration $q_a = x_a$ at time t_1 to the configuration $q_a = y_a$ at time t_2. Define the *action* by

$$J_L(\gamma) = \int_{t_1}^{t_2} L\big(q(t), \dot{q}(t), t\big)\, \mathrm{d}t.$$

A *variation* in γ with fixed endpoints is represented by a small change $\delta q_a(t)$ in the trajectory. with $\delta q_a(t_1) = 0 = \delta q_a(t_2)$. As in the proof of Euler's theorem, Proposition 2.15,

$$\delta J_L = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q_a} \delta q_a + \frac{\partial L}{\partial v_a} \dot{\delta} q_a \right) \mathrm{d}t = \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q_a} - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial v_a} \right) \right] \delta q_a \, \mathrm{d}t,$$

to the first order in δq_a. Two immediate applications are the following.

Hamilton's Principle

If $q_a = q_a(t)$ is a solution of Lagrange's then $\delta J_L = 0$ for every such variation. Conversely, if $\delta J_L = 0$ for every variation with fixed end-points, then $q_a(t)$ satisfies Lagrange's equation. This follows from Proposition 2.15 by considering variations in which all but one of the functions δq_a vanish.

So, as before, the solutions to Lagrange's equations are the critical curves of the action J_L. In other words, the actual dynamical motion from the configuration $q_a = x_a$ at time t_1 to $q_a = y_a$ at time t_2 is characterized amongst all kinematically possible motions by the fact that the action is unchanged under small variations. This is Hamilton's principle. It is also called the *principle of least action* or, more properly, *stationary action*, because the dynamical trajectory need not minimize J_L. Also one should not lose sight of the fact that there may be more than one dynamical trajectory joining the two points of $C \times \mathbb{R}$. An example can be found by considering Exercise 3.14.

Proof of Proposition 3.7

By replacing L by T, we have, as in first proof of Proposition 2.21

$$\int_{t_1}^{t_2} \left\{ \left[\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial T}{\partial v_a} \right) - \frac{\partial T}{\partial q_a} \right] - \left[\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial \tilde{T}}{\partial \tilde{v}_b} \right) - \frac{\partial \tilde{T}}{\partial q_b} \right] \frac{\partial \tilde{q}_b}{\partial q_a} \right\} \delta q_a \, \mathrm{d}t = 0.$$

Proposition 3.7 is an immediate consequence.

EXERCISES

3.13. Show that if every curve from A to B is a critical curve of J_L for a system with one degree of freedom, then

$$L(q, v, t) = \frac{\partial f}{\partial q_a} v_a + \frac{\partial f}{\partial t}$$

for some function $f = f(q, t)$ on $C \times \mathbb{R}$.

3.14. A particle of unit mass is constrained to move on the surface of a unit sphere, but is otherwise free. Show that the dynamical trajectories are great circles traversed at uniform speeds. Show that if γ is a complete circuit of a great circle in time t, then

$$J_L(v) = 2\pi^2/t.$$

Does γ minimize J_L over all curves on the sphere that start and end at a point P on the equator and take time t for the round trip from P back to P?

4 Noether's Theorem

4.1 Ignorable Coordinates

A central recurring theme in mathematical physics is the connection between symmetries and conservation laws, in particular the connection between the symmetries of Euclidean space under rotation and translation and the conservation laws for linear and angular momentum. The same connections exist in quantum theory and in relativity, where the symmetry of Minkowski space under translations and Lorentz transformations underlies the relativistic interpretation of energy, momentum, and angular momentum. The connection is particularly transparent in the Lagrangian formalism.

As an example, consider the motion of a free particle in spherical polar coordinates. This is governed by the Lagrangian

$$L = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\dot{\varphi}^2\sin^2\theta). \tag{4.1}$$

Because $\partial L/\partial\varphi = 0$, we can read off from Lagrange's equations that the angular momentum about the z-axis,

$$\frac{\partial L}{\partial\dot{\varphi}} = mr^2\dot{\varphi}\sin^2\theta \tag{4.2}$$

is conserved.

More generally, suppose that we have a system of particles subject to conservative forces and that we have expressed the Lagrangian $L = T - U$ as a

N.M.J. Woodhouse, *Introduction to Analytical Dynamics: New Edition*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-1-84882-816-2_4,

function of t and of the generalized position and velocity coordinates q_a and v_a, with $a = 1, 2, \ldots, n$. If L is independent of q_a for some a, then

$$p_a = \frac{\partial L}{\partial v_a} \tag{4.3}$$

is constant along the dynamical trajectories in $P \times \mathbb{R}$ because

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v_a}\right) = \frac{\partial L}{\partial q_a} = 0.$$

Definition 4.1

The quantity p_a is called the *generalized momentum* conjugate to q_a. A coordinate q_a such that $\partial L/\partial q_a = 0$ is said to be *ignorable* or *cyclic*.

Note that the condition

$$\frac{\partial L}{\partial q_1} = 0$$

is a property not just of the coordinate q_1, but of the entire coordinate system $q_1, \ldots, q_n$. It is possible, for example, to introduce new coordinates $\tilde{q}_a$ such that $\tilde{q}_1 = q_1$, but

$$\frac{\partial L}{\partial \tilde{q}_1} \neq \frac{\partial L}{\partial q_1}.$$

It can also happen that $\partial L/\partial \tilde{v}_1 \neq \partial L/\partial v_1$, so that the momentum conjugate to $\tilde{q}_1 = q_1$ need not be the same in the two coordinate systems.

For small ε, the quantity εp_1 is the energy needed to change the generalized velocity v_1 to $v_1 + \varepsilon$, keeping the configuration and the other generalized velocities fixed. This is the sense in which p_1 is a 'momentum'. Linear momentum is a measure of how much energy is needed to change linear velocity by a small amount, angular momentum is a measure of how much energy is required to alter angular velocity.

EXERCISES

4.1. A particle of unit mass moves under gravity on a smooth surface given in cylindrical polar coordinates z, r, θ by $z = f(r)$. Show that the motion is governed by the Lagrangian

$$L = \tfrac{1}{2}\dot{r}^2\big(1 + f'(r)^2\big) + \tfrac{1}{2}r^2\dot{\theta}^2 - gf(r).$$

Show that θ is an ignorable coordinate. Write down the conserved conjugate momentum and give its physical interpretation.

4.2 One-Parameter Transformation Groups

In a Lagrangian system with a cyclic coordinate q_a, the equations of motion are unchanged by adding a constant to q_a. So the existence of cyclic coordinates signals the presence of dynamical symmetry, that is, of the invariance of the dynamics under transformations of the extended configuration space. For example, the conservation of $mr^2\dot{\varphi}\sin^2\theta$ in our example arises from the fact that the Lagrangian and the equations of motion are unchanged by adding a constant to φ, so the dynamical behaviour is unchanged if the system is rotated about the z-axis before being set in motion.

In this example, there is nothing special about the z-axis. We could just as well measure the spherical polar coordinates from the x-axis. Then the cyclic coordinate would be the one associated with symmetry under rotations about the x-axis and the corresponding conserved quantity would be angular momentum about the x-axis.

It would be useful to be able to obtain this second conservation law without actually having to make the transformation to a new system of spherical polar coordinates, and, in general, to be able to spot the presence of a cyclic coordinate of one coordinate system when the Lagrangian is written in terms of another set of generalized coordinates and velocities.

The key to this problem is the recognition that conserved quantities should properly be associated with the action of symmetry groups, rather than with coordinate systems. Linear momenta are associated with the group of translations in space, energy with translation in time, and angular momenta with the action of the rotation group. If the Lagrangian is invariant under such an action, then there is a corresponding conserved quantity.

For the moment, we focus on *one-parameter groups of transformations*, where the group in question is the real numbers, under addition. The transformations act on the extended configuration space $C \times \mathbb{R}$ by changing the configuration of the system, possibly in different ways at different times.

Definition 4.2

A *one-parameter group of transformations* of the extended configuration space is a family of maps $\rho_s : C \times \mathbb{R} \to C \times \mathbb{R}$ labelled by $s \in \mathbb{R}$ such that

(i) ρ_0 is the identity

(ii) $\rho_{s'} \circ \rho_s = \rho_{s+s'}$ for all $s, s' \in \mathbb{R}$

(iii) $t \circ \rho_s = t$

Here t is interpreted as the map $C \times \mathbb{R} \to \mathbb{R}$ that sends (q_a, t) to t.

In coordinates, ρ_s is a family of maps of the form

$$\rho_s : (q_1, \ldots, q_n, t) \mapsto \big(x_1(q,t,s), \ldots, x_n(q,t,s), t\big).$$

So for small s, we have

$$(q_1, \ldots, q_n, t) \mapsto (q_1, \ldots, q_n, t) + s(u_1, \ldots, u_n, 0) + O(s^2),$$

where u_a is the value of $\partial x_a/\partial s$ at $s = 0$. The u_as are functions of the q_as and t, but not s. We call the them the *generators* of the group of transformations. They are the components of a *time-dependent vector field* on configuration space, but we leave this geometric interpretation until later. For the moment, we note the following two key properties.

– From the second property of ρ_s, if we ignore terms of order δs^2, then we have

$$x_a(q,t,s+\delta s) = x_a(q,t,s) + \delta s\, u_a\big(x(q,t,s),t\big).$$

Therefore, with the q_as and t held fixed, $x_a(q,t,s)$ is determined by solving

$$\frac{\mathrm{d}x_a}{\mathrm{d}s} = u_a(x_1, \ldots, x_n, t), \tag{4.4}$$

as a system of ordinary differential equations for the x_as as a functions of s, with the initial condition $x_a(q,t,0) = q_a$. This is the sense in which the u_as generate ρ_s.

– We can express the generators in a different coordinate system by making a coordinate transformation

$$\tilde{q}_a = \tilde{q}_a(q,t), \qquad \tilde{t} = t.$$

For small δs,

$$\tilde{q}_a(q + \delta s\, u, t) = \tilde{q}_a(q,t) + \delta s\, \frac{\partial \tilde{q}_a}{\partial q_b} u_b,$$

where we have ignored second order terms. Therefore in the new coordinates, the generators are

$$\tilde{u}_a = \frac{\partial \tilde{q}_a}{\partial q_b} u_b. \tag{4.5}$$

Example 4.3

A key example is that in which ρ_s is given in a coordinate system by

$$\rho_s : (q_1, \ldots, q_n, t) \mapsto (q_1 + su_1, \ldots, q_n + su_n, t) \tag{4.6}$$

for some constants u_a. Then the group translates the row vector $(q_1, \ldots, q_n)$ along the constant vector $(u_1, \ldots, u_n)$, leaving t unchanged, and the generators are the components u_a. By solving (4.4), we see that converse is also true. If the generators are constant, then the one-parameter group is given by (4.6).

Example 4.4

For a single particle moving in space, the extended configuration space is $\mathbb{R}^3 \times \mathbb{R}$, with coordinates (x, y, z, t). The one-parameter group of translations in the z direction is given by

$$(x, y, z, t) \mapsto (x, y, z + s, t).$$

The generators are $u_1 = u_2 = 0$, $u_3 = 1$.

The one-parameter group of rotations about the z-axis is given by

$$(x, y, z, t) \mapsto (x \cos s - y \sin s, x \sin s + y \cos s, z, t).$$

The generators are $u_1 = -y$, $u_2 = x$, $u_3 = 0$.

When the generators are constants, the transformations are translations of the coordinates, as in Example 4.3. In this case we say that the one-parameter group is a *dynamical symmetry* of a system with Lagrangian L if

$$u_a \frac{\partial L}{\partial q_a} = 0.$$

It then follows from Lagrange's equations that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(u_a \frac{\partial L}{\partial v_a}\right) = u_a \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v_a}\right) = u_a \frac{\partial L}{\partial q_a} = 0.$$

So the quantity

$$p = u_a \frac{\partial L}{\partial v_a}$$

is a constant of the motion. It is called the *momentum conjugate* to the transformation group. If $u_1 = 1$ and the other u_as are zero, then $p = p_1$, so this extends the terminology of 'coordinates' and 'conjugate momenta'.

When we change to a general coordinate system, the generators transform according to (4.5), and they are no longer constant. However, we do have the following.

Lemma 4.5

Let ρ_s be a one-parameter group of transformations with generators u_a in the coordinate system q_a and with generators $\tilde{u}_a$ in the coordinate system $\tilde{q}_a$, and let $L(q, v, t)$ be a function on $P \times \mathbb{R}$. Then

$$u_a \frac{\partial L}{\partial q_a} + \left(\frac{\partial u_a}{\partial q_b} v_b + \frac{\partial u_a}{\partial t}\right) \frac{\partial L}{\partial v_a} = \tilde{u}_a \frac{\partial L}{\partial \tilde{q}_a} + \left(\frac{\partial \tilde{u}_a}{\partial \tilde{q}_b} \tilde{v}_b + \frac{\partial \tilde{u}_a}{\partial t}\right) \frac{\partial L}{\partial \tilde{v}_a},$$

where $\tilde{u}_a$ and u_a are related by (4.5).

Proof

Suppose that γ is a kinematic trajectory in $C \times R$. Then $\gamma_s = \rho_s(\gamma)$ is a family of trajectories, labelled by s. When we put

$$q_a = q_a(t), \qquad v_a = \dot{q}_a(t)$$

along a dynamical trajectory, L becomes a function of t. If we make this substitution for each trajectory in the family, then the result is a function $L(s,t)$.

Now for small s, the trajectory γ_s is given by

$$q_a = q_a(t) + s u_a\big(q(t),t\big) + O(s^2),$$

where $q_a = q_a(t)$ is the equation of $\gamma_0 = \gamma$. Therefore, along γ_s

$$v_a = \dot{q}_a(t) + s\frac{\partial u_a}{\partial q_b}\dot{q}_b(t) + s\frac{\partial u_a}{\partial t} + O(s^2).$$

It follows that at $s = 0$,

$$\begin{aligned}\frac{\partial L}{\partial s} &= \frac{\partial q_a}{\partial s}\frac{\partial L}{\partial q_a} + \frac{\partial v_a}{\partial s}\frac{\partial L}{\partial v_a} \\ &= u_a\frac{\partial L}{\partial q_a} + \left(\frac{\partial u_a}{\partial q_b}v_b + \frac{\partial u_a}{\partial t}\right)\frac{\partial L}{\partial v_a}.\end{aligned}$$

But $\partial L/\partial s$ is independent of the coordinate system in which it is evaluated. So the lemma follows. □

Definition 4.6

The *derivative* of L under the action of a one-parameter group of transformations is

$$u_a\frac{\partial L}{\partial q_a} + \left(\frac{\partial u_a}{\partial q_b}v_b + \frac{\partial u_a}{\partial t}\right)\frac{\partial L}{\partial v_a}$$

on $P \times \mathbb{R}$. It is independent of the coordinate system in which it is evaluated.

The proof of the lemma gives a method for calculating the derivative. Evaluate L along a trajectory in $C \times \mathbb{R}$ and find the change in L when the trajectory is moved by the action of the group. In geometric language, the action of the group determines a vector field on $P \times \mathbb{R}$ and the derivative is the derivative along this vector field.

The lemma itself gives a means for recognizing dynamical symmetry in general coordinates. If $u_a \partial L/\partial q_a = 0$ in a coordinate system in which the u_as

are constant, then (4.6) is zero in this coordinate system and hence in every coordinate system. From (4.5), we also have that

$$u_a \frac{\partial L}{\partial v_a} = \tilde{u}_a \frac{\partial q_b}{\partial \tilde{q}_a} \frac{\partial L}{\partial v_b} = \tilde{u}_a \frac{\partial L}{\partial \tilde{v}_a}$$

so we get the same value for $p = u_a \partial L / \partial v_a$ in every coordinate system.

Definition 4.7

Let $\rho_s : C \times \mathbb{R} \to C \times \mathbb{R}$ be a one parameter group of transformations with generators u_a and let L be a Lagrangian function on $P \times \mathbb{R}$. The *momentum conjugate* to ρ_s is the function

$$p = u_a \frac{\partial L}{\partial v_a} \tag{4.7}$$

on $P \times \mathbb{R}$. The group is said to be a *dynamical symmetry* whenever the derivative of L under the action of the group vanishes.

Proposition 4.8 (Noether's Theorem)

If ρ_s is a dynamical symmetry of a system with Lagrangian $L(q, v, t)$, then its conjugate momentum is constant during the motion of the system.

Proof

By differentiating (4.7) with respect to t along the dynamical trajectories in $P \times \mathbb{R}$, we have

$$\begin{aligned} \frac{\mathrm{d}p}{\mathrm{d}t} &= \frac{\mathrm{d}u_a}{\mathrm{d}t} \frac{\partial L}{\partial v_a} + u_a \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial v_a} \right) \\ &= \left(\frac{\partial u_a}{\partial q_b} v_b + \frac{\partial u_a}{\partial t} \right) \frac{\partial L}{\partial v_a} + u_a \frac{\partial L}{\partial q_a} \\ &= 0. \end{aligned} \tag{4.8}$$

□

In fact, every dynamical symmetry is given by Example 4.3 in some coordinate system, at least in a neighbourhood of a point at which not all the u_as vanish. However, the proof above establishes Noether's theorem direct from the definition, without having to use this result.

Example 4.9

Suppose that q_1 is cyclic. Then the one-parameter transformation group

$$\rho_s : (q_1, q_2, \ldots, q_n, t) \mapsto (q_1 + s, q_2, \ldots, q_n, t)$$

is a dynamical symmetry, with generator $u = (1, 0, \ldots, 0)$. Noether's theorem gives the conservation of the conjugate momentum $p_1 = \partial L/\partial v_1$.

Example 4.10

For a free particle of mass m moving without forces, the Lagrangian is

$$L = \tfrac{1}{2}m(v_1^2 + v_2^2 + v_3^2), \tag{4.9}$$

in the coordinate system $q_1 = x$, $q_2 = y$, and $q_3 = z$. For the one-parameter group of rotations about the z-axis in Example 4.4, we have

$$u_1 = -q_2, \quad u_2 = q_1, \quad u_3 = 0. \tag{4.10}$$

This is a dynamical symmetry because

$$u_a \frac{\partial L}{\partial q_a} + \left(\frac{\partial u_a}{\partial q_b} v_b + \frac{\partial u_a}{\partial t} \right) \frac{\partial L}{\partial v_a} = -v_2 \frac{\partial L}{\partial v_1} + v_1 \frac{\partial L}{\partial v_2} = 0.$$

The conjugate conserved momentum is

$$p = u_a \frac{\partial L}{\partial v_a} = -mq_2 v_1 + mq_1 v_3, \tag{4.11}$$

which is the z-component of the angular momentum $m\boldsymbol{r} \wedge \boldsymbol{v}$.

EXERCISES

4.2. Identify the symmetry associated with the conserved quantity in Exercise 4.1.

4.3. Write down the generators of rotations about the x and y axes for the system in Example 4.10 and find the conjugate momenta. Verify that they are the x- and y-components of angular momentum.

4.3 Conservation of Energy

A conserved momentum arises when L is invariant under a group of instantaneous displacements in the state of motion. Another conservation law arises when L is independent of t. If L is such that

$$L(q, v, t + s) = L(q, v, t),$$

then the system behaves in the same way irrespective of the time at which it is set in motion and so there is symmetry under *time translation*.

We leave the general derivation of the corresponding conservation law for later and deal here only with a special case. Suppose that $L = T - U$, where $U = U(q, t)$ is the potential and T, the kinetic energy, is a homogeneous quadratic in the velocities,

$$T = \tfrac{1}{2} T_{ab}(q, t) v_a v_b. \tag{4.12}$$

Then $E = T + U$ is the *total energy* of the system.

Proposition 4.11

Let $L = T - U$, where T is a homogeneous quadratic in the velocities. If

$$\frac{\partial L}{\partial t} = 0,$$

then E is conserved.

Proof

Suppose that $\partial L/\partial t = 0$. Then $\partial U/\partial t = 0$ and $\partial T/\partial t = 0$ because T and U are homogeneous with different degrees in the velocities. Lagrange's equations are

$$\frac{\partial^2 T}{\partial v_a \partial v_b} \dot{v}_b + \frac{\partial^2 T}{\partial v_a \partial q_b} v_b - \frac{\partial T}{\partial q_a} + \frac{\partial U}{\partial q_a} = 0. \tag{4.13}$$

Hence, by multiplying by v_a and summing over a,

$$0 = \frac{\partial T}{\partial v_b} \frac{\mathrm{d} v_b}{\mathrm{d} t} + \frac{\partial T}{\partial q_b} v_b + v_a \frac{\partial U}{\partial q_a} = \frac{\mathrm{d}}{\mathrm{d} t}(T + U). \tag{4.14}$$

Here we have used

$$v_a \frac{\partial^2 T}{\partial v_a \partial v_b} = \frac{\partial T}{\partial v_b} \quad \text{and} \quad v_a \frac{\partial^2 T}{\partial v_a \partial q_b} = 2 \frac{\partial T}{\partial q_b} \tag{4.15}$$

which follow from Euler's theorem on homogeneous functions, because $\partial T/\partial v_b$ is homogeneous of degree one in the velocities and $\partial T/\partial q_b$ is homogeneous of degree two in the velocities. □

4.4 Momentum Principles

Noether's theorem states that the momentum conjugate to a symmetry is a constant of the motion. There is a weaker form of symmetry which, while it does not give rise directly to constants of the motion, is still useful in the analysis of general systems. Suppose that a system of particles has kinetic energy $T = T(q, v, t)$ in generalized coordinates.

Definition 4.12

A *kinematic symmetry* is a one-parameter group of transformations

$$\rho_s : C \times \mathbb{R} \to C \times \mathbb{R}$$

such that the derivative of T under the action of the group vanishes. That is,

$$u_a \frac{\partial T}{\partial q_a} + \left(\frac{\partial u_a}{\partial q_b} v_b + \frac{\partial u_a}{\partial t} \right) \frac{\partial T}{\partial v_a} = 0.$$

The momentum conjugate to a kinematic symmetry with generators u_a is defined to be the function $p : P \to \mathbb{R}$, where

$$p = u_a \frac{\partial T}{\partial v_a}. \tag{4.16}$$

This does not involve the dynamics of the system because the definition does not involve the forces acting on the particles.

If ρ_s is a kinematic symmetry, then the same calculation as in the proof of Noether's theorem, but using now (3.13) rather than Lagrange's equations, shows that

$$\frac{\mathrm{d}p}{\mathrm{d}t} = u_a Q_a \tag{4.17}$$

where p is the conjugate momentum and the Q_as are the generalized forces. In a constrained system subject to workless constraints, they are the q-components of the external forces. Thus a kinematic symmetry gives rise to a *momentum principle* in the form of an equation: rate of change of momentum p equals component of applied force $u_a Q_a$.

We investigate the possibilities first for a single particle with kinetic energy

$$T = \tfrac{1}{2} m \boldsymbol{v} \cdot \boldsymbol{v},$$

where $\boldsymbol{v}$ is the velocity relative an inertial frame. The *linear momentum* of the particle is the vector $\boldsymbol{p} = m\boldsymbol{v}$ and its *angular momentum* about the origin of the inertial frame is the vector

$$\boldsymbol{J} = m\boldsymbol{r} \wedge \boldsymbol{v}$$

where $\boldsymbol{r}$ is the position vector of the particle from the origin of the inertial frame.

Take the generalized coordinates to be the Cartesian coordinates of the particle in the inertial frame. The corresponding coordinates on $P \times \mathbb{R}$ are the three components of $\boldsymbol{r}$, the three components of $\boldsymbol{v}$, and t. We get evolution equations for $\boldsymbol{p}$ and $\boldsymbol{J}$ from (4.17) by considering translations and rotations about the origin of the inertial frame in this coordinate system.

Translations

Suppose that ρ_s moves the particle through distance s in the direction of a unit vector $\boldsymbol{k}$. That is

$$\rho_s : (\boldsymbol{r}, t) \mapsto (\boldsymbol{r} + s\boldsymbol{k}, t).$$

If we take the q_as to be the Cartesian coordinates of the particle, then the generators of the one-parameter group are the components of $\boldsymbol{k}$. Because these are constant and because T depends only on the generalized velocities, T is unchanged by the transformations, so ρ_s is a kinematic symmetry. We have

$$p = m\boldsymbol{k} \cdot \boldsymbol{v} = \boldsymbol{k} \cdot \boldsymbol{p}$$

is the $\boldsymbol{k}$ component of the linear momentum and

$$u_a Q_a = \boldsymbol{k} \cdot \boldsymbol{F},$$

is the $\boldsymbol{k}$-component of the force on the particle. The momentum principle is

$$\dot{p} = \boldsymbol{k} \cdot \boldsymbol{F}, \tag{4.18}$$

which is just the $\boldsymbol{k}$-component of the equation of motion. Because this holds for any $\boldsymbol{k}$, we have the linear momentum equation $\dot{\boldsymbol{p}} = \boldsymbol{F}$.

Rotations about the Origin

In this case, the transformation ρ_s rotates the position vector of the particle through an angle s in about an axis in the direction of a unit vector $\boldsymbol{k}$. To the first order in s, this changes the position and velocity of the particle by

$$(\boldsymbol{r}, \boldsymbol{v}, t) \mapsto (\boldsymbol{r}, \boldsymbol{v}, t) + s(\boldsymbol{k} \wedge \boldsymbol{r}, \boldsymbol{k} \wedge \boldsymbol{v}, t),$$

which leaves T unchanged because $\boldsymbol{v}$ is orthogonal to $\boldsymbol{k} \wedge \boldsymbol{v}$. Therefore ρ_s is a kinematic symmetry. We have

$$p = m(\boldsymbol{k} \wedge \boldsymbol{r}) \cdot \boldsymbol{v} = \boldsymbol{k} \cdot \boldsymbol{J}$$

and

$$u_a Q_a = (\boldsymbol{k} \wedge \boldsymbol{r}) \cdot \boldsymbol{F} = \boldsymbol{k} \cdot (\boldsymbol{r} \wedge \boldsymbol{F})$$

We deduce the *principle of angular momentum*

$$\dot{\boldsymbol{J}} = \boldsymbol{r} \wedge \boldsymbol{F}.$$

The right-hand side is the *moment* of the force $\boldsymbol{F}$ about the origin.

There is an immediate extension to a system of particles with kinetic energy

$$T = \tfrac{1}{2} \sum_{\alpha} m_\alpha \boldsymbol{v}_\alpha \cdot \boldsymbol{v}_\alpha.$$

The rotations and translations act on $P \times \mathbb{R}$ by acting on the position and velocity of each individual particle, and T is again invariant because each separate term in the sum is invariant. So from the action of the translations we get the *linear momentum* equation

$$\dot{\boldsymbol{p}} = \sum \boldsymbol{F}_\alpha$$

and from the action of the rotations about the origin, we get the *angular momentum equation*

$$\dot{\boldsymbol{J}} = \sum \boldsymbol{r}_\alpha \wedge \boldsymbol{F}_\alpha,$$

where

$$\boldsymbol{p} = \sum m_\alpha \boldsymbol{v}_\alpha \qquad \text{and} \qquad \boldsymbol{J} = \sum m_\alpha \boldsymbol{r}_\alpha \wedge \boldsymbol{v}_\alpha.$$

The forces on a system of particles can usually be split into a sum

$$\boldsymbol{F}_\alpha = \boldsymbol{E}_\alpha + \boldsymbol{I}_\alpha \tag{4.19}$$

where the $\boldsymbol{E}_\alpha$ are *external forces* (gravity and so on) and the $\boldsymbol{I}_\alpha$ are *internal forces* that arise from the mutual interactions between the particles. The only restriction on the internal forces is that they should do no work under small instantaneous displacements that do not alter the relative positions of the particles, that is, under rotations and translations of the whole system. For example, the forces between the particles in a rigid body are internal in this sense. In the absence of gravity and other external forces, one cannot extract energy from a rock simply by moving it or rotating it.

Only the external forces contribute to $u_a Q_a$. So we can rewrite the linear and angular momentum equations in the form

$$\dot{\boldsymbol{p}} = \sum \boldsymbol{E}_\alpha \qquad \text{and} \qquad \dot{\boldsymbol{J}} = \sum \boldsymbol{r}_\alpha \wedge \boldsymbol{E}_\alpha.$$

4.5 Relative Angular Momentum

For a system of particles, we can get a further momentum principle by considering rotations about a *moving* point C. If the point has position vector $\boldsymbol{c}(t)$ from the origin of the inertial frame, then the *angular momentum* $\boldsymbol{J}_C$ of a particle *relative to* C is defined by

$$\boldsymbol{J}_C = m(\boldsymbol{r} - \boldsymbol{c}) \wedge (\boldsymbol{v} - \dot{\boldsymbol{c}})$$

where $\boldsymbol{r}$ is the position vector from the origin of the inertial frame, $\boldsymbol{v}$ is the velocity relative to the frame, and the dot is the time derivative relative to the inertial frame. It depends on C, but is independent of the choice of inertial frame.

The relative angular momentum is related to rotations about the moving point. Let ρ_s denote the one-parameter group of transformations of the extended configuration space of a particle given by rotating the vector $\boldsymbol{r} - \boldsymbol{c}$ from C to the particle through an angle s in about an axis in the direction of a fixed unit vector $\boldsymbol{k}$. To the first order in s,

$$\rho_s : (\boldsymbol{r}, t) \mapsto (\boldsymbol{r}, t) + s\big(\boldsymbol{k} \wedge (\boldsymbol{r} - \boldsymbol{c}), t\big).$$

This changes the position and velocity of a moving particle by

$$(\boldsymbol{r}, \boldsymbol{v}, t) \mapsto (\boldsymbol{r}, \boldsymbol{v}, t) + s\big(\boldsymbol{k} \wedge (\boldsymbol{r} - \boldsymbol{c}), \boldsymbol{k} \wedge (\boldsymbol{v} - \dot{\boldsymbol{c}}), t\big).$$

Therefore

$$p = m\big(\boldsymbol{k} \wedge (\boldsymbol{r} - \boldsymbol{c})\big) \cdot \boldsymbol{v} = \boldsymbol{k} \cdot \boldsymbol{J}_C.$$

The 'momentum' conjugate to the family of rotations is the $\boldsymbol{k}$-component of the angular momentum relative to C.

For a system of particles, we similarly have $p = \boldsymbol{k} \cdot \boldsymbol{J}_C$, where $\boldsymbol{J}_C$ is now the *total angular momentum relative to* C, defined by

$$\boldsymbol{J}_C = \sum m_\alpha (\boldsymbol{r}_\alpha - \boldsymbol{c}) \wedge (\boldsymbol{v}_\alpha - \dot{\boldsymbol{c}}).$$

In general, however, ρ_s is not a kinematic symmetry. For a single particle, the infinitesimal change in T under the action of $\hat{\rho}_s$ for small s is

$$sm\big(\boldsymbol{k} \wedge (\boldsymbol{v} - \dot{\boldsymbol{c}})\big) \cdot \boldsymbol{v} = -s(\boldsymbol{k} \wedge \dot{\boldsymbol{c}}) \cdot \boldsymbol{p},$$

where $\boldsymbol{p}$ is the linear momentum. This does not vanish in general. Obvious exceptions arise when C is fixed, which gives nothing new, or when C is moving with the particle, so that $\dot{\boldsymbol{c}} = \boldsymbol{v}$, in which case the momentum equation is trivial.

For a system of particles, however, there is a more interesting case in which C is taken to be the centre of mass, so that

$$m\boldsymbol{c} = \sum_\alpha m_\alpha \boldsymbol{r}_\alpha,$$

where $m = \sum m_\alpha$ is total mass. In this case the infinitesimal change in T under $\hat{\rho}_s$ is again

$$-s(\boldsymbol{k} \wedge \dot{\boldsymbol{c}}) \cdot \boldsymbol{p},$$

where $\boldsymbol{p}$ is now the total linear momentum. However, with C as the centre of mass, $\boldsymbol{p} = m\dot{\boldsymbol{c}}$, so the change in T vanishes and therefore ρ_s is a kinematic symmetry. We conclude that the time-dependence of the angular momentum about the centre of mass is governed by

$$\dot{\boldsymbol{J}}_c = \sum (\boldsymbol{r} - \boldsymbol{c}) \wedge \boldsymbol{E},$$

where the sum is over the points at which the external forces $\boldsymbol{E}$ are applied.

It is important to remember that the equality between rate of change of angular momentum and total moment of external forces holds in general only for either a fixed point in an inertial frame or for the centre of mass.

EXERCISES

4.4. The configuration space for a particle of mass m moving in space is Euclidean space, with Cartesian coordinates x, y, z. Show that the generator of the one-parameter group of rotations about the x-axis is

$$u = (0, -z, y).$$

Hence write down an expression for the x-component of angular momentum in terms of the spherical polar coordinates defined by

$$\begin{aligned} x &= r \sin\theta \cos\varphi \\ y &= r \sin\theta \sin\varphi \\ z &= r \cos\theta. \end{aligned}$$

4.5. An astronaut is floating in empty space at rest relative to an inertial frame with her arms by her side. Explain how it is that by waving her arms and then returning them to their original position, she can rotate her body, but cannot move her centre of mass.

4.6. A system of particles with masses m_α moves in a gravitational field $\boldsymbol{g}$. Show that if $\boldsymbol{F}_\alpha = -m_\alpha \boldsymbol{g}$ where $\boldsymbol{g}$ is constant, then

$$\sum_\alpha \boldsymbol{F}_\alpha = m\boldsymbol{g} \quad \text{and} \quad \sum_\alpha \boldsymbol{r}_\alpha \wedge \boldsymbol{F}_\alpha = m\boldsymbol{c} \wedge \boldsymbol{g},$$

where $\boldsymbol{c}$ is the position vector of the centre of mass and m is the total mass. Deduce that the effect of a uniform gravitational field

on the total linear momentum and angular momentum is the same as that of a single force $m\boldsymbol{g}$ acting through the centre of mass. Show by counter-example that this is not true for a non-uniform field.

4.7. Let $\rho_s : C \times \mathbb{R} \to C \times \mathbb{R}$ be a dynamical symmetry of a system with Lagrangian L, with generators u_a. Show that under the action of ρ_s for small s, the change in $\partial L/\partial v_a$ along a kinematic trajectory is

$$\delta\left(\frac{\partial L}{\partial v_a}\right) = su_b\frac{\partial^2 L}{\partial v_a \partial q_b} + sw_b\frac{\partial^2 L}{\partial v_a \partial v_b}$$

to the first order in s, where

$$w_a = \frac{\partial u_a}{\partial q_b}v_b + \frac{\partial u_a}{\partial t}.$$

Deduce that

$$\delta\left(\frac{\partial L}{\partial v_a}\right) = -s\frac{\partial u_b}{\partial q_a}\frac{\partial L}{\partial v_b}$$

and

$$\delta\left(\frac{\partial L}{\partial q_a}\right) = -s\frac{\partial u_b}{\partial q_a}\frac{\partial L}{\partial q_b} - s\frac{\partial w_b}{\partial q_a}\frac{\partial L}{\partial v_b}.$$

Show that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial u_b}{\partial q_a}\right) = \frac{\partial w_b}{\partial q_a}.$$

Hence show that if $q_a = q_a(t)$ is a solution of Lagrange's equations, then so is $q_a = q_a(t) + su_a\big(q(t), t\big)$, to the first order in s. Deduce that if ρ_s is a dynamical symmetry, then it maps dynamical trajectories in $C \times \mathbb{R}$ to dynamical trajectories.

4.8. Establish the result of Exercise 4.7 by a variational argument.

4.10. Show that in the system in Exercise 3.7, if U depends only on the distance between the particles, then the three components of total angular momentum about the centre of mass are conserved.

5
Rigid Bodies

5.1 Rigid Body Motion

The motion of a rigid body at any instant is determined by the six components of two vectors, the angular velocity $\boldsymbol{\omega}$ and the velocity $\boldsymbol{u}$ of a chosen point O of the body. There are therefore six degrees of freedom and we should be able to describe the evolution by introducing six generalized coordinates: three for position and three for orientation.

The position coordinates are straightforward because we can use the three Cartesian coordinates of O in some inertial frame. The components of $\boldsymbol{u}$ are then the corresponding generalized velocities. As we saw in Section 1.9, it is a more complicated problem to find convenient coordinates to describe the rotational degrees of freedom. It is, in fact, impossible to find three generalized coordinates for which the three components of $\boldsymbol{\omega}$ are the corresponding generalized velocities. Whatever angular coordinates are used, the task of expressing $\boldsymbol{\omega}$ in terms of their time derivatives is always a source of complication.

Having chosen coordinates, we must express the kinetic energy as a function of q_a, v_a, and t. We can then write down the equations of motion in terms of the q-components of the external forces. In doing this, we think of the body as a collection of a large number of particles, with position vectors $\boldsymbol{r}_\alpha$, subject to the constraints

$$|\boldsymbol{r}_\alpha - \boldsymbol{r}_\beta| = \text{constant}.$$

N.M.J. Woodhouse, *Introduction to Analytical Dynamics: New Edition*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-1-84882-816-2_5,

The constraint forces are the forces between the particles and we assume that they do no work during instantaneous displacements of the body which leave unchanged the relative positions of the individual particles. They therefore make no contribution to the equations of motion for the coordinates q_a.

