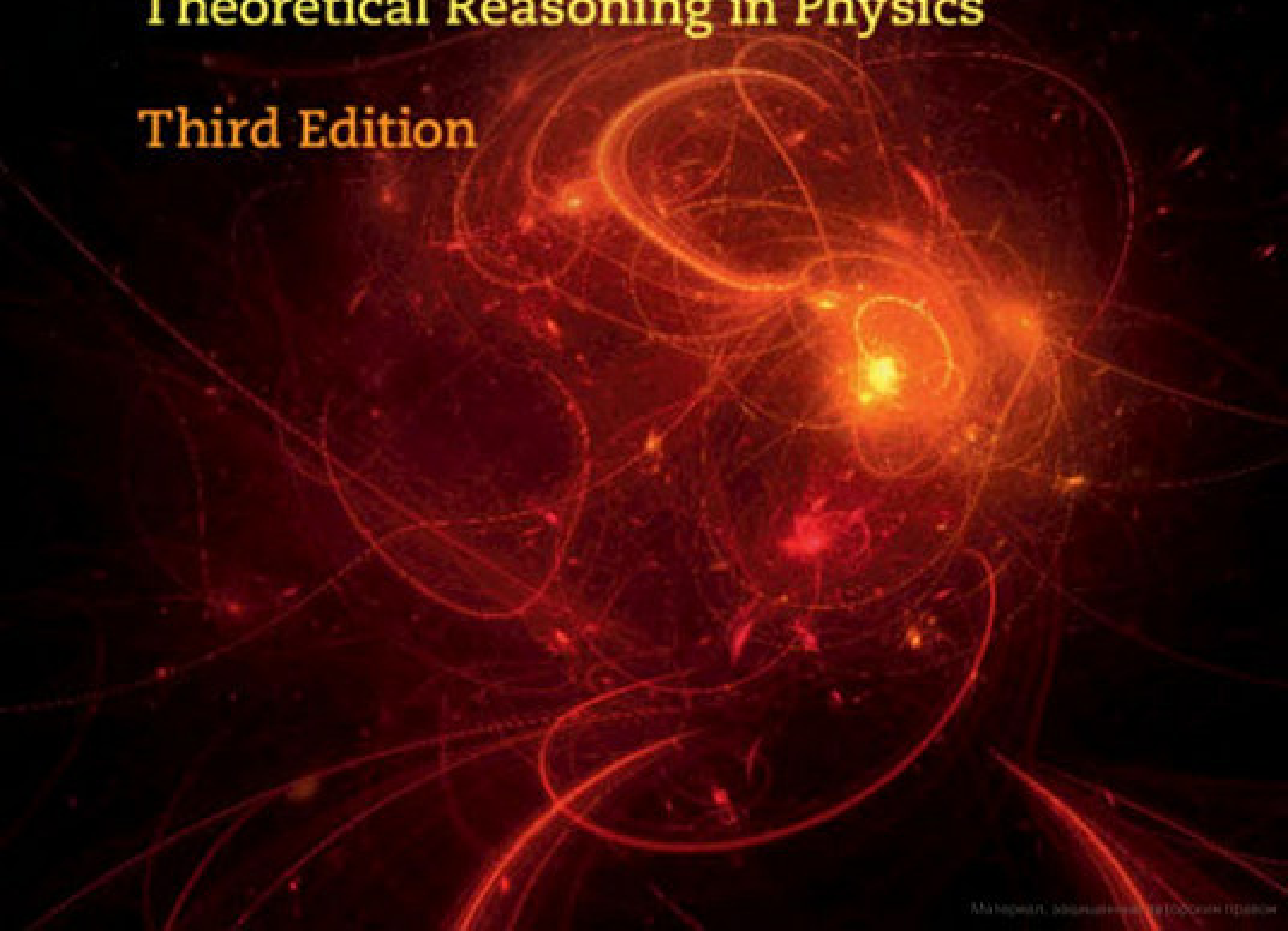
Malcolm Longair

Theoretical Concepts in Physics

An Alternative View of Theoretical Reasoning in Physics

Third Edition



Theoretical Concepts in Physics

Third Edition

In this original and integrated approach to theoretical reasoning in physics, Malcolm Longair illuminates the subject from the perspective of real physics as practised by research scientists. Concentrating on the basic insights, attitudes and techniques that are the tools of the modern physicist, this approach conveys the intellectual excitement and beauty of the subject.

Through a series of seven case studies, the contents of an undergraduate course in classical physics and the discovery of quanta are reviewed from the point of view of how the great discoveries and changes of perspective came about. This approach illuminates the intellectual struggles needed to attain understanding of some of the most difficult concepts in physics.

Longair's highly acclaimed text has been fully revised and includes new studies on the physics of fluids, Maxwell's great paper on equations for the electromagnetic field, and problems of contemporary cosmology and the very early Universe.

Malcolm Longair, CBE, FRS, FRSE, was Astronomer Royal for Scotland from 1980 to 1990 and Head of the Cavendish Laboratory from 1997 to 2005. His books with Cambridge University Press include *High Energy Astrophysics*, *The Cosmic Century*, *Quantum Concepts in Physics*, and *Maxwell's Enduring Legacy*, a scientific history of the Cavendish Laboratory.

Theoretical Concepts in Physics

An Alternative View of Theoretical Reasoning in Physics

Third Edition

MALCOLM LONGAIR
University of Cambridge

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781108484534
DOI: 10.1017/9781108613927

First published 1984
Second edition 2003
Third edition 2020

Printed in the United Kingdom by TJ International Ltd, Padstow Cornwall

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

Library of Congress Cataloging-in-Publication Data
Names: Longair, Malcolm S., 1941– author.
Title: Theoretical concepts in physics : an alternative view of theoretical reasoning in physics / Malcolm S. Longair, University of Cambridge.
Description: Third edition. | Cambridge, United Kingdom ; New York, NY : Cambridge University Press, [2020] | Includes bibliographical references and indexes.
Identifiers: LCCN 2020001145 (print) | LCCN 2020001146 (ebook) | ISBN 9781108484534 (hardback) | ISBN 9781108613927 (epub)
Subjects: LCSH: Mathematical physics.
Classification: LCC QC20 .L64 2020 (print) | LCC QC20 (ebook) | DDC 530.1–dc23
LC record available at https://lccn.loc.gov/2020001145
LC ebook record available at https://lccn.loc.gov/2020001146

ISBN 978-1-108-48453-4 Hardback

For Deborah

Contents

Preface and Acknowledgements

Origins

The inspiration for this book was a course of lectures which I delivered between 1977 and 1980 to undergraduates about to enter their final year in Physics and Theoretical Physics in the Cavendish Laboratory, the Physics Department of Cambridge University. The aim of the course was to provide a survey of the nature of theoretical reasoning in physics, which would put them in a receptive frame of mind for the very intensive courses of lectures on all aspects of physics in the final year. More details of the aims of the courses are set out in Chapter 1.

An important feature of the course was that it was entirely optional and strictly non-examinable, although the contents provided a great deal of revision and support material for the examinable courses. The lectures were delivered at 9 am every Monday, Wednesday and Friday during a four-week period in July and August, the old Cambridge Summer Term, prior to the final year of the three-year physics course. Despite the timing of the lectures, the fact that the course was not examinable, and the alternative attractions of Cambridge during the glorious summer months, the course was very well attended. I was very gratified by the positive response of the students and this encouraged me to produce a published version of the course with the inclusion of other materials from the various lecture courses I had delivered. I was not aware of any other book which covered the subject matter in quite the same way.

The first edition of the book was published in 1984 while I was in Edinburgh. The second edition, which appeared in 2002 after my return to Cambridge, included a number of significant enhancements and improvements to the story, as well as some quite new chapters. The whole book was rewritten from a more mature view point. Again, all the material was battle-tested on a succession of third-year students in the reorganised three or four year course in ‘Experimental and Theoretical Physics’. The book turned out to be popular and proved to be a steady seller, even 16 years after its publication.

In 2018, I was delighted to learn that Cambridge University Press suggested that I might consider writing a third edition. Remarkably, it coincided with the invitation to deliver the course once again to third-year students, more than 40 years after its first delivery and almost 10 years after my formal retirement from my University post! But the course remains a joy to teach and the Cambridge students are as engaged and wonderful as ever. I will never tire of expounding the remarkable ingenuity and creativity of the founders of modern physics.

My reason for welcoming this opportunity is that, since the 2002 edition, I have continued to dig deeper into essentially all the topics covered and this was an opportunity to build these insights into the book. Some of my old favorites remain more or less as before, but there are many significant new insights into many of the case studies and some new chapters. I have also tightened up the writing, reflecting my aim to be as concise and clear as possible.

Acknowledgements

The views expressed in the text are obviously all my own, but many of my Cambridge and Edinburgh colleagues played major roles in clarifying my ideas. The concept of the original course came from discussions with the late Sir Alan Cook, Volker Heine and John Waldram. I also inherited the Examples Class in Mathematical Physics from Volker Heine and the late J.M.C. Scott. Developing that class helped enormously in clarifying many of my own ideas. In later years, Brian Josephson helped with the course and provided some startling insights. The course in thermodynamics was given in parallel with one by Archie Howie and I learned a great deal from discussions with him.

In Edinburgh, Peter Brand, John Peacock and Alan Heavens contributed in important ways to my understanding. In Cambridge, many members of the Department have been very supportive of my endeavours to bring physics alive for successive generations of undergraduates. I am particularly grateful to John Waldram, David Green and Julia Riley for innumerable discussions concerning the courses we have delivered. I also acknowledge invaluable discussions with Steve Gull and Anthony Lasenby. Sanjoy Mahajan kindly took a special interest in the section on Dimensional Methods and critically reviewed what I wrote. A special debt of gratitude is due to the late Peter Harman, who kindly read my writings on Maxwell and made helpful suggestions.

Perhaps the biggest debts I owe in my education as a physicist are to the late Sir Martin Ryle and the late Peter Scheuer, who jointly supervised my research in the Radio Astronomy Group of the Cavendish Laboratory during the 1960s. I learned more from them about real physics than from anyone else. Another powerful influence was the late Sir Brian Pippard, whose penetrating understanding of physics was a profound inspiration. Although he and I had very different points of view, there is virtually no aspect of physics we discussed in which his insights did not added significantly to my understanding.

Almost as important as my debts to my Cambridge colleagues, are those to the late Yakov Borisevich Zeldovich and to my friend and colleague Rashid Alievich Sunyaev. I spent the academic year 1968–69 in Moscow as a Royal Society–USSR Academy of Sciences Exchange Research Fellow and they took a deep interest in the research I had carried out as a graduate student in the Radio Astronomy Group. During that year, it was an enormous privilege to work with them and learn so much astrophysics and cosmology from brilliant physicists brought up in a wholly different tradition of theoretical physics.

I must also acknowledge the stimulation provided over the years by the many generations of undergraduates who attended this and the other courses I have given. Their comments and enthusiasm were largely responsible for the fact that the first edition of the book appeared at all. The same remark applies to this new edition – Cambridge students are a phenomenal resource, which makes lecturing and teaching an enormous privilege and pleasure.

Grateful thank are due to the innumerable people who have helped in the preparation of all three editions of this book. In Edinburgh, Janice Murray, Susan Hooper and Marjorie Fretwell helped with the text and diagrams, while Brian Hadley and his colleagues in the Photolabs of the Royal Observatory, Edinburgh, prepared superb photographs. In Cambridge, I was greatly assisted by the late Judith Andrews and Leona Hope-Coles, as well as by the librarians in the Rayleigh Library, Nevenka Huntic and Helen Suddaby. Mark Hurn, Librarian at the Institute of Astronomy, was also very helpful in unearthing some of treasures in their library. Finally, enormous thanks are due to David Green in one of his other roles as a guru of the *LaTeX* typesetting program. Thanks to his great expertise, he found solutions to knotty typesetting issues which have added significantly to the appearance of the book.

As in all my endeavours, the debts I owe to my wife, Deborah, and our children, Mark and Sarah, cannot be adequately expressed in words.

Malcolm Longair

Figure Credits

I am most grateful to the publishers who have been most helpful in giving permission for the use of the figures from journals, books and other media for which they now hold the copyright. Every effort has been made to track down the copyright owners of all the pictures, but some of them have proved to be beyond reasonable effort. The origins of all the figures are given in the figure captions. In addition, the following list includes the original publishers of the figures and the publishers who now hold the copyright, as well as the specific forms of acknowledgement requested by them.

Astrophysical Journal. Reproduced by kind permission of the Astrophysical Journal, Figs. 20.3, 20.12 and 20.16. The images are © AAS.

arXiv astro-ph. Fig. 20.17.

Chicago University Press. Reproduced by kind permission of Chicago University Press, Figs. 3.7, VI.1, 18.1.

Creative Commons. Courtesy of Creative Commons, Figs, 9.3, 9.4, 9.13 and 10.12.

Cambridge University Press. Figs. 2.1, 2.3, 2.4, 2.7, 10.4, 10.5, 10.6, 10.7, 10.8, 10.9, 10.11, 16.2 and 16.4.

Deutsches Museum, Munich. Courtesy of the Deutsches Museum, Munich, Fig. 11.1.

Dictionary of Scientific Biography, Charles Scribner's Sons. Courtesy of Charles Scribner's Sons, Figs. 2.8 and 4.2(*a*).

European Space Agency. Courtesy of the European Space Agency and the associated Science Teams, Figs. VII.1(*a*) and 20.15.

Living Reviews. Courtesy of Living Reviews in Relativity, Fig. 19.15.
Macmillan and Company Ltd. Courtesy of Macmillan and Company Ltd, Figs. 11.3, 11.4, 14.4 and V.1.
Monthly Notices of the Royal Astronomical Society. Courtesy of the Royal Astronomical Society, Figs. 20.1 and 20.2.
NASA and the Space Telescope Science Institute. Courtesy of NASA and the Space Telescope Science Institute, Fig. VI.2.
Nature. Courtesy of Springer-Nature Publishing, Fig. 21.2.
Observatory. Courtesy of The Observatory, Fig. 20.4.
Oxford University Press. Courtesy of Oxford University Press, Figs. 4.1, 4.4, 4.5, 4.6, 10.13, 10.14, 10.15, 10.16, 10.17, 10.18, 11.7, 11.8 and 11.9.
Royal Institution of Great Britain. Courtesy of the Royal Institution of Great Britain, Figs. II.1 and 5.5.
Royal Society. Courtesy of the Royal Society, Figs. 10.2 and IV.1.
Simon and Schuster. Courtesy of Simon and Schuster publishers, Fig. 17.6.
David G. Simpson. By kind permission of Dr David G. Simpson. Fig. 10.1.
Sloan Telescope Project. By kind permission of the Sloan Telescope Project, Fig. VII.1(*b*).
Solvay International Institutes. By kind permission of the Solvay International Institutes, Fig. 17.3.
Toronto University Press. Courtesy of Toronto University Press, Figs. 3.1 and 3.5.
Viking/Penguin–Random House Publishers. Courtesy of Viking/Penguin–Random House Publishers, Fig. 10.10.

1 Introduction

1.1 Pedagogical and Real Physics

This book is for students who love physics and theoretical physics. It arises from a dichotomy which I believe pervades attempts to teach the ideal course in physics. On the one hand, there is the way in which physics and theoretical physics is presented in lecture courses and examples classes. On the other hand, there is the way we actually practise the discipline as professional physicists. In my experience, there is often little relation between these activities, which is a great misfortune.

There are, of course, good reasons why the standard lecture course has evolved into its present form. Physics and theoretical physics are not particularly easy subjects and it is important to set out the fundamentals and their applications in as clear and systematic a manner as possible. It is absolutely essential that students acquire a firm grounding in the basic techniques and concepts of physics. But we should not confuse this process with that of doing real physics. Standard lecture courses in physics and theoretical physics are basically ‘five-finger’ exercises, designed to develop technique and understanding. But such exercises are very different from performing Beethoven’s *Hammerklavier* sonata at the Royal Festival Hall. You are only really doing physics and theoretical physics when the answers *really* matter – when your reputation as a scientist hangs upon being able to reason correctly in a research context or, in more practical terms, when your ability to undertake original research determines whether you are employable, or whether your research grant is renewed. These are quite different processes from working through drill exercises, with answers provided at the back of the book.

The major problem is that there is just so much material which lecturers feel they have to include in their courses that physics syllabuses become seriously overloaded. There is little time to sit back and ask ‘What is this all about?’ Indeed, the technical aspects of the subject, which are themselves truly fascinating, can become so totally absorbing that it is generally left to the students to find out for themselves many essential truths about physics.

It is important to stress at the outset that this book is *not* a textbook on physics and theoretical physics. There is *no substitute* for the systematic development of these topics through standard courses in physics and mathematics. This book should be regarded as a supplement to the standard courses, but one which I hope may enhance your understanding, appreciation and enjoyment of physics.

This book aims to redress the balance between pedagogical and real physics through seven case studies which span much of classical physics and the beginnings of quantum physics.[1] The subjects of the case studies are as follows:

I The origins of Newton's laws of motion and of gravity.
II Maxwell's equations for the electromagnetic field.
III Mechanics and dynamics – linear and non-linear.
IV Thermodynamics and statistical physics.
V The origins of the concept of quanta.
VI Special and general relativity.
VII Cosmology.

These topics have a familiar ring, but they are treated from a rather different perspective as compared with the standard textbooks – hence the subtitle of the book *'An Alternative View of Theoretical Reasoning in Physics*. It is not just the content of the physics which I am aiming to explore, but also the leaps of imagination involved in some of the greatest discoveries in physics and theoretical physics.

At the same time, we can gain important insights into how real physics and theoretical physics are carried out. These convey some of the excitement and intense intellectual struggle involved in achieving new levels of understanding. In a number of the case studies, we will retrace the processes of discovery which were followed by the scientists themselves, using only mathematical techniques and concepts available at the time. This has an added benefit in that many of the problems which students have in understanding physics are the same as those which challenged many of the greatest physicists.

1.2 Reflections on What Every Student Should Know

Let me list some of the lessons which I hope readers will take away from the book, based on my experience of practising and teaching physics for many years.

(i) It is only too easy to lose a *global view* of the physics and theoretical physics. Professionals use the whole of physics in tackling problems and there is no artificial distinction between thermal physics, optics, mechanics, electromagnetism, quantum mechanics, and so on. A corollary of this is that in physics any problem can normally be tackled and solved in a variety of different ways. *There is often no single 'best way' of solving a problem*. Much deeper insights into how physics works are obtained if a problem is approached from very different perspectives, from thermodynamics, from electromagnetism, from quantum mechanics, and so on.

(ii) How problems are tackled and how one thinks about physics are highly personal matters. No two physicists think in exactly the same way although, when they come to work out a problem, they should come up with the same answer. The *individual physicist's response to the discipline* is an integral part of the way physics is practised to a much greater extent than students or the lecturers themselves would often like to believe. But it is the diversity of approaches to physics which provides insight into the nature of the mental processes by which physicists understand their subject. I remember vividly a splendid lecture by my colleague Douglas Gough, summarising a colloquium in Vienna entitled *Inside the Stars* in which he concluded with the following wonderful paragraph.

> I believe that one should never approach a new scientific problem with an unbiased mind. Without prior knowledge of the answer, how is one to know whether one has obtained the right result? But with prior knowledge, on the other hand, one can usually correct one's observations or one's theory until the outcome is correct. ... However, there are rare occasions on which, no matter how hard one tries, one cannot arrive at the correct result. Once one has exhausted all possibilities for error, one is finally forced to abandon a prejudice, and redefine what one means by 'correct'. So painful is the experience that one does not forget it. That subsequent replacing of the old prejudice by a new one is what constitutes a gain in real knowledge. And that is what we, as scientists, continually pursue.[2]

Douglas's dictum is the foundation of the process of discovery in research. All of us have different prejudices and personal opinions about what the solutions to problems might be and it is this diversity of approach which leads to new understandings.

(iii) It is often difficult to convey the *sheer excitement of the processes of research and discovery in physics*. Most of us spend many more hours pursuing our research than would be expected in any normal 'job'. The caricature of the 'mad' scientist is not wholly a myth in that, in carrying out frontier research, it is almost essential to become totally absorbed in the problems to the exclusion of the cares of normal life. The biographies of many of the greatest scientists illustrate their extraordinary powers of concentration. The examples of Newton and Faraday spring immediately to mind as physicists who, once embarked upon a fertile seam of research, would work unrelentingly until the inspiration was exhausted. All of us have experience of this total intellectual commitment at much more modest levels of achievement and it is only later that, on reflection, we regard these as among our best research experiences.

(iv) Key factors in these historical examples, which are familiar to all professional physicists, are the central roles of *hard work*, *experience* and, perhaps most important of all, *intuition*. Many of the most successful physicists depend very heavily upon their wide experience of a great deal of hard work in physics and theoretical physics. It would be marvellous if this experience could be taught, but I am convinced in fact it is something which can only be gained by dedicated hard work and perseverance. We all remember our mistakes and the blind alleys we have entered, and these teach us just as much about physics as our successes. I regard intuition as a distillation of all our experience as physicists. It is potentially a dangerous tool because one can make some very bad blunders by relying too much on it in frontier areas of physics. Yet it is certainly the source of many of the greatest discoveries in physics.

(v) Perhaps most important of all is the essential element of *creativity* in coming to new understandings of the laws of nature. In my view, this is not so different from creativity in the arts. The leaps of the imagination involved in discovering, say, Newton's laws of motion, Maxwell's equations, relativity and quantum mechanics are not so different in essence from the creations of the greatest artists, musicians, writers and so on. The basic difference is that physicists must be creative within a very strict set of rules and that their theories should be testable by confrontation with experiment and observation. In my view, the imagination and creativity involved in the very best experimental and theoretical physics result in a real sense of *beauty*. The great achievements of physics evoke in me, at least, the same type of response that one finds with great works of art. I think it is important to let students know when I find a piece of physics particularly beautiful – and there are many examples of these in this book. When I teach these topics, I experience the same process of rediscovery as on listening to a familiar piece of classical music – one's umpteenth rehearing of the *Eroica* symphony or of the *Sacre du Printemps*.

1.3 The Nature of Physics and Theoretical Physics

The natural sciences aim to provide a logical and systematic account of natural phenomena and to enable us to predict from our past experience to new circumstances. *Theory* is the formal basis for such endeavours. Theory need not be mathematical, but mathematics is the most powerful and general method of reasoning we possess. Therefore, we attempt to secure *data* wherever possible in a form that can be analysed *mathematically*. There are two immediate consequences for theory in physics.

The basis of all physics and theoretical physics is *experimental data* and the necessity that these data be in *quantified form*. Some would like to believe that the whole of theoretical physics could be produced by pure reason, but they are doomed to failure from the outset. As Volker Heine has remarked,

> No one starting with Schrödinger's equation would have predicted superconductivity.[3]

The great achievements of theoretical physics have been solidly based upon the achievements of experimental physics, which provides powerful constraints upon physical theory. Theoretical physicists should therefore have a good and sympathetic understanding of the methods of experimental physics, not only so that theory can be confronted with experiment in a meaningful way, but also so that new experiments can be proposed which are realisable and which can discriminate between rival theories.

A second consequence is, as stated earlier, that we must have adequate *mathematical tools* with which to tackle the problems we need to solve. Historically, the mathematics and the experiments have not always been in step. Sometimes the mathematics is available but the experimental methods needed to test the theory are unavailable. In other cases, the opposite has been true – new mathematical tools have to be developed to describe the results of experiment.

Mathematics is central to reasoning in physics but we should beware of treating it, in any sense, as the whole physical content of theory. Let me reproduce some words from the reminiscences of Paul Dirac about his attitude to mathematics and theoretical physics. Dirac sought mathematical beauty in all his work. For example, he writes:

> Of all the physicists I met, I think Schrödinger was the one that I felt to be most closely similar to myself. . . . I believe the reason for this is that Schrödinger and I both had a very strong appreciation of mathematical beauty and this dominated all our work. It was a sort of act of faith with us that any equations which describe fundamental laws of Nature must have great mathematical beauty in them. It was a very profitable religion to hold and can be considered as the basis of much of our success.[4]

Earlier, however, he had written the following:

> I completed my (undergraduate) course in engineering and I would like to try to explain the effect of this engineering training on me. Previously, I was interested only in exact equations. It seemed to me that if one worked with approximations there was an intolerable ugliness in one's work and I very much wanted to preserve mathematical beauty. Well, the engineering training which I received did teach me to tolerate approximations and I was able to see that even theories based upon approximations could have a considerable amount of beauty in them.
>
> There was this whole change of outlook and also another which was perhaps brought on by the theory of relativity. I had started off believing that there were some exact laws of Nature and that all we had to do was to work out the consequences of these exact laws. Typical of these were Newton's laws of motion. Now we learned that Newton's laws of motion were not exact, only approximations and I began to infer that maybe all the laws of nature were only approximations . . .
>
> I think that if I had not had this engineering training, I should not have had any success with the kind of work I did later on because it was really necessary to get away from the point of view that one should only deal with exact equations and that one should deal only with results which could be deduced logically from known exact laws which one accepted, in which one had implicit faith. Engineers were concerned only in getting equations which were useful for describing nature. They did not very much mind how the equations were obtained. . . .
>
> And that led me of course to the view that this outlook was really the best outlook to have. We wanted a description of nature. We wanted the equations which would describe nature and the best we could hope for was usually approximate equations and we would have to reconcile ourselves to an absence of strict logic.[5]

These are important and profound sentiments which should be familiar to the reader. There is really no strictly logical way in which we can formulate theory – we are continually approximating and using experiment to keep us on the right track. Note that Dirac was describing theoretical physics at its very highest level – concepts such as Newton's laws of motion, special and general relativity, Schrödinger's equation and the Dirac equation are the *very summits of achievement of theoretical physics* and very few of us can work creatively at that level. The same sentiments apply, however, in their various ways to all aspects of research as soon as we attempt to model quantitatively the natural world.

Most of us are concerned with applying and testing known laws to physical situations in which their application has not previously been possible, or foreseen, and we often have to make numerous approximations to make the problems tractable. The essence of our training as physicists is to develop confidence in our physical understanding of physics so that, when we are faced with a completely new problem, we can use our experience and intuition to recognise the most fruitful ways forward.

1.4 Environmental Influences

It is important to realise not only that all physicists are individuals with their own approaches and prejudices, but also that these are influenced by the tradition within which they have studied physics. I have had experience of working in a number of different countries, particularly in the USA and the former Soviet Union, and the different scientific traditions can be appreciated vividly by the approach of the physicists to research problems. These experiences have added very significantly to my understanding and appreciation of physics.

An example of a distinctively British feature of physics is the tradition of *model building*, to which we will return on numerous occasions. Model building was an especially British trait during the nineteenth and early twentieth centuries. The works of Faraday and Maxwell are full of models, as we will see, and at the beginning of the twentieth century, the variety of models for atoms was quite bewildering. J.J. Thomson's 'plum-pudding' model of the atom is perhaps one of the more vivid examples, but it is just the tip of the iceberg. Thomson was quite straightforward about the importance of model building:

> The question as to which particular method of illustration the student should adopt is for many purposes of secondary importance provided that he does adopt one.[6]

Thomson's assertion is splendidly illustrated by Heilbron's *Lectures on the History of Atomic Physics 1900-1920*.[7] Heilbron gives this splendid example of Fitzgerald's approach to the structure of the aether:

> [FitzGerald had] in mind mechanical models, that is, detailed representations of physical phenomena, especially light and electromagnetism, in terms of the motions and interactions of hypothetical particles or media . . . The representations were not meant to be taken literally. To quote Fitzgerald:
>
> To suppose that (the electromagnetic) aether is at all like the model I am about to describe (which is made from tennis balls and rubber bands) would be almost as bad a mistake as to suppose a sphere at all like $x^2 + y^2 + z^2 = r^2$ and to think that it must, in consequence, be made of paper and ink.

This approach is very different from the continental European tradition of theoretical physics – we find Poincaré remarking 'The first time a French reader opens Maxwell's book, a feeling of discomfort, and often even of distrust, is at first mingled with his admiration . . . '.[8] According to Hertz, Kirchhoff was heard to remark that he found it painful to see atoms and their vibrations wilfully stuck in the middle of a theoretical discussion.[9] It was reported to me after a lecture in Paris that one of the senior professors had commented that my presentation had not been 'sufficiently Cartesian'. I believe the British tradition of model building is alive and well. I can certainly vouch for the fact that, when I think about some topic in physics or astrophysics, I generally have some picture, or model, in my mind rather than an abstract or mathematical idea.

In my view, the development of *physical insight* is an integral part of the model-building tradition. The ability to guess correctly what will happen in a new physical situation without having to write down all the mathematics is a very useful talent and most of us develop it with time. It must be emphasised, however, that having physical insight is no substitute for producing exact mathematical answers. If you want to claim to be a theoretical physicist, you must be able to give the rigorous mathematical solution as well.

The influence of our environment applies equally to different physics departments, as well as to different countries. If we consider the term *theoretical physics*, there is a wide range of opinion as to what constitutes ‘theoretical physics’ as opposed to ‘physics’. It is a fact that in the Cavendish Laboratory in Cambridge, most of the lecture courses are strongly theoretically biassed. By this I mean that these courses aim to provide students with a solid foundation in basic theory and its development and relatively less attention is paid to matters of experimental technique. If experiments are alluded to, the emphasis is generally upon the results, rather than the experimental ingenuity by which the experimental physicists came to their answers. We expect students to acquire most of their experimental training through practical experiments. The situation is in strong contrast to the nature of the Cambridge physics courses in the early decades of the twentieth century, which were very strongly experimental in emphasis.[10]

In contrast, members of Departments of Theoretical Physics or Applied Mathematics would claim rightly that they teach much more ‘theoretical’ theoretical physics than we do. In their undergraduate teaching, I believe this is the case. There is by definition a strong mathematical bias in the teaching of these departments. They are often much more concerned about rigour in their use of mathematics than we are. In other physics departments, the bias is often towards experiment rather than theory. I find it amusing that some members of the Laboratory who are considered to be ‘experimentalists’ within the department are regarded as ‘theorists’ by other physics departments in the UK.

The reason for discussing this issue of the local environment is that it can produce a somewhat biased view of what we mean by physics and theoretical physics. My own perspective is that physics and theoretical physics are part of a continuum of approaches to physical understanding – they are different ways of looking at the same body of material. In my view, there are great advantages in developing mathematical models in the context of the experiments, or at least in an environment where day-to-day contact occurs naturally with those involved in the experiments.

1.5 Final Disclaimer

Let me emphasise at the outset that I am *not* a historian or philosopher of science. I use the history of science very much for my own purposes, which are to illuminate my own experience of how real physicists think and behave. The use of historical case studies is simply a device for conveying something of the reality and excitement of physics. I therefore apologise unreservedly to real historians and philosophers of science for using the fruits of their labours in attaining my goals. My hope is that students will gain enhanced appreciation and respect for the works of professional historians and philosophers of science from what they read in this book.

Establishing the history by which scientific discoveries were made is a hazardous and difficult business and, even in the recent past, it is often difficult to disentangle what really happened. In my background reading, I have relied heavily upon standard biographies and histories. For me, they provide vivid pictures of how science is actually carried out and I can relate them to my own research experience.

My ambition is that all advanced undergraduates in physics should be able to profit from this book, whether or not they are planning to become professional theoretical physicists. Although physics can be carried out without a deep understanding of theory, that point of view misses so much of the beauty and stimulation of the subject. Remember, however, the case of Stark, who made it a point of principle to reject almost all theories on which his colleagues had reached a consensus. Contrary to their view, he showed that spectral lines could be split by an electric field, the Stark effect, for which he won the Nobel prize.

I particularly want to convey a real appreciation of the great discoveries of physics and theoretical physics. These achievements are as great as any in the field of human endeavour.

Notes

1 Those who want to go further and find out how the concept of quanta led to the theory of quantum mechanics may enjoy my 'sequel' to this book, *Quantum Concepts in Physics: An Alternative Approach to the Understanding of Quantum Mechanics*, Cambridge: Cambridge University Press (2013). A similar approach to the present book is taken, but now analysing the turbulent years from 1900 to 1930.

2 Gough, D.O. (1993). In *Inside the Stars*, eds. W.W. Weiss and A. Baglin, IAU Colloquium No. 137, p. 775. San Francisco: Astronomical Society of the Pacific Conference Series (Vol. 40).

3 Personal communication.

4 Dirac, P.A.M.(1977). *History of Twentieth Century Physics*, Proceedings of the International School of Physics 'Enrico Fermi', Course 57, p. 136. New York and London: Academic Press.

5 Dirac, P.A.M. (1977). *op. cit.*, p. 112.

6 Thomson, J.J. (1893). *Notes on Recent Researches in Electricity and Magnetism*, vi. Oxford: Clarendon Press. (Quoted by J.L. Heilbron in note 7, p. 42.)

7 Heilbron, J.L. (1977). *History of Twentieth Century Physics*, Proceedings International School of Physics 'Enrico Fermi', Course 57, p. 40. New York and London: Academic Press.

8 Duhem, P. (1991 reprint). *The Aim and Structure of Physical Theory*, p. 85. Princeton: Princeton University Press.

9 Heilbron, J.L. (1977). *op. cit.*, p. 43.

10 For many more details, see my book *Maxwell's Enduring Legacy: A Scientific History of the Cavendish Laboratory*, Cambridge: Cambridge University Press (2016).

Case Study I

THE ORIGINS OF NEWTON'S LAWS OF MOTION AND OF GRAVITY

Our first case study encompasses essentially the whole of what can be considered the modern scientific process. Unlike the other case studies, it requires little mathematics, but a great deal in terms of intellectual imagination. It is a heroic tale of scientists of the highest genius laying the foundations of modern science. Everything is there – the role of brilliant experimental skill, of imagination in the interpretation of observational and experimental data, and the use of advanced geometrical techniques to establish the laws of planetary motion – these were to lay the foundations for the Newtonian picture of the world. As expressed by Herbert Butterfield in his *Origins of Modern Science*, the understanding of motion was one of the most difficult steps that scientists have ever undertaken.[1] In the quotation by Douglas Gough in Chapter 1, he expresses eloquently the pain experienced on being forced to discard a cherished prejudice in the sciences. How much more painful must have been the process of laying the foundations of modern science, when the concept that the laws of nature could be written in mathematical form had not yet been formulated.

How did our modern appreciation of the nature of our physical Universe come about? In Chapter 2, we set the scene for the subsequent triumphs, and tragedies, of two of the greatest minds of modern science – Galileo Galilei and Isaac Newton. Their achievements were firmly grounded in the remarkable observational advances of Tycho Brahe and Galileo's skill as an experimental physicist and astronomer. Galileo and his trial by the Inquisition are considered in Chapter 3, the emphasis being upon the scientific aspects of this controversial episode in the history of science. The scientific issues involved can be considered the touchstone for the modern view of the nature of scientific enquiry. Then, with the deep insights of Kepler and Galileo established, Newton's great achievements are placed in their historical context in Chapter 4.

It may seem somewhat strange to devote so much space at the beginning of this text to what many will consider to be ancient history, a great deal of which we now understand to be wrong or misleading. I feel very differently about it, however. It is a gripping story and full of resonances about the way we practice science today. There are other aspects

[1] Butterfield, H. *The Origins of Modern Science: 1300–1800*. London: G. Bell and Sons Ltd. (1950).

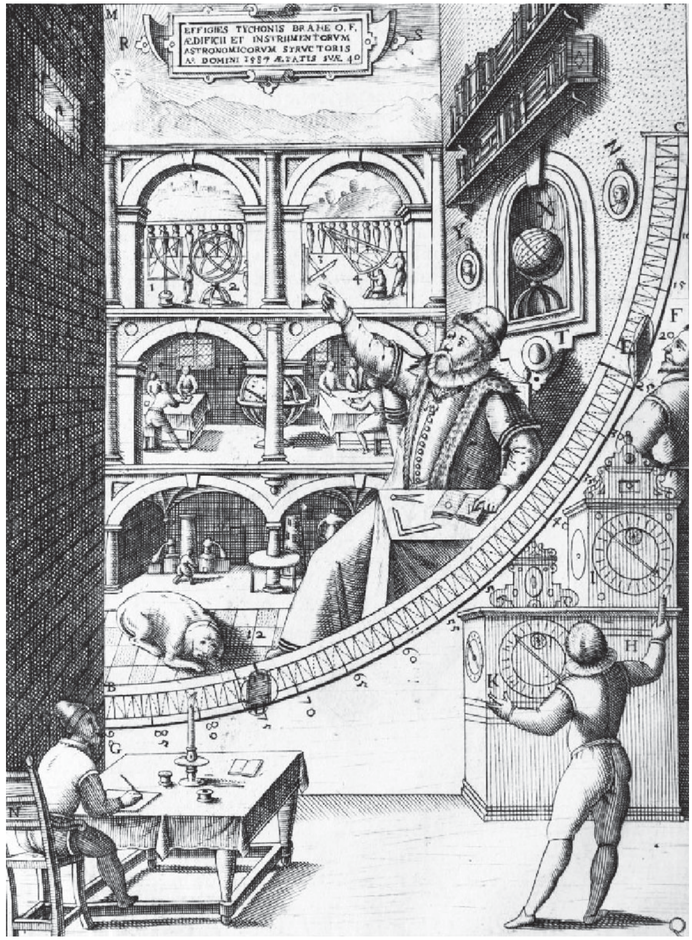


Fig. I.1 Tycho Brahe with the instruments he had constructed for accurate observations of the positions of the stars and planets. He is seated within the 'great mural transit', which produced some of the most accurate measurements of the positions of the stars and planets at that time. (From *Astronomiae Instauratae Mechanica*, 1602, p. 20, Nurnberg.)

of this story which I believe are important. The great scientists involved in this case study had complex personalities and, to provide a rounder picture of their characters and achievements, it is helpful to understand their intellectual perspectives as well as their contributions to fundamental physics.

2 From Ptolemy to Kepler: The Copernican Revolution

2.1 Ancient History

According to the historian of early physics and astronomy Olaf Pedersen, the most outstanding of all Greek astronomers was Hipparchus who was born in Nicaea in about 190 BC.[1] His monumental achievement was a catalogue of the positions and brightness of 850 stars in the northern sky, completed in 127 BC. A measure of his skill as an astronomer is that he compared his positions with those of Timocharis made in Alexandria 150 years earlier and discovered the *precession of the equinoxes*, the very slow change of the direction of the Earth's axis of rotation relative to the frame of reference of the fixed stars. We now know that this precession results from tidal torques of the Sun and Moon acting upon the slightly non-spherical Earth. At that time, however, the Earth was assumed to be stationary and so the precession of the equinoxes had to be attributed to the movement of the 'sphere of fixed stars'.

The most famous of the ancient astronomical texts is the *Almagest* of Claudius Ptolomeaus or Ptolemy, who lived in the second century AD. The word Almagest is a corruption of the Arabic translation of the title of his book, the *Megelé Syntaxis* or *The Great Composition*, which in Arabic became *al-majisti*. It consisted of 13 volumes and provided a synthesis of all the achievements of the Greek astronomers and, in particular, leant heavily upon the work of Hipparchus. In the Almagest, Ptolemy set out what became known as the *Ptolemaic System of World*, which was to dominate astronomical thinking for the next 1500 years.

How did the Ptolemaic system work? It is apparent to everyone that the Sun and Moon appear to move in roughly circular paths about the Earth. Their trajectories are traced out against the *sphere of the fixed stars*, which also appears to rotate about the Earth once per day. In addition, five planets are observable with the naked eye: Mercury, Venus, Mars, Jupiter and Saturn. The Greek astronomers understood that the planets did not move in simple circles about the Earth, but had somewhat more complex motions. Figure 2.1 shows Ptolemy's observations of the motion of Saturn in AD 137 against the background of the fixed stars. Rather than move in a smooth path across the sky, the path of the planet doubles back upon itself.

The challenge to the Greek astronomers was to work out mathematical schemes which could describe these motions. As early as the third century BC, a few astronomers, such as Heracleides of Pontus, suggested that these phenomena could be explained if the

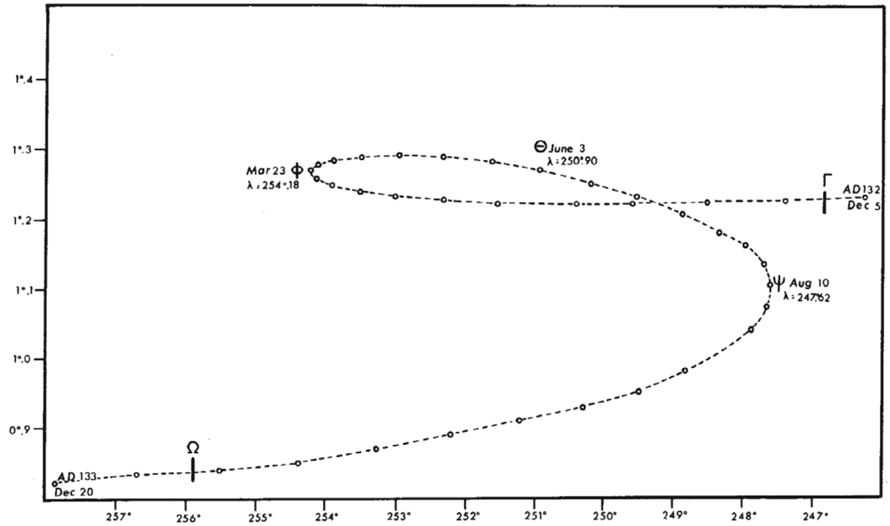


Fig. 2.1 The motion of Saturn in AD 133 as observed by Ptolemy against the background of the fixed stars. (From O. Pedersen, 1993. *Early Physics and Astronomy*, p. 61. Cambridge: Cambridge University Press.)

Earth rotates on its axis. Most remarkably, Aristarchus proposed that the planets move in circular orbits about the Sun. In his preface to *The Sand Reckoner*, Archimedes wrote to King Gelon,

> You are not unaware that by the universe most astronomers understand a sphere the centre of which is at the centre of the Earth. . . . However, Aristarchus of Samos has published certain writings on the (astronomical) hypotheses. The presuppositions found in these writings imply that the universe is much greater than we mentioned above. Actually, he begins with the hypothesis that the fixed stars and the Sun remain without motion. As for the Earth, it moves around the Sun on the circumference of a circle with centre in the Sun.[2]

These ideas were to became the inspiration for Copernicus roughly eighteen centuries later but they were rejected at that time for a number of reasons. Probably the most serious was the opposition of the upholders of Greek religious beliefs. According to Pedersen and Pihl,

> . . . in the eyes of the ordinary Greek, Aristarchus had sinned against deep-rooted ideas about Hestia's fire, and the Earth as a Divine Being. Such religious tenets could not be shaken by abstract astronomical theories incomprehensible to the ordinary person.[3]

From our perspective, the *physical arguments* against the heliocentric hypothesis are of the greatest interest. First, the idea that the Earth rotates about an axis was rejected. If the Earth rotates, then, when an object is thrown up in the air, it should not come down again in the same spot – the Earth would have moved because of its rotation before the

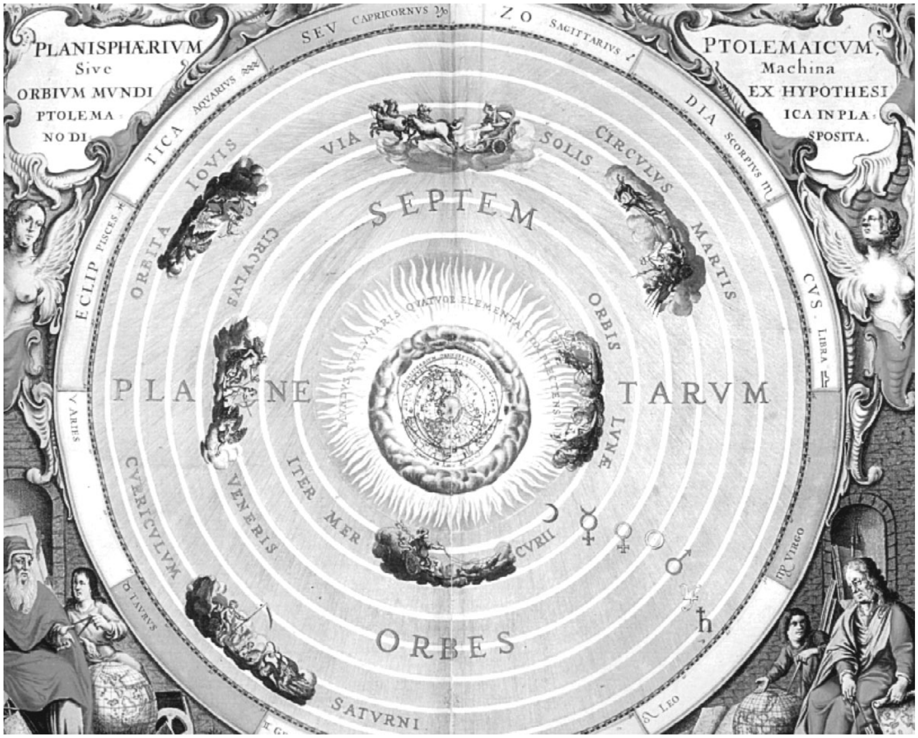


Fig. 2.2 The basic Ptolemaic system of the world showing the celestial bodies moving in circular orbits about the Earth in the order Moon, Mercury, Venus, Sun, Mars, Jupiter, Saturn and the sphere of fixed stars. (From Andreas Cellarius, 1661. *Harmonia Macrocosmica*, Amsterdam.)

object landed. No one had ever observed this to be the case and so the Earth could not be rotating. A second problem resulted from the observation that, if objects are not supported, they fall under gravity. Therefore, if the Sun were the centre of the Universe rather than the Earth, everything ought to be falling towards that centre. Now, if objects are dropped they fall towards the centre of the Earth and not towards the Sun. It follows that the Earth must be located at the centre of the Universe and so religious belief was supported by scientific rationale. Note that there was no quantitative theory of the law of gravity to provide validation or otherwise of this qualitative argument.

According to the Ptolemaic Geocentric System of the World, the Earth is stationary at the centre of the Universe and the principal orbits of the other celestial objects are circles in the order Moon, Mercury, Venus, Sun, Mars, Jupiter, Saturn and finally the sphere of the fixed stars (Fig. 2.2). This elementary Ptolemaic system could not, however, account for the details of the motions of the planets, such as the retrograde motion shown in Fig. 2.1, and so the model had to become more complex.

A central precept of Greek mathematics and philosophy was that the only allowable motions were uniform motion in a straight line and uniform circular motion. Ptolemy himself stated that uniform circular motion was the only kind of motion ‘in agreement with

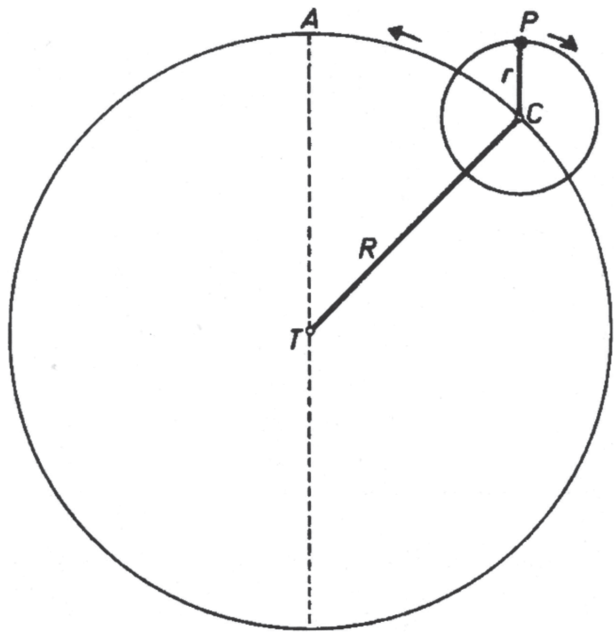


Fig. 2.3 Illustrating circular epicyclic motion about a circular orbit according to the epicyclic model of Appolonius. (From O. Pedersen, 1993. *Early Physics and Astronomy*, p. 71. Cambridge: Cambridge University Press.)

the nature of Divine Beings'. Therefore, it was supposed that, in addition to their circular orbits about the Earth, the planets, as well as the Sun and Moon, had circular motions about the principal circular orbit which were known as *epicycles* (Fig. 2.3). The type of orbit shown in Fig. 2.1 can be reproduced by selecting suitable speeds for the motions of the planets in their epicycles.

One of the basic features of *astrometry*, meaning the accurate measurement of the positions and movements of bodies on the celestial sphere, is that the accuracy with which their orbits are determined improves the longer the time span over which the observations are made. As a result, the simple epicyclic picture had to become more and more complex, the longer the time-base of the observations. To improve the accuracy of the Ptolemaic model, the centre of the circle of the planet's principal orbit could differ from the position of the Earth, each circular component of the motion remaining uniform. It was then found necessary to assume that the centre of the circle about which the epicycles took place also differed from the position of the Earth (Fig. 2.4). An extensive vocabulary was used to describe the details of the orbits.[4] The important point is that, by considerable geometrical ingenuity, Ptolemy and later generations of astronomers were able to give a good account of the observed motions of the Sun, the Moon and the planets. Although complicated, the Ptolemaic model was used in the preparation of almanacs and in the determination of the dates of religious festivals until after the Copernican revolution.

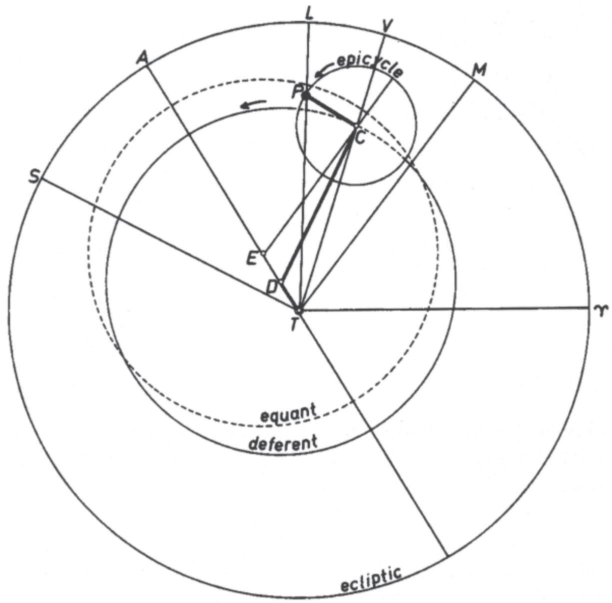


Fig. 2.4 Illustrating the complexity of the Ptolemaic model for the motion of the outer planets described in Books IX to XI of Ptolemy's *Almagest*. (From O. Pedersen, 1993. *Early Physics and Astronomy*, p. 82. Cambridge: Cambridge University Press.)

2.2 The Copernican Revolution

By the sixteenth century, the Ptolemaic system was becoming more and more complicated as a tool for predicting the positions of celestial bodies. *Nicolaus Copernicus* revived the idea of Aristarchus that a simpler model, in which the Earth orbits the Sun which is located at the centre of the Universe, might provide a simpler description of the motions of the planets. Copernicus, born in Torun in Poland in 1473, first attended the University of Kraków and then went to Bologna where his studies included astronomy, Greek, mathematics and the writings of Plato. In the early 1500s, he spent four years at Padua where he also studied medicine. By the time he returned to Poland, he had mastered all the astronomical and mathematical sciences. Copernicus made some observations himself and these were published between 1497 to 1529.

His greatest works were undoubtedly his investigations of the heliocentric Universe. When he worked out the geometrical details of this model, he found that it worked remarkably well. He restricted the motions of the Moon and the planets to uniform circular orbits, in accordance with the precepts of Aristotelian physics. In 1514 he circulated his ideas privately in a short manuscript called *De hypothesibus motuum coelestium*

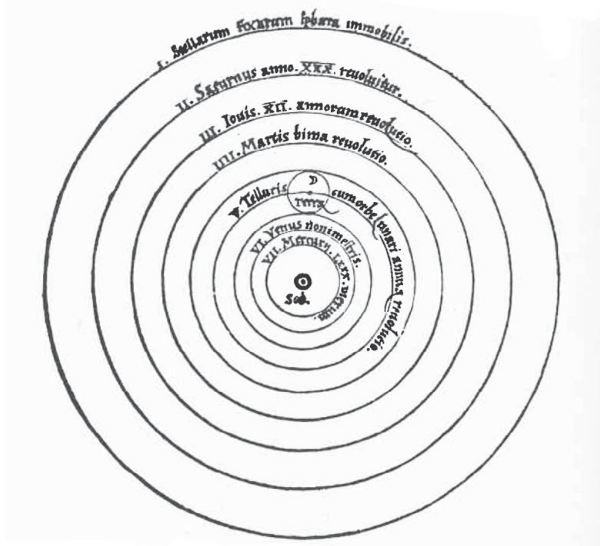


Fig. 2.5 The Copernican Universe from Copernicus's treatise *De Revolutionibus Orbium Celestium*, 1543. Opposite page 10, Nurnberg. (From the Crawford Collection, Royal Observatory, Edinburgh.)

a se constitutis commentariolus (A Commentary on the Theory of the Motion of the Heavenly Objects from their Arrangements).

In 1533, these ideas were presented to Pope Clement VII who approved of them and made a formal request in 1536 that the work be published. Copernicus hesitated, but eventually wrote his great treatise summarising what is now known as the *Copernican model of the Universe* in *De Revolutionibus Orbium Coelestium* (On the Revolutions of the Heavenly Spheres).[5] Eventually the work was published by Osiander in 1543. It is said that the first copy was brought to Copernicus on his death-bed on 24 May 1543. Osiander inserted his own foreword to the treatise stating that the Copernican model was no more than a calculating device for simplifying the predictions of planetary motions, but it is clear from the text that Copernicus was in no doubt that the Sun really is the centre of the Universe, and not the Earth.

Figure 2.5 shows the famous diagram which appears opposite page 10 of Copernicus's treatise, showing the planets in their familiar order with the Moon orbiting the Earth and the six planets orbiting the Sun. Beyond these lies the sphere of the fixed stars. The implications of the Copernican picture were profound, not only for science but also for our understanding of our place in the Universe.

The scientific implications were twofold. First, the size of the Universe was vastly increased as compared with the Ptolemaic model. If the fixed stars were relatively nearby, they ought to display parallaxes, meaning they should display apparent motions relative to more distant background stars because of the Earth's motion about the Sun. No such stellar parallax had ever been observed, and so the fixed stars must be very distant indeed.

The second fundamental insight was that something was wrong with the Aristotelian precept that all objects fall towards the centre of the Universe, now occupied by the Sun.

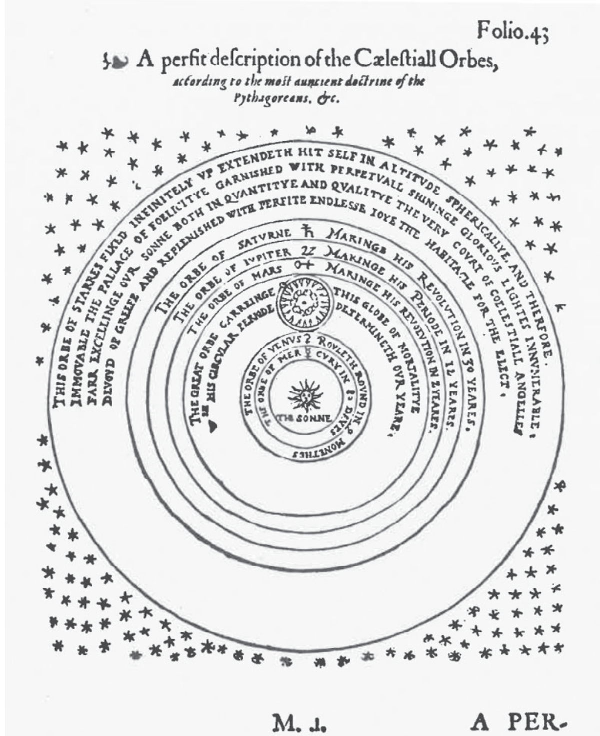


Fig. 2.6 Thomas Digges' version of the Copernican picture of the world showing the Solar System embedded in an infinite distribution of stars. (From Thomas Digges, 1576. *A Perfit Description of the Caelestiall Orbes*, London.)

The resolution of the problems of Aristotelian physics had to await the understanding of Galilean relativity and the quantification of the law of gravity by Isaac Newton.

In England, these ideas were enthusiastically adopted by the most important astronomer of the reign of Queen Elizabeth I, Thomas Digges, who translated large sections of *De Revolutionibus* into English. In his version of the Copernican model, the Universe is of infinite extent, the stars being scattered throughout space (Fig. 2.6). This remarkably prescient picture was to be adopted by Newton, but it leads to some tricky cosmological problems.[6]

2.3 Tycho Brahe: The Lord of Uraniborg

Copies of Copernicus's *De Revolutionibus Orbium Coelestium* circulated remarkably quickly throughout Europe.[7] One of the motivations behind Copernicus's researches was to create a simpler set of mathematical procedures for working out the motions of the Sun, the Moon and the planets in order to determine the exact date of the vernal equinox. This was needed to establish the correct dates of religious festivals and was one of the reasons for the favourable reception of Copernicus's heliocentric model by Pope Clement VII.

Up till the time of Copernicus, the predictions of the motions of the celestial bodies were taken from the *Alphonsine Tables*, which were derived from the Ptolemaic system, as refined by the Arabic astronomers. These tables had been prepared by the Rabbi Isaac ben Sid of Toledo and published in manuscript form in the *Libros del sabre de astronomica* in 1277 under the patronage of Alfonso X of Castile, also known as Alfonso the Wise. The tables were copied in manuscript form and quickly disseminated throughout Europe. They were only published in the modern sense in 1483. Modern scholarship has suggested that, in fact, predictions using the Copernican model were often not much better than those taken from the Alphonsine Tables. Predictions using the data in *De Revolutionibus* were made by Erasmus Reinhold who produced what became known as the *Prutenic Tables*, or Prussian tables, of the positions of the stars and planets. These were published in 1551, no more than eight years after the publication of *De Revolutionibus*.

The next hero of the story is *Tycho Brahe*, who was born into a noble family in 1546 at the family home of Knudstrup at Skåne in Denmark. He formed an early passion for astronomy but, to prepare him for the life of a nobleman, he was sent to the University of Leipzig in March 1562 to study law. He kept up his interest in astronomy, however, making observations secretly at night while his tutor slept. He also spent all the money he could save on astronomical books, tables and instruments. At this time he acquired his own copies of the Alphonsine and Prutenic Tables. The inspiration for his future work came from the predicted conjunction of Saturn and Jupiter in 1563. He found that the predictions of both sets of tables were in error, by about a month in the case of the Alphonsine Tables and by a few days using the Prutenic Tables. The need to improve the accuracy with which the planetary motions were known was one of the prime motivations for his monumental series of observations which began in the late 1570s.

Once established back in Denmark, he gave lectures at the University of Copenhagen, where he discussed the Copernican theory.

> Tycho spoke of the skill of Copernicus, whose system, although not in accord with physical principles, was mathematically admirable and did not make the absurd assumptions of the ancients.[8]

Tycho continued his international collaboration with Landgrave William IV whom he visited in Kassel in 1575. During that visit, he first became aware of the importance of the *refraction* of light rays in the Earth's atmosphere upon astronomical observations. Tycho was the first astronomer to make this correction in working out accurate stellar and planetary positions.

Tycho was determined to carry out a programme of measurement of the positions of the stars and planets to the very highest accuracy achievable. Frederick II of Denmark was persuaded that Tycho was an outstanding scientist who would bring honour to Denmark and so in 1576, to prevent Tycho 'brain-draining' to Germany, made him an offer which he could not refuse. In the words of Frederick, he provided Tycho with

> . . . our land of Hven with all our and the Crown's tenants and servants who thereon live, with all the rent and duty which comes from that . . . as long as he lives and likes to continue and follow his *studia mathematica* . . .[9]

The island of Hven lies half-way between Denmark and Sweden, and Tycho considered this peaceful island retreat ideal for his astronomical ambitions. He was allowed to use the rents for the upkeep and running of the observatories he was to build. In addition, he received regular supplementary funds from Frederick II to enable him to build his great observatories and to acquire the instruments needed for making astronomical observations. In Tycho's own words, he believed that the total enterprise had cost Frederick 'more than a tun of gold'.[10] Victor Thoren estimates that Tycho received an annual income of about 1% of the Crown's revenue throughout his years at Hven.[11] This was an early example of *big science*.

The first part of the project was to build the main observatory, which he named *Uraniborg*, or the Heavenly Castle. Besides adequate space for all his astronomical instruments, he built alchemical laboratories on the ground floor where he carried out chemical experiments. The building also included a papermill and printing press so that the results of the observations could be published promptly.

Uraniborg was completed by about 1580 and he then began the construction of a second observatory, *Stjerneborg*, which was located at ground level. He realised the value of constructing very solid foundations for the astronomical instruments which were clustered into a much more compact area. The glories of the observatories were the scientific instruments, beautifully described and illustrated in his *Astronomiae Instauratae Mechanica*, first published in 1598.[12] These instruments all predated the invention of the telescope and so all observations were made with the naked eye, consisting of measuring as accurately and systematically as possible the relative positions of the stars and planets. Two of the instruments are worthy of special mention. The first is the Great Globe, which can be seen on the ground floor of Uraniborg in Fig. I.1. On it, Tycho marked the positions of all the stars he recorded, thus forming a permanent record of their locations on the sky.

The second was the great mural quadrant being pointed out by Tycho himself in Fig. I.1. The mural quadrant is fixed in one position and had a radius $6\frac{3}{4}$ feet. Observations were made by observing stars as they transitted across the hole in the wall at the centre of the circle described by the quadrant, as well as making an accurate measurement of the time of the transit. Because of the large radius of the quadrant, very accurate positions could be measured. To keep an accurate track of time, Tycho used four clocks so that he could check if any one of them was not keeping time properly.

Tycho's technical achievements were remarkable. He was the first scientist to understand the crucial importance of taking account of *systematic errors* in his observations. We give two examples of such errors. The first, mentioned above, was taking account of atmospheric refraction. The second was the realisation that large instruments bend under gravity, producing a systematic error – if the instrument is pointed vertically, it does not bend whereas, if it is horizontal, the ends bend downwards and give incorrect angles.

Another important advance was the need to estimate how precise the observations were once the effects of systematic errors had been eliminated, in other words, the magnitude of the *random errors* in the observations. Tycho was probably the first scientist to work out the random errors in his observations. They turned out to be about ten times smaller than those of any earlier observations, an enormous increase in accuracy. Another key feature of the observations was that they were *systematic*, in the sense that they were made continuously over a period of about 20 years from 1576 to 1597 and analysed methodically by Tycho and his team of assistants. Throughout that period, he measured precisely the positions of the Sun, Moon, planets and fixed stars. His final catalogue contained the positions of 777 stars measured with an accuracy of about 1 to 2 minutes of arc. The crucial importance of knowing this figure will become apparent in the next section.

This was the summit of Tycho's achievement. He created the greatest observatory of its day, and Denmark and Hven became the centre of astronomy in Europe. After Frederick II's death in 1588, support for pure science waned under his successor, Christian IV. In addition, Tycho was mismanaging the island. Matters came to a head in 1597 when he left for exile, taking with him his observations, instruments and printing press. After two years in Wandsbek, now a borough of Hamburg, he eventually settled outside Prague under the patronage of Emperor Rudolf II at the castle of Benatek, where he set up again his magnificent set of instruments. During the remaining years of his life, a priority was the analysis of the mass of data he had secured and he had one stroke of great good fortune. One of his last acts in 1600 was to employ Johannes Kepler to reduce his observations of the planet Mars.

In about 1583, Tycho developed his own cosmology which he regarded as a synthesis of the Copernican and Ptolemaic views of the Universe (Fig. 2.7). The Earth is the centre

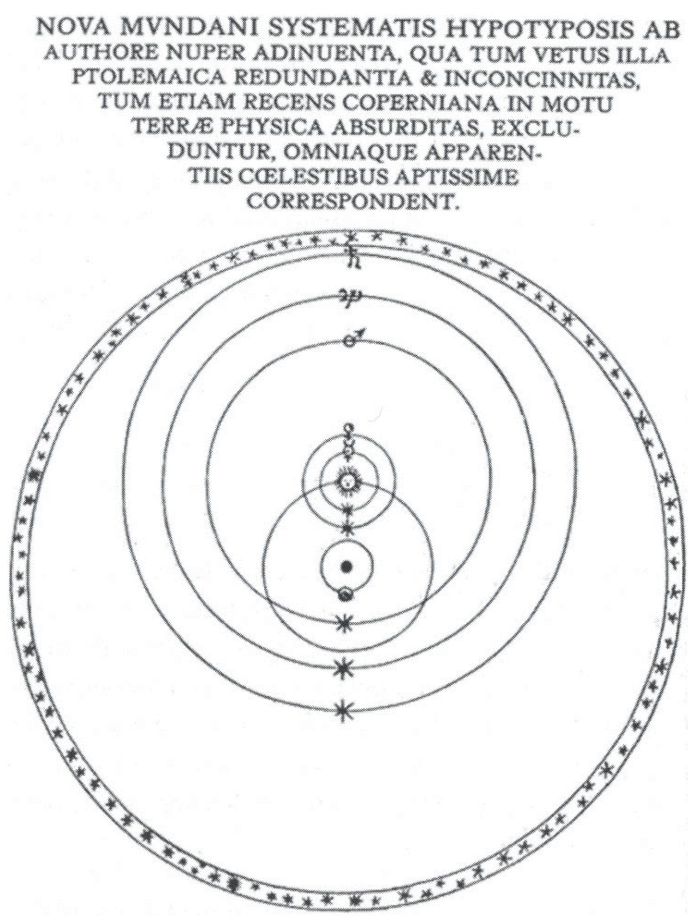


Fig. 2.7 The Tychonic System of the World. (From V.E. Thoren, 1990. *The Lord of Uraniborg*, p. 252. Cambridge: Cambridge University Press.)

of the Universe but the Moon and Sun orbit the Earth, while all the other planets orbit the Sun. Tycho was justly proud of his model, but his own observations were to lead to its demolition and the establishment of the law of gravity.

Tycho's achievements rank among the greatest in observational and experimental science. We recognise in his work all the very best hallmarks of modern experimental science, and yet the rules had not been written down in his day. The achievement is all the greater when we recall that the idea of quantitative scientific measurement scarcely existed at that time. Astronomical observations were by far the most exact of the physical sciences. Tycho's precise measurements were to play a central rôle in the Newtonian revolution.

2.4 Johannes Kepler and Heavenly Harmonies

We now introduce a very different character, *Johannes Kepler*, who was born in December 1571 in the Swabian City of Weil der Stadt. Kepler was a weak boy but a talented student, who entered the University of Tübingen in 1589. His first encounter with astronomy was through Michael Maestlin, who taught him mathematics, Euclid and trigonometry. In 1582, Maestlin published his treatise *Epitome Astronomiae*, which contained a description of Copernicus's heliocentric Universe. Maestlin was, however, a cautious individual and, under the influence of the church, remained a Ptolemaian. In contrast, Kepler had no doubts. In his words:

> Already in Tübingen, when I followed attentively the instruction of the famous magister Michael Maestlin, I perceived how clumsy in many respects is the hitherto customary notion of the structure of the universe. Hence, I was so very delighted by Copernicus, whom my teacher very often mentioned in his lectures, that I not only repeatedly advocated his views in disputations of the candidates, but also made a careful disputation about the thesis that the first motion (the revolution of the heaven of fixed stars) results from the rotation of the Earth.[13]

From the very beginning, Kepler was a passionate and convinced Copernican. His instruction at the university included both astronomy and astrology and he became an expert at casting horoscopes, which was to stand him in good stead financially.

Kepler's serious studies began in 1595 when he asked some basic questions about the Copernican picture of the Solar System. Why are there only six planets in orbit about the Sun? Why are their distances so arranged? Why do they move more slowly if they are further away from the Sun? Kepler's approach is summarised in two quotations from Max Casper. First,

> He was possessed and enchanted by the idea of harmony.[14]

Second,

> Nothing in this world was created by God without a plan; this was Kepler's principal axiom. His undertaking was no less than to discover this plan of creation, to think the thoughts of God over again . . .[15]

It was during his teaching of mathematics and geometry that the germ of an idea was planted. The moment of revelation is described in his book *Mysterium Cosmographicum* in his inimitable style.

> Behold, reader, the invention and whole substance of this little book! In memory of the event, I am writing down for you the sentence in the words from that moment of conception: the Earth's orbit is the measure of all things: circumscribe around it a dodecahedron and the circle containing it will be Mars; circumscribe around Mars a tetrahedron, and the circle containing this will be Jupiter; circumscribe about Jupiter a cube and the circle containing this will be Saturn. Now inscribe within the Earth an icosahedron and the circle contained in it will be Venus; inscribe within Venus an octahedron, and the circle contained in it will be Mercury. You now have the reason for the number of planets.[16]

How does this work? It is a well-known fact of solid geometry that there are only five regular solids, the *five platonic solids*, in which all the faces are made up of lines of equal length and in which the faces are all identical regular figures. Remarkably, by his choice of the order of the platonic solids, Kepler was able to account for the radii of the planetary orbits to an accuracy of about 5%. In 1596, he explained his model of the Solar System to the Duke of Württemburg and designs were made to construct Kepler's model (Fig. 2.8), but it was never built. Copernicus had given no special physical significance to the Sun as the centre of the Solar System, but Kepler argued that the Sun was the origin of the forces which held the planets in their orbits.

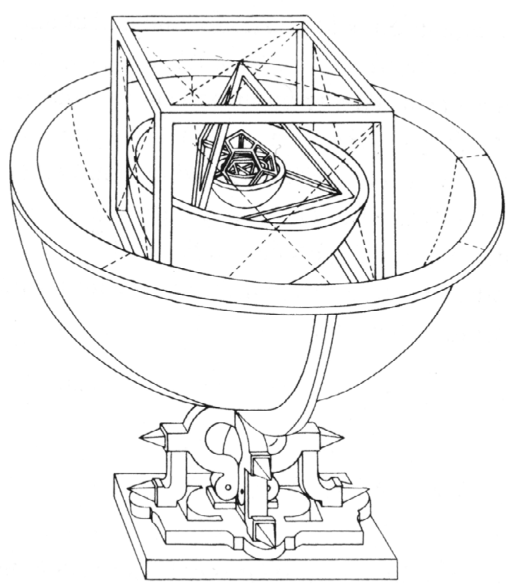


Fig. 2.8 Kepler's model of nested polyhedra which he developed to account for the number of planets and their radial distances from the Sun. (From *Dictionary of Scientific Biography, Vol. VII*, 1973, p. 292, after Kepler's original drawing in his *Mysterium Cosmographicum*, 1597.)

Now almost all Kepler's speculations are wrong and irrelevant to what follows, but there are two reasons for describing them. The first is that Kepler was seeking *physical* causes for the phenomena discovered by Copernicus. No-one had attempted to make this leap of the imagination before. The second is that this model is the first example of Kepler's fascination with harmonic and geometric principles.

With unbounded confidence and enthusiasm, Kepler published these ideas in 1596 in his *Mysterium Cosmographicum* (The Mystery of the Universe). He sent copies to many distinguished scientists of the day, including Tycho Brahe and Galileo. Galileo simply acknowledged receipt of the book, while Tycho Brahe was cautiously positive and encouraging in his response. Tycho invited Kepler to work with him in Benatek.

The *Mysterium Cosmographicum* made a considerable impact upon astronomical thinking. Looking back much later, Kepler wrote:

> For nearly all astronomical books which I published since that time were related to some one of the main chapters in this little book, presenting themselves as its more detailed argument or perfection. . . . The success which my book has had in the following years loudly testifies that no one ever produced a first book more deserving of admiration, more auspicious and, as far as its subject is concerned, more worthy.[17]

Kepler realised that what he needed to test his theory was much more accurate data on the orbits of the planets, and the only person who had acquired these was Tycho Brahe. After various toings and froings, he ended up in Tycho's employment in 1600. Tycho and Kepler were very different personalities. When Kepler moved to Benatek, Tycho was 53 and Kepler 28. Tycho was the greatest astronomer of his time and of noble origin. Kepler was the greatest mathematician in Europe and of humble origin. Tycho wanted Kepler to work on the 'Tychonic' Theory of the Solar System, whereas Kepler was already an ardent Copernican.

Just before Tycho died in 1601, he set Kepler to work on the problem of the orbit of Mars. On his deathbed, Tycho urged Kepler to complete a new set of astronomical tables to replace the Prutenic Tables. These were to be known as the *Rudolphine Tables* in honour of the emperor Rudolf II who had provided Tycho with the castle of Benatek, as well as the enormous salary of 3000 Gulden. Within two days of Tycho's death, Kepler was appointed Imperial Mathematician and the greatest period of his life's work began.

To begin with, Kepler assumed that the orbit of Mars was circular as in the Copernican picture. His first innovation was that the motion of Mars could not be explained if the motion was referred to the centre of the Earth's orbit. Rather, the motion had to be referred to the true position of the Sun. Kepler carried out many calculations to try to fit the observed orbit of Mars to circular orbits, again following the precept that only circular motions should be used to describe the orbits of the planets. After a great deal of trial and error, the best-fit orbits still disagreed with the observations by an error of 8 minutes of arc. This is where the knowledge of the random errors in Tycho's observations were critical. As Kepler stated:

> Divine Providence granted to us such a diligent observer in Tycho Brahe that his observations convicted this Ptolemaic calculation of an error of 8 minutes of arc; it is only right that we should accept God's gift with a grateful mind . . . Because these 8 minutes of arc could not be ignored, they alone have led to a total reformation of astronomy.[18]

In more mundane terms, the random errors in Tycho's final determinations of the planetary orbits amounted to only about 1 to 2 minutes of arc, whereas the minimum discrepancy which Kepler could find was at least four times this observational error. Before Tycho's observations, the errors were about ten times greater, and so Kepler would have had no problem in fitting these earlier observations to models with circular orbits.

This discrepancy could not be ignored and so Kepler had to start again. His next attempts involved the use of ovoids, egg-shaped figures, in conjunction with a magnetic theory of the origin of the forces which hold the planets in their orbits. It proved to be very complicated and tedious to work out the orbits according to this theory, and so he adopted intuitively an alternative approach in which the motions of the planets were such that they swept out equal areas in equal times. Whatever the actual shape of the orbit, the result is that the planet must move faster when it is closer to the Sun so that the area swept out by the line from the Sun to the planet is the same in equal time intervals. This theory gave excellent agreement with the observed longitudes of the planets about the Sun, and also of the Earth's orbit about the Sun. This great discovery became *Kepler's second law of planetary motion*. Formally:

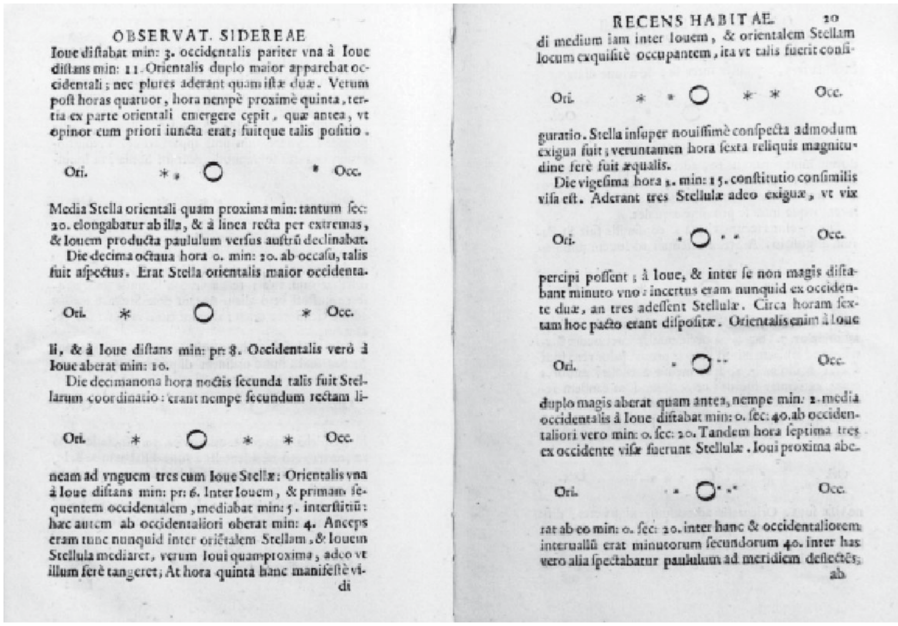
Equal areas are swept out by the line from the Sun to a planet in equal times.

Kepler proceeded with the mammoth task of fitting ovoids and the areal law to the motion of Mars, but he could not obtain exact agreement, the minimum discrepancy amounting to about 4 minutes of arc, still outside the limits of Tycho's observational errors. In parallel with these researches, he was writing his treatise *A Commentary on the Motion of Mars*, and had reached Chapter 51 before he realised that what he needed was a figure intermediate between the ovoid and the circle, the ellipse. He soon arrived at the key discovery that the orbit of Mars and indeed those of the other planets are ellipses with the Sun lying in one focus. The treatise on Mars was renamed *The New Astronomy* with the subtitle *Based on Causes, or Celestial Physics*. It was published in 1609, four years after the discovery of the law, which we now know as *Kepler's first law of planetary motion*:

The planetary orbits are ellipses with the Sun in one focus.

Notice the imaginative leap needed to place the Sun in one of the foci of the ellipses. Kepler already knew that the motion of Mars could not be referred to the centre of the Earth's orbit and so the focus is the next obvious place. Kepler had discovered a crucial fact about the orbits of the planets, but he had no physical explanation for it. It turned out to be a key discovery for the understanding of the law of gravity, but it had to await the genius of Isaac Newton before its full significance was appreciated.

The next development was due to Galileo. We have to anticipate his great discoveries made in 1609–10 with his astronomical telescope, which are described in Chapter 3. These were published in his famous book, the *Sidereus Nuncius* or the *Sidereal Messenger*, in 1610. Galileo was aware of the publication of Kepler's *Astronomia Nova* of the previous year and sent a copy of the *Sidereus Nuncius* to the Tuscan Ambassador at the Imperial Court in Prague, asking for a written opinion from Kepler. Kepler replied in a long letter on 19 April 1610, which he then published in May under the title *Dissertatio cum Nuncio Siderio* or *Conversation with the Sidereal Messenger*. As might be imagined, Kepler was wildly enthusiastic and, while Galileo presented his observations with some cautious attempts at interpretation, Kepler gave full rein to his imagination.


OBSERVAT. SIDEREAE

Ioue diſtabat min: 3. occidentalis pariter vna à Ioue diſtans min: 11. Orientalis duplo maior apparebat occidentali; nec plures aderant quam iſtæ duæ. Verum poſt horas quatuor, hora nempè proximè quinta, tertia ex parte orientali emergere cępit, quæ antea, vt opinor cum priori iuncta erat; fuitque talis poſitio.

Ori. Occ.

Media Stella orientali quam proxima min: tantum ſec: 20. elongabatur ab illa, & à linea recta per extremas, & Iouem producta paululum verſus auſtrũ declinabat.

Die decima octaua hora 0. min: 20. ab occaſu, talis fuit aſpectus. Erat Stella orientalis maior occidentali, & à Ioue diſtans min: pr: 8. Occidentalis verò à Ioue aberat min: 10.

Ori. Occ.

Die decimanona hora noctis ſecunda talis fuit Stellarum coordinatio: erant nempe ſecundum rectam lineam ad vnguem tres cum Ioue Stellæ: Orientalis vna à Ioue diſtans min: pr: 6. Inter Iouem, & primam ſequentem occidentalem, mediabat min: 5. interſtitiũ: hæc autem ab occidentaliori aberat min: 4. Anceps eram tunc nunquid inter orię̃talem Stellam, & Iouem Stellula mediaret, verum Ioui quamproxima, adeo vt illum ferè tangeret; At hora quinta hanc manifeſtè vidi

Ori. Occ.

RECENS HABITAE. 20

di medium iam inter Iouem, & orientalem Stellam locum exquiſitè occupantem, ita vt talis fuerit configuratio.

Ori. Occ.

Stella inſuper nouiſſimè conſpecta admodum exigua fuit; veruntamen hora ſexta reliquis magnitudine ferè fuit æqualis.

Die vigeſima hora 1. min: 15. conſtitutio conſimilis viſa eſt. Aderant tres Stellulæ adeo exiguæ, vt vix percipi poſſent; à Ioue, & inter ſe non magis diſtabant minuto vno: incertus eram nunquid ex occidente duæ, an tres adeſſent Stellulæ. Circa horam ſextam hoc pacto erant diſpoſitæ. Orientalis enim à Ioue duplo magis aberat quam antea, nempe min: 2. media occidentalis à Ioue diſtabat min: 0. ſec: 40. ab occidentaliori vero min: 0. ſec: 20. Tandem hora ſeptima tres ex occidente viſæ fuerunt Stellulæ. Ioui proxima aberat ab eo min: 0. ſec: 20. inter hanc & occidentaliorem interuallũ erat minutorum ſecundorum 40. inter has vero alia ſpectabatur paululum ad meridiem deflectẽs; ab

Ori. Occ.

Ori. Occ.

Ori. Occ.

Fig. 2.9 Two pages from Galileo's *Sidereus Nuncius*, showing his drawings of the movements of the four Galilean satellites. (Courtesy of the Royal Observatory, Edinburgh.)

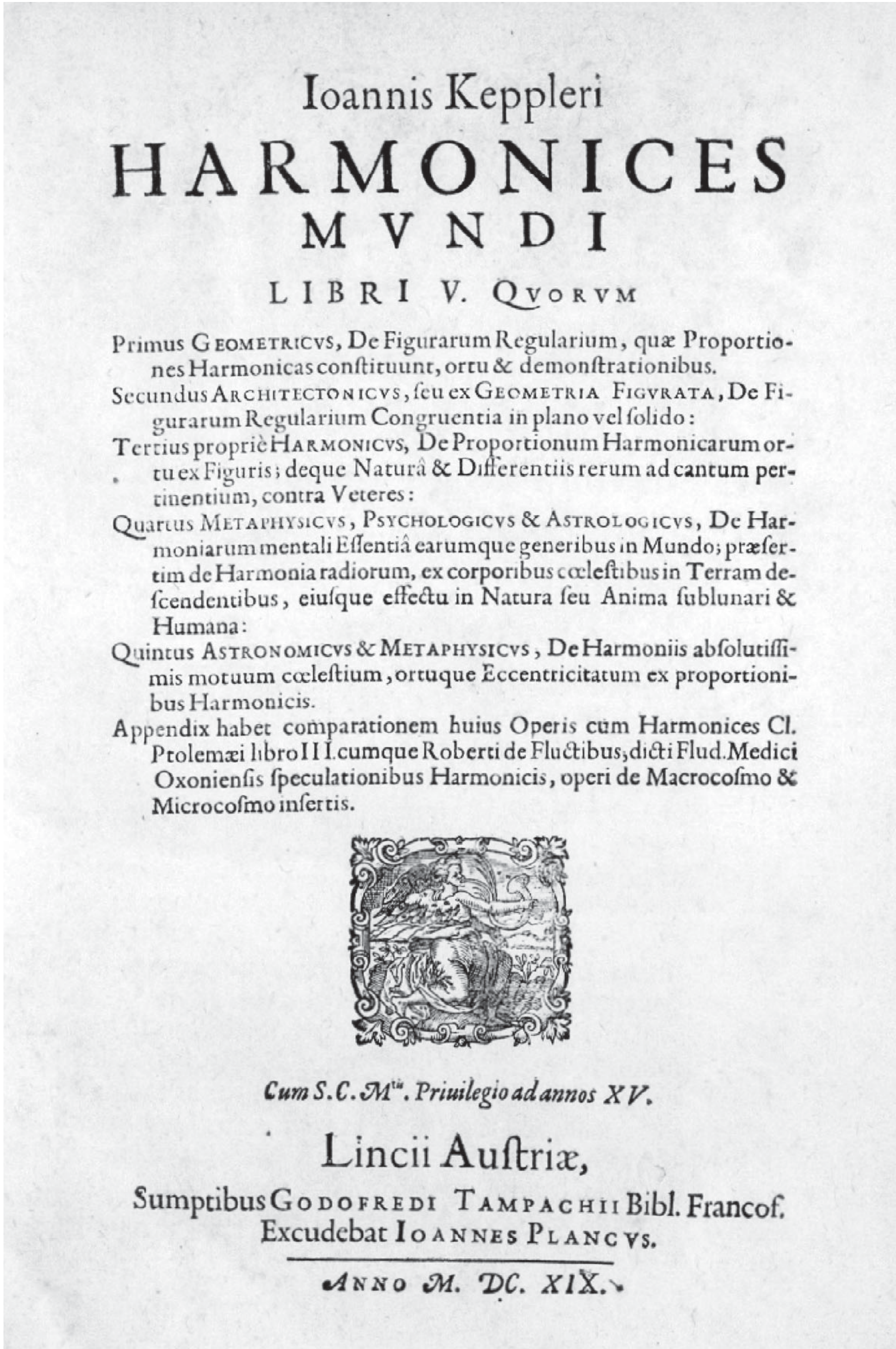
The most important discoveries for Kepler were the moons of Jupiter. Here was a miniSolar System with the four bright moons of Jupiter – Io, Europa, Ganymede and Callisto – orbiting the planet. In the *Sidereus Nuncius*, Galileo showed the motions of the four satellites, examples of his diagrams being shown in Fig. 2.9. Kepler was in full flow:

> The conclusion is quite clear. Our moon exists for us on Earth, not for the other globes. Those four little moons exist for Jupiter, not for us. Each planet in turn, together with its occupants, is served by its own satellites. From this line of reasoning we deduce with the highest degree of probability that Jupiter is inhabited.[19]

This extreme form of lateral thinking was part of the personality of the greatest mathematician in Europe at the time.

Kepler's third law was deeply buried in his treatise, the *Harmonices Mundi* or the *Harmony of the World*. According to Max Casper, this work, published in 1619, was his crowning achievement in which he attempted to synthesise all his ideas into one harmonious picture of the Universe. His harmonic theory was to encompass geometry, music, architecture, metaphysics, psychology, astrology and astronomy, as can be seen from the contents page of the five books which make up the treatise (Fig. 2.10). In modern terms, this was his 'Grand Unified Theory'.

By 1619, Kepler was no longer satisfied with an accuracy of 5% in the comparison of the radii of the planetary orbits with his harmonic theory. He now had much more accurate mean radii for the orbits of the planets and suddenly, by the time he had reached the writing of Book V, Chapter III, 8th Division of the *Harmony of the World*, he discovered what is now known as *Kepler's third law of planetary motion*:



Ioannis Keppleri

HARMONICES MVNDI

LIBRI V. QVORVM

Primus GEOMETRICVS, De Figurarum Regularium, quæ Proportiones Harmonicas conſtituunt, ortu & demonſtrationibus.

Secundus ARCHITECTONICVS, ſeu ex GEOMETRIA FIGVRATA, De Figurarum Regularium Congruentia in plano vel ſolido:

Tertius propriè HARMONICVS, De Proportionum Harmonicarum ortu ex Figuris; deque Naturâ & Differentiis rerum ad cantum pertinentium, contra Veteres:

Quartus METAPHYSICVS, PSYCHOLOGICVS & ASTROLOGICVS, De Harmoniarum mentali Eſſentiâ earumque generibus in Mundo; præſertim de Harmonia radiorum, ex corporibus cœleſtibus in Terram deſcendentibus, eiuſque effectu in Natura ſeu Anima ſublunari & Humana:

Quintus ASTRONOMICVS & METAPHYSICVS, De Harmoniis abſolutiſſimis motuum cœleſtium, ortuque Eccentricitatum ex proportionibus Harmonicis.

Appendix habet comparationem huius Operis cum Harmonices Cl. Ptolemæi libro III. cumque Roberti de Fluctibus, dicti Flud. Medici Oxonienſis ſpeculationibus Harmonicis, operi de Macrocoſmo & Microcoſmo inſertis.

Cum S.C. M^tis. Priuilegio ad annos XV.

Lincii Auſtriæ,

Sumptibus GODOFREDI TAMPACHII Bibl. Francof.
Excudebat IOANNES PLANCVS.

ANNO M. DC. XIX.

Fig. 2.10 The table of contents of Kepler's treatise *The Harmony of the World*, 1619, Linz. (From the Crawford Collection, Royal Observatory, Edinburgh.)

> *The period of a planetary orbit is proportional to the three-halves power of the mean distance of the planet from the Sun.*

This was the crucial discovery which was to lead to Newton's law of gravity. Notice that this law is somewhat different from the solution he had proposed in the *Mysterium Cosmographicum*. It is not in conflict with it, however, since there is nothing in Kepler's third law to tell us why the planets have to lie at particular distances from the Sun. It is noteworthy that Kepler's discoveries of the three laws of planetary motion were intuitive leaps of the imagination, but he had a mastery of the geometric techniques which enabled him to convert these into quantitative science.

It might be thought that the support of Kepler, a passionate Copernican, would have been invaluable to Galileo during his prosecution for his advocacy of the Copernican picture of the World. As we will see, matters were not so simple – Galileo was somewhat cautious about Kepler's support. Two years after Kepler's death on 15 November 1630, Galileo wrote:

> I do not doubt but that the thoughts of Landsberg and some of Kepler's tend rather to the diminution of the doctrine of Copernicus than to its establishment as it seems to me that these (in the common phrase) have wished too much for it . . .[20]

We have already quoted some of Kepler's purple prose which indicates part of the concern – he was ebulliently over-enthusiastic about his remarkable discoveries. In addition, the strong mystical streak in the *Harmony of the World* would not have appealed to Galileo, who was attempting to set the whole of natural philosophy on a secure mathematical foundation.

Another of Galileo's concerns is of equal interest. He wrote:

> It seems to me that one may reasonably conclude that for the maintenance of perfect order among the parts of the Universe, it is necessary to say that movable bodies are movable only circularly.[21]

Kepler's assertion that the orbits of the planets are ellipses rather than circles was intellectually unappealing to Galileo – we recognise the legacy of Aristotelian physics, according to which the only allowable motions are uniform linear and circular motions. We would now say that this view was only a prejudice on the part of Galileo, but we should not disguise the fact that we all make similar value judgements in our own work.

Eventually Kepler completed the *Rudolphine Tables* and they were published in September 1627. These set a new standard of accuracy in the predictions of solar, lunar and planetary positions. It is intriguing that, in order to simplify the calculations, he invented his own form of logarithms. Kepler had seen John Napier's *Mirifici logarithmorum canonis descritpio* of 1614 as early as 1617 – this was the first set of tables of natural logarithms.

Although it was not recognised at the time, Kepler's three laws of planetary motion were to be crucial for Newton's great synthesis of the laws of gravity and celestial mechanics. But there was another giant whose contributions were even more profound – Galileo Galilei. He has become a symbol of the birth of modern science and the struggle against received dogma. The trial of Galileo strikes right at the very heart of the modern concept of the scientific method. But he contributed much more as we will see Chapter 3.

Notes

1 Pedersen, O. (1993). *Early Physics and Astronomy*, pp. 346–347. Cambridge: Cambridge University Press.
2 Pedersen, O. (1993). *op. cit.* p. 55.
3 Pedersen, O. (1993). *op. cit.* p. 56.
4 Pedersen (1993) can be consulted for more details of the complexities of the Ptolemaic system of the world.
5 For translation, see Duncan, A.M. (1976). *Copernicus: On the Revolutions of the Heavenly Spheres. A New Translation from the Latin*. London: David and Charles; New York: Barnes and Noble Books.
6 In 1692, Richard Bentley gave the first series of Boyle Lectures, which Robert Boyle had set up 'to combat atheism'. In correspondence between Bentley and Newton, they discussed the issue of the stability of a finite or infinite Universe under the attractive force of gravity. They concluded that the Universe had to be infinite since otherwise it would collapse to its centre under gravity. They understood, however, that an infinite Universe filled with stars is gravitationally unstable. As expressed by Harrison,

> (Newton) agreed with Bentley that providence had designed a Universe of infinite extent in which uniformly distributed stars stand poised in unstable equilibrium like needles on their points.

(Harrison, E.R. (1989). *Darkness at Night: A Riddle of the Universe*, p. 73. Cambridge, Massachusetts: Harvard University Press.)
7 See Gingerich, O. (2002). *An Annotated Census of Copernicus' De revolutionibus (Nuremberg, 1543 and Basel, 1566)*. Leiden and Boston, Massachusetts: Studia Copernicana, Brill.
8 Hellman, D.C. (1970). *Dictionary of Scientific Biography*, Vol. 11, p. 403. New York: Charles Scribner's Sons.
9 Dreyer, J.L.E. (1890). *Tycho Brahe: A Picture of Scientific Life and Work in the Sixteenth Century*, pp. 86–87. Edinburgh: Adam & Charles Black.
10 Christianson, J. (1961). *Scientific American*, **204**, 118 (February issue).
11 Thoren, V.E. (1990). *The Lord of Uraniborg: A Biography of Tycho Brahe*, pp. 188–189. Cambridge: Cambridge University Press.
12 *Astronomiae Instauratae Mechanica* was published in Wandsbek in 1598. When Tycho had to flee from Denmark in 1597 taking his instruments and printing press with him, he spent two years in Wandsbek, now a borough of Hamburg, as a guest of Count Rantzau. It was here that Tycho produced the book using his printing press. A subsequent edition was published by Levinus Hulsius of Nuremberg in 1602, the year after Tycho's death.
13 Casper, M. (1959). *Kepler*, trans. C. Doris Hellman, pp. 46–47. London and New York: Abelard-Schuman.
14 Casper, M. (1959). *op. cit.*, p. 20.
15 Casper, M. (1959). *op. cit.*, p. 62.
16 Kepler, J. (1596). From *Mysterium Cosmographicum*. See *Kepleri opera omnia*, ed. C. Frisch (1858), Vol. 1, pp. 9ff.
17 Casper, M. (1959). *op. cit.*, p. 71.
18 Kepler, J. (1609). From *Astronomia Nova*. See *Johannes Kepler Gesammelte Werke*, ed. M. Casper, Vol. III, p. 178, Munich: Beck (1937).
19 Kepler, J. (1610). See *Kepler's Conversation with Galileo's Sidereal Messenger*, ed. and trans. E. Rosen (1965), p. 42. New York and London: Johnson Reprint Corporation.
20 Galileo, G. (1630). Quoted by Rupert Hall, A. (1970). *From Galileo to Newton 1630–1720. The Rise of Modern Science 2*, p. 41. London: Fontana Science.
21 Galilei, G. (1632). *Dialogues concerning the Two Chief Systems of the World*, trans. S. Drake, p. 32, Berkeley (1953).

3 Galileo and the Nature of the Physical Sciences

3.1 Introduction

There are three stories to be told. The first is Galileo as *natural philosopher*. Unlike Tycho Brahe and Kepler, Galileo was an experimental physicist whose prime concern was understanding the laws of nature in quantitative terms, from his earliest writings to his final great treatise *Discourse and Mathematical Demonstrations concerning Two New Sciences*. The second story is *astronomical* and occupies a relatively small, but crucial, period of his career from 1609 to 1612, when his astronomical discoveries had an immediate international impact.

The third is his trial and subsequent house arrest. The scientific aspects of his censure and trial by the Inquisition strike right at the heart of the nature of the physical sciences. The general view is to regard Galileo as the hero and the Catholic Church as the villain of the piece, a source of conservative reaction and blinkered authority. From the methodological point of view, Galileo made an logical error, but the church authorities made more serious blunders, which have resonated through the history of science ever since. Some acknowledgement of the Church's misjudgements was only made officially by Pope John Paul II in 1992.[1]

My reasons for devoting a whole chapter to Galileo, his science and tribulations are partly because the facts of the case need to be better known but more significantly because the story has resonances for the way in which physics as a scientific discipline is carried out. Galileo's intellectual integrity and scientific genius are an inspiration – more than anyone else, he can be credited with creating the intellectual framework for the study of physics as we know it today.

3.2 Galileo as an Experimental Physicist

Galileo Galilei was born in February 1564 in Pisa, the son of Vincenzio Galileo, the distinguished musician and musical theorist. In 1587, Galileo was appointed to the chair of mathematics at the University of Pisa, where he was not particularly popular with his colleagues, one of the main causes being his opposition to Aristotelian physics, which remained the central pillar of natural philosophy. It was apparent to Galileo that Aristotle's physics was not in accord with the way in which matter actually behaves. For example, the assertion concerning the fall of bodies of different weights reads as follows:

> If a certain weight move a certain distance in a certain time, a greater weight will move the same distance in a shorter time, and the proportion which the weights bear to each other, the times too will bear to one another; for example, if the half weight cover the distance in x, the whole weight will cover it in $x/2$.[2]

This is just wrong, as could have been demonstrated by a simple experiment. Galileo's objection is symbolised by the story of his dropping different weights from the Leaning Tower of Pisa. If they are dropped through the same height, they take the same time to reach the ground if the effects of air resistance are neglected, as was known to Galileo and earlier writers.

In 1592, Galileo was appointed to the chair of mathematics at Padua, where he was to remain until 1610. During this period he produced some of his greatest work. Initially, he was opposed to the Copernican model of the Solar System but, in 1595, he began to take it seriously in order to explain the origin of the tides in the Adriatic. He observed that the tides at Venice typically rise and fall by about 5 feet, about 1.5 metres, and therefore there must be quite enormous forces to cause this huge amount of water to be raised each day at high tide. Galileo reasoned that, if the Earth rotated on its own axis and also moved in a circular orbit about the Sun, the changes in the direction of travel of a point in the surface of the Earth would cause the sea to slosh about and so produce the effect of the tides. This is not the correct explanation for the tides, but it led Galileo to favour the Copernican picture for physical reasons.

In Galileo's printed works, the arguments are given entirely in abstract terms without reference in the conventional sense to experimental evidence. Galileo's genius as a pioneer scientist is described by Stillman Drake in his remarkable book *Galileo: Pioneer Scientist* (1990). Drake deciphered Galileo's unpublished notes and convincingly demonstrated that Galileo actually carried out the experiments to which he refers in his treatises with considerable experimental skill (Fig. 3.1).

Galileo's task was daunting – he disbelieved the basis of Aristotelian physics, but had no replacement for it. In the early 1600s, he undertook experimental investigations of the laws of fall, of the motion of balls rolling down slopes and of the motion of pendulums – these clarified the concept of *acceleration* for the first time.

A problem for physics up to the time of Galileo was that there was no way of measuring short time intervals accurately, and so he had to use considerable ingenuity in the design of his experiments. A splendid example is his experiment to investigate how a ball accelerates on rolling down a slope. He constructed a long shallow slope of length 2 metres at an angle of only 1.7° to the horizontal and cut a grove in it down which a heavy bronze ball could roll. He placed little frets on the slope so that there would be a click as the ball passed over each fret. He then adjusted the positions of the frets so that the little clicks occurred at equal time intervals (Fig. 3.2). Drake suggests that he could equalise the time intervals to about 1/64 of a second by singing a rhythmic tune and making the clicks occur at equal beats in the bar. By taking differences, he worked out the average speed between successive frets and found that it increased as the odd numbers 1, 3, 5, 7, . . . in equal time intervals.

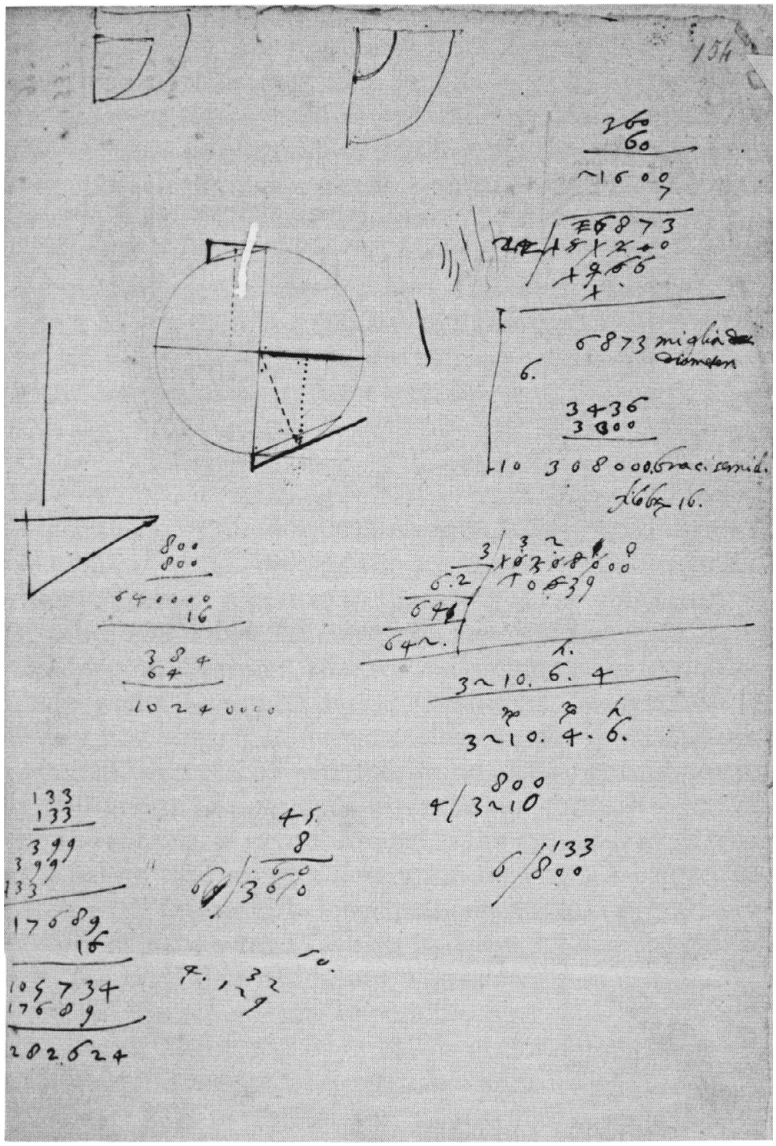


Fig. 3.1 Part of Galileo's notes concerning the laws of the pendulum. (From S. Drake, 1990. *Galileo: Pioneer Scientist*, p. 19. Toronto: University of Toronto Press.)

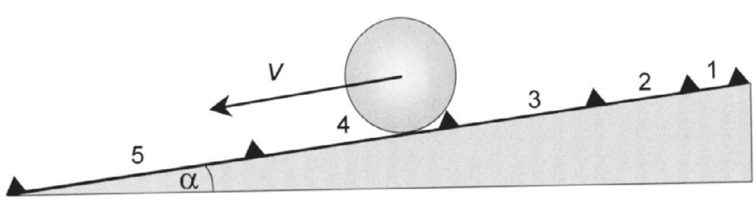


Fig. 3.2 How Galileo established the law of motion under uniform acceleration. The numbers between the frets show their relative positions in order to produce a regular sequence of clicks.

Originally, Galileo had believed that, under constant acceleration, speed is proportional to the distance travelled but, as a result of his experiments of 1604, he found rather that the speed is proportional to time. He now had two relations: the first was that between speed and distance, $x = vt$ for constant speed v, and the second related speed to time under constant acceleration $v = at$. There is no algebra in Galileo's works and the differential calculus had yet to be discovered. The simplest procedure was to break up the problem into finite time intervals. Suppose the speeds of the uniformly accelerated sphere are measured at times $0, 1, 2, 3, 4, 5, \ldots$ s, assuming it starts from rest at time 0. The acceleration is constant and so the speeds at these times are, say, $0, 1, 2, 3, 4, \ldots$ cm s^{-1}, an acceleration $a = 1$ cm s^{-2}. How far has the particle travelled after $0, 1, 2, 3, 4, 5, \ldots$ s?

At $t = 0$, no distance has been travelled. Between 0 and 1 s, the average speed is 0.5 cm s^{-1} and so the distance travelled is 0.5 cm. In the next interval between 1 and 2 s, the average speed is 1.5 cm s^{-1} and so the distance travelled in that interval is 1.5 cm and the total distance travelled from the position of rest is $0.5 + 1.5 = 2$ cm. In the next interval, the average speed is 2.5 cm s^{-1}, the distance travelled is 2.5 cm and the total distance 4.5 cm, and so on. We obtain a series of distances $0, 0.5, 2, 4.5, 8, 12.5, \ldots$ cm which can be written

$$x = \tfrac{1}{2}(0,\ 1,\ 4,\ 9,\ 16,\ 25, \ldots) = \tfrac{1}{2}(0,\ 1^2,\ 2^2,\ 3^2,\ 4^2,\ 5^2, \ldots)\ \text{cm}. \qquad (3.1)$$

This is Galileo's famous *times-squared law* for uniformly accelerated motion:

$$x = \tfrac{1}{2}at^2. \qquad (3.2)$$

This result represented a revolution in thinking about the nature of accelerated motion and was to lead to the Newtonian revolution.

He did not stop there but went on to carry out further key experiments. He next studied the law of fall, namely, if an object is dropped from different heights, how long does it take it to hit the ground? He used a form of water clock to measure time intervals accurately. Water was allowed to pour out of a tube at the bottom of a large vessel so that the amount of water which had flowed out was a measure of the time interval. By dropping objects from different heights, he established that falling objects obey the times-squared law – in other words, when objects fall freely they experience a constant acceleration, the *acceleration due to gravity*.

Having established these two results, he sought a relation between them – the answer is *Galileo's theorem*. Suppose a sphere is dropped freely through a certain distance l, which is represented by the length AB in Fig. 3.3. Construct a circle with diameter equal to the length l as shown. Now suppose the ball *slides* down an inclined plane AC without friction and, for convenience, that the top of the plane is placed at the point A. Galileo's theorem states:

> *The time it takes a ball to slide down the slope from A to the point C, where the slope cuts the circle, is equal to the time it takes the ball to fall freely from A to B.*

In other words, the time it takes a body to fall along any *chord of a circle* is the same as the time it takes to fall vertically downwards across the diameter of the circle. As the sphere slides down the slope, the component of the acceleration due to gravity $g \sin \alpha$ acts on it,

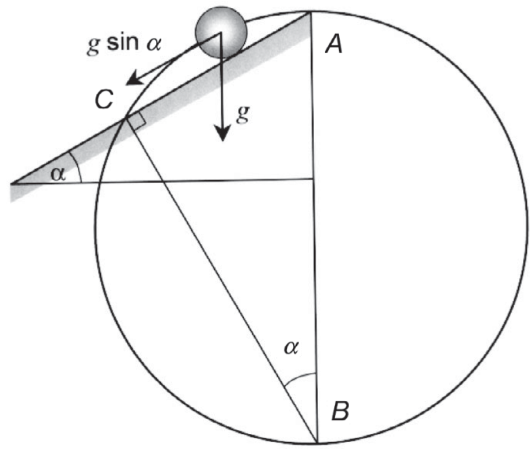


Fig. 3.3 How Galileo established Galileo's theorem.

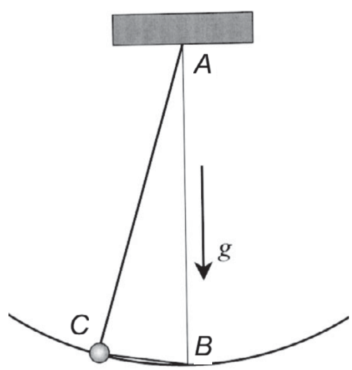


Fig. 3.4 How Galileo showed that the period of a long pendulum is independent of the amplitude of the swing.

the rest of the acceleration being balanced by the reaction force of the slope itself (Fig. 3.3). But, from the geometry of the figure, any triangle constructed on the diameter of a circle, the third point lying on the circle, is a right-angled triangle. Therefore, we can equate the angles as shown in Fig. 3.3, from which it is apparent that $AC/AB = \sin\alpha$, the ratio of the distances travelled. Since the distance travelled is proportional to the acceleration, $x = \frac{1}{2}at^2$, this proves Galileo's theorem.[3]

Galileo next recognised the relation between these phenomena and the properties of swinging pendulums. As a youth, he is said to have noticed that the period of swing of a chandelier suspended in a church is independent of the *amplitude* of its swing. Galileo made use of his law of chords to explain this observation. If the pendulum is long enough, the arc described by the pendulum is almost exactly equal to the chord across the circle joining the extreme swing of the pendulum to its lowest point (Fig. 3.4). It is therefore obvious why the period of the pendulum is independent of the amplitude of its swing – according to Galileo's theorem, the time to travel along any chord drawn from B is always

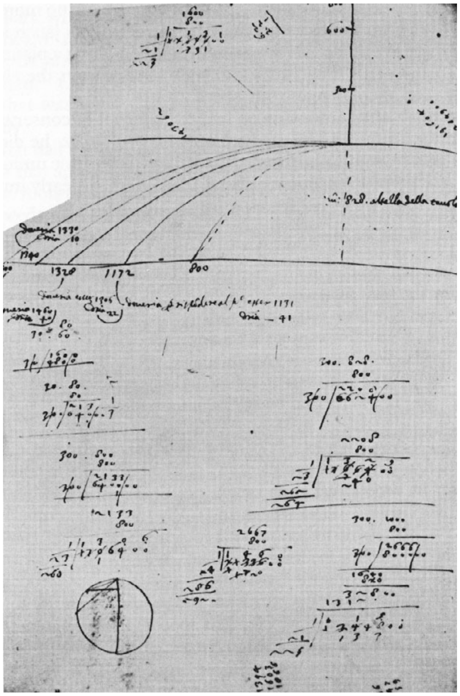


Fig. 3.5 A page from Galileo's notebooks showing the trajectories of projectiles under the combination of the acceleration under gravity and constant horizontal speed. (From S. Drake, 1990. *Galileo: Pioneer Scientist*, p. 107. Toronto: University of Toronto Press.)

the same as the time it takes the body to fall freely down twice the length of the pendulum, as can be seen in Fig. 3.3.

What Galileo had achieved was to put into mathematical form the nature of acceleration under gravity. This had immediate practical application because he could now work out the trajectories of projectiles. They travel with constant speed parallel to the ground and are decelerated by the force of gravity in the vertical direction. For the first time, he was able to work out the parabolic paths of cannon balls and other projectiles (Fig. 3.5).

Galileo began writing a systematic treatment of all these topics, showing how they could all be understood on the basis of the law of constant acceleration. In 1610, he was planning to write:

> Three books on mechanics, two with demonstrations of its principles, and one concerning its problems; and though other men have written on the subject, what has been done is not one-quarter of what I write, either in quantity or otherwise.[4]

Later he writes in the same vein:

> Three books on local motion – an entirely new science in which no one else, ancient or modern, has discovered any of the most remarkable laws which I demonstrate to exist in both natural and violent movement; hence I may call this a new science and one discovered by me from its very foundations.[5]

He was diverted from this task by news of the invention of the telescope, this event marking the beginning of his serious study of astronomy. Then, there followed his trial in 1615–16. As a result, the publication of Galileo's discoveries was delayed until the 1620s and 1630s.

3.3 Galileo's Telescopic Discoveries

The invention of the telescope is attributed to the Dutch lens-grinder Hans Lipperhey, who in October 1608 applied unsuccessfully to Count Maurice of Nassau for a patent for a device which could make distant objects appear closer. Galileo heard of this invention in July 1609 and set about constructing one himself. By August, he had achieved a magnification of 9 times, three times better than Lipperhey's telescope. This greatly impressed the Venetian Senate, who understood the military importance of such a device for a maritime nation. Galileo was immediately given a lifetime appointment at the University of Padua at a vastly increased salary.

By the end of 1609, Galileo had made a number of telescopes of increasing magnification, culminating in achieving a magnifying power of 30. In January 1610, he first turned his telescope on the skies and immediately there came a flood of remarkable discoveries. These were rapidly published in March 1610 in his *Sidereus Nuncius* or *The Sidereal Messenger*.[6] In summary, the discoveries were:

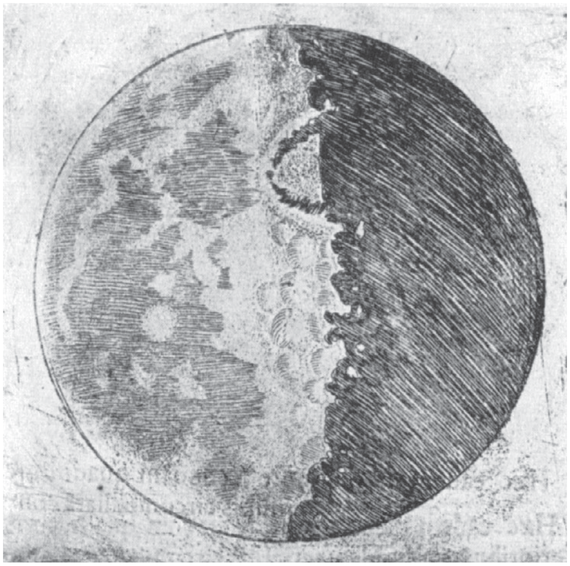
(1) The Moon is mountainous rather than being a perfectly smooth sphere (Fig. 3.6(*a*)).
(2) The Milky Way consists of vast numbers of stars rather than being a uniform distribution of light (Fig. 3.6(*b*)).
(3) Jupiter has four satellites and their motions were followed over a period of almost two months, enabling the periods of their orbits to be determined (Fig. 2.9).

The book caused a sensation throughout Europe and Galileo won immediate international fame. These discoveries demolished a number of Aristotelian precepts which had been accepted over the centuries. For example, the resolution of the Milk Way into individual stars was quite contrary to the Aristotelean view. In the satellites of Jupiter, Galileo saw a prototype for the Copernican picture of the Solar System. The immediate effect of these discoveries was that Galileo was appointed Mathematician and Philosopher to the Grand Duke of Tuscany, Cosimo de Medici, to whom the *Sidereus Nuncius* was dedicated.

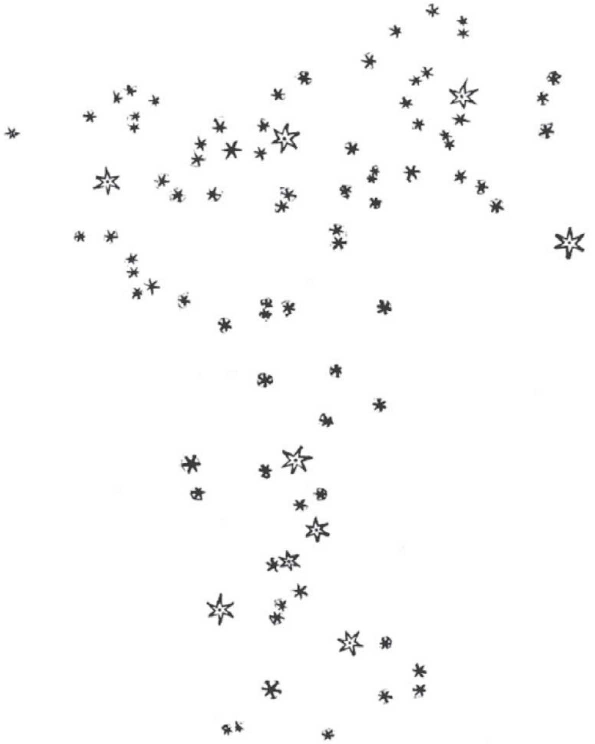
Later in 1610, he made two other key telescopic discoveries:

(4) the rings of Saturn, which he took to be close satellites of the planet;
(5) the phases of the planet Venus.

This last discovery was of the greatest significance. With his telescope, Galileo was able to resolve the disc of Venus and observe its illumination by the Sun. When Venus is on the far side of its orbit with respect to the Earth, the disc is circular in appearance, but when it is on the same side of the Sun as the Earth, it looks like a crescent Moon. This was interpreted as evidence in favour of the Copernican picture because it finds a natural explanation if



(*a*)



(*b*)

Fig. 3.6 (*a*) Galileo's drawing of the Moon as observed through his telescope. In fact, the large crater intersected by the lunar terminator is not real. The diagram should be regarded as figurative. (*b*) Galileo's sketch of the region of sky in the vicinity of Orion's belt, showing the resolution of the background light into faint stars. (From G. Galilei, 1610. *Sidereus Nuncius*, Venice. See also the translation by A. van Helden, 1989. Chicago: University of Chicago Press.).

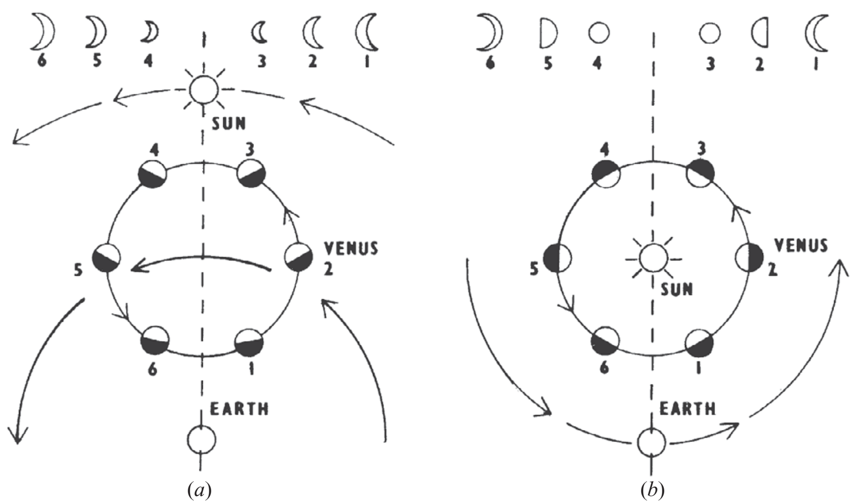


Fig. 3.7 Illustrating the phases of Venus, according to (*a*) the Geocentric and (*b*) Heliocentric pictures of the structure of the Solar System. (From A. van Helden, 1989. *Sidereus Nuncius, or The Sidereal Messenger*, p. 108. Chicago: University of Chicago Press.)

Venus and the Earth both orbit the Sun, the source of their illumination (Fig. 3.7(*b*)). If, on the other hand, Venus moved on an epicycle about a circular orbit around the Earth and the Sun moved on a more distant sphere, the pattern of illumination relative to the Earth would be quite different, as illustrated in Fig. 3.7(*a*). In 1611 these discoveries were presented by Galileo to the Pope and several cardinals, who were positively impressed by them. Galileo was elected to the Academia Lincei.

3.4 Aristotelian versus Galilean Physics: The Heart of the Matter

Before recounting the events which led to Galileo's appearance in 1632 before the Inquisition and his conviction for the second most serious crime in the papal system of justice, let us summarise some of the various facets of the debate between the Ptolemaeans, the Copernicans and the church authorities. Finocchiaro provides an excellent summary in his documentary history *The Galileo Affair*.[7] The accepted laws of physics remained Aristotelian, and only a few adventurous spirits doubted the correctness of these and of the Ptolemaic system of the world. There were problems with the Copernican picture however and so Galileo *had* to become involved in these issues which had the potential to undermine his new-found understanding of the laws of motion.

3.4.1 The Issues

The physical issues centred on the questions: (i) Does the Earth rotate on its axis with respect to the fixed stars, and (ii) do the Earth and the planets orbit the Sun?

Specifically, is the Earth in motion? This latter concept was referred to as the *geokinetic hypothesis*. Finocchiaro summarises the arguments of the Aristotelian philosophers under five headings.

(1) *The deception of the senses*. None of our senses gives us any evidence that the Earth is moving in an orbit about the Sun. If this were a fact of nature, surely it would be of such importance that our senses would make us aware of it.
(2) *Astronomical problems*. First, the heavenly bodies were believed to be composed of different forms of matter from the material of the Earth. Second, Venus should show phases similar to the Moon if it were in orbit about the Sun. Third, if the Earth moved, why didn't the stars exhibit parallaxes?
(3) *Physical arguments*. These were based upon Aristotelian physics and some have been discussed already.

    (a) If the Earth moves, falling bodies should not fall vertically. Many counter-examples could be given – rain falls vertically, objects thrown vertically upwards fall straight down again, and so on. This was in contrast to the trajectory of an object dropped from the top of a ship's mast when a ship is in motion. In this case, the object does not fall vertically downwards.
    (b) Projectiles fired in the direction of rotation of the Earth and in the opposite direction would have different trajectories. No such difference had been observed.
    (c) Objects placed on a rotating potter's wheel are flung off if they are not held down, what was referred to as the *Extruding Power of Whirling* – this is what we now call *centrifugal force*. The same phenomenon should occur if the Earth is in a state of rotation, but we are not flung off the surface of the Earth.
    (d) Next, there were purely theoretical arguments. According to Aristotelian physics, there are only two forms of motion, uniform motion in a straight line and uniform circular motion – these were the only 'natural' motions. Objects must either fall in a straight line to the centre of the Universe, or be in a state of uniform circular motion. We have already discussed the argument that objects fall towards the centre of the Earth and not towards the Sun. Furthermore, according to Aristotelian physics, simple bodies could only have one natural motion. But, according to Copernicus, objects dropped on Earth have *three* motions – downward motion under free fall, rotation of the Earth on its axis and motion in a circular orbit about the Sun.
    (e) Finally, if Aristotelian physics was to be rejected, what was to replace it? The Copernicans had to provide a better theory and none was available.

(4) The *Authority of the Bible*. There are no unambiguous statements in the Bible that assert that the Earth is stationary at the centre of the Universe. According to Finocchiaro, the most relevant statements are as follows:[8]

    (a) Psalm 104:5. 'O Lord my God . . . who laid the foundations of the Earth, that it should not be removed forever.'
    (b) Ecclesiastes 1:5. 'The Sun also riseth, and the Sun goeth down, and hasteth to the place where he ariseth.'

(c) Joshua 10:12,13. 'Then spake Joshua to the Lord in the day when the Lord delivered up the Amorites before the children of Israel, and he said in the sight of Israel, "Sun, stand thou still upon Gibeon; and thou, Moon, in the valley of Ajalon." And the Sun stood still, and the Moon stayed, until the people had avenged themselves upon their enemies.'

These are rather oblique references and it is intriguing that the Protestants were much more virulently anti-Copernican than the Catholic church because of their belief in the literal truth of the Bible. The Catholic theologians took a more sophisticated and flexible interpretation of the holy writ. However, the concept that the Earth is stationary at the centre of the Universe was also the conclusion of the *Church Fathers* – the saints, theologians and churchmen who codified Christianity. To quote Finocchiaro,

> The argument claimed that all Church Fathers were unanimous in interpreting relevant Biblical passages ... in accordance with the geostatic view; therefore the geostatic system is binding on all believers, and to claim otherwise (as Copernicus did) is heretical.[9]

(5) The most interesting argument from our perspective concerns the *hypothetical nature of Copernican theory*. It strikes at the very heart of the nature of the natural sciences. The crucial point is how we express statements concerning the success of the Copernican model. A correct statement is:

> *If* the Earth rotates on its axis and moves in a circular orbit about the Sun, and if the other planets also orbit the Sun, *then* we can describe simply and elegantly the observed motions of the Sun, Moon and planets on the celestial sphere.

What we *cannot* do logically is to reverse the argument and state that, *because* the planetary motions are explained simply and elegantly by the Copernican hypothesis, *therefore* the Earth must rotate and move in a circular orbit about the Sun. This is an elementary error of logic, because there might well be other models which mimic the success of the Copernican model.

The key point is the difference between *inductive* and *deductive* reasoning. Owen Gingerich gives a pleasant example.[10] A *deductive* sequence of arguments might run:

(*a*) If it is raining, the streets are wet.
(*b*) It is raining.
(*c*) Therefore, the streets are wet.

There is no problem here. But, now reverse (*b*) and (*c*) and we get into trouble:

(*a*) If it is raining, the streets are wet.
(*b*) The streets are wet.
(*c*) Therefore, it is raining.

This line of reasoning is obviously false since the streets could be the streets of Venice, or could have been newly washed. In other words, you cannot prove anything about the absolute truth of statements of this second type. This second type of reasoning in which we attempt to find general laws from specific pieces of evidence is called *inductive reasoning*. All the physical sciences are to a greater or lesser extent of this 'hypothetical nature'. This is in contrast to the absolute certainty of God's word as contained in the holy scriptures and its interpretation as dogma by the Church Fathers. According to Owen Gingerich, this was the issue of substance which led to the trial and censure of Galileo – the juxtaposition of the hypothetical world picture of Copernicus with the absolute truth as revealed in the Bible and codified by the Church Fathers.

3.4.2 The Galileo Affair

Prior to his great telescopic discoveries of 1610–11, Galileo was at best a cautious Copernican, but it gradually became apparent to him that his new understanding of the nature of motion eliminated the *physical problems* listed above. The new evidence was consistent with the Copernican model. Specifically, there are mountains on the Moon, just as there are on Earth, suggesting that the Earth and the Moon are similar bodies. The phases of Venus are exactly as would expected according to the Copernican picture. Thus, the physical and astronomical objections to Copernicanism could be discarded, leaving only the theological and logical problems to be debated.

As the evidence began to accumulate in favour of Copernicanism, conservative scientists and philosophers had to rely more and more upon the theological, philosophical and logical arguments. In December 1613, the Grand Duchess Dowager Christina asked Castelli, one of Galileo's friends and colleagues, about the religious objections to the motion of the Earth. Castelli responded to the satisfaction of both the Duchess and Galileo, but Galileo felt the need to set out the arguments in more detail. He suggested that there were three fatal flaws in the theological arguments. To quote Finocchiaro:

> First, it attempts to prove a conclusion (the earth's rest) on the basis of a premise (the Bible's commitment to the geostatic system) which can only be ascertained with a knowledge of that conclusion in the first place. … The business of Biblical interpretation is dependent on physical investigation, and to base a controversial physical conclusion on the Bible is to put the cart before the horse. Second, the Biblical objection is a *non sequitur*, since the Bible is an authority only in matters of faith and morals, not in scientific questions. … Finally, it is questionable whether the Earth's motion really contradicts the Bible.[11]

This letter circulated privately and came into the hands of the conservatives. Sermons were delivered attacking the heliocentric picture and its proponents were accused of heresy. In March 1615, the Dominican Friar Tommaso Caccini, who had already preached against Galileo, laid a formal charge of *suspicion of heresy* against Galileo before the Roman Inquisition. This charge was less severe than that of *formal heresy*, but was still a serious one. The Inquisition manual stated, 'Suspects of heresy are those who occasionally utter propositions that offend the listeners … Those who keep, write, read, or give others to read books forbidden in the index …' Further, there were two types of suspicion of heresy, *vehement* and *slight* suspicion of heresy, the former being considerably more serious than the latter. Once an accusation was made, there was a formal procedure which had to be followed.

Galileo responded by seeking the support of his friends and patrons and circulated three long essays privately. One of these repeated the arguments concerning the validity of the theological arguments and became known as *Galileo's letter to the Grand Duchess Christina*; the revised version was expanded from eight to forty pages. By good fortune, a Neapolitan Friar, Paolo Antonio Foscarini, published a book in the same year arguing in detail that the Earth's motion was compatible with the Bible. In December 1615, after a delay due to illness, Galileo himself went to Rome to clear his name and to prevent Copernicanism being condemned as heretical.

So far as Cardinal Roberto Bellarmine, the leading Catholic theologian of the day, was concerned, the main problem concerned the hypothetical nature of the Copernican picture. Here are his words, written on 12 April 1615 to Foscarini, after the *Letter to Christina* was circulated in Rome.

> ... it seems to me that Your Paternity (Foscarini) and Mr. Galileo are proceeding prudently by limiting yourselves to speaking suppositionally[12] and not absolutely, as I have always believed that Copernicus spoke. For there is no danger in saying that, by *assuming* the earth moves and the sun stands still, one saves all the appearances better than by postulating eccentrics and epicycles is to speak well; and that is sufficient for the mathematicians.
>
> However, it is different to want to affirm that in reality the Sun is at the centre of the world and only turns on itself without moving from east to west, and the Earth is in the third heaven and revolves with great speed around the Sun; this is a very dangerous thing, likely not only to irritate all scholastic philosophers and theologians, but also to harm the Holy Faith by rendering Holy Scripture false.[13]

Behind these remarks is a perfectly valid criticism of Galileo's support for the Copernican picture. It is not correct to infer, as Galileo did, that the observation of the phases of Venus proves that the Copernican picture is correct. For example, in Tycho's cosmology in which the planets orbit the Sun, but the Sun and the planets orbit the Earth (Fig. 2.7), exactly the same phases of Venus would be observed as in the Copernican picture. According to Gingerich, this was Galileo's crucial logical error. Strictly speaking, he could only make a hypothetical statement.

The findings of the Inquisition were favourable to Galileo personally – he was acquitted of the charge of suspicion of heresy. However, the Inquisition also asked a committee of eleven consultants for an opinion on the status of Copernicanism. On 16 February 1616, it reported unanimously that Copernicanism was philosophically and scientifically untenable and theologically heretical. This judgement was the prime cause of the subsequent condemnation of Galileo. It seems that the Inquisition had misgivings about this outcome because it issued no formal condemnation. Instead, it issued two milder instructions. First, Galileo was given a private warning by Cardinal Bellarmine to stop defending the Copernican world picture. Exactly what was said is a matter of controversy, but Bellarmine reported back to the Inquisition that the warning had been delivered and that Galileo had accepted it.

The second result was a public decree by the Congregation of the Index. First, it reaffirmed that the doctrine of the Earth's motion was heretical; second, Foscarini's book was condemned and prohibited by being placed on the Index; third, Copernicus's *De Revolutionibus* was suspended until a few offending passages were amended; fourth, all similar books were subject to the same prohibition.

Rumours circulated that Galileo had been tried and condemned by the Inquisition and, to correct matters, Bellarmine issued a brief statement to the effect that Galileo had neither been tried nor condemned, but that he had been informed of the Decree of Index and instructed not to hold or defend the Copernican picture. Although personally exonerated, the result was a clear defeat for Galileo. Although not formally a heresy, holding Copernican opinions amounted to very much the same thing.

3.5 The Trial of Galileo

For the next seven years, Galileo kept a low profile and complied with the papal instruction. In 1623, Gregory XV died and his successor, Cardinal Maffeo Barbarini, was elected Pope Urban VIII. He was a Florentine who took a more relaxed view of the interpretation of the scriptures than his predecessor. An admirer of Galileo, he adopted the position that Copernicanism could be discussed hypothetically and it might well prove to be of great value in making astronomical predictions. Galileo had six conversations with Urban VIII in spring 1624 and came to the conclusion that Copernicanism could be discussed, provided it was only considered hypothetically.

Galileo returned to Florence and immediately set about writing the *Dialogue on the Two Chief World Systems, Ptolemaic and Copernican*. He believed he had made every effort to comply with the wishes of the censors. The preface was written jointly by Galileo and the censors and, after some delay, the great treatise was published in 1632. Galileo wrote the book in the form of a dialogue between three speakers, Simplicio defending the traditional Aristotelian and Ptolemaic positions, Salviati the Copernican position and Sagredo an uncommitted observer and man of the world. Consistently, Galileo argued that the purpose was not to make judgements, but to pass on information and enlightenment. The book was published with full papal authority.

The Two Chief World Systems was well received in scientific circles, but very soon complaints and rumours began to circulate in Rome. A document dated February 1616, almost certainly a fabrication, was found in which Galileo was specifically forbidden from discussing Copernicanism in any form. By now, Cardinal Bellarmine had been dead 11 years. In fact, Galileo had not treated the Copernican model hypothetically at all, but rather as a fact of nature – Salviati is Galileo speaking his own mind. The Copernican system was portrayed in a much more favourable light than the Ptolemaic picture, contradicting Urban VIII's conditions for discussion of the two systems of the world.

The Pope was forced to take action – papal authority was being undermined at a time when the counter-reformation and the reassertion of that authority were paramount political considerations. Galileo, now 68 years old and in poor health, was ordered to come to Rome under the threat of arrest. The result of the trial was a foregone conclusion. In the end, Galileo pleaded guilty to a lesser charge on the basis that, if he had violated the conditions imposed upon him in 1616, he had done so inadvertently. The Pope insisted upon interrogation under the threat of torture. On 22 June 1633, he was found guilty of 'vehement suspicion of heresy' and forced to make a public abjuration, the proceedings being recorded in the Book of Decrees.

> I do not hold this opinion of Copernicus, and I have not held it after being ordered by injunction to abandon it. For the rest, here I am in your hands; do as you please.[14]

Galileo eventually returned to Florence where he remained under house arrest for the rest of his life – he died in Arcetri on 9 January 1642.

With indomitable spirit, Galileo set about writing his greatest work, *Discourses and Mathematical Demonstrations on Two New Sciences Pertaining to Mechanics and to Local Motion*, normally known simply as *Two New Sciences*, published in 1638. In this treatise, he brought together the understanding of the physical world which he had gained over a lifetime. The fundamental insights concern the second new science – the analysis of motion.

3.6 Galilean Relativity

The ideas expounded in *Two New Sciences* had been in Galileo's mind since 1608. One of them is what is now called *Galilean relativity*. Relativity is often thought of as something which was invented by Einstein in 1905, but that does not do justice to Galileo's great achievement. Suppose an experiment is carried out on the shore and then on a ship moving at a constant speed relative to the shore. If the effect of air resistance is neglected, is there any difference in the outcome of any experiment? Galileo answers firmly, 'No, there is not.'

The relativity of motion is illustrated by dropping an object from the top of a ship's mast (Fig. 3.8). If the ship is stationary, the object falls vertically downwards. Now suppose the ship is moving. If the object is dropped from the top of the mast, it again falls vertically downwards according to the observer on the ship. However, a stationary observer on the shore notes that, relative to the shore, the path is curved (Fig. 3.8(c)). The reason is that the ship moves at some speed v and so, relative to the shore, the object has two separate

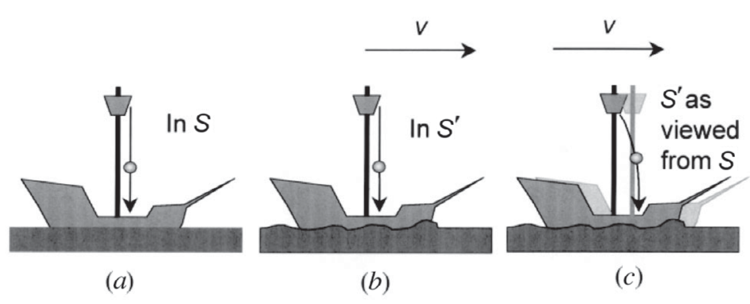


Fig. 3.8 (*a*) Dropping an object from the top of a mast in a ship, stationary in the frame of reference *S*. (*b*) Dropping an object from the top of a mast in a moving ship, viewed in the frame of reference *S′* of the ship. (*c*) Dropping an object from the top of a mast in a moving ship, as observed from the frame of reference *S*. The ship moves during the time the object falls.

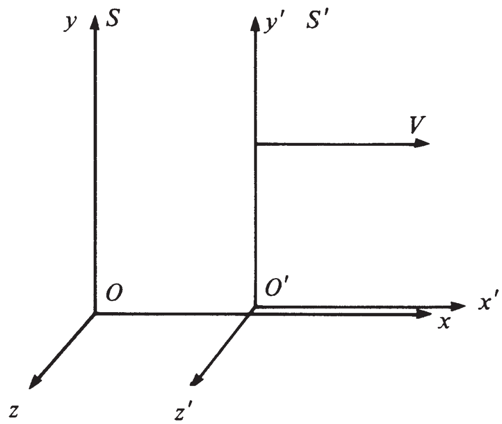


Fig. 3.9 Illustrating two Cartesian frames of reference moving at relative velocity v in the direction of the positive x-axis in 'standard configuration'.

components to its motion – vertical acceleration downwards due to gravity and uniform horizontal motion due to the motion of the ship.

This leads naturally to the concept of *frames of reference*. When the position of some object in three-dimensional space is measured, we can locate it by its coordinates in some rectangular coordinate system (Fig. 3.9). The point P has coordinates (x, y, z) in this stationary frame of reference S. Now, suppose the ship moves along the positive x-axis at some speed v. Then, we can set up another rectangular frame of reference S' on the ship. The coordinates of the point P in S', are (x', y', z'). It is now straightforward to relate the coordinates in these two frames of reference. If the object is stationary in the frame S, x is a constant but the value of x' changes as $x' = x - vt$, where t is the time, assuming the origins of the two frames of reference were coincident at $t = 0$. The values of y and y' remain the same in S and S', as do z and z'. Also, time is the same in the two frames of reference. We have derived a set of relations between the *coordinates* of objects in the frames S and S':

$$\begin{cases} x' = x - vt, \\ y' = y, \\ z' = z, \\ t' = t. \end{cases} \tag{3.3}$$

These are known as the *Galilean transformations* between the frames S and S'. Frames of reference which move at constant speed relative to one another are called *inertial frames of reference*. Galileo's great insight can be summarised by stating that *the laws of physics are the same in every inertial frame of reference*.

A corollary of this insight is the law of composition of velocities. I have used the word *speed* rather than *velocity* because I am adopting the modern usage that *speed* is a scalar quantity, whereas by *velocity* I mean the vector quantity $\boldsymbol{v}$ which has both magnitude and direction. Galileo was the first to establish the *law of composition of velocities* – if a body has components of velocity in two different directions, the motion of the body can be found by adding together the separate effects of these motions. This was how he demonstrated that the trajectories of cannon balls and missiles are parabolic (Fig. 3.5).

In *Two New Sciences* Galileo described his discoveries concerning the nature of constant acceleration, the motion of pendulums and free fall under gravity. Finally, he stated his *law of inertia*, which asserts that a body moves at a constant velocity unless some impulse or force causes it to change that velocity – notice that it is now velocity rather than speed which is constant because the direction of motion does not change in the absence of forces. This is often referred to as the *conservation of motion* – in the absence of forces, the separate components of the velocity remain unaltered. The word *inertia* is used here in the sense that it is a property of the body which resists change of motion. This law will become Newton's first law of motion. It can be appreciated why the Earth's motion caused Galileo no problems. Because of his understanding of Galilean relativity, he realised that the laws of physics would be the same whether the Earth were stationary or moving at a constant speed.

3.7 Reflections

We cannot leave this study without reflecting on the methodological and philosophical implications of the Galileo case. The Church made an error in condemning the new physics of Copernicus and Galileo. It was a further 350 years before Pope John Paul II admitted that an error had been made. In November 1979, on the occasion of the centenary of the birth of Albert Einstein, John Paul II stated that Galileo '. . . had to suffer a great deal – we cannot conceal the fact – at the hands of men and organisms of the Church.' He went on to assert that '. . . in this affair the agreements between religion and science are more numerous and above all more important than the incomprehensions which led to the bitter and painful conflict that continued in the course of the following centuries.'

For scientists, the central issue is the nature of scientific knowledge and the concept of truth in the physical sciences. Part of Cardinal Bellarmine's argument is correct. What Copernicus had achieved was a much more elegant and economical model for understanding the motions of the Sun, Moon and planets than the Ptolemaic picture, but in what sense was it the truth? If one were to put in enough effort, a Ptolemaic model of the Solar System could be created today, which would reproduce exactly the motions of the planets on the sky, but it would be of enormous complexity and provide little insight into the underlying physics. The value of the new model was not only that it provided a much simpler and coherent framework for understanding the observed motions of the celestial bodies, but also that, in the hands of Isaac Newton, it was to become the avenue for obtaining a very much deeper understanding of the laws of motion in general, leading to the unification of celestial physics, the laws of motion and the law of gravity. A scientifically satisfactory model has the capability not only of accounting economically for a large number of disparate observational and experimental phenomena, but also of being extendible to make quantitative predictions about apparently unrelated phenomena.

Notice that I use the word *model* in describing this process rather than asserting that it is the *truth*. Galileo's enormous achievement was to realise that the models which describe nature could be placed on a rigorous mathematical foundation. In perhaps his most famous remark, he stated in his treatise *Il Saggiatore* (The Assayer) of 1624:

> Philosophy is written in this very great book which always lies before our eyes (I mean the Universe), but one cannot understand it unless one first learns to understand the language and recognise the characters in which it is written. It is written in mathematical language and the characters are triangles, circles and other geometrical figures; without these means it is humanly impossible to understand a word of it; without these there is only clueless scrabbling around in a dark labyrinth.[15]

This is often abbreviated to the statement:

> The Book of Nature is written in mathematical characters.

This was the great achievement of the Galilean revolution. The apparently elementary facts established by Galileo required an extraordinary degree of imaginative abstraction. Matter does not obey the apparently simple laws of Galileo – there is always friction, experiments can only be carried out with limited accuracy and often give negative results. It needs deep insight and imagination to sweep away the unnecessary baggage and appreciate the basic simplicity of the way matter behaves. The modern approach to science is no more than the formalisation of the process begun by Galileo. It has been called the *hypothetico-deductive method* whereby hypotheses are made and consequences deduced logically from them. The model is acceptable so long as it does not run into serious conflict with the way matter actually behaves. But models are only valid within well-defined regions of parameter space. Professionals become very attached to them, and the remarks by Dirac and Douglas Gough quoted in Chapter 1 describe the need to be satisfied with approximate theories and the real pain experienced on being forced to give up a cherished prejudice.

It is rare nowadays for religious dogma to impede progress in the physical sciences. However, *scientific prejudice* and *dogma* are a common currency of the scientific debate. There is nothing particularly disturbing about this so long as we recognise what is going on. The prejudices provide the framework for carrying forward the debate and for suggesting experiments and calculations which can provide tests of the self-consistency of the ideas. We will find many examples throughout this book where the 'authorities' and 'received wisdom' were barriers to scientific progress. It takes a great deal of intellectual courage and perseverance to stand up against an overwhelming weight of conservative opinion. Through scientific patronage, scientific dogma can attain an authority to the exclusion of alternative approaches. One of the most disastrous examples was the Lysenko affair in the USSR shortly after the Second World War, when Communist political philosophy strongly impacted the biological sciences, resulting in a catastrophe for these studies in the Soviet Union.

Let me give two topical examples which we will pursue in more detail in Chapter 21. It is intriguing how the idea of *inflation* during the early stages of expansion of the Universe has almost attained the status of 'received dogma' in much of the cosmological community. There are good reasons why this idea should be taken seriously but there is no direct experimental evidence for the physics which could cause the inflationary expansion of the early Universe. Indeed, a common procedure is to work backwards and ‘derive’ the physics of inflation from the need to account for the features of the Universe we observe today. Then, theories of particle physics can be invoked to account for these requirements. There is the obvious danger of ending up with boot-strapped self-consistency without any independent experimental check of the theory. Maybe that is the best we can hope for, but some of us will maintain a healthy scepticism until there are more independent arguments to support these conjectures.

The same methodology has occurred in the theory of elementary particles with the development of string theory. The creation of self-consistent quantum field theories involving one-dimensional objects rather than point particles has been a quite remarkable achievement. The latest version of these theories involve the quantisation of gravity as an essential ingredient. Yet, these theories have not resulted in predictions which can yet be tested experimentally. Nonetheless, this is the area into which many of the most distinguished theorists have transferred all their efforts. It is an act of faith that this is the most promising way of tackling these problems, despite the fact that it might well prove to be very difficult, if not impossible, to test the theory by any experiment or observation in the foreseeable future.

Notes

1 The recent history of the Galileo affair is contained in *The Church and Galileo*, ed. E. McMullin (2006). Notre Dame, Indiana: University of Notre Dame Press. Chapter 13 ‘The Church’s Most Recent Attempt to Dispel the Galileo Myth’ by George V. Coyne, S.J. is a particularly illuminating essay.

2 Drake, S. (1990). *Galileo: Pioneer Scientist*, p. 63. Toronto: University of Toronto Press.

3 Note that if the sphere *rolled* down the slope without slipping, the acceleration would have been only $\frac{5}{7}g \sin\theta$, because the sphere acquires both linear *and* rotational energy as the gravitational potential energy becomes more negative.

4 Drake, S. (1990). *op. cit.*, p. 83.

5 Drake, S. (1990). *op. cit.*, p. 84.

6 Galilei, G. (1610). *Sidereus Nuncius*, Venice. See translation by A. van Helden (1989). *Sidereus Nuncius or The Sidereal Messenger*. Chicago: University of Chicago Press. Astonishingly, the observations and sketches were sent for publication almost before the ink was dry. The last observations were made on 2 March 1610 and the book published on 13 March of the same year.

7 Finocchiaro, M.A. (1989). *The Galileo Affair: A Documentary History*. Berkeley: University of California Press.

8 Finocchiaro, M.A. (1989). *op. cit.*, p. 24.

9 Finocchiaro, M.A. (1989). *op. cit.*, p. 24.

10 Gingerich, O. (1982). *Scientific American*, **247**, 118.

11 Finocchiaro, M.A. (1989). *op. cit.*, p. 28.

12 This word is often translated *hypothetically*.

13 Finocchiaro, M.A. (1989). *op. cit.*, p. 67.

14 Finocchiaro, M.A. (1989). *op. cit.*, p. 287.

15 Sharratt, M. (1994). *Galileo: Decisive Innovator*, p. 140. Cambridge: Cambridge University Press.

4 Newton and the Law of Gravity

4.1 Introduction

Richard Westphal's monumental biography *Never at Rest* was the product of a lifetime's study of Isaac Newton's life and work. In the preface, he writes:

> The more I have studied him, the more Newton has receded from me. It has been my privilege at various times to know a number of brilliant men, men whom I acknowledge without hesitation to be my intellectual superiors. I have never, however, met one against whom I was unwilling to measure myself so that it seemed reasonable to say that I was half as able as the person in question, or a third or a fourth, but in every case a finite fraction. The end result of my study of Newton has served to convince me that with him there is no measure. He has become for me wholly other, one of the tiny handful of supreme geniuses who have shaped the categories of human intellect, a man not finally reducible to the criteria by which we comprehend our fellow beings.[1]

In the next paragraph, he writes:

> Had I known, when in youthful self-confidence I committed myself to the task, that I would end up in similar self-doubt, surely I would never have set out.[1]

Newton's impact upon science is so all pervasive that it is worthwhile filling in some of the background to his character and extraordinary achievements. The chronology which follows is that adopted in the Introduction to the book *Let Newton Be*.[2]

4.2 Lincolnshire 1642–61

Newton was born in the hamlet of Woolsthorpe near Grantham on Christmas day 1642 (old style calendar). Newton's father was a successful farmer, who died three months before Newton's birth. When Isaac Newton was three years old, his mother married the Reverend Barnabas Smith and moved into his house, leaving Isaac behind to be brought up by his grandmother. Isaac hated his step-father as is revealed by this entry in the list of sins he had committed up to the age of 19:

> Threatening my father and mother Smith to burn them and the house over them.[3]

It has been argued that this separation from his mother at an early age was a cause of his 'suspicious, neurotic, tortured personality'.[4] He was sent to the Free Grammar School at Grantham where he lodged with an apothecary. It was probably in these lodgings that he was introduced to chemistry and alchemy, which remained a lifelong passion. Newton himself claimed that he invented toys and did experiments as a teenager. He is reported to have constructed a model of a mill powered by a mouse, clocks, ‘lanthorns’ and fiery kites, which he flew to frighten the neighbours. He was already a loner who did not get on particularly well with his schoolmates. Following his school education, it was decided that he should return to the Free Grammar School at Grantham to prepare for entry to his uncle’s old college, Trinity College, Cambridge.

4.3 Cambridge 1661–65

To begin with, Newton was a ‘subsidizer’, meaning that he earned his keep by serving the fellows and rich students at Trinity College. He took courses in Aristotelian philosophy, logic, ethics and rhetoric. His notebooks record that he was reading other subjects privately, the works of Thomas Hobbes, Henry More and René Descartes, as well as those of Kepler and Galileo. It became apparent to him that he was weak in mathematics and so he started to work ‘ferociously’ at the subject. By the end of this period, he had mastered whole areas of mathematics. He also continued to carry out experiments in a wide range of different topics. In 1664, he was elected a scholar of Trinity College, ensuring him a further four years of study without domestic duties. In 1665, he took his BA degree and then the effects of the Great Plague began to spread north to Cambridge. The University was closed and Newton returned to his home at Woolsthorpe.

4.4 Lincolnshire 1665–67

During the next two years, Newton’s burst of creative scientific activity must be one of the most remarkable ever recorded. Over 50 years later, Newton wrote:

> In the beginning of 1665, I found the method of approximating series and the rule for reducing any dignity (power) of any binomial into such a series. The same year in May, I found the method of tangents of Gregory and Slusius and in November had the direct method of fluxions and the next year in January had the theory of colours and in May following I had entrance into the inverse method of fluxions. And in the same year I began to think of gravity extending to the orbit of the Moon and (having found out how to estimate the force with which a globe revolving within a sphere presses the surface of the sphere) from Kepler’s rule of the periodical times of the planets being in sesquialternate proportion[5] of their distance from the centres of their Orbs, I deduced that the forces which keep the planets in their Orbs must [be] reciprocally as the squares of their distances from the centres about which they revolve: and thereby compared the force requisite to keep the Moon in her orb with the force of gravity at the surface of the Earth, and found them answer pretty nearly. All this was in the two plague years 1665–1666. For in those days I was in the prime of my age for invention and minded mathematics and philosophy more than at any time since.[6]

It is a quite astonishing list – Newton had laid the foundations of three quite separate areas. In mathematics, he discovered the *binomial theorem* and the *differential* and *integral calculus*. In *optics*, he discovered the decomposition of light into its primary colours. He began his unification of *celestial mechanics* with the *theory of gravity*, which was to lead ultimately to his laws of motion and the theory of universal gravitation. Let us look into his achievements in the fields of optics and of the law of gravity in more detail.

4.4.1 Optics

Newton was a skilled experimenter and undertook a number of optical experiments using lenses and prisms while he was at Trinity College and at Woolsthorpe. Many investigators had noted that, when white light is passed through a prism, all the colours of the rainbow are produced. Descartes was the proponent of the generally accepted view that, when light passes through a prism, it is modified by the material of the prism and so becomes coloured.

In 1666, Newton, in his own words, 'applied myself to the grinding of Optick glasses of other figures than spherical', and among these he produced a triangular glass prism 'to try therewith the celebrated *Phaenomena of Colours*'. Sunlight was passed through a tiny hole onto the prism and, to Newton's surprise, the coloured light produced by the prism was of oblong form. This experiment led him to carry out his famous *experimentum crucis*, illustrated in Figs. 4.1(*a*) and (*b*), the second picture showing Newton's own drawing of the experiment.

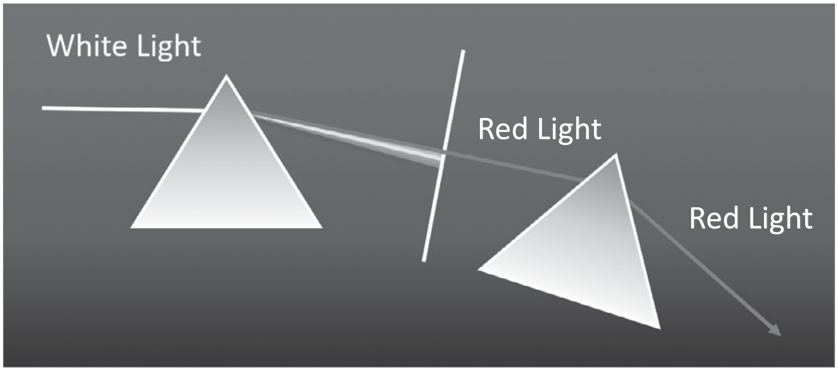
In the experiment, the coloured spectrum produced by the first prism was projected onto a plane board with a small hole to allow only one of the colours to pass through. A second prism was placed in the path of this second light beam and he discovered that there was no further decomposition of the light into more colours. The experiment was repeated for all the colours of the spectrum and consistently he found that there was no further splitting of the colours, contrary to the expectations of Descartes' theory. Newton concluded:

> Light itself is a Heterogeneous mixture of differently refrangible rays.[7]

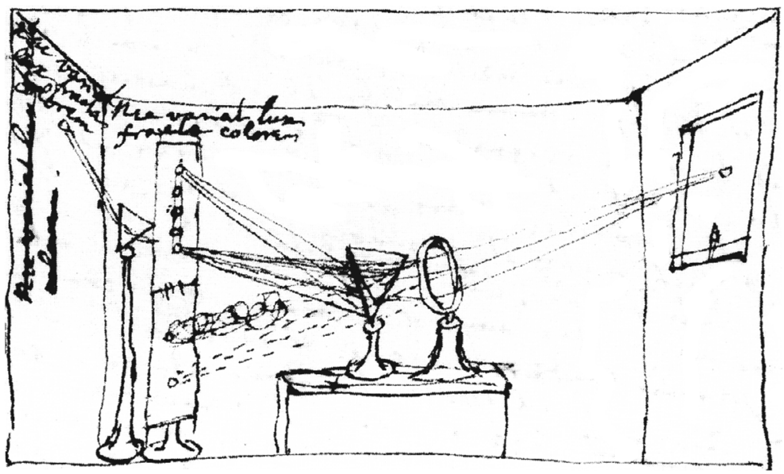
The word *refrangible* is what we would now call *refracted*. Newton had established that white light is a superposition of all the colours of the rainbow and that, when white light is passed through a prism, the different coloured rays are refracted by different amounts.

This work was not presented in public until 1672 in a paper to the Royal Society – Newton immediately found himself in the thick of a hot dispute. In addition to describing his new results, he regarded them as evidence in favour of his view that light can be considered to be made up of 'corpuscles', that is, particles which travel from the source of light to the eye. Newton did not have a satisfactory theory of 'light corpuscles', or why they should be refracted differently by material media. His problems were exacerbated by the fact that the experimentum crucis was difficult to reproduce. The ensuing unpleasant debate with Huygens and Hooke obscured the key result that white light is composed of all the colours of the rainbow.

This work led him to another important conclusion – it is not possible to build a large refracting telescope of the type pioneeered by Galileo, because white light is split up into



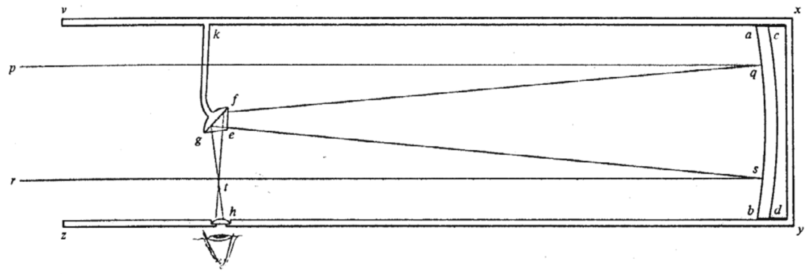
(*a*)



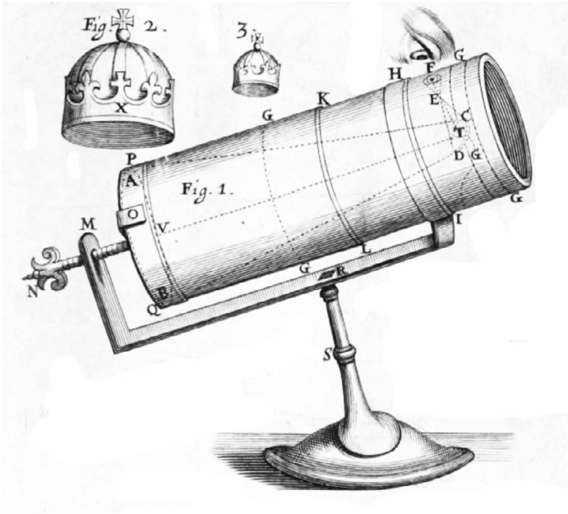
(*b*)

Fig. 4.1 (*a*) A schematic diagram showing the decomposition of white light into the colours of the spectrum. On passing through the second prism, the colours are not split up into any further colours. (*b*) Newton's sketch of his *experimentum crucis*. (From C. Hakfoort, 1988. *Let Newton Be: a New Perspective on his Life and Works*, eds. J. Fauvel, R. Flood, M. Shortland and R. Wilson, p. 87. Oxford: Oxford University Press.)

its primary colours by a refracting lens and so different colours are focussed at different positions on the optical axis of the telescope, the phenomenon known as *chromatic aberration*. Because of this problem, Newton designed and built a new type of *reflecting telescope* which would not suffer from chromatic aberration. He ground the mirrors himself and constructed the mount for the telescope and mirrors. This configuration is now called a 'Newtonian' design and is shown in Fig. 4.2(*a*). Figure 4.2(*b*) shows a picture of Newton's telescope as well as a comparison of the magnifying power of Newton's telescope with that of a refracting telescope, indicating the superiority of Newton's design. Nowadays, all large optical astronomical telescopes are reflectors, the descendants of Newton's invention.



(*a*)



(*b*)

Fig. 4.2 (*a*) The optical layout of Newton's reflecting telescope. *abcd* is the mirror made of speculum metal, *efg* is a prism and h a lens. (From I.B. Cohen, 1970. *Dictionary of Scientific Biography*, Vol. 11, p. 57. New York: Charles Scribner's Sons.) (*b*) Newton's reflecting telescope, as illustrated in the *Proceedings of the Royal Society*, Vol. 7, Plate 1 of issue 81, 1672. The two crowns show the improvement in magnification of Newton's telescope (Fig. 2) over the refracting telescope (Fig. 3), such as that used by Galileo.

These were great experimental achievements, but Newton had a pathological dislike of writing up his work for publication. The systematic presentation of his optical work was only published in 1704 as his *Opticks*, long after his discoveries and inventions were made.

4.4.2 The Law of Gravity

Newton's most famous achievement of his years at Woolsthorpe was the discovery of the law of gravity. The calculations carried out in 1665–66 were only the beginning of the story, but they contained the essence of the theory which was to appear in its full glory in his *Principia Mathematica* of 1687. As Newton himself recounted, he was aware of Kepler's third law of planetary motion, which was deeply buried in Kepler's *Harmony of*

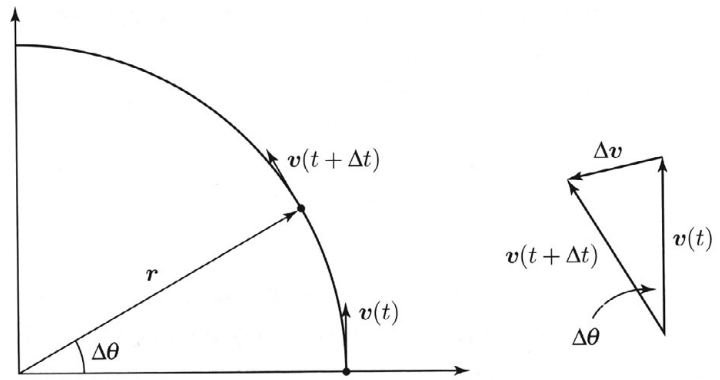


Fig. 4.3 A vector diagram illustrating the origin of centripetal acceleration.

the World. This volume was in Trinity College library and it is likely that Isaac Barrow, the Lucasian Professor of Mathematics, drew Newton's attention to it.

In Newton's own words:

> the notion of gravitation [came to my mind] as I sat in contemplative mood [and] was occasioned by the fall of an apple.[8]

Newton asked whether or not the force of gravity, which makes the apple fall to the ground, is the same force which holds the Moon in its orbit about the Earth and the planets in their orbits about the Sun. To answer this question, he needed to know how the force of gravity varies with distance. He derived this relation from Kepler's third law by a simple argument. First, he needed an expression for *centripetal acceleration*, which he rederived for himself. Nowadays, this formula can be derived by a simple geometrical argument using vectors.

Figure 4.3 shows the velocity vectors $\boldsymbol{v}$ of a particle moving at constant speed $|\boldsymbol{v}|$ in a circle with radius $|\boldsymbol{r}|$ at times t and $t + \Delta t$. In the time Δt, the angle subtended by the radius vector $\boldsymbol{r}$ changes by $\Delta\theta$. The change in the velocity vector in Δt is given by the vector triangle shown on the right. As $\Delta t \to 0$, the vector $\mathrm{d}\boldsymbol{v}$ continues to point towards the centre of the circle and its magnitude is $|\boldsymbol{v}|\Delta\theta$. Therefore, the centipetal acceleration is

$$|\boldsymbol{a}| = \frac{|\Delta \boldsymbol{v}|}{\Delta t} = \frac{|\boldsymbol{v}|\Delta\theta}{\Delta t} = |\boldsymbol{v}|\omega = \frac{|\boldsymbol{v}|^2}{|\boldsymbol{r}|}, \tag{4.1}$$

where ω is the (constant) angular velocity of the particle. This is the famous expression for the *centripetal acceleration* a of an object moving in a circle of radius r at speed v.

In 1665, Newton was well aware of the fact that the orbits of the planets are actually ellipses, and not circles, but he assumed that he could use Kepler's third law for circles, since the ellipticities of the orbits of the planets are generally small.

He also knew that accelerations are proportional to the forces which cause them and so the force f which keeps the planets in their circular orbits must be proportional to their centripetal accelerations, v^2/r, that is,

$$f \propto \frac{v^2}{r}. \tag{4.2}$$

But, Kepler's third law states that the period T of the planet's orbit is proportional to $r^{3/2}$. The speed of a planet in its orbit is $v = 2\pi r/T$ and therefore

$$v \propto \frac{1}{r^{1/2}}. \tag{4.3}$$

Now, we can substitute for v from (4.3) into (4.2) and find

$$f \propto \frac{1}{r^2}. \tag{4.4}$$

This is the primitive form of Newton's *inverse square law of gravity* – the force of gravity decreases as the inverse square of the distance from the body.

This was the key result he was looking for – if gravity were truly universal, then the same force which causes apples to fall to the ground should be the same force which keeps the Moon in its orbit about the Earth. The only difference is that the acceleration of the apple should be greater than that of the Moon because the Moon is 60 times further away from the centre of the Earth than the apple. According to the universal theory of gravity, the centripetal acceleration of the Moon should be only $1/60^2 = 1/3600$ times the acceleration due to gravity at the surface of the Earth. Newton had all the data he needed to make this comparison. Using modern numbers, the acceleration due to gravity at the surface of the Earth is on average 9.80665 m s^{-2}. The period of revolution of the Moon about the Earth is 27.32 days and its mean distance from the Earth is $r = 384\,408\,000$ m. Putting these together, the mean speed of the Moon is $v = 1023$ m s^{-1} and its centripetal acceleration $v^2/r = 2.72 \times 10^{-3}$ m s^{-2}. The ratio of this acceleration to the local acceleration due to gravity is $9.80665/2.72 \times 10^{-3} = 3600$, exactly as predicted.

Newton's calculations did not give quite such good agreement, but were sufficiently close to persuade him that the same force which holds the Moon in its orbit about the Earth and the planets in their orbits about the Sun is exactly the same force which causes the acceleration due to gravity on Earth. From this result, the general formula for the gravitational attraction between any two bodies of masses M_1 and M_2 follows immediately:

$$f = \frac{GM_1M_2}{r^2}, \tag{4.5}$$

where G is the gravitational constant and r is the distance between the bodies. The force acts along the line joining the two bodies and is always an attractive force.

This work was not published, because there were a number of steps in the calculation which needed further elaboration.

(1) Kepler had shown that the orbits of the planets are ellipses and not circles – how did that affect the calculation?
(2) Newton was uncertain about the influence of the other bodies in the Solar System upon each others' orbits.
(3) He was unable to explain the details of the Moon's motion about the Earth, which we now know is influenced by the fact that the Earth is not spherical.
(4) Probably most important of all, there is a key assumption made in the calculation that all the mass of the Earth can be located at its centre in order to work out the acceleration due to gravity at its surface and its influence upon the Moon. The same assumption was

made for all the bodies in the Solar System. In his calculations of 1665–66, he regarded this step as an approximation. He was uncertain about its validity for objects close to the surface of the Earth.

Newton laid this work aside until 1679.

4.5 Cambridge 1667–96

The University reopened in 1667 and Newton returned to Trinity College in the summer of that year. He became a fellow of the college in the autumn of 1667 and two years later, at the age of 26, he was elected Lucasian Professor of Mathematics, a position which he held for the next 32 years. As Lucasian Professor, Newton's duties were not heavy. He was required to give at least one lecture per week in every term and to deposit his lectures, properly written up, in the University Library. In the period 1670 to 1672, he deposited lectures on optics, during 1673–83 lectures on arithmetic and algebra and in 1684–85 most of book I of what was to become the *Principia Mathematica*. In 1687, the lectures were entitled 'The System of the World' and this became part III of the *Principia*. There seems to be no record of his lectures for 1686, nor from 1688 until he left Cambridge in 1696.

His lectures were not particularly successful. Humphrey Newton, his assistant from 1685 to 1690, who was no relation, wrote of Newton during the years he was preparing the *Principia*:

> He seldom left his chamber except at term time, when he read (his lectures) in the (Old S)chools as being Lucasian Professor, where so few went to hear him, and fewer that understood him, that of times he did, in a manner, for want of hearers, read to the walls. . . . (when he lectured he) usually stayed about half an hour; when he had no auditors, he commonly returned in a fourth part of that time or less.[9]

The first of Newton's publications appeared during this period. In late 1668, Nicholas Mercator published his book *Logarithmotechnica* in which he described some of the techniques for the analysis of infinite series which Newton had already worked out in greater generality. Newton set about writing up his mathematical work so that his priority could be established. He wrote hastily the work known as *De Analysi* (*On Analysis*) which Isaac Barrow was allowed to show to a London mathematician, Mr Collins. Newton insisted, however, that the work remain anonymous. The results were communicated by letter among British and continental mathematicians and so the work became widely known.

Newton incorporated the most important parts of *De Analysi* into another manuscript, *Methodus Fluxionum et Serierum Infinitarum* (*Method of Fluxions and Infinite Series*), which was not published until 1711 by William Jones. The manuscript was, however, read by Leibniz in October 1676 on a visit to London. Although Leibniz scholars have established that he only copied out the sections on infinite series, this incident was the source of the later bitter accusations that Leibniz had plagiarised Newton's discovery of the differential and integral calculus.

Newton was prompted to return to his researches on the laws of motion and gravity in 1679 by an interchange of letters with Robert Hooke. Hooke challenged Newton to work out the curve followed by a particle falling in an inverse square field of force, the law Newton had derived 14 years earlier. This stimulated Newton to derive two crucial results. As noted by Chandrasekhar in his remarkable commentary on the *Principia*,[10]

> Newton's interest in dynamics was revived sufficiently for him to realise for the first time the real meaning of Kepler's laws of areas. And as he wrote, 'I found now that whatsoever was the law of force which kept the Planets in their Orbs, the area described by a radius drawn from them to the Sun would be proportional to the times in which they were described'; and he proved the two propositions that
>
> > all bodies circulating about a centre sweep out areas proportional to the time
>
> and that
>
> > a body revolving in an ellipse . . . the law of attraction directed to a focus of the ellipse . . . is inversely as the square of the distance.

The first remarkable discovery, that Kepler's second law, the law of areas, is correct whatever the nature of the force law so long as it is a central force, is now recognised as a consequence of the law of conservation of angular momentum. Chandrasekhar noted: 'The resurrection of Kepler's law of areas in 1679 was a triumphant breakthrough from which the *Principia* was later to flow'.[11] At this point, another dispute broke out with Hooke who claimed that he was the originator of the inverse square law of gravity. Newton violently rejected this claim and would not communicate any of the results of his calculations to Hooke, or to anyone else.

In 1684, Edmond Halley travelled to Cambridge to ask Newton precisely the same question which had been posed by Hooke. Newton's immediate response was that the orbit is an ellipse, but the proof could not be found among his papers. In November 1684, Newton sent the proof to Halley. Halley returned to Cambridge where he saw an incomplete manuscript of Newton's entitled *De Motu Corporum in Gyrum*, or *On the Motion of Revolving Bodies*, which was eventually to be transformed into the first part of the *Principia*. With a great deal of persuasion, Newton agreed to set about systematising all his researches on motion, mechanics, dynamics and gravity.

Only in 1685, when he was working at white heat on the preparation of what was to become his great treatise, the *Philosophiae Naturalis Principia Mathematica*, or *Principia* for short, did he demonstrate that, for spherically symmetric bodies, the gravitational attraction can be found exactly by placing all its mass at its centre. In the words of J.W.L. Glaisher, on the occasion of the bicentenary of the publication of the *Principia*,

> No sooner had Newton proved this superb theorem – and we know from his own words that he had no expectation of so beautiful a result till it emerged from his mathematical investigation – then all the mechanism of the universe at once lay spread before him. . . . We can imagine the effect of this sudden transition from approximation to exactitude in stimulating Newton's mind to still greater effort.[12]

Nowadays, this result can be proved in a few lines using Gauss's theorem in vector calculus as applied to the inverse square law of gravity, but these techniques were not available to Newton. This was a crucial step in the development of the *Principia*.

Humphrey Newton gives us a picture of Newton during the writing of the *Principia*:

> (He) ate sparing (and often) forgot to eat at all. (He rarely dined) in the hall, except on some public days (when he would appear) with shoes down at the heel, stockings untied, surplice on, and his head scarcely combed. (He) seldom went to the Chapel (but very often) went to St. Mary's Church, especially in the forenoon.[13]

The *Principia* is one of the greatest intellectual achievements of all time. The theory of statics, mechanics and dynamics, as well as the law of gravity, are developed entirely through mathematical relations and mechanistic interpretations of the physical origins of the forces are excluded. In the very beginning, what we now call *Newton's Laws of Motion* are set out in their definitive form – we will return to these in more detail in Chapter 8. Despite the fact that Newton had developed his own version of the integral and differential calculus, the *Principia* is written entirely in terms of geometrical arguments. Often these are difficult to follow for the modern reader, partly because of Newton's somewhat convoluted style and partly because the geometrical arguments used are no longer familiar to physicists. Chandrasekhar's remarkable reconstruction of the arguments in modern geometrical terms gives some impression of the methods which Newton must have used. As an example of the economy of expression used by Newton in the *Principia*, Fig. 4.4 shows a translation of the proof of the elliptical orbits under the influence of an inverse square field of force. Comparison with Chandrasekhar's much lengthier derivation shows how terse Newton's geometric proofs were.

4.6 Newton the Alchemist

Throughout his time at Cambridge, Newton maintained a profound interest in chemistry, alchemy and in the interpretation of ancient and biblical texts. These aspects of his work were of the greatest significance for Newton. He studied chemistry and alchemy with the same seriousness he devoted to mathematics and physics. Whereas his great contributions to the latter disciplines became public knowledge, his alchemy remained very private, his papers containing over a million words on alchemical topics. From the late 1660s to the mid 1690s, he carried out extensive chemical experiments in his laboratory at Trinity College. Figure 4.5 is a contemporary engraving of Trinity College; Newton's rooms were on the first floor to the right of the gatehouse. It has been suggested that the shed which can be seen leaning against the chapel outside the College was Newton's laboratory. He taught himself all the basic alchemical operations, including the construction of his own furnaces. His assistant Humphrey Newton wrote:

> About six weeks at spring and six at the fall, the fire in the laboratory scarcely went out. . . . what his aim might be, I was not able to penetrate into. . . . Nothing extraordinary, as I can remember, happened in making his experiments, which if there did . . . I could not in the least discern it.[14]

The two Newtons kept the furnaces going continuously for six weeks at a time, taking turns to tend them overnight.

PROPOSITION XI. PROBLEM VI

If a body revolves in an ellipse; it is required to find the law of the centripetal force tending to the focus of the ellipse.

Let S be the focus of the ellipse. Draw SP cutting the diameter DK of the ellipse in E, and the ordinate Qv in x; and complete the parallelogram $QxPR$. It is evident that EP is equal to the greater semiaxis AC: for drawing HI from the other focus H of the ellipse parallel to EC, because CS, CH are equal, ES, EI will be also equal; so that EP is the half-sum of PS, PI that is (because of the parallels HI, PR, and the equal angles IPR, HPZ), of PS, PH, which taken together are equal to the whole axis 2AC. Draw QT perpendicular to SP, and putting L for the principal latus rectum of the ellipse (or for $\frac{2BC^2}{AC}$), we shall have

$$L\cdot QR : L\cdot Pv = QR : Pv = PE : PC = AC : PC,$$

$$\text{also, } L\cdot Pv : Gv\cdot Pv = L : Gv, \text{ and, } Gv\cdot Pv : Qv^2 = PC^2 : CD^2.$$

By Cor. II, Lem. VII, when the points P and Q coincide, $Qv^2 = Qx^2$, and Qx^2 or $Qv^2 : QT^2 = EP^2 : PF^2 = CA^2 : PF^2$, and (by Lem. XII) $= CD^2 : CB^2$. Multiplying together corresponding terms of the four proportions, and simplifying, we shall have

$$L\cdot QR : QT^2 = AC\cdot L\cdot PC^2\cdot CD^2 : PC\cdot Gv : CD^2\cdot CB^2 = 2PC : Gv,$$

since $AC\cdot L = 2BC^2$. But the points Q and P coinciding, 2PC and Gv are equal. And therefore the quantities $L\cdot QR$ and QT^2, proportional to these, will be also equal. Let those equals be multiplied by $\frac{SP^2}{QR}$, and $L\cdot SP^2$ will become equal to $\frac{SP^2\cdot QT^2}{QR}$. And therefore (by Cor. I and V, Prop. VI) the centripetal force is inversely as $L\cdot SP^2$, that is, inversely as the square of the distance SP. Q.E.I.

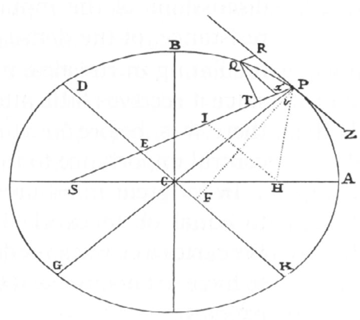


Fig. 4.4 A translation from the Latin of Newton's proof that the orbits of the planets are ellipses under the influence of an inverse square law field of gravity. (From J. Roche, 1988. *Let Newton Be: a New Perspective on his Life and Works*, eds. J. Fauvel, R. Flood, M. Shortland and R. Wilson, p. 55. Oxford: Oxford University Press.)

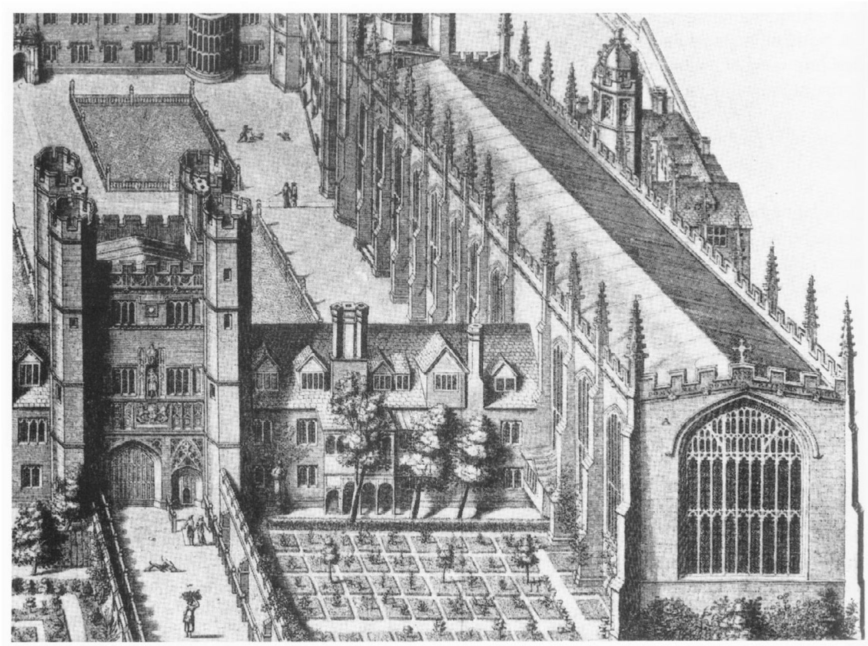


Fig. 4.5 A seventeenth century engraving of the Great Gate and Chapel of Trinity College. The shed seen in the bay of the chapel closest to the wall of Great Court may have been Newton's chemical laboratory. (From J. Golinski, 1988. *Let Newton Be: a New Perspective on his Life and Works*, eds. J. Fauvel, R. Flood, M. Shortland, and R. Wilson, p. 152. Oxford: Oxford University Press.)

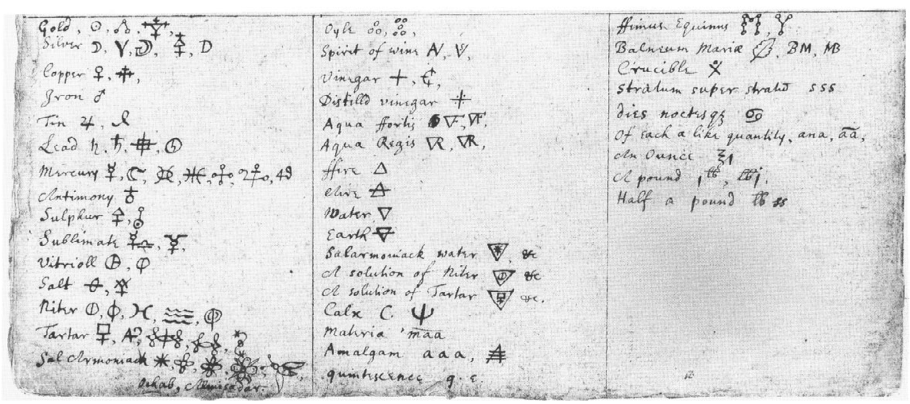


Fig. 4.6 Newton's list of alchemical symbols. (From J. Golinski, 1988. *Let Newton Be: a New Perspective on his Life and Works*, eds. J. Fauvel, R. Flood, M. Shortland and R. Wilson, p. 155. Oxford: Oxford University Press.)

From the late 1660s to the mid 1690s, Newton devoted a huge effort to systematising everything he had read on chemistry and alchemy (Fig. 4.6). His first attempt to put some order into his understanding was included in his *Chemical Dictionary* of the late 1660s, followed over the next 25 years by successive versions of his *Index Chemicus*. According to Golinski,[15] the final index cites more than 100 authors and 150 works. There are over

5000 page references under 900 separate headings. In addition, he devoted a huge effort to deciphering obscure allegorical descriptions of alchemical processes. It was part of the mystique of the alchemists that the fundamental truths should not be made known to the unworthy or the vulgar. The ultimate example of this was his belief that the biblical account of the creation was actually an allegorical description of alchemical processes. In a manuscript note of the 1680s, Newton noted:

> Just as the world was created from dark chaos through the bringing forth of the light and through the separation of the airy firmament and of the waters from the Earth, so our work brings forth the beginning out of black chaos and its first matter through the separation of the elements and the illumination of matter.

What Newton was trying to achieve is illuminated by a manuscript dating from the 1670s entitled *Of Nature's obvious laws and processes in vegetation*. To quote Golinski,

> Newton distinguished 'vulgar chemistry' from a more sublime interest in the processes of vegetation. . . . 'Nature's actions are either vegetable or purely mechanical' he wrote. The imitation of mechanical changes in nature would be common, or vulgar, chemistry, whereas the art of inducing vegetation was 'a more subtle, secret and noble way of working.' [16]

Newton's objective was no less than to isolate the essence of what gave rise to growth and life itself. Such a discovery would make man like God and so had to be kept a secret from the vulgar masses.

4.7 The Interpretation of Ancient Texts and the Scriptures

Newton devoted almost as much effort to interpreting the works of the ancients and biblical texts as he did to his chemical and alchemical studies. He convinced himself that all his great discoveries had in fact been known to the ancient Greek philosophers. In 1692, Nicholas Fatio de Duillier, Newton's protegé during the period 1689–93, wrote to Christiaan Huygens that Newton had discovered that all the chief propositions of the *Principia* had been known to Pythagoras and Plato, but that they had turned these discoveries into a 'great mystery'.

Newton believed that the ancients had kept this knowledge secret for the same reason he kept his alchemical work secret. These truths were of such great significance that they had to be preserved for those who could truly appreciate them. For this reason, the Greeks had to disguise their deep understandings in a coded language of symbolism, which only initiates such as Newton could penetrate.

Piyo Rattansi suggests that Newton's objective was to give his great discoveries an ancient and honorable lineage.[17] The *Principia Mathematica* had been well received on the continent as a text in mathematics, but not as physics. The continental physicists disliked the introduction of the 'unintelligible' force of attraction to account for the law of gravity. To them the force of gravity was a mystic force because they could not envisage any physical mechanism for causing gravitational attraction – all other physical forces could be explained by actions involving concrete bits of matter but what was it that caused gravity to act across the Solar System? The continental scientists interpreted the law of gravity as being the reintroduction of ‘occult causes’ which the whole of the revolution in science had been bent upon eliminating. Despite their objections, Newton’s *action at a distance* was to take its place as part of the foundation of classical physics. Now, Newton was claiming as ancient an authority for his discoveries as any branch of knowledge.

In addition, throughout his lifetime, he had a deep interest in the scriptures and their interpretation. His prime interest centred upon the Near East and how a proper interpretation of the texts could be shown to predict retrospectively all the significant events of history. His aim was to show that the seat of all knowledge began in Israel and from there it travelled to Mesopotamia and Egypt. The results of these researches were published in two books, *The Chronology of the Ancient Kingdoms amended* and *Observations upon the Prophecies of Daniel and the Apocalypse of St. John*. The latter work runs for 323 pages and was published in twelve editions between 1733 and 1922.

Rattansi puts Newton’s studies of the biblical texts in another revealing light:

> His biblical work was intended to vindicate the authority of the Bible against those Catholics who had tried to show that only by supplementing it with the tradition of the universal church could it be made authoritative. It served at the same time as a weapon against the free-thinkers who appealed to a purely ‘natural’ religion and thereby did away with the unique revelation enshrined in Christian religion. It did so by demonstrating that history was continually shaped according to providential design which could be shown, only after its fulfilment, to have been prefigured in the divine word.[18]

Later, Rattansi writes:

> Newton followed the Protestant interpretation of Daniel’s visions and made the Roman church of later times the kingdom of the Antichrist which would be overthrown before the final victory of Christ’s kingdom.[19]

Newton held the view that the scriptures had been deliberately corrupted in the fourth and fifth centuries. This view was at odds with the tenets of the Church of England and, as a fellow of Trinity College, he was expected to be a member of the Church. A major crisis was narrowly avoided by the arrival of a royal decree, exempting the Lucasian Professor from the necessity of taking holy orders.

In 1693, Newton had a nervous breakdown. After his recovery, he lost interest in academic studies and, in 1696, left Cambridge to take up the post of Warden of the Mint in London.

4.8 London 1696–1727

By the time Newton took up his position as Warden of the Mint, he was already recognised as the greatest living English scientist. Although the position was normally recognised as a sinecure, he devoted all his energies to the recoinage which was needed at that time to stabilise the currency. He was also responsible for prosecuting forgers of the coin of the realm, an offence which carried the death penalty. Apparently, Newton carried out this unpleasant aspect of his responsibility 'with grisly assiduity'. He was an effective manager and administrator and in 1700 was appointed Master of the Mint.

In 1703, following the death of Hooke, Newton was elected President of the Royal Society. He now held positions of great power which he used to further his own interests. Edmond Halley entered the service of the Mint in 1696 and in 1707 David Brewster was appointed as general supervisor of the conversion of the Scottish coinage into British. Newton had secured the Savillian chair of Astronomy at Oxford for David Gregory in 1692 as well as the Savillian Chair of Geometry for Halley in the early 1700s. He also ensured that William Whiston was elected to his own Lucasian Chair in Cambridge from which he had eventually resigned in 1703.

In his old age, he did not mellow. An acrimonious quarrel arose with Flamsteed, the first Astronomer Royal, who had made very accurate observations of the Moon which Newton wished to use in his analysis of its motion. Newton became impatient to use these observations before they were completed to Flamsteed's satisfaction. Newton and Halley took the point of view that, since Flamsteed was a public servant, the observations were national property and they eventually succeeded, not only in obtaining the incomplete set of observations, but also of publishing them in an unauthorised version without Flamsteed's consent in 1712. Flamsteed managed to recover about 300 of the unauthorised copies and had the pleasure of burning this spurious edition of his monumental work. He later publishing his own definitive set of observations in his *Historia Coelestis Britannica*.

The infamous dispute with Leibniz was equally nasty. It was fomented by Fatio de Duillier who brought the initial charge of plagiarism against Leibniz. Leibniz appealed to the Royal Society to set up an independent panel to pass judgement on the question of priority. Newton appointed a committee to look into the matter, but then wrote the committee's report himself in his *Commercium epistolicum*. He wrote it as if its conclusions were impartial findings which came down in Newton's favour. But he did not stop there. A review of the report was published in the *Philosophical Transactions* of the Royal Society which was also written by Newton anonymously. The story is unpleasant, to say the least, particularly since Leibniz had unquestionably made original and lasting contributions to the development of the differential and integral calculus. Indeed, the system of notation which is now used universally is Leibniz's and not Newton's.

Newton died on 20 March 1727 at the age of 85 and was buried in Westminster Abbey.

Appendix to Chapter 4: Notes on Conic Sections and Central Orbits

It is useful to recall some basic aspects of the geometry and algebra associated with conic sections and central forces.

A.1 Equations for Conic Sections

The geometric definition of conic sections is that they are the curves generated in a plane by the requirement that the ratio of the perpendicular distance of any point on the curve

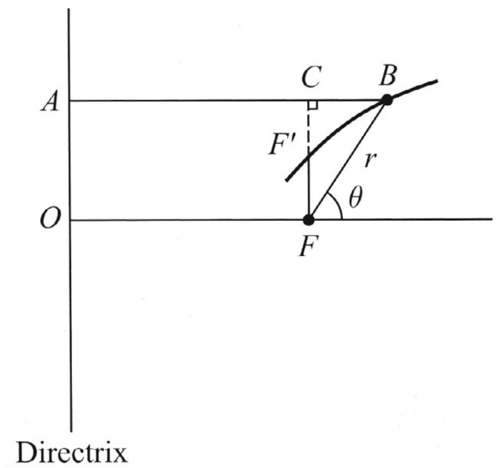


Fig. A4.1 Illustrating the geometrical construction for conic sections.

from a fixed straight line in the plane to the distance of that point on the curve from a fixed point be a constant. The fixed straight line is called the *directrix* and the fixed point F the *focus* (Fig. A4.1). This requirement can be written

$$\frac{AB}{BF} = \frac{AC + CB}{BF} = \text{constant},$$

that is,

$$\frac{AC + r\cos\theta}{r} = \text{constant}, \tag{A4.1}$$

where r and θ are polar coordinates with respect to F, the focus. Now AC and the 'constant' are independent fixed quantities and so, writing $AC = \lambda/e$ and constant $= e^{-1}$, where λ and e are constants, we can rewrite (A4.1):

$$\frac{\lambda}{r} = 1 - e\cos\theta. \tag{A4.2}$$

We find immediately the meaning of λ. When $\theta = \pi/2$ and $3\pi/2$, $r = \lambda$, that is, it is the distance FF' in Fig. A4.1. Notice that the curve is symmetric with respect to the line $\theta = 0$. λ is known as the *semi-latus rectum*. The curves generated for different values of e are shown in Fig. A4.2. If $e < 1$, we obtain an ellipse, if $e = 1$, a parabola and if $e > 1$, a hyperbola. Notice that two hyperbolae are generated in the case $e > 1$, the foci being referred to as the inner focus F_1 and the outer focus F_2 for the hyperbola to the right-hand side of the directrix.

Equation (A4.2) can be written in different ways. Suppose we choose a Cartesian coordinate system with origin at the centre of the ellipse. In the polar coordinate system of (A4.2), the ellipse intersects the x-axis at $\cos\theta = \pm 1$, that is, at radii $r = \lambda/(1 + e)$ and $r = \lambda/(1 - e)$ from the focus on the x-axis. The semi-major axis of the ellipse therefore has length

$$a = \tfrac{1}{2}\left(\frac{\lambda}{1+e} + \frac{\lambda}{1-e}\right) = \frac{\lambda}{1-e^2},$$

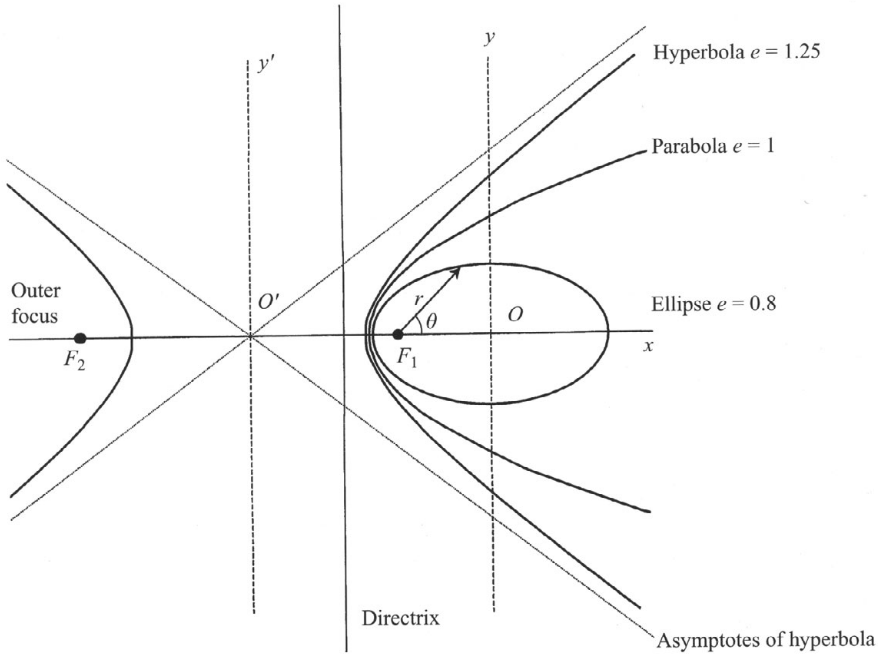


Fig. A4.2 Examples of conic sections: $e = 0$, circle (not shown); $0 < e < 1$, ellipse; $e = 1$, parabola; $e > 1$, hyperbola with two branches with inner and outer foci.

and the new centre is at distance $x = e\lambda/(1 - e^2)$ from F_1. In the new coordinate system, we can therefore write

$$x = r\cos\theta - e\lambda/(1 - e^2), \quad y = r\sin\theta.$$

With a bit of algebra, (A4.2) becomes

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \tag{A4.3}$$

where $b = a(1 - e^2)^{1/2}$. Equation (A4.3) shows that b is the length of the semi-minor axis. The meaning of e also becomes apparent. If $e = 0$, the ellipse becomes a circle. It is therefore appropriate that e is called the *eccentricity* of the ellipse.

Exactly the same analysis can be carried out for hyperbolae, for which $e > 1$. The algebra is exactly the same, but now the Cartesian coordinate system is referred to an origin at O' in Fig. A4.2 and $b^2 = a^2(1 - e^2)$ is a negative quantity. We can therefore write

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1, \tag{A4.4}$$

where $b = a(e^2 - 1)^{1/2}$.

One of most useful ways of relating the conic sections to the orbits of test particles in central fields of force is to rewrite the equations in what is known as *pedal* or *p–r form*. The variable θ is replaced by the distance coordinate p, the perpendicular distance from

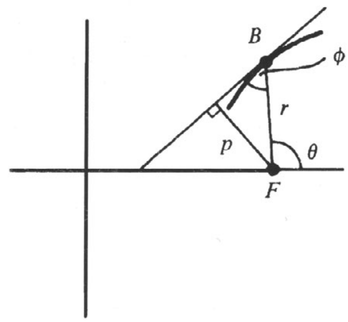


Fig. A4.3 Illustrating the definition of the pedal, or *p–r*, coordinate system.

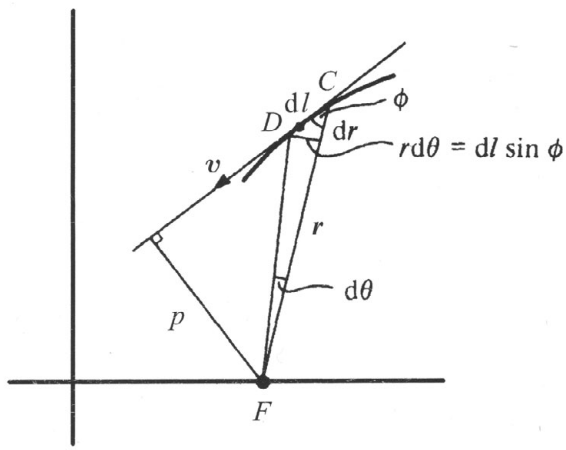


Fig. A4.4 The geometry involved in transforming from r–θ to r–p coordinates.

the tangent at a point on the curve to the focus, as illustrated in Fig. A4.3, which shows that $p = r \sin\phi$, where the angle ϕ is the angle between the tangent and the radius vector $\boldsymbol{r}$. We are interested in the tangent at the point B and so we take the derivative of θ with respect to r. From (A4.2), we find

$$\frac{\mathrm{d}\theta}{\mathrm{d}r} = -\frac{\lambda}{r^2 e \sin\theta}. \tag{A4.5}$$

Figure A4.4 shows the changes $\mathrm{d}\theta$ and $\mathrm{d}r$, as the point on the curve moves tangentially a distance $\mathrm{d}l$ along the curve. From the geometry of the small triangle defined by $\mathrm{d}l$, $\mathrm{d}r$ and $r\,\mathrm{d}\theta$, we see that

$$r\,\mathrm{d}\theta = \mathrm{d}l\ \sin\phi, \quad \mathrm{d}r = \mathrm{d}l\ \cos\phi \quad \text{and so} \quad \tan\phi = r\frac{\mathrm{d}\theta}{\mathrm{d}r}. \tag{A4.6}$$

We now have sufficient relations to eliminate θ from (A4.2) because

$$p = r\sin\phi \qquad \text{and, differentiating (A4.2),} \qquad \tan\phi = -\frac{\lambda}{re\sin\theta}. \tag{A4.7}$$

After a little algebra, we find that

$$\frac{\lambda}{p^2} = \frac{1}{A} + \frac{2}{r}, \tag{A4.8}$$

where $A = \lambda/(e^2 - 1)$. This the *pedal*, or p–r, *equation* for conic sections. If A is positive, we obtain hyperbolae, if A is negative, ellipses, and if A is infinite, parabolae. Notice that, in the case of the hyperbolae, (A4.8) refers to values of p relative to the inner focus. If the hyperbola is referred to the outer focus, the equation becomes

$$\frac{\lambda}{p^2} = \frac{1}{A} - \frac{2}{r}.$$

We can now turn the whole exercise round and state: 'If we find equations of the form (A4.8) from a physical theory, the curves must be conic sections'.

A.2 Kepler's Laws and Planetary Motion

Let us recall Kepler's three laws of planetry motion, the origins of which were discussed in detail in Chapter 3.

K1: Planetary orbits are ellipses with the Sun in one focus.
K2: Equal areas are swept out by the radius vector from the Sun to a planet in equal times.
K3: The period of a planetary orbit T is proportional to the three-halves power of its mean distance r from the Sun, that is, $T \propto r^{3/2}$.

Consider first the motion of a test particle moving in a central field of force. For an isolated system, Newton's laws of motion lead directly to the *law of conservation of angular momentum*, as shown in the standard textbooks. If $\boldsymbol{\Gamma}$ is the total net torque acting on the system and $\boldsymbol{L}$ is the total angular momentum,

$$\boldsymbol{\Gamma} = \frac{\mathrm{d}\boldsymbol{L}}{\mathrm{d}t},$$

where $\boldsymbol{L} = m(\boldsymbol{r} \times \boldsymbol{v})$ for the case of a test mass m orbiting in the field of force. For a *central* field of force originating at the point from which $\boldsymbol{r}$ is measured, $\boldsymbol{\Gamma} = \boldsymbol{r} \times \boldsymbol{f} = 0$, since $\boldsymbol{f}$ and $\boldsymbol{r}$ are parallel or antiparallel vectors, and hence

$$\boldsymbol{L} = m(\boldsymbol{r} \times \boldsymbol{v}) = \text{constant vector}, \tag{A4.9}$$

the law of conservation of angular momentum. Since there is no force acting outside the plane defined by the vectors $\boldsymbol{v}$ and $\boldsymbol{r}$, the motion is wholly confined to the $\boldsymbol{v}$–$\boldsymbol{r}$ plane.

The magnitude of $\boldsymbol{L}$ is constant and so (A4.9) can be rewritten

$$mrv \sin\phi = \text{constant}, \tag{A4.10}$$

where ϕ is the angle between the radius vector $\boldsymbol{r}$ and the velocity vector $\boldsymbol{v}$, which is in the direction of $\mathrm{d}\boldsymbol{l}$. But $r\sin\phi = p$ and hence

$$pv = \text{constant} = h. \tag{A4.11}$$

This calculation shows the value of introducing the quantity p. h is known as the *specific angular momentum*, meaning the angular momentum *per unit mass*.

Let us now work out the area swept out per unit time. In Fig. A4.4, the area of the triangle FCD is $\frac{1}{2} r\, \mathrm{d}l \sin\phi$. Therefore, the area swept out per unit time is

$$\tfrac{1}{2} r \sin\phi \frac{\mathrm{d}l}{\mathrm{d}t} = \tfrac{1}{2} r v \sin\phi = \tfrac{1}{2} p v = \tfrac{1}{2} h = \text{constant}, \tag{A4.12}$$

that is, Kepler's second law is no more than a statement of the law of conservation of angular momentum in the presence of a central force. Notice that the result does not depend upon the radial dependence of the force – it is only the fact that it is a central force which is important, as was fully appreciated by Newton.

Kepler's first law can be derived from the law of conservation of energy in the gravitational field of a point mass M. Let us work out the orbit of a test particle of mass m in this inverse square field of force. Newton's law of gravity in vector form can be written

$$\boldsymbol{f} = -\frac{GmM}{r^2}\, \mathbf{i}_r .$$

Setting $\boldsymbol{f} = -m \,\mathrm{grad}\, \phi$, where ϕ is the gravitational potential per unit mass of the particle, $\phi = -GM/r$. Therefore, the expression for conservation of energy of the particle in the gravitational field is

$$\tfrac{1}{2} m v^2 - \frac{GmM}{r} = C, \tag{A4.13}$$

where C is a constant of the motion. But we have shown that, because of conservation of angular momentum, for *any* central field of force, $pv = h =$ constant. Therefore,

$$\frac{h^2}{p^2} = \frac{2GM}{r} + \frac{2C}{m}, \tag{A4.14}$$

or

$$\frac{h^2/GM}{p^2} = \frac{2}{r} + \frac{2C}{GMm}. \tag{A4.15}$$

We recognise this equation as the pedal equation for conic sections, the exact form of the curve depending only on the sign of the constant C. Inspection of (A4.13) shows that, if C is negative, the particle cannot reach $r = \infty$ and so takes up a bound elliptical orbit; if C is positive, the orbits are hyperbolae. In the case $C = 0$, the orbit is parabolic. From the form of the equation, it is apparent that the origin of the force lies at the focus of the conic section.

To find the period of the particle in its elliptical orbit, we note that the area of the ellipse is πab and the rate at which area is swept out is $\frac{1}{2} h$ (A4.12). Therefore the period of the elliptical orbit is

$$T = \frac{2\pi ab}{h}. \tag{A4.16}$$

Comparing (A4.8) and (A4.15), the semi-latus rectum is $\lambda = h^2/GM$ and, from the analysis of Section A4.1, a, b and λ are related by

$$b = a\left(1 - e^2\right)^{1/2}; \quad \lambda = a\left(1 - e^2\right). \tag{A4.17}$$

Substituting into (A4.16), we find

$$T = \frac{2\pi}{(GM)^{1/2}}\, a^{3/2}. \tag{A4.18}$$

a is the semi-major axis of the ellipse and so is proportional to the mean distance of the test particle from the focus. We have derived Kepler's third law for the general case of elliptical orbits.

A.3 Rutherford Scattering

The scattering of α-particles by atomic nuclei was one of the great experiments of nuclear physics carried out by Rutherford and his colleagues Geiger and Marsden in Manchester in 1911. It established conclusively that the positive charge in atoms is contained within a point-like nucleus. The experiment involved firing α-particles at a thin gold sheet and measuring the angular distribution of the scattered α-particles. Rutherford was an experimenter of genius who was famous for his antipathy to theory. The theoretical calculation which follows was, however, carried out by Rutherford himself, checking with his theoretical colleagues that he had not made a blunder. No one had previously bothered to carry out this calculation, or realised its profound significance.

It is assumed that all the positive charge of an atom of atomic number Z is contained in a compact nucleus. The deflection of an α-particle of atomic number z due to the inverse square law of electrostatic repulsion between the particle and the nucleus can be determined using the formalism for orbits under an inverse square field of force. Figure A4.5 shows the dynamics and geometry of the repulsion in both pedal and Cartesian forms. From the considerations of Section A.2, we know that the trajectory is a hyperbola with the nucleus located at the outer focus. If there were no repulsion, the α-particle would have travelled along the diagonal AA' and passed by the nucleus at a perpendicular distance p_0, which is known as the *impact parameter*. The velocity of the α-particle at infinity is v_0.

The trajectory of the α-particle is given by (A4.4),

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1,$$

from which we can find the asymptotic solutions at large x and y, $x/y = \pm a/b$. Therefore, the scattering angle ϕ shown in Fig. A4.5 is

$$\tan\frac{\phi}{2} = \frac{x}{y} = \frac{a}{b} = (e^2 - 1)^{-1/2}. \tag{A4.19}$$

The pedal equation for the hyperbola with respect to its outer focus,

$$\frac{\lambda}{p^2} = \frac{1}{A} - \frac{2}{r}; \quad A = \lambda/(e^2 - 1),$$

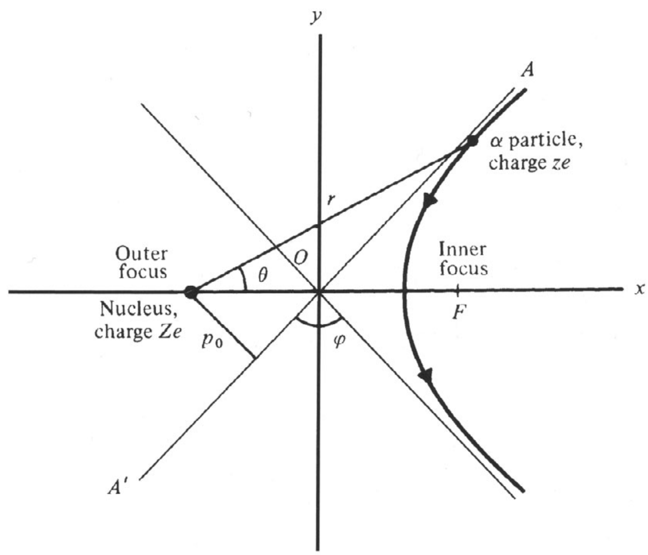


Fig. A4.5 Illustrating the geometry of the collision between an incoming α-particle and an atomic nucleus.

can be compared with the equation of conservation of energy of the α-particle in the field of the nucleus:

$$\tfrac{1}{2}mv^2 + \frac{Zze^2}{4\pi\epsilon_0 r} = \tfrac{1}{2}mv_0^2. \tag{A4.20}$$

Since $pv = h$ is a constant, (A4.20) can be rewritten

$$\left(\frac{4\pi\epsilon_0 mh^2}{Zze^2}\right)\frac{1}{p^2} = \frac{4\pi\epsilon_0 mv_0^2}{Zze^2} - \frac{2}{r}. \tag{A4.21}$$

Therefore $A = Zze/4\pi\epsilon_0 mv_0^2$ and $\lambda = 4\pi\epsilon_0 mh^2/Zze^2$. Recalling that $p_0 v_0 = h$, we can combine (A4.19) with the values of a and λ to find

$$\cot\frac{\phi}{2} = \left(\frac{4\pi\epsilon_0 m}{Zze^2}\right) p_0 v_0^2. \tag{A4.22}$$

The number of α-particles scattered through an angle ϕ is directly related to their impact parameter p_0. If a parallel beam of α-particles is fired at a nucleus, the number of particles with collision parameters in the range p_0 to $p_0 + \mathrm{d}p_0$ is

$$N(p_0)\,\mathrm{d}p_0 \propto 2\pi p_0\,\mathrm{d}p_0,$$

that is, proportional to the area of an annulus of width $\mathrm{d}p_0$ at radius p_0. Therefore, differentiating p_0 in (A4.22) with respect to ϕ, the number of particles scattered between ϕ to $\phi + \mathrm{d}\phi$ is

$$N(\phi)\,\mathrm{d}\phi \propto p_0\,\mathrm{d}p_0 \propto \left(\frac{1}{v_0^2}\cot\frac{\phi}{2}\right)\left(\frac{1}{v_0^2}\csc^2\frac{\phi}{2}\right)\mathrm{d}\phi = \frac{1}{v_0^4}\cot\frac{\phi}{2}\csc^2\frac{\phi}{2}\,\mathrm{d}\phi. \tag{A4.23}$$

This result can also be written in terms of the *probability* $p(\phi)$ that the α-particle is scattered through an angle ϕ. With respect to the incident direction, the number of particles scattered between angles ϕ and $\phi + d\phi$ is

$$N(\phi)\, d\phi = \tfrac{1}{2} \sin\phi \, p(\phi)\, d\phi. \tag{A4.24}$$

If the scattering were uniform over all solid angles, $p(\phi) = \text{constant}$. Equating the results (A4.23) and (A4.24), we find the famous result

$$p(\phi) \propto \frac{1}{v_0^4} \csc^4 \frac{\phi}{2}. \tag{A4.25}$$

This was the probability law which Rutherford derived in 1911.[20] He and his colleagues found that the α-particles scattered by a thin gold sheet followed exactly this relation for scattering angles between 5° and 150°, over which the function $\cot(\phi/2)\csc^2(\phi/2)$ varies by a factor of 40 000. From the known speeds of the α-particles, and the fact that the law was obeyed to large scattering angles, they deduced that the nucleus must be less than about 10^{-14} m in radius, that is, very much smaller than the size of atoms $\sim 10^{-10}$ m.

Pais has given an amusing account of Rutherford's discovery of the law.[21] As he points out, Rutherford was lucky to use α-particles of the right energy to observe the distribution law of scattered particles. He also remarks that Rutherford did not mention this key result at the first Solvay conference held in 1911, nor was the full significance of the result appreciated for a few years. He first used the term 'nucleus' in his book in radioactivity in 1912, stating,

> The atom must contain a highly charged nucleus.

Only in 1914, during a Royal Society discussion, did he come out forcefully in favour of the nuclear model of the atom.

Notes

1 Westphal, R.S. (1980). *Never at Rest: A Biography of Isaac Newton*, p. ix. Cambridge: Cambridge University Press.
2 Fauvel, J. *et al.* (eds.) (1988). *Let Newton Be: A New Perspective on his Life and Works*. Oxford: Oxford University Press.
3 Cohen, I.B. (1970). *Dictionary of Scientific Biography*, Vol. 11, p. 43. New York: Charles Scribner's Sons.
4 Fauvel, J. *et al.* (1988). *op. cit.*, p. 12.
5 *Sequialternate proportion* means 'to the power 3/2'.
6 Fauvel, J. *et al.* (1988). *op. cit.*, p. 14.
7 Cohen, I.B. (1970). *op. cit.*, p. 53.
8 Stukeley, W. (1752). *Memoirs of Sir Isaac Newton's Life*, ed. A.H. White, London (1936), pp. 19–20.
9 Cohen, I.B. (1970). *op. cit.*, p. 44.
10 Chandrasekhar, S. (1995). *Newton's Principia for the Common Reader*, p. 7. Oxford: Clarendon Press.
11 Chandrasekhar, S. (1995). *op. cit.*, pp. 11–12.

12 Cohen, I.B. (1970). *op. cit.*, p. 44.
13 Golinski, J. (1988). In *Let Newton Be*, *op. cit.*, p. 153.
14 Golinski, J. (1988). *op. cit.*, p. 156.
15 Golinski, J. (1988). *op. cit.*, p. 160.
16 Golinski, J. (1988). *op. cit.*, p. 151.
17 Rattansi, P. (1988). In *Let Newton Be*, *op. cit.*, p. 185.
18 Rattansi, P. (1988). *op. cit.*, p. 198.
19 Rattansi, P. (1988). *op. cit.*, p. 200.
20 Rutherford, E. (1911). *Phil. Mag.*, **21**, 669.
21 Pais, A. (1986). *Inward Bound*, pp. 188–193. Oxford: Clarendon Press.

Case Study II

MAXWELL'S EQUATIONS

Newtonian physics dominated the description of physical world in the exact sciences from the time of Newton until the second half of the nineteenth century. More and more sophisticated techniques were developed to enable the laws of motion to be applied in a huge diversity of applications, and that story is the subject of Case Study III. But the nineteenth century also saw a major shift in perspective with the realisation that the laws of physics could be described in terms of *fields* rather than according to the mechanical world picture of Newton. Central to this development were the discoveries of the laws of electromagnetism and Maxwell's equations for the electromagnetic field.

In Chapter 5, we begin by paying tribute to Michael Faraday, an experimenter of genius who, among many other achievements, discovered the phenomena of *electromagnetic induction* (Fig. II.1). His introduction of the concept of *lines of force* was crucial for the mathematisation of electromagnetism, what I have described as 'mathematics without mathematics'. This led to Maxwell's series of papers in which he introduced the concept of the *displacement current* through the use of a purely mechanical model for electromagnetic phenomena. His mathematisation of the laws of electricity and magnetism led to the demonstration that light is electromagnetic radiation, the unification of light and electromagnetism. This discovery resulted from as extreme an example of model building as I have encountered in physics. It is a remarkable example of how fruitful it can be to work by analogy, provided one is constrained by experiment. The key role of vector calculus in simplifying the mathematics of electromagnetism provides an opportunity for revising some of that material, and a number of useful results are included as an Appendix to Chapter 5.

Chapter 5 takes us up to Maxwell's papers of 1861–62, before his truly epochal paper of 1865. In his exposition of the theory of electromagnetism described in Chapter 6, Maxwell dispenses with mechanical models and deals directly with the fields themselves. This paper sets out his new perspective on the workings of nature. Fields are not just mathematical conveniences, but are physical phenomena which transmit forces through material media and the vacuum. This is quite unlike the Newtonian picture of 'action at a distance', which is implicit in Newton's Laws of Motion – Maxwell was emphatic about this distinction. Chapter 6 is devoted to an analysis of Maxwell's paper of 1865 and the subsequent validation of Maxwell's insights by the experiments of Heinrich Hertz. The implications for theoretical physics were profound. The Maxwellian revolution established the central role of fields which were to replace the mechanical world picture of Newton.

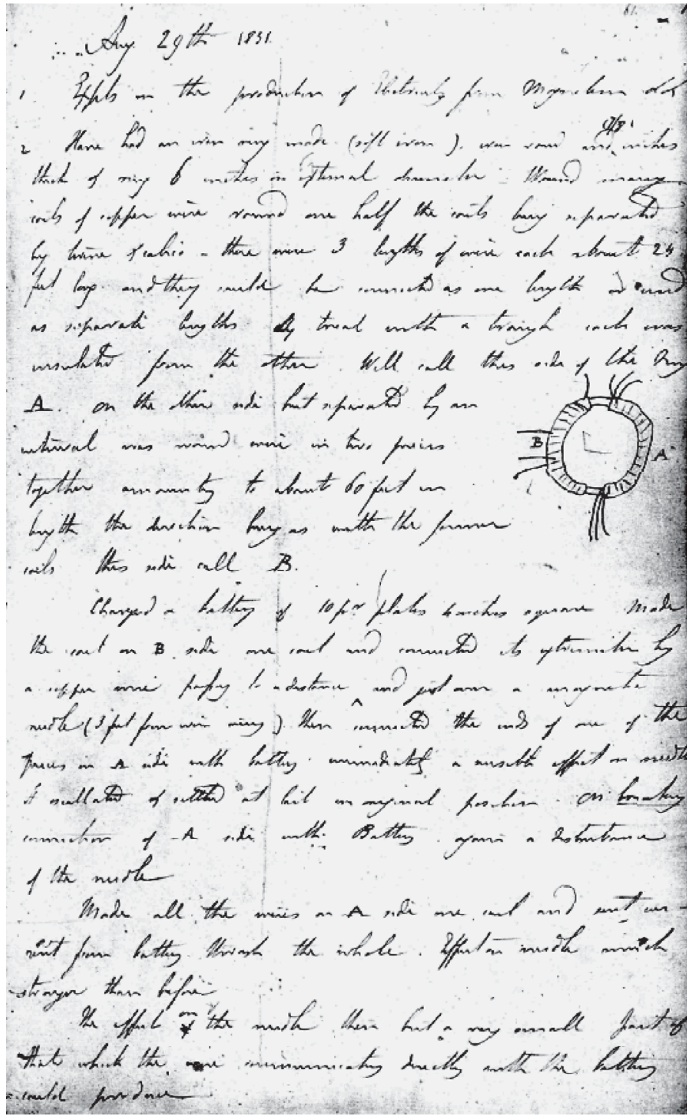


Fig. II.1 The page from Faraday's notebooks, dated 29 August 1831, in which he describes the discovery of electromagnetic induction (see Section 5.2). (Courtesy of the Royal Institution of Great Britain.)

In Chapter 7, we rewrite the history of electromagnetism by beginning with the mathematical structure of Maxwell's equations and then work out what we have to do to endow them with physical meaning with a minimum number of assumptions. This may seem a somewhat contrived exercise, but it provides insight into the mathematical structures underlying the theory and is not so different from what often has to be done in frontier areas of theoretical research. It also gives the subject much greater coherence, enabling the vast number of different aspects of classical electromagnetism to be fully encompassed by Maxwell's set of four partial differential equations.

5 The Origin of Maxwell's Equations

5.1 How It All Began

Electricity and magnetism have an ancient history. Magnetic materials are mentioned as early as 800 BC by the Greek writers, the name being derived from the mineral magnetite, which was known to attract iron in its natural state and which was mined in the Greek province of Magnesia in Thessaly. Magnetic materials were of special importance because of their use in compasses, reflected in the English word for the mineral, lodestone, meaning leading stone. Static electricity was also known to the Greeks through the electrostatic phenomena observed when amber is rubbed with fur – the Greek word for amber is *elektron*. The first systematic study of magnetic and electric phenomena was published in 1600 by William Gilbert in his treatise *De Magnete, Magneticisque Corporibus, et de Magno Magnete Tellure*. The main subject of the treatise was the Earth's magnetic field, which he showed was similar to that of a bar magnet. He also described the force between two bodies charged by friction and named it the *electric* force between them.

In addition to his famous experiments, in which he showed that lightning is an electrostatic discharge, Benjamin Franklin systematised the laws of electrostatics and defined the conventions for naming positive and negative electric charges. In the course of these studies, he also enunciated the *law of conservation of electric charge*. In 1767, Joseph Priestley showed that, inside a hollow conducting sphere, there are no electric forces. From this, he inferred that the force law for electrostatics must be of inverse square form, just as in the case of gravity. Experiments to improve the accuracy with which the inverse square law held good were carried out by Henry Cavendish in the 1770s but were not published. They were repeated with much greater precision by Maxwell and Donald MacAlister as one of the first research projects carried out in the newly founded Cavendish Laboratory in the 1870s – they showed that the exponent in the inverse square law was 2 within one part in 21 600.[1] A modern version of this experiment, carried out by Williams, Faller and Hall in 1971, established that the inverse square law holds good to better than one part in 3×10^{15}.[2] An improved limit of one part in 10^{16} was obtained by Fulcher in 1986.[3]

By the end of the eighteenth century, many of the basic experimental features of electrostatics and magnetostatics had been established. In the 1770s and 1780s, Charles-Augustin Coulomb performed very sensitive electrostatic experiments using a torsion balance to establish directly the inverse square laws of electrostatics. He undertook similar experiments in magnetostatics using very long magnetic dipoles, so that the properties of each pole of the dipole could be considered separately. The laws can be written in SI notation as:

$$f = \frac{q_1 q_2}{4\pi\epsilon_0 r^2}, \tag{5.1}$$

$$f = \frac{\mu_0 p_1 p_2}{4\pi r^2}, \tag{5.2}$$

where q_1 and q_2 are the electric charges of two point objects separated by distance r and p_1 and p_2 their magnetic pole strengths. The constants $1/4\pi\epsilon_0$ and $\mu_0/4\pi$ are included in these definitions according to SI conventions.[4] In modern vector notation the directional dependence of the electrostatic force can incorporated explicitly,

$$\boldsymbol{f} = \frac{q_1 q_2}{4\pi\epsilon_0 r^3}\,\boldsymbol{r} \quad \text{or} \quad \boldsymbol{f} = \frac{q_1 q_2}{4\pi\epsilon_0 r^2}\,\boldsymbol{i}_r, \tag{5.3}$$

where $\boldsymbol{i}_r$ is the unit vector directed radially *away* from either charge in the direction of the other. A similar pair of expressions is found for magnetostatic forces:

$$\boldsymbol{f} = \frac{\mu_0 p_1 p_2}{4\pi r^3}\,\boldsymbol{r} \quad \text{or} \quad \boldsymbol{f} = \frac{\mu_0 p_1 p_2}{4\pi r^2}\,\boldsymbol{i}_r. \tag{5.4}$$

The late eighteenth and early nineteenth century was a period of extraordinary brilliance in French mathematics. Of special importance for our story are the works of Siméon Poisson, who was a pupil of Pierre-Simon Laplace, and Joseph-Louis Lagrange. In 1812, Poisson published his famous *Mémoire sur la Distribution de l'Électricité à la Surface des Corps Conducteurs*, in which he demonstrated that many of the problems of electrostatics can be simplified by the introduction of the electrostatic potential V which is the solution of Poisson's equation:

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = -\frac{\rho_e}{\epsilon_0}, \tag{5.5}$$

where ρ_e is the electric charge density distribution. The electric field strength $\boldsymbol{E}$ is then given by

$$\boldsymbol{E} = -\text{grad}\, V. \tag{5.6}$$

In 1826, Poisson published the corresponding expressions for the magnetic flux density $\boldsymbol{B}$ in terms of the magnetostatic potential V_{mag}:

$$\frac{\partial^2 V_{mag}}{\partial x^2} + \frac{\partial^2 V_{mag}}{\partial y^2} + \frac{\partial^2 V_{mag}}{\partial z^2} = 0, \tag{5.7}$$

where the magnetic flux density $\boldsymbol{B}$ is given by

$$\boldsymbol{B} = -\mu_0\, \text{grad}\, V_{mag}. \tag{5.8}$$

Until 1820, electrostatics and magnetostatics appeared to be quite separate phenomena, but this changed with the development of the science of *current electricity*. In parallel with the development of the laws of electrostatics and magnetostatics in the latter years of the eighteenth century, the Italian anatomist Luigi Galvani discovered that electrical effects could stimulate the muscular contraction of frogs' legs. In 1791, he showed that, when

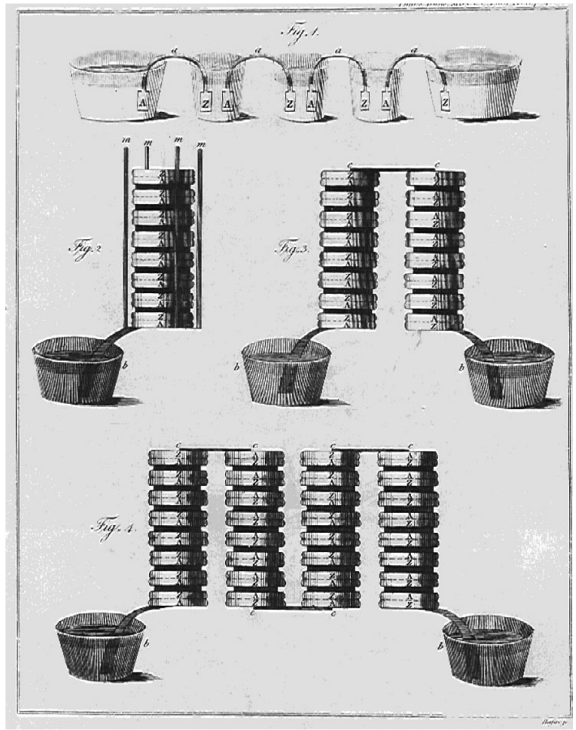


Fig. 5.1 Diagrams from Alessandro Volta's letter of 20 March 1800 to Joseph Banks, President of the Royal Society of London, showing Volta's crown of cups in Fig. 1 and examples of combinations of voltaic piles in Figs. 2 to 4. (From 'On the Electricity Excited by the Mere Contact of Conducting Substances of Different Kinds.' *Philosophical Transactions of the Royal Society of London*, 1800, **90**, 403–31.)

two dissimilar metals were used to make the connection between nerve and muscle, the same form of muscular contraction was observed. This was announced as the discovery of *animal electricity*. Alessandro Volta suspected that the electric current was associated with the presence of different metals in contact with a moist body. In 1800, he demonstrated this by constructing what became known as a *voltaic pile*, which consisted of interleaved layers of copper and zinc separated by layers of pasteboard soaked in a conducting liquid (*Figs. 2 to 4* in Fig. 5.1). With this pile, Volta was able to demonstrate all the phenomena of electrostatics – the production of electric discharges, electric shocks, and so on. By far the most important aspect of Volta's experiments was, however, the fact that he had discovered a *controllable source* of electric current. There was, however, a problem with the voltaic cell in that it had a short lifetime because the pasteboard dried out. This led Volta to invent his *crown of cups*, in which the electrodes were placed in glass vessels (*Fig. 1* in Fig. 5.1) – these were the precursors of modern batteries.

A key experimental discovery was made in 1820 by Hans-Christian Ørsted, who showed that there is always a magnetic field associated with the flow of electric current. As soon as this discovery was announced, Jean-Baptiste Biot and Félix Savart set out to discover the dependence of the strength of the magnetic field at distance $\boldsymbol{r}$ from a current element of length $\mathrm{d}\boldsymbol{l}$ in which a current I is flowing. In the same year, they found the answer, the *Biot–Savart law*, which in modern vector notation can be written

$$\mathrm{d}\boldsymbol{B} = \frac{\mu_0 I\,(\mathrm{d}\boldsymbol{l} \times \boldsymbol{r})}{4\pi r^3}. \tag{5.9}$$

Notice that the signs of the vectors are important in finding the correct direction of the field. The term $\mathrm{d}\boldsymbol{l}$ is the length of the current element in the direction of the current I and $\boldsymbol{r}$ is measured from the current element $\mathrm{d}\boldsymbol{l}$ to a point at vector distance $\boldsymbol{r}$. Next, André-Marie Ampère extended the Biot–Savart law to relate the current flowing though a closed loop to the integral of the component of the magnetic flux density around the loop. In modern vector notation, *Ampère's circuital law* in free space can be written

$$\oint_C \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{s} = \mu_0 I_{\text{enclosed}}, \tag{5.10}$$

where I_{enclosed} is the total electric current flowing through the area enclosed by the loop C.

The story developed rapidly. In 1826, Ampère published his famous treatise *Theorie des Phénomènes électro-dynamique, uniquement déduite de l'expérience*, which included the demonstration that the magnetic field of a current loop could be represented by an equivalent magnetic shell. In the treatise, he also formulated the equation for the force between two current elements, $\mathrm{d}\boldsymbol{l}_1$ and $\mathrm{d}\boldsymbol{l}_2$ carrying currents I_1 and I_2:

$$\mathrm{d}\boldsymbol{F}_2 = \frac{\mu_0 I_1 I_2\,[\mathrm{d}\boldsymbol{l}_1 \times (\mathrm{d}\boldsymbol{l}_2 \times \boldsymbol{r})]}{4\pi r^3}. \tag{5.11}$$

$\mathrm{d}\boldsymbol{F}_2$ is the force acting on the current element $\mathrm{d}\boldsymbol{l}_2, \boldsymbol{r}$ being measured from $\mathrm{d}\boldsymbol{l}_1$. Ampère also demonstrated the relation between this law and the Biot–Savart law.