5.2 Kinetic Energy

Our first task is to express the kinetic energy of a rigid body in terms of $\boldsymbol{u}$ and $\boldsymbol{\omega}$. Choose an inertial frame, choose a point O fixed in the body, and let $\boldsymbol{u}$ denote the velocity of O relative to the inertial frame. The velocity of any other point of the body is

$$\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{\omega} \wedge \boldsymbol{r}$$

where $\boldsymbol{r}$ is the vector from O to the point and $\boldsymbol{\omega}$ is the angular velocity of the body. So the total kinetic energy relative to the inertial frame is

$$\begin{aligned} T &= \sum_\alpha \tfrac{1}{2} m_\alpha (\boldsymbol{u} + \boldsymbol{\omega} \wedge \boldsymbol{r}_\alpha) \cdot (\boldsymbol{u} + \boldsymbol{\omega} \wedge \boldsymbol{r}_\alpha) \\ &= \tfrac{1}{2} m \boldsymbol{u} \cdot \boldsymbol{u} + \boldsymbol{u} \cdot \Big(\boldsymbol{\omega} \wedge \sum_\alpha m_\alpha \boldsymbol{r}_\alpha\Big) + \tfrac{1}{2} \sum_\alpha m_\alpha (\boldsymbol{\omega} \wedge \boldsymbol{r}_\alpha) \cdot (\boldsymbol{\omega} \wedge \boldsymbol{r}_\alpha), \end{aligned}$$

where the sum is over all the particles, $m = \sum_\alpha m_\alpha$ is the total mass, and the position vector $\boldsymbol{r}_\alpha$ of particle α is from O as origin. The first term,

$$T_1 = \tfrac{1}{2} m \boldsymbol{u} \cdot \boldsymbol{u},$$

is the kinetic energy of a particle of mass m moving with velocity $\boldsymbol{u}$ relative to the inertial frame. The second term can be simplified to

$$T_2 = \boldsymbol{u} \cdot \Big(\boldsymbol{\omega} \wedge \sum_\alpha m_\alpha \boldsymbol{r}_\alpha\Big) = m \boldsymbol{u} \cdot (\boldsymbol{\omega} \wedge \boldsymbol{c}),$$

where $\boldsymbol{c}$ is the position vector from O of the centre of mass, defined by

$$m \boldsymbol{c} = \sum_\alpha m_\alpha \boldsymbol{r}_\alpha.$$

To make better sense of the third term, we replace the particles by a continuous distribution of matter. Then

$$T_3 = \tfrac{1}{2} \sum_\alpha m_\alpha (\boldsymbol{\omega} \wedge \boldsymbol{r}_\alpha) \cdot (\boldsymbol{\omega} \wedge \boldsymbol{r}_\alpha)$$

becomes an integral over the body

$$T_3 = \tfrac{1}{2} \int \rho (\boldsymbol{\omega} \wedge \boldsymbol{r}) \cdot (\boldsymbol{\omega} \wedge \boldsymbol{r}) \, \mathrm{d}V = \tfrac{1}{2} \int \rho \big((\boldsymbol{\omega} \cdot \boldsymbol{\omega})(\boldsymbol{r} \cdot \boldsymbol{r})^2 - (\boldsymbol{\omega} \cdot \boldsymbol{r})^2 \big) \, \mathrm{d}V,$$

in which ρ is the density, $\boldsymbol{r}$ is the position vector from O of a general point of the body, and $\mathrm{d}V$ is the volume element. Now $\boldsymbol{\omega}$ is constant over the body at any time. So we can bring a factor of $\boldsymbol{\omega}$ out of the integral, and write $T_3 = \frac{1}{2}\boldsymbol{\omega} \cdot \boldsymbol{J}$, where

$$\boldsymbol{J} = \int \rho\big((\boldsymbol{r} \cdot \boldsymbol{r})\boldsymbol{\omega} - (\boldsymbol{\omega} \cdot \boldsymbol{r})\boldsymbol{r}\big)\,\mathrm{d}V. \tag{5.1}$$

This vector is the *angular momentum relative to* O. It is the sum over the particles of the body of

$$m_\alpha \boldsymbol{r}_\alpha \wedge (\boldsymbol{\omega} \wedge \boldsymbol{r}_\alpha) = m_\alpha \boldsymbol{r}_\alpha \wedge (\boldsymbol{v}_\alpha - \boldsymbol{u}),$$

where $\boldsymbol{v}_\alpha - \boldsymbol{u}$ is the velocity of particle α relative to O, as measured in the inertial frame. Because the body has the same angular velocity in any inertial frame, T_3 depends on the motion of the body and the choice of the point O, but not on the choice of axes. On reassembling the three components of T, we have

$$T = \tfrac{1}{2}m\boldsymbol{u} \cdot \boldsymbol{u} + m\boldsymbol{u} \cdot (\boldsymbol{\omega} \wedge \boldsymbol{c}) + \tfrac{1}{2}\boldsymbol{J} \cdot \boldsymbol{\omega}. \tag{5.2}$$

5.3 The Inertia Matrix

The relative angular momentum depends linearly on $\boldsymbol{\omega}$. That is, it is given by a linear a transformation $\boldsymbol{\omega} \mapsto \boldsymbol{J}$. Given a frame fixed in the body, the transformation is represented by a matrix called the *inertia matrix*.

Definition 5.1

Let $R = (O, \mathcal{T})$ be a rest frame of the body. The *inertia matrix* of the body in the frame R is the 3×3 symmetric matrix $\mathcal{J}$ with entries

$$\mathcal{J}_{ij} = \int \rho(r^2\delta_{ij} - r_i r_j)\,\mathrm{d}V$$

where $r^2 = \boldsymbol{r} \cdot \boldsymbol{r}$, the r_is are the components of $\boldsymbol{r}$, and δ_{ij} is the Kronecker delta, equal to one when $i = j$ and to zero otherwise.

The notation is, perhaps, less than transparent. To bring it into a more familiar form, write x, y, and z for r_1, r_2, and r_3, and put

$$\begin{aligned} A &= \int \rho(y^2 + z^2)\,\mathrm{d}V & F &= \int \rho yz\,\mathrm{d}V \\ B &= \int \rho(x^2 + z^2)\,\mathrm{d}V & G &= \int \rho xz\,\mathrm{d}V \\ C &= \int \rho(x^2 + y^2)\,\mathrm{d}V & H &= \int \rho xy\,\mathrm{d}V. \end{aligned}$$

Then

$$\mathcal{J} = \begin{pmatrix} A & -H & -G \\ -H & B & -F \\ -G & -F & C \end{pmatrix}$$

The entries are all independent of time.

Definition 5.2

The diagonal entries A, B, and C are the *moments of inertia* about the x, y, and z coordinate axes, and F, G, and H are the *products of inertia* with respect to the yz, zx, and xy coordinate planes.

Because the components of $\boldsymbol{\omega}$ are constant over the body, (5.1) is equivalent to

$$J_i = \sum_j \omega_j \int \rho(r^2\delta_{ij} - r_i r_j)\,\mathrm{d}V,$$

where the J_is and ω_is are the components of $\boldsymbol{J}$ and $\boldsymbol{\omega}$ in the frame R. Therefore

$$\begin{pmatrix} J_1 \\ J_2 \\ J_3 \end{pmatrix} = \mathcal{J} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix}. \tag{5.3}$$

Once we have evaluated the appropriate volume integrals to find the centre of mass and the entries in the inertia matrix, we can use (5.3) to determine $\boldsymbol{J}$ and hence write T in terms of $\boldsymbol{u}$ and $\boldsymbol{\omega}$. A good choice of rest frame can simplify the calculation. In order to exploit the freedom in the choice of frame, we make use of the following propositions, which determine the relationship between the inertia matrices of the body in different rest frames.

Proposition 5.3 (Parallel Axes Theorem)

Let O' be the centre of mass and let $R = (O, \mathcal{T})$ and $R' = (O', \mathcal{T})$ be two rest frames with parallel coordinate axes, R' having origin at the centre of mass. Let $\mathcal{J}_{ij}$ and $\mathcal{J}'_{ij}$ denote the entries in the corresponding matrices $\mathcal{J}$ and $\mathcal{J}'$. Then

$$\mathcal{J}_{ij} = \mathcal{J}'_{ij} + m(x^2\delta_{ij} - x_i x_j)$$

where $\boldsymbol{x}$ is the vector from O' to O, $x^2 = \boldsymbol{x}\cdot\boldsymbol{x}$ and the x_is are the $\mathcal{T}$-components of $\boldsymbol{x}$.

Proof

If $\boldsymbol{r}$ is the position vector from O of a point of the body, then its position vector from O' is $\boldsymbol{r}' = \boldsymbol{x} + \boldsymbol{r}$. Because O' is the centre of mass

$$\int \rho r_i' \, \mathrm{d}V = 0.$$

Therefore

$$\int \rho(r_i' - x_i)(r_j' - x_j) \, \mathrm{d}V = m x_i x_j + \int \rho r_i' r_j' \, \mathrm{d}V.$$

By putting $i = j$ and summing, we also have

$$\int \rho r^2 \, \mathrm{d}V = m x^2 + \int \rho r'^2 \, \mathrm{d}V.$$

The proposition follows. □

In matrix notation, the proposition is

$$\mathcal{J} = \mathcal{J}' + + m\big(X^{\mathrm{t}} X I - X X^{\mathrm{t}}\big),$$

where X is the column vector with entries x_1, x_2, x_3 and I is the identity matrix. Note that $X^{\mathrm{t}} X = \boldsymbol{x} \cdot \boldsymbol{x}$ and that $X X^{\mathrm{t}}$ is the 3×3 matrix with entries $x_i x_j$.

The parallel axes theorem states that the inertia matrix in a rest frame R is the sum of two terms: the inertia matrix of a single particle of mass m at the centre of mass and the inertia matrix of the body in a frame with axes parallel to those of R and origin at the centre of mass. For the diagonal entries, the theorem reduces to the parallel axis theorem for moments of inertia.

Proposition 5.4 (The Tensor Property)

Let $R = (O, \mathcal{T})$ and $R' = (O, \mathcal{T}')$ be two rest frames with the same origin and let H be the transition matrix from $\mathcal{T}'$ to $\mathcal{T}$. Let $\mathcal{J}$ and $\mathcal{J}'$ denote the corresponding inertia matrices. Then

$$\mathcal{J} = H \mathcal{J}' H^{\mathrm{t}}.$$

Proof

Let $\boldsymbol{r}$ denote the position vector from O of a point of the body, and let X and X' denote the two column vectors with entries equal to the components of $\boldsymbol{r}$ in, respectively, $\mathcal{T}$ and $\mathcal{T}'$. Then

$$\mathcal{J} = \int \rho(X^{\mathrm{t}} X I - X X^{\mathrm{t}}) \, \mathrm{d}V, \qquad \mathcal{J}' = \int \rho(X'^{\mathrm{t}} X' I - X' X'^{\mathrm{t}}) \, \mathrm{d}V,$$

with the integrals of the matrix-valued integrands taken entry by entry. It follows from $X = HX'$ that

$$X^{\mathrm{t}} X I - X X^{\mathrm{t}} = H\big(X'^{\mathrm{t}} X' I - X' X'^{\mathrm{t}}\big) H^{\mathrm{t}}$$

because $HH^{\mathrm{t}} = I$ and $X^{\mathrm{t}} X = \boldsymbol{r} \cdot \boldsymbol{r} = X'^{\mathrm{t}} X'$. Therefore $\mathcal{J} = H\mathcal{J}' H^{\mathrm{t}}$, because the entries in H are constant over the body.

□

We can also write $\mathcal{J}' = H^{\mathrm{t}} \mathcal{J} H$ because $H^{\mathrm{t}} = H^{-1}$. For any symmetric matrix $\mathcal{J}$, there exists an orthogonal matrix H such that $H^{\mathrm{t}} \mathcal{J} H$ is diagonal. Given O, therefore, it is possible to choose the orthonormal triad $\mathcal{T}'$ so that,

$$\mathcal{J}' = \begin{pmatrix} \mathcal{J}'_{11} & 0 & 0 \\ 0 & \mathcal{J}'_{22} & 0 \\ 0 & 0 & \mathcal{J}'_{33} \end{pmatrix}.$$

Definition 5.5

Let $\mathcal{J}$ be the inertia matrix in a rest frame $R = (O, \mathcal{T})$. A *principal axis* at O is a line through O in the direction of an eigenvector of $\mathcal{J}$; that is, in the direction of vector $\boldsymbol{x}$ such that

$$\mathcal{J} X = AX \tag{5.4}$$

for some $A \in \mathbb{R}$, where X is the column vector formed from the components of $\boldsymbol{x}$. The eigenvalue A is the corresponding *principal moment of inertia*.

The principal axes and principal moments of inertia depend on O, but not on the choice of triad. They are characterized, independently of $\mathcal{T}$, as the directions of $\boldsymbol{\omega}$ for which $\boldsymbol{J} = A\boldsymbol{\omega}$, that is, as the directions for which the angular momentum relative to O is aligned with the angular velocity.

Some familiar results from linear algebra translate into statements about principal axes and principal moments. For example, two eigenvectors of a symmetric matrix corresponding to distinct eigenvalues are necessarily orthogonal. This becomes the statement that the principal axes at O of distinct principal moments of inertia are orthogonal. If two vectors are eigenvectors of a matrix with the same eigenvalue, then any linear combination of the two is also an eigenvector with this eigenvalue. Thus if two lines through O are principal axes corresponding to the same principal moment of inertia A, then every line through O in the plane that they span is also a principal axis with principal moment of inertia A.

If $\mathcal{T}$ is chosen so that $\mathcal{J}$ is diagonal, then the coordinate axes are principal axes and the diagonal entries in $\mathcal{J}$ are the corresponding principal moments

of inertia. Conversely, if one takes three orthogonal principal axes at O as coordinate axes, then the inertia matrix is diagonal, with the principal moments of inertia as diagonal entries.

Example 5.6

†The moment of inertia of a uniform spherical shell of mass m and radius a about any axis through its centre is $\frac{2}{3}ma^2$. Hence the inertia matrix at the centre is $\frac{2}{3}ma^2I$, irrespective of the choice of triad.

The upper and lower hemispheres contribute equally to the integrals for the moments and products of inertia. Therefore the inertia matrix of a uniform hemispherical shell of mass m and radius a at its centre O is also $\frac{2}{3}ma^2I$, that is, half the matrix of a spherical shell of mass $2m$.

We use the parallel axes theorem to find the inertia matrix at a point P on the rim of the hemisphere. Because neither O nor P is the centre of mass, we must do this in two steps.

Let $\mathcal{T} = (\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k})$, where $\boldsymbol{i}$ and $\boldsymbol{j}$ are parallel to the base, with $\boldsymbol{i}$ pointing from O to P, and $\boldsymbol{k}$ is normal to the base (Figure 5.1).

The position vector from O of the centre of mass C is $-\frac{1}{2}a\boldsymbol{k}$, by Archimedes' theorem relating the area elements of a cylinder and an inscribed sphere. Therefore if we take $R' = (C, \mathcal{T})$ to be the frame with origin at the centre of mass in the parallel axes theorem, then

$$\mathcal{J}' = \tfrac{2}{3}ma^2I - \tfrac{1}{4}ma^2\left[I - \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right].$$

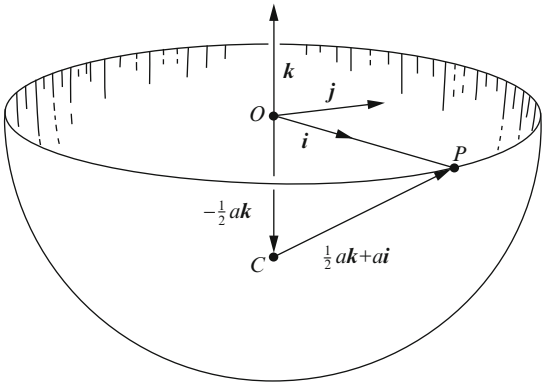

Figure 5.1

The position vector of the point P from C is $a\boldsymbol{i} + \frac{1}{2}a\boldsymbol{k}$. So if we take $R = (P, \mathcal{T})$ in a second application of the theorem, we obtain

$$\begin{aligned}\mathcal{J} &= \mathcal{J}' + \tfrac{1}{4}ma^2\left[5I - \begin{pmatrix} 4 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & 1 \end{pmatrix}\right] \\ &= \tfrac{1}{6}ma^2\begin{pmatrix} 4 & 0 & -3 \\ 0 & 10 & 0 \\ -3 & 0 & 10 \end{pmatrix}\end{aligned}$$

for the inertia matrix at P. The principal moments of inertia at P are $\frac{1}{6}ma^2 s$, where s is a root of the cubic

$$\begin{vmatrix} 4-s & 0 & -3 \\ 0 & 10-s & 0 \\ -3 & 0 & 10-s \end{vmatrix}.$$

The roots are

$$\tfrac{5}{3}ma^2, \quad \text{and} \quad \tfrac{1}{6}ma^2(7 \pm 3\sqrt{2});$$

and the principal axes are in the directions of the corresponding eigenvectors

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} -1 \\ 0 \\ 1+\sqrt{2} \end{pmatrix}, \quad \begin{pmatrix} -1 \\ 0 \\ 1-\sqrt{2} \end{pmatrix}$$

Note that the parallel axes theorem cannot be used to go directly from O to P because in each application of the theorem, one of the frames must have its origin at the centre of mass.

Suppose that the rest frame $R = (O, \mathcal{T})$ has its axes aligned with principal axes at O. Then $F = G = H = 0$ and

$$\tfrac{1}{2}\boldsymbol{J} \cdot \boldsymbol{\omega} = \tfrac{1}{2}(A\omega_1^2 + B\omega_2^2 + C\omega_3^2).$$

The expression (5.2) for T simplifies further in either of two cases.

1. The origin O of the rest frame is the centre of mass. Then $\boldsymbol{c} = 0$ and

$$T = \tfrac{1}{2}m\boldsymbol{u} \cdot \boldsymbol{u} + \tfrac{1}{2}(A\omega_1^2 + B\omega_2^2 + C\omega_3^2), \tag{5.5}$$

   with the coordinate axes of the rest frame aligned with principal axes at the centre of mass. In this case T is the sum of two parts. The first involves only the translational degrees of freedom and the second involves only the rotational degrees of freedom.

2. The origin O is at rest relative to the inertial frame. This is the case in which the body is rotating about a fixed point. We then have $\boldsymbol{u} = 0$ and

$$T = \tfrac{1}{2}(A\omega_1^2 + B\omega_2^2 + C\omega_3^2).$$

EXERCISES

5.1. Let $R = (O, \mathcal{T})$ be a rest frame of a rigid body. Denote the entries in the corresponding inertia matrix by $\mathcal{J}_{ij}$. Show that the moment of inertia about an axis through O in the direction of a unit vector with components x_i is $\mathcal{J}_{ij}x_ix_j$.

5.2. Show that the inertia matrix at the centre of a uniform solid cube with mass m and edges of length $2a$ is $\frac{2}{3}ma^2I$. Find the principal axes and principal moments of inertia at a vertex.

5.3. †Show that the principal moments of inertia at the centre of mass of a uniform solid circular cylinder, radius a, height $2h$, and mass m, are $\frac{1}{2}ma^2$ and $\frac{1}{12}m(4h^2 + 3a^2)$ (repeated). Find the principal axes and principal moments of inertia at a point distance D from the centre of mass in the plane through the centre of mass perpendicular to the axis of the cylinder.

5.4. Show that the kinetic energy of a uniform rod of mass m is

$$T = \tfrac{1}{6}m(\boldsymbol{u}\cdot\boldsymbol{u} + \boldsymbol{u}\cdot\boldsymbol{v} + \boldsymbol{v}\cdot\boldsymbol{v})$$

where $\boldsymbol{u}$ and $\boldsymbol{v}$ are the velocities of the two ends.

5.5. †Show that if a line passing through the centre of mass of a rigid body is a principal axis at one point of the line, then it is a principal axis at every point of the line.

5.6. †A rigid body is said to have *inertial symmetry* at a point P if the principal moments of inertia at P are all equal. Show that if a body has inertial symmetry at a point other than the centre of mass, then the principal moments of inertia at the centre of mass cannot all be distinct.

A uniform solid right circular cone has height h and base of radius a. For what values of h/a does the cone have inertial symmetry at its vertex?

5.7. Show that the inertia matrix at the centre of any uniform Platonic solid is a multiple of the identity matrix.

5.4 Linear and Angular Momentum

The principles of linear and angular momentum determine the motion of a rigid body under that action of external forces. Let $R = (O, \mathcal{T})$ be a rest frame of the rigid body, with its origin O at the centre of mass of the body. Consider

the motion relative to some inertial frame. The total linear momentum relative to the inertial frame is $\boldsymbol{p} = m\boldsymbol{u}$, where $\boldsymbol{u}$ is the velocity of O relative to the inertial frame. The total angular momentum about the centre of mass is

$$\boldsymbol{J}_O = \int \rho \boldsymbol{r} \wedge \dot{\boldsymbol{r}} \, \mathrm{d}\tau = \int \rho \boldsymbol{r} \wedge (\boldsymbol{\omega} \wedge \boldsymbol{r}) \, \mathrm{d}\tau, \tag{5.6}$$

where $\boldsymbol{\omega}$ is the angular velocity of the body, $\boldsymbol{r}$ is the position vector of an element of the body from O and the dot is the time derivative relative to the inertial frame. As in Section 5.2,

$$\begin{pmatrix} J_1 \\ J_2 \\ J_3 \end{pmatrix} = \mathcal{J} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix}.$$

where the J_is and ω_is are the respective $\mathcal{T}$-components of $\boldsymbol{J}_O$ and $\boldsymbol{\omega}$, and $\mathcal{J}$ is the inertia matrix at the centre of mass in the frame R. We can write down the components of $\boldsymbol{J}_O$, therefore, once we know the angular velocity $\boldsymbol{\omega}$ and the moments and products of inertia at the centre of mass.

By the principles of linear and angular momentum,

$$\dot{\boldsymbol{p}} = \sum \boldsymbol{E} \quad \text{and} \quad \dot{\boldsymbol{J}}_O = \sum \boldsymbol{r} \wedge \boldsymbol{E}. \tag{5.7}$$

In the first equation, $\sum \boldsymbol{E}$ is the sum of the external forces. In the second, the sum is over the points at which the external forces act.

In principle, the two vector equations (5.7) determine the evolution of the six degrees of freedom. The second equation can, however, be written in a more convenient form. Take the axes of R to be along principal axes at O. Then

$$J_1 = A\omega_1, \quad J_2 = B\omega_2, \quad J_3 = C\omega_3, \tag{5.8}$$

where A, B and C are the principal moments of inertia at O. It follows that the time derivative $\mathrm{D}\boldsymbol{J}_O$ of $\boldsymbol{J}_O$ with respect to R, has components $A\dot{\omega}_1$, $B\dot{\omega}_2$, and $C\dot{\omega}_3$. By the Coriolis theorem,

$$\dot{\boldsymbol{J}}_O = \mathrm{D}\boldsymbol{J}_O + \boldsymbol{\omega} \wedge \boldsymbol{J}_O.$$

Therefore

$$\begin{aligned} A\dot{\omega}_1 + (C - B)\omega_2\omega_3 &= G_1 \\ B\dot{\omega}_2 + (A - C)\omega_3\omega_1 &= G_2 \\ C\dot{\omega}_3 + (B - A)\omega_1\omega_2 &= G_3 \end{aligned}$$

where G_1, G_2, and G_3 are the components of $\sum \boldsymbol{r} \wedge \boldsymbol{E}$ in the rest frame R. These are *Euler's equations*. They determine the time-dependence of the angular velocity and hence of the orientation of the body.

5.5 Rotation About a Fixed Point

We can also take moments about a fixed point P in an inertial frame, to obtain

$$\dot{\boldsymbol{J}}_P = \sum (\boldsymbol{r} + \boldsymbol{x}) \wedge \boldsymbol{E}, \tag{5.9}$$

where $\boldsymbol{x}$ is the vector from P to the centre of mass O and $\boldsymbol{J}_P$ is the angular momentum relative to P (p. 111). It is related to angular momentum about the centre of mass by

$$\boldsymbol{J}_P = \boldsymbol{J}_O + \boldsymbol{x} \wedge \boldsymbol{p}.$$

The right-hand side of (5.9) is the moment of the external forces about P.

Equation (5.9) does not contain any new information because it is an immediate consequence of (5.7). It is particularly useful, however, in the case of a body rotating about a fixed point. In other words, when there is a point P of the body which is at rest relative to the inertial frame at all times. Then the velocity relative to the inertial frame of an element of the body with position vector $\boldsymbol{r}$ from P is $\boldsymbol{\omega} \wedge \boldsymbol{r}$. In this case,

$$\boldsymbol{J}_P = \int \rho \boldsymbol{r} \wedge (\boldsymbol{\omega} \wedge \boldsymbol{r}) \, \mathrm{d}V \tag{5.10}$$

and the components of $\boldsymbol{J}_P$ in an orthonormal triad $\mathcal{T}$ fixed in the body are related to those of $\boldsymbol{\omega}$ by (5.3), where $\mathcal{J}$ is now the inertia matrix at P in the triad $\mathcal{T}$. When $\mathcal{T}$ is aligned with principal axes at P, we once again obtain Euler's equations for the $\mathcal{T}$-components of $\boldsymbol{\omega}$, only now A, B, and C are the principal moments of inertia at P and G_1, G_2, and G_3 are the $\mathcal{T}$-components of the total moment of the external forces about P.

It is important to remember that the relationship in (5.3) between the components of the angular momentum and the angular velocity is valid in general only for

1. Angular momentum about the centre of mass in the case of arbitrary motion
2. In the case of rotation about a fixed point of an inertial frame, for the angular momentum about the fixed point

Similarly, Euler's equations for the components of $\boldsymbol{\omega}$ hold only if the rest frame R either has its origin at the centre of mass, in the case of general motion, or, in the case of rotation about a fixed point, has its origin at the fixed point.

Example 5.7 (Free Rotation)

Consider a rigid body rotating about a fixed point P. Let $R = (P, \mathcal{T})$ be a rest frame with its axes aligned with the principal axes at P and suppose that $G_1 = G_2 = G_3 = 0$. Then Euler's equations reduce to

$$\begin{aligned} A\dot{\omega}_1 + (C-B)\omega_2\omega_3 &= 0 \\ B\dot{\omega}_2 + (A-C)\omega_3\omega_1 &= 0 \\ C\dot{\omega}_3 + (B-A)\omega_1\omega_2 &= 0 \end{aligned} \tag{5.11}$$

We assume that $A < B < C$.

By multiplying the three equations first by ω_1, ω_2, and ω_3 respectively and adding, and second by $A\omega_1$, $B\omega_2$, $C\omega_3$ and adding, we find that

$$A\omega_1^2 + B\omega_2^2 + C\omega_3^2 = 2T \tag{5.12}$$

$$A^2\omega_1^2 + B^2\omega_2^2 + C^2\omega_3^2 = J^2 \tag{5.13}$$

where T and J are constants: T is the kinetic energy and J is the magnitude of the angular momentum vector $\boldsymbol{J}_P$.

The *instantaneous axis* is the line through P in the direction of $\boldsymbol{\omega}$. By studying the way in which it moves relative to R, we can discover a great deal about the dynamical behaviour of the body.

By writing $\boldsymbol{\omega} = \lambda \boldsymbol{r}$ and eliminating λ between (5.12) and (5.13), we see that the instantaneous axis is a generator of a fixed quadric cone in the body. In the rest frame R, the cone has equation

$$\alpha x^2 + \beta y^2 + \gamma z^2 = 0$$

where

$$\alpha = \frac{A}{2T} - \frac{A^2}{J^2}, \quad \beta = \frac{B}{2T} - \frac{B^2}{J^2}, \quad \gamma = \frac{C}{2T} - \frac{C^2}{J^2}.$$

With A, B, and C fixed, the shape of the cone depends on the value of $J^2/2T$, which must lie between A and C. There are three cases.

(1) $J^2 = 2BT$. Then $\beta = 0$, $\alpha > 0$, and $\gamma < 0$. The cone degenerates into a pair of planes that intersect along the y-axis.

(2) $2CT \geq J^2 > 2BT$. Then $\alpha > 0$, $\beta > 0$, and $\gamma \leq 0$. The intersection of the cone with the plane $z = 1$ is one of the family of ellipses in Figure 5.2.

(3) $2BT > J^2 \geq 2AT$. Then $\alpha \geq 0$, $\beta < 0$, and $\gamma < 0$. The intersection of the cone with the plane $x = 1$ is one of the family of ellipses in Figure 5.3.

In the figures, the limiting case of the parallel lines arises as $\beta \to 0$. The arrows on the ellipses indicate the movement of the instantaneous axis relative to the body. The directions are found by examining the signs in Euler's equations.

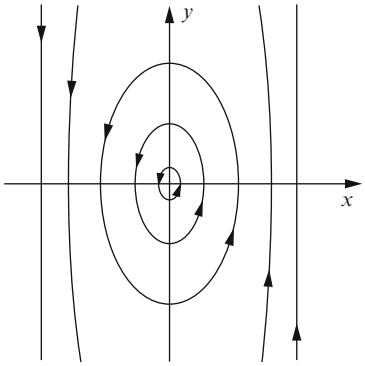

Figure 5.2 Case 2. The intersection with the plane $z = 1$ is one of the ellipses $\alpha x^2 + \beta y^2 = -\gamma$.

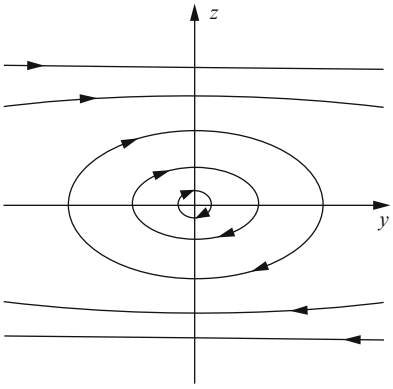

Figure 5.3 Case 3. The intersection with the plane $x = 1$ is one of the ellipses $-\beta y^2 - \gamma z^2 = \alpha$.

By combining Figures 5.2 and 5.3, we obtain Figure 5.4, which shows the quadric cone and the direction of movement of the instantaneous axis for various values of $J^2/2T$. In the limiting case $A = B$, the plane pair coalesces into the xy-plane, the quadric cones become circular and the picture reduces to Figure 5.5.

Motion with constant angular velocity is possible about any of the three coordinate axes of the rest frame, with $\boldsymbol{\omega}$ parallel to $\boldsymbol{J}_P$ and the instantaneous axis fixed both in the body and in the inertial frame. Such motion is stable if $\omega_2 = \omega_3 = 0$ or if $\omega_1 = \omega_2 = 0$ because, if the body is disturbed slightly when it is rotating about either the x-axis or the z-axis, then the instantaneous axis must remain on a narrow cone surrounding the original axis of rotation.

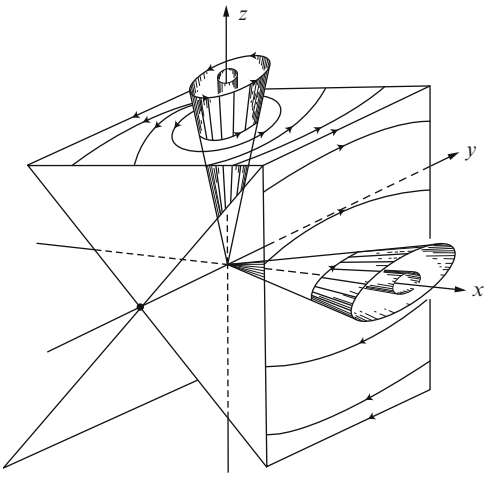

Figure 5.4

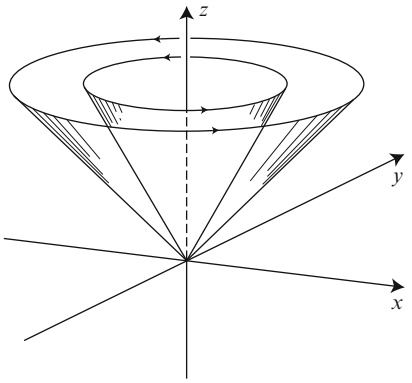

Figure 5.5

However, the motion is unstable when $\omega_1 = \omega_3 = 0$.

The case $J^2 = 2BT$ is worth investigating further. By substituting for ω_1 and ω_3 from (5.12) and (5.13) into the second equation in (5.11),

$$B^2\dot{\omega}_2^2 = \frac{(2T - B\omega_2^2)(C - B)(B - A)}{AC}. \tag{5.14}$$

This can be solved analytically, but it is possible to deduce the qualitative behaviour by sketching the trajectories in the $\omega_2, \dot{\omega}_2$-plane (Figure 5.6) without finding the explicit solutions. The trajectories are the curves given by (5.14) for the different values of T. Unstable equilibrium is possible when $\omega_2 = \pm\sqrt{2T/B}$ (rotation about the y-axis). But if the motion is disturbed, with the relationship

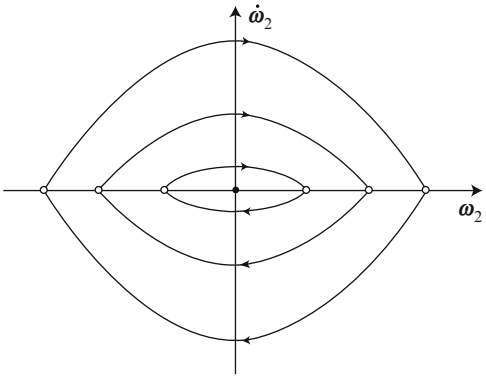

Figure 5.6

between T and J^2 maintained, then the system follows one of the curves in Figure 5.6 and ω_2 tends asymptotically to the negative of its original value. The instantaneous axis sweeps out one or other of the $J^2 = 2BT$ pair of planes in the body, eventually coming back into coincidence with the y-axis, but with ω_2 having changed sign. During the process the angular momentum vector remains fixed relative to the inertial frame. Thus the effect of the disturbance is to cause the body to flip over. If the intersection points of the y-axis of R with the surface are marked red and blue, then, seen from the inertial frame, the body ends up spinning in exactly the same way as initially, but with the red and blue points interchanged.

This somewhat implausible behaviour depends, of course, on the exact relationship between T and J^2 being maintained when the original rotation about the y-axis is disturbed.

Example 5.8 (Motion of a Top, Vector Treatment)

A (symmetric) top is a rigid body with axial symmetry about a line through its centre of mass.

Suppose that P is a point on the axis of symmetry at a distance a from the centre of mass. Consider the motion in which P is fixed relative to an inertial frame and the top is rotating about P without friction under the influence of gravity. Let $\boldsymbol{e}$ denote the unit vector along the axis of symmetry pointing from P towards the centre of mass (Figure 5.8). Then

$$\dot{\boldsymbol{e}} = \boldsymbol{\omega} \wedge \boldsymbol{e}$$

where $\boldsymbol{\omega}$ is the angular velocity of the top and the dot denotes the time derivative with respect to the inertial frame. By taking the vector product with $\boldsymbol{e}$,

$$\boldsymbol{\omega} = \boldsymbol{e} \wedge \dot{\boldsymbol{e}} + n\boldsymbol{e}$$

where $n = \boldsymbol{\omega} \cdot \boldsymbol{e}$.

The principal axes at P must be the axis of symmetry itself and all the lines through P orthogonal to $\boldsymbol{e}$. Denote the corresponding principal moments of inertia by C, for the moment about the axis of symmetry, and by A, for the moment about any orthogonal axis.

Because $\boldsymbol{e}$ is along the axis of symmetry and $\boldsymbol{e} \wedge \dot{\boldsymbol{e}}$ is orthogonal to the axis of symmetry,

$$\boldsymbol{J}_P = A\boldsymbol{e} \wedge \dot{\boldsymbol{e}} + Cn\boldsymbol{e}.$$

The external forces are the reaction force at P, which has no moment about P, and the gravitational forces on the various elements of the body. These have total moment

$$-\int \rho g \boldsymbol{r} \wedge \boldsymbol{k} \, \mathrm{d}V$$

about P, where ρ is the density, g is the acceleration due to gravity, $\boldsymbol{k}$ is a unit vector in the direction of the upward vertical and the integral is over the body. But g and $\boldsymbol{k}$ are constant over the body, so the total moment is equal to

$$-\left(\int \rho g \boldsymbol{r} \, \mathrm{d}V\right) \wedge g\boldsymbol{k} = -am\boldsymbol{e} \wedge g\boldsymbol{k},$$

where m is the mass, by using the definition of the centre of mass. Thus gravity acts as a single force $-mg\boldsymbol{k}$ through the centre of mass and therefore

$$\dot{\boldsymbol{J}}_P = A\boldsymbol{e} \wedge \ddot{\boldsymbol{e}} + C\dot{n}\boldsymbol{e} + Cn\dot{\boldsymbol{e}} = -mga\boldsymbol{e} \wedge \boldsymbol{k}. \tag{5.15}$$

By taking the scalar product with $\boldsymbol{e}$, we deduce that $\dot{n} = 0$.

We do not investigate the motion in more detail at this stage, except to remark that (5.15) has solutions for which

$$\dot{\boldsymbol{e}} = \Omega \boldsymbol{k} \wedge \boldsymbol{e}, \tag{5.16}$$

where Ω is constant. These correspond to *steady precession*, in which $\boldsymbol{e} \cdot \boldsymbol{k}$ is constant. The axis of symmetry makes a fixed angle with the vertical and rotates about the vertical with constant angular speed Ω.

On substituting for $\dot{\boldsymbol{e}}$ from (5.16) into (5.15), one finds that in steady precession, n, Ω, and the angle α between $\boldsymbol{e}$ and $\boldsymbol{k}$ are related by

$$A\Omega^2 \cos\alpha - Cn\Omega + mga = 0.$$

EXERCISES

5.8. Solve Equation (5.14) and verify the statements made about the behaviour of ω_2.

5.9. In Example 5.7, sketch and interpret the trajectories of the system in the $\omega_2, \dot{\omega}_2$-plane for different values of J^2 and T.

5.10. The surface of a rigid body is an ellipsoid with equation

$$Ax^2 + By^2 + Cz^2 = k^2$$

in a rest frame with origin O at the centre of mass and axes aligned with the principal axes at the centre of mass; A, B, and C are the principal moments of inertia at O and k is a constant.

Show that if the body rotates freely about O, with O fixed relative to an inertial frame, then the two tangent planes to the surface at the intersection points between the surface and the instantaneous axis are fixed relative to the inertial frame. Deduce that the body moves as if it were rolling between these two planes.