In 1827, Georg Simon Ohm formulated the relation between potential difference V and the current I, what is now known as *Ohm's law*, $V = RI$, where R is the resistance of the material through which the current flows. Sadly, this pioneering work was not well received by Ohm's colleagues in Cologne and he resigned from his post in disappointment. The central importance of his work was subsequently recognised, however, including the award of the Copley Medal of the Royal Society of London, its highest award, in 1841.[5]

The laws described above were known by 1830 and comprise the whole of *static electricity*, namely, the forces between *stationary* charges, magnets and currents. The essence of Maxwell's equations is that they deal with *time-varying* phenomena as well. Over the succeeding 20 years, the basic experimental features of time-varying electric and magnetic fields were established by the genius of Michael Faraday.

5.2 Michael Faraday: Mathematics without Mathematics

Michael Faraday was born into a poor family, his father being a blacksmith who moved with his family to London in early 1791. Faraday began life as a bookbinder's apprentice working in Mr Ribeau's shop and learned his early science by reading the books he had to bind.[6] These included the *Encyclopaedia Britannica*, his attention being particularly attracted by the article on electricity by James Tyler. He attempted to repeat some of the experiments himself and built a small electrostatic generator out of bottles and old wood.

In 1812, one of Mr Ribeau's customers gave Faraday tickets to Humphry Davy's lectures at the Royal Institution. Faraday sent a bound copy of his lecture notes to Davy, inquiring if there was any position available, but there was nothing at that time. In October of the same year, however, Davy was temporarily blinded by an explosion while working with the dangerous chemical 'nitrate of chlorine' and needed someone to write down his thoughts. Faraday was recommended for this task and subsequently, on 1 March 1813, took up a permanent position as assistant to Davy at the Royal Institution, where he was to remain for the rest of his life.

Soon after Faraday's appointment, Davy made a prolonged tour of several prominent scientific institutions on continental Europe and took Faraday with him as his scientific assistant. During the next 18 months, they met most of the great scientists of the day – in Paris, they met Ampère, Humboldt, Gay-Lussac, Arago and many others. In Italy, he met Volta and, while in Genoa, observed experiments conducted on the torpedo, a fish capable of giving electric shocks.

Ørsted's announcement of the discovery of a connection between electricity and magnetism in 1820 caused a flurry of scientific activity. Many articles were submitted to the scientific journals describing other electromagnetic effects and attempting to explain them. Faraday was asked to survey this mass of experiment and speculation by the editor of the *Philosophical Magazine* and, as a result, began his systematic study of electromagnetic phenomena.

Faraday proceeded to repeat all the experiments reported in the literature. In particular, he noted the movement of the poles of a small magnet in the vicinity of a current-carrying wire. It had already been noted by Ampère that the force acting on the pole of a magnet is such as to move it in a circle about the wire. Alternatively, if the magnet were kept fixed, a current-carrying wire would feel a force, moving it in a circle about the magnet. Faraday proceeded to demonstrate these phenomena by two beautiful experiments (Fig. 5.2). In one of them, illustrated in the right-hand diagram in Fig. 5.2, a magnet was placed upright in a dish of mercury with one pole projecting above the surface. A wire was arranged so that one end of it was attached to a small cork which floated on the mercury and the other end was fixed above the end of the magnet. When a current was passed through the wire, the wire was found to rotate about the axis of the magnet, exactly as Faraday had expected. In the second experiment, illustrated in the left-hand diagram of Fig. 5.2, the current-carrying wire was fixed and the magnet free to rotate about the wire. These were the first *electric motors* to be constructed.

Faraday's deep involvement in reviewing what was known about electricity and magnetism led him to the concept of *magnetic lines of force*, which sprang from the observation of the patterns which iron filings take up about a bar magnet (Fig. 5.3). A magnetic line of force represents the direction in which the force acting upon a magnetic pole acts when placed in a magnetic field. The greater the number of lines of force per unit area in the plane perpendicular to the field line, the greater the force acting upon the magnetic pole. Faraday was to lay great emphasis upon the use of this model as a means of visualising the effects of stationary and time-varying magnetic fields.

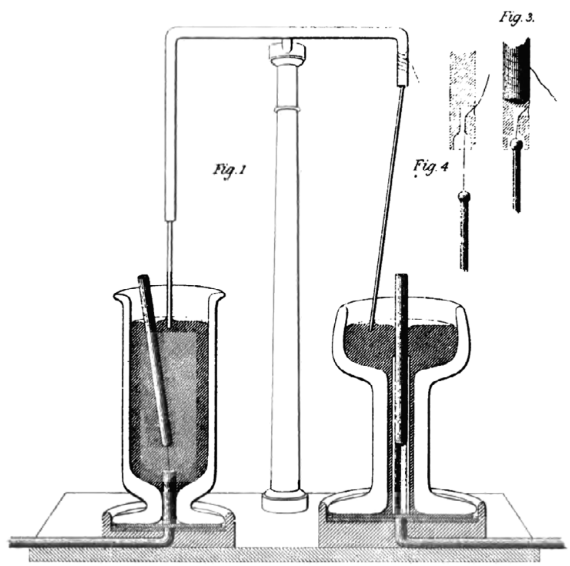


Fig. 5.2 Faraday's experiments illustrating the forces acting between a current-carrying wire and a magnet. In the left-hand half of the diagram, the current-carrying wire is fixed and the magnet rotates about the vertical axis; in the right-hand half of the diagram, the magnet is fixed and the current-carrying wire rotates about the vertical axis. These were the first electric motors to be constructed.

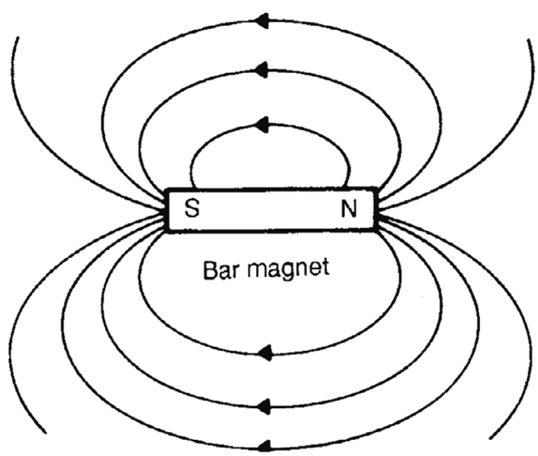


Fig. 5.3 Illustrating Faraday's concept of the magnetic lines of force about a bar magnet.

There was a problem with Faraday's picture – the force between two magnetic poles acts along the line between them. How could this behaviour be reconciled with the circular lines of force observed about a current-carrying wire? Faraday showed how he could simulate all the effects of a magnet if the current-carrying wire were bent into a loop, as illustrated in Fig. 5.4. Using the concept of lines of force, he argued that the magnetic

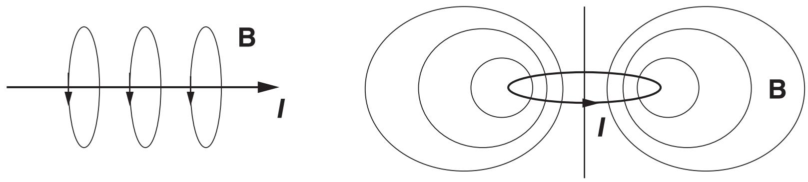


Fig. 5.4 Illustrating Faraday's reasoning concerning the equivalence of the magnetic field line distribution about a current loop and that of a bar magnet.

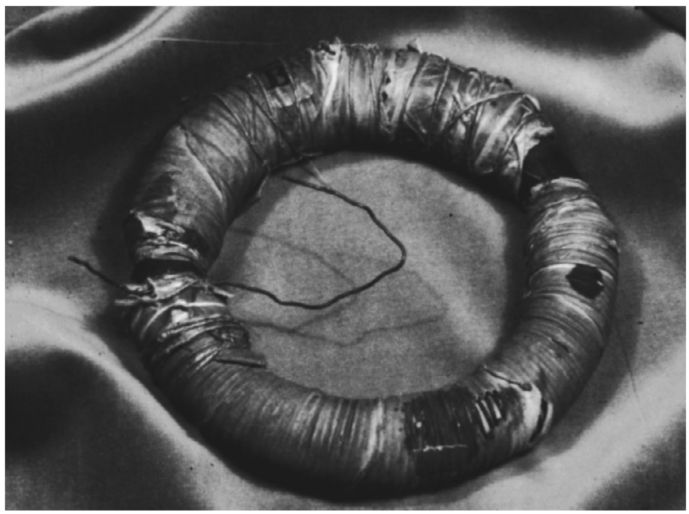


Fig. 5.5 The apparatus with which Faraday first demonstrated electromagnetic induction. (Courtesy of the Royal Institution of Great Britain.)

lines of force would be compressed within the loop with the result that one side of the loop would have one polarity and the other the opposite polarity. He proceeded to demonstrate experimentally that indeed all aspects of the forces associated with currents flowing in wires could be understood in terms of the concept of magnetic lines of force. In fact, all the laws of the forces between static magnets and currents can be derived from the exact equivalence of magnetic dipoles and current loops, as demonstrated in Appendix A5.7.

The great advance occurred in 1831. Believing in the symmetry of nature, Faraday conjectured that, since an electric current could produce a magnetic field, it is must also be possible to generate an electric current from a magnetic field. In that year, he learned of Joseph Henry's experiments in Albany, New York, using very powerful electromagnets. Faraday immediately had the idea of observing the strain in the material of a strong electromagnet caused by the lines of force. He built a strong electromagnet by winding an insulating wire, through which a current could be passed, onto a thick iron ring, thus creating a magnetic field within the ring. The effects of the strain were to be detected by another winding on the ring, which was attached to a galvanometer to measure any electric current produced. A photograph of his original apparatus is shown in Fig. 5.5.

The experiment was conducted on 29 August 1831 and is recorded meticulously in Faraday's laboratory notebooks (Fig. II.1). The effect was not at all what Faraday expected. When the primary circuit was closed, there was a displacement of the galvanometer needle in the secondary winding – an electric current had been induced in the secondary winding. Deflections of the galvanometer needle were *only* observed when the current in the electromagnet was switched on and off – there was no effect when a steady current flowed in the electromagnet. The effect only seemed to be associated with changing currents and, consequently, changing magnetic fields. This was the discovery of *electromagnetic induction*.

Over the next few weeks, there followed a series of remarkable experiments in which the nature of electromagnetic induction was established. The sensitivity of his apparatus was improved and he observed that the sense of the electric current in the second circuit changed when the electric current was switched on and off. Next, he tried coils of different shapes and sizes and found that the iron bar was not needed to create the effect. On 17 October 1831, a new experiment was carried out in which an electric current was created by sliding a cylindrical bar magnet into a long coil, or solenoid, connected to a galvanometer. In a famous experiment demonstrated at the Royal Society of London on 28 October 1831, he showed how a continuous electric current could be generated by rotating a copper disc between the poles of the 'great horse-shoe magnet' belonging to the Society. The axis and the edge of the disc were connected to a galvanometer and, as the disc rotated, the needle was continuously deflected. On 4 November 1831, he found that simply moving a copper wire between the poles of the magnet could create an electric current. Within a period of four months, he had discovered the *transformer* and the *dynamo*.

As early as 1831, Faraday had established the qualitative form of his law of induction in terms of lines of force – *the electromotive force induced in a current loop is directly related to the rate at which magnetic field lines are cut*, adding that

> By magnetic curves, I mean lines of magnetic force which would be depicted by iron filings.[7]

He now realised that the term electricity could mean a number of different things. In addition to *magneto-electricity*, which he had just discovered, electricity could be produced by friction, as had been known from ancient times as the cause of *static electricity*. It could be associated with chemical effects in a voltaic pile, *voltaic electricity*. In *thermo-electricity*, a potential difference is created when materials of different types are placed in contact and the ends at which the joins are made are maintained at different temperatures. Finally, there is *animal electricity* produced by fish such as torpedos and electric eels, which he had seen on his travels with Davy. Faraday asked 'are these different forms of electricity the same?' In 1832, he performed an elegant series of experiments in which he showed that he could produce similar chemical, magnetic and other effects, no matter what the source of the electricity might be, including electric fish.

Although the law of induction emerged at an early stage, it took Faraday many years to complete all the necessary experimental work to demonstrate the general validity of the law. In 1834, Emil Lenz enunciated the law which cleared up the issue of the direction of the induced electromotive force in the circuit – *the electromotive force acts in such a direction as to oppose the change in magnetic flux*.

Faraday was convinced that the concept of lines of force provided the key to understanding electromagnetic phenomena. In 1846, he speculated in a discourse to the Royal Institution that light might be some form of disturbance propagating along the field lines. He published these ideas in May 1846 in a paper entitled *Thoughts on Ray Vibrations*,[8] but they were received with considerable scepticism. Faraday had, however, indeed hit upon the correct concept. As will be shown in the next section, James Clerk Maxwell showed in 1864 that light is indeed a form of electromagnetic radiation. As Maxwell acknowledged in his great paper *A Dynamical Theory of the Electromagnetic Field* of 1865,

> The conception of the propagation of transverse magnetic disturbances to the exclusion of normal ones is distinctively set forth by Professor Faraday in his *Thoughts on Ray Vibrations*. The electromagnetic theory of light as proposed by him is the same in substance as that which I have begun to develope in this paper, except that in 1846 there was no data to calculate the velocity of propagation.[9]

Although Faraday could not formulate his ideas mathematically, his deep feeling for the behaviour of electric and magnetic fields provided the essential insights needed to develop the mathematical theory of the electromagnetic field. In Maxwell's words,

> As I proceeded with the study of Faraday, I perceived that his method of conceiving of phenomena was also a mathematical one, though not exhibited in the conventional form of mathematical symbols... I found, also, that several of the most fertile methods of research discovered by the mathematicians could be expressed much better in terms of ideas derived from Faraday than in their original form.[10]

Faraday had an intuitive belief in the unity of the forces of nature, and in particular between the phenomena of light, electricity and magnetism. In a series of experiments carried out towards the end of 1845, he attempted to find out if the polarisation of light was influenced by the presence of a strong electric field, but these experiments showed no positive result. Turning instead to magnetism, his initial experiments were also unsuccessful, until he passed light through lead borate glass in the presence of a strong magnetic field. He had cast these glasses himself in the years 1825 to 1830 as part of a project sponsored by the Royal Society of London to create superior optical glasses for use in astronomical instruments. These heavy glasses had the property of having large refractive indices. Faraday demonstrated the phenomenon now known as *Faraday rotation*, in which the plane of polarisation of linearly polarised light is rotated when the light rays travel along the magnetic field direction in a transparent dielectric. William Thomson, later Lord Kelvin, interpreted this phenomenon as evidence that the magnetic field caused the rotational motion of the electric charges in molecules. Following an earlier suggestion by Ampère, Kelvin envisaged magnetism as being essentially rotational in nature, and this was to influence Maxwell's model for a magnetic field in free space.

Faraday is an outstanding example of a meticulous experimenter of genius with no mathematical training who was never able to express the results of his researches in mathematical terms – there is not a single mathematical formula in his writings. He had, however, an intuitive genius for experiment, and for devising empirical conceptual models to account for his results. These models embodied the mathematics necessary to formulate the theory of the electromagnetic field.

5.3 Maxwell's Route to the Equations for the Electromagnetic Field

James Clerk Maxwell was born and educated in Edinburgh. In 1850, he went up to Cambridge where he studied for the Mathematical Tripos with considerable distinction. As James David Forbes, his Professor of Natural Philosophy at Edinburgh University wrote to William Whewell, the Master of Trinity College, in April 1852,

> Pray do not suppose that … I am not aware of his exceeding uncouthness, as well Mathematical as in other respects; … I thought the Society and Drill of Cambridge the only chance of taming him & much advised his going.[11]

Allied to his formidable mathematical abilities was a physical imagination which could appreciate the empirical models of Faraday and give them mathematical substance. Peter Harman's brilliant study *The Natural Philosophy of James Clerk Maxwell* is essential reading for appreciating Maxwell's intellectual approach and achievements.[12]

5.3.1 *On Faraday's Lines of Force (1856)*

The distinctive feature of Maxwell's thinking was his ability to work by *analogy*. As early as 1856, he described his approach in an essay entitled *Analogies in Nature* written for the Apostles club at Cambridge. The technique is best illustrated by the examples given below, but its essence can be caught in the following passage from the essay:

> Whenever [men] see a relation between two things they know well, and think they see there must be a similar relation between things less known, they reason from one to the other. This supposes that, although pairs of things may differ widely from each other, the *relation* in the one pair may be the same as that in the other. Now, as in a scientific point of view the *relation* is the most important thing to know, a knowledge of the one thing leads us a long way towards knowledge of the other.[13]

In other words, the approach consists of recognising mathematical similarities between quite distinct physical problems and seeing how far one can go in applying the successes of one theory to different circumstances. In electromagnetism, he found formal analogies between the mathematics of mechanical and hydrodynamical systems and the phenomena of electrodynamics. He acknowledged throughout this work his debt to William Thomson, who had made substantial steps in mathematising electric and magnetic phenomena. Maxwell's great contribution was not only to take this process very much further, but also to give it real physical content.

In the same year, 1856, Maxwell published the first of his major papers on electromagnetism, *On Faraday's Lines of Force*.[14] In the preface to his *Treatise on Electricity and Magnetism* of 1873, he recalled:

> before I began the study of electricity I resolved to read no mathematics on the subject till I had first read through Faraday's *Experimental Researches in Electricity*.[15]

The first part of the paper enlarged upon the technique of analogy and drew particular attention to its application in comparing the properties of incompressible fluid flow with magnetic lines of force. We will use the vector operator expressions div, grad and curl in our development, although the use of vector methods was only introduced by Maxwell into the study of electromagnetism in a paper of 1870 entitled *On the Mathematical Classification of Physical Quantities*.[16] In his paper of 1856, the partial derivatives were written out in Cartesian form.

Consider first the *continuity equation*, or the *equation of conservation of mass*, for incompressible fluid flow. The volume v is bounded by a closed surface S and, using vector notation, the mass flow per unit time through a surface element $\mathrm{d}\boldsymbol{S}$ is $\rho\boldsymbol{u}\cdot\mathrm{d}\boldsymbol{S}$ where $\boldsymbol{u}$ is the fluid velocity and ρ its density distribution. Therefore, the total mass flux through the closed surface is $\int_S \rho\boldsymbol{u}\cdot\mathrm{d}\boldsymbol{S}$. This is equal to the rate of loss of mass from v, which is

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_v \rho\,\mathrm{d}v. \tag{5.12}$$

Therefore,

$$-\frac{\mathrm{d}}{\mathrm{d}t}\int_v \rho\,\mathrm{d}v = \int_S \rho\boldsymbol{u}\cdot\mathrm{d}\boldsymbol{S}. \tag{5.13}$$

Applying the divergence theorem (A5.1) to the right-hand side of (5.13) and transposing, we find

$$\int_v \left[\mathrm{div}(\rho\boldsymbol{u}) + \frac{\partial\rho}{\partial t}\right]\mathrm{d}v = 0. \tag{5.14}$$

This result must be true for any volume element and hence

$$\mathrm{div}\,\rho\boldsymbol{u} = -\frac{\partial\rho}{\partial t}. \tag{5.15}$$

If the fluid is incompressible, ρ = constant and hence

$$\mathrm{div}\,\boldsymbol{u} = 0. \tag{5.16}$$

Maxwell took the concepts of lines and tubes of force as expounded by Faraday very seriously and drew an immediate analogy between the behaviour of magnetic field lines and the streamlines of incompressible fluid flow (Fig. 5.6). The velocity $\boldsymbol{u}$ is analogous to

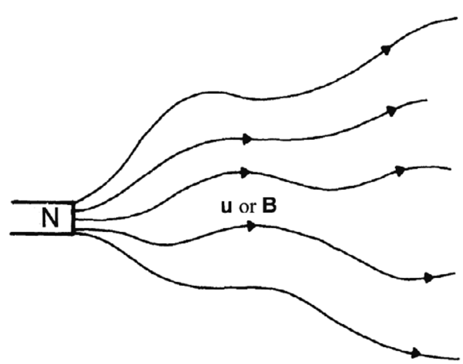


Fig. 5.6 Illustrating the analogy between magnetic field lines and the streamlines in the flow of an incompressible fluid.

the magnetic flux density $\boldsymbol{B}$. For example, if the tubes of force, or streamlines, diverge, the strength of the field decreases, as does the fluid velocity. This suggests that the magnetic field can be characterised by

$$\operatorname{div} \boldsymbol{B} = 0. \tag{5.17}$$

Maxwell recognised the important distinction between $\boldsymbol{B}$ and $\boldsymbol{H}$, associating $\boldsymbol{B}$ with magnetic fluxes, what he termed the *magnetic induction*, and $\boldsymbol{H}$ with forces, what he called the *magnetic intensity*. $\boldsymbol{u}$ is associated with the flux density of an incompressible fluid through unit area of the surface, just as $\boldsymbol{B}$, the magnetic flux density, is the flux of a vector field through unit area; $\boldsymbol{H}$, the magnetic field strength, is associated with the force on a unit magnetic pole.

One of the achievements of this paper was that all the relations between electromagnetic phenomena known at that time were expressed as a set of six 'laws', expressed in words rather than mathematically, very much akin to the style of Faraday.[17]

Faraday's law of electromagnetic induction was first put into mathematical form by Franz Ernst Neumann who, in 1845, wrote down explicitly the proportionality of the electromotive force $\mathcal{E}$ induced in a closed circuit to the rate of change of magnetic flux Φ,

$$\mathcal{E} = -\frac{\mathrm{d}\Phi}{\mathrm{d}t}, \tag{5.18}$$

where $\Phi = \int_S \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S}$ is the total magnetic flux through the circuit. This can be rewritten as

$$\int_C \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{s} = -\frac{\mathrm{d}}{\mathrm{d}t} \int_S \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S}. \tag{5.19}$$

The left-hand side is the definition of electromotive force $\mathcal{E}$ induced in the circuit and the right-hand side contains the definition of the total magnetic flux through the surface S in terms of the magnetic flux density $\boldsymbol{B}$. Now Stokes' theorem can be applied to the left-hand side of (5.19),

$$\int_S \operatorname{curl} \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{S} = -\frac{\mathrm{d}}{\mathrm{d}t} \int_S \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S}. \tag{5.20}$$

Applying this result to the elementary surface area $\mathrm{d}\boldsymbol{S}$, we find

$$\operatorname{curl} \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}. \tag{5.21}$$

Next, he rewrote the relation between the magnetic field produced by an electric current, Ampère's law, in vector form,

$$\oint_C \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{s} = I_{\text{enclosed}}, \tag{5.22}$$

where I_{enclosed} is the current flowing through the surface bounded by C. Thus,

$$\oint_C \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{s} = \int_S \boldsymbol{J} \cdot \mathrm{d}\boldsymbol{S}, \tag{5.23}$$

where $\boldsymbol{J}$ is the current density. Applying Stokes' theorem to (5.23),

$$\int_S \operatorname{curl} \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{S} = \int_S \boldsymbol{J} \cdot \mathrm{d}\boldsymbol{S}. \tag{5.24}$$

Therefore, for an elementary area $\mathrm{d}\boldsymbol{S}$, we find

$$\operatorname{curl} \boldsymbol{H} = \boldsymbol{J}. \tag{5.25}$$

We have already described Maxwell's reasoning in deriving $\operatorname{div} \boldsymbol{B} = 0$ and, in the same way, he concluded that $\operatorname{div} \boldsymbol{E} = 0$ in free space in the absence of electric charges. He knew, however, from Poisson's equation for the electrostatic potential (5.5) that, in the presence of a charge density distribution ρ_e, this equation must become

$$\operatorname{div} \boldsymbol{E} = \frac{\rho_e}{\epsilon_0}. \tag{5.26}$$

It is convenient to gather together this *primitive* and *incomplete* set of Maxwell's equations:

$$\begin{cases} \operatorname{curl} \boldsymbol{E} &= -\dfrac{\partial \boldsymbol{B}}{\partial t}, \\ \operatorname{curl} \boldsymbol{H} &= \boldsymbol{J}, \\ \operatorname{div} \epsilon_0 \boldsymbol{E} &= \rho_e, \\ \operatorname{div} \boldsymbol{B} &= 0. \end{cases} \tag{5.27}$$

The final achievement of this paper was the introduction of the *vector potential* $\boldsymbol{A}$. Such a vector had already been introduced by Neumann, Weber and Kirchhoff in order to calculate induced currents,

$$\boldsymbol{B} = \operatorname{curl} \boldsymbol{A}. \tag{5.28}$$

This definition is clearly consistent with (5.17), since $\operatorname{div} \operatorname{curl} \boldsymbol{A} = 0$. Maxwell went further and showed how the induced electric field $\boldsymbol{E}$ could be related to $\boldsymbol{A}$. Incorporating the definition (5.28) into (5.21), we obtain

$$\operatorname{curl} \boldsymbol{E} = -\frac{\partial}{\partial t}(\operatorname{curl} \boldsymbol{A}). \tag{5.29}$$

Interchanging the order of the time and spatial derivatives on the right-hand side, we find

$$\boldsymbol{E} = -\frac{\partial \boldsymbol{A}}{\partial t}. \tag{5.30}$$

5.3.2 *On Physical Lines of Force (1861–62)*

The analyses of Section 5.3.1 gave formal coherence to the theory, but Maxwell still lacked a physical model for electromagnetic phenomena. He developed his solution in 1861–62 in a remarkable series of papers entitled *On Physical Lines of Force*.[18] Since his earlier work on the analogy between $\boldsymbol{u}$ and $\boldsymbol{B}$, he had become convinced that magnetism was essentially rotational in nature. His aim was to devise a model for the medium filling all space which could account for the stresses that Faraday had associated with magnetic lines of force – in other words, a mechanical model for the *aether*, which was assumed to be the medium though which light was propagated. In his intriguing book *Innovation in Maxwell's Electrodynamics*, David Siegel demonstrated vividly the richness of Maxwell's insights in drawing physical analogies between mechanical and electromagnetic phenomena.[19]

The model was based upon the analogy between a rotating vortex tube and a tube of magnetic flux. In fact, as shown in Section 9.7.6, the exact analogy is between the magnetic flux density $\boldsymbol{B}$ and the vorticity $\boldsymbol{\omega} = \text{curl}\,\boldsymbol{u}$, that is, between the magnetic field lines and the vortex lines. If left on their own, magnetic field lines expand apart, exactly as occurs in the case of a fluid vortex tube, if the rotational centrifugal forces are not balanced. In addition, the rotational kinetic energy of the vortices can be written

$$\int_v \rho \boldsymbol{u}^2 \, \mathrm{d}v, \tag{5.31}$$

where ρ is the density of the fluid and $\boldsymbol{u}$ its rotational velocity. This expression is formally identical to that for the energy contained in a magnetic field distribution, $\int_v (\boldsymbol{B}^2/2\mu_0)\, \mathrm{d}v$. Maxwell postulated that everywhere the local magnetic flux density is proportional to the vorticity of the vortex tube, so that the angular momentum vector $\boldsymbol{L}$ is parallel to the axis of the vortex, and also to the magnetic flux density vector $\boldsymbol{B}$.

Maxwell therefore began with a model in which all space is filled with rotating vortex tubes (Fig. 5.7(*a*)). There is, however, a mechanical problem. Friction between neighbouring vortices would lead to their dissipation. Maxwell adopted the practical engineering solution of inserting 'idle wheels', or 'ball-bearings', between the vortices so that they could all rotate in the same direction without friction (Fig. 5.7(*b*)). Maxwell's published picture of the vortices, represented by an array of rotating hexagons, is shown

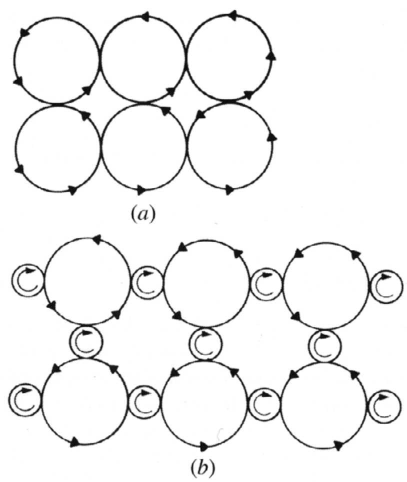


Fig. 5.7 (*a*) Maxwell's original model of rotating vortices as a representation of a magnetic field. Friction at the points where the vortices touch would lead to dissipation of the rotational energy of the tubes. (*b*) Maxwell's model with 'idle wheels' or 'ball-bearings' which prevent dissipation of the rotational energy of the vortices. If these particles are free to move, they are identified with the particles which carry an electric current in a conductor.

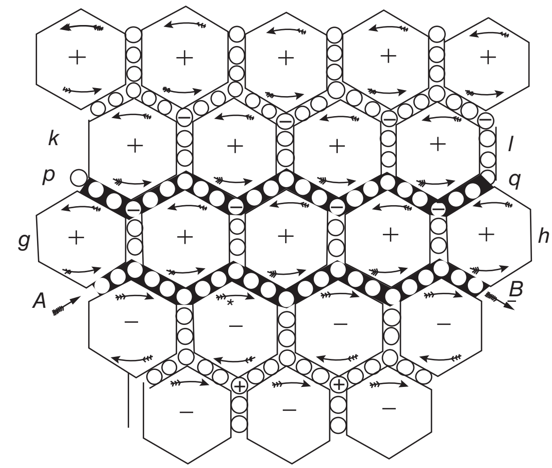


Fig. 5.8 Maxwell's picture of the dynamic interaction of the vortices (represented by hexagons) and the current-carrying particles (From *Philosophical Magazine*, 1861, Series 4, Vol. 21, Plate V, Fig. 2). A current of particles flows to the right along the shaded paths, for example from A to B. Their motion causes the upper row of vortices to rotate anticlockwise and the lower rows to rotate clockwise.

in Fig. 5.8. He then identified the idle wheels with electric particles which, if free to move, would carry an electric current as in a conductor. In insulators, including free space, they would not be free to move and so could not carry an electric current. I have no doubt that this example of the technique of working by analogy was a contributory cause of the 'feelings of discomfort, and often even of distrust' to which Poincaré alludes on first encountering Maxwell's works (see Section 1.4).

Remarkably, this mechanical model for the aether could account for all known phenomena of electromagnetism. For example, consider the magnetic field produced by an electric current flowing in a wire. The current causes the vortices to rotate as an infinite series of vortex rings about the wire as illustrated in Fig. 5.9 – the magnetic field lines form circular closed loops about the wire. Another example is the interface between two regions in which the magnetic flux densities are different but parallel. In the region of the stronger field, the vortices rotate with greater angular velocity and hence, in the interface, there must be a net force on the electric particles which drags them along the interface, as shown in Fig. 5.10, causing an electric current to flow. Notice that the sense of the current in the interface agrees with what is found experimentally.

As an example of induction, consider the effect of embedding a second wire in the magnetic field, as shown in Fig. 5.11. If the current is steady, there is no change in the current in the second wire. If, however, the current changes, an impulse is communicated through the intervening idle wheels and vortices and a reverse current is induced in the second wire.

Paper III of the series contains the flash of genius which led to the discovery of the complete set of Maxwell's equations. He now considered how insulators store electrical energy. He made the assumption that, in insulators, the idle wheels, or electric particles, can be displaced from their equilibrium positions by the action of an electric field and

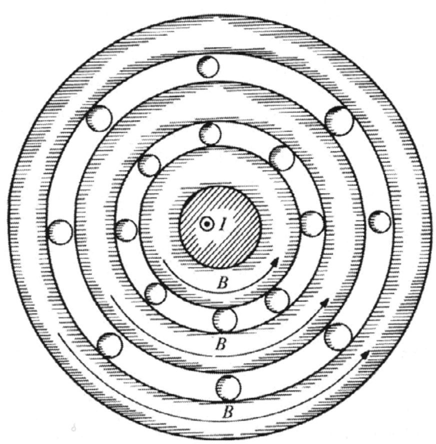


Fig. 5.9 A representation of the magnetic field about a current carrying wire according to Maxwell's model. The vortices become circular tori concentric with the axis of the wire.

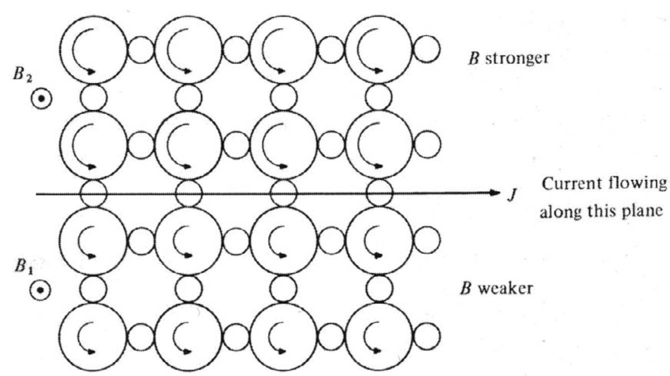


Fig. 5.10 Illustrating how flow along a current sheet ***J*** must cause a discontinuity in magnetic field strength according to Maxwell's model. The current-carrying particles flow along the line indicated. Because of friction, the lower set of vortices is slowed down whilst the upper set is speeded up. The sense of the discontinuity in ***B*** is seen to be in the correct sense according to Maxwell's equations.

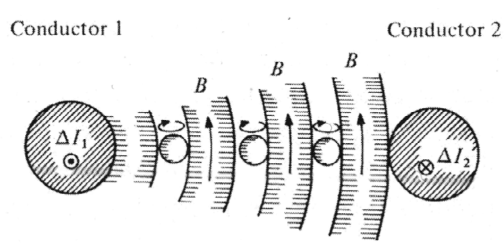


Fig. 5.11 Illustrating the phenomenon of induction according to Maxwell's model. The induced current in conductor 2, ΔI_2, is in the opposite direction to ΔI_1 as can be seen from the sense of rotation of the current carrying particles.

attributed the electrostatic energy in the medium to the elastic potential energy associated with the displacement of the electric particles. This had two immediate consequences. First, when the applied electric field is varied, there are small changes in the positions of the electric particles in the insulating medium, or vacuum, and so there are small electric currents associated with this motion. In other words, there is a current associated with the *displacement* of the electric particles from their equilibrium positions. Second, by virtue of the electric particles being bound elastically, any disturbance results in the propagation of waves though the medium. Maxwell could then carry out a straightforward calculation to find the speed at which disturbances can be propagated through the insulator, or a vacuum.

In a linear elastic medium, the displacement of the electric particles is proportional to the electric field strength

$$\boldsymbol{r} = \alpha \boldsymbol{E}. \tag{5.32}$$

When the strength of the field varies, the charges move, causing what Maxwell called a *displacement current*. If N_q is the number density of electric particles and q the charge on each of them, the *displacement current density* is

$$\boldsymbol{J}_d = qN_q\dot{\boldsymbol{r}} = qN_q\alpha\dot{\boldsymbol{E}} = \beta\dot{\boldsymbol{E}}. \tag{5.33}$$

Maxwell argued that this displacement current density should be included in (5.25), which then reads

$$\text{curl}\,\boldsymbol{H} = \boldsymbol{J} + \boldsymbol{J}_d = \boldsymbol{J} + \beta\dot{\boldsymbol{E}}. \tag{5.34}$$

The unknown constants α and β are to be determined from the known electric and magnetic properties of the medium.

First of all, let us work out the speed of propagation of a disturbance through the medium. Assuming that there are no currents, $\boldsymbol{J} = 0$, (5.21) and (5.34) reduce to

$$\begin{cases} \text{curl}\,\boldsymbol{E} = -\dot{\boldsymbol{B}}, \\ \text{curl}\,\boldsymbol{H} = \beta\dot{\boldsymbol{E}}. \end{cases} \tag{5.35}$$

The *dispersion relation* for these waves, that is, the relation between the wave vector $\boldsymbol{k}$ and the angular frequency ω, can be found by seeking wave solutions of the form $e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}$, and so replacing the vector operators by scalar and vector products according to the prescription

$$\begin{cases} \text{curl} \rightarrow \text{i}\boldsymbol{k}\times, \\ \partial/\partial t \rightarrow -\text{i}\omega, \end{cases}$$

(see Appendix A5.6). Then equations (5.35) reduce to

$$\begin{cases} \text{i}(\boldsymbol{k}\times\boldsymbol{E}) = \text{i}\omega\boldsymbol{B}, \\ \text{i}(\boldsymbol{k}\times\boldsymbol{H}) = -\text{i}\omega\beta\boldsymbol{E}. \end{cases} \tag{5.36}$$

Eliminating $\boldsymbol{E}$ from equations (5.36),

$$\boldsymbol{k}\times(\boldsymbol{k}\times\boldsymbol{H}) = -\omega^2\beta\mu\mu_0\boldsymbol{H}, \tag{5.37}$$

where we have used the linear constitutive relation $\boldsymbol{B} = \mu\mu_0\boldsymbol{H}$, μ being the permeability of the medium. Using the vector relation $\boldsymbol{A} \times (\boldsymbol{B} \times \boldsymbol{C}) = \boldsymbol{B}(\boldsymbol{A} \cdot \boldsymbol{C}) - \boldsymbol{C}(\boldsymbol{A} \cdot \boldsymbol{B})$, we find

$$\boldsymbol{k}(\boldsymbol{k} \cdot \boldsymbol{H}) - \boldsymbol{H}(\boldsymbol{k} \cdot \boldsymbol{k}) = \boldsymbol{k}(\boldsymbol{k} \cdot \boldsymbol{H}) = -\omega^2\beta\mu\mu_0\boldsymbol{H}. \tag{5.38}$$

There is no solution for $\boldsymbol{k}$ parallel to $\boldsymbol{H}$, that is, for *longitudinal waves*, since the left-hand side of (5.37) is then zero. There exist solutions, however, for *transverse waves*, for which $\boldsymbol{k} \cdot \boldsymbol{H} = |\boldsymbol{k}||\boldsymbol{H}|$. These represent plane transverse waves with the $\boldsymbol{E}$ and $\boldsymbol{H}$ vectors perpendicular to each other and to the direction of propagation of the wave. The dispersion relation for the waves is thus $k^2 = \omega^2\beta\mu\mu_0$. Since the speed of propagation of the wave is $c = \omega/k$, we find

$$c^2 = 1/\beta\mu\mu_0. \tag{5.39}$$

Notice that, because of the linear proportionality of k and ω, the phase velocity, c_{p}, and the group velocity of a wave packet, $c_{\mathrm{g}} = \mathrm{d}\omega/\mathrm{d}k$, both have the same value, c, given by (5.39).

Maxwell next evaluated the constant β. The energy density stored in the dielectric is just the work done per unit volume in displacing the electric particles a distance $\boldsymbol{r}$, that is,

$$\text{Work done} = \int \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{r} = \int N_{\mathrm{q}} q\boldsymbol{E} \cdot \mathrm{d}\boldsymbol{r}. \tag{5.40}$$

But, since $\boldsymbol{r} = \alpha\boldsymbol{E}$,

$$\mathrm{d}\boldsymbol{r} = \alpha\, \mathrm{d}\boldsymbol{E}, \tag{5.41}$$

and the work done is

$$\int_0^E N_{\mathrm{q}} q\alpha E\, \mathrm{d}E = \tfrac{1}{2}\alpha N_{\mathrm{q}} q E^2 = \tfrac{1}{2}\beta E^2. \tag{5.42}$$

But this is equal to the electrostatic energy density in the dielectric, $\frac{1}{2}\boldsymbol{D} \cdot \boldsymbol{E} = \frac{1}{2}\epsilon\epsilon_0 E^2$, where ϵ is the permittivity of the medium. Therefore $\beta = \epsilon\epsilon_0$. Inserting this value into (5.39) for the speed of the waves, we find

$$c = (\mu\mu_0\epsilon\epsilon_0)^{-1/2}. \tag{5.43}$$

In the case of a vacuum for which $\mu = 1, \epsilon = 1$, the speed of propagation of the waves is finite, $c = (\mu_0\epsilon_0)^{-1/2}$. Maxwell used Weber and Kohlrausch's experimental values for the product $\epsilon_0\mu_0$ and found, to his amazement, that c was almost exactly the speed of light. In Maxwell's letters of 1861 to Michael Faraday and William Thomson, he showed that the values agreed within about 1%. In Maxwell's own words, with his own emphasis:

> The velocity of transverse modulations in our hypothetical medium, calculated from the electro-magnetic experiments of MM. Kohlrausch and Weber, agrees so exactly with the velocity of light calculated from the optical experiments of M. Fizeau that we can scarcely avoid the inference that *light consists in the transverse modulations of the same medium which is the cause of electric and magnetic phenomena*.[20]

This remarkable calculation represented the unification of light with electricity and magnetism.

In vector form, Maxwell's equations now read

$$\begin{cases} \operatorname{curl} \boldsymbol{E} &= -\dfrac{\partial \boldsymbol{B}}{\partial t}, \\ \operatorname{curl} \boldsymbol{H} &= \boldsymbol{J} + \dfrac{\partial \boldsymbol{D}}{\partial t}, \\ \operatorname{div} \boldsymbol{D} &= \rho_{\mathrm{e}}, \\ \operatorname{div} \boldsymbol{B} &= 0. \end{cases} \tag{5.44}$$

Notice how the inclusion of what is still called the *displacement current* term, $\partial \boldsymbol{D}/\partial t$, resolves a problem with the equation of continuity in current electricity. Let us take the divergence of the second equation of (5.44). Then, because $\operatorname{div} \operatorname{curl} \boldsymbol{H} = 0$,

$$\operatorname{div}(\operatorname{curl} \boldsymbol{H}) = \operatorname{div} \boldsymbol{J} + \frac{\partial}{\partial t}(\operatorname{div} \boldsymbol{D}) = 0. \tag{5.45}$$

Since $\operatorname{div} \boldsymbol{D} = \rho_{\mathrm{e}}$, we obtain

$$\operatorname{div} \boldsymbol{J} + \frac{\partial \rho_{\mathrm{e}}}{\partial t} = 0, \tag{5.46}$$

which is the continuity equation for conservation of electric charge (see Section 7.2). Without the displacement term, the continuity equation would be nonsense and the primitive set of equations (5.27) would not be self-consistent.

One cannot but wonder that such pure gold came out of a specific mechanical model for the electromagnetic field in a vacuum. Maxwell was quite clear about the significance of the model:

> The conception of a particle having its motion connected with that of a vortex by perfect rolling contact may appear somewhat awkward. I do not bring it forward as a mode of connection existing in Nature ... It is however a mode of connection which is mechanically conceivable and it serves to bring out the actual mechanical connections between known electromagnetic phenomena.[21]

Maxwell was well aware of the fact that the theory was 'awkward' and considered it only a 'provisional and temporary hypothesis'.[22] He therefore recast the theory on a much more abstract basis without any special assumptions about the nature of the medium through which electromagnetic phenomena are propagated. In 1865, he published a paper entitled *A Dynamical Theory of the Electromagnetic Field* in which he developed the theory far beyond the mechanical analogues of the papers of 1861–62 and set out clearly the fundamental role which fields play in electromagnetism.[23] The analysis of this epochal paper is the subject of Chapter 6.

Appendix to Chapter 5: Notes on Vector Fields

These notes are intended to reinforce what you have already learned about vector fields and to summarise some useful results which are used in Chapters 6 and 7.

A5.1 The Divergence Theorem and Stokes' Theorem

The significance of these theorems is that we can either express the laws of physics in 'large-scale' form in terms of integrals over finite volumes of space or in 'small-scale' form in terms of the properties of the field in the vicinity of a particular point. The first form leads to integral equations and the second to differential equations. The fundamental theorems relating these two forms are as follows.

Divergence Theorem

$$\underbrace{\int_S \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{S}}_{\text{large-scale}} = \int_v \underbrace{\operatorname{div} \boldsymbol{A}}_{\text{small-scale}} \, \mathrm{d}v. \tag{A5.1}$$

d$\boldsymbol{S}$ is the element of surface area of the surface S which encloses volume v. The direction of the vector d$\boldsymbol{S}$ is always taken to be directed *normally outwards* to the element of surface. The volume integral on the right-hand side is over the volume enclosed by the closed surface S.

Stokes' Theorem

$$\underbrace{\int_C \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{l}}_{\text{large-scale}} = \int_S \underbrace{\operatorname{curl} \boldsymbol{A}}_{\text{small-scale}} \cdot \mathrm{d}\boldsymbol{S}, \tag{A5.2}$$

where C is a closed curve and the integral on the left is taken round the closed curve. S is *any* open surface bounded by the loop C, d$\boldsymbol{S}$ being the element of surface area. The sign of d$\boldsymbol{S}$ is decided by the right-hand corkscrew convention. Notice that if $\boldsymbol{A}$ is a vector field of force, $\int \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{l}$ is *minus* the work needed to take a particle once round the circuit.

A5.2 Results Related to the Divergence Theorem

Problem 1

By taking $\boldsymbol{A} = f \operatorname{grad} g$, derive two forms of *Green's theorem*:

$$\int_v [f\nabla^2 g + (\nabla f)\cdot(\nabla g)]\,\mathrm{d}v = \int_S f\nabla g \cdot \mathrm{d}\boldsymbol{S}, \tag{A5.3}$$

$$\int_v [f\nabla^2 g - g\nabla^2 f]\,\mathrm{d}v = \int_S (f\nabla g - g\nabla f)\cdot \mathrm{d}\boldsymbol{S}. \tag{A5.4}$$

Solution

Substitute $\boldsymbol{A} = f \operatorname{grad} g$ into the divergence theorem (A5.1):

$$\int_S f\nabla g \cdot \mathrm{d}\boldsymbol{S} = \int_v \operatorname{div}(f\nabla g)\,\mathrm{d}v.$$

Because $\nabla \cdot (a\boldsymbol{b}) = a\nabla \cdot \boldsymbol{b} + \nabla a \cdot \boldsymbol{b}$, we obtain (A5.3),

$$\int_S f\nabla g \cdot \mathrm{d}\boldsymbol{S} = \int_v (\nabla f \cdot \nabla g + f\nabla^2 g)\,\mathrm{d}v. \tag{A5.5}$$

Performing the same analysis for $g\nabla f$, we find

$$\int_S g\nabla f \cdot \mathrm{d}\boldsymbol{S} = \int_v (\nabla f \cdot \nabla g + g\nabla^2 f)\,\mathrm{d}v. \tag{A5.6}$$

Subtracting (A5.6) from (A5.5), we obtain the desired result (A5.4),

$$\int_v (f\nabla^2 g - g\nabla^2 f)\,\mathrm{d}v = \int_S (f\nabla g - g\nabla f) \cdot \mathrm{d}\boldsymbol{S}. \tag{A5.7}$$

This result is particularly useful if one of the functions is a solution of Laplace's equation $\nabla^2 f = 0$.

Problem 2

Show that

$$\int_v \frac{\partial f}{\partial x_k}\,\mathrm{d}v = \int_S f\,\mathrm{d}S_k. \tag{A5.8}$$

Solution

Take the vector $\boldsymbol{A}$ to be $f\boldsymbol{i}_k$ where f is a scalar function of position and $\boldsymbol{i}_k$ is the unit vector in some direction. Substituting into the divergence theorem,

$$\int_S f\,\boldsymbol{i}_k \cdot \mathrm{d}\boldsymbol{S} = \int_v \operatorname{div} f\,\boldsymbol{i}_k\,\mathrm{d}v.$$

The divergence on the right-hand side is $\partial f/\partial x_k$ and so

$$\int_S f\,\mathrm{d}S_k = \int_v \frac{\partial f}{\partial x_k}\,\mathrm{d}v.$$

This is the divergence theorem projected in the $\boldsymbol{i}_k$ direction.

Problem 3

Show that

$$\int_v (\nabla \times \boldsymbol{A})\,\mathrm{d}v = \int_S \mathrm{d}\boldsymbol{S} \times \boldsymbol{A}. \tag{A5.9}$$

Solution

Apply the result (A5.8) to the scalar quantities A_x and A_y, the components of the vector $\boldsymbol{A}$. Then

$$\int_v \frac{\partial A_y}{\partial x}\,\mathrm{d}v = \int_S A_y\,\mathrm{d}S_x, \tag{A5.10}$$

$$\int_v \frac{\partial A_x}{\partial y}\,\mathrm{d}v = \int_S A_x\,\mathrm{d}S_y. \tag{A5.11}$$

Subtracting (A5.11) from (A5.10),

$$\int_v \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right)\mathrm{d}v = \int_S (A_y\,\mathrm{d}S_x - A_x\,\mathrm{d}S_y),$$

that is,

$$\int_v \mathrm{curl}\,A_z\,\mathrm{d}v = -\int_S (\boldsymbol{A}\times \mathrm{d}\boldsymbol{S})_z.$$

This must be true separately for all three directions x, y and z and so we can summarise these results in the single equation

$$\int_v \mathrm{curl}\,\boldsymbol{A}\,\mathrm{d}v = \int_S (\mathrm{d}\boldsymbol{S}\times\boldsymbol{A}).$$

A5.3 Results Related to Stokes' Theorem

Problem

Show that an alternative way of writing Stokes' theorem is

$$\int_C f\,\mathrm{d}\boldsymbol{l} = \int_S \mathrm{d}\boldsymbol{S}\times \mathrm{grad}\,f, \tag{A5.12}$$

where f is a scalar function.

Solution

Write $\boldsymbol{A} = f\boldsymbol{i}_k$ where $\boldsymbol{i}_k$ is the unit vector in the k direction. Then Stokes' theorem,

$$\int_C \boldsymbol{A}\cdot \mathrm{d}\boldsymbol{l} = \int_S \mathrm{curl}\boldsymbol{A}\cdot \mathrm{d}\boldsymbol{S},$$

becomes

$$\boldsymbol{i}_k\cdot \int_C f\,\mathrm{d}\boldsymbol{l} = \int_S \mathrm{curl}\,f\boldsymbol{i}_k\cdot \mathrm{d}\boldsymbol{S}.$$

We now need the expansion of curl $\boldsymbol{A}$ where $\boldsymbol{A}$ is the product of a scalar and a vector.

$$\mathrm{curl}\,f\boldsymbol{g} = f\,\mathrm{curl}\,\boldsymbol{g} + (\mathrm{grad}\,f)\times\boldsymbol{g}. \tag{A5.13}$$

In the present case, $\boldsymbol{g}$ is the unit vector $\boldsymbol{i}_k$ and hence curl $\boldsymbol{i}_k = 0$. Therefore,

$$\boldsymbol{i}_k \cdot \int_C f \,\mathrm{d}\boldsymbol{l} = \int_S [(\mathrm{grad}\, f) \times \boldsymbol{i}_k] \cdot \mathrm{d}\boldsymbol{S}$$
$$= \boldsymbol{i}_k \cdot \int_S \mathrm{d}\boldsymbol{S} \times \mathrm{grad}\, f,$$

that is,

$$\int_C f \,\mathrm{d}\boldsymbol{l} = \int_S \mathrm{d}\boldsymbol{S} \times \mathrm{grad}\, f.$$

A5.4 Vector Fields with Special Properties

A vector field $\boldsymbol{A}$ for which curl $\boldsymbol{A} = 0$ is called *irrotational, conservative* or a *conservative field of force*. More generally, a vector field is conservative if it satisfies any one of the following conditions, all of which are equivalent.

(a) $\boldsymbol{A}$ can be expressed in the form $\boldsymbol{A} = -\mathrm{grad}\,\phi$, where ϕ is a scalar function of position, that is, ϕ depends only upon x, y and z.
(b) curl $\boldsymbol{A} = 0$.
(c) $\int_C \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{l} = 0$.
(d) $\int_A^B \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{l}$ is independent of the path from A to B.

The conditions (b) and (c) are plainly identical because of Stokes' theorem

$$\int_C \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{l} = \int_S \mathrm{curl}\, \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{S} = 0.$$

If $\boldsymbol{A} = -\mathrm{grad}\,\phi$, then curl $\boldsymbol{A} = 0$. Finally, we can write

$$\int_A^B \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{l} = -\int_A^B \mathrm{grad}\,\phi \cdot \mathrm{d}\boldsymbol{l} = -(\phi_B - \phi_A),$$

and is independent of the path between A and B. This last property is the reason for the name 'conservative field'. It does not matter what route one chooses between A and B. If $\boldsymbol{A}$ is a field of force, ϕ is a potential, equal to the work which needs to be expended to bring unit mass, charge and so on to that point in the field.

A vector field $\boldsymbol{A}$ for which div $\boldsymbol{A} = \nabla \cdot \boldsymbol{A} = 0$ is called a *solenoidal field*. If $\boldsymbol{B}$ = curl $\boldsymbol{A}$, div $\boldsymbol{B} = 0$. Conversely, if div $\boldsymbol{B} = 0$, $\boldsymbol{B}$ can be expressed as the curl of some vector $\boldsymbol{A}$ in many different ways. This is because we can add to $\boldsymbol{A}$ an arbitrary conservative vector field which always vanishes when $\boldsymbol{A}$ is curled. Thus, if $\boldsymbol{A}' = \boldsymbol{A} - \mathrm{grad}\,\phi$,

$$\boldsymbol{B} = \mathrm{curl}\,\boldsymbol{A}' = \mathrm{curl}\,\boldsymbol{A} - \mathrm{curl}\,\mathrm{grad}\,\phi = \mathrm{curl}\,\boldsymbol{A}.$$

One of the most useful results in vector analysis is the identity

$$\nabla \times (\nabla \times \boldsymbol{A}) = \nabla(\nabla \cdot \boldsymbol{A}) - \nabla^2 \boldsymbol{A}. \qquad \text{(A5.14)}$$

A5.5 Vector Operators in Curvilinear Coordinates

It is useful to have at hand a list of the vector operators grad, div, curl and ∇^2 in rectangular (or Cartesian), cylindrical and spherical polar coordinates. The standard books give the following.

grad

Cartesian $$\operatorname{grad}\Phi = \nabla\Phi = \boldsymbol{i}_x\frac{\partial\Phi}{\partial x} + \boldsymbol{i}_y\frac{\partial\Phi}{\partial y} + \boldsymbol{i}_z\frac{\partial\Phi}{\partial z},$$

Cylindrical polar $$\operatorname{grad}\Phi = \nabla\Phi = \boldsymbol{i}_r\frac{\partial\Phi}{\partial r} + \boldsymbol{i}_z\frac{\partial\Phi}{\partial z} + \boldsymbol{i}_\phi\frac{1}{r}\frac{\partial\Phi}{\partial\phi},$$

Spherical polar $$\operatorname{grad}\Phi = \nabla\Phi = \boldsymbol{i}_r\frac{\partial\Phi}{\partial r} + \boldsymbol{i}_\theta\frac{1}{r}\frac{\partial\Phi}{\partial\theta} + \boldsymbol{i}_\phi\frac{1}{r\sin\theta}\frac{\partial\Phi}{\partial\phi}.$$

div

Cartesian $$\operatorname{div}\boldsymbol{A} = \nabla\cdot\boldsymbol{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z},$$

Cylindrical polar $$\operatorname{div}\boldsymbol{A} = \nabla\cdot\boldsymbol{A} = \frac{1}{r}\left[\frac{\partial}{\partial r}(rA_r) + \frac{\partial}{\partial z}(rA_z) + \frac{\partial}{\partial\phi}(A_\phi)\right],$$

Spherical polar $$\operatorname{div}\boldsymbol{A} = \nabla\cdot\boldsymbol{A} = \frac{1}{r^2\sin\theta}\left[\frac{\partial}{\partial r}(r^2\sin\theta\, A_r) + \frac{\partial}{\partial\theta}(r\sin\theta\, A_\theta) + \frac{\partial}{\partial\phi}(rA_\phi)\right].$$

curl

Cartesian $$\operatorname{curl}\boldsymbol{A} = \nabla\times\boldsymbol{A} = \left(\frac{\partial A_y}{\partial z} - \frac{\partial A_z}{\partial y}\right)\boldsymbol{i}_x + \left(\frac{\partial A_z}{\partial x} - \frac{\partial A_x}{\partial z}\right)\boldsymbol{i}_y + \left(\frac{\partial A_x}{\partial y} - \frac{\partial A_y}{\partial x}\right)\boldsymbol{i}_z.$$

Cylindrical polar $$\operatorname{curl}\boldsymbol{A} = \nabla\times\boldsymbol{A} = \frac{1}{r}\left[\frac{\partial}{\partial z}(rA_\phi) - \frac{\partial}{\partial\phi}(A_z)\right]\boldsymbol{i}_r + \frac{1}{r}\left[\frac{\partial}{\partial\phi}(A_r) - \frac{\partial}{\partial r}(rA_\phi)\right]\boldsymbol{i}_z + \left[\frac{\partial}{\partial r}(A_z) - \frac{\partial}{\partial z}(A_r)\right]\boldsymbol{i}_\phi.$$

Spherical polar $$\operatorname{curl} \boldsymbol{A} = \nabla \times \boldsymbol{A} = \frac{1}{r^2 \sin\theta}\left[\frac{\partial}{\partial\theta}(r\sin\theta\, A_\phi) - \frac{\partial}{\partial\phi}(rA_\theta)\right]\boldsymbol{i}_r + \frac{1}{r\sin\theta}\left[\frac{\partial}{\partial\phi}(A_r) - \frac{\partial}{\partial r}(r\sin\theta\, A_\phi)\right]\boldsymbol{i}_\theta + \frac{1}{r}\left[\frac{\partial}{\partial r}(rA_\theta) - \frac{\partial}{\partial\theta}(A_r)\right]\boldsymbol{i}_\phi .$$

Laplacian

Cartesian $$\nabla^2\Phi = \frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} + \frac{\partial^2\Phi}{\partial z^2},$$

Cylindrical polar $$\nabla^2\Phi = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\Phi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2\Phi}{\partial\phi^2} + \frac{\partial^2\Phi}{\partial z^2},$$

Spherical polar $$\nabla^2\Phi = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\Phi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Phi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\Phi}{\partial\phi^2},$$

$$= \frac{1}{r}\frac{\partial^2}{\partial r^2}(r\Phi) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Phi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\Phi}{\partial\phi^2}.$$

A5.6 Vector Operators and Dispersion Relations

In finding three-dimensional solutions of wave equations, such as Maxwell's equations, it often simplifies the working considerably to use the relations $\nabla \to \mathrm{i}\boldsymbol{k}$, $\partial/\partial t \to -\mathrm{i}\omega$, if the phase factor of the wave is written in the form $\exp[\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t)]$. The following relations can be easily proved using the scalar product $\boldsymbol{k}\cdot\boldsymbol{r} = k_x x + k_y y + k_z z$:

$$\nabla\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} = \operatorname{grad} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} = \mathrm{i}\boldsymbol{k}\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}},$$

$$\nabla\cdot(\boldsymbol{A}\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}) = \operatorname{div}(\boldsymbol{A}\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}) = \mathrm{i}\boldsymbol{k}\cdot\boldsymbol{A}\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}},$$

$$\nabla\times(\boldsymbol{A}\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}) = \operatorname{curl}(\boldsymbol{A}\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}) = \mathrm{i}[\boldsymbol{k}\times\boldsymbol{A}]\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}},$$

where $\boldsymbol{A}$ is a constant vector. Thus, it can be seen that $\nabla \to \mathrm{i}\boldsymbol{k}$ carries out the vector operations for wave solutions. In a similar way, it can be shown that $\partial/\partial t \to -\mathrm{i}\omega$.

When we attempt to solve a wave equation by this technique, we find that the exponential phase factor cancels out through the wave equation, resulting in a relation between $\boldsymbol{k}$ and ω which is known as a *dispersion relation*. Notice that the vectors $\boldsymbol{k}$ and $\boldsymbol{A}$ may themselves be complex. Here is an example.

A simple wave equation with a damping term γ can be written

$$\nabla^2 u - 2\gamma\frac{\partial u}{\partial t} - \frac{1}{v^2}\frac{\partial^2 u}{\partial t^2} = 0,$$

where v is the sound speed in the medium. Recalling that $\nabla^2 u = \nabla \cdot (\nabla u)$ and using the substitutions $\nabla \to \mathrm{i}\boldsymbol{k}$, $\partial/\partial t \to -\mathrm{i}\omega$, we find

$$(\mathrm{i}\boldsymbol{k}) \cdot (\mathrm{i}\boldsymbol{k}) + 2\mathrm{i}\gamma\omega + \frac{\omega^2}{v^2} = 0,$$

$$k^2 = \frac{\omega^2}{v^2} + 2\mathrm{i}\gamma\omega. \tag{A5.15}$$

If there is no damping, $\gamma = 0$, the phase velocity of the waves is $v = \omega/k_0$ and the group velocity, $v_\mathrm{g} = \mathrm{d}\omega/\mathrm{d}k$, is also equal to v.

If $\gamma \neq 0$, the dispersion relation has an imaginary part, corresponding to the damping of the waves. If the damping is small, $(\gamma v^2/\omega) \ll 1$, (A5.15) reduces to

$$k = \frac{\omega}{v}\left(1 + \mathrm{i}\frac{\gamma v^2}{\omega}\right) = k_0\left(1 + \mathrm{i}\frac{\gamma v^2}{\omega}\right),$$

and so the expression for the propagation of the wave becomes

$$u(\boldsymbol{r}, t) = u_0 \exp\left(-\gamma v r\right)\ \mathrm{e}^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t)} = u_0 \exp\left(-\alpha t\right)\ \mathrm{e}^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t)},$$

where $\alpha = \gamma v^2$.

A5.7 How to Relate the Different Expressions for the Magnetic Fields Produced by Currents

It is simplest to begin with Faraday's brilliant experiments of the 1820s (Section 5.2). Ørsted had shown that magnetic fields are produced by electric currents and Ampère that the magnetic field lines about a current-carrying wire are circular. Faraday reasoned that, if the wire were bent into a circle, the magnetic field lines would be more concentrated inside the loop and less concentrated outside, as illustrated in Fig. 5.4. In his great series of experiments, he showed that *current loops produce identical magnetostatic effects to bar magnets* (Fig. A5.1(*a*)).

We therefore formally identify the *magnetic dipole moment* $\boldsymbol{m}$ associated with the current I flowing in an elementary current loop of area $\mathrm{d}\boldsymbol{A}$ by the relation

$$\boldsymbol{m} = I\,\mathrm{d}\boldsymbol{A}. \tag{A5.16}$$

The direction of $\boldsymbol{m}$ is given by the right-hand corkscrew rule with the direction of rotation being given by the direction of the current flow (Fig. A5.1(*b*)). Therefore, the couple acting upon a current loop located in a magnetic field is

$$\boldsymbol{G} = \boldsymbol{m} \times \boldsymbol{B} = I\,(\mathrm{d}\boldsymbol{A} \times \boldsymbol{B}). \tag{A5.17}$$

Let us show how Ampère's circuital law can be derived from the definition of the magnetic moment of a current loop. There is one crucial and obvious difference between a magnet and a current loop – *you cannot pass through a bar magnet but you can pass through a current loop.* Suppose magnetic poles existed. If we were to release one in the magnetic field of a bar magnet, it would move along the field lines, crash into the magnet and stop there. In the case of the current loop, however, it would just keep on

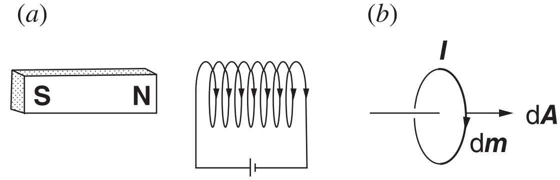


Fig. A5.1 (*a*) Illustrating the equivalence of a bar magnet and a current loop. (*b*) Illustrating the definition of the magnetic moment of a current loop.

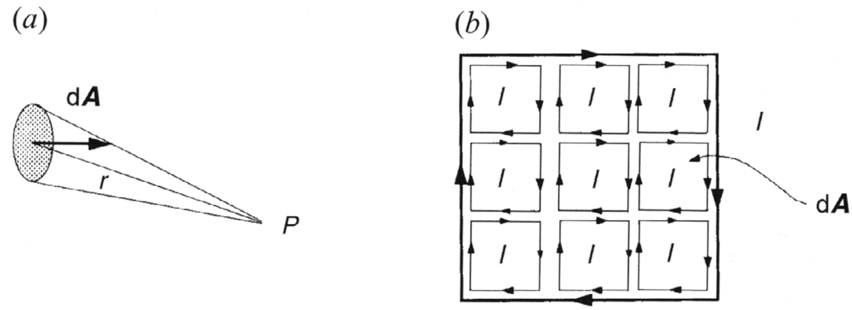


Fig. A5.2 (*a*) Evaluating the magnetic potential V_{mag} of an elementary current loop. (*b*) Illustrating the use of Ampère's magnetic shell to evaluate the magnetostatic potential of a finite current loop.

circulating round the same magnetic line of force gaining energy all the time. This does not happen because *there are no free magnetic poles*. If the experiment is repeated with a small magnetic dipole, it comes to rest inside the coil.

Let us work out the magnetic flux density in the vicinity of an elementary current loop d$\boldsymbol{A}$ (Fig. A5.2(*a*)). We can derive the magnetostatic potential V_{m} at any point P at distance $\boldsymbol{r}$ from the centre of the loop:

$$V_{\mathrm{m}} = \frac{\boldsymbol{m} \cdot \boldsymbol{i}_r}{4\pi r^2} = \frac{\boldsymbol{m} \cdot \boldsymbol{r}}{4\pi r^3}.$$

But $\boldsymbol{m} = I\,\mathrm{d}\boldsymbol{A}$ and so

$$V_{\mathrm{m}} = I\,\frac{\mathrm{d}\boldsymbol{A} \cdot \boldsymbol{i}_r}{4\pi r^2}.$$

Now, $\mathrm{d}\boldsymbol{A} \cdot \boldsymbol{i}_r/r^2 = \mathrm{d}\Omega$ is the solid angle subtended by the current loop at the point P. Therefore,

$$V_{\mathrm{m}} = \frac{I\,\mathrm{d}\Omega}{4\pi}. \tag{A5.18}$$

Now, let us extend this result to a finite sized loop. It is easiest to split up the loop into many elementary loops, each carrying the current I (Fig. A5.1(*b*)) – this is the procedure described by Ampère in his treatise of 1825. Clearly, in the interior, the oppositely flowing currents cancel out leaving only the current I flowing around the outside loop. Therefore,

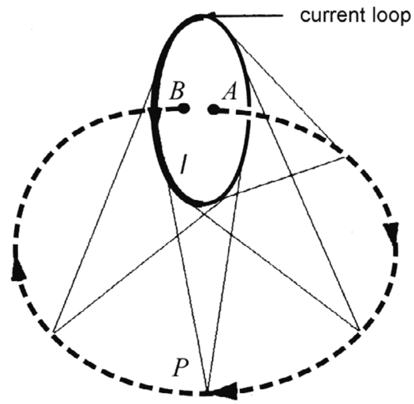


Fig. A5.3 Evaluating the change in magnetic potential in passing from one side of a current loop to the other along the path shown.

the magnetostatic potential at a point P is just the sum of the magnetostatic potentials due to all the elementary current loops:

$$V_m = \sum \frac{I\,d\Omega}{4\pi} = \frac{I\Omega}{4\pi}.$$

This is a remarkable simplification. We can now work out the magnetostatic potential V_m at any point in space and find the magnetic flux density from

$$V_m = \frac{I\Omega}{4\pi}; \quad \boldsymbol{B} = -\mu_0 \,\mathrm{grad}\, V_m,$$

or, integrating,

$$\int_A^B \boldsymbol{B} \cdot d\boldsymbol{s} = -\mu_0 \int_A^B \mathrm{grad}\, V_m \cdot d\boldsymbol{s} = -\mu_0[V_m(B) - V_m(A)].$$

Now, let us inspect what happens when we take the line integral from one side of a current loop to another. The path is shown in Fig. A5.3. The path starts at A, which is very close to the current loop. Therefore, the solid angle subtended by the loop is almost 2π sr. As the point moves round the path, we see that the solid angle decreases and, at the point P, the solid angle decreases to zero. Moving further round the loop, the solid angle becomes negative until it becomes almost -2π when it reaches B. Thus, the change in solid angle in passing from A to B is -4π sr as $B \to A$. Therefore,

$$V_m(B) - V_m(A) = \frac{I\Delta\Omega}{4\pi} = -I.$$

If this were a magnet, this change in potential would be exactly compensated by the reverse field inside the magnet and that would be the end of the story. In the case of the current loop, however, we simply pass through the loop to the point A where we started and we find

$$\oint_{\mathrm{A}}^{\mathrm{B}\to\mathrm{A}} \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{s} = -\mu_0[V_{\mathrm{m}}(\mathrm{B}) - V_{\mathrm{m}}(\mathrm{A})] = \mu_0 I. \qquad (\mathrm{A}5.19)$$

This is *Ampère's circuital theorem*.

Ampère's circuital theorem is an integral equation relating the integral of the magnetic flux density round a closed loop to the enclosed current. Next we work out the magnetic flux density at any point in space due to the current flowing in the line element $d\boldsymbol{s}$, the *Biot–Savart Law* (5.8), which is the most general form of expression for the magnetic flux density produced by a current element $I\,\mathrm{d}\boldsymbol{s}$.

We begin with the expression for the magnetic flux density in terms of the magnetostatic potential:

$$\boldsymbol{B} = -\mu_0 \,\mathrm{grad}\, V_{\mathrm{m}}.$$

We showed above that, if the current loop subtends a solid angle Ω, the magnetostatic potential is $V_{\mathrm{m}} = I\Omega/4\pi$. Then,

$$\boldsymbol{B} = -\frac{\mu_0 I}{4\pi}\,\mathrm{grad}\,\Omega.$$

Thus, the problem reduces to working out grad Ω. In other words, how does Ω change when we move a vector distance $d\boldsymbol{l}$ from P as illustrated in Fig. A5.4(a)? From the geometry of the diagram, we can write

$$\mathrm{d}\Omega = (\mathrm{grad}\,\Omega) \cdot \mathrm{d}\boldsymbol{l}.$$

It is simplest to shift the contour by $d\boldsymbol{l}$ in the opposite direction and work out how the projected surface area changes as observed from P (Fig. A5.4(b)).

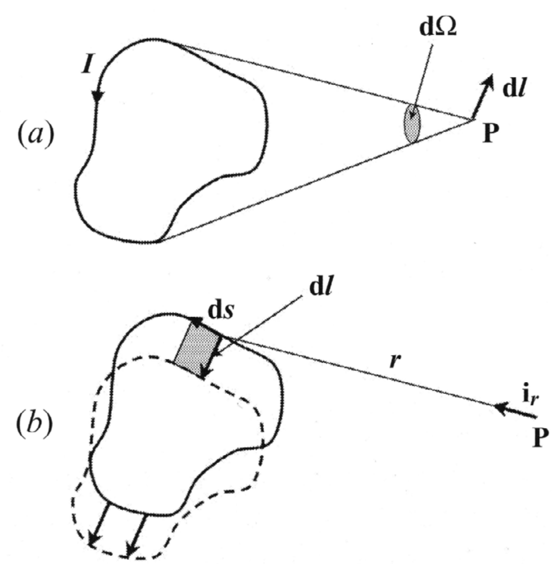


Fig. A5.4 Evaluating the change in magnetic potential on moving a distance ***dl*** from an arbitrarily shaped current loop.

d$\boldsymbol{s}$ is an element of the current loop and so the area formed by the parallelogram defined by d$\boldsymbol{s}$ and d$\boldsymbol{l}$ is

$$\mathrm{d}\boldsymbol{A} = \mathrm{d}\boldsymbol{s} \times \mathrm{d}\boldsymbol{l}.$$

If $\boldsymbol{i}_r$ is the unit vector from P to the area $d\boldsymbol{A}$, the projected area as observed from P is

$$\mathrm{d}A_{\text{proj}} = \boldsymbol{i}_r \cdot (\mathrm{d}\boldsymbol{s} \times \mathrm{d}\boldsymbol{l}).$$

The change in the element of solid angle is therefore $\mathrm{d}A_{\text{proj}}/r^2$ and so we find the total change in solid angle by integrating round the complete current loop:

$$\mathrm{d}\Omega = \oint_C \frac{\boldsymbol{i}_r \cdot (\mathrm{d}\boldsymbol{s} \times \mathrm{d}\boldsymbol{l})}{r^2} = \oint_C \frac{\boldsymbol{r} \cdot (\mathrm{d}\boldsymbol{s} \times \mathrm{d}\boldsymbol{l})}{r^3}.$$

We recall the vector identity

$$\boldsymbol{A} \cdot (\boldsymbol{B} \times \boldsymbol{C}) = \boldsymbol{C} \cdot (\boldsymbol{A} \times \boldsymbol{B}),$$

and hence

$$\mathrm{d}\Omega = \oint_C \frac{\boldsymbol{r} \cdot (\mathrm{d}\boldsymbol{s} \times \mathrm{d}\boldsymbol{l})}{r^3} = \oint_C \frac{\mathrm{d}\boldsymbol{l} \cdot (\boldsymbol{r} \times \mathrm{d}\boldsymbol{s})}{r^3} = \mathrm{d}\boldsymbol{l} \cdot \oint \frac{\boldsymbol{r} \times \mathrm{d}\boldsymbol{s}}{r^3},$$

since d$\boldsymbol{l}$ is a constant vector all round the current loop. Therefore, since

$$\mathrm{d}\Omega = (\text{grad}\,\Omega) \cdot \mathrm{d}\boldsymbol{l} = \mathrm{d}\boldsymbol{l} \cdot \oint \frac{\boldsymbol{r} \times \mathrm{d}\boldsymbol{s}}{r^3},$$

we find

$$\text{grad}\,\Omega = \oint \frac{(\boldsymbol{r} \times \mathrm{d}\boldsymbol{s})}{r^3}.$$

Therefore, since

$$\boldsymbol{B} = -\frac{\mu_0 I}{4\pi}\,\text{grad}\,\Omega,$$

we obtain the expression

$$\boldsymbol{B} = -\frac{\mu_0 I}{4\pi} \oint \frac{\boldsymbol{r} \times \mathrm{d}\boldsymbol{s}}{r^3}.$$

This integral describes the total contribution to the field at P due to all the increments of path length round the circuit and so we can remove the integral sign to find the contribution from the current element $d\boldsymbol{s}$:

$$\mathrm{d}\boldsymbol{B} = \frac{\mu_0 I}{4\pi} \frac{\mathrm{d}\boldsymbol{s} \times \boldsymbol{r}}{r^3}. \qquad \text{(A5.20)}$$

This is the proof of the *Biot–Savart Law*. Notice, however, that Biot and Savart deduced the law *experimentally*.

We have therefore four different ways of working out the magnetic field associated with currents:

(1) The magnetic moment of the current loop

$$\boldsymbol{m} = I\,\mathrm{d}\boldsymbol{A}.$$

(2) The gradient of the magnetostatic potential

$$\boldsymbol{B} = -\mu_0 \,\mathrm{grad}\, V_{\mathrm{m}}.$$

(3) Ampère's circuital theorem

$$\oint_{\mathrm{C}} \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{s} = \mu_0 I_{\mathrm{enclosed}}.$$

(4) The Biot–Savart law

$$\mathrm{d}\boldsymbol{B} = \frac{\mu_0 I}{4\pi} \frac{\mathrm{d}\boldsymbol{s} \times \boldsymbol{r}}{r^3}.$$

Notes

1 Maxwell included Henry Cavendish's results in his edition of Cavendish's papers in *The Electrical Researches of the Honourable Henry Cavendish*, Cambridge: Cambridge University Press. pp. 104–113 (1879). Cavendish found that the exponent was 2 ± 0.02. MacAlister's improved estimate appears as Note 19 in this volume.
2 Williams, E.R. *et al.* (1971). New Experimental Test of Coulomb's Law: A Laboratory Upper Limit on the Photon Rest Mass, *Physical Review Letters*, **26**, 721–724. doi: 10.1103/PhysRevLett.26.721.
3 Fulcher, L.P. (1986). Improved Result for the Accuracy of Coulomb's Law: A Review of the Williams, Faller, and Hill Experiment, *Physical Review A*, **33**, 759–761. doi: 10.1103/PhysRevA.33.759.
4 Purists might prefer the whole analysis of this chapter to be carried out in the original notation, but such adherence to historical authenticity might obscure the essence of the arguments for the modern reader.
5 The validity of Ohm's law was questioned throughout much of the nineteenth century and it does not form part of the standard presentation of Maxwell's equations. As Maxwell emphasised it 'must still be considered a purely empirical law, as it has not hitherto been deduced from the fundamental principles of dynamics.' The experimental validation of the law was provided by Maxwell and one of his early research students George Chrystal with high precision.
6 A delightful, short biography of Faraday has been written by J.M. Thomas (1991). *Michael Faraday and the Royal Institution: The Genius of Man and Place*. Bristol: Adam Hilger, IoP Publications.
7 See Harman, P.M. (1998). *The Natural Philosophy of James Clerk Maxwell*, p. 74. Cambridge: Cambridge University Press.
8 M. Faraday (1846). Thoughts on Ray Vibrations, *Philosophical Magazine*, S.3, **28**, 345–350.
9 Maxwell, J.C. (1865). In *The Scientific Papers of J. Clerk Maxwell* in two volumes (1890), ed. W.D. Niven. **1**, p. 466. Cambridge: Cambridge University Press. (Reprinted 1965. New York: Dover Publications Inc.)
10 Maxwell, J.C. (1873). *Treatise on Electricity and Magnetism*, **1**, ix. Oxford: Clarendon Press. (Reprint of 3rd edition, 1998. Oxford Classics series.)
11 Harman, P.M. (1990). *The Scientific Letters and Paper of James Clerk Maxwell*, **1**, p. 9. Cambridge: Cambridge University Press.
12 Harman, P.M. (1998). *The Natural Philosophy of James Clerk Maxwell*. Cambridge: Cambridge University Press.
13 Harman, P.M. (1990). *op. cit.*, p. 381.

14 Maxwell, J.C. (1856). *Scientific Papers*, ed. W.D. Niven. *op. cit.*, **1**, pp. 155–229.
15 Maxwell, J.C. (1873). *op. cit.*, viii.
16 Maxwell, J.C.(1870). *Scientific Papers*, ed. W.D. Niven. *op. cit.*, **2**, pp. 257–266. Maxwell invented the terms 'slope', now 'gradient', 'curl' and 'convergence', the opposite of divergence, to provide an intuitive feel for the physical meaning of these operators.
17 Maxwell, J.C. (1856). *Scientific Papers*, ed. W.D. Niven. *op. cit.*, **1**, pp. 206–207.
18 Maxwell, J.C. (1861–62). *Scientific Papers*, ed. W.D. Niven. *op. cit.*, **1**, pp. 459–512. This reference includes the four separate parts which appeared in the *Philosophical Magazine*. Part I. The Theory of Molecular Vortices applied to Magnetic Phenomena, **21**, 161–175 (1861); Part II. The Theory of Molecular Vortices applied to Electric Currents, **21**, 281–291 (1861); Part III. The Theory of Molecular Vortices applied to Statical Electricity, **23**, 12–24 (1862); Part IV. The Theory of Molecular Vortices applied to the Action of Magnetism on Polarised Light, **23**, 85–93 (1862).
19 Siegel, D.M. (1991). *Innovation in Maxwell's Electromagnetic Theory: Molecular Vortices, Displacement Current, and Light*. Cambridge: Cambridge University Press.
20 Maxwell, J.C. (1861–62). *Scientific Papers*, ed. W.D. Niven. *op. cit.*, **1**, p. 500.
21 Maxwell, J.C. (1861–62). *Scientific Papers*, ed. W.D. Niven. *op. cit.*, **1**, p. 486.
22 Harman, P.M. (1998). *op. cit.*, p. 113.
23 Maxwell, J.C. (1865). *Scientific Papers*, ed. W.D. Niven. *op. cit.*, **1**, pp. 459–512.

6 Maxwell (1865): *A Dynamical Theory of the Electromagnetic Field*

There are a few scientific papers which profoundly changed how physicists think about the world. Maxwell's great paper of 1865, *A Dynamical Theory of the Electromagnetic Field*,[1] is undoubtedly one of this select few and deserves a chapter to itself.[2]

Maxwell was well aware of the need to recast his theory of electromagnetism of 1861–62 on a much more abstract basis without any special assumptions about the nature of the medium through which electromagnetic phenomena are propagated, which he achieved in his paper of 1865. To quote Whittaker:[3]

> In this, the architecture of his system was displayed, stripped of the scaffolding by aid of which it had been first erected.

There is only one brief, somewhat apologetic, reference in a footnote to his papers *On Physical Lines of Force* of 1861–62. As Siegel has pointed out, from the beginning of his researches into electromagnetism, he had three goals:[4]

> (1) first, to develop a theory that would encompass all the known facts of electricity and magnetism;
> (2) second, to accomplish this unification by using Michael Faraday's approach to electricity and magnetism – the *lines of force* or *field* approach;
> (3) and third, to have the new theory go beyond what was known from experiment . . .

He envisaged predictions of the theory which could be confronted by future experiments.

Maxwell's own view of the significance of his paper of 1865 is revealed in what Francis Everitt calls 'a rare moment of unveiled exuberance' in a letter to his cousin Charles Cay:

> I have also a paper afloat, containing an electromagnetic theory of light, which, till I am convinced to the contrary, I hold to be great guns.[5]

The paper of 1865 is lengthy and divided into seven parts. All of them contain insights and innovations which illuminate how his thinking had developed since his papers of 1861–62. For ease of comprehension, I have translated Maxwell's notation into more familiar symbols for the contemporary reader – these translations are shown in an endnote.[6] Note that the resulting set of equations is a mixture of electrostatic and electromagnetic units, that is, in unrationalised MKS units.[7]

6.1 PART I – Introductory

Maxwell begins with a lengthy introduction in which he reviews the entire field of electromagnetism and sets out the agenda he proposes to follow in the paper. At the outset, he emphasises the difference in approach of his theory as compared with those of Weber and Neumann. Their approaches involved the Newtonian concept of 'action at a distance' without any account of how forces are transmitted from their sources to other bodies. Maxwell states that he

> . . . preferred to seek an explanation of the fact in another direction, by supposing them to be produced by actions which go on in the surrounding medium as well as in the excited bodies . . . The theory I propose may therefore be called a theory of the *Electromagnetic Field*, because it has to do with the space in the neighbourhood of the electric and magnetic bodies, and it may be called a *Dynamical* Theory, because it assumes that in that space there is matter in motion, by which the observed electromagnetic phenomena are produced.

Central to his theory was the *elasticity* of the medium through which electromagnetic phenomena are propagated, providing a physical realisation for his innovative concept of the *displacement current*. As a consequence, he also had to assume the existence of an *aetherial medium*:

> . . . pervading all bodies, and modified only in degree by their presence; that the parts of the medium are capable to being set in motion by electric currents and magnets; that this motion is communicated from one part of the medium to another by forces from the connexions of these parts.

Maxwell's logic for the necessity of the aether as a physical component of the Universe is set out with remarkable clarity in his contribution to the ninth edition of the *Encyclopaedia Britannica* in 1878.[8]

He contrasts his approach with that of other investigators such as Helmholtz and Thomson, who deduced the phenomena of induction from their mechanical actions. Maxwell proposes to follow the opposite agenda – to deduce the mechanical actions from the laws of induction. To achieve this he needs the *General Equations of the Electromagnetic Field*, which consist of 20 equations with 20 variables. The paper sets out the origins of these equations and, among their many applications, is the key result that electromagnetic radiation propagates at the speed of light – the unification of light, electricity and magnetism. Maxwell acknowledges the inspiration of Faraday's paper *Thoughts on Ray Vibrations*:[9]

> The electromagnetic theory of light as proposed by (Faraday) is the same in substance as that which I have begun to develop in this paper, except that in 1846 there was no data to calculate the velocity of propagation.

6.2 PART II – On Electromagnetic Induction

Maxwell now embarks on a lengthy discussion of electromagnetic induction. The mechanical origin of his thought informs his approach to the determination of the forms of the equations of self and mutual inductances between two inductances. Here he employs his technique of working by analogy to remarkable effect. He writes:

> Now, if the magnetic state of the field depends on motions of the medium, a certain force must be exerted in order to increase or diminish these motions, and when the motions are excited they continue, so that the effect of the connexion between the current and the electromagnetic field surrounding it, is to endow the current with a kind of momentum, just as the connexion between the driving-point of a machine and a fly-wheel endows the driving point with an additional momentum, which may be called the momentum of the fly-wheel reduced to the driving point. The unbalanced force acting on the driving-point increases this momentum, and is measured by the rate of its increase.

The inspiration for this approach came from Lagrange's *Méchanique Analytique*. Maxwell had studied Poisson's *Traité de Méchanique* at Edinburgh University and was familiar with the French school of analytic dynamics, particularly praising Lagrange's insights. He first derives an equation for the forces F_{A} and F_{B} acting on a body at driving points A and B:

$$\left.\begin{aligned} F_{\mathrm{A}} &= \frac{\mathrm{d}}{\mathrm{d}t}(L_{\mathrm{A}}v_{\mathrm{A}} + M_{\mathrm{AB}}v_{\mathrm{B}}), \\ F_{\mathrm{B}} &= \frac{\mathrm{d}}{\mathrm{d}t}(M_{\mathrm{AB}}v_{\mathrm{A}} + L_{\mathrm{B}}v_{\mathrm{B}}), \end{aligned}\right\} \tag{6.1}$$

where the terms L_{A} and L_{B} are the weighted masses of the extended body and M_{AB} is the coupling between the impressed forces by the rigid body. The significance of the coupling between these independent forces is explained by Maxwell as follows:

> If the velocity of A be increased at the rate $\mathrm{d}v_{\mathrm{A}}/\mathrm{d}t$, then in order to prevent B from moving a force, $F'_{\mathrm{B}} = \mathrm{d}/\mathrm{d}t\ (M_{\mathrm{AB}}v_{\mathrm{A}})$ must be applied to it. This effect on B, due to an increase of the velocity of A, corresponds to the electromotive force on one circuit arising from an increase in the strength of a neighbouring circuit.

Maxwell now introduces Newton's Second Law of Motion $F = \mathrm{d}p/\mathrm{d}t$, where p is the momentum. The quantities $(L_{\mathrm{A}}v_{\mathrm{A}} + M_{\mathrm{AB}}v_{\mathrm{B}})$ and $(M_{\mathrm{AB}}v_{\mathrm{A}} + L_{\mathrm{B}}v_{\mathrm{B}})$ are then defined to be the *reduced momenta* which describe the combined action of the two forces applied at the points A and B of the rigid body.

He now adds frictional terms of the form $R_{\mathrm{A}}v_{\mathrm{A}}$ and $R_{\mathrm{B}}v_{\mathrm{B}}$ to each of equations (6.1) where the Rs are the coefficients of resistance. The total forces acting at A and B, F'_{A} and F'_{B} are then:

$$\left.\begin{aligned} F'_{\mathrm{A}} &= R_{\mathrm{A}}v_{\mathrm{A}} + \frac{\mathrm{d}}{\mathrm{d}t}(L_{\mathrm{A}}v_{\mathrm{A}} + M_{\mathrm{AB}}v_{\mathrm{B}}), \\ F'_{\mathrm{B}} &= R_{\mathrm{B}}v_{\mathrm{B}} + \frac{\mathrm{d}}{\mathrm{d}t}(M_{\mathrm{AB}}v_{\mathrm{A}} + L_{\mathrm{B}}v_{\mathrm{B}}). \end{aligned}\right\} \tag{6.2}$$

F'_{A} and F'_{B} are the forces referred to the points A and B respectively.

Of particular significance in Maxwell's thinking is the equivalence between *reduced momenta* in mechanics and the corresponding quantities in electromagnetic induction, the *electromagnetic momenta*, what Faraday referred to as the *Electrotonic State*. Maxwell immediately converts the mechanical equations into the equations of induction between two current-carrying conductors. Then the equation for the electromotive force $\mathcal{E}_A$ in A is given by

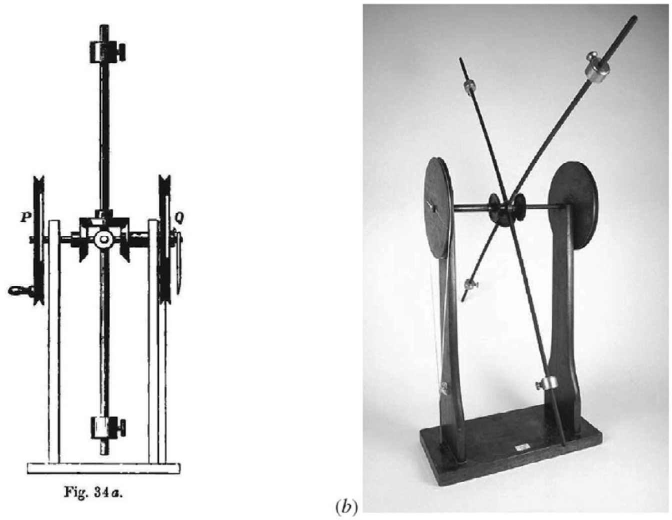
$$\mathcal{E}_A = R_A I_A + \frac{d}{dt}(L_A I_A + M_{AB} I_B), \tag{6.3}$$

and that of $\mathcal{E}_B$ in B by

$$\mathcal{E}_B = R_B I_B + \frac{d}{dt}(M_{AB} I_A + L_B I_B), \tag{6.4}$$

where $\mathcal{E}_A$ and $\mathcal{E}_B$ are the electromotive forces, I_A and I_B the currents, and R_A and R_B the resistances in A and B respectively. Mechanics has been translated by analogy directly into electromagnetism. From these, he derived the expressions for work and energy, the heat produced by the currents, the intrinsic energy of the currents and the mechanical action between conductors, and much more.

In typical Maxwellian fashion, he then designed a mechanical model which illustrates precisely the rules of electromagnetic induction in mechanical terms. Figure 6.1(*a*) shows the illustration on page 228, Volume 2 of the third (1891) edition of Maxwell's *Treatise on Electricity and Magnetism* edited by J.J. Thomson and published 12 years after Maxwell's death. The model, built by Messrs Elliot Brothers of London in 1876, is shown in Fig. 6.1(*b*). The extended rigid body C is a flywheel which consists of four long steel rods at right angles to each other to which heavy weights are attached towards their ends, giving the flywheel a large moment of inertia. Forces are applied to the flywheel to cause



(*a*) (*b*)

Fig. 6.1 (*a*) The diagram from Maxwell's *Treatise on Electricity and Magnetism*, 3rd edition, showing the mechanical model he had built to illustrate the principles of electromagnetic induction. (*b*) The original model belonging to the Cavendish Laboratory and now on loan to the Whipple Museum, Cambridge University.

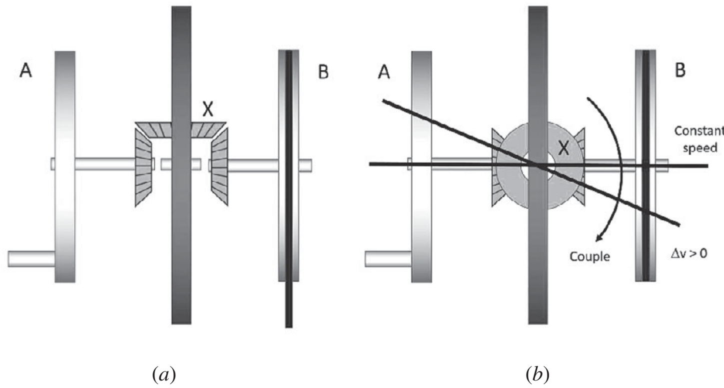


Fig. 6.2 (*a*) A schematic diagram showing more clearly than Fig. 6.1 the arrangement of the bevel gears which transmit the forces to the flywheel. (*b*) Illustrating the origin of the couple acting on the system when the disc A is accelerated.

it to rotate through the differential gear arrangement shown in Fig. 6.1(*a*). To make the arrangement clearer, I have redrawn the differential gearing schematically in Figs. 6.2(*a*) and (*b*). A and B are attached to separate axles which have bevel gears at their ends. They mesh with the horizontal bevel gear X, as shown in Fig. 6.2(*a*), which is attached to the flywheel but which is free to rotate about its own axis. The discs A and B are the origins of the forces F_A and F_B and their 'independent driving points' are the opposite sides of the bevel gear X.

Suppose B is stationary while A is rotated at the constant speed v_A. Then, the flywheel only rotates at half the speed of A. In fact, by geometry, the flywheel always rotates at the average speed of rotation of the discs A and B. Now suppose we accelerate the rotation of A. Since $F = m_A(\Delta v_A/\Delta t)$, there is a force acting on the bevel gear X. But the gear is attached to the flywheel and so there is a reaction force in the opposite sense acting on B. In other words, when the disc A is accelerated, the result is a couple acting on the bevel gear X, which causes the disc B to rotate in the opposite direction (Fig. 6.2(*b*)). But, this reaction force at B only takes place during the period when the disc A is accelerated. It is the perfect mechanical analogue for electromagnetic induction, and includes Lenz's law.

What is remarkable about this analysis is that Maxwell's electromagnetic momentum is precisely what we now call the vector potential, $\boldsymbol{A} \equiv [A_x, A_y, A_z]$. The origin of this identification is apparent from equations (6.1) which are no more than $E = \partial A/\partial t$. In his paper of 1865, Maxwell works entirely in terms of $\boldsymbol{A}$ rather than the magnetic flux density $\boldsymbol{B}$ which is found from the relation $\boldsymbol{B} = \text{curl}\,\boldsymbol{A}$.

Remarkably, exactly the same mechanical equations (6.2) with the frictional terms included appear in a later paper by Maxwell *On Governors* of 1868.[10] This paper on the stability of control systems is regarded by engineers as the pioneering paper in cybernetics.

6.3 PART III – General Equations of the Electromagnetic Field

In this part, the complete set of Maxwell's equations for the electromagnetic field are presented with clarity and compelling logic. The equations are the same as presented in his earlier papers, but their interpretation has changed in revolutionary ways. The equations are labelled alphabetically as used in the paper. In summary, these are:

Three equations of Magnetic Force (H_x, H_y, H_z)

$$\mu H_x = \frac{\mathrm{d}A_z}{\mathrm{d}y} - \frac{\mathrm{d}A_y}{\mathrm{d}z},\ \mu H_y = \frac{\mathrm{d}A_x}{\mathrm{d}z} - \frac{\mathrm{d}A_z}{\mathrm{d}x},\ \mu H_z = \frac{\mathrm{d}A_y}{\mathrm{d}x} - \frac{\mathrm{d}A_x}{\mathrm{d}y}. \qquad \text{(B)}$$

Three equations of Electric Currents (J_x, J_y, J_z)

$$\frac{\mathrm{d}H_z}{\mathrm{d}y} - \frac{\mathrm{d}H_y}{\mathrm{d}z} = 4\pi J'_x,\ \frac{\mathrm{d}H_x}{\mathrm{d}z} - \frac{\mathrm{d}H_z}{\mathrm{d}x} = 4\pi J'_y,\ \frac{\mathrm{d}H_y}{\mathrm{d}x} - \frac{\mathrm{d}H_x}{\mathrm{d}y} = 4\pi J'_z. \qquad \text{(C)}$$

Three equations of Electromotive Force (E_x, E_y, E_z)

$$\begin{aligned} E_x &= \mu\left(H_z\frac{\mathrm{d}y}{\mathrm{d}t} - H_y\frac{\mathrm{d}z}{\mathrm{d}t}\right) - \frac{\mathrm{d}A_x}{\mathrm{d}t} - \frac{\mathrm{d}\phi}{\mathrm{d}x}, \\ E_y &= \mu\left(H_x\frac{\mathrm{d}z}{\mathrm{d}t} - H_z\frac{\mathrm{d}x}{\mathrm{d}t}\right) - \frac{\mathrm{d}A_y}{\mathrm{d}t} - \frac{\mathrm{d}\phi}{\mathrm{d}y}, \\ E_z &= \mu\left(H_y\frac{\mathrm{d}x}{\mathrm{d}t} - H_x\frac{\mathrm{d}y}{\mathrm{d}t}\right) - \frac{\mathrm{d}A_z}{\mathrm{d}t} - \frac{\mathrm{d}\phi}{\mathrm{d}z}. \end{aligned} \qquad \text{(D)}$$

Three equations of Electric Elasticity (D_x, D_y, D_z)

$$E_x = kD_x,\ E_y = kD_y,\ E_z = kD_z. \qquad \text{(E)}$$

Three equations of Electric Resistance (ϱ)

$$E_x = -\varrho J_x, E_y = -\varrho J_y, E_z = -\varrho J_z. \qquad \text{(F)}$$

Three equations of Total Currents (J'_x, J'_y, J'_z)

$$J'_x = J_x + \frac{\mathrm{d}D_x}{\mathrm{d}t},\ J'_y = J_y + \frac{\mathrm{d}D_y}{\mathrm{d}t},\ J'_z = J_z + \frac{\mathrm{d}D_z}{\mathrm{d}t}. \qquad \text{(A)}$$

One equation of Free Electricity (ρ_e)

$$\rho_\mathrm{e} + \frac{\mathrm{d}D_x}{\mathrm{d}x} + \frac{\mathrm{d}D_y}{\mathrm{d}y} + \frac{\mathrm{d}D_z}{\mathrm{d}z} = 0. \qquad \text{(G)}$$

One equation of Continuity $(\mathrm{d}\rho_\mathrm{e}/\mathrm{d}t)$

$$\frac{\mathrm{d}\rho_\mathrm{e}}{\mathrm{d}t} + \frac{\mathrm{d}J_x}{\mathrm{d}x} + \frac{\mathrm{d}J_y}{\mathrm{d}y} + \frac{\mathrm{d}J_z}{\mathrm{d}z} = 0. \qquad \text{(H)}$$

The result is 20 equations for the 20 variables which are:

Electromagnetic Momentum	A_x	A_y	A_z
Magnetic Intensity	H_x	H_y	H_z
Electromotive Force	E_x	E_y	E_z
Current due to true conduction	J_x	J_y	J_z
Electric Displacement	D_x	D_y	D_z
Total Current (including variation of displacement)	J'_x	J'_y	J'_z
Quantity of free Electricity	ρ_e		
Electric Potential	ϕ		

Written in vector notation, the equations have a familiar appearance. If we identify $\mu\boldsymbol{H}$ with $\boldsymbol{B}$, the equations (A) to (H) can be written as follows:

$$\boldsymbol{B} = \mu\boldsymbol{H} = \text{curl}\,\boldsymbol{A}, \tag{B}$$

$$\text{curl}\,\boldsymbol{H} = 4\pi\boldsymbol{J}' = 4\pi\left(\boldsymbol{J} + \frac{\mathrm{d}\boldsymbol{D}}{\mathrm{d}t}\right), \tag{C}$$

$$\boldsymbol{E} = \mu(\boldsymbol{v}\times\boldsymbol{H}) - \frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t} - \nabla\phi = (\boldsymbol{v}\times\boldsymbol{B}) - \frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t} - \nabla\phi, \tag{D}$$

$$\boldsymbol{E} = k\boldsymbol{D}, \tag{E}$$

$$\boldsymbol{E} = -\rho\boldsymbol{J}, \tag{F}$$

$$\boldsymbol{J}' = \boldsymbol{J} + \frac{\mathrm{d}\boldsymbol{D}}{\mathrm{d}t}, \tag{A}$$

$$\varrho_e + \nabla\cdot\boldsymbol{D} = 0, \tag{G}$$

$$\frac{\mathrm{d}\varrho_e}{\mathrm{d}t} + \nabla\cdot\boldsymbol{J} = 0. \tag{H}$$

Thus, comparing these equations with (5.44),

Maxwell's first equation is found by omitting the $\boldsymbol{v}\times\boldsymbol{B}$ term and taking the curl of (D).
Maxwell's second equation is (C).
Maxwell's third equation is (G).[11]
Maxwell's fourth equation is found from the divergence of (B).
(A) is absorbed into (C).
(D) contains the expression for the force on unit charge, $\boldsymbol{f} = (\boldsymbol{v}\times\boldsymbol{B})$, normally referred to as the Lorentz force.
(E) is the constitutive expression relating $\boldsymbol{E}$ and $\boldsymbol{D}$.
(F) is Ohm's law, which Maxwell recognised was an empirical relation.
(H) is the continuity equation for electric charge.

Heaviside and Hertz are given credit for putting Maxwell's equations into their conventional vector form but it is apparent that only the simplest modifications were needed from Maxwell's original version.[12] This would include a consistent use of Gaussian units, to be replaced by the SI system in 1960.

There are some intriguing features of Maxwell's presentation of the set of equations:

- Whereas the electrical displacement (D_x, D_y, D_z) appeared awkwardly in his papers of 1861–62, it is now deeply embedded into the structure of electromagnetism as:

  > the opposite electrification of the sides of a molecule or particle of a body which may or may not be accompanied with transmission through the body.

  The idle-wheels were an unnecessary artefact – the phenomenon of electrical displacement must necessarily occur and its time variation contributes to the total current.
- The electromagnetic momentum $\boldsymbol{A} \equiv [A_x, A_y, A_z]$ is what we now call the vector potential. The origin of this identification is apparent from (6.2) and (6.3) and presages the four-vector notation of special relativity in which the four-vector for the electromagnetic potential is written $\mathbf{A} \equiv [\phi/c, \boldsymbol{A}] = [\phi/c, A_x, A_y, A_z]$ with the dimensions of momentum divided by electric charge. Maxwell makes liberal use of the vector potential in his development of the equations, in contrast to contemporary practice of working initially with the fields $\boldsymbol{E}, \boldsymbol{D}, \boldsymbol{B}, \boldsymbol{H}$ and $\boldsymbol{J}$.
- Maxwell includes Ohm's law in his exposition of the structure of the equations, which is avoided in the contemporary set of equations. Maxwell was well aware of the empirical basis of Ohm's law in contrast to the equations of electromagnetic induction. As already mentioned, one of his first successful projects as Cavendish Professor was to establish Ohm's law with much improved precision with his graduate student George Chrystal.

The final section of this part of the paper is the determination of the various forms of energy associated with the fields. He starts again with the electromagnetic momentum, or vector potential, A_x, A_y, A_z and finds the total energy from the expression

$$E = \frac{1}{2} \sum \left(A_x J'_x + A_y J'_y + A_z J'_z \right) \, \mathrm{d}V, \tag{6.5}$$

summing over all the space occupied by the currents. The total energy can be related to the total energy existing in the fields themselves:

$$E = \sum \left\{ \frac{1}{8\pi} (\mu H_x^2 + \mu H_y^2 + \mu H_z^2) + \frac{1}{2} (E_x D_x + E_y D_y + E_z D_z) \, \mathrm{d}V \right\}. \tag{6.6}$$

Maxwell then describes the key significance of this expression:

> (74) In speaking of the Energy of the field, however, I wish to be understood literally. All energy is the same as mechanical energy, whether it exists in the form of motion or in that of elasticity, or in any other form. The energy in electromagnetic phenomena is mechanical energy. The only question is, Where does it reside? On the old theories it resides in the electrified bodies, conducting circuits, and magnets, in the form of an unknown quality called potential energy, or the power of producing certain effects at a distance. On our theory it resides in the electromagnetic field, in the space surrounding the electrified and magnetic bodies, as well as in those bodies themselves, and is in two different forms, which may be described without hypothesis as magnetic polarization and electric polarization, or, according to a very probable hypothesis, as the motion and the strain of one and the same medium.

6.4 PART IV – Mechanical Actions in the Field

Maxwell's next task is to show how these expressions for the various fields lead to the known laws of forces between the various electromagnetic entities. Note that these force laws had not been explicitly included in the formulation of the equations of electromagnetism. He derives from the field equations the known laws of force of a conductor moving through a magnetic field, the mechanical force on a magnet and the force on an electrified body.[13]

Maxwell also points out that the coefficient of 'electric elasticity' k, which appears in equation (E) of his equations, is directly related to the ratio of electrostatic to electromagnetic units of measurement through the expression $k = 4\pi v^2$. Weber and Kohlrausch had already determined this ratio experimentally and found it to be $v = 310\,740\,000$ m s^{-1}.

The last section of this part involves trying to apply the same techniques to understand the laws of gravity. Here, Maxwell hits the problem of the universal attractive force of gravity which results in negative values for the energy of a gravitating system. He confesses,

> As I am unable to understand in what way a medium can possess such properties, I cannot go any further in this direction in searching for the cause of gravitation.

6.5 PART V – Theory of Condensers

This section is concerned with the determination of the capacity and absorption of capacitors of various construction. This was an issue of considerable importance for the laying of long lengths of submarine telegraph cables, highlighted by the failure of the project to lay the first transatlantic telegraph cable in 1858.[14] Maxwell had already been deeply involved in the issues concerning the determination of fundamental constants and of absolute standards of resistance. The 1861 meeting of the British Association had appointed a committee to oversee the determination of fundamental standards and Maxwell joined the Committee in 1862, soon after the publication of his papers on electromagnetism of 1861–62.

He became deeply involved in testing his theory by precise experiment, in particular, the determination of the ratio of electrostatic to electromagnetic units of charge. The activities of the Committee became much more mathematical and theoretical, playing directly to Maxwell's strengths. He set about supervising the design and construction of apparatus to make a very accurate determination of the ohm with his colleagues Balfour Stewart and Fleeming Jenkin at King's College, London.[15] The success of these experiments convinced the Committee that the absolute value of the ohm determined by this and similar techniques was the clearly preferred standard.

The work on standards fragmented in subsequent years, but Maxwell maintained his strong interest and leading role in the project and made it one of the central themes of the research programme of the new Cavendish Laboratory in 1874. The work on determining the absolute standard of resistance was transferred from the Kew Observatory to the Cavendish Laboratory and was to remain one of the central roles of the Laboratory until it was taken over by the National Physical Laboratory on its foundation in 1900.

6.6 PART VI – Electromagnetic Theory of Light

Maxwell now seeks to determine whether or not the waves which can be propagated through any material medium are consistent with the postulate that light can be identified with electromagnetic waves. Setting the conduction terms to zero, he derives the equations for the propagation of electromagnetic waves in the x, y, z directions by the familiar procedures (Section 5.3.2):

$$\left.\begin{aligned} k\nabla^2 \mu H_x &= 4\pi\mu \frac{\mathrm{d}^2}{\mathrm{d}t^2}\mu H_x, \\ k\nabla^2 \mu H_y &= 4\pi\mu \frac{\mathrm{d}^2}{\mathrm{d}t^2}\mu H_y, \\ k\nabla^2 \mu H_z &= 4\pi\mu \frac{\mathrm{d}^2}{\mathrm{d}t^2}\mu H_z. \end{aligned}\right\} \tag{6.7}$$

He continues

(95) If we assume that H_x, H_y, H_z are functions of $lx + my + nz - Vt = w$, the first equation becomes

$$k\mu \frac{\mathrm{d}^2 H_x}{\mathrm{d}w^2} = 4\pi\mu^2 V^2 \frac{\mathrm{d}^2 H_x}{\mathrm{d}w^2}, \tag{6.8}$$

or

$$V = \pm\sqrt{\frac{k}{4\pi\mu}}. \tag{6.9}$$

The other equations give the same value of V, so that the wave is propagated in either direction with a velocity V.

This wave consists entirely of magnetic disturbances, the direction of magnetization being in the plane of the wave. No magnetic disturbance whose direction of magnetization is not in the plane of the wave can be propagated as a plane wave at all.

Hence magnetic disturbances propagated through the electromagnetic field agree with light in this, that the disturbance at any point is transverse to the direction of propagation, and such waves may have all the properties of polarized light.

For case of air, for which $\mu = 1$, the speed of propagation of light was measured by Foucault to be 298 000 000 m s^{-1}, compared with the value derived by Weber and Kohlrausch who found $v = 310\,740\,000$ m s^{-1} from their measured value of k. These figures also agreed within experimental error with the value of the speed of light determined from the astronomical aberration of light rays.

Next, Maxwell goes on to show that the refractive index of a non-conducting material n is given by the square root of the specific inductive capacity of the medium. He writes:

> [Writing] $k = \frac{1}{D}k_0$ and $k_0 = 4\pi V_0^2$ ... Hence
>
> $$D = \frac{n^2}{\mu}, \tag{6.10}$$
>
> or the Specific Inductive Capacity is equal to the square of the index of refraction divided by the coefficient of magnetic induction.

Maxwell goes on to make a preliminary exploration of the case of anisotropic crystalline materials. He was well aware of the fact that crystals could have different refractive indices along different axes and so studies the case in which μ takes different values along the x, y, z directions.

The next application is the relation between electrical resistance and the transparency of materials. The objective was to account for the fact that, as he put it, 'Most transparent solid bodies are good insulators, whereas all good conductors are very opaque.' He investigates the propagation of light along the x-axis of the transverse disturbance A_y. Including the resistivity ϱ of the medium, the propagation equation becomes

$$\frac{\mathrm{d}^2 A_y}{\mathrm{d}x^2} = 4\pi\mu_y\left(\frac{1}{k}\frac{\mathrm{d}^2 A_y}{\mathrm{d}t^2} - \frac{1}{\varrho}\frac{\mathrm{d}A_y}{\mathrm{d}t}\right). \tag{6.11}$$

If A_y takes the form

$$A_y = \mathrm{e}^{-\alpha x}\cos(k_y x + \omega t), \tag{6.12}$$

where α is the absorption coefficient,

$$\alpha = \frac{2\pi\mu_y}{\varrho}\frac{\omega}{k_y} = \frac{2\pi\mu_y}{\varrho}\frac{V}{n}, \tag{6.13}$$

where V is the velocity of light in air, and n is the index of refraction. The fraction of the incident intensity of light transmitted through the thickness x is

$$\mathrm{e}^{-2\alpha x}. \tag{6.14}$$

If R is the resistance of a sample of the material of thickness is x, breadth b, and length l, then

$$R = \frac{l\varrho}{bx} \quad \text{and} \quad 2\alpha x = 4\pi\mu_y\frac{V}{n}\frac{l}{bR}, \tag{6.15}$$

establishing quantitatively the relation between α, n and R – if R is large, the absorption coefficient α is small, and vice versa.

6.7 PART VII – Calculation of the Coefficients of Electromagnetic Induction

The final part of the paper concerns the accurate estimation of the coefficients of electromagnetic induction. This might seem a descent from the heights of Parts 3, 4 and 6 of the paper, but these calculations were of central importance for the absolute determination of the ohm, one of Maxwell's preoccupations for the rest of his life. Suffice to say, Maxwell describes in some detail the various ways in which self and mutual inductances could be measured, in the context of the experiments which he and his colleagues were carrying out at King's College, London, and which were contained in the Report to the Committee of the British Association. The considerations were to find their application in the meticulous experiments of Rayleigh, Schuster and Sedgwick in their determination of the absolute value of the ohm after Maxwell's death.[16]

6.8 The Aftermath

The identification of light with electromagnetic radiation was a triumph, providing a physical foundation for the wave theory of light, which could successfully account for the phenomena of reflection, refraction, polarisation, and so on. It is striking, however, how long it took for Maxwell's deep insights to become generally accepted by the community of physicists. He elaborated the theory in his great *Treatise on Electricity and Magnetism* as soon as he had settled back at his home at Glenlair in the Dumfries and Galloway region of southern Scotland in 1865.

It is significant that, while he was writing the *Treatise*, he was also an examiner for the Cambridge Mathematical Tripos and realised the dire need for suitable textbooks. The two-volume *Treatise* is, however, unlike many of the other great treatises such as Newton's *Principia* in that it is not a systematic presentation of the subject but a work in progress, reflecting Maxwell's own approach to these researches. In a later conversation, Maxwell remarked that the aim of the *Treatise* was not to expound his theory finally to the world, but to educate himself by presenting a view of the stage he had reached. Disconcertingly, Maxwell's advice was to read the four parts of the *Treatise* in parallel rather than in sequence. For example, on reaching Section 585, halfway through Volume 2, Maxwell states that he is to 'begin again from a new foundation without any assumption except those of the dynamical theory ... '. This introduced new mathematical approaches to electromagnetism, including quaternions, integral theorems such as Stokes' theorem, topological concepts and Lagrangian–Hamiltonian methods of analytic dynamics.

One of the most important results appears in Section 792 of Volume 2 in which Maxwell works out the pressure which radiation exerts on a conductor on the basis of electromagnetic theory. This profound result provides a relation between the pressure p and the energy density ε of radiation derived purely from the properties of electromagnetic fields, $p = \varepsilon/3c^2$. We will derive this result using essentially Maxwell's approach in Section 13.3.2 in the context of Boltzmann's derivation of the Stefan–Boltzmann law.

Maxwell was remarkably modest about his contribution. For example, he was President of the Mathematical and Physical Sciences Section of the British Association for the Advancement of Science in 1870. His Presidential lecture was a splendid opportunity for describing his new ideas, and yet he scarcely mentioned his recent work on electromagnetism, simply referring to it in an off-hand way as, 'Another theory of electricity which

I prefer', without even mentioning that it was his own. As Freeman Dyson remarks, 'The moral of this story is that modesty is not always a virtue'.[17]

But the problems were much deeper. Not only was Maxwell's theory complex, the discovery of the equations for the electromagnetic field also required a major shift in perspective for physicists of the late nineteenth century. It is worth quoting Dyson a little further.

> There were other reasons, besides Maxwell's modesty, why his theory was hard to understand. He replaced the Newtonian universe of tangible objects interacting with one another at a distance by a universe of fields extending through space and only interacting locally with tangible objects. The notion of a field was hard to grasp because fields are intangible. The scientists of that time, including Maxwell himself, tried to picture fields as mechanical structures composed of a multitude of little wheels and vortices extending throughout space. These structures were supposed to carry the mechanical stresses that electric and magnetic fields transmitted between electric charges and currents. To make the fields satisfy Maxwell's equations, the system of wheels and vortices had to be extremely complicated. If you try to visualise the Maxwell theory with such mechanical models, it looks like a throwback to Ptolemaic astronomy with planets riding on cycles and epicycles in the sky. It does not look like the elegant astronomy of Newton.[17]

In fairness to Maxwell, the vortices and idle-wheels had disappeared by the time of his great paper of 1865, but his underlying reasoning was undoubtedly mechanical. To quote Dyson again:

> Maxwell's theory becomes simple and intelligible only when you give up thinking in terms of mechanical models. Instead of thinking of mechanical objects as primary and electromagnetic stresses as secondary consequences, you must think of the electromagnetic field as primary and mechanical forces as secondary. The idea that the primary constituents of the universe are fields did not come easily to the physicists of Maxwell's generation. Fields are an abstract concept, far removed from the familiar world of things and forces. The field equations of Maxwell are partial differential equations. They cannot be expressed in simple words like Newton's law of motion, force equals mass times acceleration. Maxwell's theory had to wait for the next generation of physicists, Hertz and Lorentz and Einstein, to reveal its power and clarify its concepts. The next generation grew up with Maxwell's equations and was at home in a universe built out of fields. The primacy of fields was as natural to Einstein as the primacy of mechanical structures had been to Maxwell.[17]

But there were more immediate concerns for the physics community. First, Maxwell was rejecting the commonly held view that the laws of motion should involve Newton's 'action at a distance'. Second, Maxwell's equations were not 'relativistic' in the sense that, according to the Galilean transformations of Newtonian mechanics, Maxwell's equations were not form invariant – they only held good in a preferred frame of reference. This story is taken up in Section 18.1.1.

Maxwell died in 1879 before direct experimental evidence was obtained for the existence of electromagnetic waves. The matter was finally laid to rest 10 years after Maxwell's death in a classic series of experiments by Heinrich Hertz, almost 30 years after

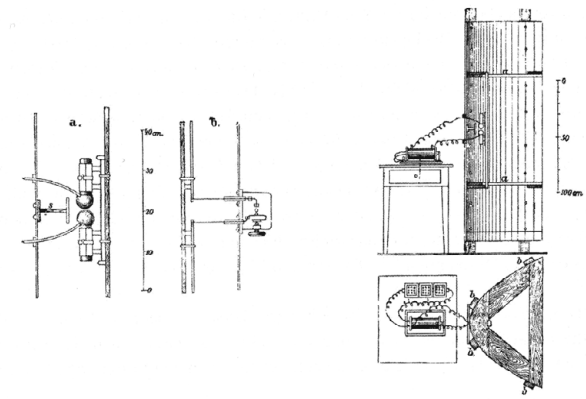


Fig. 6.3 Hertz's apparatus for the generation and detection of electromagnetic radiation. The emitter *a* produced electromagnetic radiation in discharges between the spherical conductors. The detector *b* consisted of a similar device with the jaws of the detector placed as close together as possible to achieve maximum sensitivity. The emitter was placed at the focus of a cylindrical paraboloid reflector to produce a directed beam of radiation. (From H. Hertz, 1893. *Electric Waves*, pp. 183–184, London: MacMillan & Co.)

Maxwell had identified light with electromagnetic radiation. Hertz's great monograph *On Electric Waves*[18] sets out beautifully his remarkable set of experiments.

Hertz found that he could detect the effects of electromagnetic induction at considerable distances from his apparatus. Examples of the types of emitter and detector which he used are shown in Fig. 6.3. Electromagnetic radiation was emitted when sparks were produced between the large spheres by the application of a high voltage from an induction coil. The method of detection of the radiation field was the observation of sparks in the gap between the pair of small spheres of the detector.

After a great deal of trial and error, Hertz found that for particular arrangements there was a strong resonance, corresponding to the resonant frequencies of the emitter and detector. The frequency of the waves at resonance could be found from the resonant frequency of the emitter and detector, which he took to be $\omega = (LC)^{-1/2}$, where L and C are the inductance and capacitance of the dipole. He measured the wavelength of the radiation by placing a reflecting sheet at some distance from the spark gap emitter so that standing waves were set up along the line between the emitter and the sheet. The speed of the waves could be found from the relation $c = \nu\lambda$.

The speed turned out to be almost exactly the speed of light in free space. He then began a series of experiments which demonstrated conclusively that these waves behaved in all respects exactly like light – rectilinear propagation, polarisation, reflection, refraction. Some of the experiments were quite remarkable in their execution. To demonstrate refraction he constructed a prism weighing 12 cwt (610 kg) out of 'so-called hard pitch, a material like asphalt'.

The experiments demonstrated convincingly that there exist electromagnetic waves of frequency about 1 GHz and wavelength 30 cm which behave in all respects like light. These great experiments provided unambiguous validation of Maxwell's equations.

Notes

1 Maxwell, J.C. (1865). A Dynamical Theory of the Electromagnetic Field, *Philosophical Transactions of the Royal Society of London*, **155**, 459–512.
2 The major sources used in preparing this chapter were:

Buchwald, J.Z. (1985). *From Maxwell to Microphysics: Aspects of Electromagnetic Theory in the Last Quarter of the Nineteenth Century*. Chicago: University of Chicago Press.
Campbell, L. and Garnett, W. (1882). *The Life of James Clerk Maxwell*. London: MacMillan & Co.
Flood, R., McCartney, M. and Whitaker, A. (eds.) (2014). *James Clerk Maxwell: Perspectives on his Life and Work*. Oxford: Oxford University Press.
Harman, P. (1990). *The Scientific Letters and Papers of James Clerk Maxwell. Volume I, 1846–1862* (1995), *Volume II, 1862–1873* (1998), *Volume III, 1874–1879* (2002). Cambridge: Cambridge University Press.
Harman, P. (1998). *The Natural Philosophy of James Clerk Maxwell*. Cambridge: Cambridge University Press.
Maxwell, J.C. *The Scientific Papers of J. Clerk Maxwell* in two volumes (1890), ed. W.D. Niven. Cambridge: Cambridge University Press. (Reprinted 1965. New York: Dover Publications Inc.)
Maxwell, J.C. (1873). *Treatise on Electricity and Magnetism*, in two volumes. Oxford: Clarendon Press. (Reprint of 3rd edition, 1998. Oxford Classics series.)
Siegel, D.M. (1991). *Innovation in Maxwell's Electromagnetic Theory: Molecular Vortices, Displacement Current, and Light*. Cambridge: Cambridge University Press.

3 Whittaker, E. (1951). *A History of the Theories of Aether and Electricity*, p. 255. London: Thomas Nelson & Sons.
4 Siegel, D.M. (2014). Maxwell's Contributions to Electricity and Magnetism, in *James Clerk Maxwell: Perspectives on his Life and Work*, eds. R. Flood, M. McCartney and A. Whitaker, pp. 187–203. Oxford: Oxford University Press.
5 Campbell, L. and Garnett, W. (1882). *The Life of James Clerk Maxwell*, p. 342 (Letter of 5 January 1865), London: MacMillan & Co. See also Harman, P.M. (ed.) (1995). *The Scientific Letters and Papers of James Clerk Maxwell*, **2**, p. 203. Cambridge: Cambridge University Press.
6 Translation table from Maxwell's notation to unrationalised MKS units

Quantity							
Mass of body	C			→	M_C		
Forces acting at A and B	A	B		→	F_A	F_B	
Velocities	u	v	w	→	v_A	v_B	v_C
Reduced masses ≡ inductances	L	M	N	→	L_A	M_{AB}	L_B
Resistances	R	S		→	R_A	R_B	
Electromotive Force acting on a circuit	ξ	η		→	$\mathcal{E}_A$	$\mathcal{E}_B$	
Current flowing in a circuit	x	y		→	I_A	I_B	
Quantity of Electricity	X	Y		→	Q_A	Q_B	
Heat generated by resistance	H			→	H		
Intrinsic Energy of Currents	E			→	E		
Work done when inductances change	W			→	W		
Angular frequency	p			→	ω		
Voltages	A	C	D	→	V_A	V_C	V_D

Quantity							
Electromagnetic Momentum	F	G	H	$\rightarrow$	A_x	A_y	A_z
Magnetic Intensity	α	β	γ	$\rightarrow$	H_x	H_y	H_z
Electromotive Force	P	Q	R	$\rightarrow$	E_x	E_y	E_z
Current due to true conduction	p	q	r	$\rightarrow$	J_x	J_y	J_z
Electric Displacement	f	g	h	$\rightarrow$	D_x	D_y	D_z
Total Current	p'	q'	r'	$\rightarrow$	J'_x	J'_y	J'_z
Quantity of free Electricity	e			$\rightarrow$	$\rho_{\rm e}$		
Electric Potential	Ψ			$\rightarrow$	ϕ		
Magnetic Potential	ϕ			$\rightarrow$	$\phi_{\rm m}$		
Specific Resistance	ρ			$\rightarrow$	ϱ		
Specific Capacity of Electric Inductance	D			$\rightarrow$	ϵ		

7 More details of the various systems of units used in electromagnetism can be found in the essay *Unit Systems in Electromagnetism* at http://info.ee.surrey.ac.uk/Workshop/advice/coils/unit_systems/.

8 Maxwell, J.C. (1878). Ether, in *Encyclopaedia Britannica*, 9th edition, **8**, pp. 568–572. Edinburgh: A. & C. Black.

9 Faraday, M. (1846). Thoughts on Ray-vibrations, *Faraday's Researches in Electricity*, **3**, 447–452. London: Richard Taylor and William Francis. This letter was first published in the *Philosophical Magazine, Series 3*, **53**, 345–350 (1846), following his lecture at the Royal Institution in April 1846.

10 Maxwell, J.C. (1868). On Governors, *Proceedings of the Royal Society of London*, **16**, 270–283.

11 The sign of $\varrho_{\rm e}$ is wrong – this may have been a printer's error which escaped proof-reading.

12 There is a good account of Heaviside's equations by B.J. Hunt in *The Maxwellians*, 119–128. Ithaca and London: Cornell University Press. See also, Nahin, P.J. (1988). *Oliver Heaviside: Sage in Solitude*. New York: IEEE Press.

13 Many parts of this analysis are similar to those presented in Chapter 7. Until I studied Maxwell's paper in detail, I had not appreciated that the approach of Chapter 7, which we taught in a Theoretical Physics Examples Class in Cambridge, was almost exactly what Maxwell had published in 1865.

14 The history of the laying of the transatlantic cable and its importance for promoting the use of theoretical physics techniques as applied to electrical engineering projects is splendidly told by C. Smith and N.M. Wise in their biography of William Thomson, *Energy and Empire: A Biographical Study of Lord Kelvin* (1989). Cambridge: Cambridge University Press. An abbreviated version of this story is contained in my book *Maxwell's Enduring Legacy: A Scientific History of the Cavendish Laboratory* (2016). Cambridge: Cambridge University Press.

15 Maxwell, J.C., Stewart, B. and Jenkin, F. (1863). Description of an Experimental Measurement of Electrical Resistance, Made at King's College, *British Association Reports*, 140–158.

16 Rayleigh, J.W. and Schuster, A. (1881). On the Determination of the Ohm in Absolute Measure, *Proceedings of the Royal Society of London*, **32**, 104–141. Rayleigh, J.W. and Sidgwick, E.M. (1883). Experiments, by the Method of Lorentz, for the further Determination in Absolute Value of the British Association Unit of Resistance, with an Appendix on the Determination of the Pitch of a Standard Tuning Fork, *Philosophical Transactions of the Royal Society of London*, **174**, 295–322.

17 Dyson, F. (1999). In *James Clerk Maxwell Commemorative Booklet*. Edinburgh: James Clerk Maxwell Foundation.

18 Hertz, H. (1893). *Electric Waves*. London: Macmillan & Company. Reprinted 1962. New York: Dover Publications. The original book, *Untersuchungen über die Ausbreitung der elektrischen Kraft*, was published by Johann Ambrosius Barth in Leipzig in 1892.

7 How to Rewrite the History of Electromagnetism

7.1 Introduction

Now that we have derived Maxwell's equations as he derived them, let us do everything backwards, starting with Maxwell's equations but regarding them simply as a set of vector equations relating the vector fields $\boldsymbol{E}$, $\boldsymbol{D}$, $\boldsymbol{B}$, $\boldsymbol{H}$ and $\boldsymbol{J}$. Initially, these fields have *no physical significance*. We then make a minimum number of postulates in order to give them physical significance and derive from them Maxwell's equations of electromagnetism. This approach is not in fact so different from Maxwell's procedures in his paper of 1865 and also that taken by Stratton in his book *Electromagnetic Theory*.[1] The equations can then be applied to new aspects of electromagnetic theory – the properties of electromagnetic waves, the emission of waves by accelerated charges, and so on – which provide further tests of the theory.

It might be argued that no-one would ever have derived Maxwell's equations in this way. I would argue that the procedure of starting with a mathematical structure, which is then given physical meaning, is found in other aspects of fundamental physics, for example, the theory of linear operators and quantum mechanics, tensor calculus and the special and general theories of relativity. String theory is a contemporary example of the same procedure in an attempt to unify all the forces of nature. The mathematics gives formal coherence to physical theory and in this case illuminates the remarkable symmetries and elegance of the theory of the electromagnetic field.

This study began as an examples class in mathematical physics and it is instructive to maintain that format. Much of the analysis which follows is mathematically straightforward – the emphasis is upon the clarity with which we are able to make the correspondence between the mathematics and the physics.

7.2 Maxwell's Equations as a Set of Vector Equations

We start with Maxwell's equations in the form

$$\operatorname{curl} \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \tag{7.1}$$

$$\operatorname{curl} \boldsymbol{H} = \boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t}. \tag{7.2}$$

$\boldsymbol{E}, \boldsymbol{D}, \boldsymbol{B}, \boldsymbol{H}$ and $\boldsymbol{J}$ are to be regarded as unspecified vector fields which are functions of space and time coordinates and which are intended to describe the properties of the electromagnetic field.[2] They are supplemented by a continuity equation for $\boldsymbol{J}$,

$$\operatorname{div} \boldsymbol{J} + \frac{\partial \rho_e}{\partial t} = 0. \tag{7.3}$$

We then make the first physical identification:

$$\rho_e = \textit{the electric charge density} \text{ and } \textit{charge is conserved}. \tag{7.4}$$

Problem Show from (7.3) that $\boldsymbol{J}$ is a *current density*, that is, the rate of flow of charge through unit surface area.

Proof Integrate (7.3) over a volume v bounded by a surface S, that is

$$\int_v \operatorname{div} \boldsymbol{J}\, dv = -\frac{\partial}{\partial t}\int_v \rho_e\, dv.$$

According to the divergence theorem,

$$\int_v \operatorname{div} \boldsymbol{J}\, dv = \int_S \boldsymbol{J} \cdot d\boldsymbol{S}, \tag{7.5}$$

and so is equal to the rate at which charge is lost from the volume,

$$-\frac{\partial}{\partial t}\int_v \rho_e\, dv.$$

Therefore,

$$\int_S \boldsymbol{J} \cdot d\boldsymbol{S} = -\frac{\partial}{\partial t}(\text{total enclosed charge}). \tag{7.6}$$

Thus, $\boldsymbol{J}$ must represent the rate of flow, or flux, of charge per unit area through a plane perpendicular to $\boldsymbol{J}$.

7.3 Gauss's Theorem in Electromagnetism

Problem Show that the fields $\boldsymbol{B}$ and $\boldsymbol{D}$ must satisfy the relations

$$\operatorname{div} \boldsymbol{B} = 0 \quad \text{and} \quad \operatorname{div} \boldsymbol{D} = \rho_e. \tag{7.7}$$

Derive the corresponding integral forms for these relations – these are different forms of *Gauss's theorem* in electromagnetism.

Proof Take the divergences of (7.1) and (7.2). Since the divergence of a curl is always zero, we obtain

$$\begin{cases} \operatorname{div} \operatorname{curl} \boldsymbol{E} = -\dfrac{\partial}{\partial t}(\operatorname{div} \boldsymbol{B}) = 0, \\ \operatorname{div} \operatorname{curl} \boldsymbol{H} = \operatorname{div} \boldsymbol{J} + \dfrac{\partial}{\partial t}(\operatorname{div} \boldsymbol{D}) = 0. \end{cases} \tag{7.8}$$

Using (7.3), the second of these relations becomes

$$\frac{\partial}{\partial t}(\mathrm{div}\,\boldsymbol{D}) - \frac{\partial \rho_{\mathrm{e}}}{\partial t} = \frac{\partial}{\partial t}(\mathrm{div}\,\boldsymbol{D} - \rho_{\mathrm{e}}) = 0. \tag{7.9}$$

Thus, the *partial derivatives* of div $\boldsymbol{B}$ and (div $\boldsymbol{D} - \rho_{\mathrm{e}}$) with respect to time are both zero at all points in space. Therefore,

$$\mathrm{div}\,\boldsymbol{B} = \text{constant}; \quad \mathrm{div}\,\boldsymbol{D} - \rho_{\mathrm{e}} = \text{constant}.$$

We have to establish the values of the constants. Three approaches can be taken. (i) For *simplicity*, set both of the constants to zero and see if we obtain a self-consistent story. (ii) At some time, we believe we could arrange charges and currents in the Universe to reduce div $\boldsymbol{B}$ and div $\boldsymbol{D} - \rho_{\mathrm{e}}$ to zero. If we can do it for one moment, it must always be true. (iii) Look at the real world and see what the constants should be once we have a physical identification for the vector fields. We will find that we obtain a self-consistent story if we take both constants to be zero, that is,

$$\mathrm{div}\,\boldsymbol{B} = 0; \quad \mathrm{div}\,\boldsymbol{D} - \rho_{\mathrm{e}} = 0. \tag{7.10}$$

We can now write these relations in integral form. Integrate the expressions (7.10) over a closed volume v and apply the divergence theorem:

$$\int_v \mathrm{div}\,\boldsymbol{B}\,\mathrm{d}v = 0 \qquad \text{and hence} \qquad \int_S \boldsymbol{B}\cdot\mathrm{d}\boldsymbol{S} = 0. \tag{7.11}$$

Similarly, for the $\boldsymbol{D}$ fields,

$$\int_v \mathrm{div}\,\boldsymbol{D}\,\mathrm{d}v = \int_v \rho_{\mathrm{e}}\,\mathrm{d}v, \qquad \text{and hence} \qquad \int_S \boldsymbol{D}\cdot\mathrm{d}\boldsymbol{S} = \int_v \rho_{\mathrm{e}}\,\mathrm{d}v. \tag{7.12}$$

(7.12) tells us that the fields $\boldsymbol{D}$ can originate on electric charges.

7.4 Time-Independent Fields as Conservative Fields of Force

Problem Show that, if the vector fields $\boldsymbol{E}$ and $\boldsymbol{B}$ are time independent, $\boldsymbol{E}$ must satisfy $\int_C \boldsymbol{E}\cdot\mathrm{d}\boldsymbol{s} = 0$, that is, it is a conservative field, and so can be written in the form $\boldsymbol{E} = -\mathrm{grad}\,\phi$, where ϕ is a scalar potential function. Prove that this is definitely not possible if the field $\boldsymbol{B}$ is time varying.

Proof Since $\partial\boldsymbol{B}/\partial t = 0$, (7.1) shows that curl $\boldsymbol{E} = 0$. Therefore, from Stokes' theorem (A5.2), we find

$$\oint_C \boldsymbol{E}\cdot\mathrm{d}\boldsymbol{s} = 0,$$

where the line integral of $\boldsymbol{E}$ is taken around a closed contour C. This is one of the ways of defining a *conservative* field (see Appendix A5.4). Since curl grad $\phi = 0$, where ϕ is a scalar function, $\boldsymbol{E}$ can be derived from the gradient of a scalar function,

$$\boldsymbol{E} = -\mathrm{grad}\,\phi. \tag{7.13}$$

If $\boldsymbol{B}$ is time varying, we find that

$$\operatorname{curl} \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} \neq 0.$$

Thus, if we try to derive $\boldsymbol{E}$ entirely from the gradient of a scalar potential, we find that

$$-\operatorname{curl} \operatorname{grad} \phi = 0 = \frac{\partial \boldsymbol{B}}{\partial t} \neq 0.$$

Thus, $\boldsymbol{E}$ cannot be wholly expressed as $-\operatorname{grad} \phi$, if $\boldsymbol{B}$ is time varying.

7.5 Boundary Conditions in Electromagnetism

Across any sharp boundary between two media, the properties of the fields are expected to change discontinuously. There are three cases to consider.

Case 1

From Section 7.3, show that the component of the vector field $\boldsymbol{B}$ perpendicular to the surface between the media, $\boldsymbol{B} \cdot \boldsymbol{n}$, is continuous at the boundary and likewise that $\boldsymbol{D} \cdot \boldsymbol{n}$ is continuous, if there are no surface charges on the boundary. If there are surface charges of surface charge density σ, show that $(\boldsymbol{D}_2 - \boldsymbol{D}_1) \cdot \boldsymbol{n} = \sigma$.

Proof We erect a very short cylinder across the boundary (Fig. 7.1(a)) and then apply the rules which we know are true for the fluxes of the vector fields $\boldsymbol{B}$ and $\boldsymbol{D}$ through this volume under all circumstances. The diagram shows the $\boldsymbol{B}$ and $\boldsymbol{D}$ fields passing through the boundary of the cylinder.

First use Gauss's law for the vector field $\boldsymbol{B}$ through the surface of the cylinder in integral form,

$$\int_S \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S} = 0. \qquad (7.14)$$

Now squash the cylinder until it is infinitesimally thin. Then the surface area round the edge of the cylinder $2\pi r\, \mathrm{d}l$ tends to zero as $\mathrm{d}l$ tends to zero, and only the components

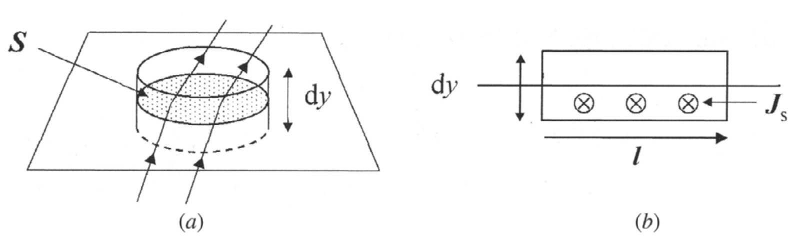


Fig. 7.1 Illustrating the boundary conditions for (a) the normal components of $\boldsymbol{B}$ and $\boldsymbol{D}$, and (b) the components of $\boldsymbol{E}$ and $\boldsymbol{H}$ parallel to the surface. The lower medium is labelled 1 and the upper medium 2.

through the upper and lower faces of the cylinder are left. If $\boldsymbol{n}$ is the unit vector normal to the surface, $\boldsymbol{B} \cdot \boldsymbol{n}$ is the flux of the vector field $\boldsymbol{B}$ per unit area. Taking medium 1 to be the lower layer and medium 2 the upper, the flux per unit area through the upper surface of the cylinder is $\boldsymbol{B}_2 \cdot \boldsymbol{n}$ and that entering through the lower surface is $-\boldsymbol{B}_1 \cdot \boldsymbol{n}$, recalling that d$\boldsymbol{S}$ is always taken to point outwards through the closed surface. Therefore, Gauss's theorem reduces to

$$(\boldsymbol{B}_2 - \boldsymbol{B}_1) \cdot \boldsymbol{n} = 0, \tag{7.15}$$

that is, the normal component of $\boldsymbol{B}$ is continuous.

In exactly the same way,

$$\int_S \boldsymbol{D} \cdot \mathrm{d}\boldsymbol{S} = \int_v \rho_\mathrm{e}\, \mathrm{d}v. \tag{7.16}$$

When we squash the cylinder, the left-hand side becomes the difference in the values of $\boldsymbol{D} \cdot \mathrm{d}\boldsymbol{S}$ on either side of the surface and $\int_v \rho_\mathrm{e}\, \mathrm{d}v$ becomes the surface charge $\sigma\, \mathrm{d}S$, that is

$$(\boldsymbol{D}_2 - \boldsymbol{D}_1) \cdot \boldsymbol{n} = \sigma. \tag{7.17}$$

If there are no surface charges, $\sigma = 0$ and the normal component of $\boldsymbol{D}$ is continuous.

Case 2

If the fields $\boldsymbol{E}$ and $\boldsymbol{H}$ are static fields (as also are $\boldsymbol{D}$, $\boldsymbol{B}$ and $\boldsymbol{J}$), show that (*a*) the tangential component of $\boldsymbol{E}$ is continuous at the boundary, and (*b*) the tangential component of $\boldsymbol{H}$ is continuous if there is no surface current density $\boldsymbol{J}_\mathrm{s}$. If there is, show that

$$\boldsymbol{n} \times (\boldsymbol{H}_2 - \boldsymbol{H}_1) = \boldsymbol{J}_\mathrm{s}.$$

Proof

(*a*) Create a long rectangular loop which crosses the surface between the two media as shown in Fig. 7.1(*b*). The sides of the loop perpendicular to the surface d$\boldsymbol{y}$ are very much shorter than the length of the horizontal arms d$\boldsymbol{l}$. The fields are static and therefore, from (7.1), curl $\boldsymbol{E} = 0$ or, applying Stokes' theorem, $\oint_C \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{s} = 0$. We apply this result to the long rectangular loop.

Let $\boldsymbol{E}_1$ and $\boldsymbol{E}_2$ be the fields on lower and upper sides respectively of the interface between the media. The loop can be squashed to infinitesimal thickness so that the ends d$\boldsymbol{y}$ make no contribution to the line integral. Projecting the electric field strengths $\boldsymbol{E}_1$ and $\boldsymbol{E}_2$ onto d$\boldsymbol{l}$, the line integrals along the upper and lower arms of the rectangle are

$$\int \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{s} = \boldsymbol{E}_2 \cdot \mathrm{d}\boldsymbol{l}; \qquad \int \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{s} = -\boldsymbol{E}_1 \cdot \mathrm{d}\boldsymbol{l}.$$

Therefore, the line integral round the complete loop is

$$\oint_C \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{s} = (\boldsymbol{E}_2 - \boldsymbol{E}_1) \cdot \mathrm{d}\boldsymbol{l} = 0.$$

We need to express this result in terms of the unit vector $\boldsymbol{n}$ normal to the interface. Let the unit vectors in the directions parallel to $\mathrm{d}\boldsymbol{l}$ be $\boldsymbol{i}_l$ and perpendicular to the loop $\boldsymbol{i}_\perp$, such that $\boldsymbol{i}_l = (\boldsymbol{i}_\perp \times \boldsymbol{n})$. Then, the above expression becomes

$$\begin{aligned}(\boldsymbol{E}_2 - \boldsymbol{E}_1) \cdot \mathrm{d}\boldsymbol{l} &= (\boldsymbol{E}_2 - \boldsymbol{E}_1) \cdot (\boldsymbol{i}_\perp \times \boldsymbol{n})\,|\mathrm{d}\boldsymbol{l}| = 0 \\ &= \boldsymbol{i}_\perp \cdot [(\boldsymbol{n} \times (\boldsymbol{E}_2 - \boldsymbol{E}_1)]\,|\mathrm{d}\boldsymbol{l}| = 0. \qquad (7.18)\end{aligned}$$

This result must be true for all orientations of the rectangular loop and so

$$\boldsymbol{n} \times (\boldsymbol{E}_2 - \boldsymbol{E}_1) = 0, \qquad (7.19)$$

that is, the tangential components of $\boldsymbol{E}$ are equal on either side of the interface.

(*b*) Now carry out the same analysis for the case of static magnetic fields for which $\mathrm{curl}\,\boldsymbol{H} = \boldsymbol{J}$. Integrating (7.2) through the area $\mathrm{d}\boldsymbol{S}$ and applying Stokes' theorem:

$$\int_S \mathrm{curl}\,\boldsymbol{H} \cdot \mathrm{d}\boldsymbol{S} = \int_S \boldsymbol{J} \cdot \mathrm{d}\boldsymbol{S}; \qquad \oint_C \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{s} = \int_S \boldsymbol{J} \cdot \mathrm{d}\boldsymbol{S}. \qquad (7.20)$$

Now apply this result to the rectangular loop shown in Fig. 7.1(*b*) as it is squashed to zero thickness. Exactly as in (*a*), the line integral of $\boldsymbol{H}$ round the loop is

$$\oint_C \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{s} = \boldsymbol{i}_\perp \cdot [\boldsymbol{n} \times (\boldsymbol{H}_2 - \boldsymbol{H}_1)]\,|\mathrm{d}\boldsymbol{l}|. \qquad (7.21)$$

The total current flowing through the elementary loop $\int_S \boldsymbol{J}\cdot\mathrm{d}\boldsymbol{S}$ can be rewritten $\boldsymbol{J}\cdot\mathrm{d}\boldsymbol{S} = (\boldsymbol{J} \cdot \boldsymbol{i}_\perp)\,|\mathrm{d}\boldsymbol{s}|\,|\mathrm{d}l| = (\boldsymbol{J}_\mathrm{s} \cdot \boldsymbol{i}_\perp)\,|\mathrm{d}\boldsymbol{l}|$, where $\boldsymbol{J}_\mathrm{s} = \boldsymbol{J}\,|\mathrm{d}\boldsymbol{s}|$ as $\mathrm{d}\boldsymbol{s} \to 0$, is the surface current density, that is, the current per unit length.

We can therefore rewrite (7.20)

$$\boldsymbol{i}_\perp \cdot [\boldsymbol{n} \times (\boldsymbol{H}_2 - \boldsymbol{H}_1)]\,|\mathrm{d}\boldsymbol{l}| = \boldsymbol{i}_\perp \cdot \boldsymbol{J}_\mathrm{s}\,|\mathrm{d}\boldsymbol{l}|,$$

or

$$\boldsymbol{n} \times (\boldsymbol{H}_2 - \boldsymbol{H}_1) = \boldsymbol{J}_\mathrm{s}. \qquad (7.22)$$

If there are no surface currents, $\boldsymbol{J}_\mathrm{s} = 0$ and $\boldsymbol{H} \times \boldsymbol{n}$, the tangential component of $\boldsymbol{H}$, is continuous.

At this stage, it is not so obvious that the relations (7.19) and (7.22) are true in the presence of time-varying fields because we have discarded the time-varying components. In fact, they are true, as we demonstrate in Case 3.

Case 3

Prove that the statements (7.19) and (7.22) in Case 2 are correct even in the presence of time-varying fields.

Proof First integrate (7.1) over a surface S and apply Stokes' theorem.

$$\int_S \mathrm{curl}\,\boldsymbol{E} \cdot \mathrm{d}\boldsymbol{S} = -\int_S \frac{\partial \boldsymbol{B}}{\partial t} \cdot \mathrm{d}\boldsymbol{S},$$

$$\oint_C \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{s} = -\int_S \frac{\partial \boldsymbol{B}}{\partial t} \cdot \mathrm{d}\boldsymbol{S}.$$

Now apply this result to the little rectangle in Fig. 7.1(*b*). The analysis proceeds as before, except that there is a time-varying component on the right-hand side. Following the same analysis which led to (7.19),

$$\boldsymbol{i}_\perp \cdot [\boldsymbol{n} \times (\boldsymbol{E}_2 - \boldsymbol{E}_1)]\,|\mathrm{d}\boldsymbol{l}| = -\frac{\partial \boldsymbol{B}}{\partial t} \cdot \boldsymbol{i}_\perp\,|\mathrm{d}\boldsymbol{l}|\,|\mathrm{d}\boldsymbol{y}| + \text{small end contributions.}$$

Now cancel out the $|\mathrm{d}\boldsymbol{l}|$s on either side and let the width of the rectangle shrink to zero:

$$\boldsymbol{n} \times (\boldsymbol{E}_2 - \boldsymbol{E}_1) = -|\mathrm{d}\boldsymbol{y}|\left(\frac{\partial \boldsymbol{B}}{\partial t}\right)_{\mathrm{d}\boldsymbol{y}\to 0}.$$

$\partial \boldsymbol{B}/\partial t$ is a finite quantity in the case of time-varying fields, but in the limit in which $\mathrm{d}\boldsymbol{y}$ tends to zero, the right-hand side makes no contribution, that is, as before

$$(\boldsymbol{E}_2 - \boldsymbol{E}_1) \times \boldsymbol{n} = 0.$$

Exactly the same analysis can be carried out for time-varying $\boldsymbol{D}$ fields, that is,

$$\int_S \operatorname{curl} \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{S} = \int_S \boldsymbol{J} \cdot \mathrm{d}\boldsymbol{S} + \int_S \frac{\partial \boldsymbol{D}}{\partial t} \cdot \mathrm{d}\boldsymbol{S},$$

$$\oint_C \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{S} = \int_S \boldsymbol{J} \cdot \mathrm{d}\boldsymbol{S} + \int_S \frac{\partial \boldsymbol{D}}{\partial t} \cdot \mathrm{d}\boldsymbol{S}.$$

Now, squash the rectangle to zero thickness, following the same reasoning which led to (7.22).

$$\boldsymbol{i}_\perp \cdot [\boldsymbol{n} \times (\boldsymbol{H}_2 - \boldsymbol{H}_1)]\,|\mathrm{d}\boldsymbol{l}| = \boldsymbol{i}_\perp \cdot \boldsymbol{J}_\mathrm{s}\,|\mathrm{d}\boldsymbol{l}| + \frac{\partial \boldsymbol{D}}{\partial t} \cdot \boldsymbol{i}_\perp\,|\mathrm{d}\boldsymbol{l}|\,|\mathrm{d}\boldsymbol{y} \quad + \text{ small end contributions,}$$

that is,

$$\boldsymbol{n} \times (\boldsymbol{H}_2 - \boldsymbol{H}_1) = \boldsymbol{J}_\mathrm{s} + \left(\frac{\partial \boldsymbol{D}}{\partial t}\right)_{\mathrm{d}\boldsymbol{y}\to 0} |\mathrm{d}\boldsymbol{y}|,$$

$$\boldsymbol{n} \times (\boldsymbol{H}_2 - \boldsymbol{H}_1) = \boldsymbol{J}_\mathrm{s},$$

even in the presence of time-varying $\boldsymbol{D}$ fields.

7.6 Ampère's Law

Problem If $\boldsymbol{H}$ is not time varying – we will eventually identify this sort of $\boldsymbol{H}$ with a magnetostatic field – show that $\boldsymbol{H}$ can be expressed in the form $-\operatorname{grad} V_{\mathrm{mag}}$ when there are only permanent magnets and magnetisable materials but *no* currents. If there are steady currents, show that $\oint_C \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{s} = I_{\mathrm{enclosed}}$. This equation is often used even for time-varying fields. Why is this so?

Proof In the absence of currents, $\boldsymbol{J} = 0$, and so, from (7.2), $\operatorname{curl} \boldsymbol{H} = 0$. $\boldsymbol{H}$ is therefore a conservative field and can be expressed as the gradient of a scalar field:

$$\boldsymbol{H} = -\operatorname{grad} V_{\mathrm{mag}}.$$

If *steady* currents are present, curl $\boldsymbol{H} = \boldsymbol{J}$. Integrating over a surface S and applying Stokes' theorem:

$$\int_S \operatorname{curl} \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{S} = \int_S \boldsymbol{J} \cdot \mathrm{d}\boldsymbol{S},$$

$$\oint_C \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{s} = (\text{total enclosed current}) = I_{\text{enclosed}}. \tag{7.23}$$

This equation can be used in the presence of time-varying $\boldsymbol{D}$ fields, *provided they are very slowly varying*, that is, $\partial \boldsymbol{D}/\partial t \ll \boldsymbol{J}$. This is very often a good approximation, but we have to check that it is valid when any of the fields or currents are varying.

7.7 Faraday's Law

Problem Derive Faraday's law, $\oint_C \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{s} = -\mathrm{d}\Phi/\mathrm{d}t$, where $\Phi = \int_S \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S}$ is the flux of the field $\boldsymbol{B}$ through the closed contour C. In this formulation, $\boldsymbol{E}$ is the *total* electric field, whereas the usual statement of Faraday's law refers only to the induced field. Explain this discrepancy.

Proof From (7.1), and using Stokes' theorem,

$$\int_S \operatorname{curl} \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{S} = -\frac{\partial}{\partial t}\left(\int_S \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S}\right),$$

that is,

$$\oint_C \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{s} = -\frac{\partial \Phi}{\partial t}. \tag{7.24}$$

Normally, Faraday's law refers only to the induced part of the field $\boldsymbol{E}$. There may also, however, be a component due to an electrostatic field, that is

$$\begin{aligned} \boldsymbol{E} &= \boldsymbol{E}_{\text{induced}} + \boldsymbol{E}_{\text{electrostatic}} \\ &= \boldsymbol{E}_{\text{induced}} - \operatorname{grad} V, \\ \operatorname{curl} \boldsymbol{E} &= \operatorname{curl} \boldsymbol{E}_{\text{induced}} - \operatorname{curl} \operatorname{grad} V \\ &= -\frac{\partial \boldsymbol{B}}{\partial t}, \end{aligned}$$

which shows that $\boldsymbol{E}$ refers to the total field including the electrostatic part, but the latter does not survive curling.

7.8 Coulomb's Law

The analysis so far has been based upon the mathematical properties of a set of equations describing vector fields. Although we have used words such as magnets, electrostatic field and so on, the properties of the equations are independent of these physical identifications. We now need to give the fields $\boldsymbol{E}$, $\boldsymbol{D}$, $\boldsymbol{H}$ and $\boldsymbol{B}$ physical meaning. First of all, we *define* $q\boldsymbol{E}$ to be the force acting on a stationary charge q. Note that this statement is following Maxwell's agenda according to which the fields are *primary* and the forces *secondary* in the sense that they are derived from the responses of material bodies to the fields.

Problem The electrostatic field in a vacuum is associated with a charge distribution ρ_e. Show that the field satisfies Poisson's equation $\nabla^2 V = -\rho_e/\epsilon_0$. From the definition of $\boldsymbol{E}$ and the results of Section 7.4, show that V is the work done in bringing unit charge from infinity to the point $\boldsymbol{r}$ in the field. Show that the solution of Poisson's equation is

$$V(\boldsymbol{r}) = \int \frac{\rho_e(\boldsymbol{r}')}{4\pi\epsilon_0|\boldsymbol{r} - \boldsymbol{r}'|}\, d^3\boldsymbol{r}'.$$

Hence derive Coulomb's law of electrostatics, $F = q_1 q_2/4\pi\epsilon_0 r^2$.

Proof We have shown that $\operatorname{div}\boldsymbol{D} = \rho_e$ and, in a vacuum, we define $\boldsymbol{D} = \epsilon_0\boldsymbol{E}$. Therefore,

$$\operatorname{div}(\epsilon_0\boldsymbol{E}) = \rho_e \quad \text{and} \quad \operatorname{div}(\epsilon_0 \operatorname{grad} V) = -\rho_e,$$

$$\nabla^2 V = -\frac{\rho_e}{\epsilon_0}. \tag{7.25}$$

This is *Poisson's equation for electrostatics in a vacuum with a charge distribution* ρ_e.

The work done is defined to be that expended against the field $\boldsymbol{E}$ to bring the charge q from infinity to $\boldsymbol{r}$, that is,

$$\text{Work done} = -\int_\infty^r \boldsymbol{f}\cdot d\boldsymbol{r} = -q\int_\infty^r \boldsymbol{E}\cdot d\boldsymbol{r} = q\int_\infty^r \operatorname{grad} V\cdot d\boldsymbol{r} = qV, \tag{7.26}$$

where the electrostatic potential V is assumed to be zero at infinity. Thus, V measures the work done against an electrostatic field in bringing unit charge from infinity to that point in the field. Note that V is *not* the amount of energy needed to set up the electric field distribution in the first place. This will be evaluated in Section 7.11.

To find the solution of Poisson's equation, it is simplest to work backwards. The reason is that *Laplace's equation* $\nabla^2 V = 0$ is a linear equation for which the principle of superposition applies. Thus, if V_1 and V_2 are two independent solutions of Laplace's equation, $a_1 V_1 + a_2 V_2$ is also a solution. Since we can imagine the charge distribution ρ_e to be the sum of a distribution of point charges in a vacuum, the field at any point is the superposition of the vacuum solutions of Laplace's equation which can be related to the sources of the field, that is, to the charges q. Let us therefore consider only an elementary volume of the charge distribution located at $\boldsymbol{r}'$:

$$\rho_e(\boldsymbol{r}')\, d^3\boldsymbol{r}' = q(\boldsymbol{r}').$$

The proposed solution of Laplace's equation is

$$V(r) = \frac{\rho_e(\boldsymbol{r}')\, d^3\boldsymbol{r}'}{4\pi\epsilon_0|\boldsymbol{r} - \boldsymbol{r}'|} = \frac{q}{4\pi\epsilon_0 r}, \tag{7.27}$$

where r is the radial distance from the charge q at point $\boldsymbol{r}'$. Let us test this vacuum solution away from the origin at $r = 0$. From the equations for $\nabla^2 V$ in Section A5.5, Laplace's equation in spherical polar coordinates is

$$\nabla^2 V = \frac{1}{r}\frac{\partial^2}{\partial r^2}(rV) + \frac{1}{r^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial V}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 V}{\partial\phi^2}\right] = 0.$$

We need only consider the radial coordinate because of the spherical symmetry of the problem. Substituting the proposed solution for $V(r)$ into $(1/r)(\partial^2/\partial r^2)(rV)$,

$$\nabla^2 V = \frac{1}{r}\frac{\partial^2}{\partial r^2}(rV) = \frac{1}{r}\frac{\partial^2}{\partial r^2}\left(\frac{q}{4\pi\epsilon_0}\right) = 0,$$

provided $r \neq 0$. Thus, the proposed solution satisfies Laplace's equation away from the origin.

Now evaluate the charge Q enclosed by a sphere centred on the origin:

$$Q = \int_v \rho_e\,dv = -\int_v \epsilon_0 \nabla^2 V\,dv.$$

$V = q/4\pi\epsilon_0 r$ and hence the total enclosed charge is

$$Q = -\frac{q}{4\pi}\int_v \operatorname{div}\left(\operatorname{grad}\frac{1}{r}\right)dv.$$

Using the divergence theorem,

$$Q = -\frac{q}{4\pi}\int_S \operatorname{grad}\left(\frac{1}{r}\right)\cdot d\boldsymbol{S} = \frac{q}{4\pi}\int_S \frac{dS}{r^2} = \frac{q}{4\pi}\int_S d\Omega = q,$$

where dS is the element of area perpendicular to the radial direction and $d\Omega$ is the element of solid angle. Thus, according to Poisson's equation, the correct value of the charge is found at the origin from the proposed solution. We conclude that the solution

$$V(\boldsymbol{r}) = \frac{\rho_e(\boldsymbol{r}')\,d^3\boldsymbol{r}'}{4\pi\epsilon_0|\boldsymbol{r} - \boldsymbol{r}'|}$$

satisfies Poisson's equation in electrostatics.

Returning again to a single particle q_1 at the origin, $\nabla^2 V = q_1/\epsilon_0$. The force on another particle q_2 is $q_2\boldsymbol{E}$, that is

$$\begin{aligned} \boldsymbol{f} &= -q_2 \operatorname{grad} V = -\frac{q_1 q_2}{4\pi\epsilon_0}\operatorname{grad}\left(\frac{1}{r}\right) \\ &= \frac{q_1 q_2}{4\pi\epsilon_0 r^2}\boldsymbol{i}_r = \frac{q_1 q_2}{4\pi\epsilon_0}\frac{\boldsymbol{r}}{r^3}. \end{aligned} \quad (7.28)$$

This is the derivation of *Coulomb's inverse square law in electrostatics*. It may look as though we have taken a rather circuitous route to reach the point at which most courses in electricity and magnetism begin. The point of interest is that we can start with Maxwell's equations and deduce Coulomb's force law rigorously with a minimum number of assumptions.

7.9 The Biot–Savart Law

Problem Steady or slowly varying currents are flowing in a vacuum. What is 'slowly varying'? Show that curl $\boldsymbol{H} = \boldsymbol{J}$.

It can be shown that the solution of this equation is

$$\boldsymbol{H}(\boldsymbol{r}) = \int \frac{\boldsymbol{J}(\boldsymbol{r}') \times (\boldsymbol{r} - \boldsymbol{r}')}{4\pi|\boldsymbol{r} - \boldsymbol{r}'|^3} \, \mathrm{d}^3\boldsymbol{r}'. \tag{7.29}$$

Show that this is the Biot–Savart law $\mathrm{d}\boldsymbol{H} = I \sin\theta \, (\mathrm{d}\boldsymbol{s} \times \boldsymbol{r})/4\pi r^3$ for the contribution to the magnetic field strength $\boldsymbol{H}$ at distance $\boldsymbol{r}$ from the current element $I\,\mathrm{d}\boldsymbol{s}$.

Proof We have already discussed the issue of slowly varying fields in Section 7.6. 'Slowly varying' means that the displacement current term $\partial\boldsymbol{D}/\partial t \ll \boldsymbol{J}$. Then, from (7.2), we obtain curl $\boldsymbol{H} = \boldsymbol{J}$.

Just as in Section 7.8, where the solution of $\nabla^2 V = \rho_e/\epsilon_0$ was required, so here we need the solution of curl $\boldsymbol{H} = \boldsymbol{J}$. The general solution (7.29) is found in the standard textbooks and need not be repeated here. Let us differentiate (7.29) so that the relation becomes one between a current element and the field at a distance $|\boldsymbol{r} - \boldsymbol{r}'|$:

$$\mathrm{d}\boldsymbol{H} = \frac{\boldsymbol{J}(\boldsymbol{r}') \times (\boldsymbol{r} - \boldsymbol{r}')}{4\pi|\boldsymbol{r} - \boldsymbol{r}'|^3} \, \mathrm{d}^3\boldsymbol{r}'. \tag{7.30}$$

Now we identify $\boldsymbol{J}(\boldsymbol{r}')\,\mathrm{d}^3\mathrm{r}'$ with the current element $I\mathrm{d}\boldsymbol{s}$ and so

$$\mathrm{d}\boldsymbol{H} = \frac{I\,\mathrm{d}\boldsymbol{s} \times \boldsymbol{r}}{4\pi r^3}, \qquad |\mathrm{d}\boldsymbol{H}| = \frac{I \sin\theta \, |\mathrm{d}\boldsymbol{s}|}{4\pi r^2}, \tag{7.31}$$

where, by $\boldsymbol{r}$, we now mean the vector from the current element to the point in the field. This is the *Biot–Savart law*.

7.10 The Interpretation of Maxwell's Equations in Material Media

We have already given a physical meaning to the electric field strength $\boldsymbol{E}$ in Section 7.6 and $\boldsymbol{J}$ is defined by (7.3) and (7.4). We still have to define $\boldsymbol{D}$, $\boldsymbol{B}$ and $\boldsymbol{H}$ but we have only two independent equations left, (7.1) and (7.2) – we have not yet closed the set of equations. In order to do so, we have to introduce a further set of definitions based upon experimental evidence. First of all, we consider the behaviour of electromagnetic fields *in vacuo*. From experiment we infer that

$$\text{in vacuo,} \quad \begin{cases} \boldsymbol{D} = \epsilon_0 \boldsymbol{E}, \\ \boldsymbol{B} = \mu_0 \boldsymbol{H}, \end{cases} \tag{7.32}$$

where ϵ_0 and μ_0 are constants, $\boldsymbol{D}$ is the electric flux density, $\boldsymbol{E}$ the electric field strength, $\boldsymbol{H}$ the magnetic field strength and $\boldsymbol{B}$ the magnetic flux density. The relations between $\boldsymbol{D}$ and $\boldsymbol{E}$ and between $\boldsymbol{B}$ and $\boldsymbol{H}$ are known as *constitutive equations* or *constitutive relations*.

Inside material media, the relations (7.32) are not correct and so we introduce vectors which describe the difference between the actual values and the vacuum definitions – these differences are referred to as the electric and magnetic *polarisation properties* of the material:[3]

$$\boldsymbol{P} = \boldsymbol{D} - \epsilon_0 \boldsymbol{E}, \tag{7.33}$$

$$\boldsymbol{M} = \frac{\boldsymbol{B}}{\mu_0} - \boldsymbol{H}. \tag{7.34}$$

Granted these definitions, our objective is to find physical meanings for the polarisation vectors, $\boldsymbol{P}$ and $\boldsymbol{M}$. Let us proceed with the mathematical analysis of the equations and see if these definitions lead to consistency with all the known laws of electricity and magnetism.

Problem Using the above definitions and the results of Section 7.3, show that Maxwell's equations can be written in the form

$$\operatorname{curl} \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \tag{7.35a}$$

$$\operatorname{div} \boldsymbol{B} = 0, \tag{7.35b}$$

$$\operatorname{div} \epsilon_0 \boldsymbol{E} = (\rho_e - \operatorname{div} \boldsymbol{P}), \tag{7.35c}$$

$$\operatorname{curl}(\boldsymbol{B}/\mu_0) = \left(\boldsymbol{J} + \frac{\partial \boldsymbol{P}}{\partial t} + \operatorname{curl} \boldsymbol{M}\right) + \frac{\partial(\epsilon_0 \boldsymbol{E})}{\partial t}. \tag{7.35d}$$

Show that, in *electrostatics*, the field $\boldsymbol{E}$ may be calculated correctly everywhere by replacing polarisable media by a vacuum together with a volume charge distribution $-\operatorname{div} \boldsymbol{P}$ and a surface charge density $\boldsymbol{P} \cdot \boldsymbol{n}$. Then write down the expression for the electrostatic potential V at any point in space in terms of these charges and show that $\boldsymbol{P}$ represents the dipole moment per unit volume within the medium. You may use the fact that the electrostatic potential of an electric dipole of dipole moment $\boldsymbol{p}$ can be written

$$V = \frac{1}{4\pi\epsilon_0} \boldsymbol{p} \cdot \nabla\left(\frac{1}{r}\right). \tag{7.36}$$

In *magnetostatics*, show that the $\boldsymbol{B}$ field may be correctly calculated by replacing polarisable bodies by a current distribution $\operatorname{curl} \boldsymbol{M}$ with surface current densities $-\boldsymbol{n} \times \boldsymbol{M}$. Then write down the expression for the magnetostatic vector potential $\boldsymbol{A}$ at any point in space in terms of these currents and show that $\boldsymbol{M}$ represents the magnetic dipole moment per unit volume within the medium. You may use the fact that the magnetostatic vector potential of a magnetic dipole of magnetic dipole moment $\boldsymbol{m}$ at distance r can be written

$$\boldsymbol{A} = \frac{\mu_0}{4\pi} \boldsymbol{m} \times \nabla\left(\frac{1}{r}\right). \tag{7.37}$$

Proof The first part is straightforward and simply involves rewriting Maxwell's equations. (7.35a) and (7.35b) are general results already discussed in Section 7.2:

$$\operatorname{curl} \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} \quad \text{and} \quad \operatorname{div} \boldsymbol{B} = 0.$$

To derive (7.35c), since $\operatorname{div}\boldsymbol{D} = \rho_e$ and $\boldsymbol{D} = \boldsymbol{P} + \epsilon_0\boldsymbol{E}$,

$$\operatorname{div}\epsilon_0\boldsymbol{E} = \rho_e - \operatorname{div}\boldsymbol{P}. \tag{7.38}$$

Similarly, from equation (7.2) and the definition $\boldsymbol{H} = (\boldsymbol{B}/\mu_0) - \boldsymbol{M}$,

$$\begin{aligned}
\operatorname{curl}\boldsymbol{H} &= \boldsymbol{J} + \frac{\partial\boldsymbol{D}}{\partial t},\\
\operatorname{curl}\left(\frac{\boldsymbol{B}}{\mu_0} - \boldsymbol{M}\right) &= \boldsymbol{J} + \frac{\partial\boldsymbol{P}}{\partial t} + \frac{\partial(\epsilon_0\boldsymbol{E})}{\partial t},\\
\operatorname{curl}\left(\frac{\boldsymbol{B}}{\mu_0}\right) &= \left(\boldsymbol{J} + \frac{\partial\boldsymbol{P}}{\partial t} + \operatorname{curl}\boldsymbol{M}\right) + \frac{\partial(\epsilon_0\boldsymbol{E})}{\partial t}.
\end{aligned} \tag{7.39}$$

Now let us consider the interpretation of the set of equations (7.35).

Electrostatics

Equation (7.35c) states that to calculate $\operatorname{div}\epsilon_0\boldsymbol{E}$ at any point in space, we have to add to ρ_e the quantity $-\operatorname{div}\boldsymbol{P}$ which we may think of as an effective charge $\rho_e^* = -\operatorname{div}\boldsymbol{P}$ in addition to the charge ρ_e.

Now consider the boundary conditions between the region containing the polarisable medium (medium 1) and the vacuum (medium 2). From Section 7.5, we know that, under all circumstances, the component of $\boldsymbol{D}$ normal to the boundary is continuous:

$$(\boldsymbol{D}_2 - \boldsymbol{D}_1)\cdot\boldsymbol{n} = \sigma, \tag{7.40}$$

where σ is the surface charge density. We suppose that there are no free charges on the surface, $\sigma = 0$. Therefore,

$$\begin{aligned}
(\boldsymbol{D}_2 - \boldsymbol{D}_1)\cdot\boldsymbol{n} &= 0,\\
[\epsilon_0\boldsymbol{E}_2 - (\epsilon_0\boldsymbol{E}_1 + \boldsymbol{P}_1)]\cdot\boldsymbol{n} &= 0,\\
[\epsilon_0\boldsymbol{E}_2 - \epsilon_0\boldsymbol{E}_1] &= \boldsymbol{P}_1\cdot\boldsymbol{n}.
\end{aligned} \tag{7.41}$$

Thus, when we replace the medium by a vacuum, there must also be an effective surface charge density $\sigma^* = \boldsymbol{P}_1\cdot\boldsymbol{n}$ per unit area at the boundary.

In calculating the $\boldsymbol{E}$ field anywhere in space, we find the correct answer if we replace the material medium by a charge distribution $\rho_e^* = -\operatorname{div}\boldsymbol{P}$ and a surface charge distribution $\sigma^* = \boldsymbol{P}\cdot\boldsymbol{n}$ per unit area. Suppose we replace the material medium by a vacuum but with a *real charge distributions* ρ_e^* and a real surface charge distribution σ^* – what is the field at any point in space?

The total electric field due to these charges is obtained from the gradient of the electrostatic potential V where

$$\begin{aligned}
V &= \frac{1}{4\pi\epsilon_0}\int_v \frac{\rho_e^*}{r}\,\mathrm{d}v + \frac{1}{4\pi\epsilon_0}\int_S \frac{\sigma^*}{r}\,\mathrm{d}S\\
&= -\frac{1}{4\pi\epsilon_0}\int_v \frac{\operatorname{div}\boldsymbol{P}}{r}\,\mathrm{d}v + \frac{1}{4\pi\epsilon_0}\int_S \frac{\boldsymbol{P}\cdot\mathrm{d}\boldsymbol{S}}{r}.
\end{aligned}$$

Applying the divergence theorem to the second integral, we obtain

$$V = \frac{1}{4\pi\epsilon_0}\int_v \left[-\frac{\operatorname{div}\boldsymbol{P}}{r} + \operatorname{div}\left(\frac{\boldsymbol{P}}{r}\right)\right] \mathrm{d}v.$$

But

$$\operatorname{div}\left(\frac{\boldsymbol{P}}{r}\right) = \frac{\operatorname{div}\boldsymbol{P}}{r} + \boldsymbol{P}\cdot\operatorname{grad}\frac{1}{r},$$

and hence

$$V = \frac{1}{4\pi\epsilon_0}\int_v \boldsymbol{P}\cdot\operatorname{grad}\left(\frac{1}{r}\right)\mathrm{d}v. \tag{7.42}$$

This is the rather beautiful result we have been seeking. We recall that the potential at distance $\boldsymbol{r}$ from an electrostatic dipole is

$$V = \frac{1}{4\pi\epsilon_0}\boldsymbol{p}\cdot\operatorname{grad}\left(\frac{1}{r}\right),$$

where $\boldsymbol{p}$ is the electric dipole moment of the dipole. Thus, we can interpret (7.42) as showing that the quantity $\boldsymbol{P}$ is the *dipole moment per unit volume* within the material. However, the dipoles at the surfaces of the material result in net positive charges on one side of the material and negative charges on the opposite side, so that the whole system remains neutral. It is these charges which give rise to the surface charge distribution of $\rho_e^* = \boldsymbol{P}\cdot\boldsymbol{n}$.

Magnetostatics

The analysis in the case of magnetostatics proceeds in exactly the same way. Equation (7.35d) states that we have to include a current density distribution $\boldsymbol{J}^* = \operatorname{curl}\boldsymbol{M}$ in the expression for $\boldsymbol{B}$ when we replace the medium by a vacuum. In the case of surfaces, (7.22) indicates that

$$\boldsymbol{n}\times(\boldsymbol{H}_2 - \boldsymbol{H}_1) = \boldsymbol{J}_s.$$

If there are no real surface currents, $\boldsymbol{J} = 0$,

$$\boldsymbol{n}\times(\boldsymbol{H}_2 - \boldsymbol{H}_1) = 0,$$

$$\boldsymbol{n}\times\left[\frac{\boldsymbol{B}_2}{\mu_0} - \left(\frac{\boldsymbol{B}_1}{\mu_0} - \boldsymbol{M}_1\right)\right] = 0.$$

Thus,

$$\boldsymbol{n}\times\left[\frac{\boldsymbol{B}_2}{\mu_0} - \frac{\boldsymbol{B}_1}{\mu_0}\right] = -(\boldsymbol{n}\times\boldsymbol{M}_1). \tag{7.43}$$

There must therefore be an effective surface current distribution $\boldsymbol{J}_s^* = -(\boldsymbol{n}\times\boldsymbol{M}_1)$.

We now carry out an analysis similar to that which illustrated the meaning of the polarisation $\boldsymbol{P}$. The expression for the vector potential $\boldsymbol{A}$ due to the current I flowing in a current distribution can be derived using (A5.20) and the vector identity (A5.13),[4]

$$\boldsymbol{A} = \frac{\mu_0}{4\pi} \int \frac{I\,\mathrm{d}\boldsymbol{s}}{r}. \tag{7.44}$$

Rewriting this expression in terms of the current density distribution $\boldsymbol{J}$, $I\,\mathrm{d}\boldsymbol{s} = \boldsymbol{J}\,\mathrm{d}\sigma\,\mathrm{d}l = \boldsymbol{J}\,\mathrm{d}v$ and so

$$\boldsymbol{A} = \frac{\mu_0}{4\pi} \int \frac{\boldsymbol{J}\,\mathrm{d}v}{r}.$$

The vector potential due to the charge distribution $\boldsymbol{J}^*$ and the surface current distribution $\boldsymbol{J}_\mathrm{s}^*$ is therefore

$$\begin{aligned}\boldsymbol{A} &= \frac{\mu_0}{4\pi} \int_v \frac{\boldsymbol{J}^*\,\mathrm{d}v}{r} + \frac{\mu_0}{4\pi} \int_S \frac{\boldsymbol{J}_\mathrm{s}^*\,\mathrm{d}S}{r} \\ &= \frac{\mu_0}{4\pi} \int_v \frac{\mathrm{curl}\,\boldsymbol{M}\,\mathrm{d}v}{r} - \frac{\mu_0}{4\pi} \int_S \frac{(\mathrm{d}\boldsymbol{S} \times \boldsymbol{M})}{r},\end{aligned} \tag{7.45}$$

since $\boldsymbol{n}\,\mathrm{d}S = \mathrm{d}\boldsymbol{S}$.

Now, from Appendix A5.2, Problem 3, we know that

$$\int_v \mathrm{curl}\,\boldsymbol{a}\,\mathrm{d}v = \int_S \mathrm{d}\boldsymbol{S} \times \boldsymbol{a}.$$

Therefore, setting $\boldsymbol{a} = \boldsymbol{M}/r$, we find

$$\int_v \mathrm{curl}\,\frac{\boldsymbol{M}}{r}\,\mathrm{d}v = \int_S \mathrm{d}\boldsymbol{S} \times \frac{\boldsymbol{M}}{r}. \tag{7.46}$$

Therefore (7.45) becomes

$$\boldsymbol{A} = \frac{\mu_0}{4\pi} \int_v \left[\frac{\mathrm{curl}\,\boldsymbol{M}}{r} - \mathrm{curl}\,\frac{\boldsymbol{M}}{r}\right] \mathrm{d}v. \tag{7.47}$$

But,

$$\mathrm{curl}\,x\boldsymbol{a} = x\,\mathrm{curl}\,\boldsymbol{a} - \boldsymbol{a} \times \mathrm{grad}\left(\frac{1}{r}\right).$$

Therefore,

$$\boldsymbol{A} = \frac{\mu_0}{4\pi} \int_v \boldsymbol{M} \times \mathrm{grad}\left(\frac{1}{r}\right) \mathrm{d}v. \tag{7.48}$$

This is the result we have been seeking. Comparison with (7.37) shows that $\boldsymbol{M}$ is the magnetic dipole moment per unit volume of the material.

7.11 The Energy Densities of Electromagnetic Fields

Problem From the definition of the electric field $\boldsymbol{E}$ in Section 7.8, show that the rate at which batteries have to do work to push charges and currents against electrostatic fields, including emfs but neglecting Ohmic heating, is

$$\int_v \boldsymbol{J} \cdot (-\boldsymbol{E})\,\mathrm{d}v,$$

and hence that the total energy of the system is

$$U = -\int_{\text{all space}} \mathrm{d}v \int_0^{\text{final fields}} \boldsymbol{J} \cdot \boldsymbol{E}\, \mathrm{d}t.$$

By taking the scalar products of (7.2) with $\boldsymbol{E}$ and (7.1) with $\boldsymbol{H}$, transform this expression into

$$U = \int_{\text{all space}} \mathrm{d}v \int_0^{\boldsymbol{D}} \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{D} + \int_{\text{all space}} \mathrm{d}v \int_0^{\boldsymbol{B}} \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{B}.$$

Proof This is one of the classic pieces of analysis in electromagnetic theory carried out by Maxwell in Part III of his paper of 1865 (see Section 6.3). We start with the work done by the electromagnetic field on a particle of charge q. In general, work is only done by the electric field because, in the case of magnetic fields, no work is done on the particle, the force acting perpendicular to the displacement of the particle $\boldsymbol{f} = q(\boldsymbol{u} \times \boldsymbol{B})$. Thus, as the particle moves from $\boldsymbol{r}$ to $\boldsymbol{r} + \mathrm{d}\boldsymbol{r}$ in unit time, the work done per second by the field is $q\boldsymbol{E} \cdot \mathrm{d}\boldsymbol{r}$, that is,

$$\text{Work done per second} = q\boldsymbol{E} \cdot \boldsymbol{u}.$$

Therefore, per unit volume, the work done per second is $qN\boldsymbol{E} \cdot \boldsymbol{u} = \boldsymbol{J} \cdot \boldsymbol{E}$, where N is the number density of charges q. Now integrate over all space to find the total work done per second by the currents:

$$\int_v \boldsymbol{J} \cdot \boldsymbol{E}\, \mathrm{d}v.$$

The origin of the power to drive these currents is batteries which must do work $\int_v (-\boldsymbol{J}) \cdot \boldsymbol{E}\, \mathrm{d}v$ on the system. Therefore, the total amount of energy supplied by the batteries is

$$U = -\int_{\text{all space}} \mathrm{d}v \int_0^t (\boldsymbol{J} \cdot \boldsymbol{E})\, \mathrm{d}t. \tag{7.49}$$

The rest of the analysis is straightforward. Now express $(\boldsymbol{J} \cdot \boldsymbol{E})$ in terms of $\boldsymbol{E}, \boldsymbol{D}, \boldsymbol{H}$ and $\boldsymbol{B}$. From (7.2),

$$\boldsymbol{E} \cdot \boldsymbol{J} = \boldsymbol{E} \cdot \left(\operatorname{curl} \boldsymbol{H} - \frac{\partial \boldsymbol{D}}{\partial t}\right). \tag{7.50}$$

Now take the scalar product of (7.1) with $\boldsymbol{H}$, add it to (7.50):

$$\boldsymbol{E} \cdot \boldsymbol{J} = \boldsymbol{E} \cdot \operatorname{curl} \boldsymbol{H} - \boldsymbol{H} \cdot \operatorname{curl} \boldsymbol{E} - \boldsymbol{E} \cdot \frac{\partial \boldsymbol{D}}{\partial t} - \boldsymbol{H} \cdot \frac{\partial \boldsymbol{B}}{\partial t}. \tag{7.51}$$

The vector relation $\nabla \cdot (\boldsymbol{a} \times \boldsymbol{b}) = \boldsymbol{b} \cdot (\nabla \times \boldsymbol{a}) - \boldsymbol{a} \cdot (\nabla \times \boldsymbol{b})$ enables us to write (7.51) as

$$\boldsymbol{E} \cdot \boldsymbol{J} = -\operatorname{div}(\boldsymbol{E} \times \boldsymbol{H}) - \boldsymbol{E} \cdot \frac{\partial \boldsymbol{D}}{\partial t} - \boldsymbol{H} \cdot \frac{\partial \boldsymbol{B}}{\partial t}.$$

Therefore, integrating with respect to time, we find the total energy supplied to the system by the batteries per unit volume,

$$-\int_0^t (\boldsymbol{J} \cdot \boldsymbol{E})\, \mathrm{d}t = \int_0^t \operatorname{div}(\boldsymbol{E} \times \boldsymbol{H})\, \mathrm{d}t + \int_0^{\boldsymbol{D}} \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{D} + \int_0^{\boldsymbol{B}} \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{B}. \tag{7.52}$$

Now integrate over all space, exchange the order of integration and apply the divergence theorem to the first integral on the right-hand side of (7.52). The latter term becomes

$$\int_0^t \int_v \operatorname{div}(\boldsymbol{E} \times \boldsymbol{H})\, \mathrm{d}v\, \mathrm{d}t = \int_0^t \int_S (\boldsymbol{E} \times \boldsymbol{H}) \cdot \mathrm{d}\boldsymbol{S}\, \mathrm{d}t.$$

The quantity $\boldsymbol{E} \times \boldsymbol{H}$ is known as the *Poynting vector* and is the rate of flow of energy in the direction normal to both $\boldsymbol{E}$ and $\boldsymbol{H}$ per unit area. Evidently, integrating the flux of $\boldsymbol{E} \times \boldsymbol{H}$ over the closed surface S represents the rate of loss of energy through the surface.

The other two terms in (7.52) represent the energies stored in the $\boldsymbol{E}$ and $\boldsymbol{H}$ fields per unit volume. We can therefore write the expression for the total energy in electric and magnetic fields:

$$U = \int_v \int_0^{\boldsymbol{D}} \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{D}\, \mathrm{d}v + \int_v \int_0^{\boldsymbol{B}} \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{B}\, \mathrm{d}v. \tag{7.53}$$

This is the answer we have been seeking. If the polarisable media are linear, $\boldsymbol{D} \propto \boldsymbol{E}$ and $\boldsymbol{B} \propto \boldsymbol{H}$, the total energy can be written

$$U = \int_v \tfrac{1}{2} (\boldsymbol{D} \cdot \boldsymbol{E} + \boldsymbol{B} \cdot \boldsymbol{H})\, \mathrm{d}v. \tag{7.54}$$

Problem Transform the first term on the right-hand side of (7.54) into the form $\int \frac{1}{2} \rho_e V\, \mathrm{d}v$ in the case in which the polarisable media are linear, that is, $\boldsymbol{D} \propto \boldsymbol{E}$.

Proof Let us do this one backwards.

$$\tfrac{1}{2} \int_v \rho_e V\, \mathrm{d}v = \tfrac{1}{2} \int_v (\operatorname{div} \boldsymbol{D}) V\, \mathrm{d}v.$$

Now $\operatorname{div}(V\boldsymbol{D}) = \operatorname{grad} V \cdot \boldsymbol{D} + V \operatorname{div} \boldsymbol{D}$ and hence substituting and using the divergence theorem

$$\begin{aligned}
\tfrac{1}{2} \int_v \rho_e V\, \mathrm{d}v &= \tfrac{1}{2} \int_v \operatorname{div}(V\boldsymbol{D})\, \mathrm{d}v - \tfrac{1}{2} \int_v \operatorname{grad} V \cdot \boldsymbol{D}\, \mathrm{d}v \\
&= \tfrac{1}{2} \int_S V\boldsymbol{D} \cdot \mathrm{d}\boldsymbol{S} + \tfrac{1}{2} \int_v \boldsymbol{E} \cdot \boldsymbol{D}\, \mathrm{d}v.
\end{aligned} \tag{7.55}$$

We now have to deal with the surface integral in (7.55). We allow the integral to extend over a very large volume and ask how V and $\boldsymbol{D}$ vary as the radius r tends to infinity. If the system is isolated, the lowest multipole electric field which could be present is associated with its net electric charge, for which $D \propto E \propto r^{-2}$ and $V \propto r^{-1}$. If the system were neutral, the lowest possible multipole of the field at a large distance would be a dipole field for which $D \propto r^{-3}, V \propto r^{-2}$. Therefore, the best we can do is to consider the lowest multipole case:

$$\tfrac{1}{2} \int_S V\boldsymbol{D} \cdot \mathrm{d}\boldsymbol{S} \propto \tfrac{1}{2} \int \frac{1}{r} \times \frac{1}{r^2} r^2\, \mathrm{d}\Omega \propto \frac{1}{r} \to 0 \text{ as } r \to \infty, \tag{7.56}$$

where $\mathrm{d}\Omega$ is the element of solid angle. Thus, in the limit, the first term on the right-hand side of (7.55) vanishes and

$$\tfrac{1}{2}\int_v \rho_e V \,\mathrm{d}v = \tfrac{1}{2}\int_v \boldsymbol{E}\cdot\boldsymbol{D}\,\mathrm{d}v. \tag{7.57}$$

The right-hand side is exactly the same as $\iint \boldsymbol{E}\cdot \mathrm{d}\boldsymbol{D}\,\mathrm{d}v$, provided the media are linear.

Problem Show that, in the absence of permanent magnets, the second term on the right-hand side of (7.53) becomes $\sum \frac{1}{2} I_n \phi_n$ where I_n is the current in the nth circuit and ϕ_n is the magnetic flux threading that circuit, again assuming the media are linear, $\boldsymbol{B} \propto \boldsymbol{H}$.

Proof It is simplest to use the vector potential $\boldsymbol{A}$ defined by $\boldsymbol{B} = \operatorname{curl}\boldsymbol{A}$ which formed such a key role in Maxwell's discovery of his equations (see Sections 6.2 and 6.3). Let us analyse the integral

$$\int_v \mathrm{d}v \int_0^{\boldsymbol{B}} \boldsymbol{H}\cdot \mathrm{d}\boldsymbol{B}. \tag{7.58}$$

Taking the differential of the definition of $\boldsymbol{A}$, $\mathrm{d}\boldsymbol{B} = \operatorname{curl}\mathrm{d}\boldsymbol{A}$. We use again the vector identity

$$\nabla\cdot(\boldsymbol{a}\times\boldsymbol{b}) = \boldsymbol{b}\cdot(\nabla\times\boldsymbol{a}) - \boldsymbol{a}\cdot(\nabla\times\boldsymbol{b}).$$

Therefore,

$$\nabla\cdot(\boldsymbol{H}\times \mathrm{d}\boldsymbol{A}) = \mathrm{d}\boldsymbol{A}\cdot(\nabla\times\boldsymbol{H}) - \boldsymbol{H}\cdot(\nabla\times \mathrm{d}\boldsymbol{A}).$$

Substituting for $\boldsymbol{H}\cdot\mathrm{d}\boldsymbol{B} = \boldsymbol{H}\cdot(\nabla\times\mathrm{d}\boldsymbol{A})$ into the integral (7.58) and applying Stokes' theorem to the second integral on the right-hand side,

$$\begin{aligned}\int \mathrm{d}v \int_0^{B} \boldsymbol{H}\cdot \mathrm{d}\boldsymbol{B} &= \int_v \mathrm{d}v \int_0^{\boldsymbol{A}} (\nabla\times\boldsymbol{H})\cdot \mathrm{d}\boldsymbol{A} - \int_v \mathrm{d}v \int_0^{\boldsymbol{A}} \nabla\cdot(\boldsymbol{H}\times \mathrm{d}\boldsymbol{A})\\ &= \int_v \mathrm{d}v \int_0^{\boldsymbol{A}} (\nabla\times\boldsymbol{H})\cdot \mathrm{d}\boldsymbol{A} - \int_S \int_0^{\boldsymbol{A}} (\boldsymbol{H}\times \mathrm{d}\boldsymbol{A})\cdot \mathrm{d}\boldsymbol{S}.\end{aligned}$$

We use the same argument which led to (7.56) to evaluate the surface integral at infinity. Since there are no magnetic poles, $A \propto r^{-2}$ and $H \propto r^{-3}$ are the best we can possibly do. It follows that $\int_S \ldots \mathrm{d}\boldsymbol{S} \to 0$ as $r\to\infty$. Therefore,

$$\int \mathrm{d}v \int_0^{\boldsymbol{B}} \boldsymbol{H}\cdot \mathrm{d}\boldsymbol{B} = \int_v \mathrm{d}v \int_0^{\boldsymbol{A}} (\nabla\times\boldsymbol{H})\cdot \mathrm{d}\boldsymbol{A}. \tag{7.59}$$

Assuming there are no displacement currents present, $\partial\boldsymbol{D}/\partial t = 0$ and hence $\nabla\times\boldsymbol{H} = \boldsymbol{J}$. The relation (7.59) therefore becomes

$$\int_v \mathrm{d}v \int_0^{\boldsymbol{B}} \boldsymbol{H}\cdot \mathrm{d}\boldsymbol{B} = \int_v \mathrm{d}v \int_0^{\boldsymbol{A}} \boldsymbol{J}\cdot \mathrm{d}\boldsymbol{A}.$$

Now consider a closed current loop carrying a current I (Fig. 7.2). Consider a short section of the current tube of length $\mathrm{d}\boldsymbol{l}$ in which the current density is $\boldsymbol{J}$ and the cross-section $\mathrm{d}\boldsymbol{\sigma}$.

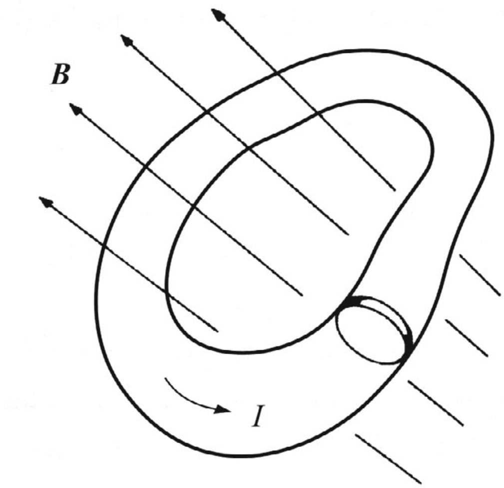


Fig. 7.2 Illustrating the magnetic flux passing through a closed current tube carrying a current *I*.

The advantage of using a closed tube is that the current I is constant, that is, $\boldsymbol{J} \cdot \mathrm{d}\boldsymbol{\sigma} = J\mathrm{d}\sigma = I =$ constant. Therefore, since $\mathrm{d}v = \mathrm{d}\sigma\,\mathrm{d}l$,

$$\int_v \mathrm{d}v \int_0^{\boldsymbol{A}} \boldsymbol{J} \cdot \mathrm{d}\boldsymbol{A} = \int_l \int_0^{\boldsymbol{A}} J\,\mathrm{d}\sigma\,\mathrm{d}\boldsymbol{A} \cdot \mathrm{d}\boldsymbol{l}$$
$$= \int_l \int_0^{\boldsymbol{A}} I\,\mathrm{d}\boldsymbol{A} \cdot \mathrm{d}\boldsymbol{l}.$$

Now, in taking the integral over d$\boldsymbol{A}$ we notice that we are, in fact, working out an energy through the product $I\mathrm{d}A$ and therefore the energy needed to attain the value $\boldsymbol{A}$ is one-half the product of I and $\boldsymbol{A}$, as occurs in electrostatics. We have disguised the dependence on the vector field $\boldsymbol{J}$ by including it in the constant current I. Therefore,

$$\int_v \mathrm{d}v \int_0^{\boldsymbol{A}} \boldsymbol{J} \cdot \mathrm{d}\boldsymbol{A} = \tfrac{1}{2} I \int_l \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{l}$$
$$= \tfrac{1}{2} I \int_S \mathrm{curl}\,\boldsymbol{A} \cdot \mathrm{d}\boldsymbol{S} = \tfrac{1}{2} I \int_S \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{S} = \tfrac{1}{2} I\Phi.$$

Now we fill up the whole of space by a superposition of current loops and flux linkages, so that

$$\int \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{B} = \tfrac{1}{2} \sum_n I_n \Phi_n.$$

7.12 Concluding Remarks

We are in real danger of writing a textbook on vector fields in electromagnetism. It really is a remarkably elegant story. What is striking is the remarkable economy of Maxwell's equations in accounting for all the phenomena of classical electromagnetism. The formalism can be extended to treat much more complex systems involving anisotropic forces, for example, in the propagation of electromagnetic waves in magnetised plasmas, and in anisotropic material media. And it all originated in Maxwell's mechanical analogue for the properties of the vacuum through which electromagnetic phenomena are propagated.

Notes

1 Stratton, J.A. (1941). *Electromagnetic Theory*. London and New York: McGraw Hill.

2 The book *Mathematical Methods for Physics and Engineering*, 3rd edition (2006) by K.F. Riley, M.P. Hobson and S.J. Bence (Cambridge: Cambridge University Press) can be warmly recommended as an excellent source for all the mathematics needed in this book. Chapters 7 and 10 on vector algebra and vector operators are particularly helpful for this chapter.

3 This terminology is advocated by Maxwell in his important paragraph (74) highlighted in Section 6.3 about the polarisation properties of material media.

4 It is simplest to demonstrate this result working backwards. Writing the relation (7.44) in differential form,

$$\mathrm{d}\boldsymbol{A} = \frac{\mu_0 I}{4\pi}\frac{\mathrm{d}\boldsymbol{s}}{r}.$$

Now,

$$\mathrm{d}\boldsymbol{B} = \mathrm{curl}\,(\mathrm{d}\boldsymbol{A}) = \frac{\mu_0 \boldsymbol{I}}{4\pi}\,\mathrm{curl}\left(\frac{\mathrm{d}\boldsymbol{s}}{r}\right).$$

The identity (A5.13) states that

$$\mathrm{curl}\, f\boldsymbol{g} = f\,\mathrm{curl}\,\boldsymbol{g} + (\mathrm{grad}\, f) \times \boldsymbol{g},$$

and so, setting $f = 1/r$ and $\boldsymbol{g} = \mathrm{d}\boldsymbol{s}$,

$$\mathrm{curl}\left(\frac{\mathrm{d}\boldsymbol{s}}{r}\right) = \frac{1}{r}\,\mathrm{curl}\,\mathrm{d}\boldsymbol{s} + \mathrm{grad}\left(\frac{1}{r}\right) \times \mathrm{d}\boldsymbol{s} = -\frac{\boldsymbol{i}_r}{r^2} \times \mathrm{d}\boldsymbol{s},$$

since $\mathrm{d}\boldsymbol{s}$ is a fixed current element. Hence, we find (A5.20),

$$\mathrm{d}\boldsymbol{B} = \frac{\mu_0 \boldsymbol{I}}{4\pi}\frac{\mathrm{d}\boldsymbol{s} \times \boldsymbol{r}}{r^3}.$$

Case Study III

MECHANICS AND DYNAMICS: LINEAR AND NON-LINEAR

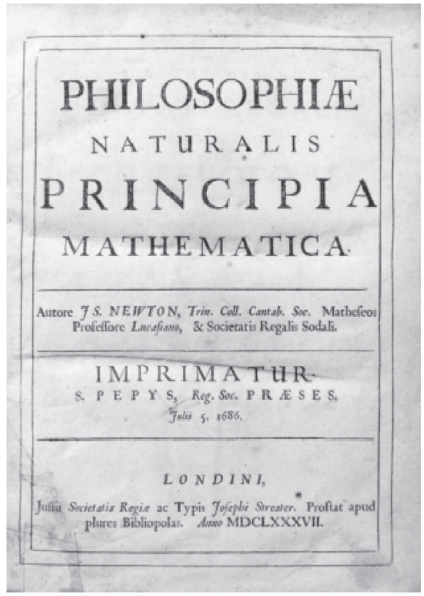
One of the key parts of any course in theoretical physics is the development of a wide range of more and more advanced procedures for treating problems in classical mechanics and dynamics. These are all extensions of the basic principles enunciated by Newton in his *Principia Mathematica* (Fig. III.1), although some of them appear to bear only a distant resemblance to the three laws of motion. As an example of the variety of ways in which the fundamentals of mechanics and dynamics can be expounded, here is a list of different approaches which I found in the textbook *Foundations of Physics*[1] by R.B. Lindsay and H. Margenau, which I read with some trepidation as a first-year undergraduate:

- Newton's Laws of Motion,
- D'Alembert's Principle,
- the Principle of Virtual Displacements, or Virtual Work,
- Gauss's Principle of Least Constraint,
- Hertz's Mechanics,
- Hamilton's Principle and the Principle of Least Action,
- Generalised Coordinates and the Methods of Lagrange,
- the Canonical Equations of Hamilton,
- Transformation Theory of Mechanics and the Hamilton–Jacobi Equations,
- Action-Angle Variables.

This is not the place to go into the details of these different approaches which are more than adequately covered in standard texts such as Goldstein's *Classical Mechanics*.[2] Rather, the emphasis in Chapter 8 is upon features of these approaches which provide insights into different aspects of dynamical systems.

It is important to appreciate that these different approaches are *fully equivalent*. A given problem in mechanics or dynamics can, in principle, be worked out by any of these techniques. Newton's laws of motion, shown in their original formulation in Fig. III.2, are not necessarily the best approach to tackle any particular problem, others often providing

[1] Lindsay, R.B. and Margenau, H. (1957). *Foundations of Physics*. London: Constable & Co. Ltd. (Dover reprint).
[2] Goldstein, H. (1950). *Classical Mechanics*. London: Addison-Wesley.



PHILOSOPHIÆ
NATURALIS
PRINCIPIA
MATHEMATICA.

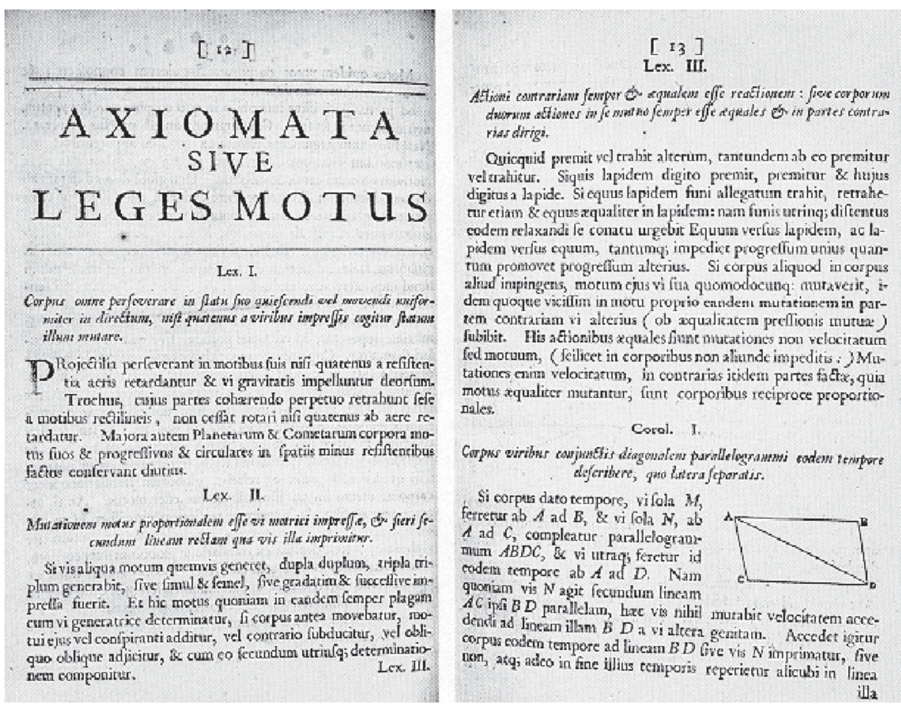
Autore *J S. NEWTON*, *Trin. Coll. Cantab. Soc.* Matheseos Professore *Lucasiano*, & Societatis Regalis Sodali.

IMPRIMATUR.
S. PEPYS, *Reg. Soc.* PRÆSES.
Julii 5. 1686.

LONDINI,

Jussu Societatis Regiæ ac Typis *Josephi Streater*. Prostat apud plures Bibliopolas. *Anno* MDCLXXXVII.

Fig. III.1 The title page of Newton's *Principia Mathematica*, 1687, London.



[12]

AXIOMATA SIVE LEGES MOTUS

Lex. I.

Corpus omne perseverare in statu suo quiescendi vel movendi uniformiter in directum, nisi quatenus a viribus impressis cogitur statum illum mutare.

Projectilia perseverant in motibus suis nisi quatenus a resistentia aeris retardantur & vi gravitatis impelluntur deorsum. Trochus, cujus partes cohærendo perpetuo retrahunt sese a motibus rectilineis, non cessat rotari nisi quatenus ab aere retardatur. Majora autem Planetarum & Cometarum corpora motus suos & progressivos & circulares in spatiis minus resistentibus factos conservant diutius.

Lex. II.

Mutationem motus proportionalem esse vi motrici impressæ, & fieri secundum lineam rectam qua vis illa imprimitur.

Si vis aliqua motum quemvis generet, dupla duplum, tripla triplum generabit, sive simul & semel, sive gradatim & successive impressa fuerit. Et hic motus quoniam in eandem semper plagam cum vi generatrice determinatur, si corpus antea movebatur, motui ejus vel conspiranti additur, vel contrario subducitur, vel oblique oblique adjicitur, & cum eo secundum utriusq; determinationem componitur.

Lex. III.

[13]

Lex. III.

Actioni contrariam semper & æqualem esse reactionem: sive corporum duorum actiones in se mutuo semper esse æquales & in partes contrarias dirigi.

Quicquid premit vel trahit alterum, tantundem ab eo premitur vel trahitur. Siquis lapidem digito premit, premitur & hujus digitus a lapide. Si equus lapidem funi allegatum trahit, retrahetur etiam & equus æqualiter in lapidem: nam funis utrinq; distentus eodem relaxandi se conatu urgebit Equum versus lapidem, ac lapidem versus equum, tantumq; impediet progressum unius quantum promovet progressum alterius. Si corpus aliquod in corpus aliud impingens, motum ejus vi sua quomodocunq; mutaverit, idem quoque vicissim in motu proprio eandem mutationem in partem contrariam vi alterius (ob æqualitatem pressionis mutuæ) subibit. His actionibus æquales fiunt mutationes non velocitatum sed motuum, (scilicet in corporibus non aliunde impeditis:) Mutationes enim velocitatum, in contrarias itidem partes factæ, quia motus æqualiter mutantur, sunt corporibus reciproce proportionales.

Corol. I.

Corpus viribus conjunctis diagonalem parallelogrammi eodem tempore describere, quo latera separatis.

Si corpus dato tempore, vi sola *M*, ferretur ab *A* ad *B*, & vi sola *N*, ab *A* ad *C*, compleatur parallelogrammum *ABDC*, & vi utraq; feretur id eodem tempore ab *A* ad *D*. Nam quoniam vis *N* agit secundum lineam *AC* ipsi *BD* parallelam, hæc vis nihil mutabit velocitatem accedendi ad lineam illam *BD* a vi altera genitam. Accedet igitur corpus eodem tempore ad lineam *BD* sive vis *N* imprimatur, sive non, atq; adeo in fine illius temporis reperietur alicubi in linea illa

Fig. III.2 Newton's laws of motion as they appear in Newton's *Principia Mathematica*.

a much more straightforward route. Furthermore, some of these procedures result in a much deeper appreciation of some of the basic features of dynamical systems. Among these, *conservation laws* and *normal modes of oscillation* are concepts of particular importance. Some of these procedures lead naturally to formalisms which can be taken over into quantum mechanics. The need to be able to develop techniques for solving dynamical problems in particular coordinate systems leads to a discussion of Hamiltonian methods and Hamilton–Jacobi theory. The objective of this discussion is not to provide a detailed account of advanced mechanics and dynamics, but to outline their value in solving technically complex problems.

In parallel with the development of advanced techniques for tackling problems in the dynamics of particle and extended bodies, the laws of Newtonian physics were extended to the study of fluids and gases by pioneers such as Leonhard Euler and Daniel Bernoulli in the eighteenth century and George Stokes, William Thomson (Lord Kelvin) and John William Strutt (Lord Rayleigh) in the nineteenth. They introduced new concepts and procedures which are not only non-trivial and subtle, but also of the greatest practical importance. These are introduced in Chapter 9, the emphasis again being upon the key physical issues.

Chapter 9 introduces the complexities of the non-linearity implicit in many fluid dynamical problems but this is only the tip of the iceberg. In Chapter 10, we go beyond the analytic to aspects of mechanical and dynamical systems which can only be tackled by the use of numerical methods using high-speed computers. Several different approaches to these problems are discussed – dimensional methods, the highly non-linear phenomena involved in chaotic dynamical systems and, the most extreme of all, the study of systems which are so non-linear that only by computer modelling can we obtain insight into the origin of the regularities which appear within the disorder. Many of these areas of study are only in their infancy.

8 Approaches to Mechanics and Dynamics

8.1 Newton's Laws of Motion

In Case Study I, we outlined the history which led to Newton's publication of his laws of motion in the *Principia Mathematica*, published by Samuel Pepys in 1687 (Fig. III.1). The *Principia* consists of three books, preceded by two short sections called *Definitions* and *Axioms, or the Laws of Motion*. The definitions are eight in number and describe concepts such as mass, momentum or impulse, impressed force, inertia and centrifugal force. The three axioms which follow are what we now call *Newton's laws of motion*. There has been considerable debate about the exact meanings of the definitions and axioms, but in the analyses of Books I and III Newton uses these quantities unambiguously with their modern meanings. The three laws of motion appear in essentially identical form to the words used in the standard textbooks.

- *Newton 1*. 'Every body continues in its state of rest, or of uniform motion in a straight line, except in so far as it is compelled by forces to change that state'. In vector notation,

$$\text{If} \quad \boldsymbol{f} = 0, \quad \frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = 0. \tag{8.1}$$

- *Newton 2*. 'Change of motion (that is, momentum) is proportional to the force and takes place in the direction of the straight line in which the force acts.' In fact, Newton means that the rate of change of momentum is proportional to force as demonstrated by his use of the law throughout the *Principia*. In modern notation,

$$\boldsymbol{f} = \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t}, \quad \boldsymbol{p} = m\boldsymbol{v}. \tag{8.2}$$

  where $\boldsymbol{p}$ is the momentum.
- *Newton 3*. 'To every action there is always an equal and contrary reaction; or, the mutual actions of any two bodies are always equal and oppositely directed along the same straight line'.

There are a number of interesting issues raised by these definitions and the three laws of motion. Newton apparently regarded a number of the concepts as self evident, for example, those of 'mass' and 'force'. It is simplest to regard the three laws of motion themselves as providing the definitions of the quantities involved and reflecting our experimental experience.

As an example, we can use Newton's second law to define *mass*. We devise a means by which a force of fixed magnitude can be applied to different bodies. By experiment, we find that the bodies suffer accelerations inversely proportional to the applied force. The masses of the bodies can therefore be defined as quantities proportional to the inverse of their accelerations. Using Newton's second and third laws, a *relative* scale of mass can be defined. Suppose two point masses A and B interact with each other. According to Newton's third law, $\boldsymbol{f}_A = -\boldsymbol{f}_B$. Notice that action and reaction always act on different bodies. If we measure their accelerations under their mutual interaction, they are found to be in a certain ratio a_{AB}/a_{BA} which we can call M_{AB}. If we then compare C with A and B separately, we will measure ratios of accelerations, M_{BC} and M_{AC}. From experiment, we find that $M_{AB} = M_{AC}/M_{BC}$. Therefore, we can define a mass scale by choosing one of them as a standard.

Now all of this may seem rather trivial but it emphasises an important point. Even in something as apparently intuitive as Newton's laws of motion, there are fundamental assumptions and results of experiment and experience upon which the whole structure of Newtonian physics is based. In practice, the equations are a very good approximation to what we observe and measure in the real world. There is, however, no purely logical way of setting up the foundations *ab initio*. This is not a problem so long as we adhere to Dirac's dictum 'that we want the equations which describe nature ... and have to reconcile ourselves to an absence of strict logic.'[1]

Various conservation laws follow from Newton's laws of motion, namely:

- *the law of conservation of momentum*,
- *the law of conservation of angular momentum*,
- *the law of conservation of energy*, once suitable potential energy functions have been defined.

We will derive these later using Euler–Lagrange techniques in Section 8.4.

One important concept is that of the *invariance* of Newton's laws of motion with respect to transformation between frames of reference in uniform relative motion. We discussed Galileo's discovery of inertial frames of reference and the invariance of the laws of physics under Galilean transformation of the coordinates in Section 3.6. According to the Newtonian picture, the coordinates in the frames of reference S and S' shown in Fig. 3.9 with uniform relative velocity V are related by

$$\begin{cases} x' = x - Vt, \\ y' = y, \\ z' = z, \\ t' = t. \end{cases} \tag{8.3}$$

Taking the second derivative of the first of the relations (8.3) with respect to time, $\ddot{x}' = \ddot{x}$, that is, the acceleration and consequently the forces are the same in any inertial frame of reference. In relativistic language, the acceleration is an *invariant* between inertial frames of reference. Time is absolute, in the sense that t keeps track of time in all inertial frames of reference. The expressions (8.3) are often referred to as *Galilean transformations*. It is intriguing that one of his motivations for advancing this argument was to demonstrate that there is nothing inconsistent with our view of the Universe if the Earth happens to be in uniform motion rather than being at rest at the centre of the Universe, in other words, it was part of his defence of the Copernican picture of the world.

8.2 Principles of 'Least Action'

Some of the most powerful approaches to mechanics and dynamics involve finding that function which minimises the value of some quantity subject to well-defined boundary conditions. In the formal development of mechanics, the procedures are stated axiomatically. Let us take as an example what Feynman refers to as the *principle of minimum action*.[2] Consider the case of the dynamics of a particle in a conservative field of force, that is, one which is derived as the gradient of a scalar potential, $\boldsymbol{F} = -\text{grad}\, V$ as, for example, in the cases of gravity and electrostatics.

We introduce a set of axioms which enables us to work out the path of the particle subject to this force field. First, we define the quantity

$$\mathcal{L} = T - V = \tfrac{1}{2}mv^2 - V, \tag{8.4}$$

the difference between the kinetic energy T and the potential energy V of the particle in the field. To derive the trajectory of the particle between two fixed endpoints in the field in a fixed time interval t_1 to t_2, we find that path which minimises the function

$$S = \int_{t_1}^{t_2} \left(\tfrac{1}{2}mv^2 - V\right) \,\mathrm{d}t = \int_{t_1}^{t_2} \left[\tfrac{1}{2}m\left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\right)^2 - V\right] \mathrm{d}t. \tag{8.5}$$

These statements are to be regarded as equivalent to Newton's laws of motion or, rather, what Newton accurately called his 'axioms'. $\mathcal{L}$ is called the *Lagrangian*.

These axioms are consistent with Newton's first law of motion. If there are no forces present, V = constant and hence we minimise $S = \int_{t_1}^{t_2} v^2\,\mathrm{d}t$. The minimum value of S must correspond to a constant velocity v between t_1 and t_2. If the particle travelled between the end points by accelerating and decelerating in such a way that t_1 and t_2 are the same, the integral must be greater than that for constant velocity between t_1 and t_2 because v^2 appears in the integral and a basic rule of analysis tells us that $\langle v^2 \rangle \geq \langle \overline{v} \rangle^2$. Thus, in the absence of forces, $\boldsymbol{v}$ = constant, Newton's first law of motion.

To proceed further, we need the techniques of the *calculus of variations*. Suppose $f(x)$ is a function of a single variable x. The minimum value of the function corresponds to the value of x at which $\mathrm{d}f(x)/\mathrm{d}x$ is zero (Fig. 8.1). Approximating the variation of the function about the minimum at $x = 0$ by a power series,

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots, \tag{8.6}$$

the function can only have $\mathrm{d}f/\mathrm{d}x = 0$ at $x = 0$ if $a_1 = 0$. At this point, the function is approximately a parabola since the first non-zero coefficient in the expansion of $f(x)$ about the minimum is the term in x^2 – for small displacements x from the minimum, the change in the function $f(x)$ is second order in x.

The same principle is used to find the path of the particle when S is minimised. If the true path of the particle is $\boldsymbol{x}_0(t)$, another path between t_1 and t_2 is given by

$$\boldsymbol{x}(t) = \boldsymbol{x}_0(t) + \boldsymbol{\eta}(t), \tag{8.7}$$

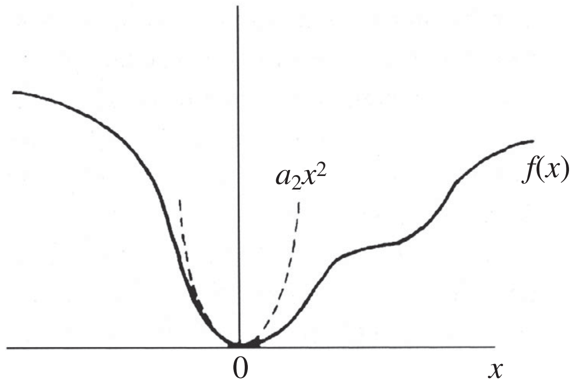


Fig. 8.1 Illustrating how a function $f(x)$ can be approximated as a_2x^2 close to a minimum.

where $\boldsymbol{\eta}(t)$ describes the deviation of $\boldsymbol{x}(t)$ from the minimum path $\boldsymbol{x}_0(t)$. Just as we can define the minimum of a function as the point at which there is no first-order dependence of the function upon $\boldsymbol{x}$, so we can define the minimum of the function $\boldsymbol{x}(t)$ as that function for which the dependence on $\eta(t)$ is at least *second order*, that is, there should be no linear term in $\eta(t)$.

Substituting (8.7) into (8.5),

$$S = \int_{t_1}^{t_2} \left[\frac{m}{2} \left(\frac{\mathrm{d}\boldsymbol{x}_0}{\mathrm{d}t} + \frac{\mathrm{d}\boldsymbol{\eta}}{\mathrm{d}t} \right)^2 - V(\boldsymbol{x}_0 + \boldsymbol{\eta}) \right] \mathrm{d}t$$

$$= \int_{t_1}^{t_2} \left\{ \frac{m}{2} \left[\left(\frac{\mathrm{d}\boldsymbol{x}_0}{\mathrm{d}t} \right)^2 + 2\frac{\mathrm{d}\boldsymbol{x}_0}{\mathrm{d}t} \cdot \frac{\mathrm{d}\boldsymbol{\eta}}{\mathrm{d}t} + \left(\frac{\mathrm{d}\boldsymbol{\eta}}{\mathrm{d}t} \right)^2 \right] - V(\boldsymbol{x}_0 + \boldsymbol{\eta}) \right\} \mathrm{d}t. \tag{8.8}$$

We now eliminate first-order quantities in $\mathrm{d}\boldsymbol{\eta}$. First of all, we drop the term $(\mathrm{d}\boldsymbol{\eta}/\mathrm{d}t)^2$ as second order in small quantities and then expand $V(\boldsymbol{x}_0 + \boldsymbol{\eta})$ to first order in $\boldsymbol{\eta}$ by a Taylor expansion,

$$V(\boldsymbol{x}_0 + \boldsymbol{\eta}) = V(\boldsymbol{x}_0) + \nabla V \cdot \boldsymbol{\eta}. \tag{8.9}$$

Substituting (8.9) into (8.8) and preserving only quantities to first order in $\boldsymbol{\eta}$, we find

$$S = \int_{t_1}^{t_2} \left[\frac{m}{2} \left(\frac{\mathrm{d}\boldsymbol{x}_0}{\mathrm{d}t} \right)^2 - V(\boldsymbol{x}_0) + m\frac{\mathrm{d}\boldsymbol{x}_0}{\mathrm{d}t} \cdot \frac{\mathrm{d}\boldsymbol{\eta}}{\mathrm{d}t} - \nabla V \cdot \boldsymbol{\eta} \right] \mathrm{d}t. \tag{8.10}$$

The first two terms inside the integral are the minimum path and so are a constant. We therefore need to ensure that the last two terms have no dependence upon $\boldsymbol{\eta}$, the condition for a minimum. We therefore need consider only the last two terms,

$$S = \int_{t_1}^{t_2} \left(m\frac{\mathrm{d}\boldsymbol{x}_0}{\mathrm{d}t} \cdot \frac{\mathrm{d}\boldsymbol{\eta}}{\mathrm{d}t} - \boldsymbol{\eta} \cdot \nabla V \right) \mathrm{d}t. \tag{8.11}$$

We integrate the first term by parts so that $\boldsymbol{\eta}$ alone appears in the integrand, that is,

$$S = m\left[\boldsymbol{\eta} \cdot \frac{\mathrm{d}\boldsymbol{x}_0}{\mathrm{d}t}\right]_{t_1}^{t_2} - \int_{t_1}^{t_2} \left[\frac{\mathrm{d}}{\mathrm{d}t}\left(m\frac{\mathrm{d}\boldsymbol{x}_0}{\mathrm{d}t}\right) \cdot \boldsymbol{\eta} + \boldsymbol{\eta} \cdot \nabla V\right] \mathrm{d}t. \tag{8.12}$$

The function $\boldsymbol{\eta}$ must be zero at t_1 and t_2 because that is where the path begins and ends and the endpoints are fixed. Therefore, the first term on the right-hand side of (8.12) is zero and we can write

$$S = -\int_{t_1}^{t_2} \left\{\boldsymbol{\eta} \cdot \left[\frac{\mathrm{d}}{\mathrm{d}t}\left(m\frac{\mathrm{d}\boldsymbol{x}_0}{\mathrm{d}t}\right) + \nabla V\right]\right\} \mathrm{d}t = 0. \tag{8.13}$$

This must be true for arbitrary perturbations about $\boldsymbol{x}_0(t)$ and hence the term in square brackets must be zero, that is,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(m\frac{\mathrm{d}\boldsymbol{x}_0}{\mathrm{d}t}\right) = -\nabla V. \tag{8.14}$$

We have recovered Newton's second law of motion since $\boldsymbol{f} = -\nabla V$, that is,

$$\boldsymbol{f} = \frac{\mathrm{d}}{\mathrm{d}t}\left(m\frac{\mathrm{d}\boldsymbol{x}_0}{\mathrm{d}t}\right) = \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t}. \tag{8.15}$$

Thus, our alternative formulation of the laws of motion in terms of an action principle is exactly equivalent to Newton's statement of the laws of motion.

These procedures can be generalised to take account of conservative *and* non-conservative forces such as those which depend upon velocity, for example, friction and the force on a charged particle in a magnetic field.[3] The key point is that we have a prescription which involves writing down the kinetic and potential energies (T and V respectively) of the system, then forming the Lagrangian $\mathcal{L}$ and finding the minimum value of the function S. The advantage of this procedure is that it is often a straightforward matter to write down these energies in some suitable set of coordinates. We therefore need rules which tell us how to find the minimum value of S in any set of coordinates. These are the *Euler–Lagrange equations*.

An important point about the principle of least action should be noted. In general, we do not have a definite prescription for finding the Lagrangian. To quote Feynman,

> The question of what the action (S) should be for a particular case must be determined by some kind of trial and error. It is just the same problem as determining the laws of motion in the first place. You just have to fiddle around with the equations that you know and see if you can get them into the form of a principle of least action.[4]

We will give an example of this once we have developed the Euler–Lagrange equations.

8.3 The Euler–Lagrange Equations

Consider the case of a system of N particles interacting through a scalar potential function V. The positions of the N particles are given by the vectors $[\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3 \ldots \boldsymbol{r}_N]$

in Cartesian coordinates. Since three numbers are needed to describe the position of each particle, for example, x_i, y_i, z_i, the vector describing the positions of all the particles has $3N$ coordinates. For greater generality, we wish to transform these coordinates into a different set of coordinates which we write as $[q_1, q_2, q_3, \dots q_{3N}]$. The set of relations between these coordinates can be written

$$q_i = q_i(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3, \dots \boldsymbol{r}_N), \quad \text{and hence} \quad r_i = r_i(q_1, q_2, q_3, \dots q_{3N}). \tag{8.16}$$

This is no more than a change of variables.

The aim of the procedure is to write down the equations for the dynamics of the particles, that is, an equation for each independent coordinate, in terms of the coordinates q_i rather than $\boldsymbol{r}_i$. We are guided by the analysis of the previous subsection on action principles to form the quantities T and V, the kinetic and potential energies respectively, in terms of the new set of coordinates and then to find the stationary value of S, that is

$$\mathcal{L} = T - V, \qquad \delta S = \delta \int_{t_1}^{t_2} (T - V)\,\mathrm{d}t = 0, \tag{8.17}$$

where δ means 'take the variation about a particular value of the coordinates' as discussed in the previous section. This formulation is called *Hamilton's principle* where $\mathcal{L}$ is the *Lagrangian* of the system. Notice that Hamilton's principle makes no reference to the coordinate system to be used in the calculation.

The kinetic energy of the system is

$$T = \sum_i m_i \dot{\boldsymbol{r}}_i^2. \tag{8.18}$$

In terms of our new coordinate system, we can write without loss of generality

$$r_i = r_i(q_1, q_2, \dots q_{3N}, t) \quad \text{and} \quad \dot{r}_i = \dot{r}_i(\dot{q}_1, \dot{q}_2, \dot{q}_3, \dots \dot{q}_{3N}, q_1, q_2, \dots q_{3N}, t). \tag{8.19}$$

Notice that we have now included explicitly the time dependence of r_i and $\dot{r}_i$. Therefore, we can write the kinetic energy as a function of the coordinates $\dot{q}_i, q_i$ and t, that is, $T(\dot{q}_i, q_i, t)$, where we understand that all the values of i from 1 to $3N$ are included. Similarly, we can write the expression for the potential energy entirely in terms of the coordinates q_i and t, that is, $V(q_i, t)$. We therefore need to find the stationary values of

$$S = \int_{t_1}^{t_2} [T(\dot{q}_i, q_i, t) - V(q_i, t)]\,\mathrm{d}t = \int_{t_1}^{t_2} \mathcal{L}(\dot{q}_i, q_i, t)\,\mathrm{d}t. \tag{8.20}$$

We repeat the analysis of Section 8.2 in which we found the condition for S to be independent of first-order perturbations about the minimum path. As before, we let $q_0(t)$ be the minimum solution and write the expression for another function $q(t)$ in the form

$$q(t) = q_0(t) + \eta(t). \tag{8.21}$$

Now we insert the trial solution (8.21) into (8.20),

$$S = \int_{t_1}^{t_2} \mathcal{L}[\dot{q}_0(t) + \dot{\eta}(t), q_0(t) + \eta(t), t]\,\mathrm{d}t.$$

Performing a Taylor expansion to first order in $\dot{\eta}(t)$ and $\eta(t)$,

$$S = \int_{t_1}^{t_2} \mathcal{L}[\dot{q}_0(t), q_0(t), t]\,\mathrm{d}t + \int_{t_1}^{t_2} \left[\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\dot{\eta}(t) + \frac{\partial \mathcal{L}}{\partial q_i}\eta(t)\right] \mathrm{d}t. \tag{8.22}$$

Setting the first integral equal to S_0 and integrating the term in $\dot{\eta}(t)$ by parts,

$$S = S_0 + \left[\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\eta(t)\right]_{t_1}^{t_2} - \int_{t_1}^{t_2} \left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right)\eta(t) - \frac{\partial \mathcal{L}}{\partial q_i}\eta(t)\right] \mathrm{d}t. \tag{8.23}$$

Because $\eta(t)$ must always be zero at end points, the first term in square brackets disappears and the result can be written

$$S = S_0 - \int_{t_1}^{t_2} \eta(t) \left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) - \frac{\partial \mathcal{L}}{\partial q_i}\right] \mathrm{d}t.$$

We require this integral to be zero for all first-order perturbations about the minimum solution. Therefore, the condition is

$$\frac{\partial \mathcal{L}}{\partial q_i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) = 0. \tag{8.24}$$

The equations (8.24) represent $3N$ second-order differential equations for the time evolution of the $3N$ coordinates and are known as the *Euler–Lagrange equations*. They are no more than Newton's laws of motion written in the q_i coordinate system which can be chosen to be the most convenient for the particular problem at hand.

Example An interesting example is that of the motion of a relativistic particle moving in an electromagnetic field. The Lagrangian is

$$\mathcal{L} = -m_0c^2/\gamma - q\Phi_\mathrm{e} + q\boldsymbol{v}\cdot\boldsymbol{A}, \tag{8.25}$$

where the Lorentz factor $\gamma = (1 - v^2/c^2)^{-1/2}$, Φ_e is the electrostatic potential and $\boldsymbol{A}$ the vector potential. Although $q\Phi_\mathrm{e}$ and $q\boldsymbol{v}\cdot\boldsymbol{A}$ have the familiar forms of potential energy, the term $-m_0c^2/\gamma$ is not the kinetic energy in relativity. The term does, however, reduce correctly to the non-relativistic form for the kinetic energy, that is, $-m_0c^2/\gamma \to \frac{1}{2}m_0v^2 - m_0c^2$ as $v \to 0$; the constant term $-m_0c^2$ does not matter since we are only interested in minimising with respect to variables in the Lagrangian.