5.11. †In Example 5.8, let $\boldsymbol{N} = \boldsymbol{e} \wedge \dot{\boldsymbol{e}}$ and let θ be the angle between $\boldsymbol{e}$ and $\boldsymbol{k}$. Show that

$$A\boldsymbol{N} \cdot \boldsymbol{N} + 2mga\cos\theta \quad \text{and} \quad A\boldsymbol{N} \cdot \boldsymbol{k} + Cn\cos\theta$$

are constant.

Show that if initially $\theta = \beta$ and $\boldsymbol{N} = 0$, then during the subsequent motion

$$2Amga/C^2n^2 \geq \cos\beta - \cos\theta \geq 0.$$

5.12. †A uniform solid sphere of radius a rolls without slipping inside a fixed sphere of radius $2a$. Show that if $\boldsymbol{e}$ is the unit vector pointing from the centre of the larger sphere towards the centre of the smaller sphere, then

$$7a\boldsymbol{e} \wedge \ddot{\boldsymbol{e}} - 2an\dot{\boldsymbol{e}} + 5g\boldsymbol{e} \wedge \boldsymbol{k} = 0$$

where n is constant and $\boldsymbol{k}$ is a unit vector in the direction of the upward vertical. The dot denotes the time derivative with respect to fixed axes.

5.13. †A hollow right circular cylinder is fixed with its axis vertical. A uniform solid sphere rolls without slipping on the inside surface of the cylinder. Show that when the centre of the sphere does not move on a vertical line, the height of the centre of the sphere performs

simple harmonic motion, and that between oscillations the plane containing the axis of the cylinder and the centre of the sphere turns through an angle $\pi\sqrt{14}$.

5.6 Lagrange's Equations

Euler's equations determine the time-dependence of the angular velocity $\boldsymbol{\omega}$ of a rigid body, but, except in special cases, they are not a good starting point for determining the evolution of the configuration. For that, we need coordinates in the configuration space and the corresponding Lagrangian equations. We need six coordinates in all, three to tie down the translational degrees of freedom and three to specify the orientation of the body.

A natural choice when considering the general motion of the body is to use the Cartesian coordinates of the centre of mass together with the Euler angles relative to some inertial frame of a triad fixed in the body. Let $\tilde{R} = (\tilde{O}, \tilde{\mathcal{T}})$ denote the inertial frame and let $R = (O, \mathcal{T})$ be a rest frame of a rigid body, with origin O at the centre of mass of the body and with axes aligned with the principal axes at O. Let x, y, and z denote the coordinates of O in $\tilde{R}$, and let θ, φ, and ψ denote the Euler angles of $\mathcal{T}$ relative to $\tilde{\mathcal{T}}$. Then x, y, z, θ, φ, and ψ are coordinates on the configuration space of the body. By combining (5.5) with (1.24), the kinetic energy of the body relative to $\tilde{R}$ is

$$\begin{aligned} T \;=\; & \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + \tfrac{1}{2}A(\dot{\theta}\sin\psi - \dot{\varphi}\sin\theta\cos\psi)^2 \\ & + \tfrac{1}{2}B(\dot{\theta}\cos\psi + \dot{\varphi}\sin\theta\sin\psi)^2 + \tfrac{1}{2}C(\dot{\psi} + \dot{\varphi}\cos\theta)^2, \end{aligned}$$

where A, B, and C are the principal moments of inertia at O.

For rotations about a point P which is fixed both in R and in the body, we instead take $R = (P, \mathcal{T})$, with $\mathcal{T}$ now aligned with the principal axes at P. Then

$$\begin{aligned} T = & \tfrac{1}{2}A(\dot{\theta}\sin\psi - \dot{\varphi}\sin\theta\cos\psi)^2 \\ & + \tfrac{1}{2}B(\dot{\theta}\cos\psi + \dot{\varphi}\sin\theta\sin\psi)^2 + \tfrac{1}{2}C(\dot{\psi} + \dot{\varphi}\cos\theta)^2, \end{aligned} \tag{5.17}$$

where A, B, and C are now the principal moments of inertia at P.

Both coordinate systems are singular when $\theta = 0$, in the same way that spherical polar coordinates are singular on the polar axis. The problem arises when the x, y-planes of the two frames become parallel, a configuration which goes by the name of 'gimbal lock'. One must use a different coordinate system to study the motion near such a configuration. One could, for example, permute the vectors in $\mathcal{T}$ before introducing the Euler angles.

Example 5.9 (Motion of a Top: Lagrange's Equations)

We now look in more detail at the motion of the top in Example 5.8. Choose $\mathcal{T} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ so that the $\boldsymbol{e}_i$s are aligned with the principal axes at P, with $\boldsymbol{e} = \boldsymbol{e}_3$ along the axis of symmetry. Choose $(\tilde{\boldsymbol{e}}_1, \tilde{\boldsymbol{e}}_2, \tilde{\boldsymbol{e}}_3)$ so that $\boldsymbol{k} = \tilde{\boldsymbol{e}}_3$. Then the Lagrangian is

$$L = \tfrac{1}{2}A(\dot{\varphi}^2 \sin^2\theta + \dot{\theta}^2) + \tfrac{1}{2}C(\dot{\psi} + \dot{\varphi}\cos\theta)^2 - mga\cos\theta,$$

by taking $A = B$ in (5.17).

Because ψ and φ are cyclic

$$\begin{aligned} \dot{\psi} + \dot{\varphi}\cos\theta &= n \\ A\dot{\varphi}\sin^2\theta + Cn\cos\theta &= j \end{aligned}$$

where n and j are constants, and because $\partial L/\partial t = 0$,

$$A\dot{\theta}^2 + A\dot{\varphi}^2\sin^2\theta + Cn^2 + 2mga\cos\theta = 2E$$

where E is constant. The constants n, j, and E are, respectively, the $\boldsymbol{e}_3$-component of $\boldsymbol{\omega}$, the angular momentum about the vertical axis through P, and the total energy.

By writing $u = \cos\theta$ and by rearranging the equations

$$\dot{\varphi} = \frac{j - Cnu}{A(1-u^2)} \qquad \text{and} \qquad A\dot{u}^2 = F(u)$$

where

$$F(u) = (2E - Cn^2 - 2mgau)(1-u^2) - \frac{(j-Cnu)^2}{A}.$$

The angles θ and φ are the spherical polar angles of $\boldsymbol{e}$. That is, θ is the angle between the axis of the top and the upward vertical and $\dot{\varphi}$ is the angular speed with which the axis rotates about the vertical.

Suppose that the top is set in motion with $\theta = \cos^{-1}(u_1)$ and $\dot{\theta} = 0$. We shall keep j and n fixed, with $n > 0$ and $0 < j/Cn < 1$, and investigate what happens for various values of u_1 by looking at the phase portrait in the $u, \dot{u}$-plane. The trajectories in the $u, \dot{u}$-plane are the curves $A\dot{u}^2 = F(u)$, the constant E in the definition of F being determined by

$$0 = (2E - Cn^2 - 2mgau_1)(1-u_1^2) - \frac{(j-Cnu_1)^2}{A}.$$

The graph of $F(u)$ for different values of u_1 and the phase portrait are shown in Figures 5.7 and 5.8. The dotted vertical line is $u = j/Cn$. To the left of the line (that is for $-1 < u < j/Cn$), $\dot{\varphi}$ is positive and to the right ($j/Cn < u < 1$), $\dot{\varphi}$ is negative. Remember that u must lie between -1 and 1. As the top moves,

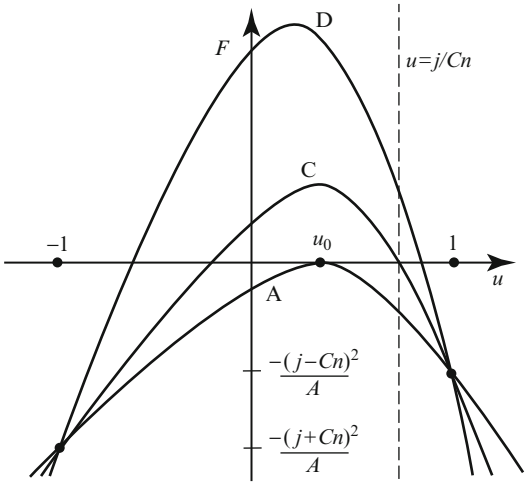

Figure 5.7

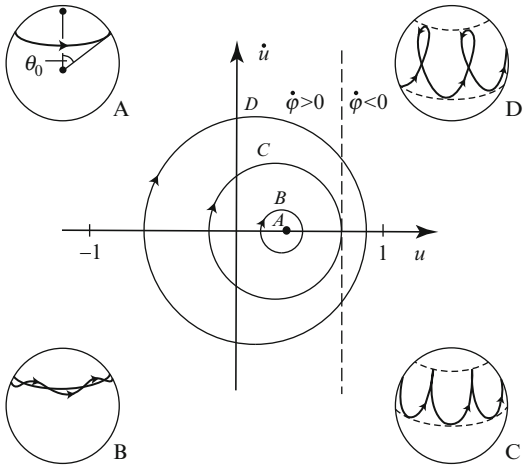

Figure 5.8

its axis traces out a curve on the unit sphere centred at P. The rough form of the curve can be seen from the phase portrait (Figure 5.8).

There is a critical value u_0 of u_1 for which $u_1 = u_0$ is a root of both $F(u)$ and $F'(u)$. Here the graph of $F(u)$ touches the u-axis at $u = u_0$ and the trajectory is $u = u_0 =$ constant. The axis precesses steadily about the vertical with θ and φ constant and $\dot{\varphi} > 0$. It traces out a horizontal circle on the sphere on which $\theta = \theta_0 = \cos^{-1}(u_0)$ (Figure 5.8, top left).

For u_1 such that

$$u_0 < u_1 < j/Cn,$$

u oscillates between the two roots of $F(u)$ on either side of u_0, but $\dot{\varphi}$ is always positive. The axis rotates anticlockwise about the vertical (seen from above), but its angle with the vertical oscillates between the two corresponding values of θ (B, bottom left). The oscillatory behaviour of θ is called *nutation*.

If $u_1 = j/Cn$, then $\dot{\varphi} = 0$ when θ reaches its minimum (C, bottom right). If

$$j/Cn < u_1 < 1,$$

then $\dot{\varphi}$ is negative during part of the motion and the curve on the sphere loops back on itself (D, top right).

Example 5.10

†A thin circular disc of radius a moves on a smooth horizontal plane. The disc makes contact with the plane at a point and can slip freely. Initially, the centre is at rest and the disc is spinning with angular speed n about its axis, which is at an angle α to the vertical. The problem is to show that the spin about the axis, that is, the component of the angular velocity along the axis, remains constant and that during the subsequent motion,

$$a\dot{\theta}^2(1 + 4\cos^2\theta) + 4an^2(\cos\alpha - \cos\theta)^2\operatorname{cosec}^2\theta + 8g(\sin\theta - \sin\alpha) = 0,$$

where θ is the angle that the axis makes with the vertical.

Because there are no horizontal forces, the horizontal coordinates of the centre of mass remain constant. In effect, the system has three degrees of freedom. Let $\tilde{\mathcal{T}} = (\tilde{\boldsymbol{e}}_1, \tilde{\boldsymbol{e}}_2, \tilde{\boldsymbol{e}}_3)$ be a triad fixed relative to the plane, with $\tilde{\boldsymbol{e}}_3$ normal to the plane, and let $\mathcal{T} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ be a triad fixed relative to the disc, with $\boldsymbol{e}_3$ normal to the disc. We use the Euler angles of $\mathcal{T}$ relative to $\tilde{\mathcal{T}}$ as coordinates (Figure 5.9).

The angular velocity of the disc is

$$\dot{\theta}\boldsymbol{j} + \dot{\psi}\boldsymbol{e}_3 + \dot{\varphi}\tilde{\boldsymbol{e}}_3.$$

with $\boldsymbol{j}$ as in (1.13). The height of the centre above the plane is $a\sin\theta$ and the velocity of the centre is $a\dot{\theta}\cos\theta\,\tilde{\boldsymbol{e}}_3$.

The principal moments of inertia of the disc at its centre are $\frac{1}{2}ma^2$ about any diameter and $\frac{1}{4}ma^2$ about the axis through the centre normal to the plane of the disc. Therefore the Lagrangian is

$$L = \tfrac{1}{8}ma^2\big(4\dot{\theta}^2\cos^2\theta + \dot{\theta}^2 + \dot{\varphi}^2\sin^2\theta + 2(\dot{\psi} + \dot{\varphi}\cos\theta)^2\big) - mga\sin\theta.$$

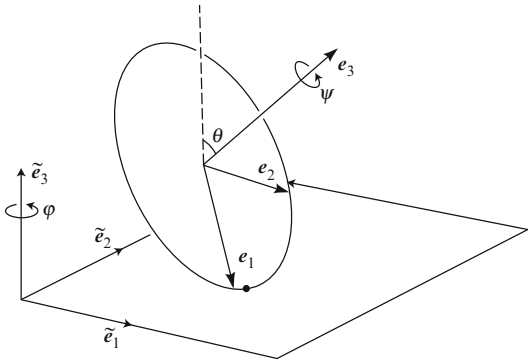

Figure 5.9

We see immediately that ψ and φ are cyclic coordinates, and therefore

$$\begin{aligned} \dot{\psi} + \dot{\varphi}\cos\theta &= \text{constant} \\ \tfrac{1}{4}\dot{\varphi}\sin^2\theta + \tfrac{1}{2}(\dot{\psi} + \dot{\varphi}\cos\theta)\cos\theta &= \text{constant}. \end{aligned}$$

Now $\dot{\psi}+\dot{\varphi}\cos\theta = \boldsymbol{e}_3\cdot\boldsymbol{\omega}$ is the spin of the disc about its axis.[1] The first equation states that the spin is constant, as required. Initially, $\theta = \alpha$, $\dot{\varphi} = \dot{\theta} = 0$, and $\boldsymbol{e}_3 \cdot \boldsymbol{\omega} = n$. Therefore

$$\dot{\psi} + \dot{\varphi}\cos\theta = n \quad \text{and} \quad \dot{\varphi}\sin^2\theta = 2n\cos\alpha - 2n\cos\theta.$$

Because $\partial L/\partial t = 0$, the total energy

$$E = \tfrac{1}{8}ma^2\big(\dot{\theta}^2(1 + 4\cos^2\theta) + 4(\cos\theta - \cos\alpha)^2 n^2 \operatorname{cosec}^2\theta + 2n^2\big) + mga\sin\theta$$

is also constant, by Proposition 4.11. Initially $E = \frac{1}{4}ma^2n^2 + mga\sin\alpha$. Hence the stated result.

Example 5.11

†Two uniform circular discs, each of radius a and mass m, are joined by a light rod AB of length $4a$ connected to the discs at their centres A, B in such a way that the planes of the discs are perpendicular to the rod and the discs can turn freely about the rod. The system is placed on a horizontal table which is rough enough to prevent slipping and which is forced to rotate with variable angular velocity Ω about a vertical axis passing through the centre O of AB. Initially the system is at rest. The problem is to find the angular velocity of the discs in terms of Ω.

[1] A common mistake in this type of problem is to equate the spin about the axis to $\dot{\psi}$, forgetting the contribution from $\dot{\varphi}$.

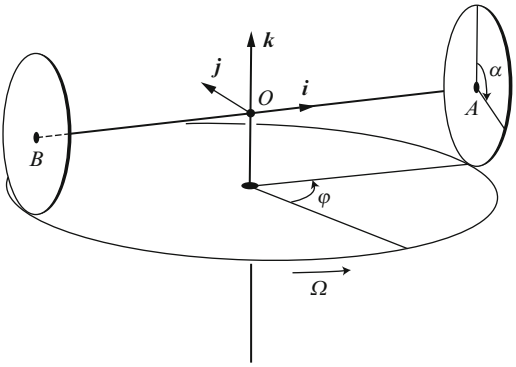

Figure 5.10

Let $\boldsymbol{i}$ be a unit vector along OA, let $\boldsymbol{k}$ denote the unit vertical vector, and let $\boldsymbol{j} = \boldsymbol{k} \wedge \boldsymbol{i}$. Let α denote the angle between a fixed line on the disc with centre A and the vertical, and let φ denote the angle between AB and a horizontal line fixed in space (see Figure 5.10).

The velocity of A is $2a\dot{\varphi}\boldsymbol{j}$ and the angular velocity of the disc with centre A is

$$\boldsymbol{\omega} = \dot{\alpha}\boldsymbol{i} + \dot{\varphi}\boldsymbol{k}.$$

Hence its kinetic energy is

$$T = \tfrac{1}{2}m(2a\dot{\varphi})^2 + \tfrac{1}{4}ma^2\dot{\alpha}^2 + \tfrac{1}{8}ma^2\dot{\varphi}^2.$$

Let F denote the $\boldsymbol{j}$-component of the friction force on the disc. Then

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{\varphi}}\right) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{17ma^2\dot{\varphi}}{4}\right) = 2aF$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{\alpha}}\right) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{ma^2\dot{\alpha}}{2}\right) = aF,$$

the generalized components of the friction force being found by considering in turn the small displacements (1) $\delta\varphi = \varepsilon$, $\delta\alpha = 0$ and (2) $\delta\varphi = 0$, $\delta\alpha = \varepsilon$. Therefore $17\dot{\varphi} = 4\dot{\alpha}$, because $\dot{\varphi}$ and $\dot{\alpha}$ are both zero initially.

The velocity of the point of the disc in contact with the table is

$$2a\dot{\varphi}\boldsymbol{j} + (\dot{\alpha}\boldsymbol{i} + \dot{\varphi}\boldsymbol{k}) \wedge (-a\boldsymbol{k}).$$

This must be equal to the velocity of the point of the table in contact with the disc, which is $2a\Omega\boldsymbol{j}$. Therefore

$$2a\dot{\varphi} + a\dot{\alpha} = 2a\Omega,$$

and so

$$\dot{\varphi} = \frac{8\Omega}{25} \quad \text{and} \quad \dot{\alpha} = \frac{34\Omega}{25}.$$

EXERCISES

5.14. In Example 5.9,

(a) With $0 < j/Cn < 1$, show that u_0 is the root of

$$G(u) = (j - Cnu)(Cn - ju) - mgaA(1 - u^2)^2$$

in the interval $(-1, 1)$. By sketching the graph of G and considering $G(-1)$, $G(j/Cn)$ and $G(1)$, show that $u_0 < j/Cn$. Show that if j and n are large, then $u_0 \sim j/Cn$. Describe the motion of the axis of a top which is set in motion with large n, but with $\dot{\theta} = \dot{\varphi} = 0$.

(b) For small oscillations about steady precession with $\dot{\varphi} = \Omega$, show by considering $F''(u_0)$ that

$$A\ddot{v} + \lambda^2 v = 0$$

where $v = u - u_0$ and $\lambda^2 = A\Omega^2 - 2mgau_0 + m^2g^2a^2/A\Omega^2$. What is the period of the oscillations when n is large?

(c) Repeat the analysis in the example for the case $n > 0$, $-1 < j/Cn < 0$.

5.15. †A thin uniform disc, mass M, radius a and centre C, has a thin uniform rod OC, mass m and length $a\sqrt{3}$, fixed to it at C, so that OC is orthogonal to the disc. The end O of the rod is fixed but freely pivoted at the centre O of a horizontal turntable, and the rim of the disc rests on the surface of the turntable. No slipping occurs.

The turntable is forced to rotate about the vertical axis through O with variable angular velocity Ω. Initially the system is at rest. Show that if P is the point of contact between the disc and the turntable, and φ is the angle between OP and a line fixed in the turntable, then

$$\dot{\varphi} = -\left(\frac{11M + 4m}{19M + 4m}\right)\Omega.$$

5.16. †A thin uniform rod of length $2a$ and mass m has a small light ring fixed at one end. The ring is threaded on a fixed vertical wire. Show that if z is the height of the centre of the rod, θ the angle the rod makes with the upward vertical, and φ the angle that the vertical

plane containing the rod makes with a fixed vertical plane, then the Lagrangian of the system is

$$L = \tfrac{1}{6}m[3\dot{z}^2 + a^2\dot{\theta}^2(1 + 3\cos^2\theta) + 4a^2\dot{\varphi}^2\sin^2\theta] - mgz.$$

Initially the rod makes an acute angle α with the vertical and its centre has velocity V perpendicular to the rod and the wire. Show that the angle the rod makes with the wire oscillates between α and $\pi - \alpha$ with period

$$\frac{a}{V}\int_{-\cos\alpha}^{\cos\alpha}\left[\frac{1+3u^2}{\cos^2\alpha - u^2}\right]^{1/2}\mathrm{d}u.$$

5.17. †A uniform hollow circular cylinder of mass m, radius a, rolls without slipping on a fixed rough horizontal plane. A similar cylinder of mass m and the same length, but radius $\frac{1}{2}a$, rolls without slipping inside the larger cylinder. The two cylinders are positioned so that their axes are parallel and their ends coincide. Consider the vertical plane through the centre of mass. Show that if θ is the angle between the downward vertical and the line in this plane joining the centre of mass of the larger cylinder to a point fixed on the rim of the larger cylinder, and if φ is the angle between the downward vertical and the line joining the centres of mass, then

$$2ma^2\dot{\theta}^2 + \tfrac{1}{4}ma^2\dot{\varphi}^2 - \tfrac{1}{2}ma^2\dot{\theta}\dot{\varphi}(1+\cos\varphi) - \tfrac{1}{2}mga\cos\varphi$$

is constant during the motion.

5.18. A top of mass m is pivoted at a point on its axis of symmetry at a distance a from its centre of mass. The three principal moments of inertia at the pivot are A, A, C, where C is the moment of inertia about the symmetry axis. Show that, with an appropriate choice of Euler angles, the Lagrangian is

$$L = \tfrac{1}{2}A(\dot{\theta}^2 + \dot{\varphi}^2\sin^2\theta) + \tfrac{1}{2}C(\dot{\psi} + \dot{\varphi}\cos\theta)^2 - mga\cos\theta\,.$$

Show that $n = \dot{\psi} + \dot{\varphi}\cos\theta$ and $k = A\dot{\varphi}\sin^2\theta + Cn\cos\theta$ are constants of the motion. Write down the total energy and explain briefly why it is also a constant of the motion.

The top is set in motion with

$$\theta = \frac{\pi}{3}, \quad \dot{\theta} = 0, \quad \dot{\varphi} = \frac{2\gamma}{\sqrt{3}}, \quad \dot{\psi} = \frac{(3A - C)\gamma}{C\sqrt{3}}\,,$$

where $\gamma = \sqrt{mga/A}$. Show that

$$\left(\frac{\mathrm{d}u}{\mathrm{d}t}\right)^2 = \gamma^2(1-u)^2(2u-1)$$

where $u = \cos\theta$. Verify that

$$\frac{1}{u} = 1 + \frac{1}{\cosh(\gamma t)}\,.$$

What happens to the axis of the top as $t \to \infty$? See [13], p. 158.

5.19. Kovalevskaya's top has Lagrangian

$$L = C\big(\dot\theta^2 + \dot\varphi^2 \sin^2\theta + \tfrac{1}{2}(\dot\psi + \dot\varphi\cos\theta)^2\big) + mga\sin\theta\cos\psi,$$

where C and m are constants. Describe the physical system that has this Lagrangian. Note that L is independent of φ and t and write down the corresponding conserved quantities. Put

$$z = C(\dot\varphi\sin\theta + \mathrm{i}\,\dot\theta)^2 + mga\sin\theta\,\mathrm{e}^{-\mathrm{i}\psi}.$$

Show that

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \mathrm{i}\,(\dot\varphi\cos\theta - \dot\psi)z$$

and deduce that $|z|^2$ is also conserved. See [13], p. 166.

5.7 Nonholonomic Constraints

We begin with an example. Consider a uniform solid sphere of radius a and mass m, which rolls without slipping on a moving horizontal plane. The plane passes through the origin of an inertial frame $\tilde R = (\tilde O, \tilde{\mathcal{T}})$ and is normal to the z-axis of the frame. It is forced to move parallel to itself with velocity

$$\boldsymbol{U} = u\tilde{\boldsymbol{e}}_1 + v\tilde{\boldsymbol{e}}_2$$

relative to the $\tilde R$, where u and v are given functions of time. We investigate the motion of the sphere under the assumption that the only external forces acting on it are the friction force and normal reaction at the point of contact, and the gravitational force $-mg\tilde{\boldsymbol{e}}_3$.

Let $R = (O, \mathcal{T})$ be a frame fixed in the sphere, with its origin at the centre. Let

$$\boldsymbol{r} = x\tilde{\boldsymbol{e}}_1 + y\tilde{\boldsymbol{e}}_2 + a\tilde{\boldsymbol{e}}_3$$

denote the vector from $\tilde O$ to O and let θ, φ, and ψ denote the Euler angles of $\mathcal{T}$ relative to $\tilde{\mathcal{T}}$ (Figure 5.11).

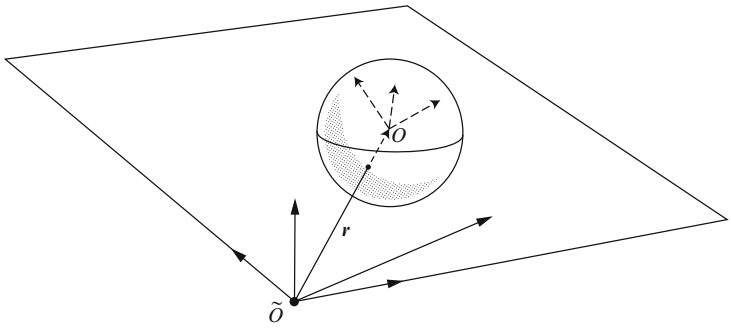

Figure 5.11

All axes through the centre of the sphere are principal axes, with moment of inertia $\frac{2}{5}ma^2$. Therefore the sphere's total kinetic energy relative to $\tilde{R}$ is

$$T = \tfrac{1}{5}ma^2(\dot{\theta}^2 + \dot{\varphi}^2 + \dot{\psi}^2 + 2\dot{\varphi}\dot{\psi}\cos\theta) + \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2).$$

When we eliminate the holonomic constraint that fixes the height of O above the plane, we are left with five degrees of freedom, corresponding to the five generalized coordinates $q_1 = x$, $q_2 = y$, $q_3 = \theta$, $q_4 = \varphi$, and $q_5 = \psi$. The equations of motion are

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial v_a}\right) - \frac{\partial T}{\partial q_a} = K_a$$

where the K_as are the q-components of the friction. Gravity and the normal reaction do not contribute. Our problem is to find the K_as, knowing only that they are the generalized forces responsible for maintaining the rolling condition

$$\dot{\boldsymbol{r}} + \boldsymbol{\omega} \wedge (-a\tilde{\boldsymbol{e}}_3) = u\tilde{\boldsymbol{e}}_1 + v\tilde{\boldsymbol{e}}_2, \tag{5.18}$$

where $\boldsymbol{\omega}$ is the angular velocity of the sphere and the dot denotes the time derivative relative to the inertial frame. The $\tilde{\mathcal{T}}$-components of $\boldsymbol{\omega}$ are given by (1.25). So by taking the $\tilde{\mathcal{T}}$-components of (5.18), we obtain two constraint equations

$$\begin{aligned} \dot{x} - a\dot{\theta}\cos\varphi - a\dot{\psi}\sin\theta\sin\varphi &= u \\ \dot{y} - a\dot{\theta}\sin\varphi + a\dot{\psi}\sin\theta\cos\varphi &= v. \end{aligned} \tag{5.19}$$

In a more general problem of this type, we could be faced with the task of solving

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial v_a}\right) - \frac{\partial T}{\partial q_a} = E_a + K_a \tag{5.20}$$

$(a = 1, 2, ..., n)$ where the E_as are the components of given external forces, and the K_as are the components of unknown constraint forces, responsible for maintaining a number of constraints of the form

$$A_a(q,t)v_a + B(q,t) = 0, \tag{5.21}$$

in which the A_as and Bs are known functions of the q_as and t.

It can happen that such a constraint is really a holonomic constraint in disguise. If $f(q,t) = 0$ throughout the motion, then

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial q_a}v_a + \frac{\partial f}{\partial t} = 0,$$

which is of the same form as (5.21). Conversely, if there exists a function $g(q,t)$ such that

$$A_a = g\frac{\partial f}{\partial q_a}, \qquad B = g\frac{\partial f}{\partial t} \tag{5.22}$$

for some function $f(q,t)$, then (5.21) is equivalent to the condition that f should be constant during the motion. In this case, we can treat $f =$ constant as a holonomic constraint (for various different values of the constant).

Definition 5.12

If (5.22) holds, then the constraint equation (5.21) is *integrable*, and the function $f(q,t)$ is an *integral* of the constraint.

It can be shown that a necessary and sufficient condition for integrability is that

$$F_{ab}A_c + F_{bc}A_a + F_{ca}A_b = 0$$

$$BF_{ab} + A_b\frac{\partial A_a}{\partial t} - A_a\frac{\partial A_b}{\partial t} + A_a\frac{\partial B}{\partial q_b} - A_b\frac{\partial B}{\partial q_a} = 0,$$

where

$$F_{ab} = \frac{\partial A_b}{\partial q_a} - \frac{\partial A_a}{\partial q_b}.$$

By using this criterion, one can show that the constraints (5.19) are *not* integrable. They are genuinely nonholonomic.

Suppose that we have k constraint equations of the form of (5.21),

$$A_{ra}(q,t)v_a + B_r(q,t) = 0, \tag{5.23}$$

where $r = 1, 2, \ldots, k$. If the constraints are integrable, with integrals $f_r(q,t)$, and if the constraint forces are workless, then

$$K_aX_a = 0 \quad \text{whenever} \quad X_a\frac{\partial f_r}{\partial q_a} = 0$$

for all r (see p. 90). This is equivalent to

$$K_aX_a = 0 \quad \text{whenever} \quad X_aA_{ra} = 0 \quad \text{for all } r,$$

a condition that makes sense even when the constraints are not integrable.

Definition 5.13

The generalized constraint forces K_a are said to be *workless* if $K_a X_a = 0$ whenever $X_a A_{ra} = 0$ for all r.

This may seem a strange condition. Its usefulness lies in the fact that it is satisfied by the constraint forces in a variety of systems and, in particular, in most rolling problems. For example, for our sphere, the condition has the following interpretation. Freeze the motion of the plane and the sphere, recording all the forces acting on the system. Choose real numbers X_a $(a = 1, 2, \ldots, 5)$ and consider the small change in the configuration given by

$$\delta x = \varepsilon X_1, \quad \delta y = \varepsilon X_2, \quad \delta\theta = \varepsilon X_3, \quad \delta\varphi = \varepsilon X_4, \quad \delta\varphi = \varepsilon X_5,$$

where ε is some small parameter and t is held fixed. With the two constraint equations (5.19), the condition $X_a A_{ra} = 0$ $(r = 1, 2)$ reads

$$\begin{aligned} X_1 - aX_3 \cos\varphi - aX_5 \sin\theta \sin\varphi &= 0 \\ X_2 - aX_3 \sin\varphi + aX_5 \sin\theta \cos\varphi &= 0, \end{aligned}$$

or, equivalently,

$$\begin{aligned} \delta x - a\delta\theta \cos\varphi - a\delta\psi \sin\theta \sin\varphi &= 0 \\ \delta y - a\delta\theta \sin\varphi + a\delta\psi \sin\theta \cos\varphi &= 0 \end{aligned}$$

to the first order. But this is simply the condition that the displacement should be produced by rolling the sphere a small distance along the plane with the plane held fixed, by reversing the argument that led to (5.19). During such a displacement, the point of the sphere initially in contact with the plane does not move to the first order in ε, which is another way of saying that the displacement is produced by rolling. It follows that if $X_a A_{ra} = 0$ $(r = 1, 2)$, then the friction forces do no work during the displacement and $K_a X_a = 0$. In this case, therefore, the constraint forces are workless in the sense of the definition. Note, however, that this is not the same as saying that the friction forces do no work during the actual motion, in which the motion of the plane is *not* frozen. That would be false.

Now return to the general case and consider the values of the functions K_a, A_{ra}, and B at some fixed values of q_a and t.

Proposition 5.14

If $X_a K_a = 0$ whenever $X_a A_{ra} = 0$, then there exist real numbers $\lambda_1, \ldots, \lambda_k$ such that

$$K_a = \lambda_1 A_{1a} + \lambda_2 A_{2a} + \cdots + \lambda_k A_{ka}. \tag{5.24}$$

Proof

The proposition is a restatement of a standard result in linear algebra. We regard X_1, X_2,..., X_n as the entries in a column vector X. Let A denote the $k \times n$ matrix

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{k1} & A_{k2} & \dots & A_{kn} \end{pmatrix}$$

and let ρ denote the rank of A, that is the dimension of the vector space $V \subset \mathbb{R}^n$ spanned by the rows of A. Let N denote the space of column vectors X such that $AX = 0$. Then the rank-nullity theorem tells us that $\dim(N) = n - \rho$.

Let N^0 denote the annihilator of N. That is, N^0 is the space of row vectors

$$C = (c_1, c_2, \dots, c_n)$$

such that $CX = 0$ whenever $X \in N$. Then

$$\dim(N^0) = n - \dim(N) = \rho.$$

Now V is certainly contained in N^0, by the definition of N. But V and N^0 have the same dimension. Therefore $V = N^0$.

The condition $K_a X_a = 0$ whenever $A_{ra} X_a = 0$ is simply the condition that the row vector $K = (K_1, \dots, K_n)$ should lie in N^0. When it holds, K lies in V and so there exist $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ such that (5.24) holds. □

Suppose now that the constraint forces in (5.20) are workless. Then, by applying the proposition at each time during the motion, we can deduce that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial v_a}\right) - \frac{\partial T}{\partial q_a} = E_a + \lambda_1 A_{1a} + \dots + \lambda_k A_{ka} \tag{5.25}$$

where λ_1,..., λ_k are now functions of time. Together with the original constraint equations (5.23), this gives a system of $n + k$ equations in the $n + k$ unknown functions of time $q_1(t), \dots, q_n(t), \lambda_1(t), \dots, \lambda_k(t)$.

When the external forces are conservative, so that $E_a = -\partial U / \partial q_a$, (5.25) becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v_a}\right) - \frac{\partial L}{\partial q_a} = \lambda_1 A_{1a} + \dots + \lambda_k A_{ka} \tag{5.26}$$

where $L = T - U$.

Definition 5.15

The functions $\lambda_1(t)$,..., $\lambda_k(t)$ are called *Lagrange multipliers*.

In our example of the sphere rolling on the moving plane, we have seven equations, the two constraints, together with

$$(x) \qquad \frac{\mathrm{d}}{\mathrm{d}t}\big(m\dot{x}\big) = \lambda$$

$$(y) \qquad \frac{\mathrm{d}}{\mathrm{d}t}\big(m\dot{y}\big) = \mu$$

$$(\theta) \qquad \frac{\mathrm{d}}{\mathrm{d}t}\big(\tfrac{2}{5}ma^2\dot{\theta}\big) + \tfrac{2}{5}ma^2\dot{\varphi}\dot{\psi}\sin\theta = -a\lambda\cos\varphi - a\mu\sin\varphi$$

$$(\varphi) \qquad \frac{\mathrm{d}}{\mathrm{d}t}\big(\tfrac{2}{5}ma^2(\dot{\varphi} + \dot{\psi}\cos\theta)\big) = 0$$

$$(\psi) \qquad \frac{\mathrm{d}}{\mathrm{d}t}\big(\tfrac{2}{5}ma^2(\dot{\psi} + \dot{\varphi}\cos\theta)\big) = -a\lambda\sin\theta\sin\varphi + a\mu\sin\theta\cos\varphi$$

where $\lambda = \lambda_1$ and $\mu = \lambda_2$ are the two Lagrange multipliers corresponding to the two constraints. We can see from the first two equations that they are in fact the two components of the friction force at the point of contact.

EXERCISES

5.20. Rework example (5.11) by treating the rolling condition as if it were a nonholonomic constraint.

5.21. A system is subject to workless nonholonomic constraints with

$$B_r = 0.$$

Show that if $T = \frac{1}{2}T_{ab}v_av_b$ where $T_{ab} = T_{ab}(q)$, then

$$\frac{\mathrm{d}T}{\mathrm{d}t} = v_aE_a.$$

5.22. Solve Exercise 5.13 by using coordinates and Lagrange multipliers.

6
Oscillations

6.1 Simple Harmonic Motion

In this chapter, we look at an analogue of simple harmonic motion in systems with many degrees of freedom. This is not only of interest in itself, but it also gives a good approximation to the behaviour near equilibrium of a general class of mechanical systems.

With just one degree of freedom, one can characterize simple harmonic motion by the equation of motion

$$\ddot{q} + \omega^2 q = 0,$$

or equivalently by the form of the Lagrangian

$$L = \tfrac{1}{2}v^2 - \tfrac{1}{2}\omega^2 q^2$$

that generates the motion. The constant ω is the angular frequency of the oscillations. Angular frequency is measured in radians per second and is related to frequency ν, which is measured in hertz (Hz), or cycles per second, by $\omega = 2\pi\nu$.

6.2 Several Degrees of Freedom

In a system with n degrees of freedom, the analogue is the motion generated by

$$L = \tfrac{1}{2}K_{ab}v_a v_b - \tfrac{1}{2}P_{ab}q_a q_b$$

N.M.J. Woodhouse, *Introduction to Analytical Dynamics: New Edition*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-1-84882-816-2_6,

where the P_{ab}s are the entries in a constant $n \times n$ symmetric matrix P and the K_{ab}s are the entries in a constant positive definite symmetric matrix K. The corresponding equations of motion are

$$K_{ab}\ddot{q}_b + P_{ab}q_b = 0 \tag{6.1}$$

or, equivalently,

$$K\ddot{q} + Pq = 0$$

where q is the column vector with entries q_a.