Consider first the case of a particle in an electric field which is derivable from a scalar potential so that the Lagrangian becomes

$$\mathcal{L} = -\frac{m_0c^2}{\gamma} - q\Phi_\mathrm{e}. \tag{8.26}$$

Considering first the x-component of the Euler–Lagrange equation (8.24),

$$\frac{\partial \mathcal{L}}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial v_x}\right) = 0, \tag{8.27}$$

and hence

$$-q\frac{\partial \Phi_\mathrm{e}}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\left[-m_0 c^2 \frac{\partial}{\partial v_x}\left(1 - \frac{v_x^2 + v_y^2 + v_z^2}{c^2}\right)^{1/2}\right] = 0,$$

$$-q\frac{\partial \Phi_\mathrm{e}}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\gamma m_0 v_x\right) = 0.$$

Similar relations are found for the y and z components and so we can add them together vectorially to write

$$q\,\mathrm{grad}\,\Phi_\mathrm{e} + \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = 0 \quad \text{or} \quad \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = -q\,\mathrm{grad}\,\Phi_\mathrm{e} = q\boldsymbol{E}, \tag{8.28}$$

where $\boldsymbol{p} = \gamma m_0 \boldsymbol{v}$ is the relativistic three-momentum (see Section 18.4.1). We have recovered the expression for the acceleration of a relativistic particle in an electric field $\boldsymbol{E}$.

Since we have already dealt with the terms $-m_0c^2/\gamma - q\Phi_\mathrm{e}$, let us consider only the Lagrangian

$$\mathcal{L} = q\boldsymbol{v}\cdot\boldsymbol{A} = q(v_x A_x + v_y A_y + v_z A_z). \tag{8.29}$$

Again, considering only the x-component of the Euler–Lagrange equation,

$$\frac{\partial \mathcal{L}}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial v_x}\right) = 0,$$

$$q\left(v_x\frac{\partial A_x}{\partial x} + v_y\frac{\partial A_y}{\partial x} + v_z\frac{\partial A_z}{\partial x}\right) - \frac{\mathrm{d}}{\mathrm{d}t}(qA_x) = 0. \tag{8.30}$$

When we add together vectorially the components in the x, y and z directions, the last term in $\mathrm{d}/\mathrm{d}t$ sums to

$$-q\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}. \tag{8.31}$$

Considering only the x-component, the term in large round brackets in (8.30) can be reorganised as follows:

$$q\left(v_x\frac{\partial A_x}{\partial x} + v_y\frac{\partial A_y}{\partial x} + v_z\frac{\partial A_z}{\partial x}\right) = q\left[\left(v_x\frac{\partial A_x}{\partial x} + v_y\frac{\partial A_x}{\partial y} + v_z\frac{\partial A_x}{\partial z}\right)\right.$$
$$\left. + v_y\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) + v_z\left(\frac{\partial A_z}{\partial x} - \frac{\partial A_x}{\partial z}\right)\right]. \tag{8.32}$$

The right-hand side of (8.32) is the x-component of the vector

$$q(\boldsymbol{v}\cdot\nabla)\boldsymbol{A} + q[\boldsymbol{v}\times(\nabla\times\boldsymbol{A})], \tag{8.33}$$

where the operator $(\boldsymbol{v}\cdot\nabla)$ is described in Sections 9.1 and 9.2. Therefore, the result of applying the Euler–Lagrange formalism to the term $q\boldsymbol{v}\cdot\boldsymbol{A}$ is

$$-q\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t} + q(\boldsymbol{v}\cdot\nabla)\boldsymbol{A} + q[\boldsymbol{v}\times(\nabla\times\boldsymbol{A})]. \tag{8.34}$$

But, as shown in Section 9.1, the total and partial derivatives are related by

$$\frac{\mathrm{d}}{\mathrm{d}t} = \frac{\partial}{\partial t} + (\boldsymbol{v}\cdot\nabla), \tag{8.35}$$

and hence (8.34) becomes

$$-q\frac{\partial \boldsymbol{A}}{\partial t} + q[\boldsymbol{v} \times (\nabla \times \boldsymbol{A})]. \tag{8.36}$$

This is the result we have been seeking. Following the reasoning which led to (5.30) in Section 5.3.1, the *induced* electric field is given by

$$\boldsymbol{E}_{\text{induced}} = -\frac{\partial \boldsymbol{A}}{\partial t}, \tag{8.37}$$

and, by definition, $\boldsymbol{B} = \text{curl}\,\boldsymbol{A} = \nabla \times \boldsymbol{A}$. Therefore, (8.36) reduces to

$$q\boldsymbol{E}_{\text{induced}} + q(\boldsymbol{v} \times \boldsymbol{B}). \tag{8.38}$$

We can now reassemble the complete Lagrangian from (8.28) and (8.38) and write the equation of motion of the particle as

$$\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = q\boldsymbol{E} + q\boldsymbol{E}_{\text{induced}} + q(\boldsymbol{v} \times \boldsymbol{B}) = q\boldsymbol{E}_{\text{tot}} + q(\boldsymbol{v} \times \boldsymbol{B}), \tag{8.39}$$

where $\boldsymbol{p} = \gamma m_0 \boldsymbol{v}$ and $\boldsymbol{E}_{\text{tot}}$ includes both electrostatic and induced electric fields. Equation (8.39) is the complete expression for the equation of motion of a relativistic charged particle moving in combined electric and magnetic fields.

8.4 Lagrangians, Conservation Laws and Symmetry

In Newtonian mechanics, conservation laws are derived from the first integral of the equations of motion. In the same way, we can derive a set of conservation laws from the first integral of the Euler–Lagrange equations. Approaching the problem from this perspective illuminates the close relations between symmetry and conservation laws, a great deal depending upon the form of the Lagrangian itself.

8.4.1 Lagrangian Not a Function of q_i

In this case $\partial\mathcal{L}/\partial q_i = 0$ and hence the Euler–Lagrange equation becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial\dot{q}_i}\right) = 0, \quad \frac{\partial\mathcal{L}}{\partial\dot{q}_i} = \text{constant}.$$

An example of such a Lagrangian is the motion of a particle in the absence of a field of force. Then $\mathcal{L} = \frac{1}{2}m\dot{\boldsymbol{r}}^2$ and

$$\frac{\partial\mathcal{L}}{\partial\dot{q}_i} = \frac{\partial\mathcal{L}}{\partial\dot{\boldsymbol{r}}} = m\dot{\boldsymbol{r}} = \text{constant}, \tag{8.40}$$

that is, Newton's first law of motion.

This calculation suggests how we can define a generalised momentum p_i. For an arbitrary coordinate system, we can define a *conjugate momentum* as

$$p_i \equiv \frac{\partial \mathcal{L}}{\partial \dot{q}_i}. \tag{8.41}$$

This conjugate momentum is not necessarily anything like a normal momentum, but has the property that, if $\mathcal{L}$ does not depend upon q_i, it is a constant of the motion.

8.4.2 Lagrangian Independent of Time

This is clearly related to energy conservation. According to the chain rule,

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}t} = \frac{\partial \mathcal{L}}{\partial t} + \sum_i \dot{q}_i \frac{\partial \mathcal{L}}{\partial q_i}. \tag{8.42}$$

From the Euler–Lagrange equation,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) = \frac{\partial \mathcal{L}}{\partial q_i}.$$

Therefore, substituting for $\partial \mathcal{L}/\partial q_i$ in (8.42),

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}t} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_i \dot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) = \frac{\partial \mathcal{L}}{\partial t},$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathcal{L} - \sum_i \dot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) = \frac{\partial \mathcal{L}}{\partial t}. \tag{8.43}$$

But the Lagrangian does not have any explicit dependence upon time and hence $\partial \mathcal{L}/\partial t = 0$. Therefore

$$\sum_i \dot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \mathcal{L} = \text{constant}. \tag{8.44}$$

This expression is the *law of conservation of energy* in Newtonian mechanics. The quantity which is conserved is known as the *Hamiltonian* H, where

$$H = \sum_i \dot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \mathcal{L}. \tag{8.45}$$

We can write this in terms of the conjugate momentum $p_i = \partial \mathcal{L}/\partial \dot{q}_i$ as

$$H = \sum_i p_i \dot{q}_i - \mathcal{L}. \tag{8.46}$$

In the case of Cartesian coordinates, the Hamiltonian becomes

$$H = \sum_i (m_i \dot{\boldsymbol{r}}_i) \cdot \dot{\boldsymbol{r}}_i - \mathcal{L} = 2T - (T - V) = T + V, \tag{8.47}$$

which shows explicitly the relation to energy conservation. Formally, the conservation law arises from the homogeneity of the Lagrangian with respect to time.

8.4.3 Lagrangian Independent of the Absolute Position of the Particles

By this statement, we mean that $\mathcal{L}$ depends only on $\boldsymbol{r}_1 - \boldsymbol{r}_2$ and not on the absolute value of $\boldsymbol{r}_1$. Suppose we change all the q_is by a same small amount ϵ. Then, the requirement that $\mathcal{L}$ remains unchanged on shifting all the coordinates of all the particles by ϵ becomes

$$\mathcal{L} + \sum_i \frac{\partial \mathcal{L}}{\partial q_i}\delta q_i = \mathcal{L} + \sum_i \epsilon \frac{\partial \mathcal{L}}{\partial q_i} = \mathcal{L}.$$

Invariance requires that

$$\sum_i \frac{\partial \mathcal{L}}{\partial q_i} = 0.$$

But, from the Euler–Lagrange equation,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) = \sum_i \frac{\partial \mathcal{L}}{\partial q_i} = 0, \quad \frac{\mathrm{d}}{\mathrm{d}t}\sum_i p_i = 0. \tag{8.48}$$

This is the law of *conservation of linear momentum* derived from the requirement that the Lagrangian be homogeneous with respect to spatial translations.

8.4.4 Lagrangian Independent of the Orientation of the System in Space

By this statement we mean that $\mathcal{L}$ is invariant under rotations. If the system is rotated through a small angle $\delta\boldsymbol{\theta}$, as in Fig. 8.2, the position and velocity vectors, $\boldsymbol{r}_i$ and $\boldsymbol{v}_i$ respectively, change by

$$\delta \boldsymbol{r}_i = \mathrm{d}\boldsymbol{\theta} \times \boldsymbol{r}_i, \quad \delta \boldsymbol{v}_i = \mathrm{d}\boldsymbol{\theta} \times \boldsymbol{v}_i.$$

We require that $\delta\mathcal{L} = 0$ under this rotation and hence

$$\delta\mathcal{L} = \sum_i \left(\frac{\partial \mathcal{L}}{\partial \boldsymbol{r}_i}\cdot \delta\boldsymbol{r}_i + \frac{\partial \mathcal{L}}{\partial \boldsymbol{v}_i}\cdot \delta\boldsymbol{v}_i\right) = 0.$$

Using the Euler–Lagrange relations,

$$\sum_i \left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{r}}_i}\right)\cdot(\mathrm{d}\boldsymbol{\theta}\times\boldsymbol{r}_i) + \frac{\partial \mathcal{L}}{\partial \boldsymbol{v}_i}\cdot(\mathrm{d}\boldsymbol{\theta}\times\boldsymbol{v}_i)\right] = 0. \tag{8.49}$$

Reordering the vector products, (8.49) becomes

$$\begin{aligned}
\sum_i \left\{\mathrm{d}\boldsymbol{\theta}\cdot\left[\boldsymbol{r}_i \times \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{r}}_i}\right)\right] + \mathrm{d}\boldsymbol{\theta}\cdot\left(\boldsymbol{v}_i \times \frac{\partial \mathcal{L}}{\partial \boldsymbol{v}_i}\right)\right\} &= 0,\\
\sum_i \mathrm{d}\boldsymbol{\theta}\cdot\left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{r}_i \times \frac{\partial \mathcal{L}}{\partial \boldsymbol{v}_i}\right)\right] &= 0,\\
\mathrm{d}\boldsymbol{\theta}\cdot\sum_i \left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{r}_i \times \frac{\partial \mathcal{L}}{\partial \boldsymbol{v}_i}\right)\right] &= 0.
\end{aligned} \tag{8.50}$$

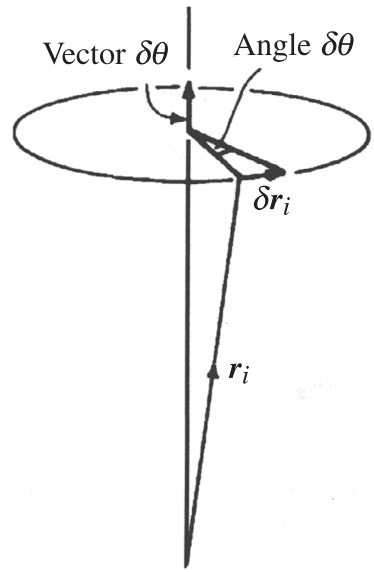


Fig. 8.2 Illustrating the rotation of the system of coordinates through a small angle d$\boldsymbol{\theta}$.

Therefore, if the Lagrangian is independent of orientation,

$$\sum_i \left(\boldsymbol{r}_i \times \frac{\partial \mathcal{L}}{\partial \boldsymbol{v}_i} \right) = \text{constant}, \quad \text{that is,} \quad \sum_i (\boldsymbol{r}_i \times \boldsymbol{p}_i) = \text{constant}. \tag{8.51}$$

This is the *law of conservation of angular momentum* and results from the requirement that the Lagrangian be invariant under rotations.

8.5 Lagrangians, Small Oscillations and Normal Modes

Let us give a simple example of the power of the Euler–Lagrange equation in analysing the dynamical behaviour of systems which undergo *small oscillations*. This leads naturally to the key concept of the *normal modes* of oscillation of the system. In analysing the particular example which follows, we reveal more general features of the behaviour of systems undergoing small oscillations about the equilibrium state.

Example A typical problem is that of a hollow cylinder, for example, a pipe with thin walls, which is suspended by strings of equal length at either end. The strings are attached to the circumference of the cylinder as shown in Fig. 8.3(a), so that it can wobble and swing at the same time. We consider first the swinging motion in which the strings and the flat ends of the cylinder remain in the same planes. An end-on view of the displaced cylinder is

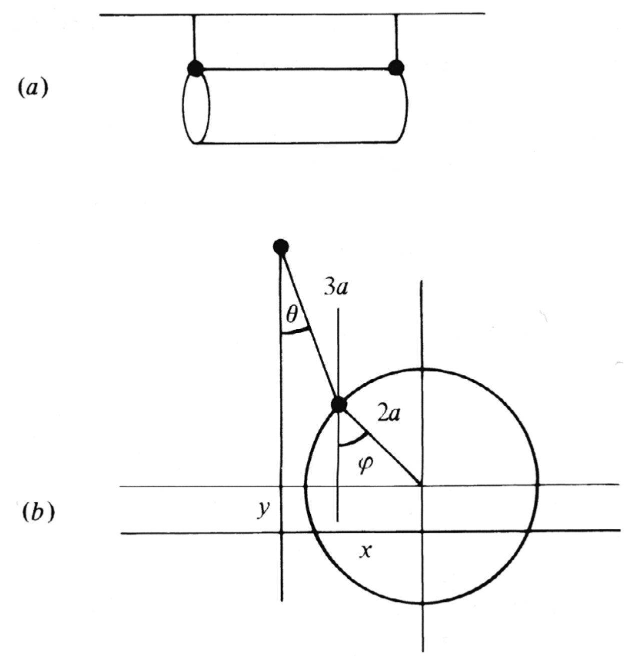


Fig. 8.3 Illustrating the swinging motion of a hollow cylinder suspended by strings at either end and the coordinate system used to analyse its motion.

shown in Fig. 8.3(*b*). For the sake of definiteness, we take the length of the string to be $3a$ and the radius of the cylinder $2a$.

First of all, we choose a suitable set of coordinates which define completely the position of the cylinder when it undergoes a small displacement from the equilibrium position, $\theta = 0, \phi = 0$. θ and ϕ fulfil this role and are the natural coordinates to use for pendulum-like motion. We now write down the Lagrangian for the system in terms of the coordinates $\dot{\theta}, \dot{\phi}, \theta, \phi$.

The *kinetic energy T* consists of two parts, one associated with the *translational motion* of the cylinder and the other with its *rotation*. In the analysis, we consider only small displacements. Therefore, the horizontal displacement of the centre of mass of the cylinder from the vertical position is

$$x = 3a \sin\theta + 2a \sin\phi = 3a\theta + 2a\phi, \tag{8.52}$$

and consequently its translational motion is $\dot{x} = 3a\dot{\theta} + 2a\dot{\phi}$. The linear kinetic energy is therefore $\frac{1}{2}ma^2(3\dot{\theta} + 2\dot{\phi})^2$, where m is the mass of the cylinder. The rotational motion of the cylinder is $\frac{1}{2}I\dot{\phi}^2$, where I is the moment of inertia of the cylinder about its horizontal axis. In this case, $I = 4a^2m$. Thus, the total kinetic energy of the cylinder is

$$T = \tfrac{1}{2}ma^2(3\dot{\theta} + 2\dot{\phi})^2 + 2a^2m\dot{\phi}^2 = \tfrac{1}{2}ma^2(9\dot{\theta}^2 + 12\dot{\theta}\dot{\phi} + 8\dot{\phi}^2). \tag{8.53}$$

The *potential energy* is entirely associated with the vertical displacement of the centre of mass of the cylinder above its equilibrium position:

$$y = 3a(1 - \cos\theta) + 2a(1 - \cos\phi) = \frac{3a\theta^2}{2} + a\phi^2, \tag{8.54}$$

for small values of θ and ϕ. Consequently, the potential energy is

$$V = \frac{mg}{2}(3a\theta^2 + 2a\phi^2). \tag{8.55}$$

The Lagrangian is therefore

$$\mathcal{L} = T - V = \tfrac{1}{2}\, ma^2(9\dot{\theta}^2 + 12\dot{\theta}\dot{\phi} + 8\dot{\phi}^2) - \frac{mg}{2}(3a\theta^2 + 2a\phi^2). \tag{8.56}$$

Notice that we have derived a Lagrangian which is quadratic in $\dot{\theta}, \dot{\phi}, \theta$ and ϕ. The reason for this is that the kinetic energy is a quadratic function of velocity. Also, the potential energy is evaluated relative to the equilibrium position, $\theta = \phi = 0$, and consequently the lowest order expansion about the minimum is second order in the displacement (see Section 8.2).

We now use the Euler–Lagrange equation

$$\frac{\partial \mathcal{L}}{\partial q_i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) = 0. \tag{8.57}$$

Let us take the $\theta, \dot{\theta}$ pair of coordinates first. Substituting (8.56) into (8.57) and taking $\dot{q}_i = \dot{\theta}$ and $q_i = \theta$, we find

$$\tfrac{1}{2} ma^2 \frac{\mathrm{d}}{\mathrm{d}t}(18\dot{\theta} + 12\dot{\phi}) = -\tfrac{1}{2} mga\, 6\theta. \tag{8.58}$$

Similarly, for the coordinate pair $\dot{q}_i = \dot{\phi}, q_i = \phi$, we find

$$\tfrac{1}{2} ma^2 \frac{\mathrm{d}}{\mathrm{d}t}(12\dot{\theta} + 16\dot{\phi}) = -\tfrac{1}{2} mga\, 4\phi. \tag{8.59}$$

The result is two differential equations:

$$\begin{cases} 9\ddot{\theta} + 6\ddot{\phi} = -\dfrac{3g}{a}\theta, \\[2mm] 6\ddot{\theta} + 8\ddot{\phi} = -\dfrac{2g}{a}\phi. \end{cases} \tag{8.60}$$

In a normal mode of oscillation, all components oscillate at the same frequency, that is, we seek oscillatory solutions of the form

$$\begin{cases} \ddot{\theta} = -\omega^2\theta, \\ \ddot{\phi} = -\omega^2\phi. \end{cases} \tag{8.61}$$

Now insert the trial solutions (8.61) into (8.60). Setting $\lambda = a\omega^2/g$,

$$\begin{cases} (9\lambda - 3)\theta + 6\lambda\phi = 0, \\ 6\lambda\theta + (8\lambda - 2)\phi = 0. \end{cases} \tag{8.62}$$

The condition that (8.62) be satisfied for all θ and ϕ is that the determinant of the coefficients should be zero, that is,

$$\begin{vmatrix} (9\lambda - 3) & 6\lambda \\ 6\lambda & (8\lambda - 2) \end{vmatrix} = 0. \tag{8.63}$$

Multiplying out this determinant,

$$6\lambda^2 - 7\lambda + 1 = 0, \tag{8.64}$$

which has solutions $\lambda = 1$ and $\lambda = \frac{1}{6}$. Therefore, the angular frequencies of oscillation are $\omega^2 = \lambda(g/a)$, $\omega_1 = (g/a)^{1/2}$ and $\omega_2 = (g/6a)^{1/2}$, that is, the ratio of frequencies of oscillation of the normal modes is $6^{1/2} : 1$.

We can now find the physical nature of these modes by inserting the solutions $\lambda = \frac{1}{6}$ and $\lambda = 1$ into (8.62). The results are

$$\begin{cases} \lambda = 1, & \omega_1 = (g/a)^{1/2}, \quad \phi = -\theta, \\ \lambda = \frac{1}{6}, & \omega_2 = (g/6a)^{1/2}, \quad \phi = \frac{3}{2}\theta. \end{cases} \tag{8.65}$$

These modes of oscillation are illustrated in Fig. 8.4. According to our analysis, if we set the cylinder oscillating in either of the modes (a) or (b) shown in Fig. 8.4, it will continue to do so for all time at frequencies ω_1 and ω_2 respectively. We also note that we can represent any initial set of conditions by a superposition of modes 1 and 2, if we choose suitable amplitudes and phases for the normal modes. Suppose we represent the x-displacement of the centre of mass of the cylinder by

$$x_1 = A_1(3a\theta + 2a\phi)\,\mathrm{e}^{\mathrm{i}(\omega_1 t+\psi_1)} = A_1 a\theta\,\mathrm{e}^{\mathrm{i}(\omega_1 t+\psi_1)},$$
$$x_2 = A_2(3a\theta + 2a\phi)\,\mathrm{e}^{\mathrm{i}(\omega_2 t+\psi_2)} = 6A_2 a\theta\,\mathrm{e}^{\mathrm{i}(\omega_2 t+\psi_2)},$$

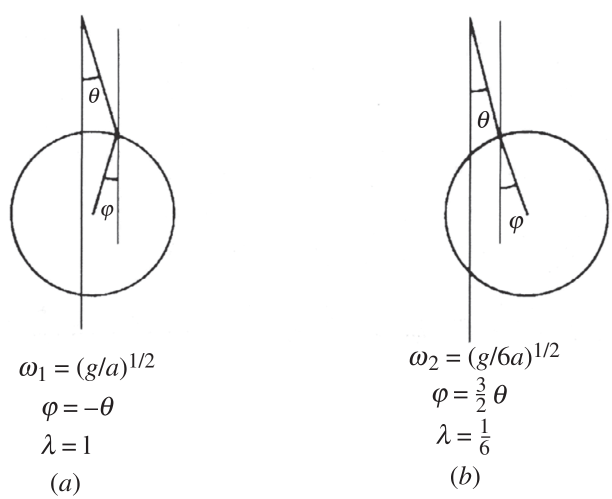


Fig. 8.4 Illustrating the normal modes of oscillation of the cylinder about its equilibrium position.

where ψ_1 and ψ_2 are the phases of the oscillations. Then, if the oscillation begins with some specific values of θ, ϕ, $\dot{\theta}$ and $\dot{\phi}$, we can find suitable values for A_1, A_2, ψ_1 and ψ_2, since we have four initial conditions and four unknowns. The beauty of this procedure is that we can find the behaviour of the system at *any* subsequent time by adding together the independent behaviour of the two normal modes of oscillation, that is

$$x = x_1 + x_2 = A_1 a\theta\, e^{i(\omega_1 t+\psi_1)} + 6A_2 a\theta\, e^{i(\omega_2 t+\psi_2)}.$$

This analysis illustrates the fundamental importance of *normal modes*. Any configuration of the system can be represented by a suitable superposition of normal modes and this enables us to predict the subsequent dynamical behaviour of the system.

Complete sets of orthonormal functions are of particular importance. The functions themselves are independent and normalised to unity so that any function can be represented by a superposition of them. A simple example is the Fourier series

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\frac{2\pi n x}{L} + \sum_{n=1}^{\infty} b_n \sin\frac{2\pi n x}{L},$$

which can describe precisely any function defined in the interval $0 < x < L$. The separate terms with coefficients a_0, a_n, b_n can be thought of as normal modes of oscillation with fixed end points at 0 and L at which appropriate boundary conditions are applied. There are many different sets of orthonormal functions which find application in a wide range of aspects of physics and theoretical physics. For example, in spherical polar coordinates, the Legendre polynomials and associated Legendre polynomials provide complete sets of orthogonal functions defined on a sphere.

In practice, the modes are not completely independent. In real physical situations, there exist small higher order terms in the Lagrangian which result in a coupling between the normal modes. These enable energy to be exchanged between the modes so that eventually energy is shared between them, even if they were very different to begin with. This idea is the basis of the equipartition theorem which will be discussed in more detail in Chapter 12.

If they are not maintained, the modes will eventually decay by dissipative processes. The time evolution of the system can be accurately determined by following the decay of each normal mode with time. In the example given above, the time evolution of the system would be

$$x = A_1 a\theta\, e^{-\gamma_1 t}\, e^{i(\omega_1 t+\psi_1)} + 6A_2 a\theta\, e^{-\gamma_2 t}\, e^{i(\omega_2 t+\psi_2)}.$$

It is evident that dynamical systems can become very complicated. Suppose, for example, the cylinder were allowed to move arbitrarily rather than being required to oscillate with the ends of the cylinder and the string in one plane. The motion would become more complex, but we can guess what the normal modes must be. There are further degrees of freedom which are indicated schematically in Fig. 8.5. These are associated with swinging motion along the axis of the cylinder (Fig. 8.5(a)) and with torsional oscillations of the cylinder about O (Fig. 8.5(b)). Indeed, in many physical problems, one can get quite a long way by guessing what the forms of the normal modes must be by inspection. The key

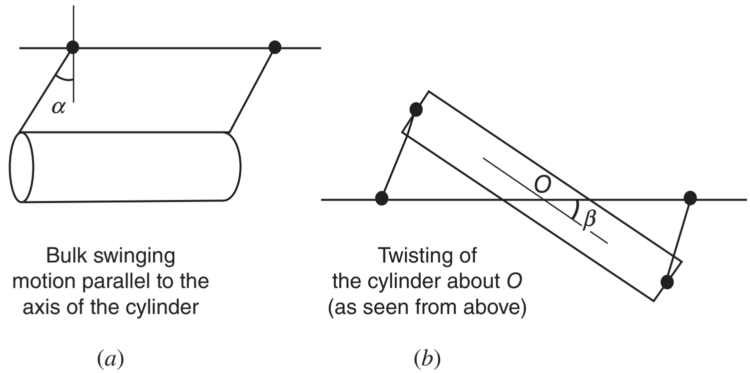


Fig. 8.5 The two other normal models of oscillation of the suspended cylinder.

point is that all the parts of the system must oscillate at the same frequency ω. We will have a great deal more to say about normal modes in our case studies on the origins of statistical mechanics and of the concept of quanta.

8.6 Hamilton's Equations

An equivalent way of developing the equations of mechanics and dynamics is to convert the Euler–Lagrange equations (8.24) into a set of 6N first-order differential equations by introducing the *generalised* or *conjugate momenta* p_i (Section 8.4.1). Taking the derivative of the Lagrangian (8.24) in Cartesian coordinates with respect to the component of velocity $\dot{x}_i$, we find

$$\frac{\partial \mathcal{L}}{\partial \dot{x}_i} = \frac{\partial T}{\partial \dot{x}_i} - \frac{\partial V}{\partial \dot{x}_i} = \frac{\partial T}{\partial \dot{x}_i} = \frac{\partial}{\partial \dot{x}_i} \sum_j \frac{1}{2} m_j \left(\dot{x}_j^2 + \dot{y}_j^2 + \dot{z}_j^2 \right) = m_i \dot{x}_i = p_i . \tag{8.66}$$

The quantity $\partial \mathcal{L}/\partial \dot{q}_i$ is defined to be p_i, the *canonical momentum*, conjugate to the coordinate q_i, by

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}. \tag{8.67}$$

The p_i defined by (8.67) does not necessarily have the dimensions of linear velocity times mass if q_i is not in Cartesian coordinates. Furthermore, if the potential is velocity dependent, even in Cartesian coordinates, the generalised momentum is not identical with the usual definition of mechanical momentum. An example of this is the motion of a charged particle in an electromagnetic field in Cartesian coordinates analysed in Section 8.3 for which the Lagrangian can be written in the non-relativistic case,

$$\mathcal{L} = \sum_i \frac{1}{2} m_i \dot{\boldsymbol{r}}_i^2 - \sum_i e_i \phi(x_i) + \sum e_i \boldsymbol{A}(x_i) \cdot \dot{\boldsymbol{r}}_i , \tag{8.68}$$

where e_i is the charge of the particle i, $\phi(x_i)$ is the electrostatic potential at the particle i and $\boldsymbol{A}(x_i)$ the vector potential at the same position. Then, the x-component of the canonical momentum of the ith particle is

$$p_{ix} = \frac{\partial \mathcal{L}}{\partial \dot{x}_i} = m_i \dot{x}_i + e_i A_x, \tag{8.69}$$

and similarly for the y and z coordinates.

As discussed in Section 8.4, a number of conservation laws can be derived. For example, if the Lagrangian does not depend upon the coordinate q_i, (8.24) shows that the generalised momentum is conserved, $\mathrm{d}\dot{p}_i/\mathrm{d}t = 0$, p_i = constant. Another result, derived in Section 8.4.2, is that the total energy of the system in a conservative field of force with no time-dependent constraints is given by the *Hamiltonian* H which can be written

$$H = \sum_i p_i \dot{q}_i - \mathcal{L}(q, \dot{q}). \tag{8.70}$$

It looks as though H depends upon $p_i, \dot{q}_i$ and q_i but, in fact, we can rearrange the equation so that H is a function of only p_i and q_i. Let us take the total differential of H, assuming $\mathcal{L}$ is time independent. Then

$$\mathrm{d}H = \sum_i p_i\, \mathrm{d}\dot{q}_i + \sum_i \dot{q}_i\, \mathrm{d}p_i - \sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \mathrm{d}\dot{q}_i - \sum_i \frac{\partial \mathcal{L}}{\partial q_i}\, \mathrm{d}q_i. \tag{8.71}$$

Since $p_i = \partial \mathcal{L}/\partial \dot{q}_i$, the first and third terms on the right-hand side cancel. Therefore,

$$\mathrm{d}H = \sum_i \dot{q}_i\, \mathrm{d}p_i - \sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i}\, \mathrm{d}q_i. \tag{8.72}$$

This differential depends only on the increments $\mathrm{d}p_i$ and $\mathrm{d}q_i$ and hence we can compare $\mathrm{d}H$ with its formal expansion in terms of p_i and q_i:

$$\mathrm{d}H = \sum_i \frac{\partial H}{\partial p_i} \mathrm{d}p_i + \sum_i \frac{\partial H}{\partial q_i}\, \mathrm{d}q_i.$$

It follows immediately that

$$\frac{\partial H}{\partial q_i} = -\frac{\partial \mathcal{L}}{\partial q_i}; \qquad \frac{\partial H}{\partial p_i} = \dot{q}_i. \tag{8.73}$$

Since

$$\frac{\partial \mathcal{L}}{\partial q_i} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right), \tag{8.74}$$

we find from the Euler–Lagrange equation,

$$\frac{\partial H}{\partial q_i} = -\dot{p}_i.$$

We thus reduce the equations of motion to the pair of relations

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}. \tag{8.75}$$

These two equations are known as *Hamilton's equations*. They are now *first-order differential equations* for each of the $3N$ coordinates. We are now treating the p_is and the q_is on the same footing. If V is independent of $\dot{q}$, H is just the total energy $T + V$ expressed in terms of the coordinates p_i and q_i. Inserting $H = T + V$ into the equation for $\dot{p}_i$ in (8.75), we immediately recover Newton's second law of motion, $\boldsymbol{f} = \dot{p}_i = -\nabla_i V$.

8.7 Hamilton's Equations and Poisson Brackets

Poisson brackets can be used in conjunction with Hamilton's equations of motion to reduce the formalism to a yet more compact form. In addition, they provide a route to the incorporation of quantum mechanics within the framework of Hamiltonian mechanics. The Poisson bracket for the functions g and h is defined to be the quantity

$$[g, h] = \sum_{i=1}^{n} \left(\frac{\partial g}{\partial p_i} \frac{\partial h}{\partial q_i} - \frac{\partial g}{\partial q_i} \frac{\partial h}{\partial p_i} \right). \tag{8.76}$$

We can also write in general

$$\dot{g} = \sum_{i=1}^{n} \left(\frac{\partial g}{\partial q_i} \dot{q}_i + \frac{\partial g}{\partial p_i} \dot{p}_i \right) \tag{8.77}$$

for the variation of any physical quantity g and hence, using Hamilton's equations (8.75) with $h = H$, we can write

$$\dot{g} = [H, g]. \tag{8.78}$$

Therefore, Hamilton's equations can be written

$$\dot{q}_i = [H, q_i], \qquad \dot{p}_i = [H, p_i].$$

The Poisson brackets have a number of useful properties. If we identify g with q_i and h with q_j in (8.76), we find

$$[q_i, q_j] = 0 \quad \text{and} \quad [p_j, p_k] = 0.$$

If $j \neq k$,

$$[p_j, q_k] = 0,$$

but if $g = p_k$ and $h = q_k$,

$$[p_k, q_k] = 1, \quad [q_k, p_k] = -1.$$

Quantities with Poisson brackets which are zero are said to *commute*. Those with Poisson brackets equal to unity are said to be *canonically conjugate*. From the relation (8.78), we note that any quantity which commutes with the Hamiltonian does not change with time. In particular, H itself is constant in time because it commutes with itself. Yet again we have returned to the conservation of energy.

These quantities played an important role in the development of quantum mechanics, as demonstrated in Dirac's classic text *The Principles of Quantum Mechanics*.[5] The key feature of quantum mechanics is that pairs of dynamical variables such as momentum and position coordinates are non-commuting and these features of Poisson brackets provide a natural language for quantum physics.

There is a delightful story in Dirac's memoirs about how he came to realise their importance. In October 1925, Dirac was worried by the fact that, according to his and Heisenberg's independent formulations of quantum mechanics, the dynamical variables did not commute, that is, for two variables u and v, uv was not the same as vu. Dirac had a strict rule about relaxing on Sunday afternoons by taking country walks. He writes:

> It was during one of the Sunday Walks in October 1925 when I was thinking very much about this $uv - vu$, in spite of my intention to relax, that I thought about Poisson brackets . . . I did not remember very well what a Poisson bracket was. I did not remember the precise formula for a Poisson bracket and only had some vague recollections. But there were exciting possibilities there and I thought that I might be getting on to some big new idea.
>
> Of course, I could not [find out what a Poisson bracket was] right out in the country. I just had to hurry home and see what I could then find about Poisson brackets. I looked though my notes and there was no reference there anywhere to Poisson brackets. The text books which I had at home were all too elementary to mention them. There was nothing I could do, because it was Sunday evening then and the libraries were all closed. I just had to wait impatiently through that night without knowing whether this idea was any good or not but still I think that my confidence grew during the course of the night. The next morning, I hurried along to one of the libraries as soon as it was open and then I looked up Poisson brackets in Whittaker's *Analytic Dynamics* and I found that they were just what I needed. They provided the perfect analogy with the commutator.[6]

This represented a major breakthrough in formulating the key role of non-commuting variables and operators in quantum mechanics.[7] Dirac's great achievement was to incorporate quantum concepts within the formalism of Hamiltonian mechanics.

8.8 The Hamilton–Jacobi Equations and Action–Angle Variables

Hamilton's equations (8.75) are often difficult to solve but they can be simplified to tackle specific problems by suitable changes of variable. It turned out that the appropriate tools were available through Carl Jacobi's extension of Hamilton's equations which resulted in the Hamilton–Jacobi equation, canonical transformations and action-angle variables. These procedures had been used with considerable success in astronomical problems, for example, in Delaunay's *Théorie du mouvement de la lune*[8] and in Charlier's *Die Mechanik des Himmels*.[9] In these cases, the transformations were used to change coordinates to those most convenient for analysing planetary and lunar orbits and then studying their stability using perturbation techniques in the new coordinate system. The old quantum theory of electron orbits about atomic nuclei took over these procedures in order to understand perturbations of their elliptical orbits. It is no surprise that Karl Schwarzschild, theoretical astrophysicist and the Director of the Potsdam Observatory, was one of the principal contributors to the introduction of Hamilton–Jacobi theory into the solution of problems in the old quantum theory.

The objective of the Hamilton–Jacobi approach is to transform from one set of coordinates to another, while preserving the Hamiltonian form of the equations of motion and ensuring that the new variables are also canonical coordinates. This approach is sometimes called the *transformation theory of mechanics*. Thus, suppose we wish to transform between the set of coordinates q_i to a new set Q_i. We need to ensure that the independent momentum coordinates simultaneously transform into canonical coordinates in the new system, that is, there should also be a set of coordinates P_i corresponding to p_i. Thus, the problem is to define the new coordinates (Q_i, P_i) in terms of the old coordinates (q_i, p_i):

$$Q_i = Q_i(q, p, t), \quad P_i = P_i(q, p, t). \tag{8.79}$$

Furthermore, the transformations should be such that Q_i and P_i are canonical coordinates, meaning that

$$\dot{Q}_i = \frac{\partial K}{\partial P_i}, \quad \dot{P}_i = -\frac{\partial K}{\partial Q_i}, \tag{8.80}$$

where K is now the Hamiltonian in the (Q_i, P_i) coordinate system. The transformations which result in (8.80) are known as *canonical transformations* or *contact transformations*. In the old system, the coordinates satisfied Hamilton's variational principle,

$$\delta \int_{t_1}^{t_2} \left[\sum p_i \dot{q}_i - H(q, p, t)\right] \mathrm{d}t = 0, \tag{8.81}$$

and so the same rule must hold true in the new set of coordinates,

$$\delta \int_{t_1}^{t_2} \left[\sum P_i \dot{Q}_i - K(Q, P, t)\right] \mathrm{d}t = 0. \tag{8.82}$$

The equations (8.81) and (8.82) must be valid simultaneously.

Thus, the problem reduces to finding the appropriate generating functions for the problem at hand. The clever trick is to note that the integrands inside the integrals in (8.81) and (8.82) need not be equal, but can differ by the total time derivative of some arbitrary function S. This works because the integral of S between the fixed end points is a constant and so disappears when the variations (8.81) and (8.82) are carried out,

$$\delta \int_{t_1}^{t_2} \frac{\mathrm{d}S}{\mathrm{d}t}\, \mathrm{d}t = \delta(S_2 - S_1) = 0. \tag{8.83}$$

Once the generating function S is known, the transformation equations (8.79) are completely determined. It looks as though S is a function of the four coordinates (q_i, p_i, Q_i, P_i), but in fact they are related by the functions (8.79) and so S can be defined by any two of (q_i, p_i, Q_i, P_i). Goldstein[3] and Lindsay and Margenau [10] show how the various transformations between the coordinate systems can be written. For example, if we choose to write the transformations in terms of p_i, P_i, the transformation equations are

$$p_i = \frac{\partial S}{\partial q_i}, \quad P_i = -\frac{\partial S}{\partial Q_i} \quad \text{and} \quad K = H + \frac{\partial S}{\partial t}. \tag{8.84}$$

Alternatively, if we choose q_i, P_i, the transformation equations become

$$p_i = \frac{\partial S}{\partial q_i}, \quad Q_i = \frac{\partial S}{\partial P_i} \quad \text{with} \quad K = H + \frac{\partial S}{\partial t}. \tag{8.85}$$

The other two sets of transformations are included in the endnote.[11]

The question is then, 'Can we find the appropriate generating functions for the problem at hand?' For the cases in which we are interested in which there is no explicit dependence upon the time t and in which the variables in the partial differential equations are separable, the answer is 'Yes'. We begin with the relation between the Hamiltonian H and the total energy of the system E. For the cases we are interested in,

$$H(q_i, p_i) = E. \tag{8.86}$$

We can now replace p_i by its definition in terms of the partial differentials of the generating function S, $p_i = \partial S/\partial q_i$, so that

$$H\left(q_i, \frac{\partial S}{\partial q_i}\right) = E. \tag{8.87}$$

Writing out the partial derivatives, this is a set of n first-order partial differential equations for the generating function S,

$$H\left(q_1, q_2, \dots q_n, \frac{\partial S}{\partial q_1}, \frac{\partial S}{\partial q_2}, \dots \frac{\partial S}{\partial q_n}\right) = E. \tag{8.88}$$

Equation (8.88) is known as the *Hamilton–Jacobi equation* in time-independent form. S is referred to as *Hamilton's principal function*.

The astronomers were interested in periodic solutions of the equations of motion to describe periodic planetary orbits under a central gravitational field of force. There is no explicit dependence of the coordinate transformations upon time and so the transformation equations (8.85) become

$$p_i = \frac{\partial S}{\partial q_i}, \quad Q_i = \frac{\partial S}{\partial P_i} \quad \text{with} \quad K = H. \tag{8.89}$$

Let us carry out the simple example of a harmonic oscillator to illustrate how the procedure works before setting out the procedure in more general terms.

Example: the Harmonic Oscillator The mass of the oscillator is m and the spring constant k. From these, we can define the angular frequency ω by $\omega^2 = k/m$. The Hamiltonian is therefore

$$H = \frac{p^2}{2m} + \frac{kq^2}{2}. \tag{8.90}$$

Now introduce a generating function defined by

$$S = \frac{m\omega}{2} q^2 \cot Q. \tag{8.91}$$

which depends only upon q and Q. We can therefore use (8.84) to find the transformations between coordinate systems. There is no explicit dependence of S upon time t and so, from the analysis which leads to the third equation of (8.84), $H = K$. We can now find the coordinates p and q in terms of the new coordinates P and Q. From the first and second equations of (8.84)

$$p = \frac{\partial S}{\partial q} = m\omega q \cot Q, \quad P = -\frac{\partial S}{\partial Q} = \frac{m\omega}{2} q^2 \operatorname{cosec}^2 Q. \tag{8.92}$$

These can be reorganised to give expressions for p and q in terms of P and Q:

$$q = \sqrt{\frac{2}{m\omega}}\sqrt{P}\,\sin Q, \quad p = \sqrt{2m\omega}\sqrt{P}\cos Q. \tag{8.93}$$

Substituting these expressions for p and q into the Hamiltonian (8.90), we find

$$H = K = \frac{p^2}{2m} + \frac{kq^2}{2} = \omega P. \tag{8.94}$$

We now find the solutions for P and Q from Hamilton's equations:

$$\dot{P} = -\frac{\partial K}{\partial Q}, \quad \dot{Q} = \frac{\partial K}{\partial P}, \tag{8.95}$$

and so, from (5.94), it follows that

$$\dot{P} = 0, \dot{Q} = \omega \quad \text{and so} \quad P = \alpha, Q = \omega t + \beta, \tag{8.96}$$

where α and β are constants.

The expressions (8.96) are the key results. The canonical transformations have resulted in a natural set of coordinates for the harmonic oscillator, namely, the amplitude of the oscillations is related to $\alpha \equiv P$ and the phase of the oscillations is $Q = \omega t + \beta$.

These solutions can now be substituted into (8.93) and we find

$$p = \sqrt{2m\omega\alpha}\,\cos(\omega t + \beta), \quad q = \sqrt{\frac{2\alpha}{m\omega}}\,\sin(\omega t + \beta). \tag{8.97}$$

For the general case of periodic orbits, we define a new coordinate

$$J_i = \oint p_i \,\mathrm{d}q_i \tag{8.98}$$

which is known as a *phase integral*, the integration of the momentum variable p_i being taken over a complete cycle of the values of q_i. In the case of Cartesian coordinates, it can be seen that J_i has the dimensions of angular momentum and is known as an *action variable*, by analogy with Hamilton's definition in his *principle of least action*. During one cycle the values of q_i vary continuously between finite values from $q_{\min}$ to $q_{\max}$ and back again to $q_{\min}$. As a result, from the definition (8.98), the action J_i is a constant. If q_i does not appear in the Hamiltonian, as occurs in the cases of interest to us, the variable is said to be *cyclic*.

Accompanying J_i there is a conjugate quantity w_i,

$$w_i = \frac{\partial S}{\partial J_i}. \tag{8.99}$$

The corresponding canonical equations are

$$\dot{w}_i = \frac{\partial K}{\partial J_i}, \quad \dot{J}_i = -\frac{\partial K}{\partial w_i}, \tag{8.100}$$

where K is the transformed Hamiltonian. w_i is referred to as an *angle variable* and so the motion of the particle is now described in terms of (J_i, w_i) or *action-angle variables*. If the transformed Hamiltonian K is independent of w_i, it follows that

$$\dot{w}_i = \text{constant} = \nu_i, \quad \dot{J}_i = 0, \tag{8.101}$$

where the ν_is are constants. Therefore,

$$w_i = \nu_i t + \gamma_i, \quad J_i = \text{constant}. \tag{8.102}$$

The constants ν_i are associated with the frequency of motion. If we take the integral of w_i round a complete cycle of q_i, we find

$$\Delta w_i = \oint \frac{\partial w_i}{\partial q_k}\, \mathrm{d}q_k = \oint \frac{\partial}{\partial q_k}\left(\frac{\partial S}{\partial J_i}\right) \mathrm{d}q_k. \tag{8.103}$$

Taking the differentiation outside the integral,

$$\Delta w_i = \frac{\partial}{\partial J_i} \oint \left(\frac{\partial S}{\partial q_k}\right) \mathrm{d}q_k = \frac{\partial J_k}{\partial J_i}, \tag{8.104}$$

the last equality resulting from the definition $p_k = \partial S/\partial q_k$. Thus, $\Delta w_i = 1$, if $i = k$ and $\Delta w_i = 0$, if $i \neq k$. In other words, each independent variable w_i changes from 0 to 1 when the corresponding q_i goes through a complete cycle.

It is apparent from these considerations that action-angle variables are the ideal tools for studying planetary orbits and perturbations to them. The works of Delaunay and Charlier were central to understanding phenomena such as perturbed Keplerian dynamics, the precession of the equinoxes, tidal locking, the instability of perturbed polar orbits (the Kozai–Lidov effect), resonances and so on. Ultimately, they lead to phenomena such as chaotic orbits, which we will return to in Chapter 10, and the Kolmogorov–Arnold–Moser (KAM) theorem concerning the persistence of quasi-periodic motions in dynamical systems under small perturbations. Part of the astronomical agenda is to address the important question, ‘Is the Solar System stable?’ These important subjects go far beyond the ambitions of this book, but they illustrate the very rich harvest of dynamical effects which can be studied by these more advanced techniques.

These techniques were also important in the early development of the old quantum theory. The mathematical techniques developed by the astronomers could be immediately adapted to study the elliptical electron orbits in atoms. Significant successes included Schwarzschild and Epstein’s explanation of the Stark effect, the splitting of spectral lines by an applied electric field.[12] As remarked by Max Jammer,[13]

> ... it almost seemed as if Hamilton’s method had expressly been created for treating quantum mechanical problems.

But the euphoria did not last long. The old quantum theory was, to quote Jammer again, a 'lamentable hodge-podge' of quantum rules bolted on to classical mechanics and dynamics – something much more profound was needed and this was to be supplied by Werner Heisenberg's great paper of 1925.[7]

8.9 A Warning

My approach to classical mechanics has intentionally been non-rigorous in order to bring out important features of different approaches to mechanics and dynamics. The subject often appears somewhat abstruse and mathematical and I have intentionally concentrated upon the simplest parts which relate most intuitively to our understanding of Newton's laws. I would emphasise that the subject can be made rigorous and such approaches are found in books such as Goldstein's *Classical Mechanics*[3] or Landau and Lifshitz's *Mechanics*.[14] My aim has been to show that there are very good physical reasons for developing these more challenging approaches to mechanics and dynamics.

Notes

1 See Section 1.3.

2 As in so many aspects of basic theoretical physics, I find Feynman's approach in Chapter 19 of Volume 2 of his *Lectures on Physics*,[4] entitled *The Principle of Least Action*, quite brilliant in exposing the basic ideas behind this approach to dynamics. Two notes are worth making about the use of the term 'minimum' or 'least action'. The first is that we will take the term generically to mean finding the stationary value of some function, subject to a fixed set of boundary conditions. This will normally mean minimising some function, but finding the stationary function, subject to the boundary conditions, is perhaps a more helpful terminology. Second, the term 'action' itself is ambiguous. It also appears in 'action-angle variables' with a specific meaning for periodic systems (see Section 8.8).

3 Goldstein, H. (1950). *Classical Mechanics*. London: Addison-Wesley.

4 Feynman, R.P. (1964). *The Feynman Lectures in Physics*, Vol. 2, eds. R.P. Feynman, R.B. Leighton and M. Sands. Chapter 19. London: Addison-Wesley.

5 Dirac, P.A.M. (1930). *The Principles of Quantum Mechanics*. Oxford: Clarendon Press. The second, third and fourth editions appeared in 1935, 1947 and 1958.

6 Dirac, P.A.M. (1977). In *History of Twentieth Century Physics*, Proceedings of the International School of Physics 'Enrico Fermi', Course 57, p. 122. New York and London: Academic Press.

7 For more details, see my book *Quantum Concepts in Physics* (2013). Cambridge: Cambridge University Press.

8 Delaunay, C.E. Théorie du mouvement de la lune, *Memoires de l'Académie des Sciences (Paris)*, Vol. I (1860) **28** and Vol. II (1867) **29**.

9 Charlier, C.V.L. (1902). *Die Mechanik des Himmels*. Liepzig: Veit.

10 Lindsay, R.B. and Margenau, H. (1957). *Foundations of Physics*. London: Constable & Co. (Dover reprint).

11 There are four transformations between the p_i, q_i, P_i and Q_i. Two of them are given in (8.92) and (8.93). The other two are:

$$q_i = -\frac{\partial S}{\partial p_i}, \quad P_i = -\frac{\partial S}{\partial Q_i} \quad \text{and} \quad K = H + \frac{\partial S}{\partial t},$$

and

$$q_i = -\frac{\partial S}{\partial p_i}, \quad Q_i = \frac{\partial S}{\partial P_i} \quad \text{with} \quad K = H + \frac{\partial S}{\partial t}.$$

12 Schwarzschild, K. (1916). Zur QuantenHypothese, *Sitzungberichte der (Kgl.) Preussischen Akademie der Wissenschaften (Berlin)*, 548–568; Epstein, P.S. (1916). Zur Theorie des Starkeffektes (On the theory of the Stark effect), *Annalen der Physik (4)*, **50** 489–520.

13 Jammer, M. 1989. *The Conceptual Development of Quantum Mechanics*, 2nd edition. New York: American Institute of Physics and Tomash Publishers.

14 Landau, L.D. and Lifshitz, E.M. (1960). *Mechanics*, Vol. 5 of *Course of Theoretical Physics*. Oxford: Pergamon Press.

9 The Motion of Fluids

In Chapter 8, the increasingly sophisticated techniques developed to solve problems in classical mechanics and dynamics for point masses and extended bodies were described. Now we tackle the laws of motion as applied to fluids – liquids and gases. Fluid dynamics is a vast subject, which can rapidly become very complex because of the intrinsic non-linearity of the equations needed to tackle many problems. Fluids play a central role in many areas of the physical, biological and medical sciences, as well as engineering. They are of central importance for everyday life – the understanding of climate and weather, water use and distribution, the flow of blood in veins and arteries, the operation of wind farms, helicopters and aircraft. The list is essentially endless and challenging.

The intention of this chapter is to provide a gentle introduction to some of the key concepts and more familiar problems as preparation for tackling major classics of the subject such as Batchelor's text *An Introduction to Fluid Dynamics*[1] and Landau and Lifshitz's *Fluid Mechanics*.[2] These challenging books repay careful study. As is so often the case, Feynman provides a clear and thoughtful presentation of the fundamentals of fluid mechanics and dynamics in Chapters 40 and 41 of Volume 2 of *The Feynman Lectures on Physics*.[3]

The inspiration for the present chapter was the book by T.E. Faber *Fluid Dynamics for Physicists*.[4] Tom Faber was an admired colleague in the Cavendish Laboratory who had a very deep understanding of physics and an enthusiasm for teaching the physics of fluids at the undergraduate level. His book is very strongly physics, rather than mathematics, based and successfully makes the physics of even the most complex fluids problems accessible to students. Chapter 1 is a remarkable tour-de-force in which Faber reveals the huge number of considerations which need to be taken into account in describing the fluid mechanics of a simple syringe and is recommended reading.

It would be pointless to attempt to cover the vast range of topics and insights contained in Faber's book. My aim is much more modest and that is to highlight the distinctive features of the physics of fluids and then to explore some classic results discovered by the pioneers of the subject. In Chapter 10 we return to some of the more challenging non-linear aspects of the subject in the context of non-linear dynamical systems.

9.1 The Equation of Continuity

First of all, we derive the equation which tells us that a fluid does not 'get lost'. By a procedure similar to that used in electromagnetism (Section 7.2), consider the net mass flux

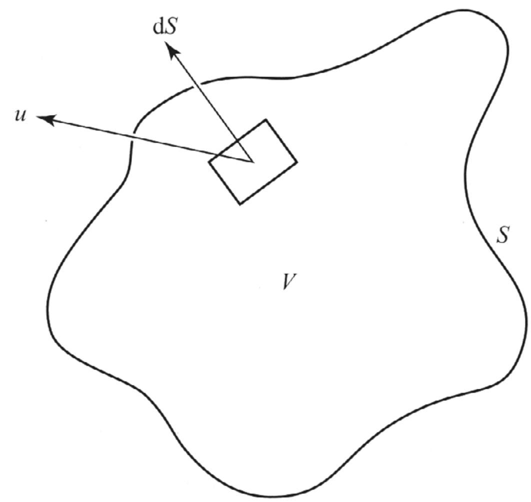


Fig. 9.1 Illustrating the mass flux of fluid from a volume V.

from a volume V bounded by a surface S. If $\mathrm{d}\boldsymbol{S}$ is an element of surface area, the direction of the vector being taken normally outwards, the mass flow through $\mathrm{d}\boldsymbol{S}$ is $\rho\boldsymbol{u}\cdot\mathrm{d}\boldsymbol{S}$, where $\boldsymbol{u}$ is the fluid velocity and ρ its density (Fig. 9.1). Integrating over the surface of the volume, the rate of outflow of mass must equal the rate of loss of mass from within v, that is,

$$\int_S \rho\boldsymbol{u}\cdot\mathrm{d}\boldsymbol{S} = -\frac{\mathrm{d}}{\mathrm{d}t}\int_v \rho\,\mathrm{d}v. \tag{9.1}$$

Using the divergence theorem,

$$\int_v \mathrm{div}(\rho\boldsymbol{u})\,\mathrm{d}v = -\frac{\mathrm{d}}{\mathrm{d}t}\int_v \rho\,\mathrm{d}v, \quad \text{that is,} \quad \int_v \left(\mathrm{div}\,\rho\boldsymbol{u} + \frac{\mathrm{d}\rho}{\mathrm{d}t}\right)\mathrm{d}v = 0.$$

This result must be true for any volume within the fluid and hence

$$\mathrm{div}\,\rho\boldsymbol{u} + \frac{\partial\rho}{\partial t} = 0. \tag{9.2}$$

This is the *equation of continuity*. Notice that we have changed from total to partial derivatives between the last two equations. The reason is that, in microscopic form, we use partial derivatives to denote variations in the properties of the fluid *at a fixed point in space*. This is distinct from derivatives which follow the properties of a particular element of the fluid. The latter are denoted by total derivatives, for example, $\mathrm{d}\rho/\mathrm{d}t$, as can be appreciated from the following argument.

We define $\mathrm{d}\rho/\mathrm{d}t$ to be the rate of change of density of a fluid element *as it moves through the fluid* (Fig. 9.2). The motion of the element is followed over the short time interval δt. If the velocity of the fluid at (x, y, z) is $\boldsymbol{u} = (u_x, u_y, u_z)$, it follows that $\mathrm{d}\rho/\mathrm{d}t$ is given by

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \lim_{\delta t\to 0}\frac{1}{\delta t}[\rho(x + u_x\,\delta t, y + u_y\,\delta t, z + u_z\,\delta t, t + \delta t) - \rho(x, y, z, t)].$$

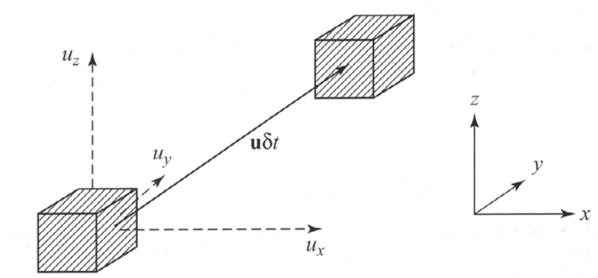


Fig. 9.2 Illustrating the relation between the total and partial derivatives, d/dt and $\partial/\partial t$ respectively.

Now perform a Taylor expansion for the first term in square brackets:

$$\begin{aligned}\frac{\mathrm{d}\rho}{\mathrm{d}t} &= \lim_{\delta t\to 0}\frac{1}{\delta t}\left[\rho(x,y,z,t)+\frac{\partial\rho}{\partial x}u_x\,\delta t+\frac{\partial\rho}{\partial y}u_y\,\delta t+\frac{\partial\rho}{\partial z}u_z\,\delta t+\frac{\partial\rho}{\partial t}\,\delta t-\rho(x,y,z,t)\right]\\ &= \frac{\partial\rho}{\partial t}+u_x\frac{\partial\rho}{\partial x}+u_y\frac{\partial\rho}{\partial y}+u_z\frac{\partial\rho}{\partial z}.\end{aligned} \tag{9.3}$$

We can write this in compact form as

$$\frac{\mathrm{d}\rho}{\mathrm{d}t}=\frac{\partial\rho}{\partial t}+\boldsymbol{u}\cdot\operatorname{grad}\rho\;=\frac{\partial\rho}{\partial t}+(\boldsymbol{u}\cdot\nabla)\,\rho, \tag{9.4}$$

where the operator $(\boldsymbol{u}\cdot\nabla)$ means

$$(\boldsymbol{u}\cdot\nabla)=u_x\frac{\partial}{\partial x}+u_y\frac{\partial}{\partial y}+u_z\frac{\partial}{\partial z}. \tag{9.5}$$

Equation (9.3) is no more than the chain rule in differentiation and this accounts for our definitions of d/dt and $\partial/\partial t$. Equation (9.4) is also a relation between differential operators:

$$\frac{\mathrm{d}}{\mathrm{d}t}=\left(\frac{\partial}{\partial t}+\boldsymbol{u}\cdot\operatorname{grad}\right), \tag{9.6}$$

which recurs throughout fluid dynamics.

We can work in either of these coordinate systems. If we use the frame of reference which follows an element of the fluid, the coordinates are called *Lagrangian coordinates*. If we work in a fixed external reference frame, they are called *Eulerian coordinates*. Generally, Eulerian coordinates are used, although there are occasions when it is simpler to use a Lagrangian approach, for example, when forces act upon an element of the fluid.

Expression (9.4) enables us to rewrite the equation of continuity (9.2) as follows:

$$\operatorname{div}\rho\boldsymbol{u}+\frac{\mathrm{d}\rho}{\mathrm{d}t}=\boldsymbol{u}\cdot\operatorname{grad}\rho.$$

Expanding $\operatorname{div}\rho\boldsymbol{u}$,

$$\boldsymbol{u}\cdot\operatorname{grad}\rho+\rho\operatorname{div}\boldsymbol{u}+\frac{\mathrm{d}\rho}{\mathrm{d}t}=\boldsymbol{u}\cdot\operatorname{grad}\rho\qquad\text{and so}\qquad\frac{\mathrm{d}\rho}{\mathrm{d}t}=-\rho\operatorname{div}\boldsymbol{u}. \tag{9.7}$$

If the fluid is incompressible, ρ = constant and the flow is described by $\operatorname{div}\boldsymbol{u}=0$. If the flow is not time dependent $\partial\rho/\partial t=0$ and is *irrotational*, meaning $\operatorname{curl}\boldsymbol{u}=0$, the velocity field $\boldsymbol{u}$ can be expressed in terms of a velocity potential ϕ defined by $\boldsymbol{u} = \text{grad}\,\phi$. Note the opposite sign convention as compared with the electrostatic potential for the electric field strength $\boldsymbol{E} = -\text{grad}\,\phi$. The velocity potential ϕ can then be found as the solution of Laplace's equation,

$$\nabla^2\phi = 0, \tag{9.8}$$

subject to satisfying the boundary conditions. The analogy with the mathematics of electrostatics and magnetostatics is a fruitful approach to both subjects and was fully exploited by Maxwell, as recounted in Chapter 5.

9.2 The Equations of Motion for an Incompressible Fluid in the Absence of Viscosity

9.2.1 Euler's Equation

We begin with the simplest case of the force acting on an element of the fluid of dimensions $\Delta x, \Delta y, \Delta z$ in Cartesian coordinates. The pressure is always assumed to be isotropic at any point in the fluid but it can vary with position. It is the differences in pressure across the fluid element which results in the force acting upon it. Consider first the force acting on the surface $\Delta A = \Delta y\,\Delta z$ in the x-direction and make a Taylor expansion for small Δx,

$$f_x = \rho\,\Delta v \frac{\mathrm{d}u_x}{\mathrm{d}t} = [p(x) - p(x + \Delta x)]\,\Delta A = \left[p(x) - \left(p(x) + \frac{\partial p}{\partial x}\Delta x\right)\right]\,\Delta y\,\Delta z = -\frac{\partial p}{\partial x}\,\Delta v,$$

$$\rho\,\frac{\mathrm{d}u_x}{\mathrm{d}t} = -\frac{\partial p}{\partial x}, \tag{9.9}$$

where $\Delta v = \Delta x\,\Delta y\,\Delta\,z$ is the volume of the fluid element. The fluid element experiences similar forces in the y and z directions and we can add these vectorially to obtain the equation

$$\rho\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} = -\text{grad}\,p = -\nabla p. \tag{9.10}$$

Often, the flow takes place in a gravitational field in which case we can add the term $-\nabla\phi$ to (9.10), so that

$$\rho\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} = -\text{grad}\,p - \rho\,\text{grad}\,\phi, \tag{9.11}$$

where ϕ is the gravitational potential per unit mass. Equation (9.11) is known as *Euler's equation*. For terrestrial applications, the gravitational potential can be written in terms of the acceleration due to gravity g acting in the vertical z-direction relative to some reference value. Then the gravitational potential can be simply written $\phi = gz$.[5]

We have been careful to write (9.11) as a Lagrangian derivative, meaning that the force acts in the frame of reference which moves with the local velocity of the fluid. Now, let us rewrite this equation in terms of *Eulerian coordinates*, that is, in terms of partial rather than total differentials. The analysis proceeds exactly as in Section 9.1, but now we consider the vector quantity $\boldsymbol{u}$, rather than the scalar quantity ρ. This is only a minor complication because the three components of the vector $\boldsymbol{u}$ (u_x, u_y, u_z), are scalar functions and so the relation (9.6) can be applied directly in each coordinate. For example,

$$\frac{\mathrm{d}}{\mathrm{d}t}u_x = \frac{\partial u_x}{\partial t} + \boldsymbol{u} \cdot \mathrm{grad}\, u_x .$$

There are similar equations for u_y and u_z and therefore, adding all three together vectorially,

$$\begin{aligned}\boldsymbol{i}_x \frac{\mathrm{d}u_x}{\mathrm{d}t} + \boldsymbol{i}_y \frac{\mathrm{d}u_y}{\mathrm{d}t} + \boldsymbol{i}_z \frac{\mathrm{d}u_z}{\mathrm{d}t} &= \boldsymbol{i}_x \frac{\partial u_x}{\partial t} + \boldsymbol{i}_y \frac{\partial u_y}{\partial t} + \boldsymbol{i}_z \frac{\partial u_z}{\partial t} \\ &+ \boldsymbol{i}_x (\boldsymbol{u} \cdot \mathrm{grad}\, u_x) + \boldsymbol{i}_y (\boldsymbol{u} \cdot \mathrm{grad}\, u_y) + \boldsymbol{i}_z (\boldsymbol{u} \cdot \mathrm{grad}\, u_z).\end{aligned}$$

Or,

$$\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} = \frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u} \cdot \mathrm{grad})\, \boldsymbol{u}. \tag{9.12}$$

Notice that the operator $(\boldsymbol{u} \cdot \mathrm{grad})$ means that we perform the operation $[u_x(\partial/\partial x) + u_y(\partial/\partial y) + u_z(\partial/\partial z)]$ on *all three of the components of the vector* $\boldsymbol{u}$ to find the magnitude of $(\boldsymbol{u} \cdot \mathrm{grad})$ in each direction. The equation of motion in Eulerian coordinates therefore becomes

$$\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u} \cdot \mathrm{grad})\, \boldsymbol{u} = -\frac{1}{\rho}\, \mathrm{grad}\, p - \mathrm{grad}\, \phi. \tag{9.13}$$

This equation indicates clearly where the complexities of fluid mechanics come from. The second term on the left-hand side $(\boldsymbol{u} \cdot \mathrm{grad})\, \boldsymbol{u}$ introduces a nasty non-linearity in the equation for the velocity $\boldsymbol{u}$ and this causes all sorts of complications when trying to find exact solutions in fluid dynamical problems. It is obvious that the subject can rapidly become one of great mathematical complexity.

9.2.2 Bernoulli's Theorem

An interesting way of rewriting (9.13) is to introduce the vector $\boldsymbol{\omega}$, called the *vorticity* of the flow, which is defined by $\boldsymbol{\omega} = \mathrm{curl}\, \boldsymbol{u} = \nabla \times \boldsymbol{u}$. Now, according to one of the standard vector identities,

$$\boldsymbol{u} \times \boldsymbol{\omega} = \boldsymbol{u} \times (\mathrm{curl}\, \boldsymbol{u}) = \tfrac{1}{2}\, \mathrm{grad}\, \boldsymbol{u}^2 - (\boldsymbol{u} \cdot \mathrm{grad})\, \boldsymbol{u}, \tag{9.14}$$

and so the equation of motion becomes

$$\frac{\partial \boldsymbol{u}}{\partial t} - (\boldsymbol{u} \times \boldsymbol{\omega}) = -\mathrm{grad}\left(\tfrac{1}{2}\, u^2 + \frac{p}{\rho} + \phi\right). \tag{9.15}$$

Equation (9.15) is useful for finding specific solutions to certain classes of fluid dynamical problems. For example, if we are concerned only with steady motion, $\partial \boldsymbol{u}/\partial t = 0$ and hence

$$\boldsymbol{u} \times \boldsymbol{\omega} = \mathrm{grad}\left(\tfrac{1}{2}u^2 + \frac{p}{\rho} + \phi\right). \tag{9.16}$$

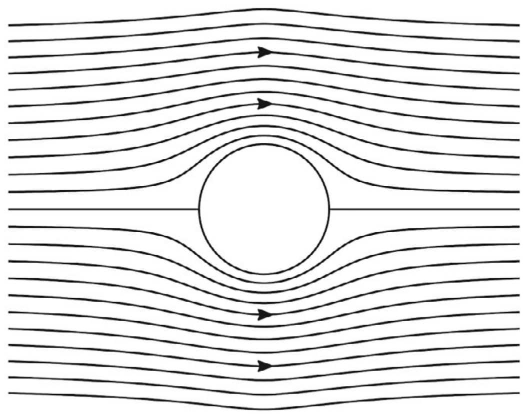


Fig. 9.3 Illustrating the streamlines associated with potential flow past a solid sphere of an incompressible fluid with zero viscosity. (Courtesy of Creative Commons Attribution-Share Alike 3.0 Unported license.)

If the flow is *irrotational*, meaning that $\boldsymbol{\omega} = \text{curl}\,\boldsymbol{u} = 0$, $\boldsymbol{u}$ can be derived from the gradient of a scalar potential (see Appendix A5.4). This is often referred to as *potential flow*. Then, the right-hand side must be zero and hence

$$\tfrac{1}{2}u^2 + \frac{p}{\rho} + \phi = \text{constant}. \tag{9.17}$$

Another important way in which this conservation law can be applied is to introduce the concept of *streamlines*, defined to be lines in the fluid which are everywhere parallel to the instantaneous velocity vector $\boldsymbol{u}$. For example, Fig. 9.3 shows the streamlines associated with the flow of an incompressible fluid with zero viscosity past a solid sphere. If we follow the flow along a streamline, the quantity $\boldsymbol{u} \times \boldsymbol{\omega}$ must be perpendicular to $\boldsymbol{u}$ and hence if we take the gradient of $(\tfrac{1}{2}u^2 + p/\rho + \phi)$ in the direction along a streamline, $\boldsymbol{u} \times \boldsymbol{\omega} = 0$. The result (9.17) is known as *Bernoulli's theorem*.

9.3 Some Applications of Bernoulli's Theorem

9.3.1 Venturi Flow

The Venturi effect is a simple application of Bernoulli's theorem to streamline flow through a tube with a constriction (Fig. 9.4). By conservation of mass, the fluid must speed up on passing through the constriction. Since Bernoulli's equation states that

$$\tfrac{1}{2}u^2 + \frac{p}{\rho} = \text{constant}, \tag{9.18}$$

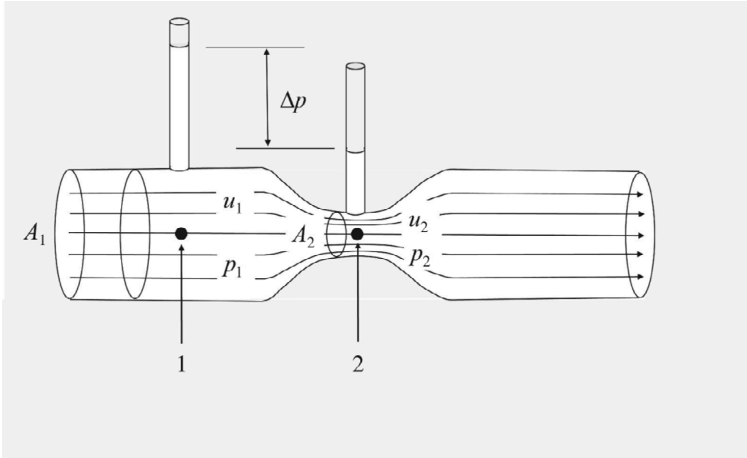


Fig. 9.4 Illustrating the Venturi effect. The pressure at location 2 is less than that at location 1, as measured by the Venturi meter which measures the difference in pressure at 1 and 2 Δp. (Courtesy of the Creative Commons Attribution-ShareAlike License.)

it follows that the pressure at the beginning of the constriction must be *less* than the pressure in the main part of the tube. This is illustrated in Fig. 9.4 by the pressure gauges in the main tube and the constriction.

This is just one of many examples in which the behaviour of the fluid seems contrary to our intuition. In this case, the pressure difference must exist in order to accelerate the fluid down the narrow section of the tube. If the end of the constriction were open, an accelerated jet of fluid would be expelled from what would become a nozzle. In the case of a de Laval nozzle, the Venturi effect is extended to hot, high pressure gases on the upstream side with supersonic jets on the downstream side, resulting the power source for jet engines.

9.3.2 When Can Compressibility and Viscosity Be Ignored?

In some of the examples which follow, we consider the application of Bernoulli's theorem to the flow of gases and make the apparently startling approximation that the gas can be considered incompressible and further that viscosity can be ignored. Following Faber, we begin with the case of flow through a constriction (Fig. 9.4). Let us take the limit in which the pressure difference is large and so we can approximate

$$p_1 - p_2 = \Delta p = \tfrac{1}{2}\,\rho U^2, \tag{9.19}$$

where U is the speed of the fluid downstream of the constriction. What is the change in density of the fluid in the case of a gas?

We introduce the *compressibility* β of the fluid by the definition

$$\beta = -\frac{1}{V}\frac{\partial V}{\partial p}. \tag{9.20}$$

Then, the change in density of the fluid is $\Delta\rho/\rho = \beta\,\Delta p$. This expression is valid for any fluid, but let us consider the case of an ideal gas with ratio of specific heat capacities $c_p/c_V = \gamma$. The compressibility becomes $\beta = 1/\gamma\rho$. Therefore,

$$\frac{\Delta\rho}{\rho} = \tfrac{1}{2}\frac{1}{\gamma p}\,\rho U^2. \tag{9.21}$$

But the speed of sound is $c_{\rm s} = (\gamma p/\rho)^{1/2}$ and so to order of magnitude

$$\frac{\Delta\rho}{\rho} \sim \left(\frac{U}{c_{\rm s}}\right)^2 = \mathcal{M}^2, \tag{9.22}$$

where $\mathcal{M}$ is the *Mach number*. Thus, for highly subsonic flow $\mathcal{M} \ll 1$, the change of the density of the gas is very small and very much less than the change in pressure or velocity of the flow along the streamlines. Faber recommends that the incompressible approximation can be used for flow speeds less that one-fifth of the sound speed, which is the case in many important gas dynamic problems.

What about viscosity? We will deal with this topic in Section 9.5 and in the next chapter but let us introduce the concept of the *Reynolds number* Re from the perspective of the present considerations. The simplest way of introducing the Reynolds number is to compare the pressure force per unit area $p_1 - p_2 = \Delta p = \frac{1}{2}\,\rho U^2$ in (9.19) with the viscous force per unit area in the presence of a shear velocity gradient $f = \eta\,\partial u/\partial z$, where η is the coefficient of *shear viscosity*. The ratio of these forces is to order of magnitude

$$\frac{\frac{1}{2}\,\rho U^2}{\eta U/L} \sim \frac{\rho L U}{\eta} = \frac{LU}{\nu} = \mathrm{Re}, \tag{9.23}$$

where L is the scale over which the velocity gradient takes place and $\nu = \eta/\rho$ is known as the *kinematic viscosity*. It is evident that if the Reynolds number is much greater than 1, viscosity can be neglected, but if less than 1, the effects of viscosity become dominant. As Faber emphasises, there are subtleties in the application of this simple relation concerning the role of viscosity in fluid flow, but the general rule that viscous forces can be neglected at high Reynolds numbers, $\mathrm{Re} \geq 100$, is adequate for our immediate purposes. The values of the Mach and Reynolds numbers need to be estimated for any given problem at the outset. One key aspect of fluid dynamics where viscosity plays an essential role is in *boundary layers*. These effects concern the transmission of stresses to the bulk of the material through 'friction' between the walls and the fluid. One way of thinking about the role of viscosity in this case is to note that the thickness of the boundary layer, L, is often very small indeed compared with the scale of the system, and so from (9.23) the Reynolds number becomes small and viscous forces need to be included in the analysis.

9.3.3 Vena Contracta, Helicopters and Windmills

The neglect of compressibility and viscosity is a great simplification in obtaining simple solutions to fluid dynamical problems. One of the useful tools in developing some of these results is the concept of the *vena contracta*, meaning the 'contracted vein'. It is illustrated

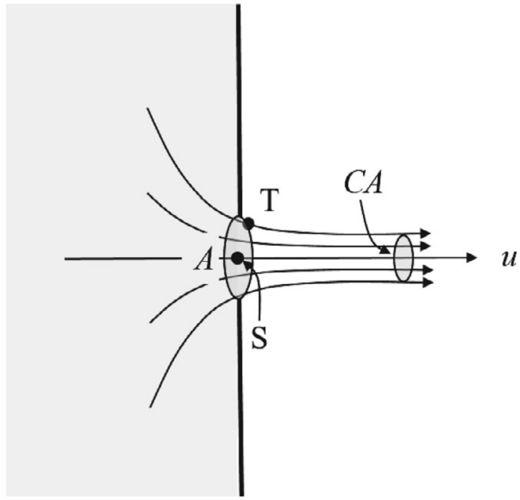


Fig. 9.5 Illustrating the vena contracta for the outflow of fluid from an orifice. The area of the orifice A is greater than the area CA of the jet where $C < 1$. (After T.E. Faber 1995.)

for the discharge of fluid from a small circular orifice of area A in the side of a container in Fig. 9.5, resulting in the formation of a horizontal jet. Over a short distance beyond the orifice, the streamlines narrow because they are curved as the fluid approaches and passes through the orifice. There is therefore a transverse pressure force acting on the fluid leaving the orifice such that the pressure at S is greater than the atmospheric pressure at T. There is therefore a *contraction coefficient* or *efflux coefficient* C which describes the area of the jet relative to the area of the orifice when the jet has attained its steady state away from the orifice itself. The area of the vena contracta is therefore CA, where A is the area of the orifice. For large Reynolds numbers, C is normally about 0.62. This is an example of a free jet.

The combination of Bernoulli's equation and the vena contracta enable us to understand the basic features of helicopters and windmills. In the analysis which follows, we adopt Faber's careful analysis in which he discusses the various approximations and subtleties involved in real examples of these machines.

In the case of the idealised helicopter, the function of the helicopter blades is to draw air from above and discharge it below (Fig. 9.6). The air flow is subsonic on either side of the blades and so the air can be taken to be incompressible. Because the air is incompressible, the air flow speed just above and below the blades at P and Q respectively must be the same. Work is done, however, by the blades and so the pressure at Q is greater than that at P. However, the air will not be drawn into the blades unless the pressure just above the blades is less than that far away from them, which will be atmospheric pressure. Furthermore, if the pressure at Q is greater than atmospheric pressure, the jet will be accelerated below the blades and narrow to form a vena contracta as illustrated in Fig. 9.6. In other words, the function of the blades is to decrease the pressure above the blades and increase it below them, both relative to atmospheric pressure at a large distance from the helicopter.

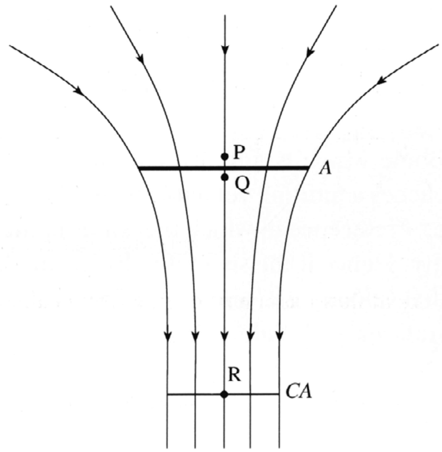


Fig. 9.6 The lines of flow above and below the rotor blades of an idealised helicopter. (After T.E. Faber 1995.)

Let us apply Bernoulli's theorem along the central flow line from atmospheric pressure far from the blades to the point P where the air flow has velocity u_P. Then,

$$p_P + \tfrac{1}{2}\,\rho u_P^2 = p_A. \tag{9.24}$$

Carrying out the same calculation from the central flow line at Q to the point R downstream, we find

$$p_Q + \tfrac{1}{2}\,\rho u_Q^2 = p_Q + \tfrac{1}{2}\,\rho u_P^2 = p_A + \tfrac{1}{2}\rho u_R^2, \tag{9.25}$$

where u_R is the downstream velocity at R. Subtracting (9.24) from (9.25), we find

$$p_Q - p_R = \tfrac{1}{2}u_R^2. \tag{9.26}$$

In a simple approximation, we take the pressure difference to apply across the whole surface area A swept out by the blades and that the velocity u_R is the same across the vena contracta at R. Thus, equating the upthrust on the helicopter $(p_Q - p_P)A$ to the downward momentum given to the jet per unit time, $\rho U_R^2 CA$,

$$p_Q - p_P = \tfrac{1}{2}u_R^2 A \;= \rho u_R^2 CA, \tag{9.27}$$

where C is the contraction factor and the area of the vena contracta is CA. Therefore $C = \frac{1}{2}$.

A similar example is the operation of a windmill, where we now seek to estimate how much wind power can be converted into the rotational motion of the sails. The idealised windmill is shown schematically in Fig. 9.7 with the upstream wind velocity U before it encounters the blades of the windmill. Having extracted power from the wind, its downstream velocity is only a fraction f of the upstream velocity. Because of conservation of mass of the incompressible air, the slower moving wind downstream expands to an area A/f. The resulting pattern for streamlines is shown in Fig. 9.7, just like the helicopter, but in reverse.

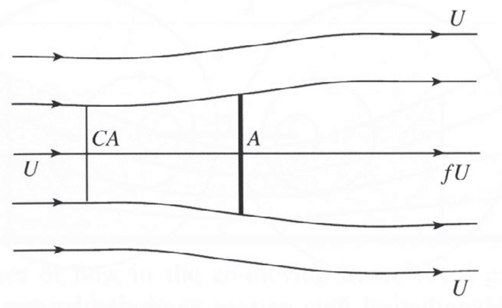


Fig. 9.7 The lines of flow passing through and around an idealised windmill. (After T.E. Faber 1995.)

The approximate analysis follows the same logic as in the case of the helicopter. We consider the central streamline on both sides of the windmill sails. Considering first the upstream side, the pressure at infinity is p_0 and at the point equivalent to P in Fig. 9.6 the pressure is p_1. Therefore, applying Bernoulli's equation,

$$p_1 + \tfrac{1}{2}\rho U^2 = p_0. \tag{9.28}$$

From the equivalent of point Q to infinity on the downstream side, Bernoulli's theorem leads to

$$p_2 + \tfrac{1}{2}\rho f^2 U^2 = p_0. \tag{9.29}$$

Subtracting, the net force acting on the windmill is

$$(p_1 - p_2)A = \tfrac{1}{2}\rho U^2(1 - f^2)A. \tag{9.30}$$

But this must equal the rate of loss of momentum by the 'jet', which is ρUCA per unit volume on the upstream side and $\rho UCAf$ on the downstream side. The rate of loss of momentum is therefore $\rho U^2 CA(1-f)$. Equating this result to (9.30), the contraction factor $C = \frac{1}{2}(1 + f)$. Thus, the rate at which kinetic energy is extracted from the wind is

$$\tfrac{1}{2}\rho U^3 CA(1 - f^2) = \tfrac{1}{4}\rho U^3(1 - f^2)(1 + f). \tag{9.31}$$

There is a maximum amount of power which can be extracted from the wind, which is found by maximising the value of (9.31) by differentiating with respect to f. The value corresponds to $f = \frac{1}{3}$ and to a maximum power

$$P = \frac{8\rho AU^3}{27}. \tag{9.32}$$

This power can be compared with the power of the incident wind upon the sails, which is $\frac{1}{2}\rho AU^3$, and so the maximum efficiency of the idealised windmill in generating energy is 16/27 or 59%.

Faber emphasises the numerous approximations and simplifications involved in these arguments, but these idealised models demonstrate the basic physics behind the forces involved in the dynamics of incompressible fluids.

9.3.4 Potential Flow about a Sphere

A simple example is the potential flow of an incompressible fluid with zero viscosity past a solid sphere. We need to find the solution of Laplace's equation $\nabla^2\phi = 0$, subject to the boundary conditions that the radial component of velocity at the surface of the sphere is zero and the flow is uniform at infinity with velocity U in the x-direction uniform. It is natural to solve the problem in spherical polar coordinates, the sphere having radius a. The flow is symmetric about the x-direction and so there is no azimuthal component. Therefore, from the expression for the Laplacian in A5.5, Laplace's equation becomes

$$\nabla^2\phi = \frac{1}{r}\frac{\partial^2}{\partial r^2}(r\phi) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\phi}{\partial\theta}\right) = 0. \tag{9.33}$$

The general solution of this equation is found by separation of variables with the result

$$\phi = \sum_{n=0}^{\infty}\left[A_n r^n + B_n r^{-(n+1)}\right] P_n(\cos\theta), \tag{9.34}$$

where A_n and B_n are constants to be determined by the boundary conditions and P_n is the Legendre polynomial of degree n. Those we will need are $P_0(\cos\theta) = 1$, $P_1(\cos\theta) = \cos\theta$ and $P_2(\cos\theta) = \frac{3}{2}\cos^2\theta - \frac{1}{2}$.

The flow is uniform in the x-direction at $r = \infty$ and so the velocity potential $\phi = Ux = Ur\cos\theta$ satisfies this requirement. This suggests that we test the trial solution with $n = 1$ for the overall flow, that is,

$$\phi = \left[Ur + B_1 r^{-2}\right]\cos\theta. \tag{9.35}$$

At the radius of the sphere, $r = a$, there can be no radial component of the fluid velocity, meaning $\mathrm{d}\phi/\mathrm{d}r = 0$. Therefore,

$$U = \frac{2B_1}{r^3} \quad \text{and} \quad \phi = U\left(r + \frac{a^3}{2r^2}\right)\cos\theta. \tag{9.36}$$

Since this is a solution of Laplace's equation and it fits the boundary conditions, it is the required unique solution. To find the velocity field, we take the gradient of (9.36) in spherical polar coordinates using the expression for the gradient in Section A5.5 and find

$$\boldsymbol{u} = \nabla\phi = \left(\frac{\partial\phi}{\partial r}, \frac{1}{r}\frac{\partial\phi}{\partial\theta}, 0\right) = \left[U\cos\theta\left(1 - \frac{a^3}{r^3}\right), -U\sin\theta\left(1 + \frac{a^3}{2r^3}\right), 0\right]. \tag{9.37}$$

The solution for $\boldsymbol{u}$ shows that the radial velocity component is indeed zero at the surface of the sphere and that the tangential velocity at $\theta = \pi/2$ is $\frac{3}{2}U$, corresponding to the speed up of the flow associated with the compression of the field lines, as shown in Fig. 9.3. Remarkably, because of the absence of viscosity in this treatment, there is no net force on the sphere. This is because the pressure forces on the sphere are perfectly symmetrical on the upstream and downstream sides of the sphere. This is often referred to as *d'Alembert's paradox*. We need to incorporate the effects of viscosity, however small, to account for the fact that there is indeed a drag force on the sphere, as we discuss in Section 9.5

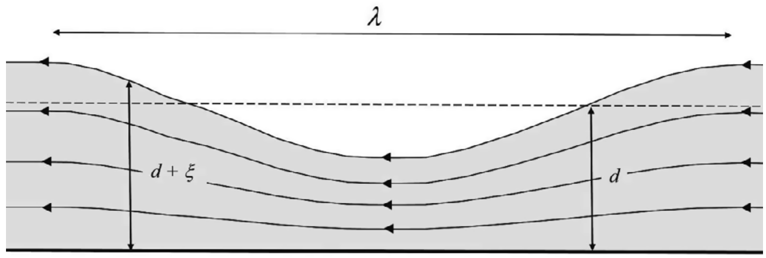


Fig. 9.8 The streamlines of flow in a shallow water gravity wave in the comoving frame of the wave. (After T.E. Faber 1995.)

9.4 Gravity Waves in Shallow and Deep Water

The dispersion relations and other properties of shallow and deep water gravity waves can be derived using Bernoulli's equation. First of all, we reproduce Faber's analysis of small amplitude waves on the surface of shallow water of depth d and wavelength $\lambda \gg d$. Surface tension and viscosity are taken to be zero so that surface effects can be neglected. It is convenient to work in the frame of reference moving with the speed of the waves, the streamlines in this frame being shown schematically in Fig. 9.8. Because of the compression and expansion of the streamlines, the flow velocity u is slightly greater than its mean value $\overline{u}$ in the wave troughs and slightly slower at the crests by an amount Δu. We can neglect the velocity change in the vertical direction since the wave amplitude ξ relative to Δu is very small. By conservation of mass, that is, using the continuity equation, the mass flux in the direction of u is a constant:

$$(d + \xi)\, u = (d + \xi)(\overline{u} + \Delta u) = \text{constant}. \tag{9.38}$$

Expanding this relation and retaining only first-order small quantities, we find

$$\Delta u\, d = -\xi\, \overline{u}. \tag{9.39}$$

At the surface of the water, the pressure is always atmospheric pressure and so is constant. Then, Bernoulli's equation (9.17) becomes

$$\tfrac{1}{2}u^2 + g(d + \xi) = \text{constant}. \tag{9.40}$$

Now expanding (9.40) to first order in small quantities, we find

$$\overline{u}\, \Delta u = -g\, \xi. \tag{9.41}$$

Dividing (9.39) by (9.41), we obtain

$$\overline{u} = \sqrt{gd}. \tag{9.42}$$

This is the speed at which the water drifts to the left in Fig. 9.8 in the frame of reference in which the wave is stationary. Transforming to the frame in which the flow is zero, it is evident that the speed of the wave to the right is $c = \sqrt{gd}$.

There are two interesting consequences of this solution. The first is that the phase velocity of the waves $c = \omega/k = \sqrt{gd}$ is a constant and so the waves are *non-dispersive*. The second is that this result explains why the wave crests of shallow water waves on a gently sloping beach are parallel to the shoreline. As the depth of water d becomes shallower, the speed of the waves decreases according to (9.42) and, just as in optics, the wave direction is refracted, the direction of the wave vector bending towards the gradient vector of the slope, as when light passes from air into water.

This behaviour is in contrast to the case of deep water surface waves, in which the depth of the water d is much greater than the wavelength of the waves, $|\xi| \ll \lambda \ll d$. To simplify the analysis, it is again assumed that surface tension and viscosity can be neglected. Since the water has no vorticity, its motion is determined by Laplace's equation. We consider waves propagating on the water surface in the positive x-direction and in the z-direction perpendicular to the undisturbed surface at $z = 0$ and the solid underwater seabed at $z = -d$. Thus, we need to solve Laplace's equation in two dimensions:

$$\nabla^2\phi = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial z^2} = 0, \tag{9.43}$$

subject to the boundary conditions. Finding the general form of the solution is straightforward. Adopting a trial solution of the form

$$\phi \propto \mathrm{e}^{\mathrm{i}k_x x}\,\mathrm{e}^{\mathrm{i}k_z z}, \tag{9.44}$$

where k_x and k_z are complex numbers, it follows that

$$k_x^2 + k_z^2 = 0. \tag{9.45}$$

We seek small amplitude wave solution for waves travelling in the positive x-direction and so k_x has to be real and k_z imaginary. We can therefore write $k_x = k$ and $k_z = \pm\mathrm{i}k$.

Now, we tackle the boundary conditions. At the flat seabed, the vertical velocity of the fluid must be zero:

$$u_z(-h) = \frac{\partial\phi}{\partial z} = 0. \tag{9.46}$$

The general form of solution for ϕ from (9.44) can be written

$$\phi = \mathrm{e}^{\mathrm{i}kx}\,\left(A\,\mathrm{e}^{kz} + B\,\mathrm{e}^{-kz}\right). \tag{9.47}$$

The derivative of ϕ with respect to z must be zero at $z = -h$ and so the expression for ϕ becomes

$$\phi = \phi_0\,\mathrm{e}^{\mathrm{i}kx}\,\left[\cosh k\,(\xi + h)\right]\mathrm{e}^{-\mathrm{i}\omega t}. \tag{9.48}$$

where we have now included the time dependence of the wave solution with the term $\mathrm{e}^{-\mathrm{i}\omega t}$.

At the surface, we can apply Bernoulli's equation since the flow is irrotational but, because of the presence of the waves, the flow is no longer steady but time dependent. The equation of motion for such a flow is given by (9.15), omitting the vorticity term:

$$\frac{\partial\boldsymbol{u}}{\partial t} = \mathrm{grad}\left(\frac{\partial\phi}{\partial t}\right) = -\mathrm{grad}\left(\tfrac{1}{2}u^2 + \frac{p}{\rho} + \Phi\right), \tag{9.49}$$

where we have written the gravitational potential Φ to distinguish it from the velocity potential ϕ.

We seek the equation for first-order perturbations about the flat unperturbed surface $z = 0$. Inspecting the terms of Bernoulli's equation in the form (9.49),

$$\tfrac{1}{2}u^2 + \frac{p}{\rho} + \Phi + \frac{\partial \phi}{\partial t} = \text{constant}. \tag{9.50}$$

The first term in the perturbed velocity u is second order in small quantities and so is dropped. The surface pressure is atmospheric pressure and is a constant along the perturbed surface. The gravitational potential is $\Phi = g\xi$ and we can neglect the constant since we take the gradient to find the velocity. So, Bernoulli's equation becomes

$$g\xi + \frac{\partial \phi}{\partial t} = 0. \tag{9.51}$$

The vertical velocity component of the surface is

$$\frac{\partial \xi}{\partial t} = u_z = \frac{\partial \phi}{\partial z}. \tag{9.52}$$

Taking the partial derivative of (9.51) with respect to time, we substitute for $\partial\xi/\partial t$ in (9.52) and find the familiar 'diffusion' equation,

$$g\frac{\partial \phi}{\partial x} = -\frac{\partial^2 \phi}{\partial t^2}. \tag{9.53}$$

Inserting the solution (9.48) into (9.53) and setting $z = 0$, we find the dispersion relation, that is, the relation between angular frequency and wavelength, for the waves,

$$\omega^2 = gk \tanh(kh). \tag{9.54}$$

For deep water waves, $kh \gg 1$, $\tanh(kh)$ tends to 1 and so the dispersion relation becomes

$$\omega = \sqrt{gk}. \tag{9.55}$$

The waves are dispersive because (9.55) is not a linear relation between frequency and wavelength. The group velocity $u_{\mathrm{gr}} = \mathrm{d}\omega/\mathrm{d}k = \frac{1}{2}\sqrt{gk}$, so that the group velocity is half the phase velocity.

The solution (9.54) also applies for shallow water waves. In the limit $kh \ll 1$, $\tanh(kh) = kh$ and hence the dispersion relation becomes $\omega = \sqrt{gh}\,k$ and the phase and group velocities are $\sqrt{gh}$ as in (9.42).

The motion of the water in deep water waves can be derived using (9.48) since

$$u_x = \frac{\partial \phi}{\partial x}; \quad u_z = \frac{\partial \phi}{\partial z}. \tag{9.56}$$

Then, noting that $\xi \ll h$,

$$u_x = \frac{\partial \phi}{\partial x} = \mathrm{i}k\,\phi_0\,\mathrm{e}^{\mathrm{i}(kx-\omega t)}\cosh kh, \tag{9.57}$$

$$u_z = \frac{\partial \phi}{\partial z} = k\,\phi_0\,\mathrm{e}^{\mathrm{i}(kx-\omega t)}\sinh kh. \tag{9.58}$$

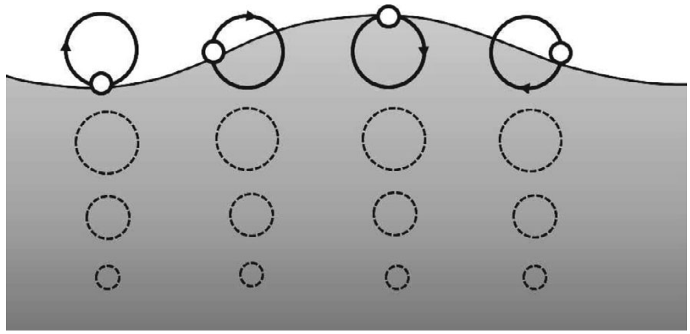


Fig. 9.9 The motions of water particles in deep water gravity waves. The surface waves are propagating to the right. (After image by Byron Inouye, Curriculum Research & Development Group (CRDG), College of Education. ©Ũniversity of Hawai'i.)

The key feature of these solutions is that the waves in the x and z directions are $\pi/2$ out of phase because of the imaginary unit i in (9.57) and so the motion of the water is, in general, ellipsoidal with ratio of amplitudes $\cosh kh/\sinh kh$ in the horizonal and vertical directions. In the limit $kh \gg 1$, $\cosh kh/\sinh kh = 1$ and so the water moves in circles, as shown in Fig. 9.9. They decrease exponentially in amplitude with depth into the water, the sinh function, $\sinh z = \frac{1}{2}(e^z - e^{-z})$, disguising the fact the amplitude of the wave decreases by a factor of 1/e in one wavelength with increasing depth. This is a common feature of wave solutions of diffusion equations of the form (9.53). Furthermore, as the depth of the water decreases towards zero, $\cosh kh \to 1$ and $\sinh kh \to 0$. As a result, the circles become ellipses and these become more and more elongated as the depth of the water decreases.

This is just the beginning of a long and intriguing story about the properties of waves in fluids. Any wave profile can be described by a Fourier series or Fourier integrals and then the evolution of the separate components followed through the linear and into the non-linear regime in which phenomena such breaking waves, hydraulic jumps, bores, solitons and so on are encountered. These topics are recommended for further reading.

9.5 The Equation of Motion of an Incompressible Fluid Including Viscous Forces

We have come almost as far as we can without writing down the stress and strain tensors for the fluid in order to derive the equations of fluid flow including the effects of viscosity. We will outline the derivation of the Navier–Stokes equation and indicate some of the subtleties involved.

The aim is to work out the viscous force per unit volume acting on a fluid with shear viscosity η. There are velocity gradients in the x, y, z directions which can vary throughout the fluid. First, we recall Newton's insight that the shear viscous force depends linearly

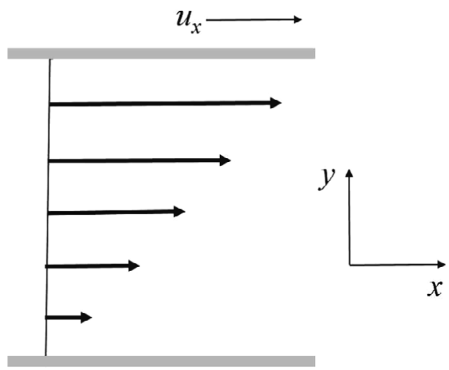


Fig. 9.10 Illustrating the action of viscous forces in unidirectional flow in the positive x-direction with a velocity gradient in the y-direction, represented by the bold arrows.

upon the velocity gradient. In the one-dimensional case shown in Fig. 9.10, there is a uniform gradient of the velocity u_x in the y-direction. Newton postulated that in this case there is a shear stress between adjacent layers of the flow. By *shear stress* we mean the force per unit area acting in a plane in the x-direction given by

$$s_{yx} = \eta \frac{\partial u_x}{\partial y}. \tag{9.59}$$

η is known as the *shear viscosity*.

We now work out the viscous forces acting on an element of the fluid of volume $\mathrm{d}V = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$ (Fig. 9.10). The viscous force acting on the bottom surface of the volume at y is given by (9.59). The force on the top surface is $\eta\,\mathrm{d}x\,\mathrm{d}z\,\partial u_x(y + \mathrm{d}y)/\partial y$. The net force on the element of fluid is the difference between these forces. Performing a Taylor expansion,

$$\eta\,\mathrm{d}x\,\mathrm{d}z \frac{\partial u_x(y + \mathrm{d}y)}{\partial y} = \eta\,\mathrm{d}x\,\mathrm{d}z \left(\frac{\partial u_x(y)}{\partial y} + \frac{\partial^2 u_x(y)}{\partial y^2}\,\mathrm{d}y \right).$$

Therefore, the net force acting on the volume V due to the velocity gradient in the y-direction is

$$\eta\,\mathrm{d}x\,\mathrm{d}z \frac{\partial^2 u_x}{\partial y^2}\,\mathrm{d}y = \eta \frac{\partial^2 u_x}{\partial y^2}\,\mathrm{d}V. \tag{9.60}$$

The equation of motion of the element of fluid is therefore

$$\rho\,\mathrm{d}V \frac{\mathrm{d}u_x}{\mathrm{d}t} = \eta \left(\frac{\partial^2 u_x}{\partial y^2} \right) \mathrm{d}V; \qquad \frac{\mathrm{d}u_x}{\mathrm{d}t} = \frac{\eta}{\rho} \left(\frac{\partial^2 u_x}{\partial y^2} \right).$$

This simple formula is suggestive of the form of the viscous term $\mu/\rho\,\nabla^2 \boldsymbol{u}$, which will appear in the full version of the analysis in three dimensions, the *Navier–Stokes equation*.

This statement, however, disguises a number of key aspects in the physics of fluid flow in the presence of viscosity. First of all, in general, the forces acting on an elementary volume of fluid with sides Δx, Δy, Δz are described by a 3×3 tensor. In the simplest case,

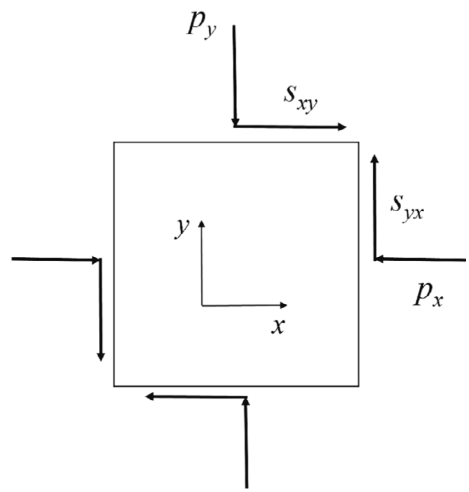


Fig. 9.11 Illustrating the action of viscous forces in two dimensions. The unlabelled vectors represent the balancing pressures and stresses in the fluid. (After T.E. Faber 1995.)

each face of the cube is subject to a normal pressure and two orthogonal shear stresses in the plane perpendicular to the normal to the surface. In general, the stress tensor s_{ij} has the form

$$s_{ij} = \begin{bmatrix} \tau_{11} & \tau_{12} & \tau_{13} \\ \tau_{21} & \tau_{22} & \tau_{23} \\ \tau_{31} & \tau_{32} & \tau_{33} \end{bmatrix}. \tag{9.61}$$

At the microscopic level, as the size d of the elementary volume tends to zero, it is essential that, even in the presence of velocity gradients, there is pressure balance across the volume – otherwise, the volume would suffer an infinite acceleration as $d \to 0$ (Fig. 9.11). In the same way, τ_{12} must equal τ_{21} or the elementary volume element would suffer an infinite angular acceleration as $d \to 0$. We avoid this situation by symmetrising the expression for the shear stress associated with velocity gradients in the x–y plane by writing the components of the stress tensor:

$$\tau_{12} = \tau_{21} = \eta \left(\frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x} \right). \tag{9.62}$$

As a result, there are only six independent elements of the stress tensor, the three diagonal pressures and the three off-diagonal terms.

9.5.1 Couette Flow

Let us demonstrate how this works in the two-dimensional case for a viscous fluid confined between two concentric rotating cylinders, what is known as *Couette flow*, named after the French physicist Maurice Couette who invented this form of viscometer. The outer cylinder has angular velocity ω_b while the inner cylinder rotates at angular frequency ω_a. The fluid moves in circular streamlines between the cylinders and so it is convenient to convert the expression for the viscous shear stress from Cartesian to polar coordinates r, ϕ using the relations $x = r\cos\phi$ and $y = r\sin\phi$. Writing the components of the velocity in cylindrical coordinates,

$$u_x = u_r \cos\phi - u_\phi \sin\phi; \quad u_y = u_r \sin\phi + u_\phi \cos\phi, \tag{9.63}$$

where u_r is the velocity in the radial direction and u_ϕ the circular velocity, which varies with radius r. The differentials in (9.62) can be converted into polar coordinates through the relations

$$\frac{\partial}{\partial x} = \cos\phi \frac{\partial}{\partial r} - \frac{\sin\phi}{r}\frac{\partial}{\partial \phi}; \quad \frac{\partial}{\partial y} = \sin\phi \frac{\partial}{\partial r} + \frac{\cos\phi}{r}\frac{\partial}{\partial \phi}. \tag{9.64}$$

Because of the cylindrical symmetry of the problem, it is convenient to evaluate the differentials at $\phi = 0$ in which case

$$\tau_{12} = \tau_{21} = \eta \left[\frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x}\right] = \frac{1}{r}\left(\frac{\partial u_x}{\partial \phi}\right)_{\phi=0} + \left(\frac{\partial u_y}{\partial r}\right)_{\phi=0}, \tag{9.65}$$

$$\frac{1}{r}\left(\frac{\partial u_x}{\partial \phi}\right)_{\phi=0} = -\frac{u_\phi}{r}; \quad \left(\frac{\partial u_y}{\partial r}\right)_{\phi=0} = \frac{\partial u_\phi}{\partial r}. \tag{9.66}$$

Hence,

$$f = \eta\left(-\frac{u_\phi}{r} + \frac{\partial u_\phi}{\partial r}\right) = \eta r \frac{\partial}{\partial r}\left(\frac{u_\phi}{r}\right) = \eta r \frac{\partial \omega}{\partial r}. \tag{9.67}$$

Thus, there is no shear stress if the fluid rotates as a solid body, $\omega(r) =$ constant, which makes sense.

Next, we work out the torque acting on an annular cylinder of the fluid of thickness dr and unit depth at radius r from the centre (Fig. 9.12). The torque acting on the inner edge of the annulus at r is

$$G = \eta r^2 (2\pi r)\frac{\partial \omega}{\partial r}. \tag{9.68}$$

The torque on the outer edge of the annulus is

$$G(r + \mathrm{d}r) = G(r) + \frac{\partial G}{\partial r}\,\mathrm{d}r, \tag{9.69}$$

and so the net torque acting on the annulus is the difference of these torques $(\partial G/\partial r)\,\mathrm{d}r$. The equation of motion of the annulus is therefore

$$\frac{\partial \mathcal{L}}{\partial t} = \frac{\partial G}{\partial r}\,\mathrm{d}r = \frac{\partial}{\partial r}\left[\eta r^2 (2\pi r)\frac{\partial \omega}{\partial r}\right]\mathrm{d}r, \tag{9.70}$$

where $\mathcal{L} = 2\pi r^2 \rho u_\phi \,\mathrm{d}r$ is the angular momentum of the annulus. Hence,

$$r^2 \frac{\partial u_\phi}{\partial t} = \frac{\eta}{\rho}\frac{\partial}{\partial r}\left[r^3 \frac{\partial \omega}{\partial r}\right]. \tag{9.71}$$

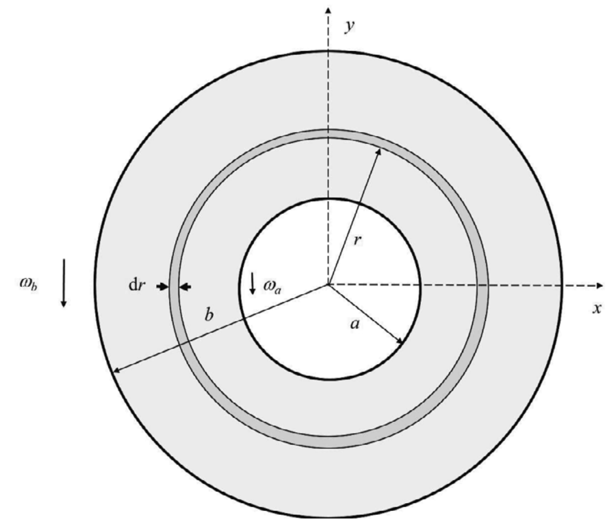


Fig. 9.12 Illustrating the forces acting on a cylindrical annulus in the flow of fluid between two rotating cylinders, what is referred to as Couette flow.

Writing $\nu = \eta/\rho$ where ν is the kinematic viscosity, the equation of motion for the circumferential velocity u_ϕ is

$$\frac{\partial u_\phi}{\partial t} = \frac{\nu}{r^2}\frac{\partial}{\partial r}\left[r^3\frac{\partial \omega}{\partial r}\right]. \tag{9.72}$$

This is the relation we have been seeking. We are interested in the steady state velocity distribution, in which case $\partial u_\phi/\partial t = 0$ and hence

$$r^3\frac{\partial \omega}{\partial r} = A, \tag{9.73}$$

where A is a constant. Integrating, we find

$$\omega = -\frac{2A}{r^2} + B. \tag{9.74}$$

Fitting the boundary conditions that the angular velocities at a and b are ω_a and ω_b respectively, we find

$$A = \frac{2a^2b^2}{b^2 - a^2}(\omega_b - \omega_a); \qquad B = \frac{b^2\omega_b - a^2\omega_a}{b^2 - a^2}. \tag{9.75}$$

Finally, we find the torque G from (9.67) and (9.72),

$$G = \eta r^2(2\pi r l)\,\frac{\partial \omega}{\partial r} = \eta r^2(2\pi r l)\,\frac{A}{r^3} = \frac{4\pi\eta l a^2 b^2}{b^2 - a^2}(\omega_b - \omega_a), \tag{9.76}$$

where l is the length of the cylinders. Notice that the torque is a constant as a function of radius, which is as expected in the steady state.

There are two merits of this apparent digression in deriving the Navier–Stokes equation. First, we have demonstrated how the viscosity of fluids can be measured using a Couette viscometer. The inner cylinder is rotated at a constant angular velocity, while the outer cylinder is forced to remain stationary by a spring balance, which measures the magnitude of the torque G. Second, we have demonstrated the self-consistency of our definitions of the expressions for the off-diagonal terms of the stress tensor.

9.5.2 The Navier–Stokes Equation

Let us write the relation between the stress tensor and the resulting strains in a form applicable for all three coordinates:

$$\tau_{ij} = \eta \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right). \tag{9.77}$$

The force acting on an elementary volume is given by the derivative of the elements in the stress tensor across opposite faces, as illustrated by the simple calculation which led to (9.60). Therefore, the total force acting on unit volume moving with the flow is

$$\sum_j \frac{\partial \tau_{ij}}{\partial x_j} = \eta \sum_j \left(\frac{\partial^2 u_i}{\partial x_j\, \partial x_j} + \frac{\partial^2 u_j}{\partial x_i\, \partial x_j} \right). \tag{9.78}$$

It is convenient to rewrite this expression in vector notation as

$$\boldsymbol{f} = \eta \left[\nabla^2 \boldsymbol{u} + \nabla(\nabla \cdot \boldsymbol{u}) \right], \tag{9.79}$$

and so we can include this term directly into the force equation (9.13) to obtain

$$\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u} \cdot \mathrm{grad})\boldsymbol{u} = -\frac{1}{\rho}\, \mathrm{grad}\, p - \mathrm{grad}\, \phi + \frac{\eta}{\rho}\, \nabla^2 \boldsymbol{u} + \frac{\eta + \eta'}{\rho} \nabla(\nabla \cdot \boldsymbol{u}), \tag{9.80}$$

where an additional term associated with resistance of the fluid to volume change, what is referred to as the *bulk viscosity* η', has been included in front of the divergence term in (9.80). This is the term which gives rise to the attenuation of sound waves passing through the medium. This expression is therefore applicable to both compressible and incompressible flow. The most common version of the force equation assumes that the fluid is incompressible, in which case $\nabla \cdot \boldsymbol{u} = 0$ and the *Navier–Stokes equation* can be written

$$\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u} \cdot \mathrm{grad})\boldsymbol{u} = -\frac{1}{\rho}\, \mathrm{grad}\, p - \mathrm{grad}\, \phi + \frac{\eta}{\rho}\, \nabla^2 \boldsymbol{u}. \tag{9.81}$$

This simplified treatment of fluid flow in the presence of stresses disguises the fact that there are many subtleties in a full derivation of (9.81). These are treated in the standard textbooks, Faber providing a strongly physical interpretation of significance of the components of the stress tensor in different circumstances.[6]

For future use, let us now convert (9.81) into the equation for the evolution of the vorticity $\boldsymbol{\omega}$ of the flow. We use the vector identity

$$(\boldsymbol{u} \cdot \nabla)\boldsymbol{u} = (\nabla \times \boldsymbol{u}) \times \boldsymbol{u} + \tfrac{1}{2}\nabla(\boldsymbol{u} \cdot \boldsymbol{u}). \tag{9.82}$$

Substituting this expression for $(\boldsymbol{u}\cdot\nabla)\boldsymbol{u}$ into (9.81) and then taking the curl of that equation, we find

$$\frac{\partial \boldsymbol{\omega}}{\partial t} = \nabla \times (\boldsymbol{u} \times \boldsymbol{\omega}) + \frac{\eta}{\rho}\nabla^2 \boldsymbol{\omega}. \tag{9.83}$$

This expression will prove to be invaluable in understanding the dynamics of fluid flow.

9.5.3 Poiseuille Flow

The Navier–Stokes equation has few analytic solutions, but it can be solved in special cases. One of the easiest is the velocity distribution in a narrow tube of inner radius a, such as the needle of a syringe in which the Reynolds number $\mathrm{Re} = Lu\rho/\eta$ is very much less than 1. As discussed above and in Section 10.2.3, in this limit the flow is dominated by viscosity rather than by the inertia of the fluid. Inspection of (9.81) shows that, since the flow down the tube is steady and the velocity gradient is perpendicular to the flow, the left-hand side of the Navier–Stokes equation is zero. We also neglect the influence of the gravitational potential ϕ and so we need to solve

$$\frac{1}{\rho}\,\mathrm{grad}\, p = \frac{\eta}{\rho}\nabla^2 \boldsymbol{u}. \tag{9.84}$$

The pressure difference between the ends of the tube of length l is taken to be P and is uniform over the cross-section of the tube. We need the Laplacian in cylindrical polar coordinates in the radial direction (see Section A5.5) and so we solve the equation

$$\frac{P}{l} = \frac{\eta}{\rho}\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\left(r\,\frac{\mathrm{d}u(r)}{\mathrm{d}r}\right). \tag{9.85}$$

Integrating twice with respect to r, we find the radial velocity distribution

$$u(r) = \frac{\rho P}{4\eta l}\left(a^2 - r^2\right). \tag{9.86}$$

The velocity distribution is parabolic with maximum velocity along the centre of the tube and zero at its walls. The volume rate of flow through the tube is found by integrating over the area of the tube:

$$\text{volume rate of flow} = \int_0^a 2\pi r\,\mathrm{d}r\,\frac{\rho P}{4\eta l}\left(a^2 - r^2\right) = \frac{\pi P a^4}{8\eta l}. \tag{9.87}$$

This formula is known as *Poiseuille's law* or the *Hagen–Poiseuille law*. Poiseuille was a physician who studied the flow of blood through arteries while Hagen was a hydraulic engineer. This formula provides a simple means of estimating the viscosity of fluids in the small Reynolds number limit.

9.6 Stokes' Formula for Highly Viscous Flow

Stokes' famous formula for the drag on a sphere moving through a highly viscous medium is one of the classics of fluid dynamics, $f_{\mathrm{D}} = 6\pi\eta a U$. It looks so simple and, as we will show in Section 10.2.3, its form can be derived by dimensional analysis, but obtaining the constant in the formula involves significant calculation. Stokes' original paper is lengthy and he applied the formula to the damping of the oscillations of a spherical pendulum bob by viscous drag in air.[7] But its application is multifold, including the formula for the terminal velocity of oil drops in Millikan's renowned experiment and in sedimentation.

We recall the result that, in the absence of viscosity, there is no drag force on the sphere (Section 9.3.4). The presence of viscous friction changes the situation dramatically. We can see this from the simple argument presented by Thorne and Blandford.[8] We will show that the Navier–Stokes equation reduces to the form $\nabla P = \eta \nabla^2 \boldsymbol{u}$, where P is the pressure difference across the sphere. Then, to order of magnitude, the pressure difference is $P \sim \eta U/a$ and so the force acting on the sphere is of order $f_{\mathrm{D}} \sim Pa^2 \sim \eta U a$. The challenge is to find the constant in this relation. Thus, the effect of viscosity is to create a pressure difference across the sphere which causes the drag force – there is no drag when the viscosity is zero, in agreement with the potential flow argument in Section 9.3.4.

The Navier–Stokes equation (9.81) for an incompressible fluid in the limit of small Reynolds number steady flow reduces to

$$\mathrm{grad}\, p = \eta \, \nabla^2 \boldsymbol{u}, \tag{9.88}$$

where we have deleted the inertial term $(\boldsymbol{u} \cdot \mathrm{grad})\boldsymbol{u}$, which is very much less than the viscous term, and the gravitational potential term. The flow must be uniform with velocity U at large values of r, just as in the case of pure potential flow. The solution is to be found for zero radial velocity at the surface of the sphere, but now there is a circumferential component in the θ-direction associated with the 'no-slip' boundary condition at $r = a$. This is a key difference from the case of potential flow – now, the fluid velocity in the circumferential direction at $r = a$ must be zero, its derivative at the surface is finite and so gives rise to the viscous drag.

We will outline the procedure here, without seeking to derive the details of the proof. In the books by Landau and Lifshitz, Faber and Thorne and Blandford, considerable effort is expended in inspecting carefully the Navier–Stokes equation to determine which quantities need to be preserved to make the equation soluble and match the boundary conditions. These discussions indicate the approximations needed in even this relatively simple problem. It is interesting to note the different approaches taken to solving (9.88).

- Perhaps the most traditional is the paper by Stokes himself, who systematically sets about solving the problem in considerable generality and detail.
- Landau and Lifshitz determine the form of the solution to the problem by physical arguments, in particular, by requiring the solution involve 'dimensionally' homogeneous combinations of axial or polar vectors.
- Faber adopts a standard approach of seeking a suitable solution involving a principal integral and a complementary function for (9.88), but also uses physical arguments and adopts informed trial solutions to fit the boundary conditions.
- Thorne and Blandford use a combination of all these arguments to find the required solution.

The result of these calculations is a pair of equations for the velocity field about the sphere of the form,

$$u_r = U\cos\theta\left(1 - \frac{3a}{2r} + \frac{a^3}{2r^3}\right), \tag{9.89}$$

$$u_\theta = -U\sin\theta\left(1 - \frac{3a}{4r} - \frac{a^3}{4r^3}\right), \tag{9.90}$$

where u_r and u_θ are the components of the velocity field in the radial $\boldsymbol{i}_r$ and polar $\boldsymbol{i}_\theta$ directions. We can now determine the excess stress acting at the surface of the sphere as a function of polar coordinates. We quote the results given by Faber in spherical polar coordinates. In the radial direction, the force per unit area, that is the excess pressure p^* as a function of θ, is[9]

$$p_r^* = p^* - 2\eta\left(\frac{\partial u_r}{\partial r}\right)_{r=a} = -\frac{3\eta U\cos\theta}{2a}, \tag{9.91}$$

where we have used the result derived by Faber that

$$p^* = -\frac{3}{2r^2}\eta\cos\theta. \tag{9.92}$$

The force per unit area parallel to the surface of the sphere is given by the θr component of the stress tensor, namely,[10]

$$s_{\theta r} = \eta a\left[\frac{\partial}{\partial r}\left(\frac{u_\theta}{r}\right) + \frac{1}{a^2}\frac{\partial u_r}{\partial\theta}\right]_{r=a} = -\frac{3\eta U\sin\theta}{2a}. \tag{9.93}$$

Finally, we take the components of the force acting parallel to the direction of the flow so that the drag force $f_{\rm D}$ is

$$f_{\rm D} = p_r^*\cos\theta + s_{\theta r}\sin\theta = -\frac{3\eta U}{2a}. \tag{9.94}$$

Remarkably, it can be seen that the force per unit area over the surface of the sphere is the same over the whole surface and so the total drag force is

$$f_{\rm D} = 4\pi a^2\,\frac{3\eta U}{2a} = 6\pi\eta a U, \tag{9.95}$$

the famous Stokes' law. The pattern of streamlines in illustrated in Fig. 9.13.

It is intriguing to compare this result with that of potential flow, analysed in Section 9.3.4. As Faber points out, the differences are considerable. They originate from the assumption in the potential flow case that, with viscosity absent, the fluid behaves like a superfluid. There is no friction between the flow and the sphere and so, at the surface of the sphere, there is no restriction on the velocity of the fluid in the $\boldsymbol{i}_\theta$ direction. Because of the squashing of the streamlines in this solution, the velocity field in the equatorial plane at $\theta = \pi/2$ *decreases* to the asymptotic value U at large values of r. In contrast, in the viscous case, that velocity is zero at the surface of the sphere and *increases* towards the asymptotic value U at large r. The solution also shows that the perturbation to the flow

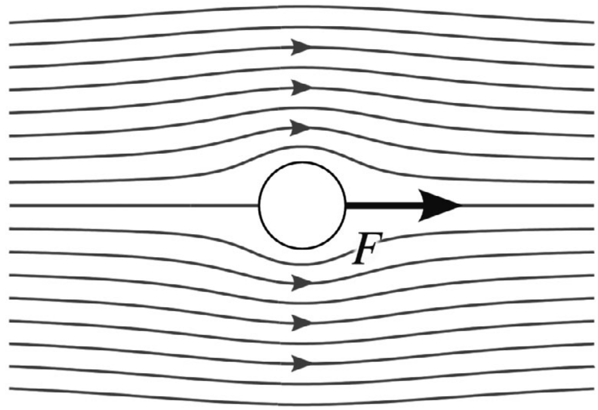


Fig. 9.13 Illustrating the streamlines associated with flow past a solid sphere of an incompressible fluid in the limit of small but finite viscosity. The pattern of streamlines should be compared with those in Fig. 9.3. (Courtesy of Creative Commons Attribution-Share Alike 3.0 Unported license.)

caused by the sphere is long range because of the slow decrease of the terms in powers of a/r with increasing r. Another way of thinking about this is that the boundary layer about the sphere is on the scale of the sphere itself and not confined to a narrow boundary region.

But this still does not explain why there is the apparent discontinuity in behaviour between a fluid with extremely small viscosity and one with none at all. As Feynman points out, we need to look at the values of the parameters included in the Reynolds number $\mathrm{Re} = \rho LU/\eta = LU/\nu$. The potential flow solution corresponds to $\eta = \nu = 0$, but in reality the viscosity remains very small but finite. Then, on scales $L \sim \nu/U$, the solution changes from potential flow to one dominated by viscosity. The problem is solved by the fact that the boundary layer through which viscosity determines the velocity gradient is very narrow indeed.

9.7 Vorticity, Circulation and Kelvin's Circulation Theorem

9.7.1 Vorticity

The concept of vorticity, $\boldsymbol{\omega} = \mathrm{curl}\,\boldsymbol{u}$, introduced in Section 9.2.2, is a central concept in fluid dynamics and gives physical insight into the properties of fluid flow. Let us consider first some examples of vorticity for the simple flows shown in Fig. 9.14.

- Case (a) shows rigid body rotation of the fluid, meaning that the velocity of rotation of the fluid is $\boldsymbol{u} = \boldsymbol{\Omega}\times\boldsymbol{r}$, where $\boldsymbol{\Omega} = \Omega_0\,\boldsymbol{i}_z$ is a constant angular velocity vector acting in the z direction out of the paper. In cylindrical polar coordinates (r, ϕ, z), $\boldsymbol{u} = (0, \Omega_0\, r, 0)$, and so using the expression for $\nabla\times\boldsymbol{u}$ given in Appendix 5.2.2, we find

$$\boldsymbol{\omega} = \nabla\times\boldsymbol{u} = \frac{1}{r}\frac{\partial(\Omega_0 r^2)}{\partial r}\,\boldsymbol{i}_z = 2\Omega_0\,\boldsymbol{i}_z, \tag{9.96}$$

that is, the vorticity is twice the angular velocity of the flow.

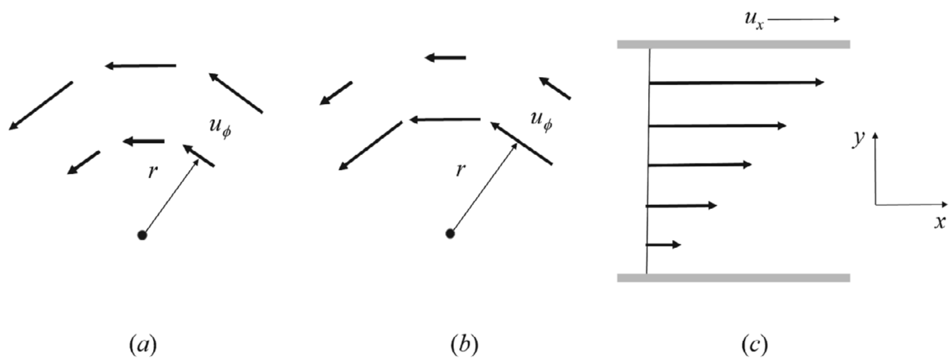


Fig. 9.14 Illustrative examples of vorticity in two-dimensional flows. (*a*) Constant angular velocity: $\boldsymbol{\Omega}$ is independent of radius $\boldsymbol{r}$. (*b*) Constant angular momentum per unit mass, $\boldsymbol{u} \times \boldsymbol{r}$ = constant. (*c*) Laminar shear flow with $u_x = ay$, as in Fig. 9.10. (After K.S. Thorne and R.D. Blandford, 2017, *Modern Classical Physics*, p. 732. Princeton NJ: Princeton University Press.)

- Case (b) illustrates cylindrical flow, in which the angular momentum per unit mass $\boldsymbol{j}$ is constant with radius, $\boldsymbol{j} = \boldsymbol{r} \times \boldsymbol{u} = |\boldsymbol{r}|\,|\boldsymbol{u}|\,\boldsymbol{i}_z$ = constant. The velocity vector is then $\boldsymbol{u} = (0, j/r, 0)$ and so the vorticity is

$$\boldsymbol{\omega} = \nabla \times \boldsymbol{u} = \frac{1}{r}\frac{\partial j}{\partial r}\,\boldsymbol{i}_z = 0. \tag{9.97}$$

  This is the type of flow associated with water flowing down a plug hole and with a tornado. Although this looks like a clear case of vortex flow, the vorticity is zero, except at $r = 0$ where the velocity diverges to infinity. The reason for zero vorticity is that, if two elements are separated by a distance x in the circumferential direction, their angular velocity is $+j/x^2$. If the elements are separated by x in the radial direction, their angular velocity is $-j/x^2$. Thus, these two angular motions cancel out. If however the integral is taken round a closed loop which includes the core of the vortex, the circulation is $2\pi j$ (see below).
- Case (c) shows laminar flow (Section 9.5) and corresponds to the one-dimensional variant of rigid body rotation discussed in (a). In this case $v_x \propto y$ and, taking the curl of the velocity in Cartesian coordinates, $\boldsymbol{\omega} = -(v_x/y)\,\boldsymbol{i}_z$. Thus, although the flow is laminar, it has a finite vorticity. We can visualise this as being associated with the rotation of a vector moving with the flow caused by the differential velocity field in the y-direction. Thus, a flow need not have 'rotation' in order to have a finite vorticity.

9.7.2 Circulation

Next, we introduce the concept of the *circulation* Γ of the flow, defined as the line integral of the velocity $\boldsymbol{u}$ around a closed curve C in the fluid:

$$\Gamma = \oint_C \boldsymbol{u} \cdot \mathrm{d}\boldsymbol{l}, \tag{9.98}$$

where $\mathrm{d}\boldsymbol{l}$ is the element of length of the closed loop. Using Stokes' theorem (Section A5.1), this definition can be converted into a surface integral for the flux of the vorticity through the surface S,

$$\Gamma = \int_S \boldsymbol{\omega} \cdot \mathrm{d}A. \tag{9.99}$$

We can now revisit case (*b*) above to determine the circulation for the differentially rotating fluid, but now taking the integral round the core of vortex. The circumferential velocity u_ϕ around the cylinder of radius r is $u_\phi = j/r$. Hence,

$$\Gamma = \oint_C \boldsymbol{u} \cdot \mathrm{d}\boldsymbol{l} = 2\pi r u_\phi = 2\pi j, \tag{9.100}$$

where j is the angular momentum per unit mass, which was taken to be a constant in this example. Formally, the vorticity diverges as $r \to 0$, but in fact at small enough scales, at which the viscosity cannot be neglected, the rotating fluid has a finite-sized core resulting in a finite value for the vorticity in that region.

9.7.3 Vortex Lines

We now introduce the concept of *vortex lines*. Let us recall the three expressions which describe the fluid flow in terms of the vorticity $\boldsymbol{\omega}$:

$$\nabla \cdot \boldsymbol{u} = 0, \tag{9.101}$$

$$\boldsymbol{\omega} = \nabla \times \boldsymbol{u}, \tag{9.102}$$

$$\frac{\partial \boldsymbol{\omega}}{\partial t} = \nabla \times (\boldsymbol{u} \times \boldsymbol{\omega}), \tag{9.103}$$

for incompressible flow with zero viscosity. By vortex lines, we mean field lines which track the vorticity vectors in the flow. The direction of the vortex lines is given by the direction of $\boldsymbol{\omega}$ and the magnitude of the vorticity is given by the number density of vortex lines per unit area normal to $\boldsymbol{\omega}$. It is important to distinguish between vortex lines and streamlines, the latter enabling us to visualise the velocity distribution $\boldsymbol{u}$. Specifically, the streamlines $\boldsymbol{u}$ associated with the vorticity are always perpendicular to $\boldsymbol{\omega}$, as illustrated by the examples in Section 9.7.1. The analogy with the magnetic field case is particulary helpful. Taking the curl of (9.101), $\mathrm{div}\,\boldsymbol{\omega} = 0$ and so, in this case, the vortex lines are similar to the lines of magnetic flux density in free space for which $\mathrm{div}\,\boldsymbol{B} = 0$.

The analogy with the dynamics of lines of magnetic flux density in a fully ionised plasma is striking. In that case, the equation for the dynamical behaviour of the magnetic flux density field $\boldsymbol{B}$ is given by the equation

$$\frac{\partial \boldsymbol{B}}{\partial t} = \nabla \times (\boldsymbol{v} \times \boldsymbol{B}) + \frac{1}{\sigma \mu_0} \nabla^2 \boldsymbol{B}. \tag{9.104}$$

The identity of the forms of this equation and (9.83) is striking. In the case of a fully ionised plasma corresponding to infinite electrical conductivity σ, the last term in (9.104) is zero. As discussed in detail in Section 11.2 of my book *High Energy Astrophysics*, equation (9.104) leads to the concept of *magnetic flux freezing* in which the magnetic field lines are 'frozen' into the plasma as it moves.[11] The magnetic field lines follow the motions in the plasma. Exactly the same phenomenon occurs in the case of the vortex lines in a fluid. In the absence of the diffusion term associated with the viscosity, the vortex lines are frozen into the fluid and move with it.

9.7.4 Kelvin's Circulation Theorem

Let us give a fluid dynamical proof of the theorem. Following Lautrup,[12] it is simplest to work in Lagrangian coordinates in which the derivatives are taken in the frame of reference which follows the flow. Specifically, C is a comoving closed curve. Then, we aim to show that the circulation obeys the rule

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}t} = \frac{\mathrm{d}\Gamma(C(t),t)}{\mathrm{d}t} = 0. \tag{9.105}$$

In other words, although the contour C is dragged along with and distorted by the fluid, the circulation Γ remains invariant.

In the small time interval δt, the change in circulation is

$$\delta\Gamma(C(t),t) = \Gamma(C(t+\delta t), t+\delta t) - \Gamma(C(t),t) \tag{9.106}$$

$$= \oint_{C(t+\delta t)} \boldsymbol{u}(\boldsymbol{r}', t+\delta t)\cdot \mathrm{d}\boldsymbol{l}' - \oint_{C(t)} \boldsymbol{u}(\boldsymbol{r},t)\cdot \mathrm{d}\boldsymbol{l}. \tag{9.107}$$

Now, we use the transformation $\boldsymbol{r}' = \boldsymbol{r} + \boldsymbol{u}(\boldsymbol{r},t)\,\delta t$ to map the curve $C(t)$ onto $C(t+\delta t)$. The line element moving with the flow becomes $\mathrm{d}\boldsymbol{l}' = \mathrm{d}\boldsymbol{l} + (\mathrm{d}\boldsymbol{l}\cdot\nabla)\boldsymbol{u}\,\delta t$ by applying (9.5) to the three components of the velocity vector $\boldsymbol{u}$:

$$\delta\Gamma(C(t),t) = \oint_{C(t)} \boldsymbol{u}(\boldsymbol{r} + \boldsymbol{u}(\boldsymbol{r},t)\,\delta t, t+\delta t)\cdot(\mathrm{d}\boldsymbol{l} + (\mathrm{d}\boldsymbol{l}\cdot\nabla)\boldsymbol{u}\,\delta t) - \oint_{C(t)} \boldsymbol{u}(\boldsymbol{r},t)\cdot\mathrm{d}\boldsymbol{l}. \tag{9.108}$$

Expanding to first order in δt and recalling that the calculation is carried out in the comoving frame, we find

$$\delta\Gamma(C(t),t) = \delta t \oint_{C(t)} \left(\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t}\right)\cdot\mathrm{d}\boldsymbol{l} + \delta t \oint_{C(t)} \boldsymbol{u}\cdot(\mathrm{d}\boldsymbol{l}\cdot\nabla)\boldsymbol{u}. \tag{9.109}$$

We now use Euler's theorem in the form (9.21) to replace the total derivative of $\boldsymbol{u}$ and rewrite the second term in more transparent form:

$$\delta\Gamma(C(t),t) = \delta t \oint_{C(t)} \left(-\nabla\phi - \frac{\nabla p}{\rho}\right)\cdot\mathrm{d}\boldsymbol{l} + \delta t \oint_{C(t)} \tfrac{1}{2}\nabla\left(\boldsymbol{u}^2\right)\cdot\mathrm{d}\boldsymbol{l} \tag{9.110}$$

$$= \delta t \oint_{C(t)} -\nabla\left(\phi + \frac{p}{\rho} - \tfrac{1}{2}\boldsymbol{u}^2\right)\cdot\mathrm{d}\boldsymbol{l} \tag{9.111}$$

$$= 0. \tag{9.112}$$

The last line follows from the fact that all three terms inside the integral (9.111) are now gradient fields and so their integrals around closed loops are zero, proving Kelvin's circulation theorem.

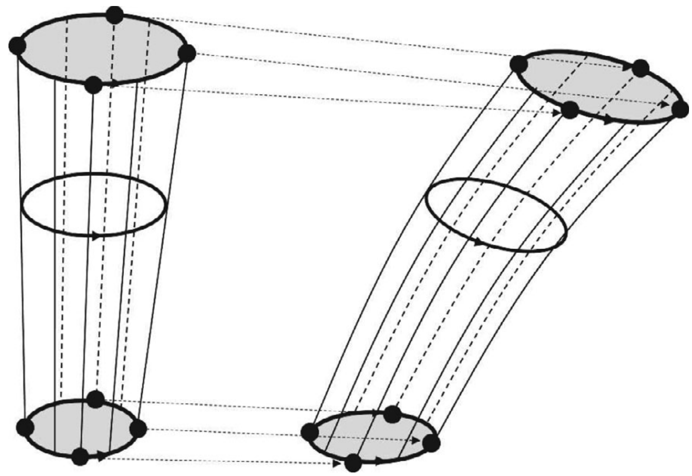


Fig. 9.15 Illustrating the motion of a vortex tube 'frozen' into the flow of in incompressible fluid in the absence of viscosity. The solid and dashed vertical lines are vortex lines defining the vortex tube. The dotted lines show the motion of particles of the fluid. The elliptical solid lines show the streamlines of the velocity field about the vortex tube.

9.7.5 Applications of Kelvin's Circulation Theorem

Kelvin's circulation theorem is a powerful tool for understanding fluid dynamical phenomena. The basic result is illustrated in Fig. 9.15, which shows a vortex tube being dragged along with the fluid. The solid and dashed lines within the vortex tube are vortex lines indicating the evolving direction and magnitude of the vorticity vectors $\boldsymbol{\omega}$. The ellipses show the streamlines of the motion of the fluid and, as expected, they are always normal to the vorticity vectors. The dotted lines show the motion of particles of the fluid, from which it can be seen that the vortex lines move as though they were 'frozen into the fluid'. Thus, within the constraints of subsonic motion and the absence of the effects of viscosity, vortex tubes are stable structures in fluid flow.

Perhaps the most familiar examples of these phenomena in nature are tornadoes, which are stable vortex tubes in the atmosphere and which can travel vast distances without losing their form and cause severe damage and destruction in their wakes. Another example is the stability of smoke rings, illustrated schematically in Fig. 9.16. The ring is a closed toroidal vortex tube, the dashed lines indicating the vortex lines and the elliptical solid lines the velocity vectors of the material of the smoke ring. Kelvin's theorem again accounts for the stability of the smoke ring. Finally, in the case of superfluid flow, in which the fluid has zero viscosity, the vortex tubes are very stable structures and play, for example, a key role in the physics of the interiors of neutron stars.

Eventually, these structures are damped out by the effects of viscosity, as indicated by the term $\eta\nabla^2\boldsymbol{\omega}$ in the full version of the Navier–Stokes equation (6.81). This term leads to the diffusion of the vortex lines and their dissipation through friction.

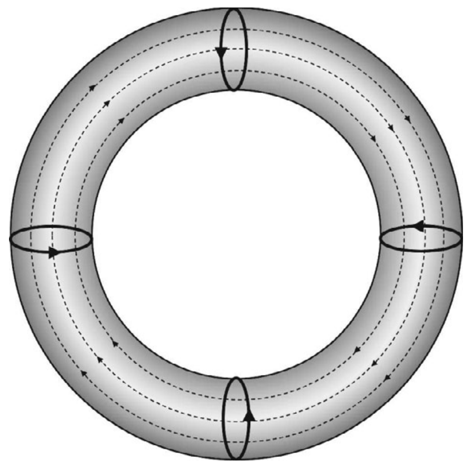


Fig. 9.16 Illustrating the fluid dynamical structure of a smoke ring. The solid ellipses represent the streamlines of the flow of the fluid. The dotted lines are the vortex lines associated with the toroidal structure of the smoke ring.

9.7.6 The Analogy between Vortex Tubes and Magnetic Flux Tubes

We have already discussed the profound impact these fluid dynamical phenomena had upon the understanding of electromagnetism during the nineteenth century in Section 5.3. As is apparent from the comparison of the equations (9.83) and (9.104), the analogy is between the magnetic flux density vector $\boldsymbol{B}$ and the vorticity vector $\boldsymbol{\omega}$:

$$\text{Vorticity :} \qquad \frac{\partial \boldsymbol{\omega}}{\partial t} = \nabla \times (\boldsymbol{u} \times \boldsymbol{\omega}) + \frac{\eta}{\rho} \nabla^2 \boldsymbol{\omega},$$

$$\text{Magnetic flux density :} \qquad \frac{\partial \boldsymbol{B}}{\partial t} = \nabla \times (\boldsymbol{v} \times \boldsymbol{B}) + \frac{1}{\sigma \mu_0} \nabla^2 \boldsymbol{B}.$$

Notice that it is the vortex lines and not the streamlines which provide the analogy between fluid flow and magnetism. These analogies were very important for William Thomson (later Lord Kelvin) and Maxwell in formulating the theory of the electromagnetic field.

It is intriguing that Kelvin and Maxwell adopted vortex models to account for the stability of atoms and molecules. In the mid-nineteenth century, there was no fundamental constant of nature which could determine the size of atoms – Planck's constant h only appeared in 1900. Their thinking was based upon the idea that atoms might be stable vortices in the frictionless aether, the medium through which it was assumed electromagnetic phenomena were propagated. Maxwell used this picture to create his model of a triatomic molecule consisting of three vortices representing the three atoms of the molecule in stable motion about each other. To produce a 'motion picture' of the model, Maxwell created a zoetrope strip of twelve images which showed the interpenetrating vortices moving in stable trajectories about each other (Fig. 9.17).

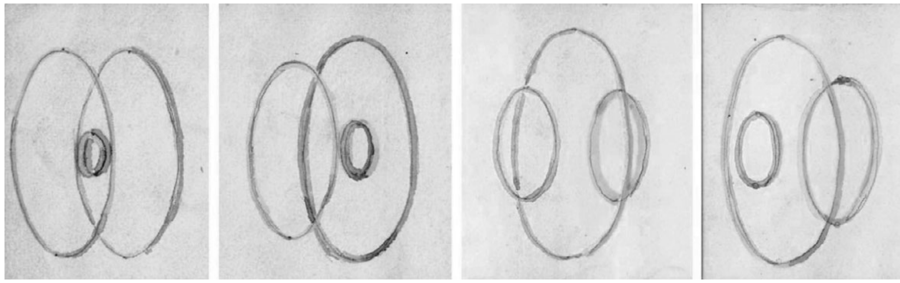


Fig. 9.17 Maxwell's model for a triatomic molecule consisting of three frictionless interacting vortices in the aether. These are four of twelve images which Maxwell drew as a zoetrope strip to illustrate their stable motion about each other. (Courtesy of the Cavendish Laboratory, University of Cambridge.)

9.8 Concluding Remarks

This brief introduction to some of the more remarkable phenomena associated with the equations of fluid flow is intended to illustrate the subtlety, complexity and enormous range of applicability of the fluid dynamics. We have barely scratched the surface of the subject which is of vital practical importance for everyday life – the flight of birds and aircraft, atmospheric physics, meteorology, oceanography to mention but a few obvious applications. This is not the place to delve more deeply into phenomena such as the stability of fluid flow, the effects of temperature, which lead, for example, to the transport of energy by convection, the physics of boundary layers and so on. Faber's *Fluid Dynamics for Physicists* provides an excellent introduction to these topics.[4] We will return to some of these topics in Chapter 10 when we consider some highly non-linear examples of fluid flow.

Notes

1 Batchelor, G.K. (1967). *An Introduction to Fluid Dynamics*. Cambridge: Cambridge University Press.
2 Landau, L.D. and Lifshitz, E.M. (1959). *Fluid Mechanics*, Vol. 5 of *Course of Theoretical Physics*. Oxford: Pergamon Press.
3 Feynman, R.P. (1964). *The Feynman Lectures on Physics*, Vol. 2, eds. R. Feynman, R.B. Leighton and M.L. Sands. Redwood City, California: Addison-Wesley Publishing Company.
4 Faber, T.E. (1995). *Fluid Dynamics for Physicists*. Cambridge: Cambridge University Press.
5 Note the point that, in the case of the gravitational field of the Earth, for example, the gravitational potential is a negative quantity and becomes less negative with increasing height, which is why we can write $\phi = gz$.
6 Chapter 6 of Faber's book (*op. cit.*) gives a careful exposition of the subtleties behind the apparently innocuous term $(\eta/\rho)\,\nabla^2\boldsymbol{u}$ in the Navier–Stokes equation.
7 G.G. Stokes (1851). On the Effect of the Internal Friction of Fluids on the Motion of Pendulums, *Transactions of the Cambridge Philosophical Society*, **9**, 8–106. Also,

Mathematical and Physical Papers of George Gabriel Stokes, Vol. 3, pp. 1–141. Cambridge: Cambridge University Press (1901).

8 Thorne and Blandford's massive *Modern Classical Physics* (Princeton: Princeton University Press, 2017) can be strongly recommended. See especially pp. 746–754 for the derivation of Stokes' law.

9 See Faber Section 6.1, equation (6.11) for the expression for p_r^*.

10 See Faber Section 6.9, equation (6.53) for the expression for $s_{\theta r}$.

11 Although the end result is the same, it is interesting that the proof of magnetic flux freezing involves considerations of the change in flux due to changes in the current within the loop and the effects of magnetic induction due to the motion of the loop.

12 B. Lautrup (2011). *Physics of Continuous Matter*. Boca Raton, Florida: CRC Press.

10 Dimensional Analysis, Chaos and Self-Organised Criticality

10.1 Introduction

The increasingly powerful mathematical to ols described in Chapter 8 provided the means for tackling complex dynamical problems in classical physics. Despite these successes, in many areas of physics, problems can become rapidly very complex and, although we may be able to write down the differential or integral equations which describe the behaviour of the system, often it is not possible to find analytic solutions.

This raises an important point about the problems which physicists and theorists set out to tackle. There is an emphasis upon problems which are mathematically soluble and these generally involve linear equations. But most of nature is not like that. Two quotations by Hermann Bondi make the point rather nicely.

> If you walk along the street you will encounter a number of scientific problems. Of these, about 80 per cent are insoluble, while $19\frac{1}{2}$ per cent are trivial. There is then perhaps half a per cent where skill, persistence, courage, creativity and originality can make a difference. It is always the task of the academic to swim in that half a per cent, asking the questions through which some progress can be made.[1]

Later, he wrote:

> (Newton's) solution of the problem of motion in the solar system was so complete, so total, so precise, so stunning, that it was taken for generations as the model of what any decent theory should be like, not just in physics, but in all fields of human endeavour. . . .
>
> I regard this as profoundly misleading. In my view, most of science is not like the Newtonian solar system, but much more like weather forecasting.[2]

The point is that the physicists and theorists simplify problems to the point at which there is a reasonable chance of finding solutions and that has been the thrust of this text so far. While Bondi was undoubtedly correct when he was writing, there have been significant advances in tackling some of these really tough problems. In particular, the exponential increase in the power of computers has led to the development of a major branch of physics, computational physics, enabling us to obtain insight into what appeared quite intractable problems a decade or more ago. The objective of this chapter is to study some of the techniques developed to tackle these complex problems, some of them so non-linear that they seem quite beyond the scope of traditional analysis.

- First, we review the techniques of *dimensional analysis*. Used with care and insight, this approach is powerful and finds many applications in pure and applied physics. We will give as examples the non-linear pendulum, fluid flow, explosions, turbulence, and so on.

- Next, we briefly study *chaos*, the analysis of which only became feasible with the development of high-speed computers. The equations of motion are deterministic and yet the outcome is extremely sensitive to the precise initial conditions.
- Beyond these examples are even more extreme systems in which so many non-linear effects come into play that it is impossible to predict the outcome of any experiment in any conventional sense. And yet regularities are found in the form of scaling laws. There must be some underlying simplicity in the way in which the system behaves, despite the horrifying complexity of the many processes involved. These topics involve *fractals* and the burgeoning field of *self-organised criticality*.
- Finally, we look briefly at what Roger Penrose has termed *non-computational* physics. Are there problems which in principle we cannot solve by mathematical techniques, no matter how powerful our computers?

10.2 Dimensional Analysis

Ensuring that equations are dimensionally balanced is part of our basic training as physicists. Dimensional analysis is, however, much deeper than simply dimensional consistency and can help us solve complicated problems without having to solve complicated equations. Let us begin with a simple example and then extend it to illustrate the power of dimensional analysis.

10.2.1 The Simple Pendulum

As a traditional example of the 'method of dimensions', consider the small amplitude oscillations of a simple pendulum. Its period τ might depend upon the length of the string l, the mass of the bob m and the gravitational acceleration g. The period of the pendulum can therefore be written

$$\tau = f(m, l, g). \tag{10.1}$$

The expression (10.1) must be dimensionally balanced and so we write it in the form

$$\tau \sim m^{\alpha} l^{\beta} g^{\gamma}, \tag{10.2}$$

where α, β and γ are constants. We adopt the convention in this chapter that the units need not match if we use the proportional sign '$\propto$', but when we use the sign '$\sim$', the equation must be dimensionally balanced. In terms of dimensions, (10.1) reads

$$[\mathrm{T}] \equiv [\mathrm{M}]^{\alpha}[\mathrm{L}]^{\beta}[\mathrm{LT}^{-2}]^{\gamma}, \tag{10.3}$$

where [L], [T] and [M] mean the dimensions of length, time and mass respectively. The dimensions must be the same on either side of the equation and so, equating the powers of [T], [M] and [L], we find

$$
\begin{array}{lll}
[\mathrm{T}]: & 1 = -2\gamma & \gamma = -\tfrac{1}{2}, \\
[\mathrm{M}]: & 0 = \alpha & \alpha = 0, \\
[\mathrm{L}]: & 0 = \beta + \gamma & \beta = +\tfrac{1}{2}.
\end{array}
$$

Therefore,

$$\tau \sim \sqrt{l/g}, \tag{10.4}$$

the correct expression for the simple pendulum. Note that the period τ is independent of the mass of the pendulum bob, m. This simple application of dimensional methods is found in many textbooks, but it does scant justice to the real power of dimensional analysis in pure and applied physics.

Let us expand the scope of dimensional analysis. We introduce a remarkable theorem enunciated by Edgar Buckingham in 1914,[3] following some derogatory remarks about the method of dimensions by the distinguished theorist Richard Tolman. Buckingham made creative use of his theorem in understanding, for example, how to scale from model ships in ship tanks to the real thing. The procedure is as follows:

- First, guess what the important quantities in the problem are, as above.
- Then apply the *Buckingham* Π *theorem*, which states that a system described by n variables, built from r independent dimensions, is described by $(n - r)$ independent dimensionless groups.[4]
- Form from the important quantities all the possible dimensionless combinations – these are known as *dimensionless groups*.
- The most general solution to the problem can be written as a function of all the independent dimensionless groups.

Let us repeat the pendulum problem. As before, we guess what the independent variables are likely to be – Table 10.1 shows a list of the variables with their dimensions.

There are $n = 5$ variables and $r = 3$ independent dimensions and so, according to the Buckingham Π theorem, we can form two independent dimensionless groups. θ_0 is already dimensionless and so $\Pi_1 = \theta_0$. Only one variable depends upon m and so *no dimensionless group can contain m*, that is, the period of the pendulum is independent of the mass of the bob.

There are many possibilities for Π_2, for example, $(\tau^2 g/l)^2$, $\theta_0 l/\tau^2 g$, and so on. There are advantages, however, in making Π_2 independent of Π_1 and as simple as possible. Then, all

Table 10.1 The non-linear pendulum

Variable	*Dimensions*	*Description*
θ_0	–	angle of release
m	[M]	mass of bob
τ	[T]	period of pendulum
g	$[\mathrm{L}][\mathrm{T}]^{-2}$	acceleration due to gravity
l	[L]	length of pendulum

other possible dimensionless groups can be formed from these two independent examples, Π_1 and Π_2. The simplest independent group we can form is, by inspection,

$$\Pi_2 = \frac{\tau^2 g}{l}.$$

Therefore, the general solution for the motion of the pendulum is

$$f(\Pi_1, \Pi_2) = f\left(\theta_0, \frac{\tau^2 g}{l}\right) = 0, \tag{10.5}$$

where f is a function of θ_0 and $\tau^2 g/l$. Notice that we have not written down any differential equation.

Now, (10.5) might well turn out to be some complicated function, but we guess that it will often be possible to write the solution as some unknown functional dependence of one of the dimensionless groups upon the other. In this case, we can write (10.5) as follows:

$$\Pi_2 = f_1(\Pi_1) \qquad \text{or} \qquad \tau = 2\pi\, f_1(\theta_0)\sqrt{l/g}. \tag{10.6}$$

Now, we might not be able to determine $f_1(\theta_0)$ from theory, but we can determine *experimentally* how the period τ depends upon θ_0 for particular values of l and g. Then, $f(\theta_0)$ is determined and will be true for *all other simple pendulums*: $f_1(\theta_0)$ is *universal*. Notice that this solution is valid for *all* values of θ_0, and not just small values, $\theta_0 \ll 1$.

In this example, we can find the exact non-linear solution for an arbitrary value of θ_0. By conservation of energy,

$$mgl(1 - \cos\theta) + \tfrac{1}{2}I\dot{\theta}^2 = \text{constant} = mgl(1 - \cos\theta_0), \tag{10.7}$$

where $I = ml^2$ is the moment of inertia of the swinging pendulum and θ_0 is its maximum amplitude. Let us first recall the linear solution which is found for small values of θ_0. The equation (10.7) becomes

$$\ddot{\theta} + \frac{g}{l}\theta = 0,$$

and so the angular frequency and period of oscillations are

$$\omega = \sqrt{g/l}, \quad \tau = 2\pi/\omega = 2\pi\sqrt{l/g}.$$

Now returning to the exact differential equation (10.7), we can write

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \sqrt{\frac{2g}{l}}(\cos\theta - \cos\theta_0)^{1/2}. \tag{10.8}$$

The period of the pendulum is four times the time it takes to swing from 0 to θ_0 and so

$$\tau = 4\int_0^{t(\theta_0)} \mathrm{d}t = 4\sqrt{\frac{l}{2g}}\int_0^{\theta_0} \frac{\mathrm{d}\theta}{(\cos\theta - \cos\theta_0)^{1/2}}. \tag{10.9}$$

Notice that the integral depends only upon θ_0. Performing a numerical integration, the function $f_1(\theta_0)$ shown in Fig. 10.1 is obtained.

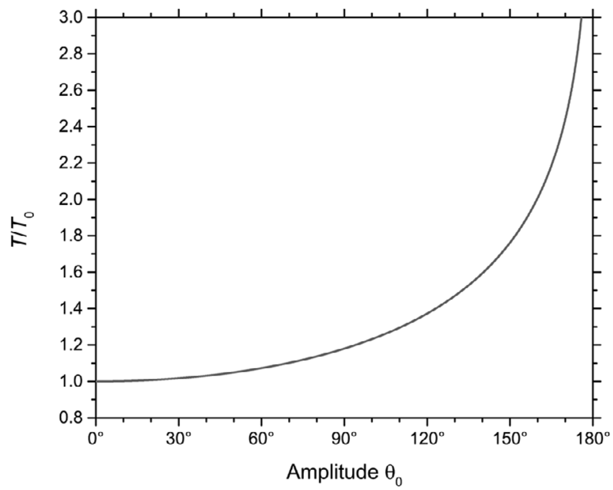


Fig. 10.1 The dependence of the period of a non-linear pendulum upon the angle of release θ_0 relative to the small-angle values of $f_1(\theta_0)$. (Courtesy of Dr. David G. Simpson, *General Physics I Lecture Notes*, http://pgccphy.net/1030/phy1030.pdf, page 285.)

We can therefore write

$$\tau = 2\pi\sqrt{l/g}\, f_1(\theta_0). \tag{10.10}$$

Just as predicted by dimensional analysis, the result only depends upon a general function of the dimensionless parameters θ_0 and $\tau^2 g/l$.

10.2.2 G.I. Taylor's Analysis of Explosions

In 1950, Geoffrey (G.I.) Taylor published his famous analysis of the dynamics of the shock waves associated with nuclear explosions. A huge amount of energy E is released within a small volume resulting in a strong spherical shock front which is propagated through the surrounding air. The internal pressure is enormous and very much greater than that of the surrounding air. The dynamics of the shock front depend, however, upon the density ρ_0 of the surrounding air which is swept up by the expanding shock front and causes its deceleration. The compression of the ambient gas by the shock front plays a role in its dynamics and depends upon the ratio of the specific heat capacities of air, $\gamma = 1.4$ for the molecular gases oxygen and nitrogen. The only other parameters in the problem are the radius of the shock front r_f and the time t. Hence, the table of variables and their dimensions is as shown in Table 10.2.

There are five variables and three independent dimensions and so, according to the Buckingham Π theorem, we can form two dimensionless groups. One of them is $\Pi_1 = \gamma$. The second dimensionless quantity is found by elimination. From the quotient of E and ρ_0, we find

$$\left[\frac{E}{\rho_0}\right] = \frac{[\mathrm{M}][\mathrm{L}]^2[\mathrm{T}]^{-2}}{[\mathrm{M}][\mathrm{L}]^{-3}} = \frac{[\mathrm{L}^5]}{[\mathrm{T}^2]} = \left[\frac{r_f^5}{t^2}\right].$$

Table 10.2 G.I. Taylor's analysis of explosions

Variable	*Dimensions*	*Description*
E	$[M][L]^2[T]^{-2}$	energy release
ρ_0	$[M][L]^{-3}$	external density
γ	–	ratio of specific heat capacities
r_f	$[L]$	shock front radius
t	$[T]$	time

Thus,

$$\Pi_2 = \frac{Et^2}{\rho_0 r_f^5}.$$

We can therefore write the solution to the problem as

$$\Pi_2 = f(\Pi_1), \quad \text{that is} \quad r_f = \left(\frac{E}{\rho_0}\right)^{1/5} t^{2/5} f(\gamma). \tag{10.11}$$

In 1941, Taylor was asked by the UK government to carry out a theoretical study of very high energy explosions. His report was presented to the UK Civil Defence Research Committee of the Ministry of Home Security in that year and was only declassified in 1949. In his detailed analysis of the expansion of the very strong shocks associated with such explosions, he derived (10.11) and showed that the constant $f(\gamma = 1.4) = 1.03$. He compared his results with high speed films of chemical explosions available at that time and found that his theoretical predictions were in reasonable agreement with what was observed.[5] In 1949, he compared his calculations with the results of higher energy explosions using TNT-RDX and again confirmed the validity of (10.11).[6] In 1947, the US military authorities released Mack's movie of the first atomic bomb explosion, which took place in the New Mexico desert in 1945, under the title *Semi-popular motion picture record of the Trinity explosion*[7] (Fig. 10.2(*a*)). Taylor was able to determine, directly from the movie, the relation between the radius of the shock wave and time (Fig. 10.2(*b*)). To the annoyance of the US government, he was then able to work out the energy released in the explosion and found it to be $E \approx 10^{14}$ J, equivalent to the explosion of about 20 kilotons of TNT.[6] This work was only published nine years after he had carried out his first calculations.

10.2.3 Fluid Dynamics: Drag in Fluids

We have already discussed in some detail analytic solutions for fluid flow about a sphere in the cases of zero viscosity (Section 9.3.4) and in the presence of a small, but finite, viscosity (Section 9.6), the latter resulting in Stokes' famous expression for the viscous drag force on a sphere. Let us now apply the techniques of dimensional analysis to incompressible fluid flow past a sphere of radius R in two limiting cases, one in which viscosity mediates the retarding force, the other in which the *inertia* of the fluid is dominant.

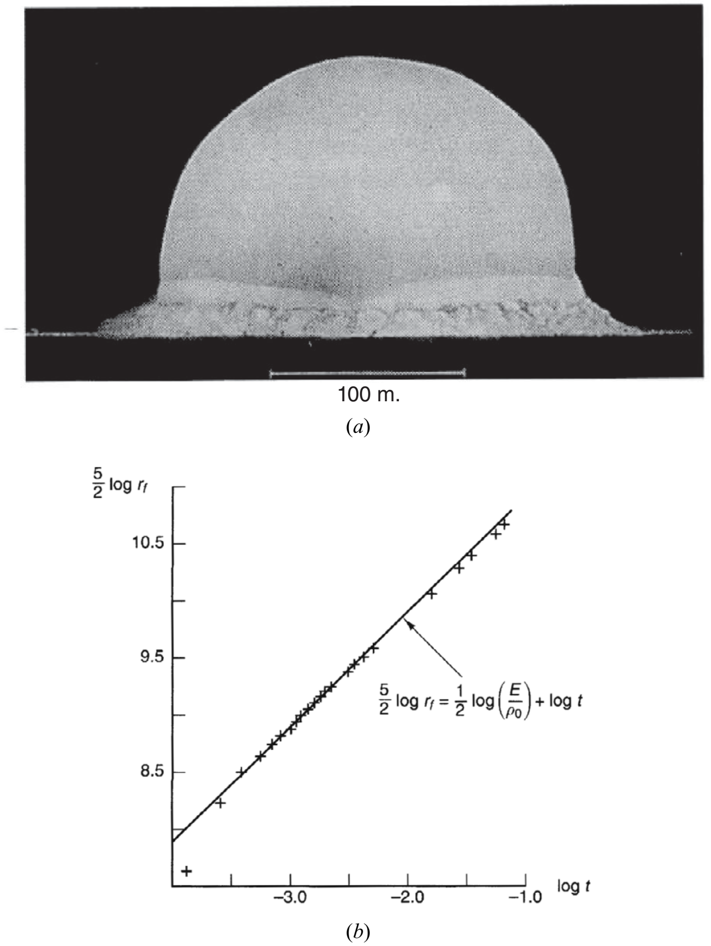


Fig. 10.2 (*a*) A frame from Mack's film of a Trinity nuclear explosion taken 15 ms after ignition. (*b*) G.I. Taylor's analysis[6] of the dynamics of the shock front of the nuclear explosion, showing that $r_f \propto t^{2/5}$.

Stokes' Law Revisited

The variables which are likely to be important in determining the drag force on a sphere moving steadily through a viscous fluid at velocity v are listed in Table 10.3. The mass of the sphere is not expected to appear in the expression for the drag force, which should only depend upon the surface area of the sphere.

According to the Buckingham Π theorem, there are five variables and three independent dimensions and so we can form two independent dimensionless groups.

The last three entries in Table 10.3 form one dimensionless group:

$$\Pi_1 = \frac{vR}{\nu} = \mathrm{Re},$$

Table 10.3 Parameters involved in estimating the drag force on a sphere in a viscous medium

Variable	*Dimensions*	*Description*
f_d	$[\mathrm{M}][\mathrm{L}][\mathrm{T}]^{-2}$	drag force
ρ_f	$[\mathrm{M}][\mathrm{L}]^{-3}$	density of fluid
R	$[\mathrm{L}]$	radius of sphere
ν	$[\mathrm{L}^2][\mathrm{T}]^{-1}$	kinematic viscosity
v	$[\mathrm{L}][\mathrm{T}]^{-1}$	velocity

where Re is the Reynolds number. The second dimensionless group involves the first two entries in the table:

$$\Pi_2 = \frac{f_\mathrm{d}}{\rho_\mathrm{f} R^2 v^2}.$$

The most general relation depends only on Π_1 and Π_2 and, since we want to find an expression for f_d, let us write

$$\Pi_2 = f(\Pi_1), \quad \text{that is,} \quad f_\mathrm{d} = \rho_\mathrm{f} R^2 v^2 f\left(\frac{vR}{\nu}\right).$$

Now, the drag force should be proportional to the kinematic viscosity ν and so the function $f(x)$ must be proportional to $1/x$. Therefore, the expression for the drag force is

$$f_\mathrm{d} = A \rho_\mathrm{f} R^2 v^2 \left(\frac{vR}{\nu}\right)^{-1} = A \nu \rho_\mathrm{f} R v, \tag{10.12}$$

where A is a constant to be found. We cannot do any better than this using dimensional techniques. The full treatment outlined in Section 9.6 results in the answer $A = 6\pi$, and so

$$f_\mathrm{d} = 6\pi \nu \rho_\mathrm{f} R v. \tag{10.13}$$

If the sphere is falling under gravity, we can find its terminal speed by equating the drag force to the gravitational force on the sphere, $f_\mathrm{g} = (4\pi/3)\rho_\mathrm{sp} R^3 g$, where ρ_sp is the mean density of the sphere. We find

$$(4\pi/3)\rho_\mathrm{sp} R^3 g = 6\pi \nu \rho_\mathrm{f} R v,$$

that is,

$$v = \frac{2}{9}\left(\frac{gR^2}{\nu}\right)\left(\frac{\rho_\mathrm{sp}}{\rho_\mathrm{f}}\right). \tag{10.14}$$

So far, we have ignored the effects of buoyancy on the terminal velocity of the sphere. This enters through another dimensionless group:

$$\Pi_3 = \frac{\rho_\mathrm{f}}{\rho_\mathrm{sp}}.$$

The effect of buoyancy is to reduce the effective weight of the sphere or, equivalently, the magnitude of the acceleration due to gravity so that

$$g \to g\left(1 - \frac{\rho_{\rm f}}{\rho_{\rm sp}}\right).$$

The reasoning behind this buoyancy correction is that the sphere would not fall under gravity if $\rho_{\rm f} = \rho_{\rm sp}$, but, if $\rho_{\rm f} \ll \rho_{\rm sp}$, the result (10.14) should be obtained. Therefore the limiting speed is

$$v = \frac{2}{9}\left(\frac{gR^2}{\nu}\right)\left(\frac{\rho_{\rm sp}}{\rho_{\rm f}} - 1\right). \tag{10.15}$$

This result was used by Millikan in his famous oil-drop experiments to determine precisely the charge of the electron (see Sections 16.2 and 17.5). As explained in Section 9.6, this result is expected to hold for small Reynolds numbers, $\mathrm{Re} \ll 1$.

High Reynolds Number Flow, $\mathrm{Re} \gg 1$

In the above analysis, it was assumed that viscous forces play a dominant role in transferring momentum to the sphere. Let us recall the terms in the Navier–Stokes equations which describe the response of the fluid to inertial and viscous forces:

$$\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} = \nu\nabla^2\boldsymbol{u}. \tag{10.16}$$

To order of magnitude, the differentials in (10.16) are

$$(\boldsymbol{u}\cdot\nabla)\boldsymbol{u} \sim \frac{v^2}{R}, \qquad \nu\nabla^2\boldsymbol{u} \sim \frac{\nu v}{R^2}. \tag{10.17}$$

The relative importance of the two terms is given by the ratio

$$\frac{\text{inertial acceleration}}{\text{viscous acceleration}} \approx \frac{v^2/R}{\nu v/R^2} = \frac{vR}{\nu} = \mathrm{Re}, \tag{10.18}$$

where Re is the Reynolds number.

Another way of looking at this result is in terms of the characteristic time for the diffusion of the viscous stresses through the fluid, which can be found to order of magnitude by approximating

$$\frac{\partial \boldsymbol{u}}{\partial t} = \nu\nabla^2\boldsymbol{u}; \quad \frac{v}{t_\nu} \approx \frac{\nu v}{R^2}; \quad t_\nu \approx \frac{R^2}{\nu}.$$

The time for the fluid to flow past the sphere is $t_v \sim R/v$. Thus, the condition that viscous forces have time to transmit momentum is

$$t_v \gg t_\nu; \quad R/v \gg \frac{R^2}{\nu}; \quad \frac{vR}{\nu} = \mathrm{Re} \ll 1.$$

In the high Reynolds number limit, $\mathrm{Re} = vR/\nu \gg 1$, the flow past the sphere becomes turbulent. Then, the viscosity is unimportant and can be dropped from our table of relevant parameters (Table 10.4).

Table 10.4 Parameters involved in estimating the drag force on a sphere in the limit Re $\gg$ 1

Variable	*Dimensions*	*Description*
f_d	$[M][L][T]^{-2}$	drag force
ρ_f	$[M][L]^{-3}$	density of fluid
R	$[L]$	radius of sphere
v	$[L][T]^{-1}$	terminal velocity

Table 10.5 Drag coefficients for objects of different shapes

Object	c_d	*Object*	c_d
Sphere	0.5	Flat plate	2.0
Cylinder	1.0	Car	0.4

There are four variables and three independent dimensions and so we can form only one dimensionless group, which we have found already:

$$\Pi_2 = \frac{f_d}{\rho_f R^2 v^2}.$$

It can be seen that Π_2 contains all the variables relevant to this case and no others can be involved. Just as in the case of Taylor's analysis of atomic explosions, the simplest solution is to try first the case in which Π_2 is taken to be a constant:

$$f_d = C\rho_f v^2 R^2, \tag{10.19}$$

where C is a constant of order unity.

There is a simple interpretation of the formula $f_d \propto \rho_f v^2 R^2$. The expression $\rho_f v^2$ is approximately the rate at which the momentum of the fluid per unit area arrives at the sphere and R^2 is the projected area of the sphere. Thus, (10.19) is roughly the rate at which the momentum of the incident fluid is pushed away by the sphere per second. To find the limiting velocity of the sphere under gravity, and taking account of the buoyancy of the medium, we find

$$f_d = C\rho_f v^2 R^2 = \frac{4\pi R^3}{3}(\rho_{sp} - \rho_f)g,$$

and so

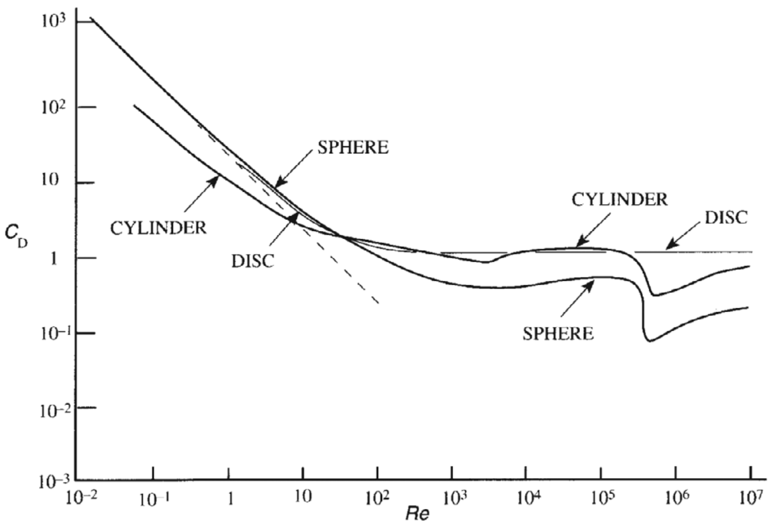
$$v \approx \sqrt{gR\frac{(\rho_{sp} - \rho_f)}{\rho_f}}.$$

Conventionally, the drag force is written

$$f_d = \tfrac{1}{2}c_d\rho_f v^2 A,$$

where A is the projected surface area of the object and c_d is the *drag coefficient*. Some values of c_d are given in Table 10.5.

(*a*)



(*b*)

Fig. 10.3 (*a*) Turbulent flow past a sphere at Re $= 117$. (*b*) The drag coefficient c_d for different objects as a function of Reynolds number Re. (From T.E. Faber, 1995. *Fluid Dynamics for Physicists*, pp. 258 and 266. Cambridge: Cambridge University Press.)

In the design of Formula 1 racing cars, the aim is to make c_d as small as possible.

While there are analytic solutions for low Reynolds numbers, when viscosity is dominant, there are none in the inertial regime at high Reynolds number. Figure 10.3(*a*) shows the turbulent vortices which develop behind a sphere at high Reynolds numbers. The variation of the drag coefficient for a wide range of Reynolds numbers is shown in Fig. 10.3(*b*). Inspection of (10.13) shows that in the limit $\mathrm{Re} \ll 1$, $f_d \propto \mathrm{Re}^{-1}$, consistent with relations shown in Fig. 10.3(*b*), but at high Reynolds numbers, the drag is independent of Re.

10.2.4 Kolmogorov Spectrum of Turbulence

Another remarkable example of the use of dimensional analysis in fluid dynamics is in the study of turbulence. As can be seen in Fig. 10.3(a), turbulent eddies form behind a sphere in flow at high Reynolds number. One of the great problems of fluid dynamics is understanding such turbulent flows, which arise from the non-linear term $(\boldsymbol{u} \cdot \nabla)\boldsymbol{u}$ in the Navier–Stokes equation (9.81). Empirically, the features of turbulence are well established. The energy of the eddies is injected on large scales, say, on the scale of the sphere seen in Fig. 10.3(a). The large-scale eddies fragment into smaller scale eddies, which in turn break up into even smaller eddies, and so on. In terms of energy transport, energy is injected on the largest scale and then cascades down though smaller and smaller eddies until the scales are sufficiently small that energy is dissipated as heat by viscous forces at the molecular level. This process is characterised by a *spectrum of turbulence*, which is defined to be the amount of energy per unit wavelength λ, or, equivalently, per unit wavenumber $k = 2\pi/\lambda$ in a steady state, when energy is continuously injected on a large scale and ultimately dissipated at the molecular level. The processes involved in the energy cascade are highly non-linear and there is no analytic theory of the spectrum of turbulence. Kolmogorov showed, however, that progress can be made by dimensional analysis.

In Kolmogorov's analysis, attention is concentrated upon the transfer of energy on scales between those at which energy is injected on the large scale l and the scale λ_{visc} at which molecular dissipation becomes important. There is no natural length scale between that at which energy is injected and that at which it is dissipated. The aim of the calculation is to determine the amount of kinetic energy in the eddies on different length scales. The natural way of describing this is in terms of the energy spectrum $E(k)\,\mathrm{d}k$, meaning the energy contained in eddies per unit mass in the wavenumber interval k to $k + \mathrm{d}k$. Such a spectrum lends itself naturally to a description in terms of a Fourier decomposition of the turbulent flow.

Suppose the rate of supply of kinetic energy per unit mass is ε_0 on the large scale. We form Table 10.6.

There are only three variables in the problem and two independent dimensions and so we can form only one dimensionless group:

$$\Pi_1 = \frac{E^3(k)k^5}{\varepsilon_0^2}.$$

Table 10.6 The variables involved in the description of homogeneous turbulence

Variable	*Dimensions*	*Description*
$E(k)$	$[\mathrm{L}]^3[\mathrm{T}]^{-2}$	energy per unit wavenumber per unit mass
ε_0	$[\mathrm{L}]^2[\mathrm{T}]^{-3}$	rate of energy input per unit mass
k	$[\mathrm{L}]^{-1}$	wavenumber

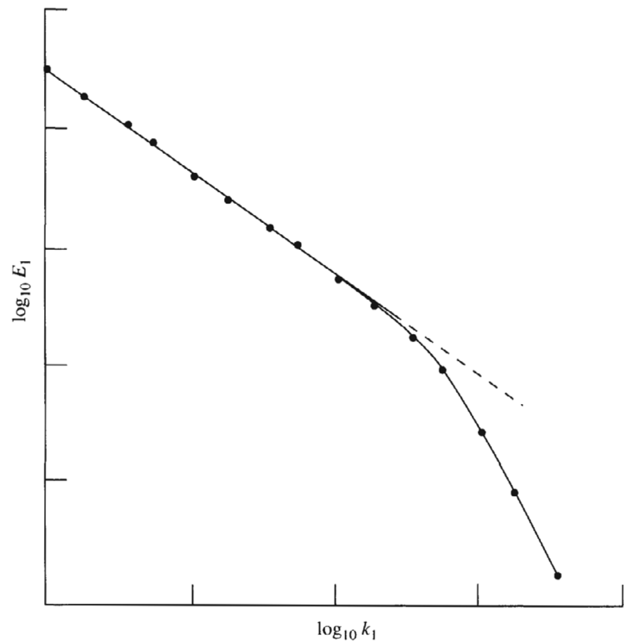


Fig. 10.4 The spectrum of homogeneous turbulence compared with the relation $E(k) \propto k^{-5/3}$ shown by the dashed line and derived from dimensional analysis. (From T.E. Faber, 1995. *Fluid Dynamics for Physicists*, p. 357. Cambridge: Cambridge University Press.)

As before, the simplest assumption is that Π_1 is a constant of order unity and so

$$E(k) \sim \varepsilon_0^{2/3} k^{-5/3}.$$

This relation, known as the *Kolmogorov spectrum of turbulence*, turns out to be a remarkably good fit to the spectrum of turbulence on scales intermediate between those on which the energy is injected and those on which dissipation by molecular viscosity is important (Fig. 10.4).

The examples in Sections 10.2.2 to 10.2.4 are only a few of the many uses of dimensionless analysis in fluid and gas dynamics. Many beautiful examples of the power of these techniques are given by G.I. Barenblatt in his classic book *Scaling, Self-Similarity, and Intermediate Asymptotics*.[8] A particularly appropriate example for the Cambridge students who attended this course is that the speed of a racing boat, or shell, is proportional to $N^{1/9}$, where N is the number of rowers.

10.2.5 The Law of Corresponding States

Let us study the application of dimensional techniques to the equation of state of imperfect gases, the aim being to derive the *law of corresponding states*. For n moles of a perfect gas, the pressure p, volume V and temperature T are related by the ideal gas law,

$$pV = nRT, \tag{10.20}$$

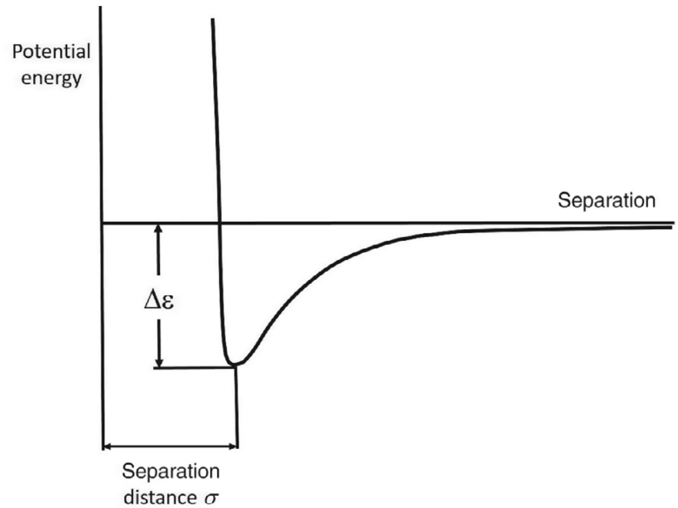


Fig. 10.5 A schematic diagram showing the intermolecular potential energy between one molecule of a gas and a single neighbour as a function of their separation. The equilibrium separation distance is σ. (After D. Tabor, 1991. *Gases, Liquids and Solids and Other States of Matter*, p. 125. Cambridge: Cambridge University Press.)

where R is the gas constant. This law provides an excellent description of the properties of all gases at low pressures and high temperatures, specifically, when the gas is far from the transition temperature to a liquid or solid.

In the case of an *imperfect gas*, we need to take account of the finite size of the molecules and the intermolecular forces acting between them. Thus, we need to relate the macroscopic properties of the gas to the microscopic properties of the molecules. The simplest way of characterising the forces acting between molecules is in terms of the potential energy of the molecules as a function of their separation (Fig. 10.5).[9]

In this model, there is an attractive potential between the molecules which becomes more negative as their separation decreases. This could, for example, be due to van de Waals forces. The potential continues to become more negative until at some scale σ the molecules are touching and then strong repulsive forces prevent them coming any closer together. Thus, σ is not only the distance of closest approach, but is also a measure of the size of the atoms or molecules. In Fig. 10.5, the properties of the intermolecular potential are characterised by the intermolecular separation σ and the depth of the attractive potential well $\Delta\varepsilon$.

All the variables which might be important in this problem, including $\Delta\varepsilon$, σ and the mass of a molecule of the gas m are listed in Table 10.7.

From the Buckingham Π theorem, there should be $(7-3) = 4$ independent dimensionless numbers. Inspection of Table 10.7 reveals that, in fact, there are rather fewer quantities with independent dimensions than appears at first sight. The quantities pV, $k_{\rm B}T$ and $\Delta\varepsilon$ all have the dimensions of energy, $[M][L]^2[T]^{-2}$. Furthermore, there is no means of creating a dimensionless quantity involving m alone, since there is no other quantity involving [T]

Table 10.7 The variables in the equation of state of an imperfect gas

Variable	*Dimensions*	*Description*
p	$[M][L]^{-1}[T]^{-2}$	pressure
V	$[L]^3$	volume
$k_B T$	$[M][L]^2[T]^{-2}$	temperature
N	–	number of molecules
m	$[M]$	mass of molecule
σ	$[L]$	intermolecular spacing
$\Delta\varepsilon$	$[M][L]^2[T]^{-2}$	depth of attractive potential

which could be used to eliminate [T] from the quantities with the dimensions of energy. Therefore, m cannot appear as a variable in the problem. Although we started with three independent dimensions, there are actually only two, which we can characterise as [L] and $[M][L]^2[T]^{-2}$. The dimensions of the six independent quantities in Table 10.7 can be constructed from [L] and $[M][L]^2[T]^{-2}$. The Buckingham Π theorem tells us that we can create (6 – 2) = 4 independent dimensionless quantities from these as before.

Let us choose the following set of four dimensionless groups to begin with:

$$\Pi_1 = N, \qquad \Pi_2 = \frac{k_B T}{\Delta\varepsilon}, \qquad \Pi_3 = \frac{V}{\sigma^3}, \qquad \Pi_4 = \frac{p\sigma^3}{\Delta\varepsilon}.$$

In this choice of the Πs, we have been guided by the need to express the macroscopic quantities p, V and T in terms of the microscopic quantities σ and $\Delta\varepsilon$.

To make progress, we now need to include some physics in the analysis. The equation of state involves only the three variables p, V and T, and yet we have four independent dimensionless groups. Suppose the gas is in some state characterised by p, V and T. p and T are *intensive* variables, which are independent of the 'extent' of the system. Suppose we cut the volume in half. The pressure and temperature remain unchanged, but the volumes of the two subvolumes and the numbers of molecules in each subvolume are both halved. Therefore, the equation of state can only involve the ratio $\Pi_5 = \Pi_3/\Pi_1 = V/N\sigma^3$.

We have reduced the problem to three dimensionless quantities, which have physical significance. Π_2 is the ratio of the typical thermal energy $k_B T$ of the molecules to their binding energy $\Delta\varepsilon$ – this ratio describes the relative importance of the kinetic energy of the molecules of the gas to their binding energies. Π_4 expresses the pressure in terms of that associated with the binding energy of the molecules and the volume over which it influences the properties of the gas. $\Pi_5 = V/N\sigma^3$ is the ratio of the volume V to the total volume occupied by the molecules – the volume $N\sigma^3$ is referred to as the *excluded volume*.

Equations of state traditionally relate the pressure to the volume and temperature, suggesting that the relation between the remaining dimensionless groups should have the form

$$\Pi_4 = f(\Pi_2, \Pi_5), \qquad \frac{p\sigma^3}{\Delta\varepsilon} = f\left(\frac{V}{N\sigma^3}, \frac{k_B T}{\Delta\varepsilon}\right),$$

$$\frac{p}{p^*} = f\left(\frac{V}{V^*}, \frac{T}{T^*}\right),$$

Table 10.8 Values of $p^* V^* / RT^*$ for different gases

Gas	$p^* V^* / RT^*$	*Gas*	$p^* V^* / RT^*$
He	0.327	Ar	0.291
Xe	0.277	H_2	0.306
O_2	0.292	N_2	0.292
Hg	0.909	H_2O	0.233
CO_2	0.277		

where

$$p^* = \frac{\Delta\varepsilon}{\sigma^3}, \quad V^* = N\sigma^3, \qquad T^* = \frac{\Delta\varepsilon}{k_B}. \tag{10.21}$$

With this choice of dimensionless quantities, $p^* V^* / N k_B T^* = 1$. The equation of state of the gas written in terms of p^*, V^* and T^* is called a *reduced equation of state*. The *Law of Corresponding States* is the statement that *the reduced equation of state is the same for all gases*.

Thus, if the shape of the potential function for gases were the same for all gases, we would expect $p^* V^* / RT^*$ to be constant for one mole of gas. The values for a selection of gases are shown in Table 10.8.

It can be seen that, despite the crudeness of the model, it is remarkably successful for gases of the same molecular family. Examples of the forms of equation of state which have proved successful include the *van der Waals* and *Dieterici* equations of state, which can be written

$$\text{van der Waals} \qquad \left(\pi + \frac{3}{\phi^2}\right)(3\phi - 1) = 8\theta, \tag{10.22}$$

and

$$\text{Dieterici} \qquad \pi(2\phi - 1) = \theta \exp\left[2\left(1 - \frac{1}{\theta\phi}\right)\right], \tag{10.23}$$

where $\pi = p/p^*$, $\phi = V/V^*$ and $\theta = T/T^*$. Notice that we gain physical insight into the meanings of the quantities p^*, V^* and T^*. Equation (10.21) shows how these quantities provide information about the separation between molecules and their binding energies.

This is only one of many examples which could be given of the use of dimensional analysis to obtain insight into physical processes where the detailed microphysics is non-trivial. Other examples include Lindemann's law of melting and critical phenomena in phase transitions, for example, the ferromagnetic transition. We will find an important application of the use of dimensional techniques in Wien's derivation of his displacement law (Section 13.4).

10.3 Introduction to Chaos

One of the most remarkable developments in theoretical physics over the last 50 years has been the study of the chaotic dynamical behaviour found in systems which follow deterministic equations, such as Newton's laws of motion. These studies were foreshadowed by the work of mathematicians such as Poincaré, but only with the development of high-speed computers has the extraordinary richness of the subject become apparent. The character of chaotic systems can be appreciated by considering a dynamical system which is started twice, but from very slightly different initial conditions, say, due to a small error.

- In the case of a *non-chaotic system*, this small difference leads to an error in prediction which grows linearly with time.
- In a *chaotic system*, the difference grows exponentially with time, so that the state of the system is essentially unknown after a few characteristic times, despite the fact that the equations of motion are entirely deterministic.

Thus, chaos is quite different from randomness – the solutions are perfectly well defined, but the non-linearities present in the equations mean that the outcome is very sensitive indeed to the precise input parameters. The study of chaotic dynamical systems rapidly became a heavy industry with several excellent books on the subject. James Gleick's *Chaos: Making a New Science*[10] is a splendid example of a popular science book communicating the real essence of a technical subject with journalistic flair; a simple analytic introduction is provided by the attractive textbook by Gregory Baker and Jerry Gollub entitled *Chaotic Dynamics: An Introduction* .[11] My intention here is to describe briefly a few important aspects of these developments which indicate the richness of the field.

10.3.1 The Discovery of Chaotic Behaviour

Chaotic behaviour in dynamical systems was discovered by Edward Lorenz of the Massachusetts Institute of Technology in the late 1950s. His interest was the study of weather systems and he derived a set of non-linear differential equations to describe their evolution. This was at the very dawn of the introduction of computers into the study of dynamical systems and, in 1959, he programmed his set of equations into his new Royal McBee electronic computer, which could carry out computations at a rate of 17 calculations per second. In his own words,

> The computer printed out the figures to three significant places, although they were carried in the computer to six, but we didn't think we cared about the last places at all. So I simply typed in what the computer printed out one time and decided to have it print it out in more detail, that is, more frequently. I let the thing run on while I went out for a cup of coffee and, a couple of hours later when it had simulated a couple of months of data, I came back and found it doing something quite different from what it had done before.
>
> It became obvious that what had happened was that I had not started off with the same conditions – I had started off with . . . the original conditions plus extremely small errors. In the course of a couple of simulated months these errors had grown – they doubled

> about every four days. This meant that, in two months, they had doubled 15 times, a factor of 30,000 or so, and the two solutions were doing something entirely different.[12]

This is the origin of what became known popularly as the *butterfly effect*, namely, even a butterfly flapping its wings in one continent could cause hurricanes in some other part of the globe. Alternatively, the term can be regarded as a metaphor for the concept of *sensitive dependence upon initial conditions*. Lorenz simplified his equations to describe the interaction between temperature variations and convective motion:

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = -\sigma x + \sigma y, \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = xz + rx - y, \\ \dfrac{\mathrm{d}z}{\mathrm{d}t} = xy - bz, \end{cases} \tag{10.24}$$

where σ, r and b are constants. This system also displayed chaotic behaviour. Lorenz's computations marked the beginning of a remarkable story which involved a number of individuals working largely in isolation, who all came across different aspects of what we would now call chaotic behaviour. The systems studied included dripping taps, electronic circuits, turbulence, populations dynamics, fractal geometry, the stability of Jupiter's great red spot, the dynamics of stellar orbits, the tumbling of Hyperion, one of the satellites of Saturn, and many other examples. This story is delightfully recounted in Gleick's book.

I particularly enjoy Baker and Gollub's short book which describes very clearly many of the essential features of chaotic systems. As they discuss, the minimum conditions for chaotic dynamics are:

- the system has at least three dynamical variables;
- the equations of motion contain a non-linear term coupling several of the variables together.

These features are satisfied by Lorenz's equations (10.24) for convective motions – there are three variables x, y, z and the coupling between them takes place through the non-linear terms xz and xy on the right hand sides of the last two equations. These conditions are necessary in order to obtain the following distinctive features of chaotic behaviour:

- the trajectories of the system in three dimensions diverge;
- the motion is confined to a finite region of the phase space of the dynamical variables;
- each trajectory is unique and does not intersect any other trajectory.

It is simplest to give a specific example of how a dynamical system can become chaotic – Baker and Gollub's analysis of a damped driven pendulum illustrates elegantly how chaotic motion comes about.

10.3.2 The Damped Driven Pendulum

Consider a damped sinusoidally-driven pendulum of length l with a bob of mass m and damping constant γ. The pendulum is driven at angular frequency ω_D. We have already written down the expression for the non-linear pendulum (10.7), but without the damping and driving terms. As shown in Section 10.2.1, the equation of motion without the additional terms is

$$mgl(1 - \cos\theta) + \tfrac{1}{2}I\dot{\theta}^2 = \text{constant}. \tag{10.25}$$

Setting $I = ml^2$ and differentiating with respect to time,

$$ml\ddot{\theta} = -mg\sin\theta. \tag{10.26}$$

The damping term is proportional to the speed of the pendulum bob and so can be included as the term $-\gamma l\dot{\theta}$ on the right-hand side of (10.26). Likewise, the driving term to be included on the right-hand side can be written $A\sin\omega_{\rm D}t$. The equation of motion is therefore

$$ml\frac{{\rm d}^2\theta}{{\rm d}t^2} + \gamma l\frac{{\rm d}\theta}{{\rm d}t} + mg\sin\theta = A\cos\omega_{\rm D}t. \tag{10.27}$$

Changing variables, Baker and Gollub rewrite (10.27) in simplified form:

$$\frac{{\rm d}^2\theta}{{\rm d}t} + \frac{1}{q}\frac{{\rm d}\theta}{{\rm d}t} + \sin\theta = g\cos\omega_{\rm D}t, \tag{10.28}$$

where q is the *quality factor* of the pendulum and g is now the amplitude of the driving oscillation, and *not* the acceleration due to gravity. The equation has been normalised so that the natural frequency of small oscillations of the pendulum ω is 1. It is simplest to regard the angular frequencies and time as dimensionless. In the cases we are about to consider, the amplitudes of the oscillations can be very large. The pendulum wire has to be regarded as stiff and weightless when the angular deviations exceed $\theta = \pm\pi/2$.

The equation (10.28) can be written as three independent first-order equations as follows:

$$\begin{cases} \dfrac{{\rm d}\omega}{{\rm d}t} = -\dfrac{1}{q}\omega - \sin\theta + g\cos\phi, \\ \dfrac{{\rm d}\theta}{{\rm d}t} = \omega, \\ \dfrac{{\rm d}\phi}{{\rm d}t} = \omega_{\rm D}. \end{cases} \tag{10.29}$$

These equations contain three independent dynamical variables, ω, θ and ϕ:

- ω is the instantaneous angular velocity;
- θ is the angle of the pendulum with respect to its vertical equilibrium position;
- ϕ is the phase of the oscillatory driving force which has constant angular frequency $\omega_{\rm D}$.

The non-linear couplings between these variables are introduced through the terms $\sin\theta$ and $g\cos\phi$ in the first of the three equations. Whether the motion is chaotic or not depends upon the values of g, $\omega_{\rm D}$ and q.

One of the problems in representing chaotic behaviour is that the diagrams necessarily have to be three dimensional in order to represent the evolution of the three dynamical variables. Fortunately, in this case, the phase of the driving oscillation ϕ is linearly proportional to time and so the dynamical behaviour can be followed in a simple two-dimensional phase diagram in which the instantaneous angular velocity ω is plotted against the angular position θ of the pendulum. The ω–θ phase diagrams for four values of g and q are shown in Fig. 10.6 in which they are drawn on the same scale. The behaviour of the pendulums in real space is shown in the left-hand column with their values of g and q.

Considering first the simplest example of the *moderately driven pendulum*, with $g = 0.5$ and $q = 2$, the phase diagram shows simple periodic behaviour, the pendulum swinging between $\theta = \pm 45^\circ$ at which the angular velocity ω is zero (Fig. 10.6(a)). This closed loop behaviour is familiar from the simple pendulum. In the terminology of dynamical systems, the origin of the phase diagram is called an *attractor*, meaning that, if there were no driving term, the trajectory of the pendulum in the phase diagram would spiral in, or be attracted to, the origin because of the damping term represented by q.

The three other diagrams show the periodic behaviour of the pendulum when the amplitude of the driving oscillator is large in the steady state.

- In Fig. 10.6(b), $g = 1.07$ $q = 2$, the pendulum follows a double-looped trajectory in the phase diagram, in one swing exceeding π radians from $\theta = 0$.
- At a larger driving amplitude, $g = 1.35$ $q = 2$ in Fig. 10.6(c), the pendulum performs a complete loop about the origin in real space. In the phase diagram, the ends of the trajectory at $\pm\pi$ radians are joined up.
- At an even greater driving amplitude (Fig. 10.6(d)), $g = 1.45$ $q = 2$, the pendulum performs two complete loops in real space in returning to its initial position in the steady state. Again, the end points at $\pm\pi$ radians in the phase diagram should be joined up.

At other values of g and q, there is, however, no steady state periodic behaviour. In Fig. 10.7, the phase diagram for $q = 2$ and $g = 1.50$ is shown. The motion is chaotic in the sense that the trajectories are non-periodic and are very sensitive to the initial conditions.

There are other revealing ways of presenting the information present on the phase diagrams. One way of characterising the behaviour of the pendulum is to plot on the phase space diagrams only the points at which the pendulum crosses the ω–θ plane at *multiples of the period of the forcing oscillation*, that is, we make *stroboscopic pictures* of the ω and θ coordinates at the forcing frequency ω_D. This form of presentation is called a *Poincaré section* and enables much of the complication of the dynamical motion to be eliminated.

For a simple pendulum, there is only one point on the Poincaré section. The examples shown in Fig. 10.8 correspond to the phase diagrams in Fig. 10.6(d) and Fig. 10.7. In the first example, the phenomenon of period doubling is apparent. This is the type of non-linear phenomenon found, for example, in dripping taps. In the chaotic example, the trajectories never cross the reference plane at the same point, but there is regular structure in the Poincaré section. When the motion becomes chaotic, the structure of the Poincaré sections becomes self-similar on different scales. As discussed by Baker and Gollub, the geometrical structure of the diagrams becomes fractal and a fractional fractal dimension can be defined. They describe how the fractal dimension of the phase space trajectories can be calculated.

Another powerful technique for analysing the motion of the pendulum is to take the *power spectrum* of its motion in the time domain. In the case of periodic motion, the power

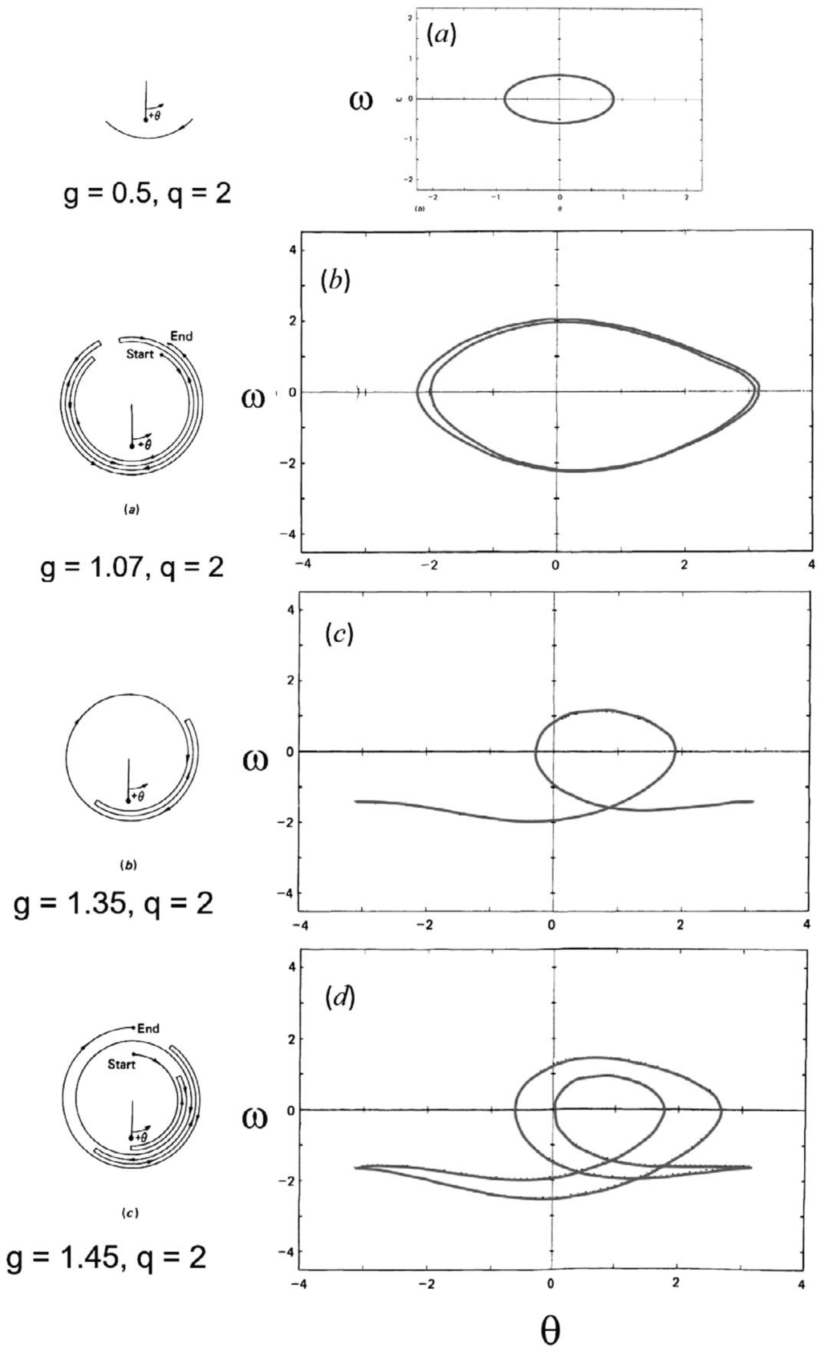


Fig. 10.6 $\omega - \theta$ phase diagrams for the non-linear damped, driven pendulum for different values of the quality factor q and the amplitude g of the driving oscillation. (After G.L. Baker and J.P. Gollub, 1990.*Chaotic Dynamics: An Introduction*, pp. 21–2. Cambridge: Cambridge University Press.)

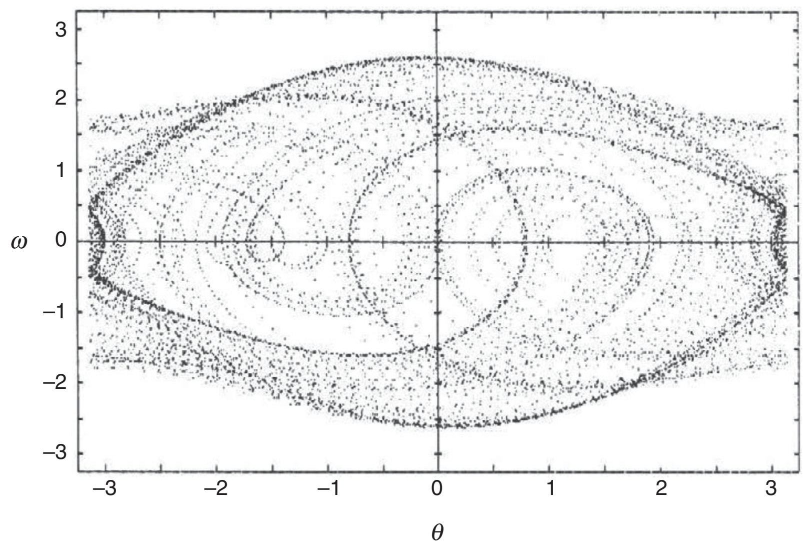


Fig. 10.7 Phase diagram for the non-linear pendulum with $q = 2$ and $g = 1.50$. The motion is chaotic. (G.L. Baker and J.P. Gollub, 1990. *Chaotic Dynamics: An Introduction*, p. 53. Cambridge: Cambridge University Press.)

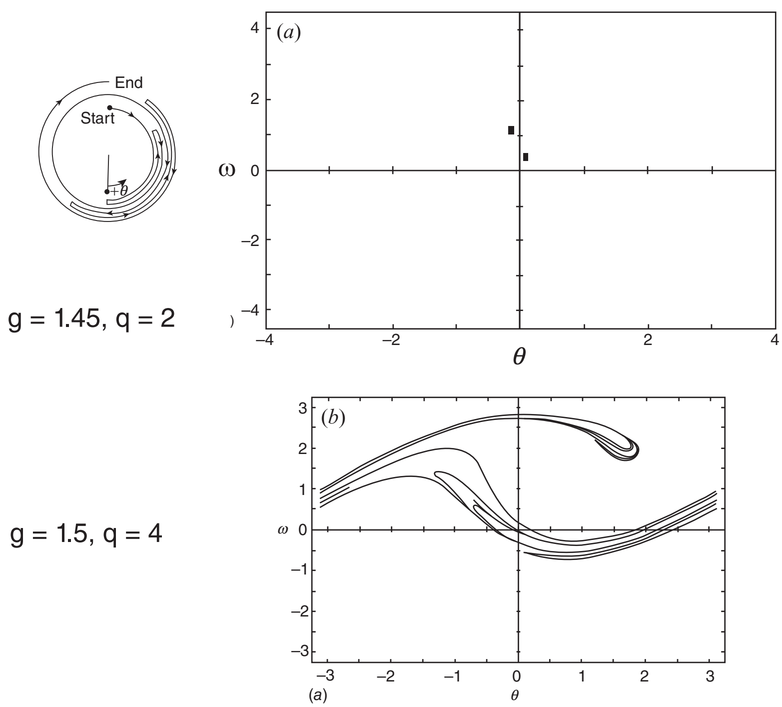


Fig. 10.8 Poincaré sections for the non-linear pendulum for (*a*) $q = 2$ and $g = 1.45$ and (*b*) $g = 1.5$, $q = 4$. In (*a*), period doubling takes place, while in (*b*), the motion is chaotic. (G.L. Baker and J.P. Gollub, 1990. *Chaotic Dynamics: An Introduction*, pp. 51 and 53. Cambridge: Cambridge University Press.)

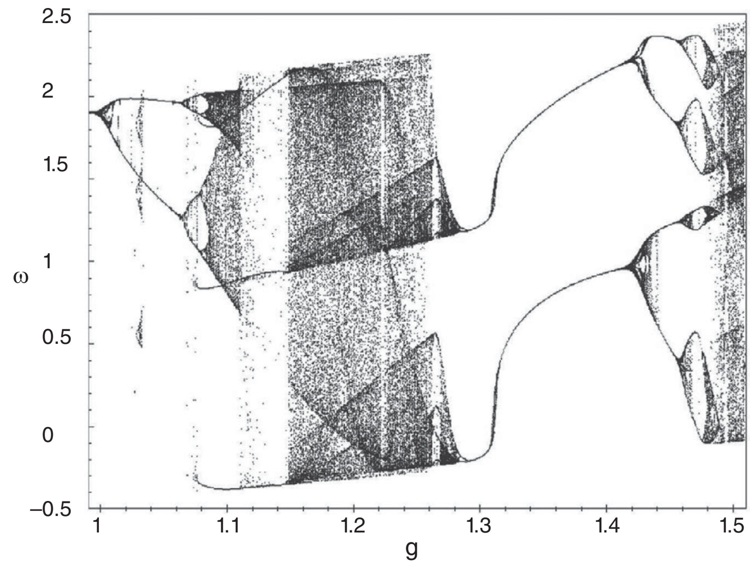


Fig. 10.9 The bifurcation diagram for a damped driven pendulum with $q = 2$ for a range of values of the amplitude of the driving oscillation g. The values of ω are those at the beginning of each driving cycle. (G.L. Baker and J.P. Gollub, 1990. *Chaotic Dynamics: An Introduction*, p. 69. Cambridge: Cambridge University Press.)

spectra display spikes corresponding to period doublings, quadruplings and so on. In the case of chaotic systems, broad-band power spectra are found.

Finally, the overall behaviour of the pendulum for different magnitudes of the forcing amplitude g can be displayed on a *bifurcation diagram* (Fig. 10.9). On the vertical axis, the magnitude of the angular velocity at a fixed phase in each period of the driving frequency is shown. The bifurcation diagram shown in Figure 10.9 displays the behaviour of the pendulum for the examples we have already studied with $q = 2$, but with a continuous range of values of g. It is useful to compare the behaviour of the pendulum in real and phase space in Figs. 10.6 and 10.7 with their location on the bifurcation diagram. Figure 10.9 makes the point very clearly that the onset of chaotic behaviour proceeds through a sequence of frequency doublings until the system becomes chaotic. This behaviour has been observed in laboratory experiments on turbulence. Figure 10.9 is a quite extraordinary diagram showing ranges of parameters in which the system has well-defined modes of periodic oscillation between regions in which the behaviour is chaotic.

All the diagrams presented above are ways of representing what are at their simplest three-dimensional diagrams on a two-dimensional sheet of paper. It is instructive to study the dynamical motion of Lorenz's three-parameter model of convection in three dimensions (Fig. 10.10). For a given set of input parameters, a unique trajectory is traced out in the three-dimensional phase space, as illustrated by the continuous line in the diagram. For a slightly different set of initial conditions, a similar trajectory would be traced out initially, but it would rapidly diverge from that shown. Notice that the evolution of the system is confined to a finite region of parameter space. In this representation, there are two attractors and the system can pass from the sphere of influence of one attractor to the other. This is an example of a *strange attractor*.

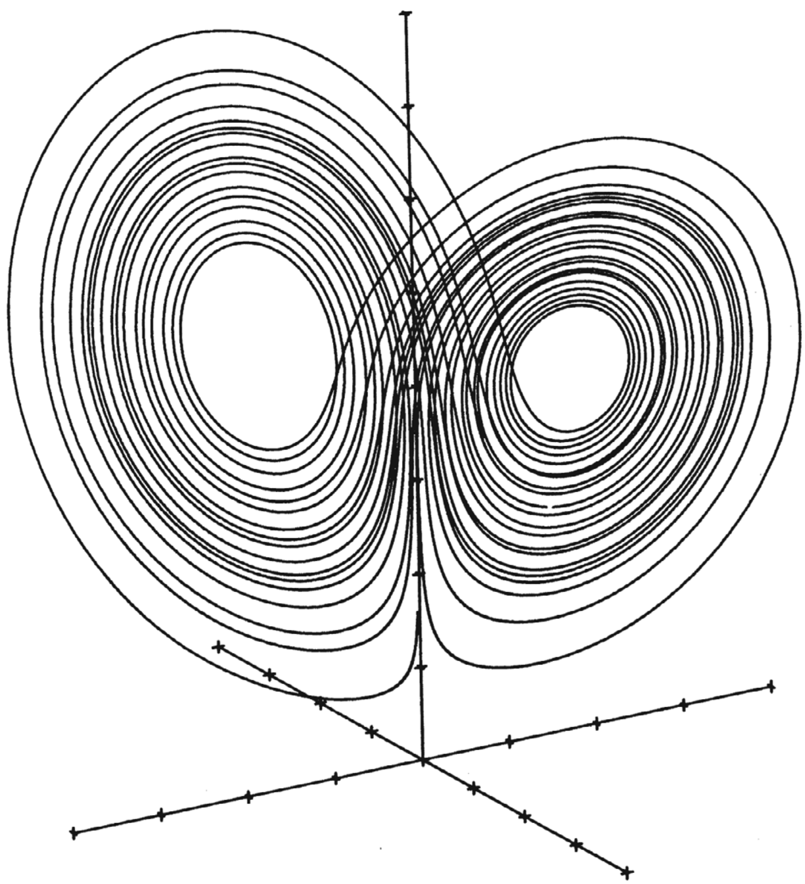


Fig. 10.10 A Lorenz attractor diagram is a three-dimensional representation of the evolution of the dynamics of a chaotic system with three independent variables. The trajectories of the system are represented by the continuous lines which never cross. Particles on neighbouring trajectories diverge rapidly resulting in chaotic motion. (From J. Gleick, 1987. *Chaos: Making a New Science*, p. 28. New York: Viking.)

10.3.3 Logistic Maps and Chaos

Remarkably, chaotic behaviour was also discovered by an entirely separate line of development. Robert May worked on the problems of population growth during the 1970s and, in particular, how a biological population reaches a stable state.[13] The *n*th generation with x_n members results in the $(n + 1)$th generation with x_{n+1} members and so we can write

$$x_{n+1} = f(x_n). \qquad (10.30)$$

In a simple picture, we might imagine, for example, that the average number of offspring for a pair of parents was the traditional 2.4 or some other number. Then, we would expect that, if $x_{n+1} = \mu x_n$, there would be unlimited, exponential growth of the size of the population if $\mu > 1$. There would therefore have to be some form of cut-off to limit the size of the population. A simple mathematical way of introducing such a cut-off would be to modify the recurrence relation to be

$$x_{n+1} = \mu x_n(1 - x_n). \tag{10.31}$$

This *difference equation* is similar to the differential equation introduced by Pierre-François Verhulst to model the development of populations in 1838:[14]

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \mu x(1 - x). \tag{10.32}$$

The behaviour of the population can be modelled on diagrams known as *logistic maps*, such as those illustrated in Fig. 10.11. In each of these diagrams, the abscissa represents the numbers of members of the nth population and the ordinate that of the $(n+1)$th generation. This latter number becomes the next value on the abscissa, to which (10.31) is next applied. In case (a), the value of μ is 2 and the sequence starts with the value $x_1 = 0.2$. Then, the next generation has population $x_2 = 0.32$. Applying the formula (10.31), the successive generations have populations $x_3 = 0.4352$, $x_4 = 0.4916$, $x_5 = 0.4999$, settling down to a stable population with $x_n = 0.5$.

In case (b), the value of μ is 3.3 and the sequence of generations shown Fig. 10.11(b) is obtained. Starting with the value $x_1 = 0.2$, the subsequent generations have populations $x_2 = 0.528$, $x_3 = 0.8224$, $x_4 = 0.4820$, $x_5 = 0.8239$, $x_6 = 0.4787, \ldots$, the population ultimately settling down to a population which oscillates between the values $x_n(\text{even}) = 0.48$ and $x_{(n+1)}(\text{odd}) = 0.83$. This is an example of the population undergoing *limit cycle* behaviour between these values. As pointed out by May,[15] this form of limit cycle may account for the population cycles of lemmings in northern regions as being the result of a predator–prey interaction, with the lemmings playing the role of predator and their food the prey. The regular oscillations of the lynx and hare populations recorded by the Hudson's Bay Trading Company in Northern Canada may be associated with a similar predator–prey interaction.

In Fig. 10.11(c), $\mu = 3.53$ and the first six values of the population are $x_1 = 0.2$, $x_2 = 0.5648$, $x_3 = 0.8677$, $x_4 = 0.4053$, $x_5 = 0.8508$, $x_6 = 0.4480, \ldots$. Eventually, the system settles down to a limit cycle between four stable states with values $x_n = 0.37$, $x_{(n+1)} = 0.52$, $x_{(n+2)} = 0.83$ and $x_{(n+3)} = 0.88$, as illustrated in the third panel. We recognise in cases (b) and (c) the phenomena of period doubling and quadrupling, as was found in the damped forced harmonic oscillator.

In the fourth example, Fig. 10.11(d), $\mu = 3.9$ and the system becomes chaotic. Just as in the case of the pendulum, we can create a bifurcation diagram to illustrate the stable values of the population (Fig. 10.12). It is remarkable that the bifurcation diagrams of Fig. 10.9 and Fig. 10.12 are so similar. There are evidently universal features of the processes of period doubling and the approach to chaotic behaviour in both examples.

In a remarkable analysis of 1978, Mitchell Feigenbaum discovered that there are regularities in the approach to chaotic behaviour through period doubling.[16] He showed that the ratio of spacings between consecutive values of μ at the bifurcations tends to a

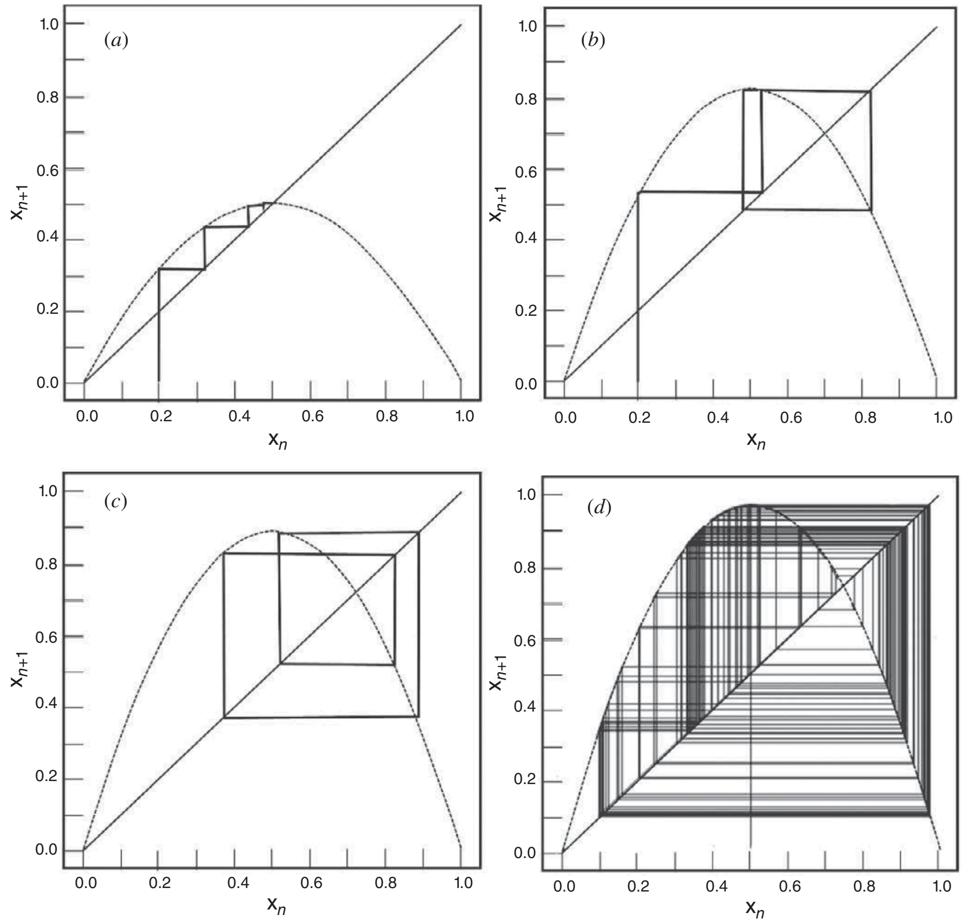


Fig. 10.11 Logistic maps for the recurrence relation $x_{n+1} = \mu x_n(1 - x_n)$ for the values $\mu = 2, 3.3, 3.53$ and 3.9.

universal constant, the *Feigenbaum number*. If the first bifurcation occurs at μ_1, the second at μ_2 and so on, the Feigenbaum number is defined to be

$$\lim_{k\to\infty} = \frac{\mu_k - \mu_{k-1}}{\mu_{k+1} - \mu_k} = 4.669\,201\,609\,102\,990\,9\ldots$$

This number is a universal constant for the period-doubling route to chaos for maps with a quadratic maximum. It is an example of *universality* in chaotic dynamical systems, by which is meant that certain features of non-linear maps are independent of the particular form of map.

There are many applications of these ideas to complex physical systems. These include the tumbling motion of the satellite Hyperion in orbit about Saturn, turbulence, lasers, chemical systems, electronic circuits with non-linear elements, and so on. The idea of chaos in non-linear systems now pervades many frontier areas of modern physics – it is no longer a mathematical curiosity. The key point is that there are universal features which underlie all chaotic phenomena and which are now an integral part of physics.

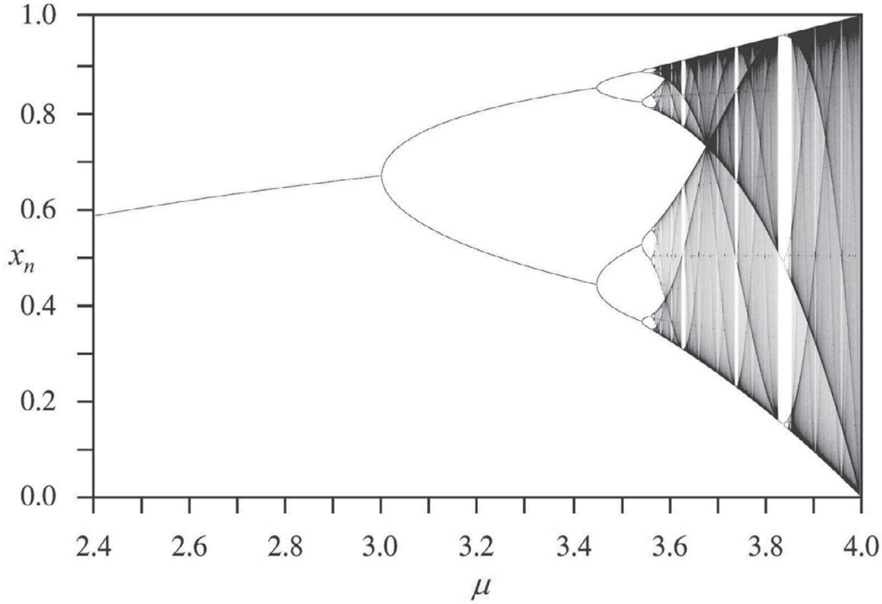


Fig. 10.12 A bifurcation diagram for the logistic maps of $x_{n+1} = \mu x_n(1 - x_n)$. (Courtesy of Creative Commons, https://commons.wikimedia.org/w/index.php?curid=323398.)

10.4 Scaling Laws and Self-Organised Criticality

The chaotic phenomena described in Section 10.3 revealed underlying regularities in non-linear dynamical systems. The examples studied have generally been relatively simple and describable by a limited number of non-linear differential equations. There are, however, many examples in nature in which non-linearities of many different forms are present and there is little hope that we can write down the equations, let alone solve them. Nonetheless, it is found that regularities appear. How does this come about? The emerging field of *self-organised criticality* may contain the seeds of a new type of physics for tackling such problems. The book *How Nature Works: The Science of Self-Organised Criticality* by Per Bak is an excellent introduction to these concepts.[17] We follow his presentation as a brief introduction to the key ideas.

10.4.1 Scaling Laws

The germ of the idea for this approach to multiply non-linear phenomena originated from the discovery that *scaling laws* are found in a number of extremely complex systems. The common feature of these examples is that they are so complex that it is essentially impossible to write down the equations which might describe the relevant aspects of the system and yet regularities appear. Here are some of the examples discussed by Bak.

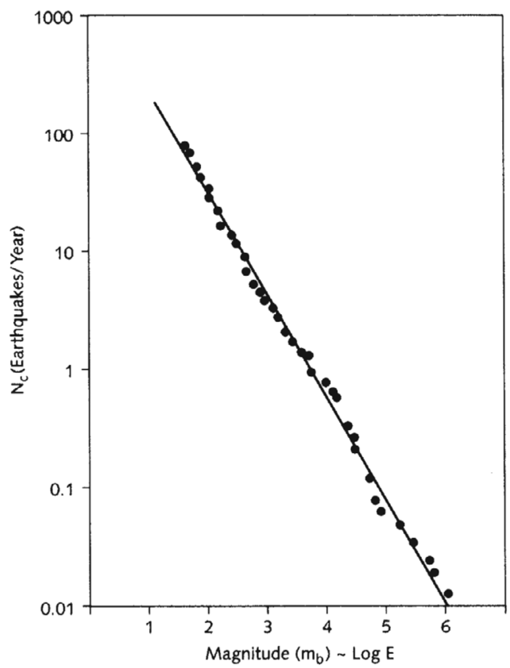


Fig. 10.13 The probability of an earthquake of magnitude greater than *m* on the Richter scale in a log-log plot. (From P. Bak, 1997. *How Nature Works: The Science of Self-Organised Criticality*, p. 13. Oxford: Oxford University Press.)

- A typical example is the *Gutenberg–Richter law for the probability of the occurrence of earthquakes* of different magnitude on the Richter scale. As an example, the earthquake statistics for the New Madrid zone in the south-east USA were studied by Arch Johnson and Susan Nava for the period 1974–83, using the Richter scale as a measure of the magnitude of the earthquake. The Richter scale is logarithmic in the energy release E of the earthquake. They confirmed the Gutenberg–Richter law that the probability of occurrence of an earthquake with magnitude greater than m is a power-law in the energy E (Fig. 10.13). It is quite extraordinary that there should be any simple relation between the probability of an earthquake occurring and its magnitude. Consider all the immensely complicated phenomena which can affect the occurrence of an earthquake – the list would include plate tectonics, mountains, valleys, lakes, geological structures, weather and so on. Whatever actually happens, it is evident that the processes are extremely non-linear, and yet the linearity of the Gutenberg–Richter law suggests that, despite the complexity, there is some underlying systematic behaviour. Large earthquakes are just extreme examples of smaller earthquakes.
- In 1963, Benoit Mandelbrot analysed the *monthly variation of cotton prices* over a 30 month period and found that the distribution was very 'spikey'. When he measured the distribution of the *changes* in the cotton price from month to month, these variations followed a power-law distribution. Other commodities, and hence the stock exchange itself, also followed this pattern. This really is an extremely non-linear problem. Consider

the many effects which determine the cotton price – supply, demand, rushes on the stock market, speculation, psychology, insider dealing, and so on. How could these result in simple power-law behaviour?

- John Sepkopski spent 10 years in libraries and museums studying *biological extinctions*. He made a huge survey of all the marine fossil records available to him in order to determine the fractions of biological families present on the Earth which became extinct over the last 600 million years. The data were grouped into periods of 4 million years and he determined what fraction of all species disappeared during each of these. The data were again very 'spikey', with huge variations in the extinction rate. Nonetheless, David Raup pointed out that the probability distribution of the fraction of extinctions in each 4 million year period followed a power-law distribution, with a few large extinction events and larger numbers of smaller ones. The famous epoch of the extinction of the dinosaurs, which occurred about 65 million years before the present epoch, was by no means the largest extinction event in the fossil record. The extinction of the dinosaurs may well have been associated in some way with the collision of an asteroid with the Earth at that epoch – the 180-km crater resulting from that collision was identified in 1991 at Chicxulub on the Yutacán peninsula of the Gulf of Mexico.
- The term *fractal* was introduced by Benoit Mandelbrot to describe geometrical structures with *features on all length scales*. In his memoirs, he confessed that he was always interested in describing mathematically the 'messiness' of the world rather than its regularities. In *The Fractal Geometry of Nature*,[18] he made the profound observation that the forms found in nature are generally of a fractal character.

  A good example is the coastline of Norway. Being buffetted by the Atlantic and the severe climate of the northern part of the globe, Norway has a very ragged shoreline containing fjords, fjords within fjords, and so on. The simplest way of quantifying its fractal nature is to ask how many square boxes of side δ are needed to enclose the complete coastline (Fig. 10.14). Remarkably, the number of boxes follows a power-law distribution $L(\delta) \propto \delta^{1-D}$, where $D = 1.52 \pm 0.01$ is known as the *fractal* or *Hausdorff dimension* of the distribution. Notice that, because it is of power-law form there is no natural length-scale with which to identify on what scale we might be looking at a photograph of the coastline. To put it another way, the coastline displays the same appearance of irregularity no matter what scale we select.
- A similar scaling law is found in the *large-scale distribution of galaxies in the Universe*. Galaxies have a strong tendency to be found in groups or clusters. These in turn are clustered into associations known as superclusters of galaxies. The simplest way of describing the clustering properties of galaxies is by means of a two-point correlation function defined by the relation

$$N(\theta)\,\mathrm{d}\Omega = N_0[1 + w(\theta)]\,\mathrm{d}\Omega,$$

where N_0 is a suitably defined average surface density and $w(\theta)$ describes the excess probability of finding a galaxy at angular distance θ within the solid angle $\mathrm{d}\Omega$ on the sky. Observations of large samples of galaxies have shown that the function $w(\theta)$ has the form

$$w(\theta) \propto \theta^{-0.8},$$

Fig. 10.14 The coast of Norway showing how the fractal nature of its coastline can be quantified by determining the number of boxes of side δ required to enclose it. In this example, 26 boxes are needed to enclose the region shown. (From P. Bak, 1997. *How Nature Works: The Science of Self-Organised Criticality*, p. 13. Oxford: Oxford University Press. After J. Feder: 1988. *Fractals*. New York: Plenum Press.)

corresponding to a spatial correlation function

$$\xi(r) \propto r^{-1.8} \quad \text{on physical scales} \quad r \leq 8 \text{ Mpc}, \tag{10.33}$$

where 1 Mpc $\approx 3\times10^{22}$ m. Thus, although clusters of galaxies are prominent structures in the Universe, large-scale clustering occurs on all scales and there are no prefered scales, just as in a fractal distribution. On scales $L \gg 8$ Mpc, the strength of the clustering decreases more rapidly with increasing scale than (10.33) – on these scales the density contrast $\delta\rho/\rho \ll 1$ and so structures with these enormous dimensions are still in their linear stage of development. On scales less than 8 Mpc, the density contrasts are greater than 1 and consequently involve non-linear gravitational interactions.