The key property of (6.1) is *linearity*. If $q_a = x_a(t)$ and $q_a = y_a(t)$ are solutions, then so is the linear combination

$$q_a = \alpha x_a(t) + \beta y_a(t)$$

for any constant α, β. By exploiting linearity, we shall break up the general solution, which can be a complicated trajectory in configuration space, into a superposition of fundamental solutions which have a more straightforward time-dependence. First, however, we need a result from linear algebra.

Proposition 6.1

Let K and P be real symmetric $n \times n$ matrices and suppose that K is positive definite. Then there exists a nonsingular matrix B such that $B^{\mathrm{t}}KB$ is the identity matrix and such that $D = B^{\mathrm{t}}PB$ is a diagonal matrix.

Proof

Because K is symmetric, there exists an orthogonal matrix H such that

$$H^{\mathrm{t}}KH = \begin{pmatrix} k_1 & 0 & 0 & \dots & 0 \\ 0 & k_2 & 0 & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \dots & k_n \end{pmatrix}.$$

Moreover, because K is positive definite, the k_as are all positive.

Let M denote the diagonal matrix with diagonal entries $k_1^{-1/2}, \dots, k_n^{-1/2}$. Then

$$M^{\mathrm{t}}H^{\mathrm{t}}KHM = I,$$

where I is the $n \times n$ identity matrix, and

$$S = M^{\mathrm{t}}H^{\mathrm{t}}PHM$$

is a symmetric matrix. Let N be an orthogonal matrix such that $N^{\mathrm{t}}SN$ is diagonal. Put $D = N^{\mathrm{t}}SN$ and $B = HMN$. Then B is non-singular and

$$\begin{aligned} B^{\mathrm{t}}KB &= N^{\mathrm{t}}M^{\mathrm{t}}H^{\mathrm{t}}KHMN = N^{\mathrm{t}}N = I \\ B^{\mathrm{t}}PB &= N^{\mathrm{t}}S^{\mathrm{t}}N = D, \end{aligned} \tag{6.2}$$

which completes the proof. □

In index notation, (6.2) reads

$$B_{ca}K_{cd}B_{db} = \delta_{ab} \quad \text{and} \quad B_{ca}P_{cd}B_{db} = D_{ab},$$

where $D_{ab} = 0$ whenever $a \neq b$. If we make a linear transformation to a new system of coordinates $\tilde{q}_a$ which are related to the q_as by

$$q_a = B_{ab}\tilde{q}_b,$$

then the Lagrangian becomes

$$\begin{aligned} L &= \tfrac{1}{2}\tilde{v}_a\tilde{v}_a - \tfrac{1}{2}D_{ab}\tilde{q}_a\tilde{q}_b \\ &= \tfrac{1}{2}(\tilde{v}_1^2 + \tilde{v}_2^2 + \cdots + \tilde{v}_n^2) - \tfrac{1}{2}(\lambda_1\tilde{q}_1^2 + \lambda_2\tilde{v}_2^2 + \cdots + \lambda_n\tilde{q}_n^2), \end{aligned} \tag{6.3}$$

where $\lambda_1 = D_{11}$, $\lambda_2 = D_{22}$, and so on.

Definition 6.2

Generalized coordinates in which the Lagrangian takes the form (6.3) are called *normal coordinates*.

In normal coordinates, the equations of motion are

$$\ddot{\tilde{q}}_1 + \lambda_1\tilde{q}_1 = 0, \quad \ddot{\tilde{q}}_2 + \lambda_2\tilde{q}_2 = 0, \quad \ldots \tag{6.4}$$

and the dynamical problem is very much simpler. The system of simultaneous linear differential equations for the q_as has been transformed into n separate equations for the $\tilde{q}_a$s, which can be solved independently of each other.

The general solution of

$$\ddot{q} + \lambda q = 0$$

takes a different form according to whether λ is positive, negative, or zero:

$$\begin{aligned} q &= E\cos\omega t + F\sin\omega t && (\lambda = \omega^2 > 0) \\ q &= E\mathrm{e}^{\omega t} + F\mathrm{e}^{-\omega t} && (\lambda = -\omega^2 < 0) \\ q &= E + Ft && (\lambda = 0), \end{aligned} \tag{6.5}$$

where E and F are constants. Thus in the general solution of (6.4), the time-dependence of each normal coordinate is of one of these forms, depending on the sign of the corresponding λ_a. There are, in particular, special solutions in which all but one of the normal coordinates vanish identically, and every solution is a linear combination of these special solutions.

Let us consider how one of the special solutions looks in the original coordinate system q_a. Suppose, for example, that $\tilde{q}_2 = \cdots = \tilde{q}_n = 0$ throughout the motion and that $\tilde{q}_1$ is of one of the three forms (6.5), with $\lambda = \lambda_1$. Then

$$\begin{aligned} q_a &= (E\cos\omega t + F\sin\omega t)A_a && (\text{if } \lambda = \lambda_1 = \omega^2 > 0) \\ q_a &= (E\mathrm{e}^{\omega t} + F\mathrm{e}^{-\omega t})A_a && (\text{if } \lambda = \lambda_1 = -\omega^2 < 0) \\ q_a &= (E + Ft)A_a && (\text{if } \lambda = \lambda_1 = 0) \end{aligned} \tag{6.6}$$

where the A_as are the entries in the first column of the matrix B. On substituting each of these into the equation of motion (6.1), we find that

$$(\lambda K_{ab} - P_{ab})A_b = 0.$$

In all three cases, therefore, there exists a column vector A, which cannot be zero because B is nonsingular, such that

$$(\lambda K - P)A = 0. \tag{6.7}$$

It follows that the matrix $\lambda K - P$ is singular and that λ must be one of the roots of the *characteristic equation*

$$\det(\lambda K - P) = 0,$$

which is a polynomial equation of degree n in λ.

Proposition 6.3

The roots of the characteristic equation are the diagonal entries in D. In particular, the roots are all real.

Proof

In the notation of Proposition 6.1,

$$\begin{aligned} (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n) &= \det(\lambda I - D) \\ &= \det\bigl(B^{\mathrm{t}}(\lambda K - P)B\bigr) \\ &= \det(\lambda K - P)\det(B)^2. \end{aligned}$$

The proposition follows because $\det(B) \neq 0$. □

The point of all this is that it gives us a way of finding solutions in the form of (6.6) without actually carrying out the transformation to normal coordinates. All that we need to do is the following. First solve the characteristic equation. Next, for each root λ, choose two constants E and F and a nonzero column vector A with entries A_aa such that (6.7) holds. Finally, define q_a as a function of t by the appropriate expression in (6.6). Then $q_a = q_a(t)$ is a solution. Moreover the existence of the transformation to normal coordinates tells us that the general solution is a linear combination of solutions of this form.

Definition 6.4

A solution of one of the forms

$$\begin{aligned} q_a &= (E\cos\omega t + F\sin\omega t)A_a \\ q_a &= (E\mathrm{e}^{\omega t} + F\mathrm{e}^{-\omega t})A_a \\ q_a &= (E + Ft)A_a, \end{aligned}$$

where E, F, ω, and A_a are constant, is called a *fundamental solution*. The three types are said to be, respectively, *oscillatory*, *exponential*, and *linear*.

Oscillatory fundamental solutions are also called *normal modes* of oscillation and the corresponding angular frequencies $\omega = \sqrt{\lambda}$ are called the *normal frequencies*. Note that a linear combination of two fundamental solutions corresponding to roots λ and μ can itself be a fundamental solution only if $\lambda = \mu$.

If the characteristic equation has any non-positive roots, then, except for very special choices of initial conditions, the solution contains terms that grow exponentially or linearly with time and the system is unstable; but if all the roots are positive, then the system is stable and the general motion is a superposition of simple harmonic motions at the normal frequencies.

Example 6.5 (Lissajous Figures)

Take $n = 2$, $\lambda_1 = \omega_1^2$ and $\lambda_2 = \omega_2^2$. Then the general solution in normal coordinates is

$$\tilde{q}_1 = C_1\cos(\omega_1 t + \varepsilon_1), \quad \tilde{q}_2 = C_2\cos(\omega_2 t + \varepsilon_2),$$

where C_1, C_2, ε_1 and ε_2 are constants. The dynamical trajectories in the corresponding configuration space are called *Lissajous figures*. Such a trajectory returns eventually to its initial configuration only if ω_1/ω_2 is rational. See Figure 6.1.

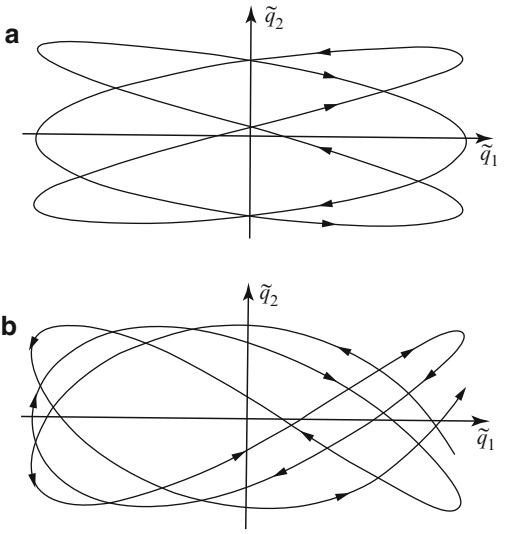

Figure 6.1 Trajectories in C (a) ω_1/ω_2 rational, (b) ω_1/ω_2 irrational.

When $C_1 = C_2 = 1$, and $\varepsilon_1 = \varepsilon_2 = 0$, $\omega_1 = 1$, and $\omega_2 = N$, where N is an integer, we have

$$\tilde{q}_1 = \cos t, \qquad \tilde{q}_2 = \cos Nt.$$

In this case, therefore,

$$\tilde{q}_2 = \cos(N \cos^{-1} \tilde{q}_1) = T_N(\tilde{q}_1),$$

and so the configuration oscillates back and forth along the graph of the function T_N between $\tilde{q}_1 = -1$ and $\tilde{q}_1 = 1$. By using trigonometric identities, one can see that T_N is a polynomial of degree N in $\tilde{q}_1$. It is called a *Chebyshev polynomial*. If the values of ε_1 and ε_2 are changed slightly, then the Lissajous figure becomes a closed curve near the graph of T_N (Figure 6.2).

Consider a fundamental solution corresponding to a root λ of the characteristic equation. The time-dependence of each of the q_as and of each of the $\tilde{q}_a$s is of the form of (6.5). But the normal coordinates $\tilde{q}_a$ must also satisfy (6.4), whatever the motion. This is a contradiction unless $\tilde{q}_a$ vanishes identically for every a for which $\lambda_a \neq \lambda$. If all the roots of the characteristic equation are distinct, then the λ_as are distinct and the fundamental solutions are precisely the solutions in which all but one of the normal coordinates vanish identically. However, if, for example, $\lambda_1 = \lambda_2$, then it is possible to have a fundamental solution in which both $\tilde{q}_1$ and $\tilde{q}_2$ are excited.

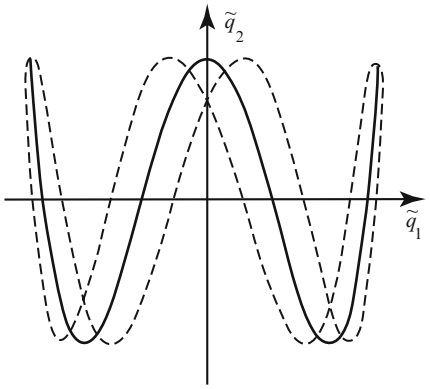

Figure 6.2 The solid curve is the graph of T_3. The broken curve is a Lissajous figure.

Example 6.6

Consider the Lagrangian

$$L = \tfrac{1}{2}(v_1^2 + v_2^2) - \tfrac{1}{2}(q_1^2 + q_2^2).$$

Here q_1 and q_2 form a system of normal coordinates, but $q_1 = \cos t$, $q_2 = \cos t$ is a fundamental solution in which both q_1 and q_2 are excited.

The procedure for finding normal coordinates is the same as that for diagonalizing a symmetric matrix, except that the role of the identity matrix is filled by the positive definite matrix K. It follows from (6.2) that if A is the kth column of B and λ is the kth diagonal entry in D, then

$$(\lambda K - P)A = 0. \tag{6.8}$$

To find the columns of B, therefore, we must first find the roots of the characteristic equation and choose corresponding solutions of (6.8). The choices cannot be made arbitrarily, however. To have $B^{\mathrm{t}}KB = I$, we must choose the different column vectors A so that they are *orthonormal* with respect to K, that is, they must be normalized in the sense that

$$A^{\mathrm{t}}KA = 1$$

and *orthogonal* in the sense that

$$A^{\mathrm{t}}KA' = 0.$$

Here A and A' are two distinct columns vectors with

$$\lambda KA = PA \quad \text{and} \quad \lambda' KA' = PA'.$$

If $\lambda \neq \lambda'$, then this latter condition is not a problem. It is an automatic consequence of

$$(\lambda - \lambda')A^{\mathrm{t}}KA' = (\lambda KA)^{\mathrm{t}}A' - A^{\mathrm{t}}(\lambda' KA') = (PA)^{\mathrm{t}}A' - A^{\mathrm{t}}PA' = 0.$$

However, if m of the roots are all equal to λ, say, then the solution space of

$$(\lambda K - P)A = 0$$

is m-dimensional, and one must choose the corresponding As to be an orthonormal basis of the solution space.

Example 6.7

Take $n = 3$ and

$$\begin{aligned} T &= \tfrac{1}{2}(2v_1^2 + v_2^2 + 3v_3^2 - 2v_1v_2 - 4v_1v_3 + 2v_2v_3) \\ U &= \tfrac{1}{2}(3q_1^2 + 2q_2^2 + 4q_3^2 - 4q_1q_2 - 6q_1q_3 + 4q_2q_3). \end{aligned}$$

Then

$$K = \begin{pmatrix} 2 & -1 & -2 \\ -1 & 1 & 1 \\ -2 & 1 & 3 \end{pmatrix}, \qquad P = \begin{pmatrix} 3 & -2 & -3 \\ -2 & 2 & 2 \\ -3 & 2 & 4 \end{pmatrix}.$$

The characteristic equation is

$$\det(\lambda K - P) = \begin{vmatrix} 2\lambda - 3 & -\lambda + 2 & -\lambda + 3 \\ -\lambda + 2 & \lambda - 2 & \lambda - 2 \\ -2\lambda + 3 & \lambda - 2 & 3\lambda - 4 \end{vmatrix} = 0.$$

Without expanding the determinant, we can spot that $\lambda = 1$ is a repeated root and that the third root is $\lambda = 2$.

With $\lambda = 1$, the solutions of $(\lambda K - P)A = 0$ are

$$A = \alpha \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \tag{6.9}$$

for any α and β. With $\lambda = 2$, the solutions are

$$A = \gamma \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \tag{6.10}$$

for any γ. If A and A' are as in (6.9), then

$$A^{\mathrm{t}} K A' = \alpha\alpha' + \beta\beta'$$

and if A is given by (6.10), then

$$A^{\mathrm{t}} K A = \gamma^2.$$

Hence we can choose the three columns of B to be as follows. For the first, we take (6.9) with $\alpha = 1$, $\beta = 0$, for the second, (6.9) with $\alpha = 0$, $\beta = 1$, and for the third (6.10) with $\gamma = 1$. This gives

$$B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

The corresponding normal coordinates are found from $\tilde{q} = B^{-1} q$. That is

$$\begin{aligned} \tilde{q}_1 &= q_3, \\ \tilde{q}_2 &= q_1 - q_3, \\ \tilde{q}_3 &= -q_1 + q_2 + q_3. \end{aligned}$$

However, B is not unique because we could take different combinations of α and β for the first two columns.

EXERCISES

6.1. In Example 6.7, show that the dynamical trajectories lie in surfaces with equations of the form

$$a\tilde{q}_1^2 + 2b\tilde{q}_1\tilde{q}_2 + c\tilde{q}_2^2 = 1.$$

Sketch some typical trajectories on the cylinder $\tilde{q}_1^2 + \tilde{q}_2^2 = 1$.

6.2. Solve the characteristic equation and find normal coordinates in the case that

$$\begin{aligned} T &= \tfrac{1}{2}(3v_1^2 + 3v_2^2 + 3v_3^2 - 2v_1v_2 - 2v_2v_3 - 2v_3v_1) \\ U &= \tfrac{1}{2}(3q_1^2 + 3q_2^2 + 3q_3^2 - 2q_1q_2 - 4q_1q_3). \end{aligned}$$

6.3 Oscillations Near Equilibrium

So far there has been no approximation. But now consider a system with Lagrangian

$$L = \tfrac{1}{2}T_{ab}v_a v_b - U,$$

in which U and the T_{ab}s are functions of the generalized coordinates q_a, but not of t. The equilibrium configurations of the system, that is the configurations in which Lagrange's equations admit solutions of the form $q_a = \text{constant}$, are given by

$$\frac{\partial U}{\partial q_a} = 0.$$

We are interested in how the system behaves when it is close to equilibrium.

Suppose that the coordinates have been chosen so that $q_a = 0$ is an equilibrium configuration. There is no loss of generality in supposing that $U(0) = 0$. Then, by ignoring terms of the third and higher orders in the coordinates and velocities,

$$L = \tfrac{1}{2}K_{ab}v_a v_b - \tfrac{1}{2}P_{ab}q_a q_b \tag{6.11}$$

where $K_{ab} = T_{ab}(0)$ and

$$P_{ab} = \left.\frac{\partial^2 U}{\partial q_a \partial q_b}\right|_{q=0}.$$

The motions close to equilibrium are those in which both the q_as and the v_as remain small. We should be able to analyse these by replacing the original Lagrangian by its approximation (6.11), although it is far from obvious that the exact solutions of the approximate equations of motion are in fact approximate solutions of the exact equations, particularly over a long period of time. We shall make no attempt to unravel this problem beyond noting that the method is clearly *invalid* if the characteristic equation of (6.11) has any non-positive roots. There are then fundamental solutions containing exponential or linear terms, which certainly do not remain close to the equilibrium configuration.

If there are negative roots, then the original system is unstable, and if all the roots are positive, then it is stable. The presence of zero roots, on the other hand, indicates no more than that the approximation is invalid. The stability question is left unresolved.

The following example indicates how the analysis goes in practice.

Example 6.8

†A uniform rod of mass m and length $2L$ is suspended by two light elastic strings of natural length a and modulus $\frac{1}{2}mg$, as shown in Figure 6.3. It is free

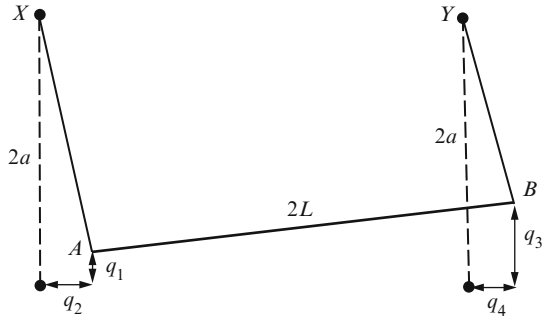

Figure 6.3

to swing in the vertical plane through the points of suspension. The problem is to find the normal modes of oscillation near equilibrium.

In equilibrium, the strings have length $2a$. Take the equilibrium position of A as the origin and introduce axes with the x-axis horizontal and the y-axis vertically upwards. Let q_1 and q_2 denote, respectively, the y and x coordinates of A; and let q_3 and $q_4 + 2L$ denote the y and x coordinates of B.

In fact the system has only three degrees of freedom because q_4 can be eliminated by using the constraint equation

$$(2L + q_4 - q_2)^2 + (q_3 - q_1)^2 = 4L^2,$$

which fixes the length of the rod. This reduces to

$$q_4 = q_2 + O(q^2)$$

in the first-order approximation.

The height of the centre of the rod above its equilibrium position is

$$\tfrac{1}{2}(q_1 + q_3).$$

and the lengths of the strings XA and YB are

$$XA = \sqrt{(2a - q_1)^2 + q_2^2}, \qquad YB = \sqrt{(2a - q_3)^2 + q_4^2}. \tag{6.12}$$

Hence the potential, which is the sum of the elastic and gravitational potential energies, is

$$U = \tfrac{1}{2}mg(q_1 + q_3) + \frac{mg}{4a}(XA - a)^2 + \frac{mg}{4a}(YB - a)^2 - \frac{mga}{2},$$

the final constant being chosen to make $U(0) = 0$. We are always free to add constants to U without changing the equations of motion.

The theory demands that we should express U as a function of q_1, q_2, and q_3 and then calculate the entries in the matrix P by taking the second partial derivatives. Of course, this works, but it is not quick. It is simpler to expand U up to the second order in the q_as. From (6.12),

$$\begin{aligned} XA &= 2a\sqrt{1 - \frac{q_1}{a} + \frac{q_2^2 + q_1^2}{4a^2}} \\ &= 2a - q_1 + \frac{q_2^2}{4a} + O(q^3) \end{aligned}$$

with a similar expression for YB. Therefore

$$(XA - a)^2 = a^2 - 2aq_1 + \tfrac{1}{2}q_2^2 + q_1^2 + O(q^3).$$

Hence, to the second order in q_1, q_2, and q_3,

$$\begin{aligned} U &= \frac{mg}{4a}(q_1^2 + \tfrac{1}{2}q_2^2 + q_3^2 + \tfrac{1}{2}q_4^2) \\ &= \frac{mg}{4a}(q_1^2 + q_2^2 + q_3^2) \end{aligned}$$

because $q_4 = q_2 + O(q^2)$. We have been lucky. Had the expression for U contained linear terms in q_4, we would not have got away with the substitution $q_4 = q_2$.

It is easier to deal with the kinetic energy. We have only to write down an expression for T *in the equilibrium configuration*. In equilibrium, we have $v_2 = v_4$ without approximation and

$$T = \tfrac{1}{6}m(v_1^2 + 3v_2^2 + v_3^2 + v_1v_3),$$

from the formula for the kinetic energy of a rod. See Exercise 5.4.

The characteristic equation is

$$\begin{vmatrix} 2a\lambda - 3g & 0 & a\lambda \\ 0 & 6a\lambda - 3g & 0 \\ a\lambda & 0 & 2a\lambda - 3g \end{vmatrix} = 0$$

Hence the normal angular frequencies are $\omega_1 = \sqrt{g/a}$, $\omega_2 = \sqrt{g/2a}$, and $\omega_3 = \sqrt{3g/a}$. The corresponding normal modes are, first,

$$\begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} = C_1 \cos\bigl(\omega_1 t + \varepsilon_1\bigr) \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},$$

where C_1 and ε_1 are constants. In this mode, the strings remain vertical and the two ends of the rod oscillate up and down in phase. Second,

$$\begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} = C_2 \cos\bigl(\omega_2 t + \varepsilon_2\bigr) \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

The lengths of the strings are constant and the rod swings back and forth in the vertical plane. Third,

$$\begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} = C_3 \cos\bigl(\omega_3 t + \varepsilon_3\bigr) \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

The strings remain vertical and the two ends oscillate up and down, exactly out of phase.

EXERCISES

6.3. †A uniform rod AB of mass m and length $2a$ is suspended by two light inextensible strings PA and QB, each of length $2a$. The fixed points P and Q are at the same level, distance $2a$ apart; and in equilibrium, PA and QB are vertical. The system performs small oscillations about its equilibrium configuration. Determine the normal frequencies and describe the normal modes.

6.4. †Four uniform rods AB, BC, CD, DA, each of mass m and length $2a$, are smoothly jointed together at A, B, C, and D. They are suspended by four light inextensible strings PA, QB, RC, SD, each of length a, attached to fixed points P, Q, R, and S. In equilibrium the strings are vertical and the rods lie in a horizontal plane in the form of a square. Show that the normal frequencies are $\sqrt{g/a}$ and $\sqrt{3g/2a}$. Describe the normal modes.

6.5. †A particle of mass m is attached by identical light elastic strings of natural length a and modulus of elasticity λ to four points A, B, C, and D, which lie at the corners of a square of side $2a$ in a horizontal plane. In equilibrium, the particle hangs under gravity at a distance $a\sqrt{2}$ below $ABCD$. Show that $mg = \lambda\sqrt{8}$. Find the normal frequencies for small oscillations about equilibrium and describe the normal modes.

6.6. †A particle A of mass m is suspended from a fixed point O by a light string of length $3a$. A second particle B of mass m is suspended from A by a second light string of length $2b$ and a third particle C of mass m is suspended from B by a third light string of length c. The system can move in a vertical plane through O.

Let x, y, z denote the horizontal displacements of A, B, C from their equilibrium positions. Show that an approximate Lagrangian

for small displacements of the system from equilibrium is

$$L = \frac{m}{2}\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) - \frac{mg}{2}\left(\frac{x^2}{a} + \frac{(y-x)^2}{b} + \frac{(z-y)^2}{c}\right).$$

Show that there is a normal mode in which $y(t) = 0$ if and only if

$$\frac{1}{a} + \frac{1}{b} - \frac{1}{c} = 0\,.$$

6.7. †A double pendulum consists of a bob of mass m suspended under gravity from a fixed point by a light string of length a, and a second bob, also of mass m, suspended from the first by a further light string, also of length a. Show that if the motion is near equilibrium and is confined to a vertical plane through the fixed point then the normal frequencies are $\sqrt{(2 \pm \sqrt{2})g/a}$. Describe the corresponding normal modes.

Without doing further calculation, describe the normal modes when the motion is not confined to a vertical plane.

7 Hamiltonian Mechanics

7.1 The Legendre Transformation

Lagrange's equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v_a}\right) - \frac{\partial L}{\partial q_a} = 0, \qquad \frac{\mathrm{d}q_a}{\mathrm{d}t} = v_a \tag{7.1}$$

determine the trajectories in the extended phase space of a holonomic system subject to conservative forces. We have seen that it is possible to simplify the dynamical analysis by making coordinate transformations of the form

$$\begin{aligned} \tilde{q}_a &= \tilde{q}_a(q,t) \\ \tilde{v}_a &= \tilde{v}_a(q,v,t) = \frac{\partial \tilde{q}_a}{\partial q_b} v_b + \frac{\partial \tilde{q}_a}{\partial t} \\ \tilde{t} &= t, \end{aligned} \tag{7.2}$$

a technique that proved particularly useful for handling constraints.

Here the transformation of the configuration coordinates is arbitrary, but the transformation of the velocity coordinates is fixed by the choice made for $\tilde{q}_a = \tilde{q}_a(q,t)$. One might hope to achieve further simplification by allowing transformations that mix up the velocity and configuration coordinates. But because of the asymmetric way in which the q_as and v_as appear in (7.1), any general substitution

$$\tilde{q}_a = \tilde{q}_a(q,v,t), \qquad \tilde{v}_a = \tilde{v}_a(q,v,t), \qquad \tilde{t} = t$$

completely destroys Lagrange's simple form of the equations of motion.

N.M.J. Woodhouse, *Introduction to Analytical Dynamics: New Edition*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-1-84882-816-2_7,

The first step towards introducing more general transformations, therefore, is to replace q_a and v_a by new coordinates that appear more on the same footing in the dynamical equations. The way in which the time derivative of $\partial L/\partial v_a$ enters the first set of equations suggests using the following

$$\tilde{q}_a = q_a, \qquad \tilde{p}_a = \frac{\partial L}{\partial v_a}, \qquad \tilde{t} = t.$$

That is, the velocity coordinates v_a are replaced by the generalized momenta. Together, $\tilde{q}_a$, $\tilde{p}_a$, and $\tilde{t}$ form a set of $2n+1$ coordinates on the extended phase space $P \times \mathbb{R}$. The transformation to these new coordinates is known as the *Legendre transformation*. Because the momenta corresponding to cyclic coordinates are conserved, we anticipate that the equations of motion should take a simple form in these new coordinates.

As usual, the tildes are used to allow us to distinguish between derivative operators that might otherwise be confused, such as

$$\frac{\partial}{\partial q_a} \quad \text{and} \quad \frac{\partial}{\partial \tilde{q}_a}.$$

The first is the partial derivative holding fixed t, the remaining q coordinates, and all the v_as. The second is the partial derivative holding fixed $\tilde{t}$, the remaining $\tilde{q}$ coordinates, and all the $\tilde{p}_a$s.

Example 7.1

Suppose that $n = 1$ and that

$$L = \tfrac{1}{2}v_1^2 - q_1 v_1.$$

Then the new coordinates are

$$\tilde{q}_1 = q_1, \quad \tilde{p}_1 = v_1 - q_1, \quad \tilde{t} = t.$$

Put $f = v_1$. Expressed in the new coordinates, $f = \tilde{p}_1 + \tilde{q}_1$. So we have

$$\frac{\partial f}{\partial q_1} = 0 \quad \text{but} \quad \frac{\partial f}{\partial \tilde{q}_1} = 1.$$

The notational distinction between the partial derivatives with respect to q_1 and $\tilde{q}_1$ is necessary, despite the fact that $q_1 = \tilde{q}_1$.

Provided that the matrix with entries

$$\frac{\partial^2 L}{\partial v_a \partial v_b}$$

is invertible, the Legendre transformation itself can be inverted, at least in some open neighbourhood in the extended phase space. The inverse transformation is of the form

$$q_a = \tilde{q}_a, \qquad v_a = v_a(\tilde{q}, \tilde{p}, \tilde{t}), \qquad t = \tilde{t}.$$

Therefore

$$\frac{\partial q_b}{\partial \tilde{q}_a} = \delta_{ab}, \quad \frac{\partial t}{\partial \tilde{q}_a} = 0, \quad \frac{\partial q_b}{\partial \tilde{p}_a} = 0, \quad \frac{\partial t}{\partial \tilde{p}_a} = 0.$$

It follows that the q-derivatives of the Lagrangian in the two coordinate systems are related by

$$\frac{\partial L}{\partial \tilde{q}_a} = \frac{\partial L}{\partial q_b}\frac{\partial q_b}{\partial \tilde{q}_a} + \frac{\partial L}{\partial v_b}\frac{\partial v_b}{\partial \tilde{q}_a} + \frac{\partial L}{\partial t}\frac{\partial t}{\partial \tilde{q}_a} = \frac{\partial L}{\partial q_a} + \tilde{p}_b \frac{\partial v_b}{\partial \tilde{q}_a}.$$

In general they are not equal, as is to be expected from the example. We can, however, make use the independence of $\tilde{q}_a$s and $\tilde{p}_a$s to write the relationship between them in the suggestive form

$$\frac{\partial L}{\partial q_a} = \frac{\partial}{\partial \tilde{q}_a}\big(L - \tilde{p}_b v_b\big) = -\frac{\partial h}{\partial \tilde{q}_a}, \tag{7.3}$$

where $h : P \times \mathbb{R} \to \mathbb{R}$ is the *Hamiltonian*, defined by

$$h = \tilde{p}_a v_a - L. \tag{7.4}$$

In a similar manner, we have

$$\frac{\partial L}{\partial \tilde{p}_a} = \frac{\partial L}{\partial q_b}\frac{\partial q_b}{\partial \tilde{p}_a} + \frac{\partial L}{\partial v_b}\frac{\partial v_b}{\partial \tilde{p}_a} + \frac{\partial L}{\partial t}\frac{\partial t}{\partial \tilde{p}_a} = \tilde{p}_b \frac{\partial v_b}{\partial \tilde{p}_a},$$

from which it follows that

$$v_a = \frac{\partial}{\partial \tilde{p}_a}\big(\tilde{p}_b v_b - L\big) = \frac{\partial h}{\partial \tilde{p}_a}. \tag{7.5}$$

A note of warning: h must be written entirely in terms of the coordinates $\tilde{q}_a$, $\tilde{p}_a$, $\tilde{t}$ before taking the partial derivative with respect to $\tilde{q}_a$ in (7.3). The tilde notation is unambiguous in its guidance on this point, but because the definition of h involves a mixture of the old and new coordinates, it is only too easy to forget the warning in solving problems, and to differentiate h before expressing the v_as in terms of the new coordinates.

7.2 Hamilton's Equations

In the new coordinates, the equations of motion are

$$\frac{\mathrm{d}\tilde{q}_a}{\mathrm{d}t} = \frac{\partial h}{\partial \tilde{p}_a}, \qquad \frac{\mathrm{d}\tilde{p}_a}{\mathrm{d}t} = -\frac{\partial h}{\partial \tilde{q}_a}.$$

The first follows from (7.5) because

$$v_a = \frac{\mathrm{d}q_a}{\mathrm{d}t} = \frac{\mathrm{d}\tilde{q}_a}{\mathrm{d}t}.$$

The second follows from Lagrange's equations and (7.3) because

$$-\frac{\partial h}{\partial \tilde{q}_a} = \frac{\partial L}{\partial q_a} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v_a}\right) = \frac{\mathrm{d}\tilde{p}_a}{\mathrm{d}t}.$$

There is now an obvious symmetry between the coordinates $\tilde{p}_a$ and $\tilde{q}_a$. The equations of motion in this form are called *Hamilton's equations* and the coordinates $\tilde{p}_1, \ldots, \tilde{p}_n, \tilde{q}_1, \ldots, \tilde{q}_n$ are called *canonical coordinates*.

Hamilton's equations are very important in understanding the general nature of mechanical systems and also the connections between classical and quantum mechanics, but they are not normally of much use in tackling specific problems. The following proposition, however, often gives a quick way of using the Hamiltonian theory to find a constant of the motion. It can be helpful in cases in which Hamilton's equations themselves do not open a useful avenue of progress.

Proposition 7.2

If $\partial L/\partial t = 0$, then h is a constant of the motion.

Proof

By differentiating h along the trajectories in $P \times \mathbb{R}$, we obtain

$$\begin{aligned}\frac{\mathrm{d}h}{\mathrm{d}t} &= \frac{\partial h}{\partial \tilde{q}_a}\frac{\mathrm{d}\tilde{q}_a}{\mathrm{d}t} + \frac{\partial h}{\partial \tilde{p}_a}\frac{\mathrm{d}\tilde{p}_a}{\mathrm{d}t} + \frac{\partial h}{\partial \tilde{t}} \\ &= \frac{\partial h}{\partial \tilde{t}}\end{aligned}$$

from Hamilton's equations. But, from (7.4),

$$\frac{\partial h}{\partial \tilde{t}} = \frac{\partial v_a}{\partial \tilde{t}}\tilde{p}_a - \frac{\partial L}{\partial \tilde{t}}$$

$$= \frac{\partial v_a}{\partial \tilde{t}} \tilde{p}_a - \frac{\partial L}{\partial q_a}\frac{\partial q_a}{\partial \tilde{t}} - \frac{\partial L}{\partial v_a}\frac{\partial v_a}{\partial \tilde{t}} - \frac{\partial L}{\partial t}\frac{\partial t}{\partial \tilde{t}}$$

$$= -\frac{\partial L}{\partial q_a}\frac{\partial q_a}{\partial \tilde{t}} - \frac{\partial L}{\partial t}\frac{\partial t}{\partial \tilde{t}}$$

because $\tilde{p}_a = \partial L/\partial v_a$. When q_a, v_a, and t are expressed as functions of $\tilde{q}_a$, $\tilde{p}_a$, and $\tilde{t}$, we have

$$\frac{\partial q_a}{\partial \tilde{t}} = 0 \quad \text{and} \quad \frac{\partial t}{\partial \tilde{t}} = 1 \tag{7.6}$$

Hence

$$\frac{\partial h}{\partial \tilde{t}} = -\frac{\partial L}{\partial t} \tag{7.7}$$

and so $\mathrm{d}h/\mathrm{d}t = 0$ whenever $\partial L/\partial t = 0$. □

The proof illustrates the importance of keeping track of the distinction between the partial derivatives $\partial/\partial t$ and $\partial/\partial \tilde{t}$.

Example 7.3

Take $n = 1$ and

$$L = \tfrac{1}{2}v_1^2 + v_1 t.$$

Then $\tilde{p}_1 = v_1 + t$, $\tilde{t} = t$, and

$$h = \tfrac{1}{2}v_1^2 = \tfrac{1}{2}(\tilde{p}_1 - \tilde{t})^2.$$

We have

$$\frac{\partial h}{\partial \tilde{t}} = -(\tilde{p}_1 - \tilde{t}) = -v_1 = -\frac{\partial L}{\partial t}.$$

In generalized coordinates adapted to holonomic constraints, the Lagrangian of a system of particles with n degrees of freedom is

$$L = \tfrac{1}{2}K_{ab}v_a v_b + A_a v_a + C,$$

where K_{ab}, A_a, and C are all functions of the q_a's and t, but not of the v_a's. The generalized momenta are

$$\tilde{p}_a = K_{ab}v_b + A_a.$$

Hence

$$h = \tfrac{1}{2}K_{ab}v_a v_b - C. \tag{7.8}$$

Thus the Hamiltonian is obtained from the Lagrangian by reversing the signs of the terms that do not involve the velocities, deleting the terms that are linear

in the velocities, and leaving unchanged the terms that are quadratic in the velocities.

We saw in Section 4.3 that if $\partial L/\partial t = 0$ and if $L = T - U$ where T is a homogeneous quadratic in the velocities and $U = U(q)$, then the total energy

$$E = T + U$$

is a constant of the motion. When L has this form, h is also equal to $T+U$. Hence the energy conservation result in Section 4.3 is a special case of Proposition 7.2. But it is important to note that in general $h \neq E$, and that it is possible for h to be a constant of the motion in a system in which the total energy itself is not conserved. See Exercise 7.1.

Now that they have served their purpose, we shall now drop the tildes on the coordinates in Hamilton's equations. It is safe to do this provided that we remember to express h as a function of q_a, $p_a = \partial L/\partial v_a$, and t *before* taking the partial derivative of h with respect to q_a.

Example 7.4

Consider the motion of a particle P of mass m moving in the plane under the influence of a force of magnitude $\lambda m/r^2$ directed towards a fixed point O, where r is the distance from O to P. The motion is governed by the Lagrangian

$$L = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) + \frac{\lambda m}{r}$$

where r and θ are polar coordinates with origin O. On putting $q_1 = r$ and $q_2 = \theta$, we have, without the tildes,

$$p_1 = m\dot{r}, \qquad p_2 = mr^2\dot{\theta}$$

and

$$\begin{aligned} h &= p_1\dot{r} + p_2\dot{\theta} - L \\ &= \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - \frac{\lambda m}{r} \\ &= \frac{1}{2m}\left(p_1^2 + \frac{p_2^2}{q_1^2}\right) - \frac{\lambda m}{r}. \end{aligned}$$

In this case Hamilton' s equations are

$$\begin{aligned} \dot{q}_1 &= \frac{p_1}{m} & \dot{q}_2 &= \frac{p_2}{mq_1^2} \\ \dot{p}_1 &= \frac{p_2^2}{mq_1^3} - \frac{\lambda m}{q_1^2} & \dot{p}_2 &= 0. \end{aligned}$$

Because $\partial L/\partial t = 0$, the Hamiltonian is a constant of the motion.

Whittaker points out in his *Analytical dynamics* [13] that the 'Hamiltonian', which appeared in a paper by Hamilton published in 1834, also made an earlier appearance in a paper by Poisson in 1809. Moreover, Poisson obtained half of 'Hamilton's equations'. The full set appeared in a paper by Lagrange in 1810, in which he considered the effect of a perturbing force on solutions of dynamical problems, with the role of h played by the potential of the disturbing force. The equations also arose in the context of work on the characteristics of first-order partial differential equations in papers by Pfaff (1814–1815) and Cauchy (1819). Hamilton's 1834 paper dealt only with the case in which h is a function of p_a and q_a, but not t. The extension to the time-dependent case was carried out in papers by Ostrogradsky (published 1848–1850) and Donkin (1854).

EXERCISES

7.1 Obtain the constant of the motion in Exercise 1.14 by the Hamiltonian method.

7.2 Obtain Hamilton's equations for a particle moving in space under an inverse-square-law central force, taking the q_as to be spherical polar coordinates.

7.3 The ends A and B of a thin uniform rod of mass m and length $2a$ can slide freely, A along a smooth horizontal wire OX and B along a smooth vertical wire OZ, with OZ pointing vertically upwards. The wire frame OXZ is made to rotate with constant angular velocity Ω about OZ. Show that if B is above O and the angle OBA is θ, then

$$h = \tfrac{2}{3}ma^2(\dot{\theta}^2 - \Omega^2 \sin^2\theta) + mga\cos\theta$$

and that h is a constant of the motion. Is h equal to the total energy?

What is the Hamiltonian if, instead of being rotated about OZ, the frame OXZ is made to rotate with constant angular velocity ω about the horizontal axis OX? Is it conserved?

7.3 Poisson Brackets

Suppose that $f : P \times \mathbb{R} \to \mathbb{R}$. Then

$$\begin{aligned} \frac{\mathrm{d}f}{\mathrm{d}t} &= \frac{\partial f}{\partial q_a}\dot{q}_a + \frac{\partial f}{\partial p_a}\dot{p}_a + \frac{\partial f}{\partial t} \\ &= \frac{\partial f}{\partial q_a}\frac{\partial h}{\partial p_a} - \frac{\partial f}{\partial p_a}\frac{\partial h}{\partial q_a} + \frac{\partial f}{\partial t}, \end{aligned} \tag{7.9}$$

where we have dropped the tildes on the coordinates q_a, p_a, and t. The combination of q_a and p_a derivatives on the right-hand side is of central importance in the search for coordinate transformations that preserve Hamilton's equations.

Definition 7.5

Let f and g be functions on $P \times \mathbb{R}$. The *Poisson bracket* of f and g is the function

$$[f, g] = \frac{\partial f}{\partial q_a}\frac{\partial g}{\partial p_a} - \frac{\partial f}{\partial p_a}\frac{\partial g}{\partial q_a}.$$

The Poisson bracket is skew-symmetric, that is $[f, g] = -[g, f]$. It also satisfies the *Jacobi identity*. That is, for any three functions f, g, and k,

$$\big[f, [g, k]\big] + \big[g, [k, f]\big] + \big[k, [f, g]\big] = 0.$$

The proof is left as an exercise.

When written in terms of the Poisson bracket, Hamilton's equations become

$$\dot{q}_a = [q_a, h] \quad \text{and} \quad \dot{p}_a = [p_a, h],$$

by taking f in (7.9) to be each of the coordinate functions in turn. This suggests that a good starting point in the search for general transformations that will simplify the equations of motion is to look for coordinate transformations that preserve Poisson brackets.

EXERCISES

7.4. †The motion of a particle in three dimensions under an inverse-square-law force is governed by the Lagrangian

$$L = \tfrac{1}{2}m\dot{\boldsymbol{r}}.\dot{\boldsymbol{r}} + \frac{k}{r},$$

where $r = |\boldsymbol{r}|$. Show by vector methods, or otherwise, that the components J_i, R_i of the two vectors

$$\boldsymbol{J} = m\boldsymbol{r} \wedge \dot{\boldsymbol{r}} \qquad \text{and} \qquad \boldsymbol{R} = \dot{\boldsymbol{r}} \wedge \boldsymbol{J} - \frac{k\boldsymbol{r}}{r}$$

are constants of the motion.

Show that $[J_1, r] = 0$. Find $[J_1, J_2]$ and $[J_1, R_2]$ in terms of components of $\boldsymbol{J}$ and $\boldsymbol{R}$.

7.4 Canonical Transformations

Suppose that new coordinates q_a'', p_a'', and t'' are defined on the extended phase space by expressions of the form

$$q_a'' = q_a''(q, p, t), \quad p_a'' = p_a''(q, p, t), \quad \text{and} \quad t'' = t.$$

Let $f, g : P \times \mathbb{R} \to \mathbb{R}$ and put

$$[f, g]'' = \frac{\partial f}{\partial q_a''}\frac{\partial g}{\partial p_a''} - \frac{\partial g}{\partial q_a''}\frac{\partial f}{\partial p_a''}.$$

We want to understand the conditions under which $[f, g] = [f, g]''$ for every f and g. The key to this problem is the introduction of a third system of coordinates q_a', p_a', t', which is intermediate between the unprimed system q_a, p_a, t and the double-primed system q_a'', p_a'' , t''. It is defined by

$$q_a' = q_a, \quad p_a' = p_a'', \quad t' = t = t''. \tag{7.10}$$

The transformation from the unprimed to the double-primed system is thus broken into two stages. First, there is the transformation from q_a, p_a, t to the primed system q_a', p_a', t', which is of the form

$$q_a = q_a', \quad p_a = p_a(q', p', t'), \quad t = t'. \tag{7.11}$$

Second, there is the transformation from q_a', p_a', t' to q_a'', p_a'', t'', which is of the form

$$q_a'' = q_a''(q', p', t'), \quad p_a'' = p_a', \quad t'' = t'. \tag{7.12}$$

With three coordinate systems in play, we must take care in the interpretation of partial derivatives. We use the familiar device to identify which variables are held fixed. When a partial derivative is taken with respect to a variable of one type, unprimed, primed, or double-primed, the values of the other variables of the same type are held fixed. For example, $\partial/\partial q_1$ is the partial derivative

with respect to q_1 with $q_2, \ldots, q_n$, $p_1, \ldots, p_n$, and t held fixed, while $\partial/\partial q_1'$ is the partial derivative with $q_2', \ldots, q_n'$, $p_1', \ldots, p_n'$, and t' held fixed. With this convention, the expression

$$\frac{\partial}{\partial q_1'}\left(\frac{\partial f}{\partial q_1}\right)$$

is legitimate and unambiguous. To unpack it, first express f as a function of q_a, p_a, and t, and take the partial derivative with respect to q_1. Then express the result as a function of q_a', p_a', and t' and take the partial derivative with respect to q_1', with p_a', t', and the remaining q_a' held fixed. In general,

$$\frac{\partial}{\partial q_1'}\left(\frac{\partial f}{\partial q_1}\right) \neq \frac{\partial}{\partial q_1}\left(\frac{\partial f}{\partial q_1'}\right),$$

so the notation $\partial^2 f/\partial q_1' \partial q_1$ is very dangerous and should be avoided when the two coordinate variables belong to different systems. Of course partial derivatives with respect to coordinates of the same system can be interchanged, so that, for example,

$$\frac{\partial^2 f}{\partial q_1' \partial p_1'} = \frac{\partial^2 f}{\partial p_1' \partial q_1'},$$

and it is safe to denote both sides by $\partial^2 f/\partial q_1' \partial p_1'$.

There is also a technical problem, over which we should take some care. It is not obvious that the primed variables will be a good system of coordinates on $P \times \mathbb{R}$ because it may not be possible to express p_a and q_a as smooth functions of the q_a's, p_a's, and t'. For example, the double-primed system might be defined by

$$q_a'' = p_a, \qquad p_a'' = -q_a,$$

in which case $q_a' = -p_a'$ everywhere in $P \times \mathbb{R}$, and the values of the primed variables cannot be used as unambiguous labels for the points of $P \times \mathbb{R}$. This is a special situation, however. In general, (7.11) and (7.12) will hold and the primed variables can be used as coordinates, in which case we say that the coordinate systems q_a, p_a, t and q_a'', p_a'', t'' are *transversal*.

The implicit function theorem gives as a necessary and sufficient condition for transversality that the two $n \times n$ matrices L and M with entries

$$L_{ab} = \frac{\partial p_b''}{\partial p_a} \quad \text{and} \quad M_{ab} = \frac{\partial q_b}{\partial q_a''}$$

should be nonsingular everywhere.

Proposition 7.6

Suppose that the coordinate system q_a, p_a, t is transversal to the coordinate system q''_a, p''_a, t''. If there exists a function $F : P \times \mathbb{R} \to \mathbb{R}$ such that when F is expressed as a function of q'_a, p'_a, and t',

$$p_a = \frac{\partial F}{\partial q'_a} \quad \text{and} \quad q''_a = \frac{\partial F}{\partial p'_a}, \tag{7.13}$$

then $[f, g] = [f, g]''$ for all f and g.

Proof

Let f and g be two functions on $P \times \mathbb{R}$. When applied to the first transformation (7.11), the chain rule gives

$$\frac{\partial g}{\partial q'_a} = \frac{\partial g}{\partial q_a} + \frac{\partial g}{\partial p_b}\frac{\partial p_b}{\partial q'_a} \quad \text{and} \quad \frac{\partial f}{\partial p'_a} = \frac{\partial f}{\partial p_b}\frac{\partial p_b}{\partial p'_a} \tag{7.14}$$

because

$$\frac{\partial q_b}{\partial q'_a} = \delta_{ab}, \qquad \frac{\partial t}{\partial q'_a} = 0, \qquad \frac{\partial q_b}{\partial p'_a} = 0, \qquad \frac{\partial t}{\partial p'_a} = 0.$$

In the same way, we obtain from the second transformation (4.2.9),

$$\frac{\partial g}{\partial q'_a} = \frac{\partial g}{\partial q''_b}\frac{\partial q''_b}{\partial q'_a} \quad \text{and} \quad \frac{\partial f}{\partial p'_a} = \frac{\partial f}{\partial p''_a} + \frac{\partial f}{\partial q''_b}\frac{\partial q''_b}{\partial p'_a}. \tag{7.15}$$

Hence, by multiplying the first equations in each of (7.14) and (7.15) by $\partial f/\partial p_a$ and summing over a,

$$\frac{\partial f}{\partial p_a}\left(\frac{\partial g}{\partial q_a} + \frac{\partial g}{\partial p_b}\frac{\partial p_b}{\partial q'_a}\right) = \frac{\partial f}{\partial p_a}\frac{\partial g}{\partial q''_b}\frac{\partial q''_b}{\partial q'_a}. \tag{7.16}$$

Similarly, by multiplying the second in each pair by $\partial g/\partial q''_a$ and summing,

$$\left(\frac{\partial f}{\partial p''_a} + \frac{\partial f}{\partial q''_b}\frac{\partial q''_b}{\partial p'_a}\right)\frac{\partial g}{\partial q''_a} = \frac{\partial f}{\partial p_b}\frac{\partial g}{\partial q''_a}\frac{\partial p_b}{\partial p'_a}. \tag{7.17}$$

From (7.13), we have

$$\begin{aligned}
\frac{\partial p_a}{\partial q'_b} - \frac{\partial p_b}{\partial q'_a} &= \frac{\partial^2 F}{\partial q'_a \partial q'_b} - \frac{\partial^2 F}{\partial q'_b \partial q'_a} = 0, \\
\frac{\partial q''_a}{\partial p'_b} - \frac{\partial q''_b}{\partial p'_a} &= \frac{\partial^2 F}{\partial p'_a \partial p'_b} - \frac{\partial^2 F}{\partial p'_b \partial p'_a} = 0,
\end{aligned}$$

and

$$\frac{\partial q''_b}{\partial q'_a} = \frac{\partial^2 F}{\partial q'_a \partial p'_b} = \frac{\partial p_a}{\partial p'_b}. \tag{7.18}$$

So, by subtracting (7.16) from the same equation with f and g interchanged,

$$[f,g] + \left(\frac{\partial p_a}{\partial q'_b} - \frac{\partial p_b}{\partial q'_a}\right)\frac{\partial f}{\partial p_a}\frac{\partial g}{\partial p_b} = [f,g] = \frac{\partial q''_b}{\partial q'_a}\left(\frac{\partial g}{\partial p_a}\frac{\partial f}{\partial q''_b} - \frac{\partial f}{\partial p_a}\frac{\partial g}{\partial q''_b}\right). \quad (7.19)$$

Similarly, from (7.17),

$$[f,g]'' = \frac{\partial p_a}{\partial p'_b}\left(\frac{\partial g}{\partial p_a}\frac{\partial f}{\partial q''_b} - \frac{\partial f}{\partial p_a}\frac{\partial g}{\partial q''_b}\right).$$

Therefore $[f,g] = [f,g]''$ for all f and g. □

The converse is also true. That is, if $[f,g] = [f,g]''$ for all f, g, then there exists a function F such that the two coordinate systems are related by (7.13). This is proved in Section 8.9.

Definition 7.7

The transformation from the coordinates q_a, p_a, t to the coordinates q''_a, p''_a, $t'' = t$ is said to be *canonical* if

$$[f,g] = [f,g]''$$

for every f and g. A function $F : P \times \mathbb{R} \to \mathbb{R}$ such that (7.13) holds is called a *generating function* for the transformation and the new coordinates are said to be *canonical*.

The sense in which F generates the coordinate transformation is not entirely straightforward. The first point to note is that it is not true that one can construct a canonical transformation by picking an arbitrary function F on $P \times \mathbb{R}$. In order to use (7.13) to obtain the relations between the new double-primed coordinates and the old unprimed coordinates, one must first express F in terms of the intermediate, primed, system of coordinates, and, of course, that is not possible unless one knows the transformation in advance. Instead, one must set about things in a slightly different order.

(1) Pick a family of functions on $C \times \mathbb{R}$

$$S = S(q_1, \ldots, q_n, k_1, \ldots, k_n, t)$$

labelled by n parameters $k_1, k_2, \ldots, k_n$.

(2) Construct functions $p'_1, \ldots, p'_n$ on $P \times \mathbb{R}$ by solving the system of equations

$$p_a = \left.\frac{\partial S}{\partial q_a}\right|_{k_a = p'_a} \quad (7.20)$$

for the p'_as as functions of the q_as, p_as, and t. Take care over the partial differentiation: on the right-hand side, differentiate S with respect to q_a, holding the other q coordinates, t, and the parameters k_a constant, and *then* substitute $p'_1, \ldots, p'_n$ for $k_1, \ldots, k_n$ in the resulting expression.

(3) Define the intermediate coordinate system q'_a, p'_a, t' by putting $q'_a = q_a$ and $t' = t$ and define $F : P \times \mathbb{R} \to \mathbb{R}$ by setting

$$F(q'_1, \ldots, q'_n, p'_1, \ldots, p'_n, t') \tag{7.21}$$

equal to the value of S at $q_a = q'_a$, $k_a = p'_a$, $t = t'$.

(4) Construct the final coordinate system q''_a, p''_a, t'' by putting

$$p''_a = p'_a, \quad q''_a = \frac{\partial F}{\partial p'_a}, \quad t'' = t' = t.$$

The three coordinate systems, the original unprimed system, the intermediate primed system, and the final double-primed system-are then related by (7.10) and (7.13), so that the transformation from q_a, p_a, t to q''_a, p''_a, t'' is canonical. There is a sense in which F and S are the same, but it is not the obvious one because S is a family of functions on $C \times \mathbb{R}$, while F is a function on $P \times \mathbb{R}$. The 'generating' is really being done by S rather than by F.

There is again a problem over 'transversality'. The second step will work only if one can, in fact, invert (7.20). By the implicit function theorem, this requires that the $n \times n$ matrix with entries

$$\frac{\partial^2 S}{\partial q_a \partial k_b}$$

should be nonsingular.

Example 7.8

With $n = 2$, take

$$S = \tfrac{1}{2} k_1 q_1^2 + k_2 q_2.$$

Then $p_1 = p'_1 q_1$ and $p_2 = p'_2$. Hence $p'_1 = p_1/q_1$ and $p'_2 = p_2$. In terms of the primed coordinates

$$F = \tfrac{1}{2} p'_1 (q'_1)^2 + p'_2 q'_2.$$

The double-primed coordinates are

$$\begin{aligned} p''_1 &= p'_1 & \qquad p''_2 &= p'_2 \\ q''_1 &= \frac{\partial F}{\partial p'_1} = \tfrac{1}{2}(q'_1)^2 & \qquad q''_2 &= \frac{\partial F}{\partial p'_2} = q'_2. \end{aligned}$$

Hence

$$p_1'' = p_1/q_1, \quad p_2'' = p_2', \quad q_1'' = \tfrac{1}{2}q_1^2, \quad q_2'' = q_2.$$

Note that $[q_1'', p_1'']'' = 1$, which is consistent with

$$[\tfrac{1}{2}q_1^2, p_1/q_1] = 1.$$

Example 7.9

The identity transformation has the generating function $F = p_a q_a = p_a' q_a'$, or, equivalently, $S = k_a q_a$.

Example 7.10

The time-dependent transformation

$$q'' = q + t, \quad p'' = p, \quad t'' = t$$

is canonical, with generating function $F = p(q + t) = p'(q' + t')$.

EXERCISES

7.5. With $n = 1$, find the transformation generated by $S = \frac{1}{6}k^2q^3$ and check that it is canonical.

7.6. With $n = 1$, show that canonical transformations preserve volumes in $P \times \mathbb{R}$; that is

$$\frac{\partial(q'', p'', t'')}{\partial(q, p, t)} = 1.$$

7.7. With $n = 1$, show that if $S = s(kq)$, where s is a function of a single variable, then $p''q'' = pq$.

7.8. With $n = 2$, find S such that $q_1'' = q_2$, $q_2'' = q_1$, $p_1'' = p_2$, $p_2'' = p_1$.

7.9. Show that if $q_a'' = q_a''(q)$ and

$$p_a = \frac{\partial q_b''}{\partial q_a} p_b'',$$

then the transformation from q_a, p_a, t to q_a'', p_a'', $t'' = t$ is canonical. Show that a generating function is $F = q_a'' p_a''$.

7.5 Infinitesimal Canonical Transformations

Consider a family of canonical transformations which are close to the identity, with $F = p'_a q'_a + \varepsilon f(q', p', t')$, where ε is a small parameter. Then

$$p_a = p'_a + \varepsilon \frac{\partial f}{\partial q'_a} \quad \text{and} \quad q''_a = q'_a + \varepsilon \frac{\partial f}{\partial p'_a},$$

so that

$$p''_a = p_a - \varepsilon \frac{\partial f}{\partial q'_a} \quad \text{and} \quad q''_a = q_a + \varepsilon \frac{\partial f}{\partial p'_a}.$$

Thus the new double-primed coordinates differ from the old unprimed coordinates by terms of order ε. Hence

$$\frac{\partial f}{\partial q'_a} = \frac{\partial f}{\partial q_a} + O(\varepsilon) \quad \text{and} \quad \frac{\partial f}{\partial p'_a} = \frac{\partial f}{\partial p_a} + O(\varepsilon).$$

By ignoring terms of order ε^2, therefore,

$$p''_a = p_a - \varepsilon \frac{\partial f}{\partial q_a} \quad \text{and} \quad q''_a = q_a + \varepsilon \frac{\partial f}{\partial p_a}. \tag{7.22}$$

For small ε, we can construct the transformation directly for any choice of f without going through all the steps in the previous section. So it *is* true that an arbitrary function f on $P \times \mathbb{R}$ generates an *infinitesimal* canonical transformation.

Note that the derivatives of f that appear in (7.22) are the same as those in Hamilton's equations. If we take $f = h$, then the infinitesimal canonical transformation is given by moving along the trajectories of the system in $P \times \mathbb{R}$ through $\delta t = \varepsilon$.

7.6 The Hamilton–Jacobi Equation

We now consider how Hamilton's equations behave under the canonical transformation with generating function $F(q'_a, p'_a, t')$. As before, q_a, p_a, t is the original coordinate system, q''_a, p''_a, t'' is the transformed system, and q'_a, p'_a, t' is the intermediate system defined by (7.10).

Proposition 7.11

For any function f on $P \times \mathbb{R}$,

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial t''} + [f, g] \quad \text{where} \quad g = \frac{\partial F}{\partial t'}.$$

Proof

We use the notation in the proof of Proposition 7.6. For any function f on $P \times \mathbb{R}$,

$$\frac{\partial f}{\partial t'} = \frac{\partial f}{\partial p_a}\frac{\partial p_a}{\partial t'} + \frac{\partial f}{\partial t} = \frac{\partial f}{\partial p_a}\frac{\partial g}{\partial q'_a} + \frac{\partial f}{\partial t}$$

because $p_a = \partial F/\partial q'_a$. Similarly, from (7.12),

$$\frac{\partial f}{\partial t'} = \frac{\partial f}{\partial q''_a}\frac{\partial q''_a}{\partial t'} + \frac{\partial f}{\partial t''} = \frac{\partial f}{\partial q''_a}\frac{\partial g}{\partial p'_a} + \frac{\partial f}{\partial t''}$$

because $q''_a = \partial F/\partial p'_a$. Hence

$$\begin{aligned}\frac{\partial f}{\partial t} - \frac{\partial f}{\partial t''} &= \frac{\partial f}{\partial q''_a}\frac{\partial g}{\partial p'_a} - \frac{\partial f}{\partial p_a}\frac{\partial g}{\partial q'_a} \\ &= \frac{\partial f}{\partial q''_a}\frac{\partial g}{\partial p_b}\frac{\partial p_b}{\partial p'_a} - \frac{\partial f}{\partial p_a}\frac{\partial g}{\partial q''_b}\frac{\partial q''_b}{\partial q'_a} \\ &= \frac{\partial q''_b}{\partial q'_a}\left(\frac{\partial g}{\partial p_a}\frac{\partial f}{\partial q''_b} - \frac{\partial f}{\partial p_a}\frac{\partial g}{\partial q''_b}\right) \\ &= [f, g],\end{aligned}$$

by using (7.10) and (7.11) to get the second equality, (7.18) to get the third, and (7.19) to get the last. □

Now consider the time derivative of f along the trajectories in $P \times \mathbb{R}$. We know from (7.9) that

$$\frac{\mathrm{d}f}{\mathrm{d}t} = [f, h] + \frac{\partial f}{\partial t}.$$

But $[f, h] = [f, h]''$ because the transformation is canonical. Hence, by applying the proposition,

$$\frac{\mathrm{d}f}{\mathrm{d}t} = [f, h] + [f, g] + \frac{\partial f}{\partial t''} = [f, h'']'' + \frac{\partial f}{\partial t''},$$

where $h'' = h + g$. In particular

$$\frac{\mathrm{d}q''_a}{\mathrm{d}t} = \frac{\partial h''}{\partial p''_a}, \qquad \frac{\mathrm{d}p''_a}{\mathrm{d}t} = -\frac{\partial h''}{\partial q''_a}.$$

Therefore the equations of motion of the system in the new coordinates are again of Hamilton's form, but with Hamiltonian $h'' = h + g$.

Canonical transformations offer wide scope for simplifying the equations of motion. In fact, if we can find a transformation such that $h'' = 0$, then the trajectories in the extended phase space will be given by constant values of the coordinates q''_a and p''_a, so the dynamical problem will be trivial.

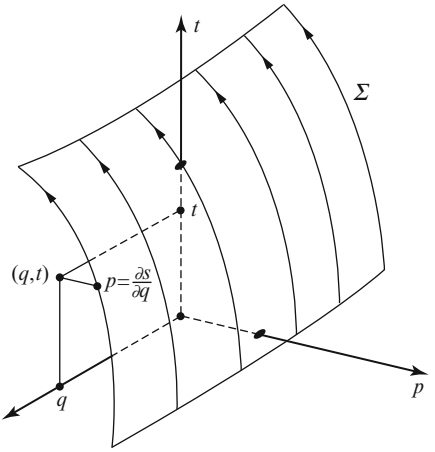

Figure 7.1

Suppose that the transformation is constructed from a family of functions $S(q, k, t)$ by the sequence of operations set out in Section 7.4. Fix, for the moment, the values of the parameters $k_1, \ldots, k_n$ and let Σ denote the set of points in $P \times \mathbb{R}$ at which

$$p_a = \frac{\partial S}{\partial q_a}.$$

Then Σ is also the subset of $P \times \mathbb{R}$ on which $p'_a = k_a$, or equivalently, $p''_a = k_a$.

The q_as and t can be used as coordinates on Σ with (q_a, t) labelling the point with unprimed coordinates

$$q_a, \quad p_a = \frac{\partial S}{\partial q_a}, \quad t$$

(Figure 7.1). On Σ, the Hamiltonian h is given as a function of q_a and t by

$$h = h\left(q_1, \ldots, q_n, \frac{\partial S}{\partial q_1}, \cdots \frac{\partial S}{\partial q_n}, t\right),$$

or in other words, by substituting $\partial S/\partial q_a$ for p_a in the expression for h in the unprimed coordinate system q_a, p_a, t on $P \times \mathbb{R}$. Similarly, g is given as a function of q_a and t on Σ by

$$g = \frac{\partial S}{\partial t},$$

because it follows from (7.21) that differentiating F with respect to t' with p'_a and q'_a held fixed is equivalent to differentiating S with respect to t with q_a and

k_a held fixed. Therefore h'' vanishes on Σ if S satisfies the partial differential equation

$$h\left(q_1, \ldots, q_n, \frac{\partial S}{\partial q_1}, \ldots, \frac{\partial S}{\partial q_n}, t\right) + \frac{\partial S}{\partial t} = 0. \tag{7.23}$$

This is the *Hamilton–Jacobi equation*.

As the k_a's vary, the corresponding subsets fill out $P \times \mathbb{R}$. We can conclude, therefore, that a family of solutions $S(q, k, t)$ of the Hamilton–Jacobi equation gives rise to a canonical transformation for which h'' vanishes.

The catch is that it is generally much harder to solve the Hamilton–Jacobi equation, which is a first-order nonlinear partial differential equation, than it is to solve Hamilton's equations. In fact, except in very special systems, the only generally applicable method is the 'method of characteristics', which is based on the reverse construction, the first step being the return to Hamilton's equations.

For this reason, canonical transformations and the Hamilton–Jacobi equation are rarely of practical use for finding analytic solutions to mechanical problems. They are, nevertheless, very important for a number of reasons. For example, they provide the connecting link between classical and quantum mechanics, through the WKB approximation. They provide significant insights into the qualitative behaviour of systems that defy analytical treatment. Finally, they play a key part in the theory of partial differential equations.

Even when one can find an n-parameter family of solutions, the best way to proceed is not always to carry out the canonical transformation explicitly. If Σ is the subset of $P \times \mathbb{R}$ corresponding to a particular set of values for the parameters k_a, then

$$p_a'' = p_a' = k_a = \text{constant}$$

on Σ. But the trajectories in $P \times \mathbb{R}$ are given in the double-primed coordinates by constant values of q_a'' and p_a''. It follows that a trajectory that passes through one point of Σ lies in Σ. Moreover, the coordinates q_a'' are also constant along the trajectories. On Σ, q_a'' is given as a function of the q_a's and t by

$$q_a'' = \frac{\partial S}{\partial k_a},$$

by combining (7.13) and (7.21). Hence the trajectories that lie in Σ are the curves given in the coordinate system q_a, t by

$$\frac{\partial S}{\partial k_a} = \text{constant}. \tag{7.24}$$

By rewriting these in the form $q_a = q_a(t)$, one obtains a family of solutions of the original dynamical problem. The members of the family are labelled by the constants on the right-hand side.

Example 7.12 (The Harmonic Oscillator)

Consider the harmonic oscillator Hamiltonian

$$h = \tfrac{1}{2}p^2 + \tfrac{1}{2}q^2.$$

The Hamilton–Jacobi equation is

$$\frac{\partial S}{\partial t} + \tfrac{1}{2}\left(\frac{\partial S}{\partial q}\right)^2 + \tfrac{1}{2}q^2 = 0.$$

It has a one-parameter family of solutions $S = -kt + s(q,k)$, where

$$s(q,k) = \int \sqrt{2k - q^2}\,\mathrm{d}q$$

and the constant k labels the different solutions in the family.

We construct a canonical transformation from S by applying the instructions set out in Section 7.4. The intermediate coordinates are $q' = q$, $t' = t'$, and p', where p' is determined as a function of q and p by solving

$$\begin{aligned} p &= \left.\frac{\partial S}{\partial q}\right|_{k=p'} \\ &= \sqrt{2p' - q^2}. \end{aligned}$$

The generating function is

$$F = S|_{k=p'} = -p't' + \int (2p' - q'^2)^{1/2}\,\mathrm{d}q'.$$

Therefore the double-primed coordinates (Figure 7.2) are

$$\begin{aligned} p'' &= p' = \tfrac{1}{2}(p^2 + q^2) \\ q'' &= \frac{\partial F}{\partial p'} = -t' + \int \frac{\mathrm{d}q'}{\sqrt{2p' - q'^2}} = -t + \sin^{-1}(q/\sqrt{2p'}). \end{aligned}$$

In this case, the extended phase space is $\mathbb{R}^3$, with coordinates q, p, t, and it is easy to solve the equations of motion. The dynamical trajectories are the helices

$$q = \sqrt{2E}\sin(t + \varepsilon), \qquad p = \sqrt{2E}\cos(t + \varepsilon)$$

where ε and E are constants.

Let r and θ denote the polar coordinates defined by

$$p = r\cos\theta, \qquad q = r\sin\theta.$$

Then the trajectories can be written more simply as

$$r = \sqrt{2E}, \qquad \theta = t + \varepsilon.$$

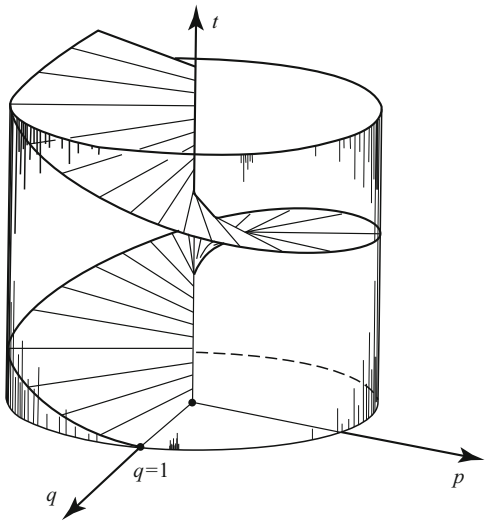

Figure 7.2 The cylinder is the surface $p'' = \frac{1}{2}$ and the helicoid is $q'' = \pi/2$. Their intersection, the helix through $(q, p, t) = (1, 0, 0)$, is a trajectory of the system in $P \times \mathbb{R}$.

In terms of r and θ, the new canonical coordinates are

$$p'' = \tfrac{1}{2}r^2, \quad q'' = \theta - t, \qquad t'' = t.$$

As the theory predicts, the dynamical trajectories are given by constant values of p'' and q''.

Example 7.13 (Inverse-square-law central force)

In the inverse-square-law problem, Example 7.1, the Hamiltonian is

$$h = \frac{1}{2m}\left(p_1^2 + \frac{p_2^2}{q_1^2}\right) - \frac{\lambda m}{q_1}$$

where $q_1 = r$ and $q_2 = \theta$. So the Hamilton–Jacobi equation is

$$\frac{\partial S}{\partial t} + \frac{1}{2m}\left(\frac{\partial S}{\partial q_1}\right)^2 + \frac{1}{2mq_1^2}\left(\frac{\partial S}{\partial q_2}\right)^2 - \frac{\lambda m}{q_1} = 0. \tag{7.25}$$

We solve this by separating the variables. That is, we look for solutions of the form

$$S = \tau(t) + s_1(q_1) + s_2(q_2).$$

On substituting into (7.25),

$$-\frac{\mathrm{d}\tau}{\mathrm{d}t} = \frac{1}{2m}\left(\frac{\mathrm{d}s_1}{\mathrm{d}q_1}\right)^2 + \frac{1}{2mq_1^2}\left(\frac{\mathrm{d}s_2}{\mathrm{d}q_2}\right)^2 - \frac{\lambda m}{q_1}.$$

The left-hand side is a function of t alone and the right-hand side of q_1 and q_2 alone. Therefore both sides must be equal to a constant, which we shall denote by k_1.

We then have $\tau = -k_1 t + \text{constant}$, and

$$\frac{\mathrm{d}s_2}{\mathrm{d}q_2} = q_1\left[2mk_1 - \left(\frac{\mathrm{d}s_1}{\mathrm{d}q_1}\right)^2 + \frac{2m\lambda m^2}{q_1}\right]^{1/2}.$$

Again both sides must be equal to a constant, which we shall denote k_2. It follows that

$$s_2 = k_2 q_2 + \text{constant}, \qquad \frac{\mathrm{d}s_1}{\mathrm{d}q_1} = f(q_1)$$

where

$$f(q_1) = \left(\frac{2\lambda m^2}{q_1} - \frac{k_2^2}{q_1^2} + 2k_1 m\right)^{1/2}.$$

Therefore

$$S(q,k,t) = -k_1 t + k_2 q_2 + \int f\,\mathrm{d}q_1$$

is a two-parameters family of solutions. There is an additional additive constant of integration, which has been dropped because it has no effect on the transformation generated by S.

Rather than look directly at the canonical transformation, which is complicated, we shall use (7.24) to deduce that the dynamical trajectories are given by

$$\frac{\partial S}{\partial k_1} = -t + \int \frac{m\,\mathrm{d}q_1}{f} = \text{constant} \tag{7.26}$$

and

$$\frac{\partial S}{\partial k_2} = q_2 - \int \frac{k_2\,\mathrm{d}q_1}{q_1^2 f} = \text{constant} \tag{7.27}$$

for the different values of k_1 and k_2.

These become more familiar when we restore $q_1 = r$ and $q_2 = \theta$ and remember that

$$\tfrac{1}{2}m\left(\dot{r}^2 + \frac{j^2}{m^2 r^2}\right) - \frac{\lambda m}{r} = E$$

where E and $j = mr^2\dot{\theta}$ are constant. If we identify k_2 with j and k_1 with E, then

$$\dot{r} = \frac{1}{m}f(r), \qquad \frac{\mathrm{d}\theta}{\mathrm{d}r} = \frac{j}{r^2 f(r)},$$

which integrate to give (7.26) and (7.27).

The Hamilton–Jacobi theory has provided a sophisticated derivation of a familiar and straightforward result, which is typical of its track record as an analytical technique. There are occasional exceptions, however. For example, it was used to find the orbits of bodies orbiting a rotating black hole in general relativity [2].

EXERCISES

7.10. Consider a system with one degree of freedom and Hamiltonian $h(q, p, t)$. Show that the dynamical trajectories in $PT = \mathbb{R}^3$ are tangent to the vector field

$$\boldsymbol{x} = \frac{\partial h}{\partial p}\boldsymbol{i} - \frac{\partial h}{\partial q}\boldsymbol{j} + \boldsymbol{k}$$

where $\boldsymbol{i}$, $\boldsymbol{j}$, and $\boldsymbol{k}$ are unit vectors along the q, p, and t axes.

Let Σ be a surface in $\mathbb{R}^3$ given by

$$p = \frac{\partial S}{\partial q}$$

where $S = S(q, t)$. Show that if S is a solution of the Hamilton–Jacobi equation, then $\boldsymbol{x}$ is tangent to Σ. Show conversely that if $\boldsymbol{x}$ is tangent to Σ, then

$$\frac{\partial}{\partial q}\left[\frac{\partial S}{\partial t} + h\left(q, \frac{\partial S}{\partial q}, t\right)\right] = 0.$$

7.11. Solve the Hamilton–Jacobi equation by separating the variables for a particle moving in space under an inverse-square-law central force, taking the q_as to be spherical polar coordinates.

7.12. A particle P of mass m is moving in the plane under the influence of two inverse-square-law forces: $\lambda(PA)^{-2}$ directed towards the point A and $\lambda(PB)^{-2}$ directed towards the point B, where A and B are separated by a distance $2b$. Solve the Hamilton–Jacobi equation in the coordinates θ and φ, where

$$2b\cosh\varphi = PA + PB, \qquad 2b\cos\theta = PA - PB.$$

8
Geometry of Classical Mechanics

8.1 Coordinate Choices

The analysis of the dynamics of a mechanical system begins with the introduction of generalized coordinates to label the configurations. There is a great deal of freedom in the choice of coordinates and a good choice can greatly simplify the work. It is rarely the case, however, that a single coordinate system can be used for all possible configurations. For example, the polar coordinates used to label the configurations of the spherical pendulum (p. 67) are *singular* when the rod is vertical, in the sense that these configurations do not determine unique values of θ and φ. When $\theta = 0$ or $\theta = \pi$, we have

$$(x, y, z) = (0, 0, 1), \qquad \text{or} \qquad (x, y, z) = (0, 0, -1)$$

irrespective of the value of φ. In these special configurations, the coordinate labels are not unique. A similar problem arises when Euler angles (p. 132) are used to label the orientations of a rigid body. When the z-axes of the inertial frame and of the body frame are aligned, which happens when $\theta = 0$ or $\theta = \pi$, the orientation is completely determined by the values $\varphi + \psi$. It is unchanged when φ and ψ are replaced by $\varphi + k$ and $\psi - k$, respectively. This is the 'gimbal lock' problem.