- $1/f$ *noise* is found in many different circumstances. It was originally discovered in electronic circuits in which it manifests itself as spikes far exceeding the expectations of random Johnson noise (see Section 17.4.2). As my colleague Adrian Webster once remarked, ‘$1/f$ noise is the spectrum of trouble’. Its origin has not been established.

Two typical examples of $1/f$ noise are the global temperature variations since 1865 and the time variability of the luminosities of quasars. Both are qualitatively similar. Generally, the $1/f$ noise spectrum has the form

$$I(f) \propto f^{-\alpha} \quad \text{where} \quad 0 < \alpha < 2$$

and f is the frequency. In both cases, the observed behaviour is thought to be the result of a large number of non-linear processes which end up producing this systematic behaviour. In the case of electronic circuits, precautions must be taken so that the presence of these large spikes does not upset the operation of the device.

- In his book *Human Behaviour and the Principle of Least Effort*,[19] George Kingsley Zipf noted a number of remarkable scaling laws in the patterns of human behaviour. For example, he showed that there is systematic behaviour in the ranking of *city sizes*, just as there is in the case of earthquakes. The cities of the world can be ranked in order of their populations and it is found that they obey a power-law distribution, similar to the Gutenberg–Richter law. A large number of non-linear physical and social effects determine the size of any city, and yet these all conspire to create a power-law size distribution.
- Zipf also showed that the frequency with which words are used in the English language follows a power-law scaling law. Words in the English language can be ranked in order of their frequency of use. For example, 'the' is the most common, followed by 'of', 'and', 'to' and so on. What are the myriad social and perceptual processes which result in a power-law rather than some other distribution?

In all these examples, massively non-linear phenomena are important and there appears at first sight to be little prospect of finding any analytic theory to account for the appearance of regular scaling laws in the real world about us. And yet there is hope from a somewhat unexpected direction.

10.4.2 Sand Piles and Rice Piles

A clue to understanding these different examples of complex non-linear systems may be found in the behaviour of sand or rice piles. Every child is familiar with the problems of building sand castles. If the walls of a sand castle become too steep, avalanches of various sizes occur until its angle is reduced to some value at which they no longer occur. In a controlled experiment, dry sand is poured steadily onto the top of the pile and the sand pile maintains its regular conical shape by undergoing a continuous series of avalanches of different sizes. The remarkable result is that, as it grows, the angle of the sand pile remains the same. This example has become the archetype for the processes involved in what is called *self-organised criticality*. It is not feasible to work out analytically the physics of the avalanches which occur in sand or rice piles. There are just too many non-linear phenomena coming into play simultaneously. Figure 10.15 shows how complex it would be to try to make any prediction for a rice pile.

Per Bak and his colleagues constructed a pleasant computer simulation which helps explain the behaviour of rice and sand piles. The model consists of a 5×5 grid of squares

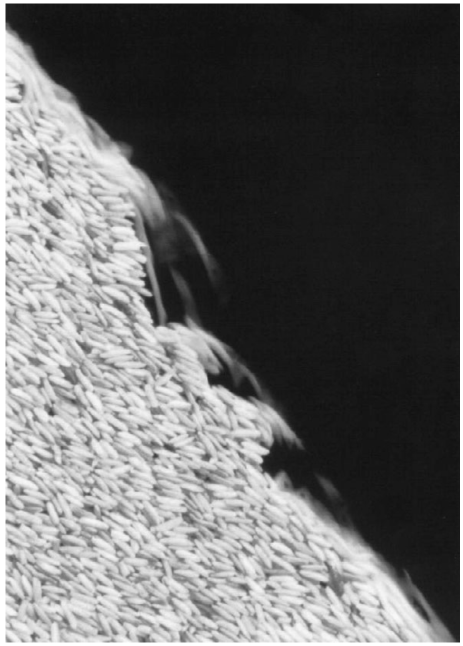


Fig. 10.15 Illustrating the mini-avalanches involved in the formation of a rice pile. (From P. Bak, 1997. *How Nature Works: The Science of Self-Organised Criticality*, Plate 4. Oxford: Oxford University Press.)

and to begin with 0, 1, 2 or 3 grains are placed randomly on each square. Then, grains are added randomly to the grid with the following rules. When a square has four grains, these are redistributed, one to each of the four immediately neighbouring squares and the number on the original square is set to zero. If this process results in any of the neighbouring squares having four grains, the process is repeated, and this sequence is repeated over and over again, until there are no more than three grains on each square – at this point, the avalanche has stopped. Figure 10.16 illustrates this process for a specific initial configuration of grains. One grain is added to the central square and then the evolution takes place through the next seven iterations as shown. In this particular example, it can be seen that the grain added to the centre results in nine toppling events in seven iterations of the algorithm. The filled area of eight squares in the centre of the grid at the bottom of Fig. 10.16 shows those squares which have suffered at least one toppling event.

Exactly the same procedure has been repeated for much larger arrays such as the 50×50 grid shown in Fig. 10.17. The grey central area represents boxes in which at least one toppling event has taken place. The four frames a, b, c and d represent a single avalanche at different stages in its evolution.

Although this is a highly idealised model, it has a number of suggestive features.

- If the simulations are run many times, the probability distribution of the sizes of the avalanches can be found and this turns out to be of power-law form, of exactly the same form as the Gutenberg–Richter law for earthquakes.
- The rice pile becomes ‘critical’ without any fine-tuning of the parameters.

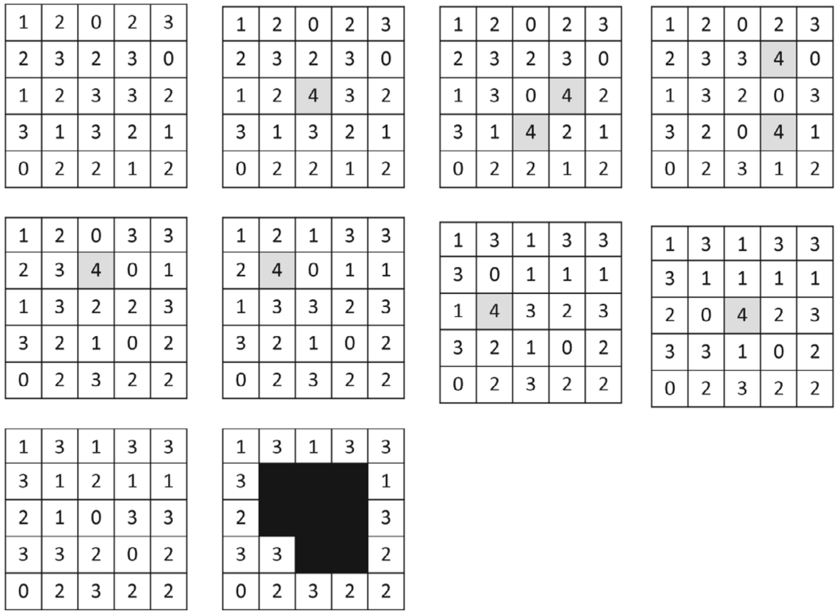


Fig. 10.16 Illustrating the evolution of a mini-avalanche in a computer model of a rice-pile. (From P. Bak, 1997. *How Nature Works: The Science of Self-Organised Criticality*, p. 53. Oxford: Oxford University Press.)

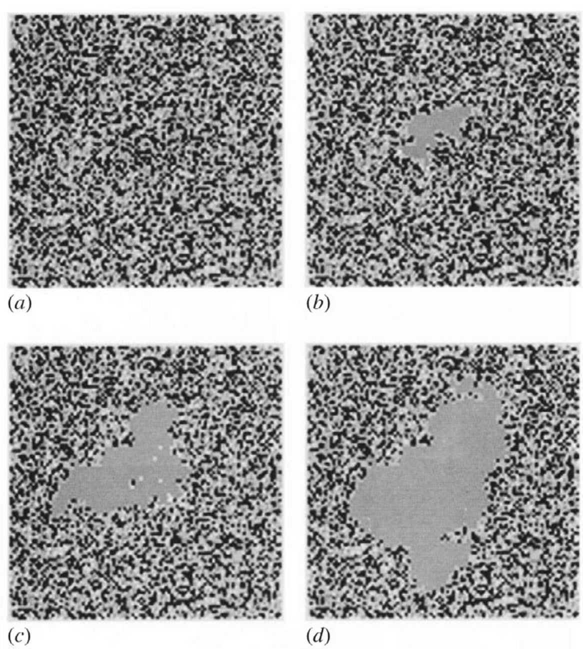


Fig. 10.17 Illustrating the evolution of a mini-avalanche in a computer model of a rice-pile on a 50×50 grid. The central grey areas in these frames show the evolution of the area in which at least one toppling event has occurred. (From P. Bak, 1997. *How Nature Works: The Science of Self-Organised Criticality*, Plate 1. Oxford: Oxford University Press.)

- The outline of the perimeter of the rice pile is very ragged and the same type of analysis used to show that the coastline of Norway is fractal shows that the outline of the rice pile is also fractal.
- The rules governing the redistribution of the rice grains can be changed quite significantly, but the same type of generic behaviour is found.
- It has not been possible to find an analytic model for any self-organised critical system.

The last bullet point is particularly significant. Problems of this type could only be tackled once high-speed computers became available so that very large numbers of simulations could be run and the computational tools made available to analyse the data. The same types of analysis have been used to study the acoustic emission from volcanoes, crystal, cluster or colloid growth, forest fires, cellular automata and so on. The characteristic of these self-organised critical phenomena is that they involve massively non-linear phenomena and yet end up producing scaling laws with no preferred scales. Bak makes the case that this type of behaviour is central to the processes of evolution in nature.

10.5 Beyond Computation

It might seem that this would be about as far as one could go in a physics text about the use of computation to solve complex problems and yet there might be even more extreme examples which go beyond computation. These thoughts are prompted by Roger Penrose's ideas expounded in *The Emperor's New Mind*[20] and *Shadows of the Mind*.[21] The inspiration comes from a somewhat unexpected source and involves the problem of tiling an infinite plane by small numbers of tiles of fixed geometric shapes. It is obvious that an infinite plane can be tiled by simple tiles of regular geometry, for example, by squares, triangles, and

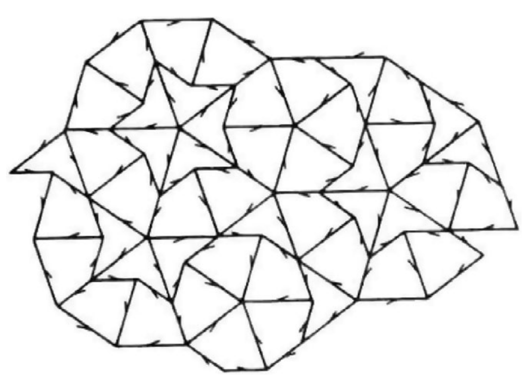


Fig. 10.18 Illustrating tiling of a plane by two regular figures which Penrose calls 'kites' and 'darts'. An infinite plane can be tiled by there two shapes and the tiling is non-periodic. (From R. Penrose, 1989. *The Emperor's New Mind*, p. 178. Oxford: Oxford University Press.)

so on. What is much less obvious is that there exist shapes which can tile an infinite plane without leaving any holes, but which have the extraordinary property that they are *non-periodic*. This means that the pattern of tiles never repeats itself anywhere on an infinite plane. I find this a quite staggeringly non-intuitive result.

Penrose has been able to reduce the numbers of shapes required to tile an infinite plane non-periodically to only two. One example of a pair of tiles which have this property is shown in Fig. 10.18. Penrose refers to the two shapes as ‘kites’ and ‘darts’. The strange patterns seen in Fig. 10.18 continue to infinity and the configuration of the tilings cannot be predicted at any point in the plane since the patterns never recur in exactly the same way. Furthermore, how would a computer be able to establish whether or not the tiling was periodic or not? It might be that analogues of the tiling problem turn up in different aspects of physical problems. It will be intriguing to find out if indeed there are problems in physics which end up being truly non-computational.

Notes

1 Bondi, H. (1983). The Making of a Scientist, *Journal of the Royal Society of Arts*, June 1983, 406.
2 Bondi, H. (1988). *Let Newton Be: A New Perspective on his Life and Works*, J. Fauvel *et al.* (eds.), pp. 245-246. Oxford: Oxford University Press.
3 Buckingham, E. (1914). On Physically Similar Systems: Illustrations of the Use of Dimensional Equations, *Physical Review*, **4**, 345–376. doi = 10.1103/PhysRev.4.345.
4 I am indebted to Sanjoy Mahajan for introducing me to the remarkable power of this theorem. His book with Peter Goldreich and Strel Phinney, *Order of Magnitude Physics* (1999), is strongly recommended. It can be found at: www.inference.org.uk/sanjoy/oom/book-a4.pdf.
5 Taylor, G.I. (1950). The Formation of a Blast Wave by a Very Intense Explosion. I. Theoretical Discussion, *Proceedings of the Royal Society of London A*, **201**, 159–174.
6 Taylor, G.I. (1950). The Formation of a Blast Wave by a Very Intense Explosion. II. The Atomic Explosion of 1945, *Proceedings of the Royal Society of London A*, **201**, 175–186.
7 Mack, J.E. (1947). *Semi-popular motion picture record of the Trinity explosion*, MDDC221, US Atomic Energy Commission.
8 Barenblatt, G.I. (1995). *Scaling, Self-similarity, and Intermediate Asymptotics*. Cambridge: Cambridge University Press.
9 See, for example, D. Tabor (1991). *Gases, Liquids and Solids and Other states of Matter*, p. 125. Cambridge: Cambridge University Press.
10 Gleick, J. (1987). *Chaos: Making a New Science*. New York: Viking.
11 Baker, G.L. and Gollub, J.P. (1996). *Chaotic Dynamics: An Introduction*. Cambridge: Cambridge University Press.
12 From television interview in the programme *Chaos*, produced by Chris Haws for Channel 4 by InCA (1988).
13 R.M. May (1976). Simple Mathematical Models with Very Complicated Dynamics, *Nature*, **261**, 459–467.
14 Verhulst, P.-F. (1838). Notice sur la loi que la population suit dans son accroissement. *Correspondance mathématique et physique*, **10**, 113–121.

15 R.M. May (1972). Limit Cycles in Predator–Prey Communities, *Science*, **177**, 900–902.
16 M.J. Feigenbaum. Quantitative Universality for a Class of Non-Linear Transformations, *Journal of Statistical Physics*, **19**, 25–52. doi:10.1007/BF01020332.
17 Bak, P. (1997). *How Nature Works: The Science of Self-Organised Criticality*. Oxford: Oxford University Press.
18 Mandelbrot, B. (1983). *The Fractal Geometry of Nature*. New York: Freeman.
19 Zipf, G.K. (1949). *Human Behaviour and the Principle of Least Effort*. Cambridge, Massachusetts: Addison-Wesley Publishers.
20 Penrose, R. (1989). *The Emperor's New Mind*. Oxford: Oxford University Press.
21 Penrose, R. (1994). *Shadows of the Mind*. Oxford: Oxford University Press.

Case Study IV

THERMODYNAMICS AND STATISTICAL PHYSICS

Thermodynamics is the science of how the properties of matter and systems change with temperature. The system may be viewed on the *microscopic scale*, in which case we study the interactions of atoms and molecules and how these change with temperature. In this approach, we need to construct physical models for these interactions. The opposite approach is to study the system on the *macroscopic scale* and then the unique status of *classical thermodynamics* becomes apparent. In this approach, the behaviour of matter and radiation *in bulk* is studied and we *quite explicitly deny that they have any internal structure at all*. In other words, the science of classical thermodynamics is solely concerned with relations between macroscopic measurable quantities.

Now this may seem to make classical thermodynamics a rather dull subject but, in fact, it is quite the opposite. In many physical problems, we may not know in detail the correct microscopic physics, and yet the thermodynamic approach can provide answers about the macroscopic behaviour of the system which is independent of the unknown detailed microphysics. Another way of looking at the subject is to think of classical thermodynamics as providing the boundary conditions which any microscopic model must satisfy. The thermodynamic arguments have validity independent of the model adopted to explain any particular phenomenon.

It is remarkable that these profound statements can be made on the basis of the *two laws of thermodynamics*. Let us state them immediately. The *first law of thermodynamics* is a statement about the conservation of energy:

> Energy is conserved when heat is taken into account.

The *second law of thermodynamics* tells us how thermodynamic systems evolve with time. This can be deduced from the statement of the law due to Clausius:

> No process is possible whose sole result is the transfer of heat from a colder to a hotter body.

These laws are no more than reasonable hypotheses formulated on the basis of practical experience. They prove, however, to be hypotheses of the greatest power. They have been applied to myriads of phenomena and time and again have proved to be correct. Examples include the extreme physical conditions such as matter in bulk at nuclear

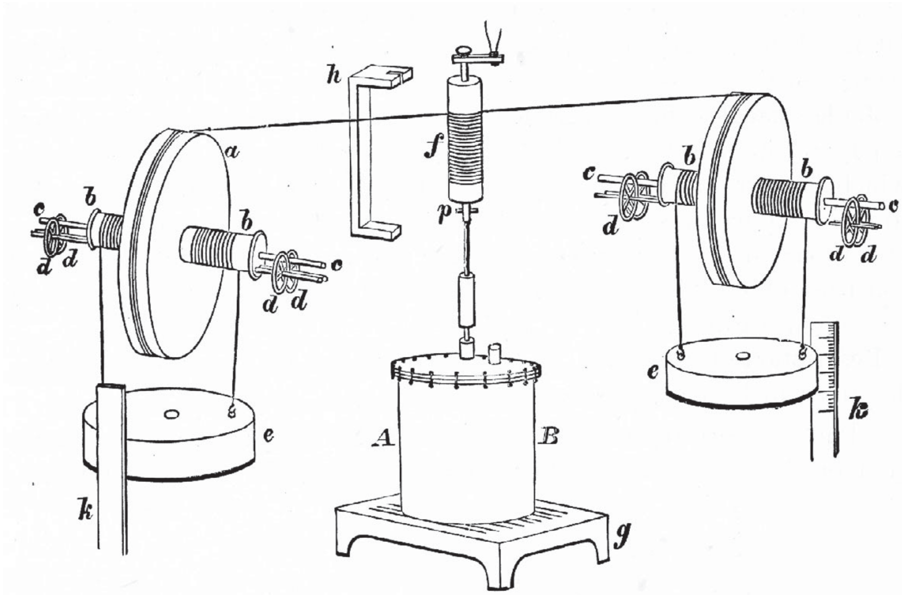


Fig. IV.1 Joule's famous paddle wheel experiment with which he determined the mechanical equivalent of heat. (From J.P. Joule, *Philosophical Transactions of the Royal Society*, 1850, **140**, opposite p. 64.)

densities inside neutron stars and in the early stages of the Big Bang model of the Universe ($\rho \sim 10^{18}$ kg m^{-3}), and at ultra-low temperatures in laboratory experiments. There is no way in which the laws of thermodynamics can be proved. Just like Newton's laws of motion, they are axioms which express our common experience of the thermal properties of matter and radiation. It is intriguing that, when Max Planck was struggling to understand the nature of the spectrum of black-body radiation in 1900, he stated:

> The two laws (of thermodynamics), it seems to me, must be upheld under all circumstances. For the rest, I was ready to sacrifice every one of my convictions about the physical laws.[1]

Indeed, the two laws of thermodynamics have never been found to be wanting in any physical problem.

The objective of Chapter 11 is twofold. First of all, the history of the discovery of the laws of thermodynamics illuminates a number of conceptual problems which confronted the nineteenth century pioneers. The resolution of these problems helps clarify concepts such as heat, energy, work, and so on. Second, when we study the nature of the second law and the concept of entropy, the early history helps explain how Carnot came to his understanding of the importance of studying perfect heat engines. For much of the argument, we deal with the *ideal* behaviour of perfect machines or systems, and then contrast that with what happens in the real imperfect world.

[1] Planck, M. (1931). Letter from M. Planck to R.W. Wood. See Hermann, A. (1971), *The Genesis of Quantum Theory (1899–1913)*, pp. 23–24. Cambridge, Massachusetts: MIT Press.

We should distinguish this *thermodynamic* approach from that of *model building*, in which we seek to interpret the nature of the laws in terms of microscopic processes. While Chapter 11 will remain strictly *model-free*, in Chapter 12 we study the *kinetic theory of gases* as enunciated by Clausius and Maxwell and the *statistical mechanics* of Boltzmann and Gibbs. These theories are explanatory in a sense that thermodynamics is not. For many students, thermodynamics is more accessible through statistical physics than through the subtleties of classical thermodynamics. In many applications, these statistical theories are very successful indeed and yet there remain problems which are too complicated to be treated by well-understood physics. A good example is the case of high temperature superconductivity, in which the physics of strongly correlated electrons requires the development of new theoretical insights which are not yet available. Then the classical thermodynamic approach comes into its own.

It comes as no surprise to learn that, in fact, the science of classical thermodynamics developed in parallel with advances in the understanding of the atomic and molecular interpretation of the properties of matter. We find many of the same pioneers of classical thermodynamics – Thomson, Clausius, Maxwell and Boltzmann – reappearing in the history of the understanding of the microscopic structure of solids, liquids and gases. We will tell two stories in Chapter 12 – the kinetic theory of gases, and then the way in which the successes of that theory led to statistical mechanics. In fact, these two approaches form part of one single story because, despite the fact that there were good reasons why the kinetic theory was by no means readily accepted, it clearly suggested that, as first recognised by Maxwell, the law of increase of entropy is in fact a statistical result.

The development of these concepts is admirably described by Peter Harman in his book *Energy, Force and Matter: The Conceptual Development of Nineteenth Century Physics*.[2] I should also confess to deriving enormous pleasure from an old textbook, which I picked up second-hand a long time ago when I was in my last year at school, *A Text-book of Heat* by H.S. Allen and R.S. Maxwell.[3] The book is charming and full of splendid historical anecdotes, as well as presenting beautifully clearly the experimental ingenuity which went into establishing many of the results we now take for granted. It is a shame that books like this are no longer published.

[2] Harman, P.M. (1982). *Energy, Force and Matter: The Conceptual Development of Nineteenth Century Physics*. Cambridge: Cambridge University Press.

[3] Allen, H.S. and Maxwell, R.S. (1939). *A Text-book of Heat*, Parts I and II. London: MacMillan & Co.

11 Basic Thermodynamics

11.1 Heat and Temperature

Like so many aspects of the revolution which led to the birth of modern science, the origins of the scientific study of heat and temperature can be traced to the early years of the seventeenth century.[1] The quantitative study of heat and temperature depended upon the development of instruments which could give a quantitative measure of concepts such as 'the degree of hotness'. The first appearance of the word *thermomètre* occured in 1624 in the volume *Récréaction Mathématique* by the French Jesuit Jean Leurechon. This was translated into English in 1633 by William Oughtred, who wrote 'Of the thermometer, or an instrument to measure the degrees of heat or cold in aire'. The *Oxford English Dictionary* gives 1633 as the date of the first appearance of the word *thermometer* in the English literature.

There had been earlier descriptions of the use of the expansion of gases to measure temperature by Galileo and others, but the first thermometers which resemble their modern counterparts were constructed in the 1640s to 1660s. Coloured alcohol within a glass bulb was used as the expanding substance and this liquid extended into the long stem of the thermometer, which was sealed off to prevent evaporative loss of the fluid. In 1701, Daniel Fahrenheit constructed an alcohol in glass thermometer in which the sealed tube was evacuated. Then, in 1714, he extended the temperature range of the thermometer by using mercury rather than fluids such as alcohol, which boiled at relatively low temperatures. In the period 1714 to 1724, he determined the fixed points of what became known as the *Fahrenheit temperature scale* as the boiling point of water at 212 °F and the temperature of melting ice as 32 °F. In 1742, Anders Celsius, Professor of Astronomy and founder of the Uppsala Observatory in Sweden, presented his mercury in glass thermometer to the Swedish Academy of Sciences adopting a centigrade temperature scale, in which the temperature of melting ice was taken to be 0 °C and the boiling point of water 100 °C. The C is referred to as *degrees Celsius* and has been adopted internationally as the standard SI unit of temperature.

Joseph Black, who was appointed Professor of Chemistry and Medicine, first in Glasgow and then in Edinburgh, was a key figure in unravelling the nature of heat and temperature. It is intriguing that one of Black's assistants, who helped carry out his experiments in the 1760s, was James Watt, who will play a central role in the subsequent story. Black's great achievements included distinguishing quantitatively between the concepts of heat and temperature and defining what we now refer to as *specific heat capacities*. He also discovered and defined the concept of *latent heat*.

The experiments he undertook were simple but effective. To quantify the study of heat, Black used the *method of constant heat supply*, in which heat is supplied at a constant rate to a body and its temperature rise measured. By careful experimentation, Black showed that the heat Q supplied to a mass m of a substance is proportional to the temperature difference ΔT, that is, $Q = mc\Delta T$, where c is defined to be the *specific heat capacity* of the substance, the word *specific* meaning, as usual, per unit mass. His great realisation was that the specific heat capacities are different for different materials, and this provided the basis for the development of the mathematical theory of heat.

In his study of the latent heat associated with the melting of ice, he noted that the only effect of adding heat to the ice was to convert it into liquid water *at the same temperature*. As he expressed it, the heat that entered the water was

> absorbed and concealed within the water, so as not to be discoverable by the application of a thermometer.[2]

In one of his classic experiments, he set up in a large empty hall two glass vessels, each four inches in diameter. The first contained ice at its melting point and the other water at the same temperature. After half an hour, the water had increased in temperature by 7 °F, whereas the ice in the second vessel was only partially melted. In fact, the ice in the second vessel was only completely melted after ten and a half hours. These data can be used to show that the latent heat of melting of ice is 80 cal gm^{-1}. Similar ingenuity was used in determining the latent heat of steam as it condenses into water at its boiling point.

11.2 Heat as Motion *versus* the Caloric Theory of Heat

Towards the end of the eighteenth century, a major controversy developed concerning the nature of heat. The idea of heat as a form of motion can be traced back to the atomic theory of the Greek philosopher Democritus in the fifth century BC, who considered matter to be made up of indivisible particles whose motions gave rise to the myriad different forms in which matter is found in nature.

These ideas found resonances in the thoughts of scientists and philosophers during the seventeenth century. Francis Bacon wrote, for example:

> Heat itself, its essence and quiddity, is motion and nothing else.[3]

This view was affirmed by Robert Boyle who was persuaded of the kinetic, or dynamic, theory by the heat generated when a nail is hammered into a block of wood. He concluded that the action of the hammer impressed:

> a vehement and variously determined agitation of the small parts of the iron, which, being a cold body before, by that superinduced commotion of all its small parts, becomes in divers senses hot.[4]

These opinions were shared by Thomas Hooke who referred to fluidity as being:

> nothing else but a certain pulse or shake of heat; for heat being nothing else but a very brisk and vehement agitation of the parts of a body, the parts of the body are thereby made so loose from one another, that they may easily move any way, and become a fluid.[5]

Similar views were held by Isaac Newton, Gottfried Leibniz, Henry Cavendish, Thomas Young and Humphry Davy, as well as by the seventeenth century philosophers Thomas Hobbes and John Locke.

During the eighteenth century, however, an alternative picture was developed, its origin being derived from the concept of *phlogiston*. The idea was that this hypothetical element could be combined with a body to make it combustible. Following the discovery of oxygen by Joseph Priestley in 1774, Antoine-Laurent Lavoisier, the father of modern chemistry, demonstrated that the process of combustion should rather be regarded as the process of *oxidation* and that it was unnecessary to involve phlogiston in the process. Nonetheless, there was a strong body of opinion that heat should be considered a physical substance. Lavoisier and others named this 'matter of heat' *caloric*, which was to be thought of as an 'imponderable fluid', that is, a massless fluid, which was conserved when heat flowed from one body to another. Its flow was conceived as being similar to the flow of water in a waterwheel, in which mass is conserved. To account for the diffusion of heat, caloric was assumed to be a self-repellent, elastic fluid, although it was attracted by matter. If one body were at a higher temperature than another and they were brought into thermal contact, caloric was said to flow from the hotter to the cooler body until they came into equilibrium at the same temperature. The hotter the body, the more caloric it contained and hence, because it was self-repellent, its volume would increase, accounting for its thermal expansion. In his *Traite Élémentaire de Chimie* of 1789, Lavoisier included *Caloric* in his Table of Simple Substances.[6]

There were, however, problems with the caloric theory – for example, when a warm body is brought into contact with ice, caloric flows from the warm body into the ice but, although ice is converted into water, the temperature of the ice–water mixture remains the same. It had to be supposed that caloric could combine with ice to form water. The theory attracted the support of a number of distinguished scientists including Joseph Black, Pierre-Simon Laplace, John Dalton, John Leslie, Claude Berthollet and Johan Berzelius. When Sadi Carnot came to formulate his pioneering ideas on the operation of perfect heat engines, he used the caloric theory to describe the flow of heat in a heat engine.

The two theories came into conflict at the end of the eighteenth century. Among the most important evidence against the caloric theory were the experiments of Count Rumford. Rumford was born Benjamin Thomson in the small rural town of Woburn in Massachusetts, USA, and had a quite extraordinary career, nicely summarised in the title of G.I. Brown's biography *Scientist, Soldier, Statesman, Spy: Count Rumford*.[7] In addition to these accomplishments, he was a founder of the Royal Institution of Great Britain and contracted an unhappy marriage to Lavoisier's widow. Despite his many practical contributions to society, he was not a well-liked figure. Brown quotes contemporary views of him as 'unbelievably cold-blooded, inherently egotistic and a snob', while another summarises his character as 'the most unpleasant personality in the history of science since Isaac Newton'.

Rumford's great contribution to physics was his realisation that heat could be created in unlimited quantities by the force of friction. This phenomenon was mostly vividly demonstrated by his experiments of 1798 in which cannons were bored with a blunt drill (Fig. 11.1). In his most famous experiment, he heated about 19 pounds of water in a box

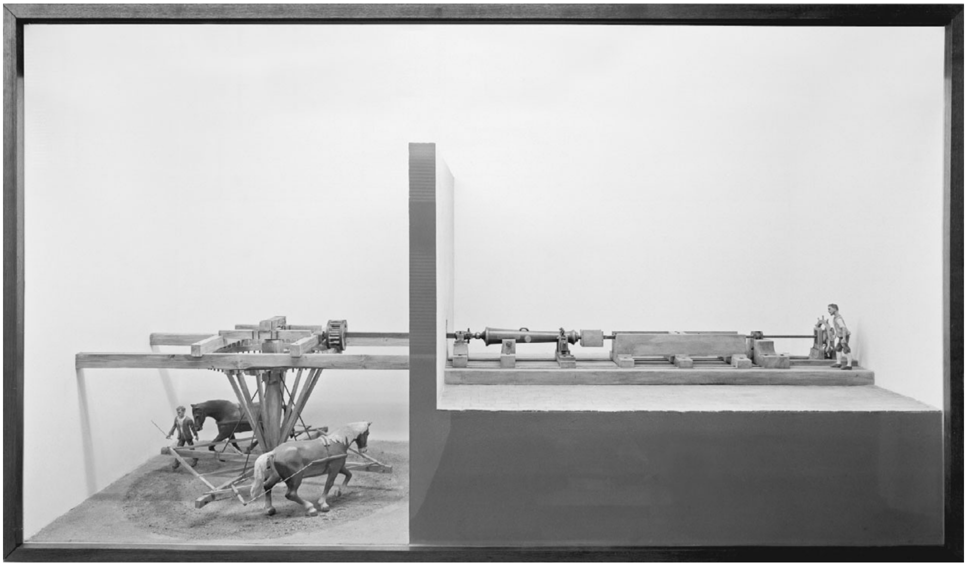


Fig. 11.1 A model of Rumford's cannon boring experiments carried out at the arsenal at Munich in 1797–98. (Courtesy of the Deutsches Museum, Munich.)

surrounding the cylinder of the cannon and after two and a half hours, the water actually boiled. In his own inimitable style, he remarked:

> It would be difficult to describe the surprise and astonishment expressed in the countenances of the bystanders, on seeing so large a quantity of water heated, and actually made to boil, without any fire. ...
>
> In reasoning on this subject, we must not forget *that most remarkable circumstance*, that the source of heat generated by friction in these experiments appeared evidently to be *inexhaustible*. It is hardly necessary to add, that anything which any *insulated* body or system of bodies can continue to furnish *without limitation* cannot possibly be a material substance; it appears to me to be extremely difficult, if not quite impossible, to form any distinct idea of anything capable of being excited and communicated in those experiments, except it be *MOTION*.[7]

The summit of Rumford's achievement was the first determination of the mechanical equivalent of heat, which he found from the amount of work done by the horses in rotating the drilling piece. In modern terms, he found a value of 5.60 J cal^{-1} in 1798 – it was to be more than 40 years before Mayer and Joule returned to this problem.

Among the contributions which were to prove to be of key importance was Fourier's treatise *Analytical Theory of Heat*[8] published in 1822. In this volume, Fourier worked out the mathematical theory of heat transfer in the form of differential equations which did not require the construction of any specific model for the physical nature of heat. Fourier's methods were firmly based in the French tradition of rational mechanics and gave mathematical expression to the *effects* of heat without enquiring into its causes.

The idea of conservation laws, as applied to what we would now call energy, for mechanical systems had been discussed in eighteenth century treatises, in particular, in the work of Leibniz. He argued that 'living force' or *'vis viva'*, which we now call kinetic energy, is conserved in mechanical processes. By the 1820s, the relation of the kinetic energy to the work done was clarified.

In the 1840s, a number of scientists independently came to the correct conclusion about the interconvertibility of heat and work, following the path pioneered by Rumford. One of the first was Julius Mayer, who in 1840 was appointed ship's doctor on the three-masted ship *Java* for a voyage to the Far East. He had plenty of time for reflection on scientific matters, among these being the production of heat in living organisms. At Surabaya, several of the crew fell ill and when he bled them, their venous blood was much redder in the tropics than it was in more moderate climes. The locals claimed that this was due to the need for less oxygen to maintain the body at its normal temperature than in colder climates. Mayer reasoned that the energy produced by the food one eats is partly converted into heat to maintain the body temperature, and partly into the work done by one's muscles. On returning to Europe, Mayer proposed the equivalence of work and energy in 1842 on the basis of experiments on the work done in compressing and heating a gas. He elevated his results to the principle of conservation of energy and derived an estimate for the mechanical equivalent of heat from the adiabatic expansion of gases of 3.58 J cal^{-1}. Mayer had great difficulty in getting his work published and his pioneering works were somewhat overshadowed by those of James Joule, who claimed that Mayer's work was 'an unsupported hypothesis, not an experimentally established quantitative law'. Joule proceeded to carry out meticulous experiments to test the hypothesis.

James Prescott Joule was born into a family which had grown wealthy through the foundation of a brewery by his grandfather. Joule's pioneering experiments were carried out in a laboratory which his father had built for him next to the brewery. His genius was as a meticulous experimenter and the reason for singling out his name from among the other pioneers of the 1840s is that he put the science of heat on a firm experimental basis. Perhaps the most important aspect of his work was his ability to measure accurately very small temperature changes in his experiments.

It is a delight to track Joule's dogged progress towards a reliable estimate of the mechanical equivalent of heat.[9] For the sake of historical authenticity, let us depart from strict adherence to SI units and quote Joule's results in the units he used. Thus, the unit of thermal energy is the British Thermal Unit (BTU), which is the heat needed to raise 1 pound of water by 1°F at standard temperature and pressure (STP). The unit of mechanical energy was the foot-poundal. For reference, the present value of the mechanical equivalent of heat is 4.1868 J cal^{-1}, corresponding to 778.2 foot-poundal BTU^{-1}.

1843 In Joule's dynamo experiments, the work done in driving a dynamo was compared with heat generated by the electric current, and resulted in a publication entitled, *Production of Heat by Voltaic Electricity*. In his earliest experiments, he established that the quantity of heat produced is proportional to RI^2 where R is the resistance and I the current. By 1843, he was able to derive from these electrical experiments a value for the mechanical equivalent of heat of 838 foot-poundal BTU^{-1}.

1843 In a postscript to his paper, Joule remarked that, in other experiments, he had shown that 'heat is evolved by the passage of water through narrow tubes' and quoted a result of 770 foot-poundal BTU^{-1}. These experiments hint at the origin of his most famous experiments, the paddle wheel experiments.

1845 Joule carried out his own version of Mayer's experiment in which the change of temperature in the adiabatic expansion of air was measured. A volume of air was compressed and then the energy loss measured when air was allowed to expand. By this means, he estimated the mechanical equivalent of heat to be 798 foot-poundal BTU^{-1}.

1847 In 1845, Joule realised:

> If my views be correct, a fall [of a cascade] of 817 feet will of course generate one degree of heat, and the temperature of the river Niagara will be raised about one-fifth of a degree (Fahrenheit) by its fall of 160 feet.[10]

The delightful sequel to this proposal was recounted by William Thomson after he had met Joule at the Meeting of the British Association for the Advancement of Science in Oxford in 1847. In Thomson's words:

> About a fortnight later, I was walking down from Chamounix to commence the tour of Mont Blanc, and whom should I meet walking up but Joule, with a long thermometer in his hand, and a carriage with a lady in it not far off. He told me that he had married since we parted at Oxford and he was going to try for the elevation of temperature in waterfalls.[11]

Prospective spouses of physicists should be warned that this is by no means an isolated example of physicists doing physics during their honeymoons.

1845–78 The first paddle wheel experiment was carried out in 1845, a drawing of the experimental arrangement being published in 1850 (Fig. IV.1). The work done by the weights in driving the paddle wheel goes into heat through the frictional force between the water and the vanes of the paddle wheel. In 1845, he presented his preliminary value of 890 foot-poundal BTU^{-1}. Improved estimates of between 781 and 788 foot-poundal BTU^{-1} were reported in 1847. Yet more refined versions of the paddle wheel experiments continued until 1878, when Joule gave his final result of 772.55 foot-poundal BTU^{-1}, corresponding to 4.13 J cal^{-1} – this can be compared with the adopted modern value of 4.1868 J cal^{-1}.

It was the results of the paddle wheel experiments which greatly excited William Thomson in 1847 when he was only 22. At the same Oxford meeting referred to above, Thomson heard Joule describe the results of his recent experiments on the mechanical equivalent of heat. Up till that time, he had adopted the assumption made by Carnot that heat (or caloric) is conserved in the operation of heat engines. He found Joule's results astonishing and, as his brother James wrote, Joule's 'views have a slight tendency to unsettle one's mind'.[12] Joule's results became widely known in continental Europe and, by 1850, Helmholtz and Clausius had formulated what is now known as the *law of conservation of energy*, or the *first law of thermodynamics*. Helmholtz, in particular, was the first to express the conservation laws in mathematical form in such a way as to incorporate mechanical and electrical phenomena, heat and work. In 1854, Thomson invented the word *thermo-dynamics* to describe the new science of heat to which he was to make such major contributions.

11.3 The First Law of Thermodynamics

The aim of this section is to describe the content of the first law as clearly as possible. My approach to both laws has been strongly influenced by Brian Pippard's *The Elements of Classical Thermodynamics*,[13] which can be warmly recommended for the clarity and depth of its exposition.

11.3.1 The Zeroth Law and the Definition of Empirical Temperature

To begin with, we need to establish what is meant by the term *empirical temperature* and how it follows from what is called the *zeroth law of thermodynamics*. Immediately, we encounter one of the distinctive features of thermodynamic reasoning. In many of these, statements are made of the form 'It is a fact of experience that . . .' and from such axioms the appropriate mathematical structures are derived.

We begin with the thermal properties of *fluids*, meaning liquids or gases, because the pressure is isotropic at all points in the fluid and also because we can change the shape of the containing vessel without changing its volume and so no work is done. At this stage, we have no definition of temperature. We now state our first axiom:

> It is a fact of experience that the properties of a fluid are entirely determined by only two properties, the pressure p and the volume V of the vessel.

It is assumed that the fluid is not subjected to other influences, such as placing it in an electric or magnetic field. If the system is wholly defined by two properties, it is known as a two-coordinate system. Most of the systems we will deal with are two-coordinate systems, but it is relatively straightforward to develop the formalism for multi-coordinate systems.

Notice carefully what this assertion means. Support a certain volume of fluid is taken through a sequence of processes such that it ends up having pressure p_1 and volume V_1. Then, suppose we take another identical quantity fluid and perform a totally different set of operations so that it also ends up with coordinates p_1 and V_1. We assert that these two fluids are totally indistinguishable in their physical properties.

Now take two isolated systems consisting of fluids with coordinates p_1, V_1 and p_2, V_2. They are brought into thermal contact and left for a very long time. In general, they will change their properties such that they reach a state of *thermodynamic equilibrium*, in which there is no net heat transfer between them. The concept of thermodynamic equilibrium is crucial – all components which make up the system have been allowed to interact thermally until after a very long time no further changes in the bulk properties of the system are found. In attaining this state, in general, heat is exchanged and work is done. In thermodynamic equilibrium, the thermodynamic coordinates of the two systems are p_1, V_1 and p_2, V_2. Now, it is clear that these four values cannot be arbitrary. It is a fact of experience that two fluids cannot have arbitrary values of p_1, V_1 and p_2, V_2 and be in thermodynamic equilibrium. There must be some mathematical relation between the four quantities in thermodynamic equilibrium and so

$$F(p_1, V_1, p_2, V_2) = 0.$$

This equation tells us the value of the fourth quantity if the other three are known.

So far we have not used the word *temperature*. We will find a suitable definition from another common fact of experience, which is so central to the subject that it has been raised to the status of a law of thermodynamics – the *zeroth law*. The formal statement is:

> If two systems, 1 and 2, are separately in thermal equilibrium with a third, 3, then they must also be in thermal equilibrium with each another.

We can write this down mathematically. System 1 being in equilibrium with system 3 means that

$$F(p_1, V_1, p_3, V_3) = 0,$$

or, expressing the pressure p_3 in terms of the other variables,

$$p_3 = f(p_1, V_1, V_3).$$

System 2 being in equilibrium with system 3 means, in the same way, that

$$p_3 = g(p_2, V_2, V_3).$$

Therefore, in thermodynamic equilibrium,

$$f(p_1, V_1, V_3) = g(p_2, V_2, V_3). \tag{11.1}$$

But the zeroth law tells us that 1 and 2 must also be in equilibrium, and consequently there must exist a function such that

$$h(p_1, V_1, p_2, V_2) = 0. \tag{11.2}$$

Equation (11.2) tells us that (11.1) must be of such a form that the dependence on V_3 cancels out on either side, that is, it must be a function of the form $f(p_1, V_1, V_3) = \phi_1(p_1, V_1)\zeta(V_3)+\eta(V_3)$. Therefore, substituting into (11.1), we find

$$\phi_1(p_1, V_1) = \phi_2(p_2, V_2) = \phi_3(p_3, V_3) = \theta = \text{constant}, \tag{11.3}$$

in thermodynamic equilibrium. This is the logical consequence of the zeroth law – there exists a function of p and V which takes a constant value for all systems in thermodynamic equilibrium with each other at a given temperature. The different equilibrium states are characterised by different constants. The constant which characterises the equilibrium state is called a *function of state*, that is, it is a quantity which takes a definite value for a particular equilibrium state. This constant is defined to be the *empirical temperature* θ. Consequently, we have defined an *equation of state*, which relates p and V to the empirical temperature θ in thermodynamic equilibrium:

$$\phi(p, V) = \theta. \tag{11.4}$$

We can now find from experiment all the combinations of p and V which correspond to a fixed value of the empirical temperature θ. Notice that we have three quantities, p, V and θ which define the equilibrium state, but any two of them are sufficient to define that state completely. Lines of constant θ are called *isotherms*.

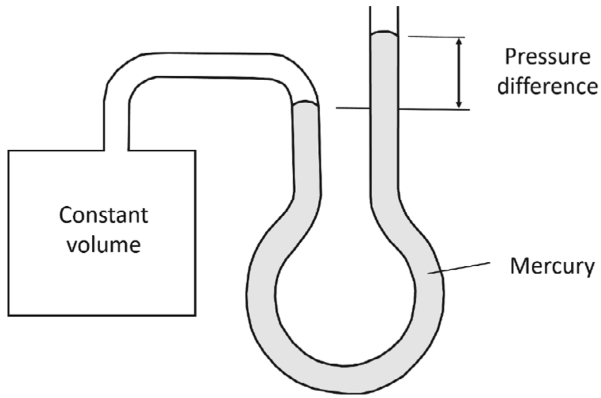


Fig. 11.2 Illustrating the principle of the constant volume gas thermometer. The thermometer makes use of the fact that, at low pressures, the gas follows closely the ideal gas law $pV = RT$ and so, if the volume is fixed, the pressure difference provides a direct measure of the temperature.

At this stage, the empirical temperature looks nothing like what we customarily recognise as temperature. To put the subject on a firm experimental foundation, we need to decide upon a *thermometric scale*. Once this scale is fixed for one system, it will be fixed for all others by virtue of the fact that all systems have the same value of the empirical temperature when they are in thermodynamic equilibrium.

The subject of thermometric scales is a vast one. We will only note the importance of the *constant volume gas thermometer*, illustrated in Fig. 11.2. Although bulky and unwieldy, the constant volume gas thermometer is of special importance because it was found empirically that all gases have the same variation of pressure with volume at low pressures, that is, at a fixed temperature, the product pV is the same for one mole of all gases at low enough pressures. In addition, such gases are found to be, to a very good approximation, *ideal* or *perfect gases* and one can then go further and relate the pressure and volume measures of these gases to the thermodynamic temperature T. The relation between these quantities is the *perfect gas law*,

$$pV = RT, \tag{11.5}$$

where R is the universal gas constant for 1 mole of gas. We will show in Section 11.5.5 that this scale can be related to the thermodynamic scale of temperature, which provides a rigorous theoretical basis for the study of heat and so, in the limit of low pressures, gas thermometers can be used to measure *thermodynamic temperatures* directly. Symbolically, we write

$$T = \lim_{p \to 0} (pV)/R. \tag{11.6}$$

Notice that, in general, p and V are much more complicated functions of T, in particular, at high pressures and close to phase transitions.

11.3.2 The Mathematical Expression of the First Law of Thermodynamics

We have already stated the *first law of thermodynamics*:

Energy is conserved if heat is taken into account.

To give mathematical substance to the law, we must define precisely what we mean by the terms heat, energy and work. The last two are straightforward. The definition of the work done is

$$W = \int_{\boldsymbol{r}_1}^{\boldsymbol{r}_2} \boldsymbol{f} \cdot \mathrm{d}\boldsymbol{r}, \tag{11.7}$$

where the action of the force $\boldsymbol{f}$ is to move the body an increment of distance $\mathrm{d}\boldsymbol{r}$. When work is done on a body, its *energy* is increased. Let us note some of the different ways in which work can be done on a system.

- *Work done in compressing a fluid*. When a fluid is compressed under the action of an external pressure p, the work done is regarded as a positive quantity. Therefore, since the volume decreases when we do mechanical work on the fluid, the amount of (positive) work done is

  $$\mathrm{d}W = -p\,\mathrm{d}V. \tag{11.8}$$

  If the fluid does work by expanding, the work done on the surroundings is positive and the work done by the system negative. Thus, the sign of the volume increment is important.
- *Work done in stretching a wire by* $\mathrm{d}\boldsymbol{l}$: $\mathrm{d}W = \boldsymbol{f} \cdot \mathrm{d}\boldsymbol{l}$.
- *Work done by an electric field* $\boldsymbol{E}$ *on a charge* q: $\mathrm{d}W = q\boldsymbol{E} \cdot \mathrm{d}\boldsymbol{r}$
- *Work done against surface tension in increasing the surface area by* $\mathrm{d}A$: $\mathrm{d}W = \gamma\,\mathrm{d}A$, where γ is the coefficient of surface tension.
- *Work done by a couple, or torque,* $\boldsymbol{G}$ *in causing an angular displacement* $\mathrm{d}\boldsymbol{\theta}$: $\mathrm{d}W = \boldsymbol{G} \cdot \mathrm{d}\boldsymbol{\theta}$.
- *Work done on a dielectric per unit volume by an electric field* $\boldsymbol{E}$: $\mathrm{d}W = \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{P}$ where $\boldsymbol{P}$ is the polarisation, that is, the dipole moment per unit volume (see Section 7.11).
- *Work done by a magnetic field* $\boldsymbol{B}$ *per unit volume on a magnetisable medium*: $\mathrm{d}W = \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{M}$, where $\boldsymbol{M}$ is the magnetisation of the medium, that is, its magnetic dipole moment per unit volume (see Section 7.11).

Thus, the work done is the scalar product of a generalised force $\boldsymbol{X}$ and a generalised displacement $\mathrm{d}\boldsymbol{x}$, $\mathrm{d}W = \boldsymbol{X} \cdot \mathrm{d}\boldsymbol{x}$. Notice also that the work done is always the product of an *intensive variable* $\boldsymbol{X}$, by which we mean a property defined at a particular point in the material, and an *extensive variable* $\mathrm{d}\boldsymbol{x}$, which describes the ‘displacement’ of the system under the action of the intensive variable.

Now consider an isolated system in which there is no thermal interaction with the surroundings. It is a fact of experience that, if work is done on the system in some way, the system attains a new equilibrium state and it does not matter how the work is done. For example, we may compress a volume of gas, or stir it with a paddle wheel, or pass an electric current through it. Joule's great contribution to thermodynamics was to demonstrate by precise experiment that this is indeed the case. The result is that, by these very different processes, the energy of the system is increased. We say that there is an increase in the *internal energy* U of the system by virtue of work having been done. Since it does not matter how the work is done, U must be a *function of state*. For this isolated system,

$$W = U_2 - U_1 \qquad \text{or} \qquad W = \Delta U. \tag{11.9}$$

Now suppose the system is not isolated, but there is a thermal interaction between the surroundings and the system. Then, the system will reach a new internal energy state and this will not all be due to work being done. We therefore define the *heat supplied* Q to be

$$Q = \Delta U - W. \tag{11.10}$$

This is our definition of *heat*. It may seem a rather roundabout way of proceeding, but it has the great benefit of logical consistency. It completely avoids specifying exactly what heat is, a problem which was at the root of many of the conceptual difficulties which existed until about 1850 in incorporating heat into the conservation laws.

It is convenient to write (11.10) in differential form:

$$\mathrm{d}Q = \mathrm{d}U - \mathrm{d}W. \tag{11.11}$$

It is also useful to distinguish between those differentials which refer to functions of state and those which do not. $\mathrm{d}U$ is the differential of a function of state, as are $\mathrm{d}p, \mathrm{d}V$ and $\mathrm{d}T$, but $\mathrm{d}Q$ and $\mathrm{d}W$ are not because we can pass from U_1 to U_2 by adding together different amounts of $\mathrm{d}Q$ and $\mathrm{d}W$. We write the differentials of these quantities as $đQ$, $đW$. Thus,

$$đQ = \mathrm{d}U - đW. \tag{11.12}$$

Equation (11.12) is the formal mathematical expression of the *first law of thermodynamics*. We are now in a position to apply the law of conservation of energy as expressed by (11.12) to many different types of problem.

11.3.3 Some Applications of the First Law of Thermodynamics

(i) *Specific heat capacities.* U is a function of state and we have asserted that we can describe the properties of a gas entirely in terms of only two functions of state. Let us express U in terms of T and V, $U = U(V, T)$. Then the differential of U is

$$\mathrm{d}U = \left(\frac{\partial U}{\partial T}\right)_V \mathrm{d}T + \left(\frac{\partial U}{\partial V}\right)_T \mathrm{d}V. \tag{11.13}$$

Substituting into (11.12) and using (11.8),

$$đQ = \left(\frac{\partial U}{\partial T}\right)_V \mathrm{d}T + \left[\left(\frac{\partial U}{\partial V}\right)_T + p\right] \mathrm{d}V. \tag{11.14}$$

We can now define mathematically the concept of the *heat capacity* C, first described experimentally by Joseph Black in the 1760s. At *constant volume*,

$$C_V = \left(\frac{đQ}{dT}\right)_V = \left(\frac{\partial U}{\partial T}\right)_V .$$

At *constant pressure*,

$$C_p = \left(\frac{đQ}{dT}\right)_p = \left(\frac{\partial U}{\partial T}\right)_V + \left[\left(\frac{\partial U}{\partial V}\right)_T + p\right]\left(\frac{\partial V}{\partial T}\right)_p . \tag{11.15}$$

These expressions determine the temperature rise for a given input of heat. Notice that these heat capacities do not refer to any particular volume or mass. It is conventional to use *specific heat capacities* or *specific heats*, where the word 'specific' takes its usual meaning of 'per unit mass'. Conventionally, specific quantitites are written in lower case letters, that is,

$$c_V = C_V/m; \quad c_p = C_p/m. \tag{11.16}$$

Subtracting, we find

$$C_p - C_V = \left[\left(\frac{\partial U}{\partial V}\right)_T + p\right]\left(\frac{dV}{dT}\right)_p . \tag{11.17}$$

The interpretation of this equation is straightforward. The second term in square brackets on the right-hand side describes how much work is done in pushing back the surroundings at constant p. The first term clearly has something to do with the internal properties of the gas because it describes how the internal energy changes with volume. It must be associated with the work done against different types of intermolecular force within the gas. Thus, $C_p - C_V$ provides information about $(dU/dV)_T$.

(ii) *The Joule expansion.* One way of determining the relation between C_p and C_V is to perform a *Joule expansion*, the free expansion of the gas into a larger volume. The system is thermally isolated from its surroundings and so there is no heat exchange. The walls of the vessels are fixed and so no work is done in the free expansion. The careful experiments carried out by Joule in 1845 are shown in Fig. 11.3. Vessel A contained dry air at a pressure

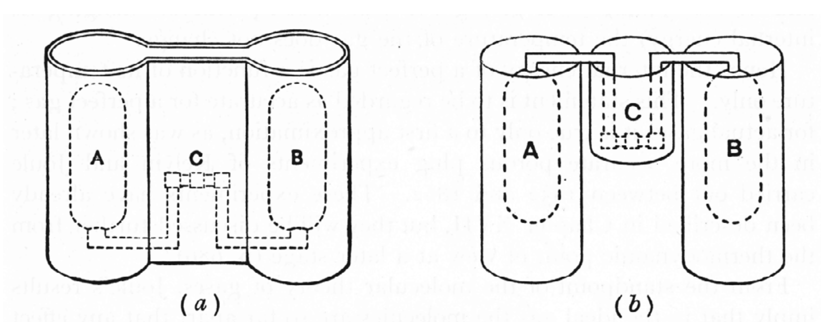


Fig. 11.3 Illustrating Joule's experiments on the expansion of gases. In (*a*), the vessels A, B and the stopcock C are surrounded by a water bath. In (*b*), the vessels and the stopcock each have separate water baths. (From H.S. Allen and R.S. Maxwell, 1939. *A Text-book of Heat,* Part II, p. 575. London: MacMillan & Co.)

of 22 atmospheres and B was evacuated. C was a carefully constructed stopcock. In the first experiment, the whole experiment was enclosed in a water bath (Fig. 11.3(*a*)). When the stopcock was opened, the air flowed from A into B, but although he could measure temperature changes as small as 0.005 °F, he could detect no change in the temperature of the water bath. In his words,

> no change of temperature occurs when air is allowed to expand in such a manner as not to develop mechanical power.[14]

In the second experiment, A, B and C were surrounded by separate water baths (Fig. 11.3(*b*)). When the stopcock was opened, the temperatures of B and C increased, while the temperature of A decreased. This occurred because the gas in A does work in pushing the gas into B. If the gas in A and B were then allowed to come into thermal contact following the expansion and reach a new equilibrium state, there would be no change in the temperature, according to Joule's experiments. Such an expansion is known as a *Joule expansion*. Now, $\Delta U = đQ + đW$. There is no heat transfer in or out of the system, $đQ = 0$, and no external work has been done by the gas $đW = 0$. Therefore, there can have been no change in the internal energy U, that is, $(\mathrm{d}U/\mathrm{d}V)_T = 0$. This is how a *perfect gas* is defined classically:

(*a*) Its equation of state is the perfect gas law, $pV = RT$ for one mole of gas.
(*b*) In a Joule expansion, there is no change in internal energy U.

For such a perfect gas, there is a simple relation between C_p and C_V. From (*b*),

$$\left(\frac{\mathrm{d}U}{\mathrm{d}V}\right)_T = 0.$$

From (*a*),

$$\left(\frac{\partial V}{\partial T}\right)_p = \left[\frac{\partial}{\partial T}\left(\frac{RT}{p}\right)\right]_p = \frac{R}{p}.$$

Therefore, from (11.17),

$$C_p - C_V = R. \tag{11.18}$$

The importance of this result is that it shows that the internal energy of a perfect gas is *a function only of temperature*. The Joule expansion demonstrates that the internal energy is independent of volume. It must also be independent of pressure, because the pressure certainly decreases in a Joule expansion. Since the state of the system is entirely determined by these two functions of state, it follows that the internal energy is only a function of temperature for a perfect gas.

For real imperfect gases, however, there is a change in U, and consequently T, with volume – physically, this is because work is done against the intermolecular van der Walls forces. Also, at very high pressures, there is an effective repulsive force because the molecules cannot be squeezed, what is known as hard core repulsion. Thus, in practice, there is a small change in temperature in the expansion. The *Joule coefficient* is defined to be $(\partial T/\partial V)_U$, that is, the change in T as volume increases at constant internal energy U, and we can relate this to other properties of the gas.

(iii) *The enthalpy and the Joule–Kelvin expansion.* The heat capacity at constant volume C_V involves the differentials of the functions of state and so we may ask if there is a differential of another function of state which corresponds to C_p. Let us write $U = U(p, T)$ instead of $U(V, T)$, remembering that we need only two coordinates to specify the state of the gas:

$$\mathrm{d}U = \left(\frac{\partial U}{\partial p}\right)_T \mathrm{d}p + \left(\frac{\partial U}{\partial T}\right)_p \mathrm{d}T.$$

Now proceed as before:

$$\begin{aligned} đQ &= \mathrm{d}U + p\,\mathrm{d}V \\ &= \left(\frac{\partial U}{\partial p}\right)_T \mathrm{d}p + p\,\mathrm{d}V + \left(\frac{\partial U}{\partial T}\right)_p \mathrm{d}T, \\ \left(\frac{đQ}{\mathrm{d}T}\right)_p &= p\left(\frac{\partial V}{\partial T}\right)_p + \left(\frac{\partial U}{\partial T}\right)_p \\ &= \left[\frac{\partial}{\partial T}(pV + U)\right]_p . \end{aligned} \tag{11.19}$$

The quantity $pV + U$ is entirely composed of functions of state and hence this new function must also be a function of state – it is known as the *enthalpy H*. Thus,

$$\begin{aligned} H &= U + pV, \\ C_p &= \left(\frac{đQ}{\mathrm{d}T}\right)_p = \left(\frac{\partial H}{\partial T}\right)_p = \left[\frac{\partial}{\partial T}(U + pV)\right]_p . \end{aligned} \tag{11.20}$$

The enthalpy often appears in flow processes and, in particular, in a *Joule–Kelvin expansion*. In this case, gas is transferred from one cylinder to another, the pressures in both cylinders being maintained at constant values p_1 and p_2 (Fig. 11.4). Suppose a certain mass of gas is pushed through one or more small holes or narrow tubes, what is often referred to as a 'porous plug', from A to B. The gas in A has initially internal energy U_1, volume V_1 and pressure p_1. After the piston in A has pushed the gas at constant pressure p_1 through the plug into B, the gas in B ends up with pressure p_2, volume V_2 and temperature T_2. The system is thermally isolated and hence we can apply the law of conservation of energy to

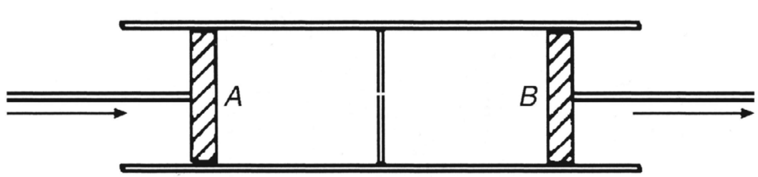


Fig. 11.4 Illustrating the Joule–Kelvin experiment on the expansion of gases. In Joule's experiments, the fluid was forced through one or more small holes or narrow tubes, often referred to as a porous plug, in such a way that the velocity of the fluid was very small on passing through the plug. Engineers refer to such a flow as a *throttled expansion*. (From H.S. Allen and R.S. Maxwell, 1939. *A Text-book of Heat*, Part II, p. 687. London: MacMillan & Co.)

the gas. The energy which is involved in pushing the gas through the plug consists of the internal energy U_1 plus the work done on the gas in A, $\int p\,\mathrm{d}V = p_1 V_1$, and this must equal the internal energy U_2 plus the work done on the gas in B, $\int p\,\mathrm{d}V = p_2 V_2$.

$$p_1 V_1 + U_1 = p_2 V_2 + U_2 \quad \text{or} \quad H_1 = H_2. \tag{11.21}$$

For a perfect gas, $H = pV + U = RT + U(T)$. But $U(T) + RT$ is a unique function of temperature and hence T must be the same in both volumes. Thus, for a perfect gas, there is no change in temperature in a Joule–Kelvin expansion. In real gases, however, there is a temperature change because of intermolecular forces. The temperature change may be either positive or negative depending upon the pressure and temperature. The Joule–Kelvin coefficient is defined to be $(\partial T/\partial p)_H$. The Joule–Kelvin experiment is a more sensitive method for determining deviations from perfect gas law behaviour than the Joule expansion.

We can now derive a more general conservation equation, in which we take into account other contributions to the total energy, for example, the kinetic energy and the potential energy of the gas if it is in a gravitational field. Let us consider the flow through a black box and include these energies as well (Fig. 11.5).

We consider the steady flow of a given mass m of gas or fluid as it enters and leaves the black box. The law of conservation of energy becomes

$$H_1 + \tfrac{1}{2}mv_1^2 + m\phi_1 = H_2 + \tfrac{1}{2}mv_2^2 + m\phi_2,$$
$$p_1 V_1 + U_1 + \tfrac{1}{2}mv_1^2 + m\phi_1 = p_2 V_2 + U_2 + \tfrac{1}{2}mv_2^2 + m\phi_2, \tag{11.22}$$

that is,

$$\frac{p}{m/V} + \frac{U}{m} + \tfrac{1}{2}v^2 + \phi = \text{constant},$$
$$\frac{p}{\rho} + u + \tfrac{1}{2}v^2 + \phi = \text{constant}, \tag{11.23}$$

where u is the *specific energy density*. This is one of the equations of fluid flow. In particular, for an incompressible fluid $u_1 = u_2$ and so we obtain *Bernoulli's equation*,

$$\frac{p}{\rho} + \tfrac{1}{2}v^2 + \phi = \text{constant}, \tag{11.24}$$

which we derived fluid mechanically in Section 9.2.2.

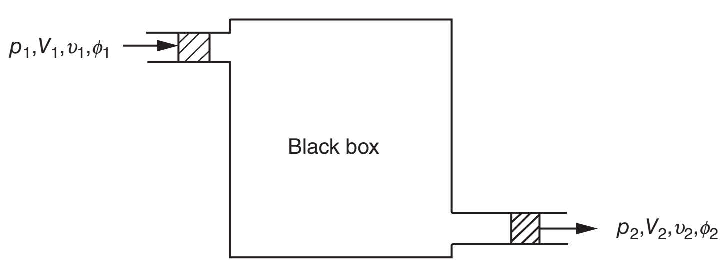


Fig. 11.5 Illustrating the conservation of energy in fluid flow in the presence of a gravitational field.

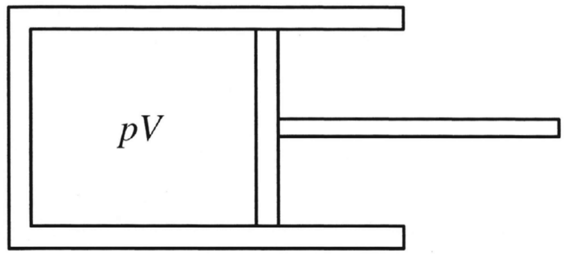


Fig. 11.6 Illustrating an adiabatic expansion within a perfectly insulated cylinder and piston.

Notice that we assumed that the additional terms present in Bernoulli's equation were absent in the Joule–Kelvin expansion. It is assumed that the Joule–Kelvin expansion takes place very slowly so that the kinetic energy terms can be neglected and the two volumes are at the same gravitational potential.

(iv) *Adiabatic expansion.* In an adiabatic expansion, the volume of the gas changes without any thermal contact between the system and its surroundings. The classical method of illustration is the expansion or compression of the gas within a perfectly insulated cylinder and piston (Fig. 11.6). An important point about expansions of this type is that they take place very slowly so that the system passes through an infinite number of equilibrium states between the initial and final thermodynamic coordinates. We will return to this key concept in our discussion of reversible processes in Section 11.5.1. We can therefore write

$$đQ = \mathrm{d}U + p\,\mathrm{d}V = 0. \tag{11.25}$$

If $C_V = (\partial U/\partial T)_V$ is taken to refer to 1 mole of gas, $\mathrm{d}U = nC_V\,\mathrm{d}T$ for n moles. During the expansion, the gas passes through an infinite series of equilibrium states for which the perfect gas law applies, $pV = nRT$, and hence, from (11.25),

$$\begin{aligned} nC_V\,\mathrm{d}T + \frac{nRT}{V}\,\mathrm{d}V &= 0, \\ \frac{C_V}{R}\frac{\mathrm{d}T}{T} &= -\frac{\mathrm{d}V}{V}. \end{aligned} \tag{11.26}$$

Integrating,

$$\frac{V_2}{V_1} = \left(\frac{T_2}{T_1}\right)^{-C_V/R} \quad \text{or} \quad VT^{C_V/R} = \text{constant}. \tag{11.27}$$

Since $pV = nRT$ at all stages in the expansion, this result can also be written

$$pV^{\gamma} = \text{constant},$$

where

$$\gamma = 1 + \frac{R}{C_V}.$$

We have already shown that, for one mole of gas, $C_V + R = C_p$ and consequently

$$1 + \frac{R}{C_V} = \frac{C_p}{C_V} = \gamma. \tag{11.28}$$

γ is the *ratio of specific heat capacities*, or the *adiabatic index*. For a monatomic gas, $C_V = \frac{3}{2}R$ and consequently $\gamma = \frac{5}{3}$.

(v) *Isothermal expansion.* In this case, there must be heat exchange with the surroundings so that the gas within the piston remains at the same temperature, $T =$ constant. In the expansion, work is done in pushing back the piston and this must be made up by a corresponding inflow of heat, that is, the work done is

$$\int_{V_1}^{V_2} p\,\mathrm{d}V = \int_{V_1}^{V_2} \frac{RT}{V}\,\mathrm{d}V = RT\,\ln\left(\frac{V_2}{V_1}\right). \tag{11.29}$$

This is the amount of heat which must be supplied from the surroundings to maintain an isothermal expansion. This result will find important applications in the analysis of heat engines.

(vi) *Expansions of different types.* Let us summarise the four different types of expansion described in this section.

- *Isothermal expansion*, $\Delta T = 0$. Heat must be supplied or removed from the system to maintain $\Delta T = 0$ (see item (v)).
- *Adiabatic expansion*, $\Delta Q = 0$. No heat exchange takes place with the surroundings (see item (iv)).
- *Joule expansion*, $\Delta U = 0$. *For a perfect gas*, the free expansion to a larger volume with fixed walls involves no change in internal energy (item (ii)).
- *Joule–Kelvin expansion*, $\Delta H = 0$. When a perfect gas passes from one volume to another and the pressures in the two vessels are maintained at p_1 and p_2 during the transfer, enthalpy is conserved (item (iii)).

The basic principle behind these apparently different phenomena is simply the conservation of energy – they are no more than elaborations of the first law of thermodynamics.

11.4 The Origin of the Second Law of Thermodynamics

It always comes as a surprise that the first law of thermodynamics proved to be so much more troublesome historically than the second law. After 1850, the first law enabled a logically self-consistent definition of heat to be formulated and the law of conservation of energy was placed on a firm conceptual basis. There must, however, be other restrictions upon thermodynamic processes. In particular, we do not yet have any rules about the direction in which heat flows, or about the direction in which thermodynamic systems evolve. These are derived from the *second law of thermodynamics*, which can be expressed in the form presented by Rudolf Clausius:

> No process is possible whose sole result is the transfer of heat from a colder to a hotter body.

It is remarkable that such powerful results can be derived from this simple statement. Equally remarkable is the way in which the fundamentals of thermodynamics were derived from the operation of steam engines. As L.J. Henderson has remarked,

> Until 1850, the steam engine did more for science than science did for the steam engine.[15]

Let us review the history behind this perceptive statement.

11.4.1 James Watt and the Steam Engine

The invention of the steam engine was perhaps the most important development which laid the foundations for the industrial revolution which swept through developed countries from about 1750 to 1850.[16] Before that time, the principal sources of power were waterwheels, which had to be located close to rivers and streams, and wind power, which was unreliable. Horses were used to provide power, but they were expensive. The importance of the steam engine was that power could be generated at the location where it was needed.

The first commercially viable and successful steam engine was built by Thomas Newcomen in 1712 in order to pump water out of coal mines and to raise water to power waterwheels. His steam engine (Fig. 11.7) operated at atmospheric pressure. The principle of operation was that the cylinder, fitted with a piston, was filled with steam, which was then condensed by injecting a jet of cold water. The resulting vacuum in the cylinder caused the piston to descend, resulting in a power stroke, which was communicated to the pump by

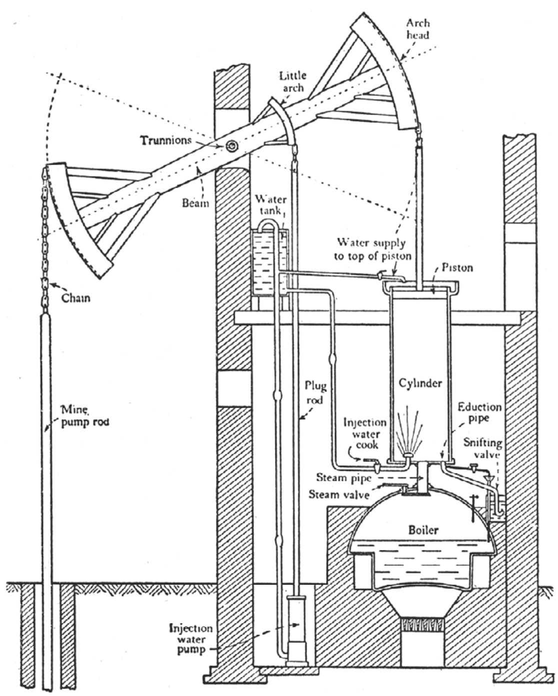


Fig. 11.7 Newcomen's atmospheric steam engine of 1712. (From H.W. Dickenson, 1958. In *A History of Technology*, Vol. IV, eds. C. Singer, E.J. Holmyard, A.R. Hall and T.I. Williams, p. 176. Oxford: Clarendon Press.)

Table 11.1 The efficiencies of steam engines[17]

Date	*Builder*	*Duty*	*Percentage thermal efficiency*
1718	Newcomen	4.3	0.5
1767	Smeaton	7.4	0.8
1774	Smeaton	12.5	1.4
1775	Watt	24.0	2.7
1792	Watt	39.0	4.5
1816	Woolf compound engine	68.0	7.5
1828	Improved Cornish engine	104.0	12.0
1834	Improved Cornish engine	149.0	17.0
1878	Corliss compound engine	150.0	17.2
1906	Triple-expansion engine	203.0	23.0

the mechanical arrangement shown in Fig. 11.7. When the power stroke was finished, the steam valve was opened and the weight of the pump rod pulled the piston back to the top of the cylinder. Through the next half century, the efficiency of Newcomen's design improved steadily thanks to the efforts of John Smeaton who, with better design and engineering, more than doubled the efficiency of Newcomen's original steam engine (Table 11.1).

One of the great heroes of this story is James Watt, who was trained as a scientific instrument maker in Glasgow. While repairing a model Newcomen steam engine in 1764, he was impressed by the waste of steam and heat in its operation. In May 1765, he made the crucial invention, which was to lead to the underpinning of the whole of thermodynamics. He realised that the efficiency of the steam engine would be significantly increased if the steam condensed in a separate chamber from the cylinder. The invention of the *condenser* was perhaps his greatest innovation. He took out his famous patent for this invention in 1768.

Figure 11.8 shows one of the steam engines built by Watt in 1788. The engine operated at atmospheric pressure. The key innovation was the separation of the condenser F from the cylinder and piston E. The cylinder was insulated and kept at a constant high temperature T_1 while the condenser was kept cool at a lower temperature T_2. The condenser was evacuated by the air pump H. The steam in E condensed when the condenser valve was opened and the power stroke took place. Steam was then allowed to fill the upper part of the cylinder and, once the stroke was finished, steam was allowed into the lower section of the cylinder, so that the weight of the pump rod restored the piston to its original position.

Pressure to increase the efficiency of steam engines continued from the mine owners, who were interested in reducing the costs of pumping operations. In 1782, Watt patented a double-action steam engine in which steam power was used in both the power and return strokes of the piston, resulting in twice the amount of work done. This required a new design of the parallel motion bars, so that the piston could push as well as pull. Watt referred to this as 'one of the most ingenious, simple pieces of mechanism I have

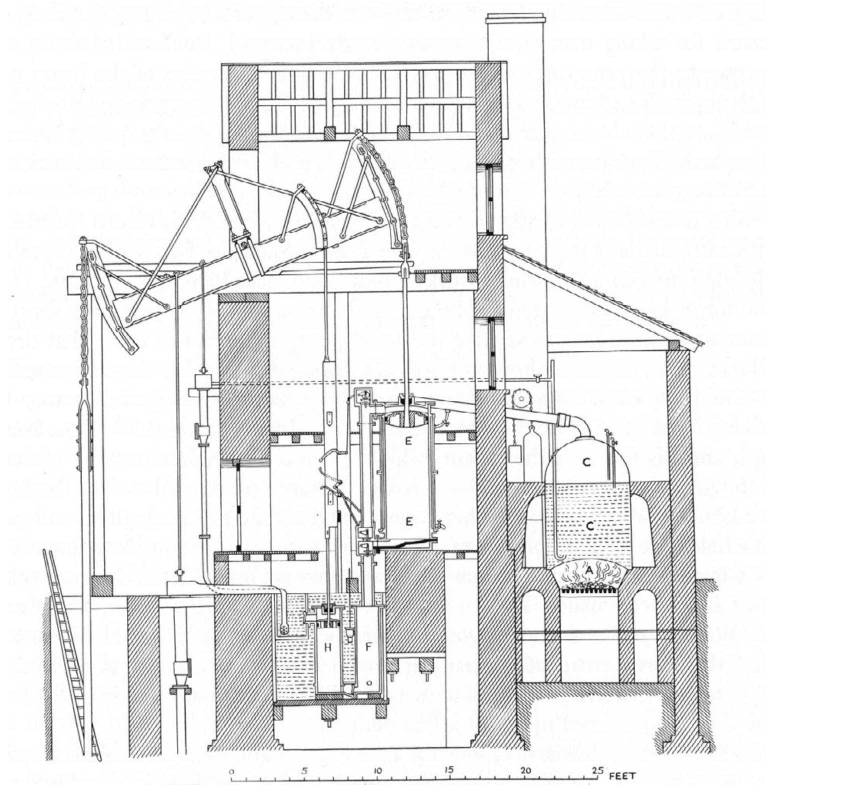


Fig. 11.8 James Watt's single-acting steam engine of 1788. The boiler C is placed in an outhouse. The cylinder E is kept at a high temperature all the time. F is the separate condenser and H an air pump. (From H.W. Dickenson, 1958. In *A History of Technology*, Vol. IV, eds. C. Singer, E.J. Holmyard, A.R. Hall and T.I. Williams, p. 184. Oxford: Clarendon Press.)

contrived.' There was also a governor to regulate the supply of steam. In the version of the steam engine shown in Fig. 11.9, the reciprocating motion of the piston was converted into rotary motion by the sun-and-planet gear.

The next significant advance occurred in the early 1800s with Richard Trevithick's steam engine which operated with high pressure steam, rather than at atmospheric pressure. As a consequence, more power could be generated from a smaller steam engine, and these were built into an early steam vehicle and an early version of the steam locomotive. Table 11.1 shows how the efficiencies of steam engines improved over the period from Newcomen to the beginning of the twentieth century. It is amusing that the 'duty' of the engines is defined in units of one million foot-poundals of energy per bushel of coal, where 1 bushel is 84 pounds, surely the ultimate non-SI unit. Table 11.1 was of the greatest importance to manufacturers and industrialists who were to drive through the industrial revolution. There was great interest in trying to understand what the maximum efficiency of a heat engine could be and these very practical considerations led to Sadi Carnot's interest in the efficiencies of heat engines.

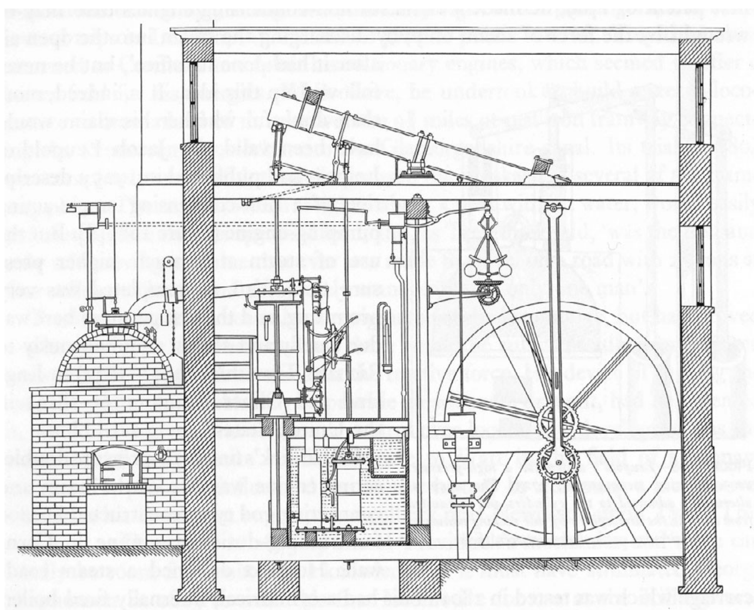


Fig. 11.9 James Watt's double-acting rotative steam engine of 1784. The parallel motion bars are seen on the raised end of the beam. The sun-and-planet gear converts the reciprocating action of the engine into rotary motion. (From H.W. Dickenson, 1958. In *A History of Technology*, Vol. IV, eds. C. Singer, E.J. Holmyard, A.R. Hall and T.I. Williams, p. 187. Oxford: Clarendon Press.)

11.4.2 Sadi Carnot and the *Réflexions*

Nicolas Léonard Sadi Carnot was the eldest son of Lazare Carnot, who was a member of the Directory after the French Revolution and later, during the Hundred Days in 1815, Napoleon's Minister of the Interior. After 1807, Lazare Carnot devoted much of his energies to the education of his sons. Sadi Carnot was educated at the elite École Polytechnique, where his teachers included Poisson, Gay-Lussac and Ampère. After a period as a military engineer, from 1819 onwards he was able to devote himself wholly to research. Carnot was concerned about the design of efficient steam engines, considering that France had fallen behind the United Kingdom in technology. He was also impressed by the fact that, although reliable estimates of the efficiencies of heat engines had been presented, those who had developed them were inventors and engineers, rather than theorists. His great work *Réflections sur la Puissance Motrice du Feu et sur les Machines Propres à Developper cette Puissance*[18] was published in 1824; it is normally translated as *Reflections on the Motive Power of Fire*. The treatise concerned a theoretical analysis of the maximum efficiency of heat engines. In his approach, he was strongly influenced by his father's work on steam engines. The treatise is, however, of much greater generality and is an intellectual achievement of the greatest originality.

Earlier work on the maximum efficiency of steam engines had involved empirical studies such as comparing the fuel input with the work output, or else theoretical studies on the basis of specific models for the behaviour of the gases in heat engines. Carnot's aims were unquestionably practical in intent, but his basic perceptions were entirely novel. In my view, the imaginative leap involved is one of genius.

In seeking to derive a completely general theory of heat engines, he was guided by his father's premise in his study of steam engines of the impossibility of perpetual motion. In the *Réflexions*, he adopted the caloric theory of heat and assumed that caloric was conserved in the cyclic operation of heat engines. He postulated that it is the transfer of caloric from a hotter to a colder body which is the origin of the work done by a heat engine. The flow of caloric was envisaged as being analogous to the flow of fluid which, as in a waterwheel, can produce work when it falls down a potential gradient.

Carnot's basic insights into the operation of heat engines were twofold. First, he recognised that a heat engine works most efficiently if the transfer of heat takes place as part of a *cyclic process*. The second was that the crucial factor in determining the amount of work which can be extracted from a heat engine is the temperature difference between the source of heat and the sink into which the caloric flows, the T_1 and T_2 prefigured in Watt's invention of the condenser. It turns out that these basic ideas are independent of the particular model of the heat flow process.

By another stroke of imaginative insight, he devised the cycle of operations which we now know as the *Carnot cycle* as an idealisation of the behaviour of any heat engine. A key feature of the Carnot cycle is that all the processes are carried out *reversibly* so that, by reversing the sequence of operations, work can be done on the system and caloric transferred from a colder to a hotter body. By joining together an arbitrary heat engine, and a reversed Carnot heat engine, he was able to demonstrate that no heat engine can ever produce more work than a Carnot heat engine. If it were otherwise, by joining the two engines together, we could either transfer heat from a colder to a hotter body without doing any work, or we could produce a net amount of work without any net heat transfer, both phenomena being in violation of common experience. The influence of Lazare Carnot's premise about the impossibility of perpetual motion is apparent. We will demonstrate these results formally in Section 11.5.2.

It is tragic that Sadi Carnot died of cholera at the age of 36 in August 1832, before the profound significance of his work was fully appreciated. In 1834, however, Émile Clapeyron reformulated Carnot's arguments analytically and related the ideal Carnot engine to the standard pressure–volume indicator diagram. There the matter rested until William Thomson studied certain aspects of Clapeyron's paper and went back to Carnot's original version in the *Réflexions*. The big problem for Thomson and others at that time was to reconcile Carnot's work, in which caloric was conserved, with Joule's experiments which demonstrated the interconvertibility of heat and work. The matter was resolved by Rudolf Clausius, who showed that Carnot's theorem concerning the maximum efficiency of heat engines was correct, but that the assumption of no heat loss was wrong. In fact, there is a conversion of heat into work in the Carnot cycle. This reformulation by Clausius constitutes the bare bones of the second law of thermodynamics.[19] As we will show, however, the law goes far beyond the efficiency of heat engines. It serves not only to define the thermodynamic temperature scale, but also to resolve the problem of how systems evolve thermodynamically. Let us develop these concepts more formally, bringing out the basic assumptions made in formulating the second law mathematically.

11.5 The Second Law of Thermodynamics

Let us first discuss the crucial distinction between *reversible* and *irreversible* processes.

11.5.1 Reversible and Irreversible Processes

A reversible process is one which is carried out infinitely slowly so that, in passing from the state A to state B, the system passes through an infinite number of equilibrium states. Since the process takes place infinitely slowly, there is no friction or turbulence and no sound waves are generated. At no stage are there unbalanced forces. At each stage, we make only an infinitesimal change to the system. The implication is that, by reversing the process precisely, we can get back to the point from which we started and nothing will have changed in either the system or its surroundings. Clearly, if there were frictional losses, we could not get back to where we started without extracting some energy from the surroundings.

Let us emphasise this point by considering in detail how we could carry out a reversible isothermal expansion. Suppose we have a large heat reservoir at temperature T and a cylinder with gas in thermal contact with it also at temperature T (Fig. 11.10). No heat flows if the two volumes are at the same temperature. But if we make an infinitesimally small movement of the piston outwards, the gas in the cylinder cools infinitesimally and so an infinitesimal amount of heat flows into the gas by virtue of the temperature difference. This small amount of energy brings the gas back to T. The system is reversible because, if we compress the gas at T slightly, it heats up and heat flows from the gas into the reservoir. Thus, provided we consider only infinitesimal changes, the heat flow process occurs reversibly.

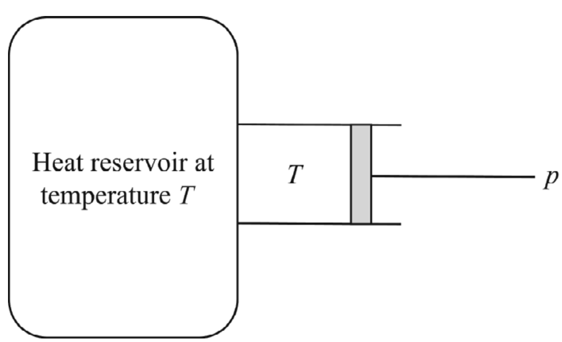


Fig. 11.10 Illustrating a reversible isothermal expansion.

Clearly, this is not possible if the reservoir and the piston are at different temperatures. In this case, we cannot reverse the direction of heat flow by making an infinitesimal change in the temperature of the cooler object. This makes the important point that, in reversible processes, the system must be able to evolve from one state to another by passing through an infinite set of equilibrium states which we join together by infinitesimal increments of work and energy flow.

To reinforce this point further, let us repeat the argument for an adiabatic expansion. The cylinder is completely thermally isolated from the rest of the Universe. Again, we perform each step infinitesimally slowly. There is no flow of heat in or out of the system and there is no friction. Therefore, since each infinitesimal step is reversible, we can perform the whole expansion by adding lots of them together.

Let us contrast this behaviour with the other two expansions we have described. In the Joule expansion, the gas expands into a large volume and this cannot take place without all sorts of non-equilibrium processes taking place. Unlike the adiabatic and isothermal expansions, there is no way in which we can design a series of equilibrium states through which the final state is reached. The Joule–Kelvin expansion is a case in which there is a discontinuity in the properties of the gas on passing into the second cylinder. We do not reach the final state by passing through a series of equilibrium states infinitely slowly.

This long introduction stresses the point that reversible processes are highly idealised, but they provide a benchmark against which all real processes can be compared. It is important to appreciate exactly what we mean by the term 'infinitely slowly' in the above arguments. In practice, what this means is that the time-scale over which the change in the system takes place must be very much longer than the time it takes thermal equilibrium to be re-established at each stage in the expansion by elementary physical processes within the system. If this is the case, the system can approach ideal behaviour, if enough care is taken.

11.5.2 The Carnot Cycle and the Definition of Thermodynamic Temperature

Clausius's statement of the second law is another 'fact of experience' which is asserted without proof:

> No process is possible whose sole result is the transfer of heat from a colder to a hotter body.

The implication is that heat cannot be transferred from a cold to a hot body without some interaction with the surroundings of the system. Notice that the law assumes that we can define what we mean by the terms 'hotter' and 'colder' – all we have available so far is an empirical temperature scale defined on the basis of the properties of perfect gases. We will show that this is identical to the thermodynamic temperature scale.

It is a matter of common experience that it is easy to convert work into heat, for example, by friction, but it is much more difficult to devise an efficient means of converting heat into work, as is illustrated by the efficiencies of steam engines listed in Table 11.1. Carnot's deep insights into the operation of an ideal heat engine are worth repeating.

(1) In any efficient heat engine, there is a working substance which is used *cyclically* in order to minimise the loss of heat.

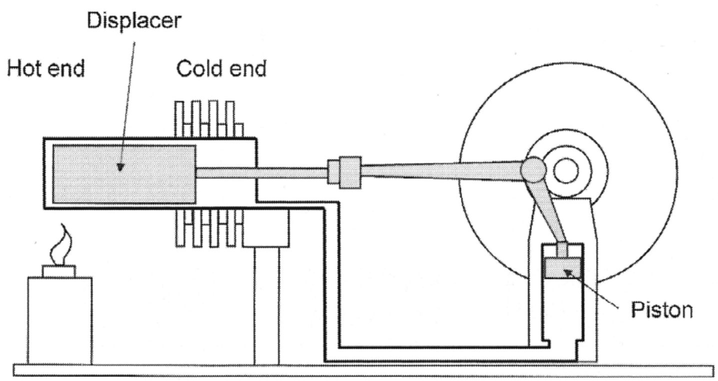


Fig. 11.11 Illustrating the construction of a simple Stirling engine. (After the 'Super Vee' Stirling engine, available from the All Hot Air Company.)

(2) In all real heat engines, heat enters the working substance at a high temperature T_1, the working substance does work and then returns heat to a heat sink at a lower temperature T_2.
(3) The best we can achieve is to construct a heat engine in which all processes take place *reversibly*, that is, all the changes are carried out infinitely slowly so that there are no dissipative losses, such as friction or turbulence.

These are idealisations of the operation of real heat engines. A simple illustration of how these principles work in practice is provided by the *Stirling heat engine*, which was invented by the Revd Dr Robert Stirling of Kilmarnock in Scotland in 1816 (Fig. 11.11). The working substance, air, fills the whole of the volume indicated by a heavy line. At the beginning of the cycle, the gas is heated to a high temperature by the heater at the 'hot end' and the piston at the right end of the diagram is pushed upwards resulting in the configuration shown in Fig. 11.11. The air then cools as a result of which its pressure drops below atmospheric pressure and this pressure difference pushes the piston back down the cylinder. Stirling's clever invention was to move the working substance from the heat source to a heat sink by means of a *displacer*, which is shown in Fig. 11.11. The hot gas comes into contact with the 'cool end' of the volume, which is maintained at room temperature by cooling fins or by the flow of water through cooling pipes. The hot gas gives up heat at the cold end so that the pressure drops and the return stroke of the piston takes place. As shown in the diagram, the piston is attached to the displacer, which now moves to the cold end of its cylinder, thus pushing the cooled gas to the hot end where it is reheated and the next power stroke can begin. By careful design, the efficiency of the Stirling engine can approach the theoretical maximum value discussed below.

This sequence of operations is almost identical to that studied by Carnot in his great paper of 1824. The working substance acquires heat at the hot end at temperature T_1, it does work on the piston and then gives up heat at the cold end at temperature T_2. The working substance is then reheated and the cycle begins again.

Carnot's famous cycle of operations of a perfect heat engine is illustrated in Fig. 11.12(a). I strongly recommend Feynman's careful description of the cycle in Chapter 44 of Volume 1 of his *Lectures on Physics*.[20] My exposition is modelled on his.

There are two very large heat reservoirs 1 and 2 which are maintained at temperatures T_1 and T_2. The working substance is a gas which is contained within a cylinder with a frictionless piston. We carry out the following reversible sequence of operations which simulate how real heat engines do work.

(1) The cylinder is placed in thermal contact with the reservoir at the higher temperature T_1 and a very slow reversible isothermal expansion is carried out. As described above, in a reversible isothermal expansion, heat flows from the reservoir into the gas in the cylinder. At the end of the expansion, a quantity of heat Q_1, has been absorbed by the gas and a certain amount of work has been done by the working substance on the surroundings. This change corresponds to moving from A to B on the p–V indicator diagram (Fig. 11.12(b)).
(2) The cylinder is now isolated from the heat reservoirs and an infinitely slow adiabatic expansion is performed so that the temperature of the working substance falls from T_1 to T_2, the temperature of the second reservoir. This process is reversible and again work is done on the surroundings.
(3) The cylinder is now placed in thermal contact with the second reservoir at temperature T_2 and the gas in the cylinder is compressed, again infinitely slowly and reversibly at temperature T_2. In this process, the temperature of the gas is continuously increased infinitesimally above T_2 and so heat flows into the second reservoir. In this process, heat Q_2 is returned to the second reservoir and work is done on the working substance. The isothermal compression continues until the isotherm at T_2 intersects the adiabatic curve which will take the working substance back to its initial state.
(4) The cylinder is again isolated from the heat reservoirs and an infinitely slow adiabatic compression is carried out, bringing the gas back to its original state. Again work is done on the working substance by the surroundings as its temperature increases under adiabatic conditions.

This is the sequence of operations of an ideal heat engine. The net effect of this cycle is that the working substance gained heat energy Q_1 from reservoir 1 at T_1, and returned Q_2 to reservoir 2 at T_2. In the process, the heat engine has done work. The amount of work done is just $W = \oint p\,\mathrm{d}V$ taken round the cycle. From the diagram it is clear that the cyclic integral is equal to the area of the closed curve described by the cycle of operations in the p–V indicator diagram (Fig. 11.13). We also know from the first law of thermodynamics that this work must be equal to the net heat supplied to the working substance, $Q_1 - Q_2$, that is,

$$W = Q_1 - Q_2. \qquad (11.30)$$

Note incidentally that the adiabatic curves must be steeper than the isotherms because for the former $pV^\gamma = \text{constant}$ and $\gamma > 1$ for all gases. We can draw this machine schematically as shown in Fig. 11.14.

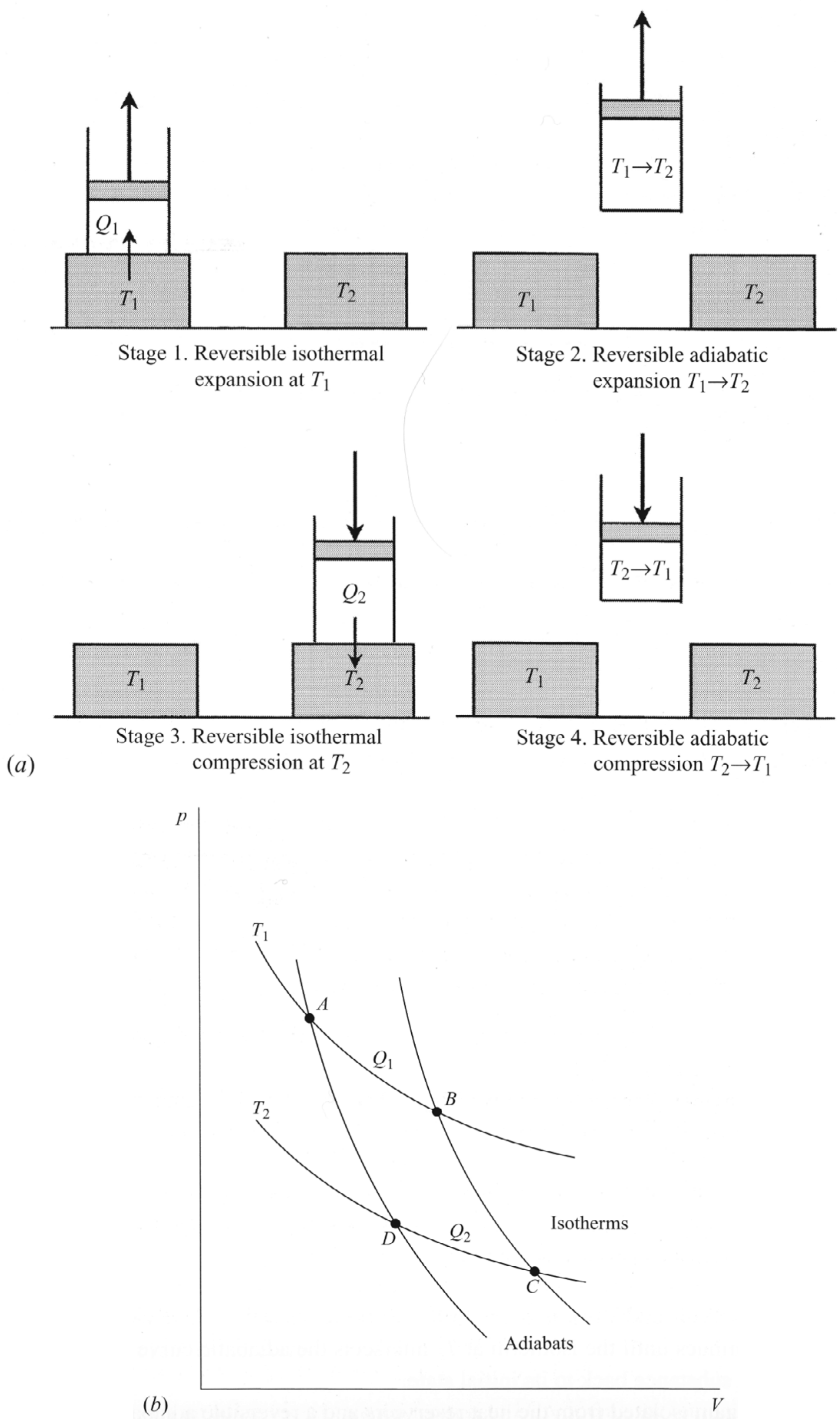


Fig. 11.12 (*a*) Illustrating schematically the four stages in the Carnot cycle of an ideal heat engine. (*b*) The four stages of operation of an ideal Carnot engine on a p–V, or indicator, diagram.

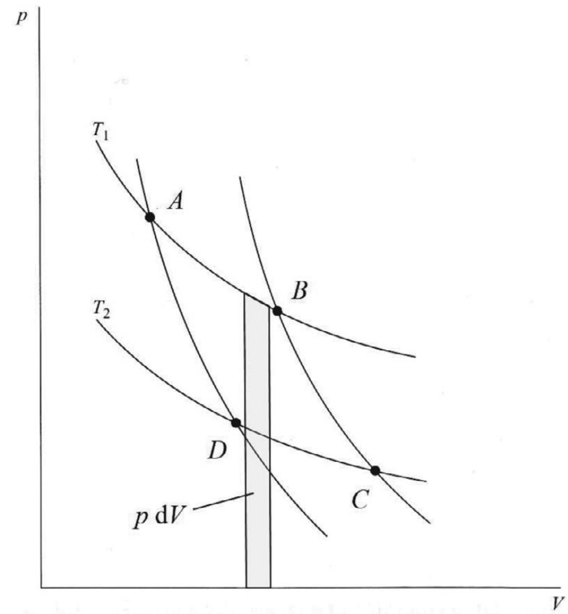


Fig. 11.13 Illustrating that the area within the closed cycle in the indicator diagram represents the total work done, namely $\oint p\,\mathrm{d}V =$ area enclosed by *ABCD*.

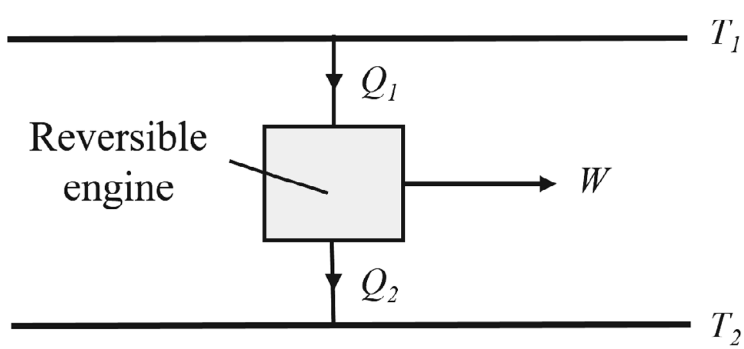


Fig. 11.14 A representation of a reversible Carnot heat engine.

The beautiful aspect of this argument is that all stages of the Carnot cycle are reversible and so we can run the whole sequence of actions backwards (Fig. 11.15). The *reversed Carnot cycle* is therefore:

(1) Adiabatic expansion from A to D, reducing the temperature of the working substance from T_1, to T_2.
(2) Isothermal expansion at T_2 in which heat Q_2 is withdrawn from the reservoir.
(3) Adiabatic compression from C to B so that the working substance returns to temperature T_1.

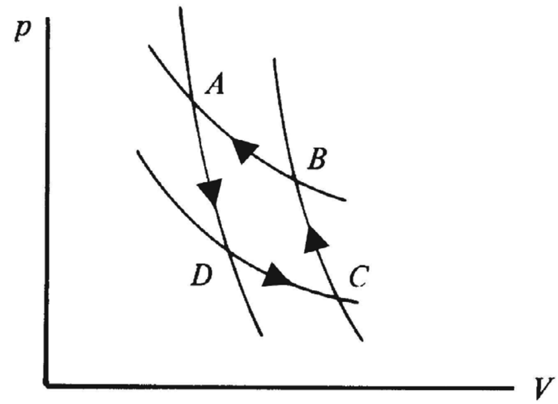


Fig. 11.15 A reversed Carnot cycle representing a refrigerator or a heat pump.

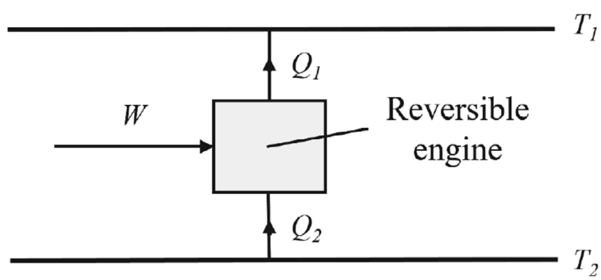


Fig. 11.16 A representation of a reversed Carnot cycle acting as a refrigerator or heat pump.

(4) Isothermal compression at T_1 so that the working substance gives up heat Q_1 to the reservoir at T_1.

The reversed Carnot cycle describes the operation of an ideal refrigerator or heat pump in which heat is extracted from the reservoir at the lower temperature and passed to the reservoir at the higher temperature. This refrigerator or heat pump is shown schematically in Fig. 11.16. Notice that, in the reversed cycle, work has been done in order to extract the heat from T_2 and deliver it to T_1.

We can now define the efficiencies of heat engines, refrigerators and heat pumps. For the standard heat engine running forwards, the efficiency

$$\eta = \frac{\text{work done in a cycle}}{\text{heat input}} = \frac{W}{Q_1} = \frac{Q_1 - Q_2}{Q_1}.$$

For a refrigerator,

$$\eta = \frac{\text{heat extracted from reservoir 2}}{\text{work done}} = \frac{Q_2}{W} = \frac{Q_2}{Q_1 - Q_2}.$$

For a heat pump, the cycle is run backwards supplying the heat Q_1 at T_1 by doing work W,

$$\eta = \frac{\text{heat supplied to reservoir 1}}{\text{work done}} = \frac{Q_1}{W} = \frac{Q_1}{Q_1 - Q_2}.$$

We can now do three things – prove Carnot's theorem, show the equivalence of the Clausius and Kelvin statements of the second law of thermodynamics and derive the concept of thermodynamic temperature.

11.5.3 Carnot's Theorem

Carnot's theorem states

> Of all heat engines working between two given temperatures, none can be more efficient than a reversible heat engine.

Suppose the opposite were true, namely, that we could construct an irreversible heat engine which has efficiency greater than that of a reversible engine working between the same temperatures. Then we could use the work produced by the irreversible engine to drive the reversible engine backwards (Fig. 11.17). Let us now regard the system consisting of the two combined engines as a single engine. All the work produced by the irreversible engine is used to drive the reversible engine backwards. Therefore, since it has been stated that $\eta_{\rm irr} > \eta_{\rm rev}$, we conclude that $W/Q_1' > W/Q_1$, and consequently $Q_1 - Q_1' > 0$. Thus, overall, the only net effect of this combined engine is to produce a transfer of energy from the lower to the higher temperature without doing any work, which is forbidden by the second law. Therefore, no irreversible engine can have efficiency greater than that of a reversible engine operating between the same two temperatures. This is the theorem which Carnot published in his *Réflexions* of 1824, 26 years before the formal statement of the laws of thermodynamics by Clausius in 1850.

11.5.4 The Equivalence of the Clausius and Kelvin Statements of the Second Law

Kelvin's statement of the second law is

> No process is possible whose sole result is the complete conversion of heat into work.

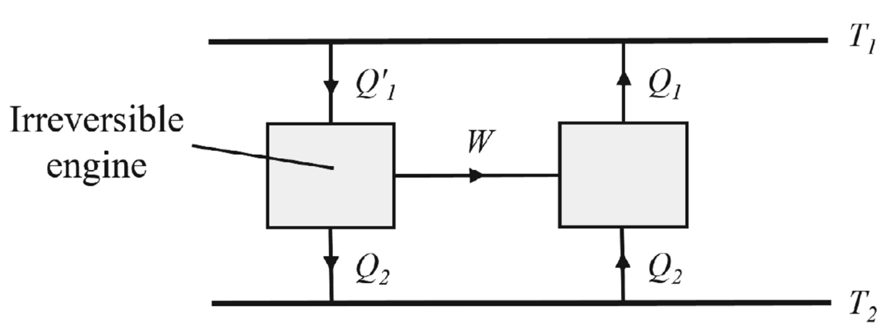


Fig. 11.17 Illustrating the proof of Carnot's theorem.

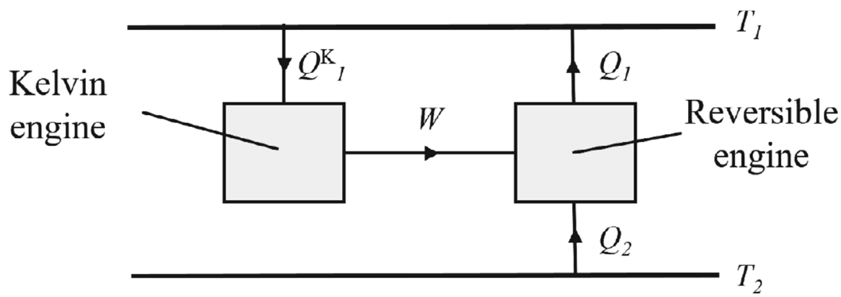


Fig. 11.18 Illustrating the equivalence of Kelvin's and Clausius's statements of the second law using a hypothetical Kelvin engine which converts all the heat supplied at T_1, into work.

Suppose it were possible for an engine K to convert the heat Q_1 completely into work W. Again, we could use this work to drive a reversible Carnot engine backwards as a heat pump (Fig. 11.18). Then, regarded as a single system, no net work would be done, but the total amount of heat delivered to the reservoir at T_1 is

$$Q_1 - Q_1^{\mathrm{K}} = Q_1 - W = Q_1 - (Q_1 - Q_2) = Q_2,$$

where Q_1^{K} means the heat supplied to the hypothetical Kelvin engine which is completely converted into work W. There is, therefore, a net transfer of heat Q_2 to the reservoir T_1 without any other change in the Universe and this is forbidden by the Clausius statement. Therefore, the Clausius and Kelvin statements are equivalent.

11.5.5 Thermodynamic Temperature

At last we can define *thermodynamic temperature*, the clue coming from Carnot's theorem. A reversible heat engine working between two temperatures has the maximum efficiency η and this is a unique function of the two temperatures which so far we have called T_1 and T_2 without defining what they are. Expressing this result in terms of the heat input and output,

$$\eta = \frac{Q_1 - Q_2}{Q_1}, \qquad \frac{Q_1}{Q_2} = \frac{1}{1-\eta}. \tag{11.31}$$

Let us now be somewhat more precise in the logic of the argument. Let the ratio Q_1/Q_2 be some function $f(\theta_1, \theta_2)$ of the empirical temperatures θ_1 and θ_2. Now join together two heat engines in series as shown in Fig. 11.19. The engines are connected so that the heat delivered at temperature θ_2 is used to supply the heat Q_2 to a second heat engine, which then delivers the work W_2 and transfers the heat Q_3 to the reservoir at θ_3. Thus,

$$\frac{Q_1}{Q_2} = f(\theta_1, \theta_2) \quad \text{and} \quad \frac{Q_2}{Q_3} = f(\theta_2, \theta_3). \tag{11.32}$$

On the other hand, we can consider the combined system as a single engine operating between θ_1 and θ_3, in which case,

$$\frac{Q_1}{Q_3} = f(\theta_1, \theta_3). \tag{11.33}$$

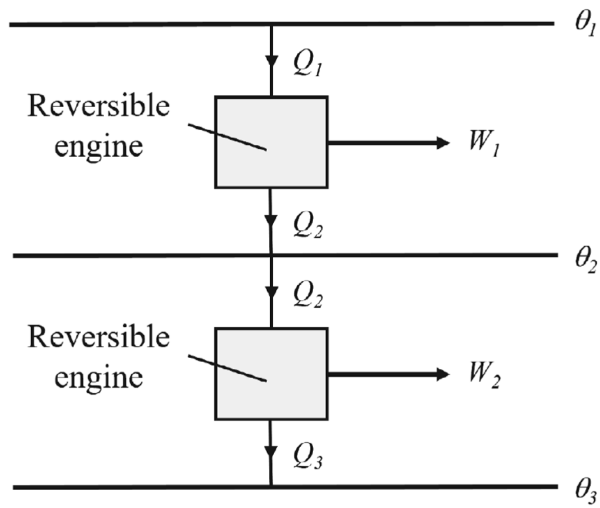


Fig. 11.19 Illustrating the origin of the definition of thermodynamic temperature.

Thus, since

$$\frac{Q_1}{Q_3} = \frac{Q_1}{Q_2} \times \frac{Q_2}{Q_3},$$

$$f(\theta_1, \theta_3) = f(\theta_1, \theta_2) \times f(\theta_2, \theta_3). \tag{11.34}$$

Consequently, the function f must have the form

$$f(\theta_1, \theta_3) = \frac{g(\theta_1)}{g(\theta_3)} = \frac{g(\theta_1)}{g(\theta_2)} \times \frac{g(\theta_2)}{g(\theta_3)}. \tag{11.35}$$

We then adopt a definition of *thermodynamic temperature* T consistent with this requirement, that is

$$\frac{g(\theta_1)}{g(\theta_3)} = \frac{T_1}{T_3}, \quad \text{that is} \quad \frac{Q_1}{Q_3} = \frac{T_1}{T_3}. \tag{11.36}$$

Now we can show that this definition is identical to that corresponding to the perfect gas temperature scale. We write the perfect gas law temperature scale as $T^{\rm P}$ and hence $pV = RT^{\rm P}$. For the Carnot cycle with a perfect gas, from (11.29),

$$\begin{cases} Q_1 = \int_A^B p\,dV = RT_1^{\rm P} \ln\left(\frac{V_B}{V_A}\right), \\ Q_2 = \int_D^C p\,dV = RT_2^{\rm P} \ln\left(\frac{V_C}{V_D}\right). \end{cases} \tag{11.37}$$

Along the adiabatic legs of the cycle,

$$pV^\gamma = \text{constant} \quad \text{and} \quad T^{\rm P} V^{\gamma-1} = \text{constant}. \tag{11.38}$$

Therefore,

$$\begin{cases} (V_B/V_A)^{\gamma-1} = (T_2^{\text{P}}/T_1^{\text{P}}), \\ (V_D/V_C)^{\gamma-1} = (T_1^{\text{P}}/T_2^{\text{P}}). \end{cases} \tag{11.39}$$

Multiplying the two equations in (11.39) together, we find that

$$\left(\frac{V_B V_D}{V_A V_C}\right) = 1, \tag{11.40}$$

that is,

$$\frac{V_B}{V_A} = \frac{V_C}{V_D}, \tag{11.41}$$

and, consequently, from the relations (11.37) and (11.36) we see that

$$\frac{Q_1}{Q_2} = \frac{T_1^{\text{P}}}{T_2^{\text{P}}} = \frac{T_1}{T_2}. \tag{11.42}$$

At last, we have derived a rigorous thermodynamic definition of temperature based upon the operation of perfect heat engines, the line of thought initiated by Carnot.