The configuration space is not exactly synonymous with the range of possible values of the coordinate n-tuple $(q_1, q_2, \ldots, q_n)$. The former is the set of all possible configurations of the system. The latter is a subset of $\mathbb{R}^n$ and in general is identified only with a subset of the configuration space. In the spherical

N.M.J. Woodhouse, *Introduction to Analytical Dynamics: New Edition*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-1-84882-816-2_8,

pendulum, for example, the configuration space is a complete sphere while the polar coordinates identify the interior of a rectangle,

$$\{(\theta, \varphi) \,:\, 0 < \theta < \varphi, 0 < \varphi < 2\pi\} \subset \mathbb{R}^2,$$

with the complement of a great semi-circle joining the two vertical configurations. In geographical language, one must exclude the north pole, where the rod points vertically upwards, the south pole, where it points vertically downwards, and the Greenwich meridian, on which $\varphi = 0$. It is an ancient problem for cartographers that one cannot identify the whole sphere with a subset of the plane by a one-to-one mapping without tearing. Charts are intrinsically local.

A full analysis of a mechanical system should therefore allow for the possibility of using different coordinate systems in different parts of the configuration space. One should think in terms of covering the space with a mosaic of 'coordinate charts'.

It is useful to develop a language in which different coordinate systems can be considered simultaneously and in which the dynamics can be modelled in a way that is independent of the particular choice of coordinates, preferably without mentioning coordinates at all. Fortunately, this does not have to be done from scratch: differential geometry and differential topology provide exactly what is needed in the language of differentiable manifolds. This allows a very precise clarification of the relationship between the global structure of the configuration space and the choice of local coordinates. It also gives new insights into the mathematical structures underlying some of the theory developed in the previous chapters. Not surprisingly, given the power of the analogies above, the language borrows extensively from the vocabulary of geography and map-making.

8.2 Manifolds

In the language of differential geometry, a mechanical configuration space is a *differentiable manifold*.[1] The definition of a manifold captures the idea that the coordinate systems are local, and ties down the permitted transformations between local coordinates.[2] There are many possibilities, but we shall only

[1] The following paragraphs are adapted from [15]. It is not surprising that they also fit seamlessly here because the mathematical theory needed to deal with coordinate transformations in space-time are exactly the same as those we need for the same task in classical mechanics.

[2] In relativity and many geometric contexts, it is conventional to write coordinates with upper rather than lower indices, so that one can make a systematic distinction between the components of covariant and contravariant tensors. It is not particu-

allow smooth, that is infinitely differentiable, transformations. Our manifolds are therefore of class C^∞.

Definition 8.1

An n-dimensional *differentiable manifold* is

(1) a connected Hausdorff topological space M, together with

(2) a collection of *charts* or *coordinate patches* (U, x_a), where $U \subset M$ is an open set and the x_as are n functions on U such that the map

$$\boldsymbol{x} : U \to \mathbb{R}^n : m \mapsto \big(x_1(m), x_2(m), \ldots, x_{n-1}(m)\big)$$

is a homeomorphism from U to an open subset $V \subset \mathbb{R}^n$.

Two conditions must hold. First, every point of M must lie in a coordinate patch. Second, if (U, x_a) and $(\tilde{U}, \tilde{x}_a)$ are charts, then the $\tilde{x}_a$s can be expressed as functions of the x_as on the intersection. It is required that

$$\boldsymbol{x}(U \cap \tilde{U}) \to \tilde{\boldsymbol{x}}(U \cap \tilde{U}) : (x_1, \ldots, x_n) \mapsto (\tilde{x}_1, \ldots, \tilde{x}_n)$$

should be infinitely differentiable and one-to-one, with

$$\det\left(\frac{\partial \tilde{x}_a}{\partial x_b}\right) \neq 0. \tag{8.1}$$

The qualification "differentiable" will often be omitted as understood.

The topological condition on M is required to rule out pathological behaviour. In fact further technical conditions, such as 'paracompactness', are needed to get sensible models. We should also specify completeness for the atlas (the set of charts). We do not dwell on such matters here because they play no critical part in the elementary development of the theory.

Because M is a topological space, it makes sense to talk about continuous functions on M. These are the continuous maps $f : M \to \mathbb{R}$. The additional structure of coordinate charts and infinitely differentiable coordinate transformations picks out a sub-class of *smooth* or infinitely differentiable functions. These are maps $f : M \to \mathbb{R}$ that can be expressed everywhere as infinitely differentiable functions of the coordinates. The set of all smooth functions on

larly helpful to do that in a dynamical context because we are not making use of raising and lowering operations and so there is little to gain from using notation that easily confuses indices with exponents. Also the standard conventions become rather awkward when applied to momentum phase space. Momenta are covector components, so should have lower indices, but also coordinates on the cotangent bundle, and so should have upper indices.

M is denoted by $C^\infty(M)$. One can add and multiply elements of $C^\infty(M)$. In formal terms, $C^\infty(M)$ has the structure of a ring. It includes the constant functions, but it is not immediately obvious that it includes anything else. In fact it is possible to construct a smooth function that takes any given value at a given point m and which vanishes outside some given open neighbourhood of m. This may seem a somewhat technical point to worry about, but it is not true for holomorphic functions and complex manifolds, which are defined in the same way, but with complex coordinates and holomorphic coordinate transformations.

It also makes sense to talk about smooth maps between manifolds. A continuous map $\rho : M' \to M$ from one manifold to another is *smooth* if in local coordinates on the two manifolds, the coordinates x_a of $\rho(m')$ are smooth functions of the coordinates x'_a of $m' \in M'$. It is a *diffeomorphism* if it is bijective and the determinant of the Jacobian matrix

$$\left(\frac{\partial x_a}{\partial x'_b}\right)$$

is nonzero everywhere that it is defined for any pair of coordinate systems. This can only be the case if the manifolds have the same dimension, so that there are as many x'_as as x_as and the matrix is square. If there is a diffeomorphism between them, then M and M' are *diffeomorphic*, which means that they are indistinguishable as manifolds.

The manifolds that we consider are built from basic examples by two constructions.

Submanifolds

Let M' be a subset of an n-dimensional manifold M. Suppose that there is a positive integer $n' \leq n$ and that for each $m \in M'$, there is a chart (U, x_a) on M such that $m \in U$ and $M' \cap U$ is the subset on which

$$x_a = 0 \qquad a = n' + 1, \ldots, n. \tag{8.2}$$

Then M' is said to be an (embedded) *submanifold* of M. A submanifold is a manifold, with the first n' coordinates in these special charts providing the coordinate labels.

Products

If M and $\tilde{M}$ are manifolds, of dimensions n and $\tilde{n}$, then the Cartesian product $M \times \tilde{M}$ is a manifold of dimension of $n + \tilde{n}$. We construct a coordinate system

in a neighbourhood of $(m, \tilde{m}) \in M \times \tilde{M}$ by picking systems of local coordinates x_a in a neighbourhood of $m \in M$ and $\tilde{x}_a$ in a neighbourhood of $\tilde{m} \in \tilde{M}$, and labelling the points of the product by $(x_1, \ldots, x_n, \tilde{x}_1, \ldots, \tilde{x}_n)$.

EXERCISES

8.1. Show that the sphere in $\mathbb{R}^3$ with equation

$$x^2 + y^2 + y^2 = 1$$

is a manifold by defining suitable charts. Show that it is a submanifold of $\mathbb{R}^3$ by using spherical polar coordinates.

8.3 Tangent Vectors

The configuration space of a mechanical system is a differentiable manifold. Each point represents a possible configuration of the system. The corresponding phase space is also a manifold, of twice the dimension. Its points represent particular configurations in particular states of motion. They are labelled by the generalized coordinates $q_1, \ldots, q_n$ of the configuration together with the generalized velocities $v_1, \ldots, v_n$ of the motion. In a short time δt, the configuration changes from q_a to $q_a + \delta t\, v_a$, so it is natural to think of the v_as as the components of a *vector* v, pointing from the first configuration to a nearby one.

This idea makes sense in a general setting and finds its expression in the definition of a *tangent vector* at a point of a manifold. A tangent vector should be pictured as an arrow pointing from one point on the manifold to a nearby one. As in Euclidean space, one specifies a vector v by giving its *components* v_a in a coordinate system. Their key property, on which the general definition is based, is their behaviour under change of coordinate system. In configuration space, when the q_as are replaced by a new coordinate system $\tilde{q}_a$, corresponding small displacements in the coordinates are related by

$$\delta\tilde{q}_a = \frac{\partial \tilde{q}_a}{\partial q_b} \delta q_b.$$

It is this behaviour that is built into the following.

Definition 8.2

Let M be an n-dimensional manifold and let $m \in M$. A *tangent vector* v at m is an object that assigns a set of n components v_a to each coordinate system

x_a on a neighbourhood of m, subject to the transformation rule

$$\tilde{v}_a = \frac{\partial \tilde{x}_a}{\partial x_b} v_b \tag{8.3}$$

under change of coordinates, with the partial derivatives evaluated at m.

Two vectors at a point can be added by adding their components in a coordinate system. Because the transformation rule is linear, the result is the same whichever coordinate system is used. A tangent vector v can also be multiplied by a scalar $\lambda \in \mathbb{R}$ by multiplying its components by λ. Again the product λv is well defined independently of the choice of coordinates by the linearity of the transformation. With these operations, the set of tangent vectors at m form a vector space, which is called the *tangent space at* m and is denoted by $T_m M$. Each choice of local coordinates gives an identification of $T_m M$ with $\mathbb{R}^n$, so the tangent space is n dimensional.

From this definition, it is a simple step to the notion of a vector field on a manifold, or a subset of a manifold. It is a map v that assigns an element $v(m) \in T_m M$ to each $m \in M$, with the property that the components of v at each point are smooth functions of the coordinates in any local coordinate system. Of course if they are smooth functions in one chart, then they are also smooth functions in the overlap of any overlapping chart. The set of all vector fields on M is denoted by $\mathfrak{X}(M)$, and the set of vector fields on $U \subset M$ by $\mathfrak{X}(U)$. If u and v are vector fields, then so is their sum $u + v$, which is formed by adding their components at each point. We can also multiply a given vector field v by a smooth function f. In formal terms, $\mathfrak{X}(M)$ is a module over the ring of smooth functions.

Vectors as Differential Operators

Given a vector field v and a smooth function f, we define a second smooth function $v(f)$, called the *derivative of* f *along* v, by

$$v(f) = v_a \frac{\partial f}{\partial x_a} \tag{8.4}$$

in local coordinates. The transformation rule for vector components combined with the chain rule ensures that the right-hand side is independent of the choice of coordinate system. It has two important properties

(1) $v(rf + sg) = rv(f) + sv(g)$ for all $f, g \in C^\infty(M)$, $r, s \in \mathbb{R}$,

(2) $v(fg) = fv(g) + v(f)g$ for all $f, g \in C^\infty(M)$.

The first property is *linearity*. The second is the *Leibniz property*. It is an immediate consequence of the Leibniz rule for partial derivatives, and it justifies the use of the term "derivative". In fact these properties can be taken as the basis of an alternative and more elegant definition of a vector field.

Basis Vectors

Given a coordinate system x_a, we can define n basis vector fields, that is vector fields that form a basis for the tangent space at each point of the coordinate patch, by taking the ath basis vector field to have components in the coordinate system

$$(0, \ldots, 0, 1, 0, \ldots, 0)$$

with a 1 in the ath place and zeros elsewhere. The derivative of f along the vector field is simply the partial derivative

$$\frac{\partial f}{\partial x_a}$$

with respect to the ath coordinate. For this reason, the ath basis vector field is denoted by ∂_a, in a notation that incorporates the view that a vector field is really the same thing as the derivative along a vector field.

Trajectories

A *trajectory* in a manifold M is a map $\gamma : I \to \mathbb{M}$, where $I \subset \mathbb{R}$ is an open interval, given in local coordinates by n smooth functions

$$s \mapsto x_a(s)$$

of a parameter s. The *tangent vector* at $\gamma(s)$ is the vector v with components

$$v_a = \frac{\mathrm{d}x_a}{\mathrm{d}s}.$$

If the map is given by smooth functions in one chart, then this is also true for any other chart, in the overlap of the charts. The chain rule ensures that the v_as obey the transformation rule for vector components.

Submanifolds

Suppose that $M' \subset M$ is a submanifold and that $m \in M'$. Then the tangent space $T_m M'$ is identified in a natural way with a subspace of $T_m M$, so we can write $T_m M' \subset T_m M$. The identification is made by mapping the tangent vector at m to a trajectory $\gamma : I \to M'$ through m to its tangent vector at m as a trajectory in M. Alternatively, in a special chart on M in which M' is given by (8.2), we identify the vector $v \in T_m M'$ with components

$$v_1, \ldots, v_{n'}$$

with the vector in $T_m M$ with components $v_1, \ldots, v_{n'}, 0, \ldots, 0$.

Derived Map

A smooth map $\rho : M' \to M$ from one manifold to another maps a small displacement on the first manifold to a small displacement on the second, and therefore induces a *derived map*

$$\rho_* : T_{m'} M' \to T_{\rho(m')} M$$

for each $m' \in M'$. In coordinates the derived map sends the vector $v' \in T_{m'} M'$ with components $v'_1, \ldots, v'_{n'}$ to the vector $v \in T_{\rho(m')} M$ with components

$$v_a = \sum_{b=1}^{n'} \frac{\partial x_a}{\partial x'_b} v'_b \qquad a = 1, \ldots, n.$$

The transformation rule ensures that the v_as behave correctly under change of coordinates on M' and on M.

Lie Bracket

Suppose that u, v are vector fields. Then a third vector field $[u, v]$ can be defined by specifying its components as

$$[u, v]_a = u_b \frac{\partial v_a}{\partial x_b} - v_b \frac{\partial u_a}{\partial x_b}. \tag{8.5}$$

For any smooth function f, we have

$$[u, v](f) = u\big(v(f)\big) - v\big(u(f)\big).$$

The vector $[u, v]$ is called the *Lie bracket* of u and v. The last equation can be used as the definition, if one first establishes that it has the two properties of the derivative along a vector field.

EXERCISES

8.2. Show that the definition of $v(f)$ is independent of the choice of vector fields.

8.3. Show that the quantities on the right-hand side of (8.5) have the correct transformation rule for vector components.

8.4. Show that if $\rho : M' \to M$ is a smooth map, then $\rho_*([u', v']) = [\rho_* u', \rho_* v']$ for any vector fields u', v' on M'.

8.4 The Tangent Bundle

At each point of a manifold M, there is a tangent vector space. The union of all the tangent spaces

$$TM = \bigcup_{m \in M} T_m M$$

is itself a manifold of twice the dimension of M. It is called the *tangent bundle*. Each point of TM is a pair (m, v) consisting of a point $m \in M$ and a tangent vector $v \in T_m M$. Local coordinates are provided by the coordinates x_a of m in some chart on M and the corresponding components v_a of v.

In mechanics, a tangent vector at a point of the configuration manifold determines a state of motion and the tangent bundle is the phase space. The extended configuration space and the extended phase space are also manifolds. The former has local coordinates $(q_1, \ldots, q_n, t)$ and the latter has coordinates $(q_1, \ldots, q_n, v_1, \ldots, v_n, t)$. So the extended configuration space differs from the configuration space by the inclusion of time as an additional coordinate. For example, in the case of a single particle moving in Euclidean space, the configuration space is $\mathbb{R}^3$ and the extended configuration space is *space-time*, with coordinates (x, y, z, t).

A trajectory in configuration space is a possible evolution of the system. If the parameter along the trajectory is time, then the components of its tangent vector are the generalized velocities. In a mechanical context, we distinguish between a *kinematic trajectory*, which represents some possible motion of the system, and a *dynamical trajectory*, which represents the actual motion under the influence of a particular set of forces. It is also useful in general to draw a terminological distinction between a *curve*, which is the image of a *trajectory* with no particular parametrization, and a trajectory, which is a curve together with a particular labelling of points by a parameter. A curve in an extended configuration space $C \times \mathbb{R}$ has a natural parametrization by time, provided that t does not have stationary points on the curve.

8.5 Functions and Differential Forms

A number of constructions in classical mechanics, particularly the Hamiltonian formalism and the theory of canonical transformations, become more transparent when restated in terms of objects on manifolds called *differential forms*. The simplest of these are smooth functions on the manifold M. In this context, a smooth function is regarded as a 0-*form*.

Next we have 1-forms. A *covector* at a point m of an n-dimensional differentiable manifold M is a element of the dual space to T_mM. To mirror the definition of the tangent space, we define a covector p at m to be an object that assigns a set of n components p_a to each coordinate system on a neighbourhood of m, subject to the transformation rule

$$\tilde{p}_a = \frac{\partial x_b}{\partial \tilde{x}_a} p_b, \tag{8.6}$$

with the partial derivatives evaluated at m. Again the covectors form an n-dimensional vector space, called the *cotangent space* at m and denoted by T_m^*M. If $v \in T_mM$ and $p \in T_m^*M$, then the real number $p_a v_a$ is denoted by $p(v)$. It is independent of the choice of coordinates because

$$\frac{\partial x_a}{\partial \tilde{x}_b} \frac{\partial \tilde{x}_b}{\partial x_c} = \delta_{ac}.$$

So a covector is alternatively defined as a map $p : T_mM \to \mathbb{R}$ that is linear in the sense that

$$p(u+v) = p(u) + p(v), \qquad p(\lambda u) = \lambda p(u)$$

for all $u, v \in T_mM$, $\lambda \in \mathbb{R}$. This is the standard definition of the dual space in linear algebra.

Example 8.3 (Generalized Forces)

The transformation rule (3.14) for generalized forces establishes that they are the components of a covector Q. This covector can be evaluated on a vector v to give $Q(v) \in \mathbb{R}$. If v represents a small displacement in the configuration of a system, then $Q(v)$ represents the work done under the displacement. Proposition 3.7 establishes that the quantities

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial v_a}\right) - \frac{\partial T}{\partial q_a}$$

are also the components of a covector at each point of a trajectory in configuration space.

The set of all pairs (m, p), where $m \in M$ and $p \in T_m^* M$ is again a manifold, with local coordinates $x_1, \ldots, x_n, p_1, \ldots, p_n$. The x_as label the point $m \in M$ and the p_as are the components of p. This manifold is called the *cotangent bundle*.

A 1-form α is an object that assigns an element $\alpha(m) \in T_m^* M$ to each point $m \in M$, in such a way that the components α_a are smooth functions of the local coordinates. The set of all 1-forms on M is denoted by $\Omega^1(M)$.

Example 8.4 (The gradient 1-form)

Let $f \in C^\infty(M)$ and put $\alpha_a = \partial f / \partial x_a$. By the chain rule, under a coordinate change we have

$$\frac{\partial f}{\partial \tilde{x}_a} = \frac{\partial x_b}{\partial \tilde{x}_a} \frac{\partial f}{\partial x_b}.$$

So the α_as have the correct transformation rule for the components of a 1-form. The 1-form with components $\alpha_a = \partial f / \partial x_a$ is called the *gradient 1-form*, and is denoted $\mathrm{d}f$. If $v \in \mathfrak{X}(M)$, then

$$\mathrm{d}f(v) = v_a \frac{\partial f}{\partial x_a} = v(f).$$

Example 8.5 (Submanifolds)

Suppose that M' is a subset of an n-dimensional manifold M, and that in some open neighbourhood U of each point $x \in M$, there are functions

$$f_{n'+1}, f_{n'+2}, \ldots, f_n : U \to \mathbb{R}$$

such that $M' \cap U$ is the subset on which all the f_is vanish and such that the gradient 1-forms $\mathrm{d}f_i$s are linearly independent at each point of U. Then M' is an n'-dimensional submanifold of M.

To introduce coordinates in a neighbourhood of $m \in M'$, we use a chart to identify a neighbourhood of m in M with an open set in $\mathbb{R}^n$. In $\mathbb{R}^n$ we can form the gradient vectors of functions in the usual sense by taking the components to be the partial derivatives with respect to the coordinates. Consider the gradient vectors of the functions f_i at m. These are $n - n'$ linearly independent vectors in $\mathbb{R}^n$. So we can pick exist n' further vectors to make up a basis for $\mathbb{R}^n$. The further vectors are the gradients of n' linear functions $x_1, \ldots, x_{n'}$. If we take

$$x_a = f_a \qquad a = n' + 1, \ldots, n,$$

then together all the x_as can be used as local coordinates on a neighbourhood U of m in M, and M' is given in this neighbourhood by the vanishing of the last $n - n'$ coordinates.

Example 8.6

If x_a is a coordinate system, then the gradient 1-forms $\mathrm{d}x_a$ have the property that

$$\mathrm{d}x_a(\partial_b) = \frac{\partial x_a}{\partial x_b} = \delta_{ab}.$$

So the $\mathrm{d}x_a$s make up the *dual basis* in T_m^*M to the coordinate basis in T_mM at each point m.

One should perhaps own up here to a loose use of terminology, which can be avoided at the expense of perhaps excessive formality. Because the coordinates are defined only on an open subset $U \subset M$, the $\mathrm{d}x_a$s are not defined on the whole of M, but only on this subset. So they are really elements of $\Omega^1(U)$, not of $\Omega^1(M)$, with $\Omega^1(U)$ being defined by treating U as a manifold in its own right. The offence here is really no worse than the universal transgression involved in discussing functions without specifying their domains. In this example as well, honesty is achieved at the expense of clarity.

A 1-form α determines a map $\alpha : \mathfrak{X}(M) \to C^\infty(M)$ by

$$\alpha(v) = \alpha_a v_a \qquad v \in \mathfrak{X}(M),$$

with the linearity property that

$$\alpha(fu + gv) = f\alpha(u) + g\alpha(v)$$

for every $u, v \in \mathfrak{X}(M)$ and $f, g \in C^\infty(M)$. Forms of higher degree are analogous *multilinear* maps. That is, they depend on a number of vector fields and are separately linear in each argument. Objects of this sort are examples of *tensors*. Differential forms have the additional property of being skew-symmetric, which means that the value changes sign when two arguments are interchanged. The reason for restricting to this special class of tensors is twofold. First, it provides us with the objects we need to bring out the geometric structure of classical mechanics. Second, there is a powerful differential calculus, called the *exterior calculus*, for forms on any manifold. The derivative operator d defined above extends to forms of any degree. It is also possible to integrate forms over appropriate subsets of M, although we shall not make any use of the general theory of integration. General tensor calculus, without skew symmetry, requires additional structure on a manifold, for example a *connection*.

Definition 8.7

Let k be a positive integer. A *differential form of degree k* or *k-form* on a manifold M is a map

$$\alpha : \overbrace{\mathfrak{X}(M) \times \cdots \times \mathfrak{X}(M)}^{k} \to C^\infty(M) : (u, \ldots, w) \mapsto \alpha(u, \ldots, w)$$

with the following properties

(1) It is linear over $C^\infty(M)$ in each argument. That is

$$\alpha(\ldots, fu + gv, \ldots) = f\alpha(\ldots, u, \ldots) + g\alpha(\ldots, v, \ldots)$$

for all $f, g \in C^\infty(M)$ and $u, v \in \mathfrak{X}(M)$.

(2) It changes sign if any two arguments are interchanged. That is

$$\alpha(\ldots, u, \ldots, v, \ldots) = -\alpha(\ldots, v, \ldots, u, \ldots).$$

An element of $C^\infty(M)$ is a 0-form.

A k-form α can be evaluated at any point $m \in M$ on any ordered set $u, v, \ldots \in T_m M$ to give a real number $\alpha(u, v, \ldots)$.

If x_a is a coordinate system, and α is a k-form then the functions

$$\alpha_{\underbrace{ab\cdots c}_{k}} = \alpha(\partial_a, \partial_b, \ldots, \partial_c)$$

are called the *components* of α in the coordinate system. The components determine α because, by multi-linearity, we have

$$\alpha(u, v, \ldots, w) = u_a v_b \cdots w_c \alpha_{ab\cdots c}.$$

By skew-symmetry, $\alpha_{ab\ldots c}$ changes sign when two of the suffices are interchanged. So the only nonzero components are those for which the suffices are all distinct. It follows that on an n-dimensional manifold, any k-form with $k > n$ vanishes identically.

EXERCISES

8.5. Show that if v is a vector and p is a covector at some point of a manifold, then $p_a v_a$ is independent of the coordinate system in which it is evaluated.

8.6 Operations on Differential Forms

Addition and Multiplication by Scalars

Two differential forms of the same degree can be added, so that if $\alpha, \beta \in \Omega^k(M)$, then

$$(\alpha + \beta)(u, v, \ldots, w) = \alpha(u, v, \ldots, w) + \beta(u, v, \ldots, w).$$

Forms can also be multiplied by functions, with the product of a k-form α and a 0-form $f \in C^\infty(M)$ defined by

$$(f\alpha)(u, v, \ldots, w) = f\alpha(u, v, \ldots, w).$$

Exterior Product

The last definition extends to a product, called the *exterior product* of two forms of different degrees. If $\alpha \in \Omega^k(M)$ and $\beta \in \Omega^l(M)$, then their exterior product is the $(k+l)$-form denoted by $\alpha \wedge \beta$ and defined by

$$\alpha \wedge \beta(v_1, v_2, \ldots, v_{k+l}) = \frac{1}{(k+l)!} \sum_\sigma \operatorname{sign}(\sigma) \alpha(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) \beta(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+l)}). \quad (8.7)$$

This requires explanation. In (8.7), $v_1, v_2, \ldots, v_{k+l}$ are $k+l$ elements of $\mathfrak{X}(M)$. They are $k+l$ separate vector fields, not the components of a single vector. The sum is over all permutations σ of $k+l$ objects and $\operatorname{sign}(\sigma)$ is equal to 1 for even permutations and -1 for odd permutations. In terms of components

$$(\alpha \wedge \beta)_{ab\cdots f} = \frac{1}{(k+l)!} \sum_\sigma \operatorname{sign}(\sigma) \alpha_{\sigma(a)\sigma(b)\cdots\sigma(c)} \beta_{\sigma(d)\sigma(e)\cdots\sigma(f)}.$$

In fact we do not really need the exterior product except for very low degree forms. It is almost sufficient to restrict attention to the exterior product of two 1-forms α, β, which is defined by

$$(\alpha \wedge \beta)(u, v) = \tfrac{1}{2}\big(\alpha(u)\beta(v) - \beta(u)\alpha(v)\big) \qquad u, v \in \mathfrak{X}(M).$$

The components of $\alpha \wedge \beta$ in this case are $\frac{1}{2}(\alpha_a \beta_b - \alpha_b \beta_a)$.

Interior Product

Given $\alpha \in \Omega^k(M)$ $(k > 0)$ and $u \in \mathfrak{X}(M)$, the *interior product* $\mathrm{i}_u\alpha$ is the $(k-1)$-form defined by

$$(\mathrm{i}_u\alpha)(v, \ldots, w) = k\alpha(u, v, \ldots, w).$$

In terms of components,

$$(\mathrm{i}_u\alpha)_{b\cdots c} = ku_a\alpha_{ab\cdots c}.$$

Exterior Derivative

The exterior derivative bears the same relation to the exterior product as the operator 'curl' does to the vector product in three dimensions. Indeed 'curl' can be interpreted as a special case of the exterior derivative.

Let α be a k form. The *exterior derivative* $\mathrm{d}\alpha$ is the $(k+1)$-form defined by

$$(\mathrm{d}\alpha)_{abc\cdots e} = \frac{1}{(k+1)!}\sum_\sigma \mathrm{sign}(\sigma)\partial_{\sigma(a)}\alpha_{\sigma(b)\sigma(c)\cdots\sigma(e)}, \tag{8.8}$$

where the sum is over the permutations of $k+1$ letters and $\partial_a = \partial/\partial x_a$. When we take the exterior derivative a second time, the result is a $(k+2)$-form $\mathrm{d}^2\alpha = \mathrm{d}(\mathrm{d}\alpha)$ with components

$$(\mathrm{d}^2\alpha)_{abc\ldots g} = \frac{1}{(k+1)!}\sum_\sigma \mathrm{sign}(\sigma)\partial_{\sigma(a)}\partial_{\sigma(b)}\alpha_{\sigma(c)\ldots\sigma(g)}.$$

The terms in the sum on right can be paired by grouping together each pair of terms in which the indices on α are the same, and in the same order. For example,

$$\partial_a\partial_b\alpha_{c\cdots e} \quad \text{and} \quad -\partial_b\partial_a\alpha_{c\cdots e}.$$

However, $\partial_a\partial_b = \partial_b\partial_a$, so the terms in each pair cancel.

This is perhaps a little clearer in the cases $k = 0$ and $k = 1$, which are the only ones that we shall use. For $f \in C^\infty(M)$,

$$(\mathrm{d}^2 f)_{ab} = \tfrac{1}{2}(\partial_a\partial_b f - \partial_b\partial_a f) = 0.$$

For $\alpha \in \Omega^1(M)$,

$$(\mathrm{d}^2\alpha)_{abc} = \tfrac{1}{6}(\partial_a\partial_b\alpha_c - \partial_b\partial_a\alpha_c + \partial_b\partial_c\alpha_a - \partial_c\partial_b\alpha_a + \partial_c\partial_a\alpha_b - \partial_a\partial_c\alpha_b) = 0.$$

To show that the exterior derivative is well defined, it is necessary to establish that the right-hand side of (8.8) transforms correctly under change of

coordinates. For 0-forms, this is already done in the context of defining the gradient 1-form. For a 1-form α, the components α_a and $\tilde{\alpha}_a$ in two coordinate systems x_a and $\tilde{x}_a$ are related by

$$\tilde{\alpha}_a = \frac{\partial x_c}{\partial \tilde{x}_a}\alpha_c.$$

Therefore

$$\begin{aligned}\tilde{\partial}_a\tilde{\alpha}_b &= \frac{\partial^2 x_c}{\partial\tilde{x}_a\partial\tilde{x}_b}\alpha_b + \frac{\partial x_c}{\partial\tilde{x}_a}\frac{\partial\alpha_c}{\partial\tilde{x}_b} \\ &= \frac{\partial^2 x_c}{\partial\tilde{x}_a\partial\tilde{x}_b}\alpha_b + \frac{\partial x_c}{\partial\tilde{x}_a}\frac{\partial x_d}{\partial\tilde{x}_b}\frac{\partial\alpha_c}{\partial x_d},\end{aligned}$$

where $\tilde{\partial}_a = \partial/\partial\tilde{x}_a$. Because the second partial derivative is symmetric under interchange of a and b, it follows that

$$\tfrac{1}{2}(\tilde{\partial}_a\tilde{\alpha}_b - \tilde{\partial}_b\tilde{\alpha}_a) = \tfrac{1}{2}\frac{\partial x_c}{\partial\tilde{x}_a}\frac{\partial x_d}{\partial\tilde{x}_b}(\partial_c\alpha_d - \partial_d\alpha_c),$$

which is the 2-form transformation rule for the components of $\mathrm{d}\alpha$. The general case works in exactly the same way, and similarly hinges on the symmetry of the second partial derivatives of the old coordinates with respect to the new ones.

A k-form such that $\mathrm{d}\alpha = 0$ is said to be *closed* and one that is equal to $\mathrm{d}\beta$ for some $(k-1)$-form β is said to be *exact*. The argument from the interchangeability of partial derivatives establishes that every exact form is closed. In fact the converse of this is very nearly true. The *Poincaré lemma* states that if α is a closed k-form ($k > 1$), then there is an open neighbourhood of each point in which $\alpha = \mathrm{d}\beta$ in the neighbourhood for some 1-form β. Again this is a direct extension of the proposition in three-dimensional vector calculus that $\operatorname{curl}\boldsymbol{v} = 0$ implies, locally, that there exists a potential function f such $\boldsymbol{v} = \operatorname{grad} f$.

The fact that in the large there can exist closed forms that are not exact is the basis of *de Rham cohomology*, which is an important tool for exploring the topology of manifolds.

Restriction

Let $M' \subset M$ be a submanifold and let α be a k-form on M. Its *restriction* to M' is a k-form on M', denoted by $\alpha|_{M'}$, with the value at $m' \in M'$ defined by

$$\alpha|_{M'}(u, v, \ldots) = \alpha(u, v, \ldots),$$

for $u, v, \ldots \in T_{m'}M'$. In the special coordinates in which M' is given by (8.2), the components of the restriction of α are simply the components $\alpha_{ab\cdots c}$ of α, with $a, b, \ldots, c$ restricted to values in $\{1, 2, \ldots, n'\}$ and with $\alpha_{ab\cdots c}$ evaluated at points of M'. There are two aspects of 'restriction', to points of M' and to vectors tangent to M'.

Pull-Back

Let $\rho : M' \to M$ be a smooth map and let α be a k-form on M. The *pull-back* $\rho^*\alpha$ is the k-form on M with the defining property

$$\rho^*\alpha(u, v, \ldots) = \alpha(\rho_* u, \rho_* v, \ldots).$$

Again, we have $\rho^*(\mathrm{d}\alpha) = \mathrm{d}\rho^*(\alpha)$. If $M' \subset M$ is a submanifold and ρ is the inclusion map $M' \to M$, then ρ^* is given by restriction.

The Lie Derivative

The *Lie derivative* of a 0-form f along a vector field v is simply the derivative $v(f)$ defined by (8.4). For a k-form α, the Lie derivative along v is denoted by $\mathcal{L}_v\alpha$ and is defined by

$$\mathcal{L}_v\alpha = \mathrm{i}_v\mathrm{d}\alpha + \mathrm{d}(\mathrm{i}_v\alpha).$$

It has two key properties, which hold for any $u, v \in \mathfrak{X}(M)$ and for any forms α and β:

$$\mathcal{L}_u(\mathrm{i}_v\alpha) = \mathrm{i}_{[u,v]}\alpha + \mathrm{i}_v\mathcal{L}_u\alpha, \quad \text{and} \quad \mathcal{L}_u(\alpha \wedge \beta) = (\mathcal{L}_u\alpha) \wedge \beta + \alpha \wedge \mathcal{L}_u\beta.$$

Integration

If $\alpha \in \Omega^1(M)$ and if $\gamma : [s_1, s_2] \subset \mathbb{R} \to M$ is a trajectory in M given by

$$x_a = x_a(s)$$

then the *integral* of α along γ is defined by

$$\int_\gamma \alpha = \int_{s_1}^{s_2} \alpha_a \frac{\mathrm{d}q_a}{\mathrm{d}s}\,\mathrm{d}s.$$

The result is independent of the coordinate system, and by the standard change of variables formula, it is also unchanged when the parameter s is replaced any increasing function of s.

More generally, one can define the integral[3]

$$\int_{M'} \alpha$$

of a k-form α over a k-dimensional submanifold $M' \subset M$. Stokes' theorem has a very beautiful generalization in this context that connects integration and exterior differentiation.

EXERCISES

8.6. Show that if α, β, and γ are 1-forms, then $\alpha \wedge \beta \wedge \gamma$ has components

$$\tfrac{1}{6}(\alpha_a\beta_b\gamma_c + \alpha_b\beta_c\gamma_a + \alpha_c\beta_a\gamma_b - \alpha_a\beta_c\gamma_b - \alpha_b\beta_a\gamma_c - \alpha_c\beta_b\gamma_a).$$

8.7. Let α be a 1-form with components α_a and let ω be a 2-form with components ω_{ab}. Show that $\mathrm{d}\alpha$ and $\mathrm{d}\omega$ have respective components

$$\begin{aligned}(\mathrm{d}\alpha)_{ab} &= \tfrac{1}{2}(\partial_a\alpha_b - \partial_b\alpha_a)\\ (\mathrm{d}\omega)_{abc} &= \tfrac{1}{6}(\partial_a\omega_{bc} + \partial_b\omega_{ca} + \partial_c\omega_{ab} - \partial_a\omega_{cb} - \partial_b\omega_{ac} - \partial_c\omega_{ba})\\ &= \tfrac{1}{3}(\partial_a\omega_{bc} + \partial_b\omega_{ca} + \partial_c\omega_{ab}).\end{aligned}$$

8.8. Let M be an n-dimensional manifold and let $\alpha \in \Omega^k(M)$. Show that in local coordinates

$$\alpha = \alpha_{ab\cdots c}\mathrm{d}x_a \wedge \mathrm{d}x_b \wedge \cdots \wedge \mathrm{d}x_c.$$

Deduce that if $\alpha \neq 0$ then $k \leq n$.

8.9. One can define k-forms in a slightly different way by associating with each tangent space T_mM the vector space $\bigwedge^k T_m^*M$. The elements of $\bigwedge^k T_m^*M$ are maps

$$\alpha : \overbrace{T_mM \times \cdots \times T_mM}^{k} \to \mathbb{R}$$

with the properties

(a) they linear over $\mathbb{R}$ in each argument.

(b) they change sign when any two arguments are interchanged.

[3] This is not quite the whole story: M' must have an *orientation*.

Show that on an n-dimensional manifold, $\bigwedge^k T^*_m M$ is a vector space and that

$$\dim\left(\bigwedge^k T^*_m M\right) = \binom{n}{k}.$$

Restate the definition of differential forms in terms of objects that map points of M to elements of the corresponding spaces $\bigwedge^k T^*_m M$.

8.10. Show that

$$(\mathrm{d}\alpha)|_{M'} = \mathrm{d}(\alpha|_{M'}).$$

8.11. Show that for a 1-form α and for a 2-form ω, the components of the Lie derivative along $v \in \mathfrak{X}(M)$ are respectively

$$\begin{aligned}(\mathcal{L}_v\alpha)_a &= v_b\partial_b\alpha_a + \alpha_b\partial_a v_b\\ (\mathcal{L}_v\omega)_{ab} &= v_c\partial_c\omega_{ab} + \omega_{cb}\partial_a v_c + \omega_{ac}\partial_b v_c.\end{aligned}$$

Hence check the key properties of the Lie derivative for 1-forms and 2-forms.

8.7 Vector Fields and Flows

A very helpful analogy to keep in mind when thinking about vector fields on manifolds comes from fluid dynamics. One can picture a vector field as an assignment of a velocity to each point, like the velocity vector field of a fluid. This suggests that one should also think of the manifold as 'flowing' along the vector field, by analogy with the way in which, in the steady case, particles of a fluid flow along the streamlines.

In more formal terms, given a vector field v on a manifold M, each point is the starting point of a *trajectory* of the field and the *flow* is constructed as a family of maps by moving points along their trajectories. For each $m \in M$, the trajectory

$$s \mapsto \rho_s(m)$$

is determined in local coordinates by solving the system of differential equations

$$\frac{\mathrm{d}x_a}{\mathrm{d}s} = v_a(x) \tag{8.9}$$

for the x_as as functions of s, with the initial condition $\rho_s(0) = m$. That is, the trajectory starting at m is characterized by the condition that its tangent vector is v at each point. Picard's theorem for differential equations ensures that it is unique, and the characterization implies that it is independent of the choice of coordinates. In the fluid analogy, the trajectories are the streamlines.

The trajectories determine the *flow* of the vector field, which is the family of maps

$$\rho_s : M \to M : m \mapsto \rho_s(m), \qquad s \in \mathbb{R} \tag{8.10}$$

with the defining property that for each $m \in M$, the trajectory $s \mapsto \rho_s(m)$ is the solution of (8.9) with initial point m at $s = 0$. In fact, generally it is only possible to solve (8.10) with given initial conditions for s in some interval in $\mathbb{R}$ containing the origin. Thus the flow is only defined on some open neighbourhood of $M \times \{0\}$ in $M \times \mathbb{R}$. When it can be defined globally for all s, the vector field is said to be *complete*.

Example 8.8

Suppose that $M = \mathbb{R}^2$, with coordinates x_1, x_2, and that $v_1 = -x_2$, $v_2 = x_1$. Then the trajectories are determined by

$$\frac{\mathrm{d}x_1}{\mathrm{d}s} = -x_2, \qquad \frac{\mathrm{d}x_2}{\mathrm{d}s} = x_1.$$

These imply that

$$\frac{\mathrm{d}^2 x_1}{\mathrm{d}s^2} = -x_1,$$

and hence that $x_1 = A\cos s + B\sin s$, $x_2 = A\sin s - B\cos s$ for constant A, B. Therefore the flow is

$$\rho_s : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} \cos s & -\sin s \\ \sin s & \cos s \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Here the vector complete, and the flow is given by rotation through s about the origin.

Example 8.9

Suppose that $M = \mathbb{R}_+$, with coordinate $x > 0$, and v has component x^2. The trajectories are given by

$$\frac{\mathrm{d}x}{\mathrm{d}s} = x^2,$$

which has solutions $x = (A - s)^{-1}$ for constant A. Here the flow is

$$\rho_s : x \mapsto \left(x^{-1} - s\right)^{-1} = \frac{x}{1 - sx}.$$

The vector field in this case is not complete The trajectory starting at x is defined only for $s < x^{-1}$.

More generally, one can construct trajectories and define a flow for a *time-dependent vector field* v. This is an object that assigns a vector in T_mM to each $m \in M$ and $t \in \mathbb{R}$. Its components $v_a(x,t)$ are functions of the local coordinates x_a and the time t. They obey the transformation rule (8.3) for vector components and are required to depend smoothly on the x_as and t.

In the fluid analogy, a time-dependent vector field is the velocity field of a non-steady flow, one in which the fluid velocity at each point changes with time. In non-steady flows, there is a distinction between streamlines, which are the curves tangent to the velocity field at a particular fixed time, and the paths actually followed by the particles of the fluid. In steady flows, they coincide, but in non-steady flows they generally do not.

In the same way, we note the distinction between two systems of ordinary differential equations determined by a time-dependent vector field v. At each fixed value of t, we have the *autonomous system*

$$\frac{\mathrm{d}x_a}{\mathrm{d}s} = v_a\big(x(s),t\big). \tag{8.11}$$

We also have the *non-autonomous system*

$$\frac{\mathrm{d}x_a}{\mathrm{d}t} = v_a\big(x(t),t\big). \tag{8.12}$$

For a time-independent vector field, they differ only by the substitution of t for the parameter s, but in the general case the solutions are different. In fluid flow, the first determines the streamlines and the second determines the actual motion of the particles of the fluid.

The distinction becomes somewhat clearer when we associate with v two vector fields on $M \times \mathbb{R}$. The first, which we also denote by v, has components

$$(v_1, \ldots, v_n, 0). \tag{8.13}$$

We call its flow $\rho_s : M \times \mathbb{R} \to M \times \mathbb{R}$ the *fixed-time flow* of v, or more often, simply the *flow* of v. It is given by solving (8.11). It has the property that t is constant along the flow, hence 'fixed-time'.

The second vector field has components

$$(v_1, \ldots, v_n, 1).$$

Its flow on $M \times \mathbb{R} \to M \times \mathbb{R}$ is called the *dynamical flow* of v. In the fluid analogy, the fixed-time flow gives the streamlines and the dynamical flow gives the actual motion of the fluid particles. In dynamics, the vector field gives the state of motion and the dynamical flow gives the actual evolution of the system under the influence of given forces.

Example 8.10

With $M = \mathbb{R}^2$, consider the time-dependent vector field with components

$$(v_1, v_2) = (1, t)$$

The autonomous system in this case is

$$\frac{\mathrm{d}x_1}{\mathrm{d}s} = 1, \qquad \frac{\mathrm{d}x_2}{\mathrm{d}s} = t.$$

So the fixed time flow $\rho_s : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}^2 \times \mathbb{R}$ sends (x_1, x_2, t) to $(x_1 + s, x_2 + st, t)$. In fluid terms, the curves $s \mapsto (x_1 + s, x_2 + st)$ for fixed x_1, x_2, and t are the streamlines at time t. The non-autonomous system on the other hand is

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = 1, \qquad \frac{\mathrm{d}x_2}{\mathrm{d}t} = t.$$

Its solutions give the dynamical flow.

A time-dependent function f on M is a function that depends on a point $m \in M$ and a time t, in other words a function on $M \times \mathbb{R}$. We define the derivative of f along a time-dependent vector field v in the same way as in the time-independent case, by

$$v(f) = v_a \frac{\partial f}{\partial x_a},$$

or in a coordinate-independent way, by

$$v(f)(m) = \frac{\mathrm{d}}{\mathrm{d}s}\Big(f\big(\rho_s(m)\big)\Big)\bigg|_{s=0}.$$

There is also a notion of *time-dependent* differential form on M, which is simply a differential form on $M \times \mathbb{R}$. We do not develop the theory of these objects in any detail, except to note that if v is a time-dependent vector field and α is a time-dependent form, then $\mathrm{i}_v\alpha$ is a well-defined time-dependent form, given by taking the interior product of α, as a form on $M \times \mathbb{R}$, with the vector field in (8.13).

8.8 Commuting Flows

Suppose that we have two time-independent vector fields v, v' on a manifold M and we construct the corresponding flows ρ_s and $\rho'_{s'}$ by solving

$$\frac{\mathrm{d}x_a}{\mathrm{d}s} = v_a(x) \quad \text{and} \quad \frac{\mathrm{d}x_a}{\mathrm{d}s'} = v'_a(x).$$

Then in general it is not true that

$$\rho_s \circ \rho'_{s'} = \rho'_{s'} \circ \rho_s. \tag{8.14}$$

When (8.14) *does* hold, the flows are said to *commute*. In this case, we define

$$\rho_{s,s'}(x) = \rho_s\big(\rho_{s'}(x)\big) = \rho_{s'}\big(\rho_s(x)\big)$$

at least for (s, s') in some neighbourhood of the origin in $\mathbb{R}^2$, which may depend on x. In coordinates, we have $x_a = x_a(s, s')$, with

$$\frac{\partial x_a}{\partial s} = v_a, \qquad \frac{\partial x_a}{\partial s'} = v'_a.$$

These two equations must be compatible, which implies that

$$\frac{\partial^2 x_a}{\partial s \partial s'} = \frac{\partial v'_a}{\partial s} = \frac{\partial v_a}{\partial s'}$$

But

$$\frac{\partial v'_a}{\partial s} = \frac{\partial x_b}{\partial s}\frac{\partial v'_a}{\partial x_b} = v_b \frac{\partial v'_a}{\partial x_b}$$

and

$$\frac{\partial v_a}{\partial s'} = \frac{\partial x_b}{\partial s'}\frac{\partial v_a}{\partial x_b} = v'_b \frac{\partial v_a}{\partial x_b}.$$

It follows that a necessary condition for the flows to commute is that

$$v_b \frac{\partial v'_a}{\partial x_b} - v'_b \frac{\partial v_a}{\partial x_b} = 0.$$

In fact, this condition is also sufficient. In the terminology of Section 8.3, the condition is that the Lie bracket vector field $[v, v']$ should vanish.

8.9 Lagrangian and Hamiltonian Dynamics

In this geometric language, the problem of understanding the dynamics of a mechanical system is the problem of determining the dynamical trajectories in its configuration manifold C and in its phase space $P = TC$. We can view the Lagrangian theory as a means of constructing of a time-dependent vector field on the phase space. Its dynamical flow in the extended phase space then gives the dynamics.

A choice of generalized coordinates q_a on C extends to a coordinate system

$$(q_1, \ldots, q_n, v_1, \ldots, v_n, t)$$

on $TC \times \mathbb{R}$. The q_as label the configurations in C and the v_as are the components of the tangent vectors to C, which label the states of motion. In such a coordinate system, a time-dependent vector field has components

$$(v_1, \ldots, v_n, w_1, \ldots, w_n),$$

where the w_as are functions of the q_as, v_as, and t. The corresponding dynamical trajectories are the solutions to the system differential equations

$$\frac{\mathrm{d}q_a}{\mathrm{d}t} = v_a(q, v, t), \qquad \frac{\mathrm{d}v_a}{\mathrm{d}t} = w_a(q, v, t).$$

When the vector field is determined from a Lagrangian, the w_as are found by writing Lagrange's equations in the form

$$\frac{\partial^2 L}{\partial v_a \partial v_b} w_b + \frac{\partial^2 L}{\partial v_a \partial q_b} v_b + \frac{\partial^2 L}{\partial v_a \partial t} - \frac{\partial L}{\partial q_a} = 0,$$

which is a set of n linear equations in the unknown w_as. Provided that the Lagrangian is *nondegenerate*, in the sense that the matrix with entries

$$\frac{\partial^2 L}{\partial v_a \partial v_b}$$

is everywhere nonsingular, it uniquely determines the w_as, and hence a time-dependent vector field v_L on TC. Without this condition, Lagrange's equations do not generate sensible dynamics, or rather, they give rise to constraints. This is a significant point in field theories, particularly gauge theories, but in our exploration of classical mechanics, we always assume implicitly that the Lagrangian is nondegenerate.

One finds the dynamical trajectories of v_L and hence the dynamical trajectories of the system by solving Lagrange's equations. The construction of v_L is somewhat more transparent in canonical coordinates, in which one again labels the configurations by the q_as, but uses the generalized momenta

$$p_a = \frac{\partial L}{\partial v_a}$$

in place of the v_as to label the states of motion. In these coordinates, the dynamical trajectories are determined by Hamilton's equations

$$\dot{q}_a = \frac{\partial h}{\partial p_a}, \qquad \dot{p}_a = -\frac{\partial h}{\partial q_a}.$$

After a transformation of the configuration coordinates q_a to a new system $\tilde{q}_a$, the generalized momenta are p_a are replaced by new coordinates

$$\tilde{p}_a = \frac{\partial L}{\partial \tilde{v}_a} = \frac{\partial q_b}{\partial \tilde{q}_a} \frac{\partial L}{\partial v_b} = \frac{\partial q_b}{\partial \tilde{q}_a} p_a,$$

by using (3.8). But this is the transformation rule for the components of a covector on C. So we should regard the q_as and p_as as coordinates on the cotangent bundle T^*C. From this point of view, the Legendre transformation

$$(q_1, \ldots, q_n, v_1, \ldots, v_n, t) \mapsto (q_1, \ldots, q_n, p_1, \ldots, p_n, t), \quad p_a = \frac{\partial L}{\partial v_a}$$

is interpreted as a map

$$\lambda : TC \times \mathbb{R} \to T^*C \times \mathbb{R}.$$

The Lagrangian is seen as a function on $TC \times \mathbb{R}$ and the Hamiltonian as a function on $T^*C \times \mathbb{R}$. The Legendre transformation maps the dynamical trajectories generated by L on $TC \times \mathbb{R}$ to those generated by h on $T^*C \times \mathbb{R}$. In a well behaved system with a nondegenerate Lagrangian,

$$\lambda : TC \times \mathbb{R} \to T^*C \times \mathbb{R}$$

is invertible and is therefore a diffeomorphism. It identifies the two manifolds in a natural way, so we can simply regard both as alternative representations of the extended phase space. However one has to be careful. When the Lagrangian depends on time, TC and T^*C are not *naturally* identified. Standard terminology uses 'phase space' for the latter, rather than, as in our usage, for the former.

In the Hamiltonian picture, with the extended phase space represented as $T^*C \times \mathbb{R}$, the dynamical evolution is given by the time-dependent vector field v_h on T^*C with components

$$\left(\frac{\partial h}{\partial p_1}, \ldots, \frac{\partial h}{\partial p_n}, -\frac{\partial h}{\partial q_1}, \ldots, -\frac{\partial h}{\partial q_n} \right).$$

An altogether more pleasing characterization of v_h is made by introducing the *symplectic 2-form* ω on T^*C, defined by

$$\omega = \mathrm{d}p_a \wedge \mathrm{d}q_a. \tag{8.15}$$

The interior product of v_h with ω is the 1-form

$$\mathrm{i}_{v_h}\omega = -\frac{\partial h}{\partial q_a}\mathrm{d}q_a - \frac{\partial h}{\partial p_a}\mathrm{d}p_a.$$

But this is simply minus the gradient 1-form of h, regarded as a time-dependent function on T^*C. Therefore v_h is determined by

$$\mathrm{i}_{v_h}\omega + \mathrm{d}h = 0,$$

which is the coordinate-independent version of Hamilton's equations.

Proposition 8.11

The symplectic 2-form is invariant under canonical transformations.

Proof

Suppose that the coordinate systems q_a, p_a and q''_a, p''_a on T^*C are related by a canonical transformation. We have to show that

$$\mathrm{d}p_a \wedge \mathrm{d}q_a = \mathrm{d}p''_a \wedge \mathrm{d}q''_a. \tag{8.16}$$

The left-hand side is the exterior derivative of $p_a\,\mathrm{d}q_a$, and the right-hand side is the exterior derivative of $-q''_a\,\mathrm{d}p''_a$. So this is equivalent to showing that

$$\mathrm{d}(p_a\,\mathrm{d}q_a + q''_a\,\mathrm{d}p''_a) = 0. \tag{8.17}$$

But, in the notation of Section 7.4,

$$p_a\,\mathrm{d}q_a + q''_a\,\mathrm{d}p''_a = \frac{\partial F}{\partial p'_a}\mathrm{d}q'_a + \frac{\partial F}{\partial p'_a}\mathrm{d}p'_a = \mathrm{d}F. \tag{8.18}$$

So the proposition follows from the fact that $\mathrm{d}(\mathrm{d}F) = 0$. □

In particular, ω is unchanged under transformations of the configuration coordinates q_a (see Exercise 7.9), so it is a *natural* structure on T^*C. Its construction requires only the manifold structure of C, and is independent of the choice of Lagrangian or Hamiltonian.

The proposition can be understood in another way by using the fact that ω determines Poisson brackets. If f is a smooth function on T^*C, then we define the *Hamiltonian vector field* v_f by

$$\mathrm{i}_{v_f}\omega + \mathrm{d}f = 0.$$

That is by taking the components of v_f to be

$$\left(\frac{\partial f}{\partial p_1}, \ldots, \frac{\partial f}{\partial p_n}, -\frac{\partial f}{\partial q_1}, \ldots, -\frac{\partial f}{\partial q_n}\right).$$

For any other function g

$$v_f(g) = -[f, g], \tag{8.19}$$

This works equally well if f is a time-dependent function, in which case v_f is a time-dependent vector field.

It follows from (8.19) that the symplectic form determines Poisson brackets. Thus a transformation to a new coordinate system q''_a, p''_a which is transverse to p_a, q_a and satisfies (8.16) must preserve Poisson brackets and therefore be

canonical. With this in mind, we widen slightly our earlier definitions and remove the awkwardness over the 'transversality' by *defining* a canonical coordinate system on T^*C to be *any* local coordinate system $q_1, \ldots, q_n, p_1, \ldots, p_n$ in which

$$\omega = \mathrm{d}p_a \wedge \mathrm{d}q_a \tag{8.20}$$

and by designating *any* transformation between canonical coordinate systems as canonical. Any system of generalized coordinates on C gives rise to a canonical coordinate system, but the power of the Hamiltonian theory stems from the fact that there is a much wider choice of canonical coordinates.

Finally, note that if (8.16) holds, then there exists a function F such that (8.18) holds, at least locally by the Poincaré lemma. Thus in outline we have a proof of the converse to Proposition 7.6.

8.10 Symmetry and Conservation Laws

Noether's theorem gives a connection between one-parameter groups of dynamical symmetries and conserved momenta. The connection in Chapter 4 is revealed more clearly in geometric language by observing that (4.5) characterizes the generators u_a of a one-parameter transformation group ρ_s as the components of a time-dependent vector field u on configuration space. With this interpretation, it follows from (4.4) that ρ_s is the fixed-time flow of u on C.

The condition (4.6) that ρ_s should be a dynamical symmetry also takes on a simpler form. If we put

$$w_a = \frac{\partial u_a}{\partial q_b} v_b + \frac{\partial u_a}{\partial t},$$

then

$$u_1, \ldots, u_n, w_1, \ldots, w_n \tag{8.21}$$

are the components of a time-dependent vector field $\hat{u}$ on TC, called the *lift* of u. With this notation, Equation (4.6) reads more simply as

$$\hat{u}(L) = 0.$$

The conjugate momentum to ρ_s is

$$p = u_a \frac{\partial L}{\partial v_a} = u_a p_a.$$

Lemma 8.12

Suppose that $\rho_s : C \times \mathbb{R} \to C \times \mathbb{R}$ is a dynamical symmetry of a system with Lagrangian L, generated by a time-dependent vector field u on C. Let p be the conjugate momentum. Then

$$\mathrm{i}_{\hat{u}} \mathrm{d}\theta + \mathrm{d}p = 0 \qquad \text{where} \qquad \theta = \frac{\partial L}{\partial v_a} \mathrm{d}q_a.$$

Proof

Because $\hat{u}(L) = 0$, we have

$$\hat{u}(p_a) = \hat{u}\left(\frac{\partial L}{\partial v_a}\right) = -\frac{\partial u_b}{\partial v_a}\frac{\partial L}{\partial q_b} - \frac{\partial w_b}{\partial v_a}\frac{\partial L}{\partial v_b} = -\frac{\partial u_b}{\partial q_a}\frac{\partial L}{\partial v_b} = -\frac{\partial u_b}{\partial q_a} p_b.$$

Therefore

$$\mathrm{i}_{\hat{u}}\, \mathrm{d}\theta = \hat{u}(p_a)\, \mathrm{d}q_a - \hat{u}(q_a)\, \mathrm{d}p_a = -\frac{\partial u_b}{\partial q_a} p_b\, \mathrm{d}q_a - u_a \mathrm{d}p_a = -\mathrm{d}p.$$

□

When transferred to $T^*C \times \mathbb{R}$ by the Legendre transformation, this allows a very elegant geometric restatement of Noether's theorem: the momentum conjugate to a dynamical symmetry is conserved and the symmetry is the flow along Hamiltonian vector field of the conserved momentum. The conservation of p is expressed by

$$\frac{\mathrm{d}p}{\mathrm{d}t} = [p, h] + \frac{\partial p}{\partial t} = 0.$$

EXERCISES

8.12. Show that the components of $\hat{u}$ transform as vector components under change of generalized coordinates.

8.11 Symplectic Manifolds

The phase space of a mechanical system is an example of a *symplectic manifold*. More generally, we adopt the following definition.

Definition 8.13

A *symplectic manifold* is a manifold M on which there is given a *symplectic 2-form* ω. That is, a 2-form $\omega \in \Omega^2(M)$ which is

(1) *Closed* in the sense that $\mathrm{d}\omega = 0$ and

(2) *Nondegenerate* in the sense that $\mathrm{i}_v\omega = 0$ for $v \in \mathfrak{X}(M)$ only if $v = 0$.

A diffeomorphism $\rho : M' \to M$ from one symplectic manifold (M', ω') to a second symplectic manifold (M, ω) is a *canonical transformation* if $\omega' = \rho^*\omega$.

Example 8.14

The Euclidean space $\mathbb{R}^{2n}$ with the symplectic form

$$\omega = \mathrm{d}p_a \wedge \mathrm{d}q_a$$

is a symplectic manifold. Darboux's theorem, below, shows that it is a local model for all $2n$-dimensional symplectic manifolds.

Example 8.15

Let C be an n-dimensional manifold. There is a natural 1-form θ on T^*C, called the *canonical 1-form*. In local coordinates q_a on C, it is defined by

$$\theta = p_a \, \mathrm{d}q_a,$$

where $q_1, \ldots, q_n, p_1, \ldots, p_n$ is the corresponding coordinate system on M. That the definition does not depend on the choice of coordinates follows from the transformation rule (8.6) for the p_as, or else by defining θ intrinsically by the formula

$$\mathrm{i}_v\theta = p(\pi_* v) \quad \text{at } (q, p) \in T^*C$$

where $\pi : M \to C : (q, p) \mapsto q$ and v is a tangent vector to T^*C at (q, p). Thus θ is the structure on T^*C that gives expression to the fact that points of T^*C are covectors on C.

The exterior derivative $\omega = \mathrm{d}\theta$ is a natural symplectic form on T^*C, given in local coordinates by

$$\omega = \mathrm{d}p_a \wedge \mathrm{d}q_a.$$

It is called the *canonical 2-form*.

Example 8.16

Let C be an n dimensional manifold and let $L : TC :\to \mathbb{R}$ be a nondegenerate time-independent Lagrangian. Put

$$\theta_L = \frac{\partial L}{\partial v_a}\,\mathrm{d}q_a \in \Omega^1(TC).$$

Then θ_L is independent of the choice of coordinates and its exterior derivative is a symplectic form on TC. Because L is time-independent, the Legendre transformation in this case can be interpreted as the map $\lambda : TC \to T^*C$ given by

$$\lambda : (q, v) \mapsto (q, p) \qquad p_a = \frac{\partial L}{\partial v_a}.$$

Under the pull-back,

$$\lambda^*(\theta) = \theta_L, \qquad \lambda^*(\omega) = \omega_L = \mathrm{d}\theta_L\,.$$

When $L = T - U$, where the kinetic energy T is a homogeneous quadratic in the velocities, λ is a diffeomorphism, and hence a canonical transformation.

Example 8.17

Let M be the unit sphere in $\mathbb{R}^3$, with

$$\omega = \sin\theta\,\mathrm{d}\theta \wedge \mathrm{d}\varphi,$$

Then (M, ω) is a compact symplectic manifold. It is a very simple example of *coadjoint orbit* associated with a Lie group, in this case the rotation group. Any real Lie group acts as a group of linear transformations of the dual space to its Lie algebra. The orbits of the group under this action have the natural structure of symplectic manifolds. In the case of the rotation group, the Lie algebra and its dual are both R^3, the space of angular velocity vectors, and the unit sphere is one of the orbits.

The dimension of a symplectic manifold is necessarily even because the nondegeneracy condition implies that the determinant of the matrix (ω_{ab}) is nonzero. But a skew-symmetric matrix can only have nonvanishing determinant if it has an even number of columns.

On a symplectic manifold, we have the same relationship between functions and Hamiltonian vector fields. Given $f \in C^\infty(M)$, we define the Hamiltonian vector field generated by f as before by

$$\mathrm{i}_{v_f}\omega + \mathrm{d}f = 0.$$

We also extend the definition of the Poisson bracket by putting

$$[f, g] = -v_f(g) = -2\omega(v_f, v_g) = -[g, f],$$

for $f, g \in C^\infty(M)$. A diffeomorphism $\rho : M' \to M$ from one symplectic manifold to another is a canonical transformation if and only if it preserves Poisson brackets. That is, if and only if

$$[f \circ \rho, g \circ \rho]' = [f, g] \circ \rho \qquad \forall f, g \in C^\infty(M),$$

where $[\,\cdot, \cdot\,]$ and $[\,\cdot, \cdot\,]'$ are the respective Poisson brackets on M and M'.

If v is the Hamiltonian vector field of some function, then the 1-form $\mathrm{i}_v\omega$ is closed, but the converse is not true. On a general symplectic manifold, it is possible that $\mathrm{i}_v\omega$ is closed, but not exact. Indeed we can construct such a vector field from a closed 1-form α by putting

$$\mathrm{i}_v\omega + \alpha = 0.$$

If α is not exact, then v is not Hamiltonian. A vector field such that $\mathrm{i}_v\omega$ is closed is said to be *locally Hamiltonian*. Because closed 1-forms become exact when restricted to suitable open neighbourhoods in M, locally Hamilton vector fields become Hamiltonian when restricted in the same way. Hence the terminology.

Lemma 8.18

Suppose that u and v are locally Hamiltonian vector fields. Then $[u, v]$ is the Hamiltonian vector field of the function $h = 2\omega(u, v)$.

Proof

Because u and v are locally Hamiltonian, we have

$$\mathcal{L}_u\omega = \mathrm{d}(\mathrm{i}_u\omega) = 0 \quad \text{and} \quad \mathcal{L}_v\omega = \mathrm{d}(\mathrm{i}_v\omega) = 0$$

Therefore

$$\begin{aligned} \mathrm{i}_{[u,v]}\omega &= \mathcal{L}_u(\mathrm{i}_v\omega) - \mathrm{i}_v\mathcal{L}_u\omega \\ &= \mathrm{i}_u\mathrm{d}(\mathrm{i}_v\omega) + \mathrm{d}(\mathrm{i}_u(\mathrm{i}_v\omega)) \\ &= -2\mathrm{d}\big(\omega(u, v)\big). \end{aligned}$$

□

An immediate corollary is that for any $f, g \in C^\infty(M)$, we have

$$[v_f, v_g] = -v_{[f,g]}.$$

A second corollary is that the Poisson bracket on a symplectic manifold satisfies the Jacobi identity. In fact, the Jacobi identity is an equivalent form of the closure condition for a nondegenerate 2-form ω.

Proposition 8.19

Let (M, ω) be a symplectic manifold and let $f, g, h \in \mathbb{C}^\infty(M)$. Then the *Jacobi identity* holds. That is

$$\big[f, [g, k]\big] + \big[g, [k, f]\big] + \big[k, [f, g]\big] = 0.$$

Proof

By Lemma 8.18,

$$2v_f\big(\omega(v_g, v_h)\big) = \big[f, [g, h]\big].$$

But $\mathcal{L}_{v_f}\omega = 0$, so we also have

$$\begin{aligned} 2v_f\big(\omega(v_g, v_h)\big) &= 2\omega\big([v_f, v_g], v_h\big) + 2\omega\big(v_g, [v_f, v_h]\big) \\ &= 2\omega\big(v_h, v_{[f,g]}\big) - 2\omega\big(v_g, v_{[f,h]}\big) \\ &= -\big[h, [f, g]\big] - \big[g, [h, f]\big]. \end{aligned}$$

The proposition follows. □

The product of two symplectic manifolds (M_1, ω_1) and (M_2, ω_2) is also a symplectic manifold. Its symplectic form is the sum of the symplectic forms on M_1 and M_2, pulled back to $M_1 \times M_2$ by the projection maps

$$M_1 \times M_2 \to M_1, \qquad M_1 \times M_2 \to M_2.$$

In contrast, a submanifold M' of symplectic manifold (M, ω) need not be symplectic. The restriction of

$$\omega' = \omega|_{M'}$$

is a closed 2-form, but it may be degenerate. At each $m' \in M'$, we have a tangent subspace

$$K_{m'} = \{v \in T_{m'}M' \mid \mathrm{i}_v\omega' = 0\} \subset T_{m'}M'.$$

This can contain nonzero vectors. When it does not, that is when $\dim K$ is everywhere zero, M' is symplectic. At the other extreme, when $\omega' = 0$, it is the whole tangent space at each point and M' is said to *isotropic*. In this case,

$$\dim M' \leq \tfrac{1}{2}\dim M.$$

If M' is isotropic and of half the dimension on M, the maximal case, then it is called a *Lagrangian submanifold*.

Example 8.20

Suppose that ω is given by (8.15).

(1) If M' is given by $p_n = q_n = 0$, then it is symplectic and

$$\omega' = \sum_1^{n-1} \mathrm{d}p_a \wedge \mathrm{d}q_a.$$

(2) If M' is given by $p_n = q_1 = q_2 = \cdots = q_n = 0$, then it is isotropic.

(3) If M' is given by $q_1 = q_2 = \cdots = q_n = 0$, then it is Lagrangian.

(4) If $S(q_1, \ldots, q_n)$ is any smooth function of the q_as, and if M' is given by

$$p_a = \frac{\partial S}{\partial q_a}, \tag{8.22}$$

then M' is Lagrangian.

Although a submanifold of a symplectic manifold is not usually itself symplectic, it is often possible to construct a symplectic manifold from it by *symplectic reduction*. The idea, which we shall not explain in detail here, is that when the dimension of K is constant on M', K is an integrable distribution, as a result of the closure of ω. If the quotient space of leaves is a manifold, then ω descends to the quotient, where is it closed and nondegenerate. In physical terms, reduction removes irrelevant degrees of freedom. For example, for an isolated system in space, the motion of the centre of mass is irrelevant. One obtains a reduced phase space that describes the internal degrees of freedom by fixing the linear momenta to define M', and factoring out overall translations, to form the quotient space. See Marsden and Weinstein [7].

8.12 Darboux's Theorem

In fact, general symplectic manifolds are not very different from the phase spaces we have already encountered, except in their large-scale structure, because it is always possible to find local coordinates $q_1, \ldots, q_n, p_1, \ldots, p_n$ such that

$$\omega = \mathrm{d}p_a \wedge \mathrm{d}q_a. \tag{8.23}$$

This is the content of the following proposition.

Proposition 8.21 (Darboux)

Let M be a symplectic manifold with symplectic form ω. Then in some neighbourhood of each point, there exists a *canonical coordinate system* p_a, q_a in which $\omega = \mathrm{d}p_a \wedge \mathrm{d}q_a$.

A very elegant and informative proof is given by Weinstein [12], but it requires further preliminary work on basic differential geometry. Instead, we sketch a different proof.

Suppose that α is a closed 2-form on a manifold. At each point m, denote by K_m the space of vectors such that $\mathrm{i}_v\alpha = 0$, and suppose that the dimension of K is constant. We call K the *kernel* of α. If u, v are vector fields taking values in K at each point, then so is $[u, v]$. This follows from the closure of α by using

$$\mathrm{i}_{[u,v]}\alpha = \mathcal{L}_u(\mathrm{i}_v\alpha) - \mathrm{i}_v\mathcal{L}_u\alpha = -\mathrm{i}_v\mathrm{d}(\mathrm{i}_u\alpha) - \mathrm{i}_v(\mathrm{i}_u\mathrm{d}\alpha) = 0.$$

The *Frobenius integrability theorem* [6] then implies the existence, locally, of a nonconstant function f such that $v(f) = 0$ whenever v takes values in the kernel.

For a closed 2-form ω of constant rank $2r$, we prove by induction on r that there exist functions $p_1, \ldots, p_r, q_1, \ldots, q_r \in C^\infty(M)$ such that

$$\omega = \sum_1^r \mathrm{d}p_a \wedge \mathrm{d}q_a.$$

Given ω with rank $2r$, the argument above establishes the existence of a function p such that $v(p) =$ for all v taking values in the kernel of ω. The argument is then used again with the closed 2-form

$$\alpha = \omega + \mathrm{d}t \wedge \mathrm{d}p$$

on $M \times \mathbb{R}$ to deduce the existence of a function f on $M \times \mathbb{R}$ such that

$$v(f) + g\,\partial_t f = 0 \quad \text{whenever} \quad \mathrm{i}_v\omega + g\,\mathrm{d}p = 0 \tag{8.24}$$

with $v \in \mathfrak{X}(M)$ and $g \in \mathbb{C}^\infty(M)$. Now pick a constant value c for f and solve the equation $f = c$ for t to obtain $t = q$, where $q \in C^\infty(M)$. The induction step replaces ω by

$$\omega - \mathrm{d}p \wedge \mathrm{d}q.$$

This has lower rank than ω because its kernel contains every vector satisfying the second equation in (8.24) for some g. The details of this last step, as well as some technical points in the proof, are left as exercises.

A coordinate system in which ω is given by (8.23) is said to be *canonical*. A necessary and sufficient condition for a set of functions p_a, q_a to form a system of canonical coordinates is that their Poisson brackets should satisfy

$$[p_a, p_b] = 0, \qquad [q_a, q_b] = 0, \qquad [q_a, p_b] = \delta_{ab}. \tag{8.25}$$

8.13 Integrability

In many of the classical examples encountered in earlier chapters, it is possible to solve the dynamical equations explicitly, or at least in terms of integrals, because there is a 'full' set of constants of the motion. Such systems are 'integrable' in a sense that we now try to capture in the general context of Hamiltonian systems.

A *time-independent Hamiltonian system* consists of a $2n$-dimensional symplectic manifold (M, ω) and a Hamiltonian function $h \in C^\infty(M)$. The dynamics of the system are given by the flow of the Hamiltonian vector field v_h on M. We have in mind the example in which $M = T^*C$ is the cotangent bundle of configuration space and h is the Hamiltonian constructed from a time-independent Lagrangian, but the same ingredients, symplectic manifold and Hamiltonian, are also found in other contexts.

Suppose that $f \in C^\infty(M)$ is some other function on M, with Hamiltonian vector field v_f. Then

$$[f, h] = -2\omega(v_f, v_h) = v_h(f),$$

and by Lemma 8.18,

$$[v_f, v_h] = -v_{[f,h]}.$$

We therefore have the following lemma

Lemma 8.22

If $v_h(f) = 0$, then $[v_f, v_h] = 0$ and the flows of v_f and v_h commute.

So when $[f, h] = 0$, we can find a family of maps

$$\rho_{(s,t)} : M \to M \qquad s, t \in \mathbb{R}$$

such that, with m and t constant,

$$s \mapsto \rho_{(s,t)}(m)$$

gives the trajectories of v_f, and with m and s constant,

$$t \mapsto \rho_{(s,t)}(m)$$

gives the trajectories of v_h. In particular, the flow of v_f maps dynamical trajectories h to dynamical trajectories of h.

A Hamiltonian system is *integrable* if there exist n functions $k_a \in C^\infty(M)$ such that

(1) The vectors $v_{k_1}, \ldots, v_{k_n}$ are everywhere linearly independent, and

(2) $[k_a, k_b] = 0 = [k_a, h]$ for all a, b.

Functions with vanishing Poisson brackets are said to be *in involution* and functions with linearly independent gradients are said to be *independent*, so the definition requires the existence of n independent constants in involution. When the conditions are satisfied, the commuting flows of the vector fields v_{k_a} give a family of maps,

$$\rho_{\boldsymbol{s}} : M \to M \qquad \boldsymbol{s} = (s_1, \ldots, s_n).$$

The dynamical trajectories are contained within the n-dimensional submanifolds on which the k_as are constant, and the flows along the vector fields v_{k_a} map dynamical trajectories to dynamical trajectories.

Suppose that we are given an integrable system. Locally it is possible to find functions $s_1, \ldots, s_n$ such that together the s_as and k_as form a canonical system. That is, so that

$$\omega = \mathrm{d}s_a \wedge \mathrm{d}k_a.$$

The s_as are constructed by choosing a Lagrangian submanifold Σ transversal to the vector fields v_{k_a}, which means that at each $m' \in \Sigma$, the vector space $T_{m'}M$ is the direct sum of $T_{m'}\Sigma$ and the span of the vectors v_{k_a}. Such submanifolds are found by introducing local canonical coordinates p_a, q_a and defining Σ by (8.22). Transversality holds for a generic choice of coordinates and of S.

For each point m in M near Σ, there is a unique vector $\boldsymbol{s} = (s_1, \ldots, s_n)$ near the origin in $\mathbb{R}^n$ and a unique point $m' \in \Sigma$ such that

$$m = \rho_{\boldsymbol{s}}(m').$$

We take $s_a(m) = s_a$, so the Lagrangian submanifold Σ is given $s_a = 0$. Because $\rho_{\boldsymbol{s}}$ preserves ω for each $\boldsymbol{s}$, all the submanifolds

$$s_a = \text{constant}$$

are also Lagrangian. Therefore $[s_a, s_b] = 0$. We also have $[k_a, k_b] = 0$, and, by construction,

$$[s_a, k_b] = v_{k_b}(s_a) = \delta_{ab}.$$

Hence (8.25) holds and therefore the s_as and k_as form a system of canonical coordinates.

Because $v_{k_a}(h) = [h, k_a] = 0$, the Hamiltonian in these coordinates is independent of the s_as. It follows that Hamilton's equations take the form

$$\dot{k}_a = 0, \qquad \dot{s}_a = -\frac{\partial h}{\partial k_a}.$$

Therefore the dynamical evolution is given by

$$s_a = \ell_a - s\frac{\partial h}{\partial k_a}$$

with the k_a and the ℓ_as constant. So the dynamical trajectories are straight lines in the coordinates s_a.

There are some rough edges to the definition of integrability which make the concept somewhat elusive. Locally every Hamiltonian system is integrable in the sense of the definition, at least in a neighbourhood of a point at which $v_h \neq 0$, because it is easy to adapt the inductive proof of Darboux's theorem to construct a canonical coordinate system in which $h = p_n$. The momentum coordinates then provide the conserved quantities k_a, at least locally.

What makes the definition nontrivial and the concept interesting is the requirement that the k_as should be defined globally on M. We need to ensure here that M is the 'whole phase space', otherwise the conditions in the definition could be trivially satisfied by discarding all but a small portion of M. This is done by requiring in addition that

(3) The vector fields v_{k_a} are complete.

Still the concept is not captured in an entirely satisfactory way, as the following example illustrates.

Example 8.23

The 'Kepler system' consists of the symplectic manifold $M = T^*C$, where

$$C = \mathbb{R}^3 \setminus \{0\}$$

and the Hamiltonian

$$h = \tfrac{1}{2}(p_1^2 + p_2^2 + p_3^2) - \frac{1}{\sqrt{q_1^2 + q_2^2 + q_3^3}}.$$

With suitable normalization of the physical constants, it describes the dynamics of a particle moving in space under the influence of an inverse-square-law force. Here there are several conserved quantities: the Hamiltonian itself, the components of the angular momentum vector

$$\boldsymbol{J} = \boldsymbol{q} \wedge \boldsymbol{p}$$

and the components of the *Runge-Lenz vector*

$$\boldsymbol{R} = \boldsymbol{p} \wedge \boldsymbol{J} - \frac{\boldsymbol{q}}{|\boldsymbol{q}|}$$

(see Exercise 7.4). Moreover, the equations of motion can be integrated explicitly, to obtain Kepler's laws. According to any reasonable definition, the system is integrable.

The origin is excluded from the configuration space because otherwise the Hamiltonian is not defined everywhere. But the result is that the Hamiltonian vector field is not complete. Radial trajectories, that is, the orbits of particles falling directly towards the centre of force, reach the origin in finite time, and so the corresponding trajectories of v_h cannot be extended to all values of t.

In this example, one can in fact 'regularize' the problem by adding additional points to M (Moser [8]), but in general our definition of integrability needs to be relaxed to allow some incomplete trajectories. We do not attempt to explore further exactly how this should be done.

Example 8.24 (Action-Angle Variables)

A subset $\Lambda \subset M$ on which the k_as are constant is a Lagrangian submanifold of M. At each $m \in \Lambda$, vectors $v_{k_a}(m)$ form a basis for $T_m M$.

An obvious circumstance in which the completeness condition is satisfied is when all these Lagrangian submanifolds are actually compact. In this case, each Λ is an n-torus. That is, it is diffeomorphic to a product

$$C_1 \times C_2 \times \cdots \times C_n$$

of n circles. We do not prove this here, but it is a consequence of the fact that Λ admits n everywhere linearly independent vector fields.

The constants k_a label the different tori. We can change this labelling without changing the Λs or disturbing the conditions for integrability. In particular,

if ω is exact, so that $\omega = \mathrm{d}\theta$ for some $\theta \in \Omega^1(M)$, then we can take the k_as to be *action variables*, defined by integrating θ around the circles C_a:

$$k_a = \frac{1}{2\pi} \oint_{C_a} \theta.$$

The functions s_a that make up a canonical coordinate system with the k_as are called *angle variables*. The ath angle variable runs from 0 to 2π around the circle C_a, while the others are single-valued on this circle.

To see this, write θ in terms of its components in the canonical coordinates s_a, k_a

$$\theta = \alpha_a \, \mathrm{d}s_a + \beta_a \, \mathrm{d}k_a.$$

With $u_a = v_{s_a}$, we then have

$$\begin{aligned}
\oint_{C_a} \mathrm{d}s_b &= -\oint_{C_a} \mathrm{i}_{u_b}\omega \\
&= -\oint_{C_a} \mathrm{i}_{u_b} \mathrm{d}(\alpha_c \, \mathrm{d}s_c + \beta_c \, \mathrm{d}k_c) \\
&= \oint_{C_a} \left(\frac{\partial \alpha_c}{\partial k_b} \, \mathrm{d}s_c - \frac{\partial \beta_b}{\partial s_c} \, \mathrm{d}s_c \right) \\
&= \frac{\partial}{\partial k_b} \oint_{C_a} \alpha_c \, \mathrm{d}s_c - \oint_{C_a} \mathrm{d}\beta_b \\
&= 2\pi \frac{\partial k_a}{\partial k_b} \\
&= 2\pi \delta_{ab},
\end{aligned}$$

because the β_bs are single-valued, and therefore the second integral in the penultimate line vanishes. In an integrable system, the action-angle variables form a natural canonical coordinate system. The action variables label the Lagrangian tori and the angle variables label the points on the circles making up the tori.

Example 8.25 (Separation of the Hamilton–Jacobi Equation)

Let $M = T^*C$ be a cotangent bundle, with its canonical symplectic structure, and suppose that

$$h(q,p) = \tfrac{1}{2} T_{ab}(q) p_a p_b + U(q)$$

is the sum of the kinetic energy and the potential in some time-independent system. If the system is integrable, then the transformation to the canonical coordinates s_a, k_a is determined by a generating function

$$S(q_1, \ldots, q_n, k_1, \ldots, k_n)$$

(see Section 7.4). The condition that h should be a function only of the k_as, and therefore that the k_as should be constants of the motion, is that S should be a family of solutions to the *time-independent Hamilton–Jacobi equation*

$$h\left(q_1, \ldots, q_n, \frac{\partial S}{\partial q_1}, \ldots, \frac{\partial S}{\partial q_n}\right) = \text{constant}, \tag{8.26}$$

with the individual solutions labelled by the k_as. For each fixed set of values of the k_as, the equations

$$p_a = \frac{\partial S}{\partial q_a}$$

determine the Lagrangian submanifold of constant k_a. The k_as are expressed as functions of the p_as and q_as, and thence as elements of $C^\infty(T^*C)$, by inverting these equations.

It is not enough simply to show the existence of a local family of solutions to establish integrability, because this only establishes the existence of a local canonical coordinate system in which the Hamiltonian depends only on the k-coordinates. Such systems always exist, although finding them explicitly is not easy. To demonstrate integrability, it is necessary to find global constants of the motion.

One case in which this can be done is that in which (8.26) is *completely separable*, which means that all the coordinates can be separated in the equation. To say that the first coordinate q_1 can be separated means that there is a family of solutions of the form

$$S = S_1(q_1, k_1, k_2, \ldots, k_n) + S_2(q_2, q_3, \ldots, q_n, k_1, k_2, \ldots, k_n)$$

where S_1 depends on $k_1, \ldots, k_n$ and q_1, but not on the other q variables, and S_2 is independent of q_1. In this case, we can replace one of the k_as by a constant of the motion which is either linear or quadratic in the momenta.

To do this, we put

$$w(k_1, k_2, \ldots, k_n) = \left.\frac{\partial S}{\partial q_1}\right|_{q_1=0}.$$

Because w is a function only of the k_as, it is also a constant of the motion.

Now write h as a quadratic in p_1,

$$h = Ap_1^2 + Bp_1 + C,$$

with coefficients that depend on the other p_as and the q_as, and put

$$b = \frac{B(p_2, \ldots, p_n, \kappa, q_2, \ldots, q_n)}{A(p_2, \ldots, p_n, \kappa, q_2, \ldots, q_n)}, \quad c = \frac{C(p_2, \ldots, p_n, \kappa, q_2, \ldots, q_n) - h}{A(p_2, \ldots, p_n, \kappa, q_2, \ldots, q_n)},$$

where κ is some chosen constant. Then at $q_1 = \kappa$,

$$w^2 + bw + c = 0. \tag{8.27}$$

In the coordinate system $q_1, \dots, q_n, k_1, \dots, k_n$, none of the terms in this equation depend on q_1. It therefore holds everywhere.

On taking the Poisson bracket with h, we obtain

$$[b, h]w + [c, h] = 0$$

and hence from (8.27), either

$$[c, h]\big([c, h] - b[b, h]\big) = -c[b, h]^2 \qquad \text{or} \qquad [b, h] = [c, h] = 0.$$

All the terms in the first equation are polynomial in the p_as. So if it holds, then $[b, h]$ divides $[c, h]$, and w is a linear function of the momenta. In the second case c, which is a quadratic function of the moment, and b, which is a linear function of the momenta, are constants of the motion. It is possible that the Hamiltonian contains no cross terms between p_1 and the other momenta, in which case b vanishes identically, but c certainly does not vanish identically. We conclude that if q_1 separates, then there exists a constant of the motion that is either quadratic or linear in the p_as.

By repeating this argument for the other coordinates, we see that when the time-independent Hamilton–Jacobi equation is completely separable, the k_as can be replaced by constants that all are either quadratic or linear in the momenta. The completeness conditions are not yet obviously satisfied, but at least the result is a global description of the Lagrangian submanifolds.

For example, the r-coordinate in the Hamiltonian of the Kepler system in spherical polar coordinates can be separated. Here

$$h = \tfrac{1}{2}\left(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\varphi^2}{r^2 \sin^2\theta}\right) - \frac{1}{r}$$

and so

$$c = \frac{1}{\kappa^2}\left(p_\theta^2 + \frac{p_\varphi^2}{\sin^2\theta}\right) - \frac{2}{\kappa} - 2h$$

is a constant of the motion, quadratic in the momenta. It is a combination of the total energy and the square of the angular momentum. The coordinate φ also separates, giving rise to a linear constant.

Example 8.26 (Tops)

The symmetric top in Example 5.9, with Hamiltonian

$$h = \tfrac{1}{2}A(\dot\varphi^2 \sin^2\theta + \dot\theta^2) + \tfrac{1}{2}C(\dot\psi + \dot\varphi\cos\theta)^2 + mga\cos\theta,$$

is integrable, with constants of the motion

$$\begin{aligned} n &= \dot{\psi} + \dot{\varphi}\cos\theta \\ j &= A\dot{\varphi}\sin^2\theta + Cn\cos\theta, \end{aligned}$$

together with h itself. The constants can be found from the separability of the Hamilton–Jacobi equation.

Another classical example is Kovaleskaya's top, which has inertia matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$$

and centre of mass at unit distance from the fixed point, in the plane orthogonal to the axis of symmetry. See Exercise 5.19. This case is interesting because it is not one in which the Hamilton–Jacobi equation is separable.

9
Epilogue: Relativity and Quantum Theory

9.1 The Relevance of Classical Mechanics

Classical mechanics is superseded as a physical theory by relativity and quantum theory, but it still has a critical part to play in our understanding of the world. First, and most obviously, as an approximation. The fact that tests of relativity and quantum theory require delicate experiment and very precise observation is testament to the great precision of the Newtonian theory as a description of everyday phenomena. It still underpins a huge range of engineering and scientific applications.

Second, as a paradigm of a good theory. Other applications of mathematics, in economics, biology, and so forth, seek to emulate its self-consistency, clear mathematical framework, and strong predictive power.

Third, even in the areas in which we now know that it is inapplicable, classical mechanics is still very important as a source of technical tools and, in quantum theory, as the framework for interpretation.

9.2 Relativity

Einstein's special theory begins with a critical examination of the concept of 'inertial frame' and of Galileo's principle that all inertial frames are equivalent for the description of physical phenomena. The difficulty that Einstein addressed

N.M.J. Woodhouse, *Introduction to Analytical Dynamics: New Edition*,
Springer Undergraduate Mathematics Series, DOI 10.1007/978-1-84882-816-2_9,

was that the principle did not seem to apply to electromagnetism. Maxwell's equations are not invariant under transformations between uniformly moving frames of reference, despite the fact that electromagnetic phenomena seem unaffected by motion. His resolution kept the principle of relativity, but changed the transformation between the coordinate systems of different frames: under the new transformation, the *Lorentz transformation*, the time coordinate has a more complicated behaviour than the simple translation

$$t \mapsto t + \text{constant}$$

in the classical picture. Both the time intervals between events and the notion of simultaneity depend on the choice of frame. They are 'relative'. A clear understanding of the role of moving frames in classical theory is needed to understand what has changed in the new theory.

The idea of using the calculus of variations and Hamilton's principle to determine the dynamical trajectories of a mechanical system is one that has rich and central applications in relativistic physics. Relativistic particle motion can be characterized in much the same way in variational terms. For example, in both special and general relativity, the history of a free particle in space-time maximizes proper time between two nearby events.

Classical field equations can similarly be derived from variations of an 'action functional', which is an integral over space-time of a Lagrangian density that depends on the field variables and their derivatives. This is true of the electromagnetic field, of gauge fields and of the gravitational field itself. Such action principles not only give a very compact and transparent way to encode field equations that may be far from simple when written out in full, but also guarantee that the equations are independent of the choice of space-time coordinates, and are therefore compatible with the principle of relativity. The way in which Hamilton's principle implies the invariance of Lagrange's equations under change of generalized coordinates provides the model for such arguments.

9.3 Quantum Theory

The connection between classical mechanics and quantum theory is more intimate. Classical mechanics provides the starting point for the construction of quantum theories through 'quantization', as well as the arena in which its predictions are interpreted. It also determines the limiting behaviour in contexts in which the value of Planck's constant is negligible.

The connections between the classical and quantum worlds are deep and varied. Indeed the formulation of classical mechanics in terms symplectic geometry enables the construction of a dictionary of analogies between the mathematical objects that appear in quantum theory and corresponding classical geometrical objects: the quantum Hilbert space corresponds to the classical phase space, quantum operators representing physical observables correspond to functions on the classical phase space, commutators correspond to Poisson brackets, quantum symmetries correspond to canonical transformations, and irreducible representations correspond to coadjoint orbits.

Quantization methods exploit this network of analogies in seeking to reconstruct a quantum theory from its classical limit. Dirac's original method starts with canonical coordinates p_a, q_a and replaces them with operators $\hat{p}_a, \hat{q}_a$ satisfying the canonical commutation relations

$$[\hat{q}_a, \hat{p}_b] = \mathrm{i}\hbar\delta_{ab}$$

More generally, "geometric quantization" seeks to recover a quantum theory from the Hamiltonian formulation of the classical theory. The idea is to seek to construct a Hilbert space from a symplectic manifold (M, ω), and to associate operators on the Hilbert space with at least some classical observables, represented by elements of $C^\infty(M)$ act as operators. If $f, g \in C^\infty(M)$, then the corresponding operators, identified by carets, are required to satisfy the *quantization condition*

$$[\hat{f}, \hat{g}] = \mathrm{i}\hbar\widehat{[f, g]},$$

where the bracket on the left is the commutator and that on the right is the Poisson bracket. Thus classical functions with non-vanishing Poisson brackets correspond to quantum observables that cannot be measured simultaneously.

Dirac's quantization starts with the Hamiltonian formulation. Feynman's path-integral method instead starts from the Lagrangian. Given the Lagrangian $L(q, v, t)$ of a classical system, one constructs an *amplitude*

$$\langle (x, t_1), (y, t_2) \rangle$$

for a pair of configurations, $q_a = x_a$ at time t_1 and $q_a = y_a$ at time t_2. The square of its modulus is the probability of observing the configuration y at time t_2 given the configuration x at time t_1. Feynman's amplitude is given by a "sum over histories"

$$\langle (x, t_1), (y, t_2) \rangle = \int \exp\left(\frac{\mathrm{i}}{\hbar}\int_{t_1}^{t_2} L(q, \dot{q}, t)\,\mathrm{d}t\right)\,\mathrm{d}\gamma$$

where the outer integral is over all trajectories from the first configuration at time t_1 to the second at time t_2. It is a very difficult problem to make good

sense of this integral over a space of paths, but whatever its formal definition, an analogy with the principle of stationary phase suggests that the dominant contribution should come from the stationary points of the action

$$\int_{t_1}^{t_2} L(q, \dot{q}, t)\, \mathrm{d}t$$

and therefore, by Hamilton's principle, from the classical dynamical trajectories.

In the other direction, Hamilton–Jacobi theory provides the connecting link between the quantum description of particle motion by Schrödinger's equation and the limiting classical behaviour. For example, a particle moving in space under the influence of a potential U, Schrödinger's equation is

$$\mathrm{i}\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi + U\psi.$$

In the WKB approximation, one substitutes

$$\psi = A(q, t, \hbar) \exp\left(\frac{i}{\hbar} S(q, t)\right),$$

where A is an asymptotic series in powers of $\hbar$, and considers the behaviour as $\hbar \to 0$. The leading terms in the equation give

$$\frac{\partial S}{\partial t} + h\left(q, \frac{\partial S}{\partial q}, t\right) = 0$$

where

$$h = \frac{1}{2m}\left(p_1^2 + p_2^2 + p_3^2\right) + U$$

is the classical Hamiltonian. So the leading approximation as $\hbar \to 0$ is the Hamilton–Jacobi equation of the corresponding classical Hamiltonian. This connection between quantum evolution equations and the limiting classical dynamics is universal.

Appendix A: Notes on Exercises

1.1 Solutions: (a) 1, (b) 3, (c) 2, (d) 4.

1.3 The rotations permute the coordinate axes: think what happens when you rotate a cube through 120° about a long diagonal. So the matrices are

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

1.8 It saves work to remember that $\operatorname{tr}(ABC) = \operatorname{tr}(CAB)$.

1.20 Put $\boldsymbol{r} = \alpha\boldsymbol{\omega} + \beta\,\tilde{\mathrm{D}}\boldsymbol{\omega} + \gamma\boldsymbol{\omega} \wedge \tilde{\mathrm{D}}\boldsymbol{\omega}$. Show that if $\tilde{\mathrm{D}}\boldsymbol{\omega} \wedge \boldsymbol{r} + \boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge \boldsymbol{r}) = 0$, then $\alpha = \beta = \gamma = 0$.

1.23 Consider the motion relative to the frame $(O, (\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}))$. The radius is $ab\Omega\big(a^2(n-\Omega)^2 + b^2\Omega^2\big)^{-1/2}$.

3.6 For the first part, multiply the equation of motion by v_a and sum over a.

3.12 Let O denote the midpoint of AB. Take the x-axis along OB and the y-axis vertically upwards. Then the coordinates of P are $x = b\cosh\varphi\cos\theta$, $y = -b\sinh\varphi\sin\theta$.

3.13 Write (2.36) in the form

$$\frac{\partial^2 L}{\partial v_a \partial q_b} v_b + \frac{\partial^2 L}{\partial v_a \partial v_b} \dot{v}_b + \frac{\partial^2 L}{\partial v_a \partial t} - \frac{\partial L}{\partial q_a} = 0. \tag{A.1}$$

This holds for all values of q_a, v_a, $\dot{v}_a$ and t. Deduce that $\partial^2 L/\partial v_a \partial v_b = 0$ and hence that $L = A_a(q,t)v_a + B(q,t)$. Substitute into (A.1) to get

$$\frac{\partial A_a}{\partial q_b} = \frac{\partial A_b}{\partial q_a}, \qquad \frac{\partial A_a}{\partial t} = \frac{\partial B}{\partial q_a}.$$

These imply that $A_a = \partial f/\partial q_a$ and $B = \partial f/\partial t$ for some function $f = f(q,t)$.

3.14 Use the polar angles θ and φ as coordinates. Without loss of generality, choose the polar axis so that $\theta = \pi/2$, $\dot\theta = 0$ initially.

4.7 For the second part, differentiate

$$u_a \frac{\partial L}{\partial q_a} + \left(\frac{\partial u_a}{\partial q_b} v_b + \frac{\partial u_a}{\partial t} \right) \frac{\partial L}{\partial v_a} = 0$$

with respect to v_b.

5.4 Write

$$T = \tfrac{1}{2} \int_0^1 m\big(t\boldsymbol{u} + (1-t)\boldsymbol{v}\big) \cdot \big(t\boldsymbol{u} + (1-t)\boldsymbol{v}\big)\, \mathrm{d}t.$$

5.6 For the last part, consider a circular disc of radius a and find a point of inertial symmetry on the axis of the disc.

5.8 Substitute $\omega_2 = K \tanh u$, where K is an appropriate constant.

5.9 Use (5.12) and (5.13) to write ω_1 and ω_3 in terms of T, J^2, and ω_2. Now substitute for ω_1 and ω_3 in $B^2\dot\omega_2^2 = (A-C)^2\omega_3^2\omega_1^2$.

5.10 The intersection of the instantaneous axis with the surface has coordinates $(\lambda\omega_1, \lambda\omega_2, \lambda\omega_3)$ where $\lambda = \pm k\sqrt{2T}$. The tangent planes at these points are given by

$$A\omega_1 x + B\omega_2 y + C\omega_3 z = \lambda(A\omega_1^2 + B\omega_2^2 + C\omega_3^2)$$

or, alternatively, by $\boldsymbol{J}_0 \cdot \boldsymbol{r} = \pm k\sqrt{2T}$. But $\boldsymbol{J}_0$ is fixed relative to the inertial frame and T is constant.

The shape of the surface is not important: the ellipsoid can be an imaginary surface in the body with equation $Ax^2 + By^2 + Cz^2 = k^2$. The motion is then such that the imaginary surface appears to be rolling between two fixed plane – a result due to Poinsot.

5.12 The angular momentum of the smaller sphere about its centre of mass is $\frac{2}{5}ma^2\boldsymbol{\omega}$ where m is its mass and $\boldsymbol{\omega}$ is its angular velocity. Hence if $\boldsymbol{R}$ is the force at the point of contact, then

$$ma\ddot{\boldsymbol{e}} = -mg\boldsymbol{k} + \boldsymbol{R}, \qquad \tfrac{2}{5}ma^2\dot{\boldsymbol{\omega}} = a\boldsymbol{e} \wedge \boldsymbol{R},$$

by the principles of linear and angular momentum. The rolling condition at the point of contact is $a\dot{\boldsymbol{e}} + \boldsymbol{\omega} \wedge (a\boldsymbol{e}) = 0$.

5.13 C. E. Easthope gives a full discussion in [4]. He also makes an interesting remark about golf.

Denote the mass of the sphere by m, its radius by a and let the radius of the cylinder be $a + c$. The centre of the sphere has position vector $\boldsymbol{r} = z\boldsymbol{k} + c\boldsymbol{e}$ from a fixed origin on the axis of the cylinder, where $\boldsymbol{k}$ and $\boldsymbol{e}$ are orthogonal unit vectors, with $\boldsymbol{k}$ pointing vertically upwards. Let$\boldsymbol{f} = \boldsymbol{k} \wedge \boldsymbol{e}$.

The triad $(\boldsymbol{e}, \boldsymbol{f}, \boldsymbol{k})$ is orthonormal and has angular velocity $\Omega\boldsymbol{k}$ with respect to fixed axes, where Ω is some function of time.

Let $\boldsymbol{\omega}$ denote the angular velocity of the sphere. Put $n = \boldsymbol{\omega} \cdot \boldsymbol{e}$ and $N = \boldsymbol{\omega} \cdot \boldsymbol{f}$.

With the dot denoting the time derivative with respect to fixed axes, the equations of motion are

$$\tfrac{2}{5}ma^2\dot{\boldsymbol{\omega}} = a\boldsymbol{e} \wedge \boldsymbol{R}, \qquad m(\ddot{z}\boldsymbol{k} + c\ddot{\boldsymbol{e}}) = \boldsymbol{R} - mg\boldsymbol{k},$$

where $\boldsymbol{R}$ is the force at the point of contact. The rolling condition is

$$\dot{z}\boldsymbol{k} + c\dot{\boldsymbol{e}} + a\boldsymbol{\omega} \wedge \boldsymbol{e} = 0.$$

The following steps lead to the stated result.

(1) From all three equations

$$\tfrac{2}{5}a\dot{\boldsymbol{\omega}} = \boldsymbol{e} \wedge (-a\dot{\boldsymbol{\omega}} \wedge \boldsymbol{e} - a\boldsymbol{\omega} \wedge \dot{\boldsymbol{e}} + g\boldsymbol{k}).$$

Hence $\boldsymbol{e} \cdot \dot{\boldsymbol{\omega}} = 0$ and $\frac{7}{5}a\dot{\boldsymbol{\omega}} = (\Omega na - g)\boldsymbol{f}$.

(2) From the rolling condition

$$\dot{z} = aN \quad \text{and} \quad 0 = c\Omega + a\boldsymbol{\omega} \cdot \boldsymbol{k}.$$

(3) $\dot{\boldsymbol{\omega}} \cdot \boldsymbol{k} = 0$ and hence Ω is constant.

(4) By considering $\dot{\boldsymbol{\omega}} \cdot \boldsymbol{f}$,

$$7\ddot{N} + 2\Omega\dot{n} = 0.$$

(5) By considering $\dot{\boldsymbol{\omega}} \cdot \boldsymbol{e}$, $\dot{n} = \Omega N$.

(6) $7\ddot{z} + 2\Omega^2\dot{z} = 0$.

5.14 Part (b): remember that $A\Omega^2 2u_0 - Cn\Omega + mga = 0$ and that

$$\Omega = \frac{j - Cnu_0}{A(1 - u_0^2)}.$$

5.16 Use the fact that φ is cyclic to show that

$$\dot{\varphi} = \frac{V \sin\alpha}{a \sin^2\theta}.$$

Show that

$$\dot{\theta}^2(1 + 3\cos^2\theta) + 4\dot{\varphi}^2 \sin^2\theta$$

is constant. Write $u = \cos\theta$ and find an expression for $\mathrm{d}t/\mathrm{d}u$.

5.17 Let α denote the angle between the downward vertical and a line joining the centre of the smaller cylinder to a point on its rim and let x denote the horizontal distance of the centre of the larger cylinder from a fixed point. Let $(\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k})$ be an orthonormal triad, with $\boldsymbol{k}$ vertical and $\boldsymbol{i}$ orthogonal to the axes of the cylinders. Then the angular velocities of the cylinders are $\dot{\theta}\boldsymbol{j}$ (large) and $\dot{\alpha}\boldsymbol{j}$ (small). The velocities of the centres are $\dot{x}\boldsymbol{i}$ (large) and

$$\dot{x}\boldsymbol{i} - (\tfrac{1}{2}a\cos\varphi)\dot{\varphi}\boldsymbol{i} + (\tfrac{1}{2}a\sin\varphi)\dot{\varphi}\boldsymbol{k}$$

(small). The vector $\boldsymbol{e} = -\cos\varphi\boldsymbol{k} - \sin\varphi\boldsymbol{i}$ is a unit vector pointing from the centre of the larger cylinder to the centre of the smaller cylinder. The rolling conditions are

$$\dot{x} - a\dot{\theta} = 0;$$

and

$$\dot{x}\boldsymbol{i} + \dot{\theta}\boldsymbol{j} \wedge (a\boldsymbol{e}) = \dot{x}\boldsymbol{i} - (\tfrac{1}{2}a\cos\varphi)\dot{\varphi}\boldsymbol{i} + (\tfrac{1}{2}a\sin\varphi)\dot{\varphi}\boldsymbol{k} + \dot{\alpha}\boldsymbol{j} \wedge (\tfrac{1}{2}a\boldsymbol{e}),$$

which gives $\dot{\alpha} = 2\dot{\theta} - \dot{\varphi}$ and hence $\alpha = 2\theta - \varphi + \text{constant}$.

Take θ and φ as generalized coordinates, and choose origins such that $x = a\theta$ and $\alpha = 2\theta - \varphi$. Then the total kinetic energy is

$$\tfrac{1}{2}ma^2\dot{\theta}^2 + \tfrac{1}{2}m(a\dot{\theta} - \tfrac{1}{2}a\dot{\varphi}\cos\varphi)^2 + \tfrac{1}{8}ma^2\dot{\varphi}^2\sin^2\varphi + \tfrac{1}{2}ma^2\dot{\theta}^2 + + \tfrac{1}{8}ma^2(2\dot{\theta} - \dot{\varphi})^2.$$

The result follows from conservation of energy ($\partial L/\partial t = 0$).

5.22 In the notation of (5.13) above: use as coordinates θ, φ, ψ, χ, and z, where θ, φ, and ψ are the Euler angles of a triad fixed relative in the sphere relative to the triad $(\boldsymbol{e}, \boldsymbol{f}, \boldsymbol{k})$, χ is the angle between $\boldsymbol{e}$ and a fixed horizontal line, and z is the height of the centre of the sphere above a fixed origin on the axis of the cylinder.

Then

$$\begin{aligned} \boldsymbol{\omega} \cdot \boldsymbol{k} &= \dot{\chi} + \dot{\varphi} + \dot{\psi}\cos\theta \\ n &= -\dot{\theta}\sin\varphi + \dot{\psi}\sin\theta\cos\varphi \\ N &= \dot{\theta}\cos\varphi + \dot{\psi}\sin\theta\sin\varphi. \end{aligned}$$

The Lagrangian is

$$\begin{aligned} L \;=\; & \tfrac{1}{5}ma^2\big(\dot\theta^2+\dot\psi^2+(\dot\varphi+\dot\chi)^2+2\dot\psi(\dot\varphi+\dot\chi)\cos\theta\big) \\ & +\tfrac{1}{2}m(\dot z^2+c^2\dot\chi^2)-mgz \end{aligned}$$

and the rolling conditions are

$$\begin{aligned} \dot z - a(\dot\theta\cos\varphi+\dot\psi\sin\theta\sin\varphi) \;&=\; 0 \\ c\dot\chi + a(\dot\varphi+\dot\chi+\dot\psi\cos\theta) \;&=\; 0, \end{aligned}$$

corresponding to which we have the Lagrange multipliers λ and μ.

The φ and χ equations give that $\lambda = 0$ and that $\dot\chi = \Omega$ is constant. After some manipulation, the ψ and θ equations give

$$\begin{aligned} \tfrac{2}{5}ma\ddot\psi\sin\theta+\tfrac{4}{5}ma\dot\psi\dot\theta\cos\theta+\tfrac{2}{5}mc\dot\chi\dot\theta \;&=\; -\mu\sin\varphi \\ \tfrac{2}{5}ma\ddot\theta-\tfrac{2}{5}ma\dot\psi^2\sin\theta\cos\theta-\tfrac{2}{5}mc\dot\chi\dot\psi\sin\theta \;&=\; -\mu\cos\varphi. \end{aligned}$$

The z equation is $m\ddot z + mg = \mu$. A little further work leads to $\dot n = \Omega N$ and $7\dot N + 2\Omega n + 5g/a = 0$ and hence $7\;\dddot z + 2\Omega^2\dot z = 0$, as before.

6.2 The roots are $\lambda = 1$, $\lambda = 1/2$, and $\lambda = 3/2$.

6.3 One might be tempted to begin by introducing four coordinates, the spherical polar angles θ_1 and φ_1 of A (with polar axis PA), and the two spherical polar angles θ_2 and φ_2 of B (with polar axis QB). This will not work, however. The polar coordinates are singular in the equilibrium configuration, as is reflected by the fact that large changes in φ_1 and φ_2 can correspond to small displacements in the system.

6.4 Note that the system has four degrees of freedom.

6.5 Take the origin to be the equilibrium position of the particle, with axes chosen so that the position vectors of A, B, C, and D are $\boldsymbol{a} = (1, 1, \sqrt{2})$, $\boldsymbol{b} = a(1, -1, \sqrt{2})$, $\boldsymbol{c} = a(-1, -1, \sqrt{2})$, and $\boldsymbol{d} = a(-1, 1, \sqrt{2})$. Suppose that the particle is at the point P with position vector $\boldsymbol{r} = (x, y, z)$. To the second order in x, y, z,

$$PA = 2a\left(1 - \frac{\boldsymbol{a}\cdot\boldsymbol{r}}{4a^2} + \frac{\boldsymbol{r}\cdot\boldsymbol{r}}{8a^2} - \frac{(\boldsymbol{a}\cdot\boldsymbol{r})^2}{32a^4}\right).$$

Therefore, also to the second order, the elastic potential energy of the string PA is

$$\frac{\lambda(PA-a)^2}{2a} = \frac{\lambda}{2a}\left(a^2 - \boldsymbol{a}\cdot\boldsymbol{r} + \tfrac{1}{2}\boldsymbol{r}\cdot\boldsymbol{r} + \frac{1}{8a^2}(\boldsymbol{a}\cdot\boldsymbol{r})^2\right)$$

We have

$$(\boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}+\boldsymbol{d})\cdot\boldsymbol{r}=4a\sqrt{2}z$$

and

$$(\boldsymbol{a}\cdot\boldsymbol{r})^2+(\boldsymbol{b}\cdot\boldsymbol{r})^2+(\boldsymbol{c}\cdot\boldsymbol{r})^2+(\boldsymbol{d}\cdot\boldsymbol{r})^2=4a^2(x^2+y^2+2z^2).$$

Hence

$$U=\frac{\lambda a}{2}\left(4-\frac{4\sqrt{2}z}{a}+\frac{5x^2}{2a^2}+\frac{5y^2}{2a^2}+\frac{3z^2}{a^2}\right)+mgz.$$

7.6 Consider

$$\frac{\partial(q,p,t)}{\partial(q',p',t')}\quad\text{and}\quad\frac{\partial(q'',p'',t'')}{\partial(q',p',t')}.$$

7.10 The vector field $\boldsymbol{x}$ is tangent to Σ if and only if

$$\boldsymbol{x}\cdot\ \text{grad}\left(p-\frac{\partial S}{\partial q}\right)=0$$

on Σ, where grad is the gradient operator

$$\text{grad}\ =\boldsymbol{i}\frac{\partial}{\partial q}+\boldsymbol{j}\frac{\partial}{\partial p}+\boldsymbol{k}\frac{\partial}{\partial t}.$$

This is equivalent to

$$\begin{aligned}0\ &=\ -\frac{\partial h}{\partial q}-\frac{\partial^2 S}{\partial q^2}\frac{\partial h}{\partial p}-\frac{\partial^2 S}{\partial q\partial t}\\ &=\ -\frac{\partial}{\partial q}\left(\frac{\partial S}{\partial t}+h\left(q,\frac{\partial S}{\partial q},t\right)\right).\end{aligned}$$

7.12 See Exercise (3.12).

8.12 Suppose that $\gamma\subset C\times\mathbb{R}$ is a kinematic trajectory given by $q_a=q_a(t)$. Then γ determines a *lifted trajectory* $\hat{\gamma}\subset TC\times\mathbb{R}$, given by

$$q_a=q_a(t),\qquad v_a=\dot{q}_a(t).$$

The flow of ρ_s carries a kinematic trajectory in $\gamma\subset C\times\mathbb{R}$ into a family γ_s of trajectories labelled by s, while the flow

$$\hat{\rho}_s:TC\times\mathbb{R}\to TC\times\mathbb{R}$$

of $\hat{u}$ maps $\hat{\gamma}$ to the family of lifted trajectories corresponding to $\hat{\gamma}_s$. If one uses this property to characterize the flow $\hat{\rho}_s$, and takes $\hat{u}$ to be the generating vector field, then it is simple to see that the components of $\hat{u}$ are given by (8.21) and hence that they must transform as the components of a vector on $TC\times\mathbb{R}$ vector under change of coordinates on $C\times\mathbb{R}$.

8.10 Use the special coordinate system.

Bibliography

[1] V. I. Arnol'd *Mathematical methods of classical mechanics*. Translated by K. Vogtmann and A. Weinstein. Springer-Verlag, New York, 1978.

[2] B. Carter *Hamilton–Jacobi and Schrödinger separable solutions of Einstein's equations*. Commun. Math. Phys. **10**, 280–310, 1968.

[3] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. I. Interscience, New York, 1953.

[4] C. E. Easthope *Three dimensional dynamics, a vectorial treatment*. 2nd edition. Butterworths, London, 1964.

[5] The quotation is from S. Drake's translation of Galileo Galilei *Dialogue concerning the two chief world systems—Ptolemaic and Copernican*. 2nd revised edition. University of California, Berkeley, California, 1967.

[6] S. Kobayashi and K. Nomizu *Foundations of Differential Geometry*, Volume I. Wiley, New York, 1963.

[7] J. E. Marsden and A. Weinstein *Reduction of symplecic manifolds with symmetry*. Rep. Math. Phys. **5**, 121–30, 1974.

[8] J. K. Moser *Regularization of Kepler's probelm and the avergaing method on a manifold*. Commun. Pure Appl. Math. **23**, 609–36, 1970.

[9] The quotations are from Motte's translation of the *Principia*, updated by F. Cajori: Isaac Newton: *The mathematical principles of natural philosophy and his system of the world*. University of California, Berkeley, 1934; they are reproduced in *Newton's philosophy of nature* (ed. H. S. Thayer). Hafner, New York, 1953.

[10] J. L. Synge and B. A. Griffith *Principles of mechanics*. 3rd edition: International student edition. McGraw-Hill Kogakusha, Tokyo, 1970.

[11] D. W. Jordan and P. Smith *Nonlinear ordinary differential equations* (Second edition). Oxford University Press, 1987.

[12] A. Weinstein *Symplectic manifolds and their Lagrangian submanifolds*. Adv. in Math. **6**, 329–46, 1971.

[13] E. T. Whittaker *A treatise on the analytical dynamics of particles and rigid bodies*. Cambridge, 1904, with many subsequent editions.

[14] E. P. Wigner *The unreasonable effectiveness of mathematics*. Commun. Pure Appl. Math. **13**, 1–14, 1960.

[15] N. M. J. Woodhouse *General Relativity*. Springer-Verlag, London, 2007.

Index