Now we can rewrite the maximum efficiencies of heat engines, refrigerators and heat pumps in terms of the thermodynamic temperatures between which they work.

$$\begin{cases} \text{Heat engine}: & \eta = \dfrac{Q_1 - Q_2}{Q_1} = \dfrac{T_1 - T_2}{T_1}, \\ \text{Refrigerator}: & \eta = \dfrac{Q_2}{Q_1 - Q_2} = \dfrac{T_2}{T_1 - T_2}, \\ \text{Heat pump}: & \eta = \dfrac{Q_1}{Q_1 - Q_2} = \dfrac{T_1}{T_1 - T_2}. \end{cases} \tag{11.43}$$

11.6 Entropy

You may have noticed something rather remarkable about the relation we derived in coming to our definition of thermodynamic temperature. Taking the heats Q_1 and Q_2 to be positive quantities, (11.42) shows that

$$\frac{Q_1}{T_1} - \frac{Q_2}{T_2} = 0. \tag{11.44}$$

This can be rewritten in the following form:

$$\int_A^C \frac{đQ}{T} - \int_C^A \frac{đQ}{T} = 0, \tag{11.45}$$

since heat is only suppled or removed along the isothermal sections of the Carnot cycle shown in Fig. 11.12. Since these sections of the cycle are reversible, this means that if we pass from A to C down either leg of the cycle,

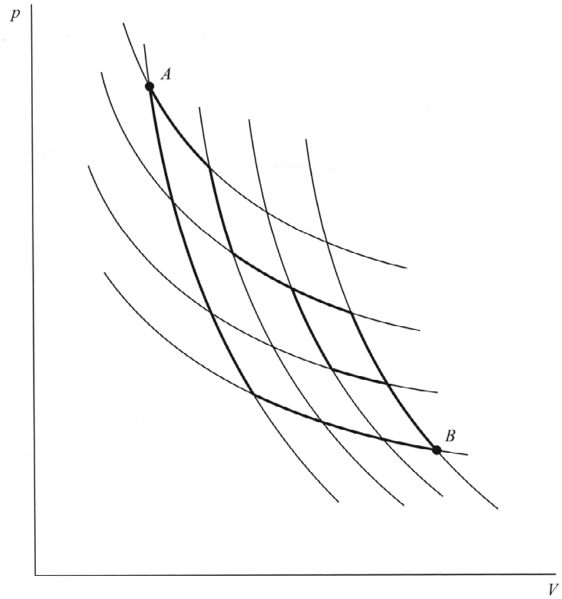


Fig. 11.20 Illustrating different paths between *A* and *B* in an indicator diagram.

$$\underbrace{\int_A^C \frac{đQ}{T}}_{\text{via } B} = \underbrace{\int_A^C \frac{đQ}{T}}_{\text{via } D}. \tag{11.46}$$

This strongly suggests that we have discovered another function of state – it does not matter how we get from A to C, as can be seen from the following argument.

For any two-coordinate system, we can always move between any two points in the indicator diagram by a large number of mini-Carnot cycles, all of them reversible as illustrated in Fig. 11.20. Whichever path we take, we will always find the same result. Mathematically, for any two points A and B,

$$\sum_A^B \frac{đQ}{T} = \text{constant}. \tag{11.47}$$

Writing this in integral form,

$$\int_A^B \frac{đQ}{T} = \text{constant} = S_B - S_A. \tag{11.48}$$

This new function of state S is called the *entropy* of the system. Notice that it is defined for reversible processes connecting the states A and B. By T, we mean the temperature at which the heat is supplied to the system, which in this case is the same as the temperature of the system itself, because the processes of heat exchange are performed reversibly. An amusing footnote to the controversy about caloric and the nature of heat is that, in a sense, *entropy* was Carnot's *caloric fluid* since entropy, rather than heat, is conserved in the Carnot cycle.

Now for any real engine, the efficiency must be less than that of an ideal Carnot engine working between the same temperatures and therefore

$$\eta_{\rm irr} \le \eta_{\rm rev},$$
$$\frac{Q_1 - Q_2}{Q_1} \le \frac{Q_1({\rm rev}) - Q_2({\rm rev})}{Q_1({\rm rev})}.$$

Therefore,

$$\frac{Q_2}{Q_1} \ge \frac{Q_2({\rm rev})}{Q_1({\rm rev})} = \frac{T_2}{T_1},$$
$$\frac{Q_2}{T_2} \ge \frac{Q_1}{T_1}.$$

Thus, round the cycle

$$\frac{Q_1}{T_1} - \frac{Q_2}{T_2} \le 0.$$

This suggests that, in general, when we add together a large number of mini-Carnot cycles,

$$\oint \frac{đQ}{T} \le 0, \qquad (11.49)$$

where the $đQ$s are the heat inputs and extractions in the real cycle. Notice that the equals sign only applies in the case of a reversible heat engine. Notice also the sign convention for heat. When heat enters the system, the sign is positive; when it is removed, it has a negative sign.

The relation $\oint đQ/T \le 0$ is fundamental to the whole of thermodynamics and is known as *Clausius's theorem*. This proof of the theorem is adequate for our present purposes. A limitation is that it only applies to two-coordinate systems and, in general, we need to be able to deal with multi-coordinate systems. This question is treated nicely in Pippard's book, where it is shown that the same result is true for multi-coordinate systems.

11.7 The Law of Increase of Entropy

Notice that the change in entropy is defined for a reversible sequence of changes from the state A to the state B. When we take the system from state A to state B in the real imperfect world, we can never effect precisely reversible changes and so the heat involved is not directly related to the entropy difference between the initial and final states. Let us compare what happens in a reversible and irreversible change between two states. Suppose an irreversible change takes place from A to B (Fig. 11.21). Then we can complete the cycle by taking any *reversible* path back from B to A. Then, according to Clausius's theorem,

$$\oint \frac{đQ}{T} \le 0 \quad \text{or} \quad \underbrace{\int_A^B \frac{đQ}{T}}_{\rm irrev} + \underbrace{\int_B^A \frac{đQ}{T}}_{\rm rev} \le 0,$$

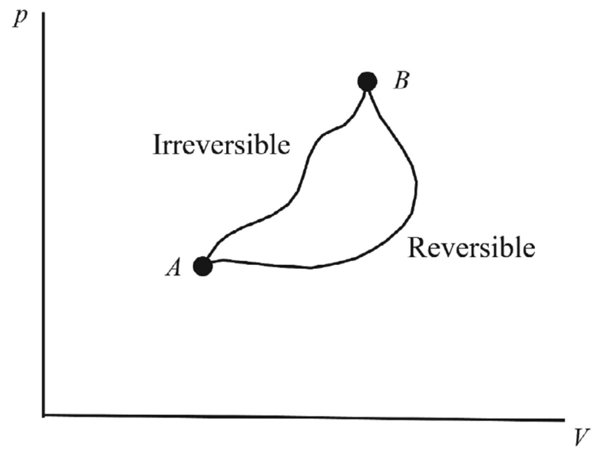


Fig. 11.21 Reversible and irreversible paths between the points *A* and *B* in an indicator diagram.

that is,

$$\underbrace{\int_A^B \frac{đQ}{T}}_{\text{irrev}} \le \underbrace{\int_A^B \frac{đQ}{T}}_{\text{rev}} .$$

But, because the second change is reversible,

$$\underbrace{\int_A^B \frac{đQ}{T}}_{\text{rev}} = S_B - S_A,$$

according to the definition of entropy. Consequently,

$$\underbrace{\int_A^B \frac{đQ}{T}}_{\text{irrev}} \le S_B - S_A,$$

or, writing this for a differential irreversible change,

$$\frac{đQ}{T} \le \Delta S. \tag{11.50}$$

Thus, we obtain the general result that for any differential change $đQ/T \le \Delta S$, where the equality holds only in the case of a reversible change.

This is a key result because at last we can understand how we can describe quantitatively what determines the direction in which physical processes evolve. Notice that, if we are dealing with reversible processes, the temperature at which heat is supplied is that of the system itself. This is not necessarily the case in irreversible processes.

For an *isolated system*, there is no thermal contact with the outside universe and hence in the above inequality, $đQ = 0$. Therefore

$$\Delta S \ge 0. \tag{11.51}$$

Thus, the entropy of an isolated system cannot decrease. A corollary of this is that, in approaching equilibrium, the entropy of the isolated system must tend towards a maximum and the final equilibrium configuration will be that for which the entropy is the greatest.

For example, consider two bodies at temperatures T_1 and T_2 and suppose that a small amount of heat ΔQ is exchanged between them. Then heat flows from the hotter to the colder body and we find an entropy decrease of the hotter body $\Delta Q/T_1$ and an entropy increase of the colder body $\Delta Q/T_2$ whilst the total entropy change is $(\Delta Q/T_2)-(\Delta Q/T_1) > 0$, which is positive, meaning that we were correct in our inference about the direction of heat flow.

So far we have only dealt with isolated systems. Now let us consider the case in which there is thermal contact between the system and its surroundings. As a first example, take a system round a complete cycle of operations so that it ends up in exactly the same state as at the beginning of the cycle, which may well involve irreversible processes as occurs in real systems. Since the system ends up in exactly the same state in which it started, all the functions of state are exactly the same and hence, for the system, $\Delta S = 0$. According to Clausius' theorem, this must be greater than or equal to $\oint đQ_{\text{sys}}/T$, that is,

$$0 = \Delta S \geq \oint \frac{đQ_{\text{sys}}}{T}.$$

In order to return to the initial state, heat must have been communicated to the system from its surroundings. The very best we can do is to transfer the heat reversibly to and from the surroundings. Then, in each stage of the cycle, $đQ_{\text{sys}} = -đQ_{\text{surr}}$ and hence

$$0 = \Delta S \geq \oint \frac{đQ_{\text{sys}}}{T} = -\oint \frac{đQ_{\text{surr}}}{T}.$$

Since the heat transfer with the surroundings is reversible, we can equate the last quantity $\oint đQ_{\text{surr}}/T$ with ΔS_{surr}. Thus, we find

$$0 \leq -\Delta S_{\text{surr}} \quad \text{or} \quad \Delta S_{\text{surr}} \geq 0.$$

Thus, although there was no change in entropy of the system itself, there was an increase in the entropy of the surroundings so that, for the Universe as a whole, the entropy increased.

As a second example, consider an irreversible change in going from states 1 to 2. Again we assume that the heat transfer with the surroundings takes place reversibly. Then, as above,

$$\Delta S_{\text{sys}} \geq \int_1^2 \frac{đQ_{\text{irr}}}{T} = -\int \frac{đQ_{\text{surr}}}{T} = -\Delta S_{\text{surr}},$$

$$\Delta S_{\text{sys}} + \Delta S_{\text{surr}} \geq 0.$$

Notice the implication of these examples. When irreversible processes are present in the cycle, although there might not be a change in the entropy of the system itself, the entropy of the Universe as a whole has increased. It is always important to consider the surroundings as well as the system itself in this type of argument.

These examples illustrate the origin of the popular expression of the two laws of thermodynamics due to Clausius, who in 1865 invented the word entropy, from the Greek word for transformation:

(1) The energy of the Universe is constant.
(2) The entropy of the Universe tends to a maximum.

The entropy change need not necessarily involve heat exchange. It is a measure of the *irreversibility* of the process. The Joule expansion of a perfect gas is a good example in that the gas expands to fill a larger volume whilst doing no work, that is, its internal energy U, and consequently its temperature T, are constant. The entropy change of the perfect gas at temperature T is $\int_A^B đQ/T$. Now

$$đQ = \mathrm{d}U + p\,\mathrm{d}V,$$

and hence for 1 mole of gas

$$đQ = C_V\,\mathrm{d}T + \frac{RT}{V}\mathrm{d}V.$$

Therefore,

$$\int_A^B \frac{đQ}{T} = C_V \int_A^B \frac{\mathrm{d}T}{T} + R\int_A^B \frac{\mathrm{d}V}{V},$$

that is,

$$S - S_0 = C_V \ln\frac{T}{T_0} + R\ln\frac{V}{V_0}. \qquad (11.52)$$

Since $T =$ constant in the Joule expansion of a perfect gas,

$$S - S_0 = R\ln\frac{V}{V_0}. \qquad (11.53)$$

Thus, although there is no heat flow, the system has undergone an irreversible change and, as a result, there must be an increase in entropy in reaching the new equilibrium state from the old. Equation (11.53) is a key result which we will meet again in the study of the origin of the concept of quanta. Notice also that the increase in entropy is associated with the change in the volume of real space accessible to the system. We will find a corresponding result for the more general case of phase space in Chapter 12.

Finally, let us write down the entropy change of a perfect gas in an interesting form. From equation (11.52),

$$\begin{aligned} S - S_0 &= C_V \ln\frac{pV}{p_0V_0} + R\ln\frac{V}{V_0} \\ &= C_V \ln\frac{p}{p_0} + (C_V + R)\ln\frac{V}{V_0} \\ &= C_V \ln\frac{pV^\gamma}{p_0V_0^\gamma}. \end{aligned} \qquad (11.54)$$

where, as usual, $\gamma = (C_V + R)/C_V$. Thus, if the expansion is adiabatic, $pV^\gamma =$ constant, there is no entropy change. For this reason, adiabatic expansions are often called *isentropic* expansions. Note also that this provides us with an interpretation of the entropy function – *isotherms* are curves of constant temperature while *adiabats* are curves of constant entropy.

11.8 The Differential Form of the Combined First and Second Laws of Thermodynamics

At last, we have arrived at the equation which synthesises the physical content of the first and second laws of thermodynamics. For reversible processes,

$$\mathrm{d}S = \frac{đQ}{T},$$

and hence, combining this with the relation $đQ = \mathrm{d}U + p\,\mathrm{d}V$,

$$T\mathrm{d}S = \mathrm{d}U + p\,\mathrm{d}V. \tag{11.55}$$

More generally, if we write the work done on the system as

$$\sum_i X_i\,\mathrm{d}x_i,$$

we find that

$$T\mathrm{d}S = \mathrm{d}U + \sum_i X_i\,\mathrm{d}x_i. \tag{11.56}$$

The remarkable thing about this formula is that it combines the first and second laws entirely in terms of functions of state, and hence the relation must be true for all changes.

We will not take this story any further, but we should note that the relations (11.55) and (11.56) are the origins of some very powerful results which we will need later. In particular, for any gas, we can write

$$\left(\frac{\partial S}{\partial U}\right)_V = \frac{1}{T}, \tag{11.57}$$

that is, the partial derivative of the entropy with respect to U at constant volume defines the thermodynamic temperature T. This is of particular significance in statistical mechanics, because the concept of entropy S enters very early in the development of statistical mechanics and the internal energy U is one of the first things we need to determine statistically. Thus, we already have the key relation needed to define the concept of temperature in statistical mechanics.

Appendix to Chapter 11: Maxwell's Relations and Jacobians

This appendix includes some mathematical tools which are useful in dealing with problems in classical thermodynamics. There is no pretence at completeness or mathematical rigour in what follows.

A11.1 Perfect Differentials in Thermodynamics

The use of perfect differentials occurs naturally in classical thermodynamics. The reason is that, for much of the discussion, we consider two-coordinate systems, in which the thermodynamic equilibrium state is completely defined by only two functions of state. When the system changes from one equilibrium state to another, the change in any function of state depends only upon the initial and final coordinates. Thus, if z is a function of state which is defined completely by two other functions of state, x and y, the differential change of state can be written

$$\mathrm{d}z = \left(\frac{\partial z}{\partial x}\right)_y \mathrm{d}x + \left(\frac{\partial z}{\partial y}\right)_x \mathrm{d}y. \tag{A11.1}$$

The fact that the differential change dz does not depend upon the way in which we make the incremental steps in dx and dy enables us to set restrictions upon the functional dependence of z upon x and y. This relation is most simply demonstrated by the two ways in which the differential change dz illustrated in Fig. A11.1 can be made. We can reach C from A either via D or B. Since z is a function of state, the change dz must be the same along the paths of ABC and ADC. Along the path ADC, we first move dx in the x direction and then dy from the point $x + \mathrm{d}x$, that is

$$\begin{aligned} z(C) &= z(A) + \left(\frac{\partial z}{\partial x}\right)_y \mathrm{d}x + \left\{\frac{\partial}{\partial y}\left[z + \left(\frac{\partial z}{\partial x}\right)_y \mathrm{d}x\right]\right\}_x \mathrm{d}y \\ &= z(A) + \left(\frac{\partial z}{\partial x}\right)_y \mathrm{d}x + \left(\frac{\partial z}{\partial y}\right)_x \mathrm{d}y + \frac{\partial^2 z}{\partial y\,\partial x}\mathrm{d}x\,\mathrm{d}y, \end{aligned} \tag{A11.2}$$

where by $\partial^2 z/\partial y\,\partial x$ we mean

$$\left[\frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right)_y\right]_x .$$

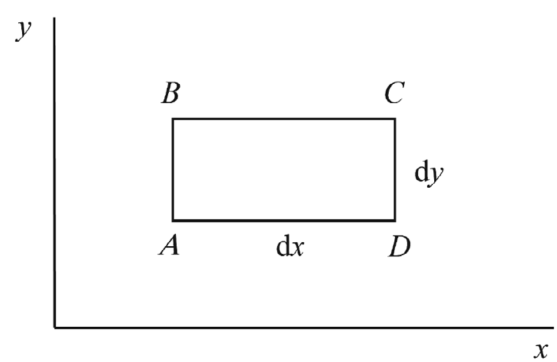


Fig. A11.1 Illustrating two ways of reaching the state C from A.

Along the path ABC, the value of $z(C)$ is

$$z(C) = z(A) + \left(\frac{\partial z}{\partial y}\right)_x \mathrm{d}x + \left\{\frac{\partial}{\partial x}\left[z + \left(\frac{\partial z}{\partial y}\right)_x \mathrm{d}y\right]\right\}_y \mathrm{d}x$$

$$= z(A) + \left(\frac{\partial z}{\partial y}\right)_x \mathrm{d}y + \left(\frac{\partial z}{\partial x}\right)_y \mathrm{d}x + \frac{\partial^2 z}{\partial x\,\partial y}\,\mathrm{d}x\,\mathrm{d}y. \quad \text{(A11.3)}$$

Notice that, in (A11.2) and (A11.3), the order in which the double partial derivatives is taken is important. Since equations (A11.2) and (A11.3) must be identical, the function z must have the property that

$$\frac{\partial^2 z}{\partial y\,\partial x} = \frac{\partial^2 z}{\partial x\,\partial y}, \quad \text{that is,} \quad \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right). \quad \text{(A11.4)}$$

This is the mathematical condition that z is a perfect differential of x and y.

A11.2 Maxwell's Relations

We have introduced quite a number of functions of state: p, V, T, S, U, H, and so on. Indeed, we could go on to define an infinite number of functions of state by combinations of these functions. When differential changes in the state of a system are made, there are corresponding changes in all the functions of state. The four most important of these differential relations are called *Maxwell's relations*. Let us derive two of them.

Consider first the equation (11.55) which relates differential changes in S, U and V:

$$T\,\mathrm{d}S = \mathrm{d}U + p\,\mathrm{d}V \quad \text{and so} \quad \mathrm{d}U = T\,\mathrm{d}S - p\,\mathrm{d}V. \quad \text{(A11.5)}$$

Since $\mathrm{d}U$ is the differential of a function of state, it must be a perfect differential and hence

$$\mathrm{d}U = \left(\frac{\partial U}{\partial S}\right)_V \mathrm{d}S + \left(\frac{\partial U}{\partial V}\right)_S \mathrm{d}V. \quad \text{(A11.6)}$$

Comparison with (A11.5) shows that

$$T = \left(\frac{\partial U}{\partial S}\right)_V; \quad p = -\left(\frac{\partial U}{\partial V}\right)_S. \quad \text{(A11.7)}$$

Since $\mathrm{d}U$ is a perfect differential, we also know from (A11.4) that

$$\frac{\partial}{\partial V}\left(\frac{\partial U}{\partial S}\right) = \frac{\partial}{\partial S}\left(\frac{\partial U}{\partial V}\right). \quad \text{(A11.8)}$$

Substituting (A11.7) into (A11.8),

$$\left(\frac{\partial T}{\partial V}\right)_S = -\left(\frac{\partial p}{\partial S}\right)_V. \quad \text{(A11.9)}$$

This is the first of four similar relations between T, S, p and V. We can follow exactly the same procedure for the enthalpy H:

$$\begin{aligned} H &= U + pV, \\ \mathrm{d}H &= \mathrm{d}U + p\,\mathrm{d}V + V\,\mathrm{d}p \\ &= T\,\mathrm{d}S + V\,\mathrm{d}p. \end{aligned} \quad \text{(A11.10)}$$

By exactly the same mathematical procedure, we find

$$\left(\frac{\partial T}{\partial p}\right)_S = \left(\frac{\partial V}{\partial S}\right)_p . \tag{A11.11}$$

We can repeat the analysis for two other important functions of state, the *Helmholtz free energy* $F = U - TS$ and the *Gibbs free energy* $G = U - TS + pV$ and find the other two Maxwell relations:

$$\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial p}{\partial T}\right)_V , \tag{A11.12}$$

$$\left(\frac{\partial V}{\partial T}\right)_p = -\left(\frac{\partial S}{\partial p}\right)_T . \tag{A11.13}$$

The functions F and G are particularly useful in studying processes which occur at constant temperature and constant pressure respectively.

The set of relations (A11.9), (A11.11), (A11.12) and (A11.13) are known as *Maxwell's relations* and are very useful in many thermodynamic problems. We will use them in subsequent chapters. It is useful to have a mnemonic by which to remember them and Jacobians provide one way of remembering them.

A11.3 Jacobians in Thermodynamics

Jacobians are found in transforming products of differentials between different systems of coordinates. In a simple example, suppose the volume element $\mathrm{d}V = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$ is to be transformed into some other coordinate system in which the differentials of the coordinates are $\mathrm{d}u$, $\mathrm{d}v$ and $\mathrm{d}w$. Then, the volume element can be written

$$\mathrm{d}V_{uvw} = \left|\frac{\partial(x,y,z)}{\partial(u,v,w)}\right| \mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w, \tag{A11.14}$$

where the Jacobian is defined to be[21]

$$\frac{\partial(x,y,z)}{\partial(u,v,w)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & \frac{\partial z}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & \frac{\partial z}{\partial v} \\ \frac{\partial x}{\partial w} & \frac{\partial y}{\partial w} & \frac{\partial z}{\partial w} \end{vmatrix} . \tag{A11.15}$$

We only need the two-coordinate version of these relations for use with Maxwell's equations. If the variables x, y and u, v are related by

$$\begin{cases} x = x(u,v), \\ y = y(u,v), \end{cases} \tag{A11.16}$$

the Jacobian is defined to be the determinant

$$\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial y}{\partial u} \\ \dfrac{\partial x}{\partial v} & \dfrac{\partial y}{\partial v} \end{vmatrix}. \tag{A11.17}$$

Then it is straightforward to show from the properties of determinants that

$$\begin{aligned} &\frac{\partial(x,y)}{\partial(x,y)} = -\frac{\partial(y,x)}{\partial(x,y)} = 1, \\ &\frac{\partial(v,v)}{\partial(x,y)} = 0 = \frac{\partial(k,v)}{\partial(x,y)} \quad \text{if } k \text{ is a constant,} \\ &\frac{\partial(u,v)}{\partial(x,y)} = -\frac{\partial(v,u)}{\partial(x,y)} = \frac{\partial(-v,u)}{\partial(x,y)} = \frac{\partial(v,-u)}{\partial(x,y)}, \end{aligned} \tag{A11.18}$$

and so on. There are similar rules for changes 'downstairs'. Notice, in particular, that

$$\frac{\partial(u,y)}{\partial(x,y)} = \left(\frac{\partial u}{\partial x}\right)_y, \tag{A11.19}$$

$$\frac{\partial(u,v)}{\partial(x,y)} = \frac{\partial(u,v)}{\partial(r,s)} \times \frac{\partial(r,s)}{\partial(x,y)} = \frac{1}{\partial(x,y)/\partial(u,v)}, \tag{A11.20}$$

and that

$$\left(\frac{\partial y}{\partial x}\right)_z \left(\frac{\partial z}{\partial y}\right)_x \left(\frac{\partial x}{\partial z}\right)_y = -1. \tag{A11.21}$$

Riley, Hobson and Bence give the general n-coordinate relations equivalent to (A11.17) to (A11.21).[22]

The value of the Jacobian notation is that we can write all four of Maxwell's equations in the compact form

$$\frac{\partial(T,S)}{\partial(x,y)} = \frac{\partial(p,V)}{\partial(x,y)}, \tag{A11.22}$$

where x and y are different variables. Let us give an example of how this works. From (A11.19), Maxwell's relation

$$\left(\frac{\partial T}{\partial V}\right)_S = -\left(\frac{\partial p}{\partial S}\right)_V \tag{A11.23}$$

is exactly the same as

$$\frac{\partial(T,S)}{\partial(V,S)} = -\frac{\partial(p,V)}{\partial(S,V)} = \frac{\partial(p,V)}{\partial(V,S)}. \tag{A11.24}$$

We can generate all four Maxwell relations by introducing the four appropriate combinations of T, S, p and V into the denominators of (A11.22). By virtue of (A11.20), (A11.22) is exactly the same as

$$\frac{\partial(T,S)}{\partial(p,V)} = 1. \tag{A11.25}$$

The Jacobian (A11.25) is the key to remembering all four relations. We need only remember that the Jacobian is +1 if T, S, p and V are written in the above order. One way of remembering this is that the intensive variables, T and p, and the extensive variables, S and V, should appear in the same order in the numerator and denominator if the sign is positive.

Notes

1 The textbook by Allen and Maxwell contains splendid historical material and quotations used in this chapter, as well as excellent illustrations of many of the experiments which laid the foundation for classical thermodynamics. Allen, H.S. and Maxwell, R.S. (1939). *A Text-book of Heat*, Parts I and II. London: MacMillan & Co.
2 Allen, H.S. and Maxwell, R.S. (1939). *A Text-book of Heat*, Part I, p. 12. London: MacMillan & Co.
3 Allen, H.S. and Maxwell, R.S. (1939). *op. cit.*, p. 14.
4 Allen, H.S. and Maxwell, R.S. (1939). *op. cit.*, p. 14.
5 Allen, H.S. and Maxwell, R.S. (1939). *op. cit.*, pp. 14–15.
6 Lavoisier, A. (1789). *Traité Élémentaire de Chimie, présenté dans un ordre nouveau, et d'aprs des découvertes modernes*, 1st edition. Paris: Cuchet, Libraire.
7 Brown, G.I. (1999). *Scientist, Soldier, Statesman, Spy: Count Rumford*. Stroud, Gloucestershire, UK: Sutton Publishing.
8 Fourier, J.B.J (1822). *Analytical Theory of Heat*, trans. A. Freeman (reprint edition, New York, 1955).
9 Joule, J.P. (1843). See *The Scientific Papers of James Prescott Joule*, 2 volumes. London 1884–1887 (reprint, London, 1963).
10 Allen, H.S. and Maxwell, R.S. (1939). *op. cit.*, p. 288.
11 Allen, H.S. and Maxwell, R.S. (1939). *op. cit.*, p. 288.
12 Harman, P.M. (1952). *Energy, Force and Matter: The Conceptual Development of Nineteenth Century Physics*, p. 45. Cambridge: Cambridge University Press.
13 Pippard, A.B. (1966). *The Elements of Classical Thermodynamics*. Cambridge: Cambridge University Press.
14 Allen, H.S. and Maxwell, R.S. (1939). *op. cit.*, Vol. II, p. 575.
15 Forbes, R.J. (1958). In *A History of Technology*, Vol. IV, eds. C. Singer, E.J. Holmyard, A.R. Hall and T.I. Williams, p. 165. Oxford: Clarendon Press.
16 For splendid articles on the development of the steam engine during the industrial revolution, see the articles by R.J. Forbes and H.W. Dickenson in *A History of Technology*, Vol. IV, eds. C. Singer, E.J. Holmyard, A.R. Hall and T.I. Williams (1958). Oxford: Clarendon Press.
17 Forbes, R.J. (1958). *op. cit.*, p. 164.
18 Carnot, N.L.S. (1924). *Réflexions sure la Puissance Motrice du Feu et sur les Machine Propres à Developper cette Puissance*. Paris: Bachelier.
19 Peter Harman's authoritative study provides key material about the history of the intellectual background to many of the theoretical issues discussed in this chapter. Harman, P.M. (1982). *Energy, Force and Matter: The Conceptual Development of Nineteenth Century Physics*. Cambridge: Cambridge University Press.
20 Feynman, R.P (1963). *The Feynman Lectures on Physics*, Vol. 1, eds. R.P. Feynman, R.B. Leighton and M. Sands, Chapter 44. Redwood City, California: Addison-Wesley Publishing Co.
21 See Riley, K.F., Hobson, M.P. and Bence, S.J. (2006). *Mathematical Methods for Physics and Engineering*, 3rd edition, pp. 200–207. Cambridge: Cambridge University Press.
22 See Riley, K.F., Hobson, M.P. and Bence, S.J. (2006). *op. cit.*, pp. 200–207.

12 Kinetic Theory and the Origin of Statistical Mechanics

12.1 The Kinetic Theory of Gases

The controversy between the caloric and the kinetic, or dynamic, theories of heat was resolved in favour of the latter by James Joule's experiments of the 1840s and by the establishment of the two laws of thermodynamics by Rudolph Clausius in the early 1850s. Before 1850, various kinetic theories of gases had been proposed, in particular, by John Herapath and John James Waterston. Joule had noted that the equivalence of heat and work could be interpreted in terms of the kinetic theory and, in his first formulation of the two laws of thermodynamics, Clausius described how the laws could be interpreted within this framework, although he emphasised that the laws are quite independent of the particular microscopic theory.

The first systematic account of the kinetic theory was published by Clausius in 1857 in a classic paper entitled *The Nature of the Motion which we call Heat*.[1] This contained a simple derivation of the perfect gas law, assuming that the molecules of a gas are elastic spheres which exert a pressure on the walls of the containing vessel by colliding with them.

12.1.1 Assumptions Underlying the Kinetic Theory of Gases

The objective of the kinetic theory is to apply *Newtonian physics* to large assemblies of atoms or molecules.[2] It turns out that many of the key results are independent of the velocity distribution of the molecules, the approach adopted by Clausius.

The basic postulates of the kinetic theory are as follows:

- Gases consist of elementary entities, molecules, which are in constant motion. Each molecule has kinetic energy $\frac{1}{2}mv^2$ and their velocity vectors point in random directions.
- The molecules are taken to be very small solid spheres with finite diameters a.
- The long-range forces between molecules are weak and undetectable in a Joule expansion and so they are taken to be zero. The molecules can, however, collide and when they do so they collide *elastically*, meaning that there is no loss of the total kinetic energy in each collision.
- The *pressure* on the wall of a vessel is the force per unit area due to the elastic collisions of enormous numbers of molecules with the wall – no kinetic energy is lost if the wall is stationary.

- The temperature is related to the average kinetic energy of the molecules of the gas. This makes sense in that, if work is done on the gas, the kinetic energies of the molecules are increased, thus increasing the internal energy and hence the temperature.

One key aspect of Clausius's model was the role of *collisions* between molecules. In his paper of 1857, he found that the typical speeds of the molecules of air were 461 and 492 m s^{-1} for oxygen and nitrogen respectively. The Dutch meteorologist Buys Ballot criticised this aspect of the theory, since it is well known that pungent odours take minutes to permeate a room, and not seconds. Clausius's response was to point out that the molecules of air collide with each other and therefore they must *diffuse* from one location to another, rather than travel in straight lines. In his paper of 1858, Clausius introduced the concept of the *mean free path* of the molecules of the gas.[3] Thus, in the kinetic theory, it is supposed that there are continually collisions between the molecules.

An important consequence of these collisions is that there must inevitably be a *velocity dispersion* among the molecules of a gas. If the collision were head-on, all the kinetic energy would be transferred from the incident to a stationary molecule. If it were a glancing collision, there would be very little transfer of energy. As a result, we cannot consider all the molecules of the gas to have the same speed – random collisions inevitably lead to a velocity dispersion. The collisions also have the effect of *randomising* the velocity vectors of the molecules and so the pressure of the gas is *isotropic*.

12.2 Kinetic Theory of Gases: First Version

Let us begin with an elementary analysis. Consider the force acting on the flat wall of a cubical enclosure resulting from the impact of the molecules of a gas. According to Newton's second law of motion, the pressure is the rate of change of momentum per unit area of all the molecules striking the wall elastically per second.

For elastic collisions of the molecules with unit area of, say, the y–z plane, the pressure is associated with changes in the x-components of their momentum vectors, $\Delta p_x = 2p_x = 2mv_x$. All the molecules with positive values of v_x within a distance v_x of the wall, arrive at the wall in one second. Therefore, the pressure p is found by summing over all the molecules within distance v_x of the wall which are moving towards it, that is,

$$p = \sum_i 2mv_x \times v_x = \sum_i 2mv_x^2 = n(+v_x)2m\overline{v_x^2}, \qquad (12.1)$$

where $n(+v_x)$ is the number density of molecules with v_x components in the positive x-direction and $\overline{v_x^2}$ is the mean square velocity of the molecules in the x-direction. Because of the isotropy of their motion, only half of the molecules have positive values of v_x. Therefore, if n is the total number density of molecules $n(+v_x) = n/2$. Hence, $p = nm\overline{v_x^2}$. Because the distribution of velocities is isotropic,

$$\overline{v^2} = \overline{v_x^2} + \overline{v_y^2} + \overline{v_y^2} \qquad \text{and} \qquad \overline{v_x^2} = \overline{v_y^2} = \overline{v_y^2}.$$

Therefore, $\overline{v_x^2} = \frac{1}{3}\overline{v^2}$ and

$$p = \tfrac{1}{3}nm\overline{v^2}. \tag{12.2}$$

We recognise that $\frac{1}{2}nm\overline{v^2}$ is the total kinetic energy of the molecules of the gas per unit volume. The perfect gas law for one mole of gas is $pV = RT$, or

$$p = (R/V)T = nk_{\rm B}T, \tag{12.3}$$

where $k_{\rm B} = R/N_0$ and N_0 is Avogadro's number, the number of molecules per mole; $k_{\rm B}$ is the gas constant per molecule, or *Boltzmann's constant*, and has the value $k_{\rm B} = 1.38 \times 10^{-23}$ J K^{-1}. Comparing (12.2) and (12.3),

$$\tfrac{3}{2}k_{\rm B}T = \tfrac{1}{2}m\overline{v^2}. \tag{12.4}$$

This result provides the direct link between the temperature T of the gas and the *mean kinetic energy* $\frac{1}{2}m\overline{v^2}$ *of the molecules*.

12.3 Kinetic Theory of Gases: Second Version

Let us carry out an improved calculation in which we integrate over all the angles at which molecules arrive at the walls of the vessel. The determination of the flux of molecules arriving at the wall can be split into two parts.

12.3.1 Probability of Molecules Arriving from a Particular Direction

This calculation turns up in many different problems in physics. We choose a particular direction in space and ask,

> What is the probability distribution of molecules having velocity vectors pointing within the range of angles θ to $\theta + {\rm d}\theta$ with respect to a chosen direction?

Because of collisions, the velocity vectors of the molecules point in random directions and so there is an equal probability of the velocity vectors pointing in any direction. Therefore, the probability of the molecules having velocity vectors lying within some solid angle Ω can be written in terms of the area A subtended by the solid angle on the surface of a sphere of radius r:

$$p(A) = \frac{A}{4\pi r^2}. \tag{12.5}$$

The probability $p(\theta)\,{\rm d}\theta$ of the molecules having velocity vectors within angles θ to $\theta + {\rm d}\theta$ of our chosen direction is given by the ratio of the surface area of the annulus shown in Fig. 12.1(a), ${\rm d}A = (2\pi r \sin\theta) \times (r\,{\rm d}\theta) = 2\pi r^2 \sin\theta\,{\rm d}\theta$ to the surface area of the sphere $4\pi r^2$, that is,

$$p(\theta)\,{\rm d}\theta = \frac{{\rm d}A}{4\pi r^2} = \tfrac{1}{2}\sin\theta\,{\rm d}\theta. \tag{12.6}$$

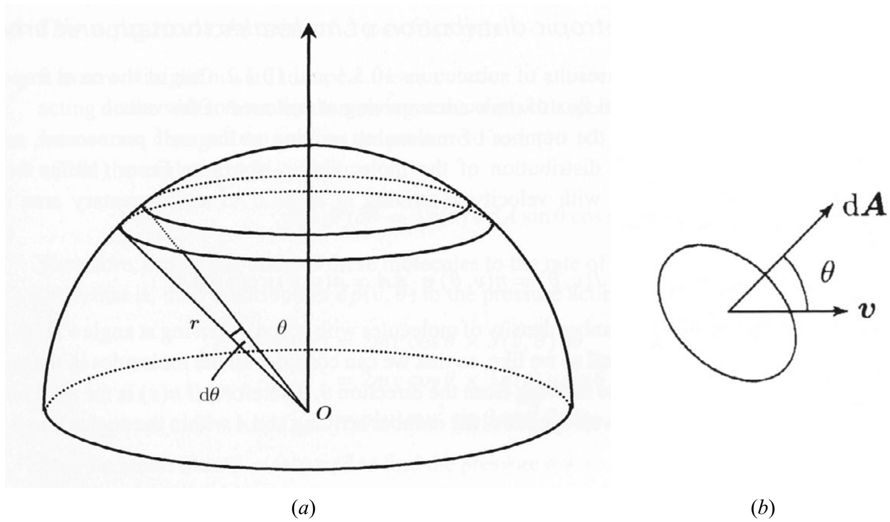


Fig. 12.1 (*a*) Evaluating the probability of molecules approaching the origin from directions in the range θ to $\theta + \mathrm{d}\theta$. (*b*) Illustrating the flux of molecules with velocity vectors $\boldsymbol{v}$ passing through an element of surface area $\mathrm{d}\boldsymbol{A}$. The convention is used of associating a vector $\mathrm{d}\boldsymbol{A}$ with the area, normal to the plane of the area and with magnitude $|\mathrm{d}\boldsymbol{A}|$.

12.3.2 The Flux of Molecules Arriving from a Particular Direction

The second part of the calculation concerns the flux of molecules, all moving with velocity $\boldsymbol{v}$, through an element of surface area $\mathrm{d}\boldsymbol{A}$, as illustrated in Fig. 12.1(*b*); the number density of the molecules is n. The component of velocity parallel to the plane of the elementary area is $|\boldsymbol{v}| \sin\theta$, but this results in no flow of molecules through $\mathrm{d}\boldsymbol{A}$. The component of velocity parallel to $\mathrm{d}\boldsymbol{A}$, that is, perpendicular to the surface, is $|\boldsymbol{v}| \cos\theta$ and so the *flux of molecules* through $\mathrm{d}\boldsymbol{A}$, that is, number of molecules passing through $\mathrm{d}\boldsymbol{A}$ per second with velocity vectors $\boldsymbol{v}$, is

$$J = n|\boldsymbol{v}||\mathrm{d}\boldsymbol{A}| \cos\theta = n\boldsymbol{v} \cdot \mathrm{d}\boldsymbol{A}. \tag{12.7}$$

12.3.3 The Flux of an Isotropic Distribution of Molecules through Unit Area

We now combine the results of Sections 12.3.1 and 12.3.2. First, we evaluate the number of molecules arriving at the wall of a box per second, assuming the velocity distribution of the molecules is isotropic. From (12.7), the flux of molecules $J(v, \theta)$ with velocity $\boldsymbol{v}$ arriving at angle θ to an elementary area $\mathrm{d}\boldsymbol{A}$ is

$$J(v, \theta) = n(v, \theta)\, \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{A} = n(v, \theta)\, v \cos\theta \, \mathrm{d}A, \tag{12.8}$$

where $n(v, \theta)$ is the number density of molecules with speed v arriving from the direction θ. $\mathrm{d}\theta$ can be taken to be infinitesimally small and so, if $n(v)$ is the total number density of molecules with speeds v, the number arriving within the angles θ to $\theta + \mathrm{d}\theta$ is

$$n(v,\theta)\,\mathrm{d}\theta = n(v)p(\theta)\,\mathrm{d}\theta = \frac{n(v)}{2}\sin\theta\,\mathrm{d}\theta. \tag{12.9}$$

Substituting (12.9) into (12.8),

$$J(v,\theta)\,\mathrm{d}\theta = \frac{n(v)}{2}\sin\theta\,\mathrm{d}\theta \times v\cos\theta\,\mathrm{d}A. \tag{12.10}$$

The total flux of molecules with speed v arriving at the area dA of the wall is therefore the integral of $J(v,\theta)$ from $\theta = 0$ to $\theta = \pi/2$, since those molecules with velocity vectors between $\pi/2$ and π are moving away from the wall. Therefore,

$$J(v) = \int_0^{\pi/2} \frac{n(v)\,v\,\mathrm{d}A}{2}\sin\theta\cos\theta\,\mathrm{d}\theta = \frac{n(v)\,v\,\mathrm{d}A}{4}. \tag{12.11}$$

The flux of molecules per unit area with speed v is therefore

$$J(v) = \frac{n(v)\,v}{4}. \tag{12.12}$$

We now add together the contributions from molecules with all speeds to find the total flux arriving at the wall per unit area:

$$J = \sum_v \frac{n(v)\,v}{4} = \tfrac{1}{4}n\,\overline{v}, \tag{12.13}$$

where $\overline{v}$ is their mean speed,

$$\overline{v} = \frac{1}{n}\sum_v n(v)v, \tag{12.14}$$

and n their total number density. If the distribution of speeds were given by a continuous probability distribution such that the probability of the speed lying in the range v to $v + \mathrm{d}v$ is $f(v)\,\mathrm{d}v$ and $\int_0^\infty f(v)\,\mathrm{d}v = 1$, the average speed would be

$$\overline{v} = \int_0^\infty v f(v)\,\mathrm{d}v. \tag{12.15}$$

12.3.4 The Equation of State of a Perfect Gas

It is now straightforward to repeat the calculation of Section 12.3.3 to find the pressure acting on the walls. Each molecule contributes a momentum change $\Delta p_x = 2mv\cos\theta$ to the pressure. From (12.10), the flux of molecules arriving within the angles θ to $\theta + \mathrm{d}\theta$ is

$$J(v,\theta)\,\mathrm{d}\theta = \frac{n(v)\,v\,\mathrm{d}A}{2}\sin\theta\cos\theta\,\mathrm{d}\theta. \tag{12.16}$$

Therefore, the contribution of these molecules to the rate of change of momentum *per unit area* acting on the walls of the vessel is

$$\begin{aligned}\mathrm{d}p(v,\theta)\,\mathrm{d}\theta &= \Delta p_x \times J(v,\theta)\,\mathrm{d}\theta = 2mv\cos\theta \times \frac{n(v)\,v}{2}\sin\theta\cos\theta\,\mathrm{d}\theta \\ &= n(v)\,mv^2\sin\theta\cos^2\theta\,\mathrm{d}\theta.\end{aligned} \tag{12.17}$$

Integrating from $\theta = 0$ to $\pi/2$, the pressure due to all the molecules with speed v is

$$p(v) = n(v)\, mv^2 \int_0^{\pi/2} \sin\theta \cos^2\theta \, \mathrm{d}\theta = \tfrac{1}{3} n(v)\, mv^2. \tag{12.18}$$

Now sum over all speeds to find the total pressure of the gas:

$$p = \sum_v \tfrac{1}{3} n(v)\, mv^2 = \tfrac{1}{3} nm\overline{v^2}, \tag{12.19}$$

where

$$\overline{v^2} = \frac{1}{n} \sum_v n(v) v^2 \tag{12.20}$$

is the *mean square speed of the molecules*, and n is their total number density. This is a more satisfactory derivation of (12.2). If the distribution of molecular speeds were continuous and given by the probability distribution $f(v)\,\mathrm{d}v$, the mean square speed would be

$$\overline{v^2} = \int_0^{\infty} v^2 f(v)\, \mathrm{d}v. \tag{12.21}$$

As before, comparing (12.19) with the perfect gas law $p = k_{\mathrm{B}}T$, we find $\frac{3}{2}k_{\mathrm{B}}T = \frac{1}{2}m\overline{v^2}$.

If the vessel contained a mixture of different gases at the same temperature, each molecule would attain the same kinetic energy because of molecular collisions. Therefore, the total pressure from the different species i is

$$p = \sum_i \tfrac{1}{3} n_i m_i \overline{v_i^2} = \sum p_i, \tag{12.22}$$

where p_i is the pressure contributed by each species. This is *Dalton's law of partial pressures*.

It has been assumed that the internal energy of the gas is entirely associated with the kinetic energies of the molecules, that is, for one mole,

$$U = \tfrac{1}{2} N_0 m\overline{v^2} = \tfrac{3}{2} N_0 k_{\mathrm{B}} T, \tag{12.23}$$

where N_0 is Avogadro's number. Therefore,

$$U = \tfrac{3}{2} RT, \tag{12.24}$$

and the heat capacity at constant volume is

$$C_V = \left(\frac{\partial U}{\partial T}\right)_V = \tfrac{3}{2}. \tag{12.25}$$

The ratio of specific heat capacities is therefore

$$\gamma = \frac{C_V + R}{C_V} = \tfrac{5}{3}. \tag{12.26}$$

This argument is so taken for granted nowadays that one forgets what an achievement it represented in 1857. Whilst accounting for the perfect gas law admirably, it did not give good agreement with the known values of γ for common molecular gases such as oxygen and nitrogen for which $\gamma \approx 1.4$. There must therefore exist other ways of storing energy within molecular gases which can increase the internal energy per molecule, an important point clearly recognised by Clausius in the last sentences of his great paper.

Two aspects of this story are worth noting. First, Clausius knew that there had to be a dispersion of speeds $f(v)$ of the molecules of the gas, but he did not know what it was and so worked in terms of mean squared values only. Second, the idea that the pressure of the gas should be proportional to $nm\overline{v^2}$ had been worked out 15 years earlier by Waterston. As early as 1843, he wrote in a book published in Edinburgh,

> A medium constituted of elastic spherical molecules that are continually impinging against each other with the same velocity, will exert against a vacuum an elastic force that is proportional to the square of this velocity and to its density.[4]

In December 1845, Waterston developed a more systematic version of his theory and gave the first statement of the equipartition theorem for a mixture of gases of different molecular species. He also worked out the ratio of specific heat capacities, although his calculation contained a numerical error. Waterston's work was submitted to the Royal Society for publication in 1845, but was harshly refereed and remained unpublished until 1892, eight years after his death. The manuscript was discovered by Lord Rayleigh in the archives of the Royal Society of London and it was then published with an introduction by Rayleigh himself. He wrote,

> Impressed by the above passage [from a later work by Waterston] and with the general ingenuity and soundness of Waterston's views, I took the first opportunity of consulting the Archives, and saw at once that the memoir justified the large claims made for it, and that it makes an immense advance in the direction of the now generally received theory. The omission to publish it was a misfortune, which probably retarded the development of the subject by ten or fifteen years.[5]

Later he notes,

> The character of the advance to be dated from this paper will be at once understood when it is realised that Waterston was the first to introduce into the theory the conception that heat and temperature are to be measured by 'vis viva' (that is, kinetic energy). . . . In the second section, the great feature is the statement (VII) that in mixed media the mean square velocity is inversely proportional to the specific weight of the molecules. The proof which Waterston gave is doubtless not satisfactory; but the same may be said of that advanced by Maxwell fifteen years later.[6]

One final quotation is worth pondering:

> The history of this paper suggests that highly speculative investigations, especially by an unknown author, are best brought before the scientific world through some other channel than a scientific society, which naturally hesitates to admit into its printed records matter of uncertain value. Perhaps one may go further and say that a young author who believes himself capable of great things would usually do well to secure the favourable recognition of the scientific world by work whose scope is limited, and whose value is easily judged, before embarking on greater flights.[7]

I fear this is not an isolated example of pioneering efforts being overlooked until they are promoted by a recognised authority. The case studies in this book bring vividly to light the struggles of even the greatest scientists in convincing the community of physicists of revolutionary new ideas.

12.4 Maxwell's Velocity Distribution

Both of Clausius's papers were known to Maxwell when he turned to the study of the kinetic theory of gases in 1859. His paper entitled *Illustrations of the Dynamical Theory of Gases* published in 1860 is characteristically novel and profound.[8] The amazing achievement was that, in this single paper, he derived the correct formula for the velocity distribution $f(v)$ of the molecules of a gas and at the same time introduced statistical concepts into the kinetic theory of gases and thermodynamics. Francis Everitt has written that this derivation of what we now know as Maxwell's velocity distribution marks the beginning of a new epoch in physics.[9] From Maxwell's analysis follow directly the concepts of the statistical nature of the laws of thermodynamics, the key to Boltzmann's statistics and the modern theory of statistical mechanics.

Maxwell's derivation of the distribution occupies no more than half a dozen short paragraphs. The problem is stated as Proposition IV of his paper:

> To find the average number of particles whose velocities lie between given limits, after a great number of collisions among a great number of equal particles.

We follow closely Maxwell's exposition with only minor changes of notation. The total number of molecules is N and the x, y and z components of their velocities are v_x, v_y and v_z. Maxwell supposed that the velocity distribution in the three orthogonal directions must be the same after a great number of collisions, that is,

$$N f(v_x)\,\mathrm{d}v_x = N f(v_y)\,\mathrm{d}v_y = N f(v_z)\,\mathrm{d}v_z, \tag{12.27}$$

where f is the same function. Now the three perpendicular components of the velocity are independent and hence the number of molecules with velocities in the range v_x to $v_x + \mathrm{d}v_x$, v_y to $v_y + \mathrm{d}v_y$ and v_z to $v_z + \mathrm{d}v_z$ is

$$N f(v_x) f(v_y) f(v_z)\,\mathrm{d}v_x\,\mathrm{d}v_y\,\mathrm{d}v_z. \tag{12.28}$$

But, the total velocity of any molecule v is $v^2 = v_x^2 + v_y^2 + v_z^2$. Because large numbers of collisions have taken place, the probability distribution for the total velocity v, $\phi(v)$, must be isotropic and depend only on v, that is,

$$f(v_x) f(v_y) f(v_z) = \phi(v) = \phi[(v_x^2 + v_y^2 + v_z^2)^{1/2}], \tag{12.29}$$

where $\phi(v)$ has been normalised so that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi(v)\,\mathrm{d}v_x\,\mathrm{d}v_y\,\mathrm{d}v_z = 1.$$

The relation (12.29) is a *functional equation* and we need to find forms of the functions $f(v_x)$, $f(v_y)$, $f(v_z)$ and $\phi(v)$ consistent with (12.29). By inspection, a suitable solution is

$$f(v_x) = C\,\mathrm{e}^{Av_x^2}, \quad f(v_y) = C\,\mathrm{e}^{Av_y^2}, \quad f(v_z) = C\,\mathrm{e}^{Av_z^2},$$

where C and A are constants. The beauty of using exponentials is that the powers to which they are raised are additive when the functions are multiplied together, as required by (12.29). Therefore,

$$\phi(v) = f(v_x)f(v_y)f(v_z) = C^3 e^{A(v_x^2+v_y^2+v_z^2)},$$
$$= C^3 e^{Av^2}. \tag{12.30}$$

The distribution must converge as $v \to \infty$ and hence A must be negative. Maxwell wrote $A = -\alpha^{-2}$ and so

$$\phi(v) = C^3 e^{-v^2/\alpha^2}, \tag{12.31}$$

where α is a constant to be determined. For each velocity component, the total probability of the molecule having some value of v_x must be unity and so

$$\int_{-\infty}^{\infty} C\, e^{-v_x^2/\alpha^2}\, dv_x = 1. \tag{12.32}$$

This is a standard integral, $\int_{-\infty}^{\infty} e^{-x^2}\, dx = \pi^{1/2}$, and hence

$$C = \frac{1}{\alpha\pi^{1/2}}. \tag{12.33}$$

This leads directly to four conclusions which we quote in Maxwell's own words, but using our notation.

1st. *The number of particles whose velocity, resolved in a certain direction, lies between* v_x *and* $v_x + dv_x$ *is*

$$N\frac{1}{\alpha\pi^{1/2}} e^{-v_x^2/\alpha^2}\, dv_x. \tag{12.34}$$

We have just proved this result.

2nd. *The number whose actual velocity lies between* v *and* $v + dv$ *is*

$$N\frac{4}{\alpha^3\pi^{1/2}} v^2 e^{-v^2/\alpha^2}\, dv.$$

This is found by transforming the product of the three expressions for $f(v_x)$, $f(v_y)$, $f(v_z)$ from Cartesian to polar coordinates in velocity space as follows:

$$\phi(v)\, dv_x\, dv_y\, dv_z = f(v_x)f(v_y)f(v_z)\, dv_x\, dv_y\, dv_z,$$
$$= \frac{1}{\alpha^3\pi^{3/2}} e^{-v^2/\alpha^2} dv_x\, dv_y\, dv_z. \tag{12.35}$$

But contributions to the volume elements dv_x, dv_y, dv_y associated with the velocity v all lie within the spherical shell of radius v and width dv in velocity space (Fig. 12.2), that is, the volume of velocity space in the range v to $v + dv$, which is $4\pi v^2\, dv$. Therefore,

$$\phi(v)\, dv = \frac{1}{\alpha^3\pi^{3/2}}\, e^{-v^2/\alpha^2} 4\pi v^2\, dv = \frac{4}{\alpha^3\pi^{1/2}}\, v^2 e^{-v^2/\alpha^2} dv, \tag{12.36}$$

as stated by Maxwell.

3rd. *To find the mean value of* v*, add the velocities of all the particles together and divide by the number of particles; the result is*

$$\textit{mean velocity} = \frac{2\alpha}{\pi^{1/2}}. \tag{12.37}$$

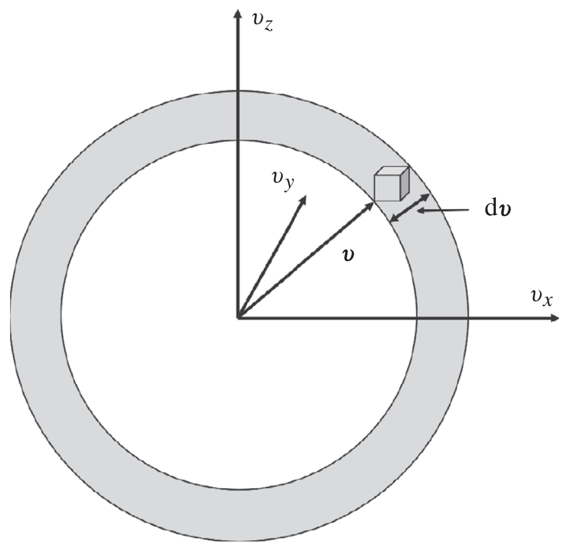


Fig. 12.2 Conversion of the volume integral over $\mathrm{d}v_x \mathrm{d}v_y \mathrm{d}v_z$ in velocity space to one over the total velocity v. The volume element of velocity space is $4\pi v^2\,\mathrm{d}v$.

This is a straightforward calculation, given the probability distribution $\phi(v)\,\mathrm{d}v$.

$$\overline{v} = \frac{\int_0^\infty v\,\phi(v)\,\mathrm{d}v}{\int_0^\infty \phi(v)\,\mathrm{d}v} = \int_0^\infty \frac{4}{\alpha^3\pi^{1/2}}\,v^3\,\mathrm{e}^{-v^2/\alpha^2}\,\mathrm{d}v.$$

This is a standard integral, $\int_0^\infty x^3 \mathrm{e}^{-x^2}\,\mathrm{d}x = 1/2$. Therefore, changing variables to $x = v/\alpha$, we find

$$\overline{v} = \frac{4}{\alpha^3\pi^{1/2}}\,\frac{\alpha^4}{2} = \frac{2\alpha}{\pi^{1/2}}.$$

4th. *To find the mean value of v^2, add all the values together and divide by N,*

$$\textit{mean value of } v^2 = \tfrac{3}{2}\alpha^2.$$

To find this result, take the variance of v by the usual procedure:

$$\overline{v^2} = \int_0^\infty v^2\phi(v)\,\mathrm{d}v = \int_0^\infty \frac{4}{\alpha^3\pi^{1/2}}\,v^4\mathrm{e}^{-v^2/\alpha^2}\,\mathrm{d}v. \tag{12.38}$$

This can be reduced to a standard form by integration by parts. Changing variables to $x = v/\alpha$, the integral (12.38) becomes

$$\overline{v^2} = \tfrac{3}{2}\alpha^2. \tag{12.39}$$

Maxwell immediately noted

> that the velocities are distributed among the particles according to the same law as the errors are distributed among the observations in the theory of the *method of least squares*.

Thus, in this very first paper, the direct relation to errors and statistical procedures was established.

To put Maxwell's distribution into its standard form, we compare the result (12.39) with the value for $\overline{v^2}$ deduced from the kinetic theory of gases (12.4) and find

$$\overline{v^2} = \tfrac{3}{2}\alpha^2 = \frac{3k_BT}{m}; \qquad \alpha = \sqrt{\frac{2k_BT}{m}}. \tag{12.40}$$

We can therefore write Maxwell's distribution (12.35) in its definitive form:

$$\phi(v)\,dv = 4\pi\left(\frac{m}{2\pi k_BT}\right)^{3/2} v^2 \exp\left(-\frac{mv^2}{2k_BT}\right) dv. \tag{12.41}$$

This is a quite astonishing achievement. It is a remarkably simple derivation of Maxwell's distribution which scarcely acknowledges the fact that we are dealing with molecules and gases at all. It is in striking contrast to the derivation which proceeds from the Boltzmann distribution. The Maxwell distribution is shown in Fig. 12.3 in dimensionless form.

Maxwell then applied this distribution law in a variety of circumstances and, in particular, he showed that, if two different types of molecule are present in the same volume, the mean *vis viva*, or kinetic energy, must be the same for each of them.

In the final section of the paper, Maxwell addressed the problem of accounting for the ratio of specific heat capacities of gases, which for many molecular gases had measured values $\gamma = 1.4$. Clausius had already noted that some additional means of storing *vis viva* was needed and Maxwell proposed that it be stored in the kinetic energy of rotation of the molecules, which he modelled as rough particles. He found that, in equilibrium,

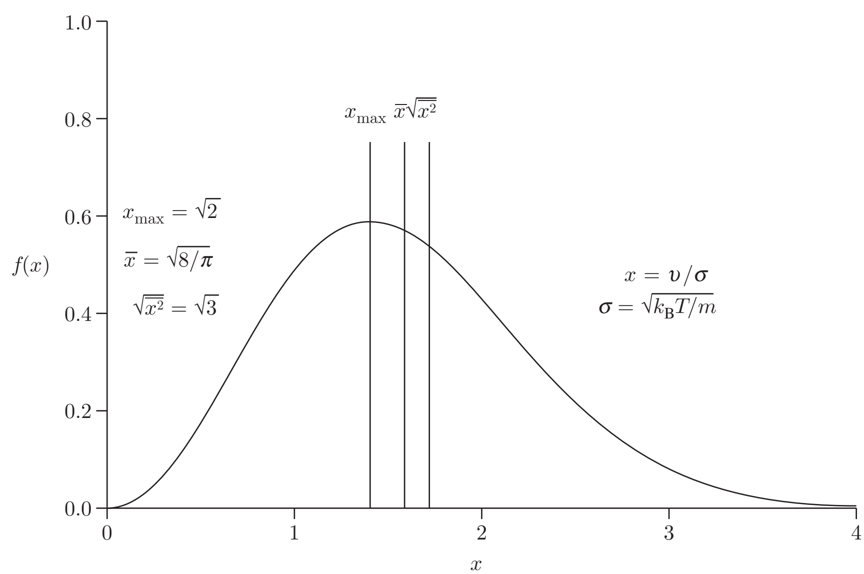


Fig. 12.3 The Maxwell distribution plotted in dimensionless form, $f(x)\,dx = \sqrt{2/\pi}x^2\,e^{-x^2/2}\,dx$, where $x = v/\sigma$ and σ is the standard deviation of the distribution. The mean, mean squared and maximum values of the distribution are indicated on the diagram.

as much energy could be stored in rotational as in translation motion, that is, $\frac{1}{2}I\omega^2 = \frac{1}{2}mv^2$ and, consequently, he could derive a value for the ratio of specific heat capacities. Instead of $U = \frac{3}{2}RT$, he found $U = 3RT$, since as much energy is stored in rotational as in translational motion in thermal equilibrium, and consequently, following through the analysis of equations (12.24) to (12.26),

$$\gamma = \tfrac{4}{3} = 1.333.$$

This value is almost as bad as the value 1.667 which follows if only translational motion is considered. This profoundly depressed Maxwell. The last sentence of his great paper reads:

> Finally, by establishing a necessary relation between the motions of translation and rotation of all particles not spherical, we proved that a system of such particles could not possibly satisfy the known relation between the two specific heats of all gases.

His inability to explain the value $\gamma = 1.4$ was a grave disappointment. In his report to the British Association for the Advancement of Science of 1860, he stated that this discrepancy 'overturned the whole hypothesis'.[10]

The second key concept which resulted directly from this paper was the principle of the equipartition of energy. This was to prove to be a highly contentious issue until it was finally resolved by Einstein's application of quantum concepts to the average energy of an oscillator (see Section 16.4). The principle states that in equilibrium energy is shared equally among all the independent modes in which energy can be stored by molecules. The problem of accounting for the value $\gamma = 1.4$ was aggravated by the discovery of spectral lines in the spectra of gases, which were interpreted as resonances associated with the internal structure of molecules. There were so many lines that if, each of them were to acquire its share of the internal energy of the gas, the ratio of specific heat capacities of gases would tend to 1. These two fundamental problems cast grave doubt upon the validity of the principle of equipartition of energy and in consequence upon the whole idea of a kinetic theory of gases, despite its success in accounting for the perfect gas law.

These issues were at the centre of much of Maxwell's subsequent work. In 1867, he presented another derivation of the velocity distribution from the perspective of its evolution as a result of molecular collisions.[11] He considered the outcome of collisions between molecules of the same mass which have velocities in the range $\boldsymbol{v}_1$ to $\boldsymbol{v}_1 + \mathrm{d}\boldsymbol{v}_1$ with those in the range $\boldsymbol{v}_2$ to $\boldsymbol{v}_2 + \mathrm{d}\boldsymbol{v}_2$. Conservation of momentum requires

$$\boldsymbol{v}_1 + \boldsymbol{v}_2 \to \boldsymbol{v}_1' + \boldsymbol{v}_2'.$$

Because of conservation of energy,

$$\tfrac{1}{2}mv_1^2 + \tfrac{1}{2}mv_2^2 = \tfrac{1}{2}mv'^2_1 + \tfrac{1}{2}mv'^2_2. \tag{12.42}$$

Since the molecules are assumed to have reached an equilibrium state, the joint probability distribution for the velocities before and after the collisions must satisfy the relation

$$f(\boldsymbol{v}_1)\,\mathrm{d}\boldsymbol{v}_1\, f(\boldsymbol{v}_2)\,\mathrm{d}\boldsymbol{v}_2 = f(\boldsymbol{v}_1')\,\mathrm{d}\boldsymbol{v}_1'\, f(\boldsymbol{v}_2')\,\mathrm{d}\boldsymbol{v}_2'. \tag{12.43}$$

To appreciate the force of this argument, it is simplest to inspect Maxwell's velocity distribution in Cartesian form (12.35). Conservation of kinetic energy ensures that the exponential terms on either side of (10.43) remain the same,

$$\exp\left[-\frac{m}{2k_{\mathrm{B}}T}\left(v_1^2 + v_2^2\right)\right] = \exp\left[-\frac{m}{2k_{\mathrm{B}}T}\left(v'^2_1 + v'^2_2\right)\right]. \tag{12.44}$$

In addition, Maxwell showed that the elements of phase space, or velocity space, before and after the collisions are related by

$$[\mathrm{d}v_x\,\mathrm{d}v_y\,\mathrm{d}v_z]_1 \times [\mathrm{d}v_x\,\mathrm{d}v_y\,\mathrm{d}v_z]_2 = [\mathrm{d}v'_x\,\mathrm{d}v'_y\,\mathrm{d}v'_z]_1 \times [\mathrm{d}v'_x\,\mathrm{d}v'_y\,\mathrm{d}v'_z]_2.$$

This is a special case of *Liouville's theorem*, according to which, if a certain group of molecules occupies a volume of phase space $\mathrm{d}v_x\,\mathrm{d}v_y\,\mathrm{d}v_z\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$, then, as they change their velocities and positions under the laws of classical mechanics, the volume of phase space is conserved. This result was known to Boltzmann as he struggled to establish the principles of statistical mechanics. Thus, Maxwell's velocity distribution can be retrieved from the requirement that the distribution be stationary following collisions between the molecules of the gas.

There is an intriguing footnote to this story. The exponential factor in Maxwell's velocity distribution is just $\exp(-E/k_{\mathrm{B}}T)$, where $E = \frac{1}{2}mv^2$ is the kinetic energy of the molecules. From (12.41), the ratio of the numbers of molecules in the velocity intervals v_i to $v_i + \mathrm{d}v_i$ and v_j to $v_j + \mathrm{d}v_j$ is

$$\frac{N(E_i)}{N(E_j)} = \frac{\exp(-E_i/k_{\mathrm{B}}T) \times 4\pi v_i^2\,\mathrm{d}v_i}{\exp(-E_j/k_{\mathrm{B}}T) \times 4\pi v_j^2\,\mathrm{d}v_j}. \tag{12.45}$$

This expression can be interpreted in terms of the probabilities that the molecules have energy E at temperature T and the volumes of velocity space available to those with velocities in the range v to $v + \mathrm{d}v$. This is a version of the much more general result which can be written

$$\frac{N(E_i)}{N(E_j)} = \frac{g_i}{g_j}\,\frac{\exp(-E_i/k_{\mathrm{B}}T)}{\exp(-E_j/k_{\mathrm{B}}T)}, \tag{12.46}$$

where the g_i and g_j describe the total numbers of available states in which the molecules would have energies E_i and E_j. In the language of statistical physics, the g_i describe the *degeneracy* of the state i, meaning the number of different states of the same energy E_i. In thermal equilibrium, the number of molecules of energy E not only depends upon the Boltzmann factor $p(E) \propto \exp(-E/k_{\mathrm{B}}T)$, but also upon the number of available states with energy E. We will provide a more rigorous derivation of this key result in Section 15.2 of Case Study V.

Despite the problems with the kinetic theory, Maxwell never seriously doubted its validity or that of the velocity distribution he had derived. Among his most compelling reasons were the results of his studies of the viscosity of gases.

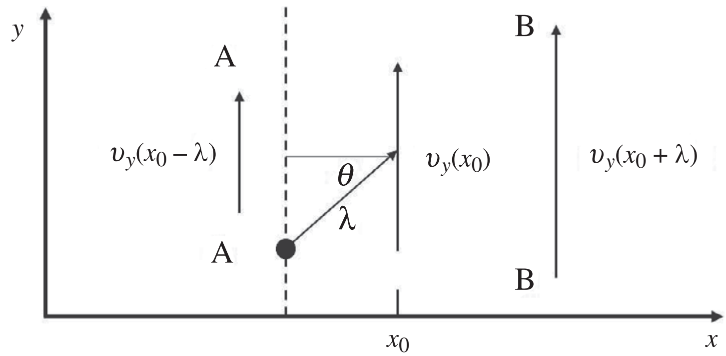


Fig. 12.4 Illustrating the geometry used to evaluate the coefficient of viscosity from the kinetic theory of gases.

12.5 The Viscosity of Gases

Maxwell immediately applied the kinetic theory of gases to their transport properties – diffusion, thermal conductivity and viscosity. His calculation of the coefficient of viscosity of gases was of special importance, the standard diagram for the transport of viscous stresses being shown in Fig. 12.4.

Gas flows in the positive y-direction and there is a velocity gradient in the x-direction, as indicated by the relative lengths of the arrows at AA, x_0 and BB. For illustrative purposes, the planes are separated by one mean free path λ. The object of the calculation is to work out the shear stress τ_{x_0}, that is, the force per unit area acting on the y–z plane in the y-direction at x_0. As discussed in Section 9.5, the stress depends upon the velocity gradient according to the expression

$$\tau_{x_0} = \eta \left(\frac{\mathrm{d}v_y}{\mathrm{d}x} \right)_{x_0}, \tag{12.47}$$

where η is the viscosity of the gas. At the molecular level, the stress is the net rate at which momentum is delivered to the interface per unit area per second by the random motions of the molecules on either side of the plane at x_0.

The molecules make many collisions with each other with mean free path λ, which is related to their collision cross-section σ by the usual relation $\lambda = (n\sigma)^{-1}$. In each collision it is a good approximation to assume that the angles though which the molecules are scattered are random. Therefore, the momentum transferred to the y–z plane at x_0 can be considered to have originated from a distance λ at some angle θ to the normal to the plane, as indicated in Fig. 12.4. From (12.10), the flux of molecules with speed v arriving in the angular range θ to $\theta + \mathrm{d}\theta$ per second is

$$J(v, \theta)\, \mathrm{d}\theta = \frac{n(v)\, v\, \mathrm{d}A}{2} \sin\theta \cos\theta\, \mathrm{d}\theta.$$

Each molecule brings with it a component of bulk momentum in the y-direction which it acquired at perpendicular distance $\lambda\cos\theta$ from the plane, $mv_y(x_0 - \lambda\cos\theta)$, and so the rate of transfer of momentum from such molecules per unit area is

$$\begin{aligned}\frac{\mathrm{d}p_y(v,\theta)\,\mathrm{d}\theta}{\mathrm{d}t} &= mv_y(x_0 - \lambda\cos\theta) \times \frac{n(v)\,v}{2}\sin\theta\cos\theta\,\mathrm{d}\theta \\ &= m\left[v_y(x_0) - \lambda\cos\theta\frac{\mathrm{d}v_y}{\mathrm{d}x}\right] \times \frac{n(v)\,v}{2}\sin\theta\cos\theta\,\mathrm{d}\theta,\end{aligned} \tag{12.48}$$

where we have made a Taylor expansion in deriving the second equality. The pleasant aspect of this calculation is that we can now integrate over angles from $\theta = \pi$ to $\theta = 0$ and so take account of the random arrival of molecules from both sides of the y–z plane at x_0:

$$\frac{\mathrm{d}p_y(v)}{\mathrm{d}t} = \tfrac{1}{2}mvn(v)\left[v_y(x_0)\int_\pi^0 \sin\theta\cos\theta\,\mathrm{d}\theta - \lambda\frac{\mathrm{d}v_y}{\mathrm{d}x}\int_\pi^0 \sin\theta\cos^2\theta\,\mathrm{d}\theta\right]. \tag{12.49}$$

The first integral in square brackets is zero and the second −2/3. Therefore, the net rate of transport of momentum per unit area at x_0 is

$$\tau_{x_0}(v) = \frac{\mathrm{d}p_y(v)}{\mathrm{d}t} = \tfrac{1}{3}mvn(v)\lambda\frac{\mathrm{d}v_y}{\mathrm{d}x}. \tag{12.50}$$

We now take averages over the speeds of the molecules,

$$\tau_{x_0} = \tfrac{1}{3}m\lambda\frac{\mathrm{d}v_y}{\mathrm{d}x}\int vn(v)\,\mathrm{d}v = \tfrac{1}{3}nm\overline{v}\lambda\frac{\mathrm{d}v_y}{\mathrm{d}x}, \tag{12.51}$$

where $\overline{v}$ is their mean speed. Therefore, from (12.47), the expression for the coefficient of viscosity is

$$\eta = \tfrac{1}{3}\lambda\overline{v}nm = \tfrac{1}{3}\lambda\overline{v}\rho, \tag{12.52}$$

where ρ is the density of the gas. Alternatively, the kinematic viscosity, $\nu = \eta/\rho = \frac{1}{3}\lambda\overline{v}$.

We can now work out how the coefficient of viscosity is expected to change with pressure and temperature. Substituting $\lambda = (n\sigma)^{-1}$ into (12.52), we find

$$\eta = \tfrac{1}{3}\lambda\overline{v}nm = \tfrac{1}{3}\frac{m\overline{v}}{\sigma}. \tag{12.53}$$

Maxwell was astonished to find that the coefficient of viscosity is *independent of the pressure*, since there is no dependence upon n in (12.53). The reason is that, although there are fewer molecules per unit volume as n decreases, the mean free path increases as n^{-1}, enabling a larger momentum increment to be transported to the plane at x_0 per mean free path.

Equally remarkably, as the temperature of the gas increases, $\overline{v}$ increases as $T^{1/2}$. Therefore, the viscosity of a gas should increase with temperature, unlike the behaviour of liquids. This somewhat counter-intuitive result was the subject of a brilliant set of experiments which Maxwell carried out with his wife from 1863 to 1865 (Fig. 12.5).

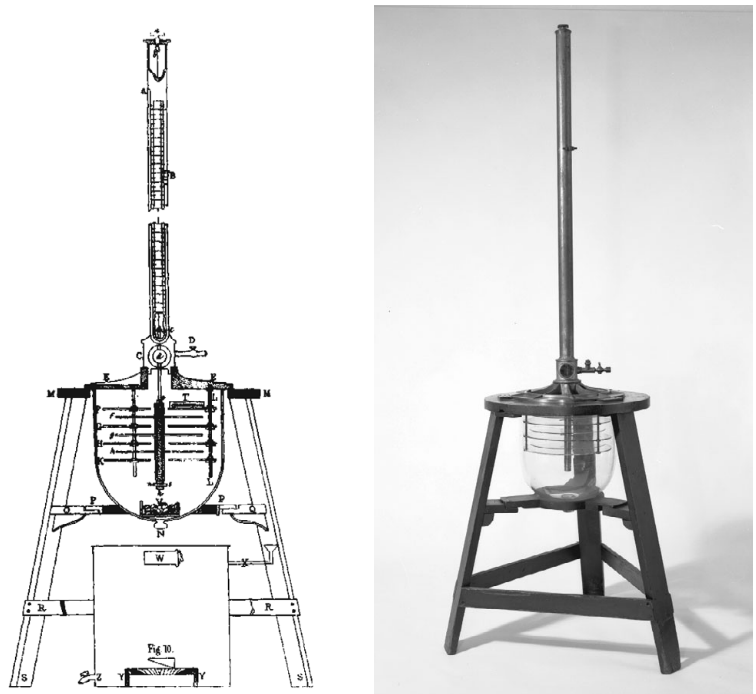


Fig. 12.5 Maxwell's apparatus for measuring the viscosity of gases.[11] (*a*) The gas fills the chamber and the glass discs oscillate as a torsion pendulum. The viscosity of the gas is found by measuring the rate of decay of the oscillations of the torsion balance. The oscillations of the torsion balance were measured by reflecting light from the mirror attached to the suspension. The pressure and temperature of the gas could be varied. The oscillations were started magnetically since the volume of the chamber had to be perfectly sealed. (*b*) A photograph of the apparatus, taken in the 1950s.

He fully confirmed the prediction that the viscosity of gases is independent of the pressure and fully expected to discover the $T^{1/2}$ law, but in fact he found a stronger dependence, $\eta \propto T$.

In his great paper of 1867,[11] he interpreted this discrepancy as indicating that there must be a weak repulsive force between the molecules which varied with their spacing r as r^{-5}, foreshadowing the concept of van der Waals forces between molecules of 1873.[12] This was a profound discovery since it meant that there was no longer any need to consider the molecules to be 'elastic spheres of definite radius'. The repulsive force, proportional to r^{-5}, meant that encounters between molecules would take the form of deflections through different angles, depending upon the impact parameter (see Section A4.3). Maxwell showed that it was much more appropriate to think in terms of a *relaxation time*, roughly the time it would take the molecule to be deflected through 90° as a result of random encounters with other molecules.

According to Maxwell's analysis, the molecules could be replaced by the concept of centres of repulsion, or, in his words, 'mere points, or pure centres of force endowed with inertia' – it was no longer necessary to make any special assumption about molecules as hard, elastic spheres. But the result is much deeper than this. Maxwell had replaced the mechanical collisions between molecules by an interaction between the fields of force of the molecules. This was part of the Maxwellian revolution in which mechanical forces were replaced by interactions between fields, as Maxwell had affirmed forcefully in his paper on electromagnetism of 1865 (see Section 6.3).

12.6 Pedagogical Digression (1): A Numerical Approach to the Boltzmann and Maxwell Distributions

My colleague Paul Alexander and I gave parallel courses to first-year students in Cambridge for several years and we were not happy about the various fudges we adopted to make the Boltzmann distribution plausible. We therefore developed a number of numerical simulations to illustrate what is going, as a trailer for the results derived using the method of undetermined multipliers in later years of the course (see Section 15.2). The value of these simulations is that they bring out vividly important features of the content of statistical physics, which are usually not mentioned in the standard approach.

12.6.1 The Boltzmann Distribution

The first set of simulations involved establishing the properties of the Boltzmann distribution. Consider first a rectangular array of 10×10 boxes, each box representing a molecule of the gas and insert five energy quanta into each box. Thus, the simulation starts with 100 molecules, all with five units of energy (Fig. 12.6(*a*)). We now adopt the following rules of procedure for the evolution of the energies of the particles:

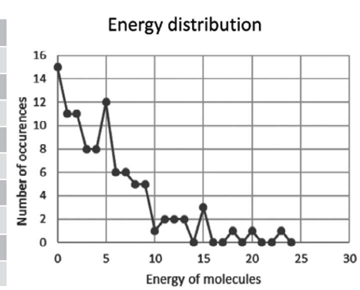
(1) Choose cell i of the array at random and move one quantum to another randomly chosen cell j.

5	5	5	5	5	5	5	5	5	5
5	5	5	5	5	5	5	5	5	5
5	5	5	5	5	5	5	5	5	5
5	5	5	5	5	5	5	5	5	5
5	5	5	5	5	5	5	5	5	5
5	5	5	5	5	5	5	5	5	5
5	5	5	5	5	5	5	5	5	5
5	5	5	5	5	5	5	5	5	5
5	5	5	5	5	5	5	5	5	5
5	5	5	5	5	5	5	5	5	5

(*a*)

0	10	2	4	15	11	2	13	2	1
5	8	2	0	0	12	15	18	2	6
8	7	2	6	12	6	0	1	9	0
0	4	3	8	5	1	0	4	0	5
7	4	1	1	2	9	9	2	20	9
9	3	8	5	3	5	23	1	8	0
3	1	5	6	6	2	7	0	5	5
3	3	5	3	4	7	5	0	1	1
0	7	2	0	4	13	0	5	4	2
5	15	3	1	0	1	6	11	4	7

(*b*)



(*c*)

Fig. 12.6 (*a*) The initial distribution showing 100 boxes with five energy quanta per box. (*b*) The distribution of energies among the boxes after 1000 random exchanges of quanta. (*c*) The energy distribution after 1000 random exchanges of quanta.

(2) This process transfers one unit of energy randomly to j from i, unless there are no quanta in cell i, in which case there is no transfer.
(3) Repeat many times randomly until the system has reached a state of statistical equilibrium.

The result of 1000 random exchanges is shown in Fig. 12.6(b), as well as a histogram of the energy distribution Fig. 12.6(c).[13] It can be seen that, starting with a distribution in which all the molecules have five units of energy, the distribution spreads out into a roughly exponential distribution with significant fluctuations. Once the system has settled down to a stationary distribution, the broad range of energies is maintained. Notice that, through purely random processes, some of the particles have increased in energy to far beyond their initial energies, in this example one of the particles attaining 23 energy units having started with five.

The simulation is repeated, but now with very much greater statistics. There are now 10 000 cells, or molecules, each with 20 quanta. Figure 12.7 shows how the system evolves after 0, 300 000, 1 800 000 and 4 000 000 random exchanges. The same evolution is observed as before, but the larger statistics make a number of key points more apparent.

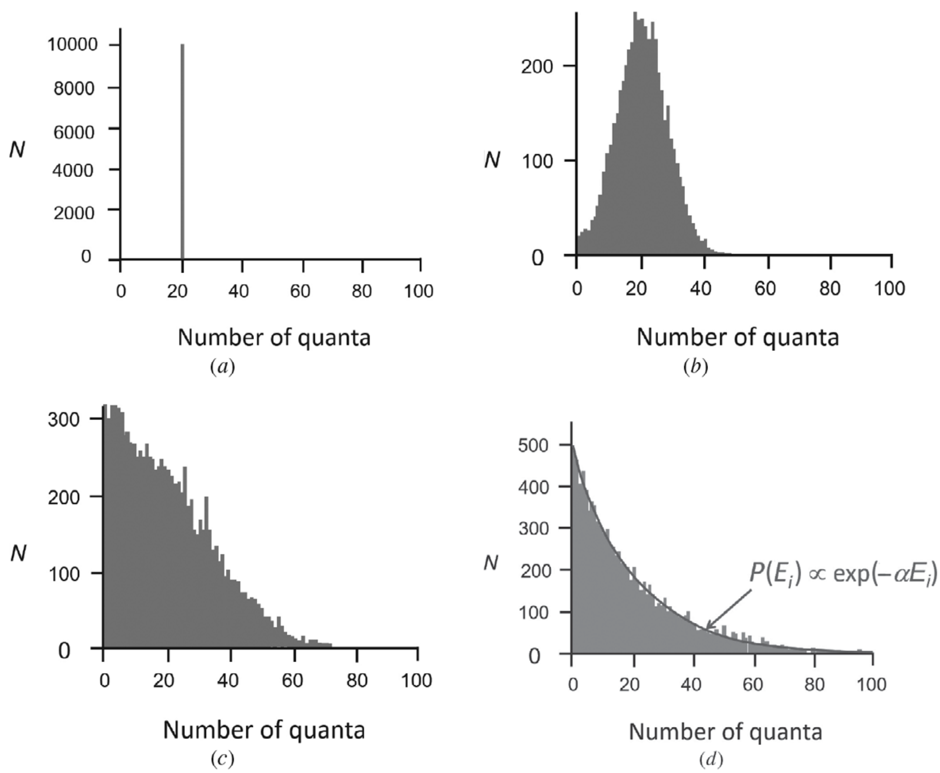


Fig. 12.7 (a) The initial distribution of 10 000 boxes with 20 energy quanta per box. (b) The distribution of energies among the boxes after 300 000 random exchanges of quanta. (c) The energy distribution after 1 800 000 random exchanges of quanta. (d) The energy distribution after 4 000 000 random exchanges of quanta.

After 300 000 exchanges, the distribution of energies takes up a more-or-less bell-shaped Gaussian shape, as expected since the energies are randomly redistributed about a mean value of 20 units. However, when the distribution hits the origin, the particle energies cannot become negative and so the energy distribution begins to 'climb up' the ordinate axis (Fig. 12.7(*c*)) and eventually settles down to an exponential distribution of energies, the Boltzmann distribution (Fig. 12.7(*d*)).

Notice again that some of the particles have attained very much greater energies than their original values. This is of great importance in many physical situations – there must be particles with energies very much greater than the mean and so are able to initiate processes at much lower temperatures than would be expected by simply adopted the mean kinetic energy of the molecules. An important example of this phenomenon for life on Earth is the fact that the nuclear reactions which power the Sun take place at a temperature about 20 times lower than would be expected if we simply set $E = \frac{3}{2}k_B T$.

When the simulation is run continuously, the random nature of the particle energies in the equilibrium state is clearly observed. The distribution constantly fluctuates about the mean exponential relation and that is a key part of the nature of statistical equilibrium. Also, if a particular molecule is tagged, it can be seen constantly changing its energy randomly, sometimes being much greater than the mean and sometimes much smaller.

12.6.2 Particle Collisions and the Equipartition of Energy

A second interesting example is to carry out numerically Maxwell's demonstration that a set of molecules comes to equipartition of energy as a result of molecular collisions.[11] First, we carry out an elementary calculation of a random two-particle collision in two-dimensions. The momenta of the particles before the collision are $\boldsymbol{p}_1 = (p_{1x}, p_{1y})$ and $\boldsymbol{p}_2 = (p_{2x}, p_{2y})$. The standard procedure is adopted of transforming these momenta into the centre of momentum frame of reference of the collision, rotating the transformed momentum vectors in that frame through some angle θ to find the new momenta $\boldsymbol{p}_1' = (p_{1x}', p_{1y}')$ and $\boldsymbol{p}_2' = (p_{2x}', p_{2y}')$ and then transforming back into the laboratory frame to find $\boldsymbol{p}_1'' = (p_{1x}'', p_{1y}'')$ and $\boldsymbol{p}_2'' = (p_{2x}'', p_{2y}'')$. The result of that calculation is that for particle 1:

$$p_{1x}'' = \tfrac{1}{2}[(p_{1x} - p_{2x})\cos\theta + (p_{1y} - p_{2y})\sin\theta + (p_{1x} + p_{2x})],$$
$$p_{1y}'' = \tfrac{1}{2}[-(p_{1x} - p_{2x})\sin\theta + (p_{1y} - p_{2y})\cos\theta + (p_{1y} + p_{2y})].$$

For particle 2, the result is

$$p_{2x}'' = \tfrac{1}{2}[(p_{2x} - p_{1x})\cos\theta + (p_{2y} - p_{1y})\sin\theta + (p_{1x} + p_{2x})],$$
$$p_{2y}'' = \tfrac{1}{2}[-(p_{2x} - p_{1x})\sin\theta + (p_{2y} - p_{1y})\cos\theta + (p_{1y} + p_{2y})].$$

The randomness is only associated with the choice of the angle θ in the centre of momentum frame of reference. We set up a two-dimensional array of boxes, each box containing the values of p_{xi} and p_{yi}, as illustrated in Fig. 12.8. The computational procedure is then as follows:

(1) Choose two molecules of the array, i and j, at random.
(2) Allow them to collide, choosing the angle θ at random.

-12, 12	-11, 7	-10, 2	-9, -3	-8, -8
-7, 11	-6, 6	-5, 1	-4, -4	-3, -9
-2, 10	-1, 5	0, 0	1, -5	2, -10
3, 9	4, 4	5, -1	6, -6	7, -11
8, 8	9,3	10, -2	11, -7	12, -12

Fig. 12.8 Illustrating an array of boxes, or molecules, showing their x and y momenta, $[p_{xi}, p_{yi}]$.

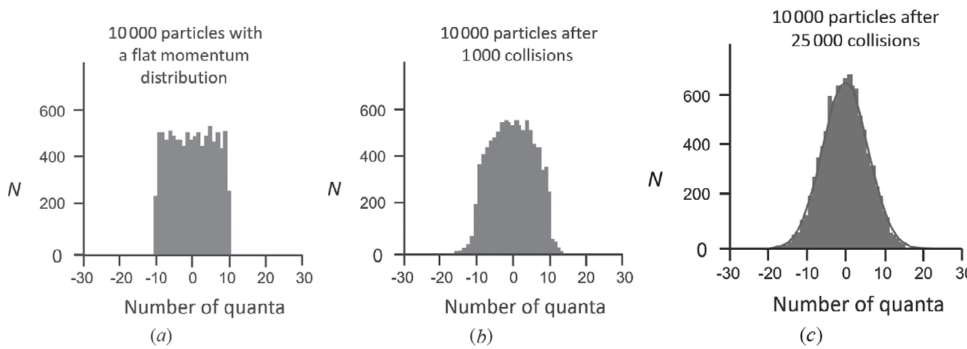


Fig. 12.9 Illustrating the evolution of the momentum distribution of particles after different numbers of collisions: (*a*) 10 000 particles with an initial flat total momentum momentum. (*b*) The total momentum distribution for 10 000 particles after 1000 collisions. (*c*) 10 000 particles after 25 000 collisions.

(3) Evaluate $[p''_{x1}, p''_{y1}]$ and $[p''_{x2}, p''_{y2}]$ for the molecules i and j and replace the old values of $[p_{x1}, p_{y1}]$ and $[p_{x2}, p_{y2}]$ for the molecules i and j by their new values $[p''_{x1}, p''_{y1}]$ and $[p''_{x2}, p''_{y2}]$.
(4) Choose a new pair of cells at random and repeat this process.

The results of such a procedure for an array of 10 000 molecules starting with a flat momentum distribution is shown in Fig. 12.9. The distribution relaxes very rapidly to a Maxwellian distribution in two dimensions. Again, the key element in this simulation is that the choice of the angle θ is made at random in the centre of momentum frame of reference. Notice again that, once the simulation reaches statistical equilibrium, it fluctuates about the mean Gaussian distribution.

Let us finally repeat the calculation with a mixture of molecules of masses m and $4m$. The evolution of the momentum distributions, both starting with the same flat momentum distribution, is illustrated in Fig. 12.10. Once again, the randomness is only associated with choice of the angle θ of the molecular collisions in the centre of momentum frame of reference. The collisions are constantly transferring momentum between molecules and it does not take many interactions to reach stationary momentum distributions for both species. Again, both distributions fluctuate about the Gaussian distributions shown in Fig. 12.10(c). Furthermore, the standard deviation of the momentum distribution of the $4m$ species w is half that of the lighter species, $2w$. Therefore, the average kinetic energies of the two sets of particles, $\overline{p_i^2}/2m_i$, are the same for both species. This is the equipartition theorem according to which, in thermal equilibrium, the particles end up with the same mean kinetic energy, mediated by particle collisions.

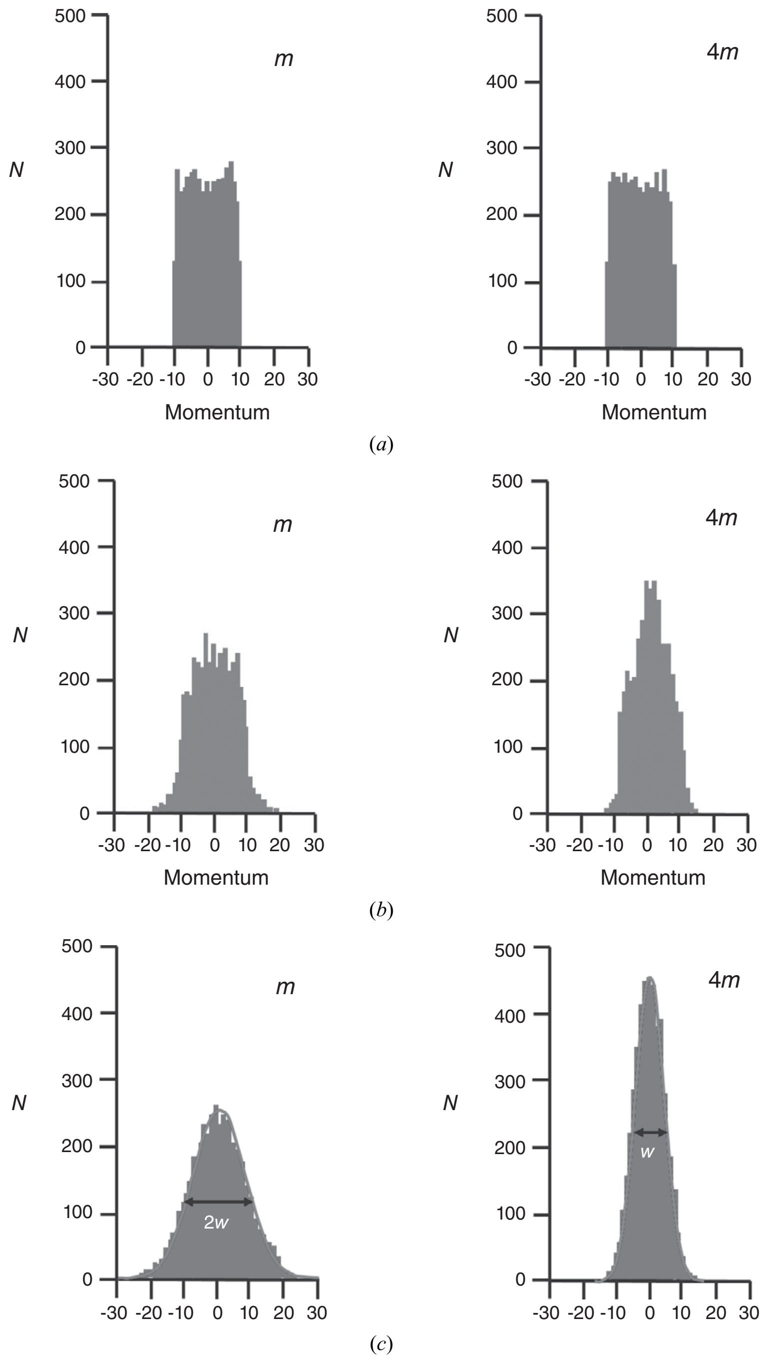


Fig. 12.10 Illustrating the evolution of the momentum distribution of particles with masses m and $4m$ due to random collisions between the two species. There are 5000 particles with mass m and 5000 with mass $4m$. (*a*) 5000 particles with mass m and 5000 with mass $4m$ both with the same flat momentum distributions. (*b*) The total momentum distribution for the particles after 10 000 random collisions. (*c*) The total momentum distribution for the particles after 50 000 random collisions.

12.7 The Statistical Nature of the Second Law of Thermodynamics

In 1867, Maxwell first presented his famous argument in which he demonstrated how, as a consequence of the kinetic theory of gases, it is possible to transfer heat from a colder to a hotter body without doing any work. He considered a vessel divided into two halves, A and B, the gas in A being hotter than that in B. Maxwell imagined a small hole drilled in the partition between A and B and a 'finite being' who watched the molecules as they approach the hole. The finite being has a shutter which can close the hole and adopts the strategy of allowing only fast molecules to pass from B to A and slow molecules from A into B. By this means, the hot molecules in the tail of the Maxwell distribution of the cool gas in B heat the hot gas in A and the cold molecules in the low energy tail of the distribution of hot gas in A cool the cool gas in B. The finite being thus enables the system to violate the second law of thermodynamics. Thomson referred to the finite being as 'Maxwell's demon', a name to which Maxwell objected, asking his old friend and Professor of Natural Philosophy at Edinburgh University, Peter Guthrie Tait, to

> call him no more a demon but a valve.[14]

Maxwell's remark illuminates a key point about the statistical nature of the second law of thermodynamics. Quite independent of finite beings or demons, there is a small but finite probability that, from time to time, exactly what Maxwell describes does indeed occur spontaneously. Whenever one fast molecule moves from B to A, heat is transferred from the colder to the hotter body without the influence of any external agency. Now, it is overwhelmingly more likely that hot molecules move from A to B and in this process heat flows from the hotter to the colder body with the consequence that the entropy of the combined system increases. There is, however, no question but that, according to the kinetic theory of gases, there is a very small but finite probability that the reverse happens and entropy decreases in this natural process. Maxwell was quite clear about the significance of his argument. He remarked to Tait that his argument was designed

> to show that the second law of thermodynamics has only a statistical certainty.[14]

This is a brilliant and compelling argument, but it depends upon the validity of the kinetic theory of gases.

The essentially statistical nature of the second law was emphasised by another argument of Maxwell's. In the late 1860s, both Clausius and Boltzmann attempted to derive the second law of thermodynamics from mechanics, an approach which was known as the dynamical interpretation of the second law. In this approach, the dynamics of individual molecules were followed in the hope that they would ultimately lead to an understanding of the origin of the second law. Maxwell refuted this approach as a matter of principle because of the simple but powerful argument that Newton's laws of motion, and indeed Maxwell's equations for the electromagnetic field, are completely time reversible. Consequently the irreversibility implicit in the second law cannot be explained by a dynamical theory. At the molecular level, the second law could only be understood as a consequence of the random behaviour of an immense number of molecules, as was demonstrated in Section 12.6. The controversy about the nature of the second law will reappear in Planck's attack on the problem of understanding the nature of the spectrum of black-body radiation.

Boltzmann originally belonged to the dynamical school of thought, but was fully aware of Maxwell's work. Among the most significant of his contributions during these years was a reworking of Maxwell's analysis of the equilibrium distribution of velocities in a gas, including the presence of a potential term $\phi(r)$ which describes the potential energy of the molecule in the field. Conservation of energy requires that

$$\tfrac{1}{2}mv_1^2 + \phi(\boldsymbol{r}_1) = \tfrac{1}{2}mv_2^2 + \phi(\boldsymbol{r}_2),$$

and the corresponding probability distribution has the form

$$f(v) \propto \exp\left[-\frac{\tfrac{1}{2}mv^2 + \phi(\boldsymbol{r})}{k_{\mathrm{B}}T}\right].$$

We recognise the appearance of the Boltzmann factor $\exp(-E/k_{\mathrm{B}}T)$ in this analysis.

Eventually, Boltzmann accepted Maxwell's doctrine concerning the statistical nature of the second law and set about working out the formal relationship between entropy and probability:

$$S = C \ln W, \tag{12.54}$$

where S is the entropy of a particular equilibrium state and W is the maximum probability of the system being found in that state. The value of the constant of proportionality C was not known. Boltzmann's analysis was of considerable mathematical complexity and this was one of the problems which stood in the way of a full appreciation the deep significance of what he had achieved.

12.8 Entropy and Probability

The formal relationship between entropy and probability involves a number of tricky issues and it would take us too far afield to look into the many subtleties of the subject. John Waldram's book *The Theory of Thermodynamics*[15] can be thoroughly recommended as an exposition of the fundamentals of statistical thermodynamics with no punches pulled – it is salutory to be reminded of the care needed in setting up the underlying infrastructure of statistical thermodynamics. The objective of the remainder of this chapter is to make plausible the formal relation between entropy and probability and to illuminate its close relationship with *information theory*.

The law of the increase of entropy tells us that systems evolve in such a way as to become more uniform. To put it another way, the molecules which make up the system become more randomised – there is less organised structure. We have already met examples in Chapter 11. When bodies at different temperatures are brought together, heat is exchanged so that they come to the same temperature – temperature inequalities are evened out. In a Joule expansion, the gas expands to fill a greater volume and thus produces uniformity throughout a greater volume of space. In both cases, there is an increase in entropy.

Maxwell's arguments strongly suggest that the entropy increase is a statistical phenomenon, although he did not attempt to quantify the nature of the relation between the second law and statistics. The advance made by Boltzmann was to describe the degree of disorder in systems in terms of the probability that they could arise by chance and to relate this probability to the entropy of the system.

Suppose we have two systems in different equilibrium states and p_1 and p_2 are the probabilities that they could arise independently by chance. The probability that they both occur is the product of the two probabilities $p = p_1 p_2$. If we are to associate entropies S_1 and S_2 with each system, entropies are additive and consequently the total entropy is

$$S = S_1 + S_2. \tag{12.55}$$

Thus, if entropy and probability are related, there must be a logarithmic relation of the form $S = C \ln p$ between them, where C is some constant.

Let us apply this statistical definition of entropy to the Joule expansion of a perfect gas. Consider the Joule expansion of 1 mole of gas from volume V to $2V$, noting that the system is in equilibrium states at the beginning and end of the expansion. What is the probability that, if the molecules were allowed to move freely throughout the volume $2V$, they would all occupy only a volume V? For each molecule, the probability of occupying the volume V is one-half. If there were two molecules, the probability would be $(\frac{1}{2})^2$, if three molecules $(\frac{1}{2})^3$, if four molecules $(\frac{1}{2})^4$ ... and, for one mole of gas with $N_0 \approx 6 \times 10^{23}$ molecules, Avogadro's number, the probability is $(\frac{1}{2})^{N_0}$ – this last probability is very small indeed. Notice that this number represents the relative probability of the two equilibrium states since it is assumed that in both $V/2$ and V the distributions of particles are uniform.

Let us adopt the definition $S = C \ln p$ to find the entropy change associated with $p_2/p_1 = 2^{N_0}$. The entropy change is then

$$S_2 - S_1 = \Delta S = C \ln(p_2/p_1) = C \ln 2^{N_0} = C N_0 \ln 2. \tag{12.56}$$

We have, however, already worked out the entropy change for a Joule expansion according to classical thermodynamics (11.53):

$$\Delta S = R \ln V_2/V_1 = R \ln 2.$$

It immediately follows that

$$R = CN \qquad \text{or} \qquad C = R/N = k_B.$$

We have recovered Boltzmann's fundamental relation between entropy and probability, with the bonus of having determined the constant relating S and $\ln p$, Boltzmann's constant k_B:

$$\Delta S = S_2 - S_1 = k_B \ln(p_2/p_1). \tag{12.57}$$

The sign has been chosen so that the more likely event corresponds to the greater entropy.

Let us perform another calculation which illustrates how the statistical approach works. Consider two equal volumes, each $V/2$, containing gas at different pressures p_1 and p_2 separated by a partition. On one side of the partition, there are r moles of gas and on the other $1-r$ moles. According to classical thermodynamics, the entropy change when 1 mole of gas changes its equilibrium state from V_0, T_0 to V, T is given by (11.52)

$$S_1(T,V) = C_V \ln\left(\frac{T}{T_0}\right) + R \ln\left(\frac{V}{V_0}\right) + S_1(T_0, V_0),$$

the subscript 1 referring to 1 mole of gas. Now suppose we add together m equal volumes so that there are in total m moles of gas. Since entropies are additive,

$$S_m(T,V) = mC_V \ln\left(\frac{T}{T_0}\right) + mR \ln\left(\frac{V}{V_0}\right) + mS_1(T_0, V_0),$$

if we keep all the other variables the same. The only catch is that V and V_0 still refer to 1 mole of gas. The volume of the m moles is $mV = V_m$. Therefore the expression for the total entropy can be written

$$S_m(T,V_m) = mC_V \ln\left(\frac{T}{T_0}\right) + mR \ln\left(\frac{V_m}{mV_0}\right) + mS_1(T_0, V_0).$$

We can now apply this result to the two volumes $V_m = V/2$ which contain r and $1-r$ moles, that is,

$$\begin{cases} S_1 = rC_V \ln\left(\dfrac{T}{T_0}\right) + rR \ln\left(\dfrac{V}{2rV_0}\right) + rS_1(T_0, V_0), \\ S_2 = (1-r)C_V \ln\left(\dfrac{T}{T_0}\right) + (1-r)R \ln\left[\dfrac{V}{2(1-r)V_0}\right] + (1-r)S_1(T_0, V_0). \end{cases} \tag{12.58}$$

Adding these together, the total entropy is

$$\begin{aligned} S = S_1 + S_2 &= C_V \ln\left(\frac{T}{T_0}\right) + R \ln\left(\frac{V}{2V_0}\right) + S_1(T_0, V_0) + (1-r)R \ln\left(\frac{1}{1-r}\right) + rR \ln\left(\frac{1}{r}\right) \\ &= C_V \ln\left(\frac{T}{T_0}\right) + R \ln\left(\frac{V}{2V_0}\right) + S_1(T_0, V_0) + \Delta S(r), \end{aligned} \tag{12.59}$$

where we have absorbed all the terms in r into $\Delta S(r)$, that is,

$$\Delta S(r) = -R[(1-r)\ln(1-r) + r\ln r]. \tag{12.60}$$

The term $\Delta S(r)$ describes how the entropy depends upon the quantity of gas on either side of the partition.

Now let us look at the problem from the statistical perspective. We can use simple statistical procedures to work out the number of different ways in which N identical objects can be distributed between two boxes. The necessary tools are developed in Section 14.2. Referring to that section, the number of ways in which m objects can be arranged in one box and $N-m$ in the other is

$$g(N,m) = \frac{N!}{(N-m)!\,m!}.$$

Let us refer these numbers to the case of equal numbers of objects in each box:

$$g(N,x) = \frac{N!}{[(N/2)-x]!\,[(N/2)+x]!},$$

where $x = m - (N/2)$. Now this is not a probability, but the exact number of microscopic ways in which we can end up with $[(N/2) - x]$ objects in one box and $[(N/2) + x]$ in the other. It is, however, a probability if we divide by the total number of possible ways of distributing the objects between the boxes, and this is a constant. The part of the probability which depends upon x is the above expression for $g(N, x)$. Therefore, adopting Boltzmann's expression for the entropy,

$$S = k_B \ln g(N, x) = k_B\{\ln N! - \ln[(N/2) - x]! - \ln[(N/2) + x]!\}. \tag{12.61}$$

Now approximate the logarithms using Stirling's approximation in the form

$$\ln M! \approx M \ln M - M. \tag{12.62}$$

After a bit of simple manipulation of (12.61), we find

$$S/k_B = N \ln 2 - \left[\left(\frac{N}{2} - x\right)\ln\left(1 - \frac{2x}{N}\right) + \left(\frac{N}{2} + x\right)\ln\left(1 + \frac{2x}{N}\right)\right] \tag{12.63}$$

$$= -N\left[\left(\tfrac{1}{2} - \frac{x}{N}\right)\ln\left(\tfrac{1}{2} - \frac{x}{N}\right) + \left(\tfrac{1}{2} + \frac{x}{N}\right)\ln\left(\tfrac{1}{2} + \frac{x}{N}\right)\right]. \tag{12.64}$$

Now N is the number of molecules in 1 mole, Avogardo's number N_0. Therefore,

$$r = \tfrac{1}{2} - \frac{x}{N_0} \quad \text{and} \quad 1 - r = \tfrac{1}{2} + \frac{x}{N_0}.$$

It follows that

$$S = -N_0 k_B [r \ln r + (1 - r)\ln(1 - r)]. \tag{12.65}$$

This is exactly the result we have obtained from classical thermodynamics (12.60) with the correct constant in front of the logarithms, $k_B = R/N_0$. Thus, the definition $S = k_B \ln g(N, x)$ can account rather beautifully for the classical result through statistical arguments.

Let us now work out the probability of small deviations from equal numbers of molecules in each volume. To do this, expand the function $g(N_0, x)$ for small values of x. From (12.63),

$$\frac{S}{k_B} = \ln g(N_0, x) = N_0 \ln 2 - \frac{N_0}{2}\left[\left(1 - \frac{2x}{N_0}\right)\ln\left(1 - \frac{2x}{N_0}\right) + \left(1 + \frac{2x}{N_0}\right)\ln\left(1 + \frac{2x}{N_0}\right)\right].$$

Expanding the logarithms to order x^2,

$$\ln(1 + x) = x - \tfrac{1}{2}x^2,$$

the expression for $\ln g(N_0, x)$ reduces to

$$\ln g(N_0, x) = N_0 \ln 2 - \frac{N_0}{2}\left[\left(1 - \frac{2x}{N_0}\right)\left(-\frac{2x}{N_0} - \frac{1}{2}\frac{4x^2}{N_0^2}\right) + \left(1 + \frac{2x}{N_0}\right)\left(\frac{2x}{N_0} - \frac{1}{2}\frac{4x^2}{N_0^2}\right)\right]$$

$$= N_0 \ln 2 - \frac{2x^2}{N_0} + \cdots$$

$$g(N_0, x) = 2^{N_0} \exp\left(-\frac{2x^2}{N_0}\right). \tag{12.66}$$

Now the total number of ways in which the objects can be divided between the boxes is 2^{N_0} and hence, in this approximation, the probability distribution is

$$p(N_0, x) = \exp\left(-\frac{2x^2}{N_0}\right), \tag{12.67}$$

that is, a Gaussian distribution with mean value 0 with probability 1! We note that the standard deviation of the distribution about the value $x = 0$ is $(N_0/2)^{1/2}$, that is, to order of magnitude, $N_0^{1/2}$. Thus, the likely deviation about the value $x = 0$ corresponds to very tiny fluctuations indeed. Since $N_0 \sim 10^{24}$, $N_0^{1/2} \sim 10^{12}$ or, in terms of the fractional fluctuations, $N_0^{1/2}/N_0 \sim 10^{-12}$. Our analysis of Section 17.2.1 shows that this is no more than the statistical fluctuation about the mean value.

This example illustrates how statistical mechanics works. We deal with huge ensembles of molecules, $N_0 \sim 10^{24}$, and hence, although it is true that there are statistical deviations from the mean behaviour, in practice, they are very, very tiny. Indeed, although it is possible for the statistical entropy to decrease spontaneously, the likelihood of this happening is negligibly small because N_0 is so large.

Let us show briefly how entropies depend upon the volume of the system, not only in real three-dimensional space, but also in *velocity* or *phase space* as well. The entropy change of the perfect gas in passing from coordinates V_0, T_0 to V, T is

$$\Delta S(T, V) = C_V \ln\left(\frac{T}{T_0}\right) + R \ln\left(\frac{V}{V_0}\right). \tag{12.68}$$

Now express the first term on the right-hand side in terms of molecular velocities rather than temperatures. According to kinetic theory, $T \propto \overline{v^2}$ and hence, since $C_V = \frac{3}{2}R$,

$$\begin{aligned}\Delta S(T, V) &= \tfrac{3}{2}R \ln\left(\frac{\overline{v^2}}{\overline{v_0^2}}\right) + R \ln\left(\frac{V}{V_0}\right)\\ &= R\left[\ln\left(\frac{\overline{v^2}}{\overline{v_0^2}}\right)^{3/2} + \ln\left(\frac{V}{V_0}\right)\right].\end{aligned} \tag{12.69}$$

We can interpret this formula as indicating that when we change T and V, there are two contributions to the entropy change. The volume of real space changes in the ratio V/V_0. In addition, the first term in square brackets indicates that the molecules occupy a larger volume of velocity or phase space by the factor $(\overline{v^2}/\overline{v_0^2})^{3/2} \sim (v^3/v_0^3)$. Thus, we can interpret the formula for the entropy increase in terms of the increase in the accessible volumes of both real and velocity space.

12.9 Entropy and the Density of States

Let us take the story one step further. In the reasoning which led to (12.61), rather than work in terms of probabilities, we simply used the total number of different ways in which

m objects can be distributed in one box and $N - m$ in the other. This leads naturally to the theory of thermodynamics as formulated by John Waldram, which has strongly influenced the approach taken in this section. We begin with one of his examples.

Just as in the route pioneered by Boltzmann, the theory begins by assuming the existence of discrete energy levels, as in any quantum system, and it is convenient to consider this to be a set of identical harmonic oscillators, which have the pleasant feature of possessing equally spaced energy levels. For present purposes, we can neglect the zero-point energy $\frac{1}{2}\hbar\omega$ of each oscillator. Suppose the system consists of 100 oscillators and we can add quanta to them, the energy of each quantum being equal to the spacing of the energy levels. We add Q quanta to the system of oscillators and ask, 'In how many different ways W can we distribute Q quanta among the $N = 100$ oscillators?' If there were no quanta, $Q = 0$, there is only one way of distributing them and so $W = 1$. If $Q = 1$, we can place one quantum in any of the 100 ground states and so $Q = 100$. If $Q = 2$, we can place two quanta in different ground states in $(100 \times 99)/2 = 4950$ ways, plus the 100 ways in which the two quanta can be allocated to the same oscillator. Hence, in total, $W = 5050$. Notice that, in carrying out this sum, we avoid duplications – it is the total number of different ways which is enumerated.

The extension of this procedure to distributing Q quanta among N oscillators is carried out in Section 15.2, with the result that the total number of different configurations is

$$W = \frac{(Q + N - 1)!}{Q!\,(N - 1)!} \tag{12.70}$$

We can now create a table (Table 12.1) which shows how rapidly W increases with increasing Q for the case of 100 oscillators.

We notice that $W \sim 10^Q$ as Q becomes large. It is left as an exercise to the reader to show that, using Stirling's approximation (12.63), if $Q = 10^{23}$, $W \sim 10^{10^{23}}$. Thus, the number of ways of obtaining a total energy Q increases very rapidly indeed with increasing Q. We say that the *density of states* W increases very rapidly with energy.

Table 12.1 The number of different ways in which Q quanta can be distributed among N oscillators

Q	W
0	1
1	100
2	5050
3	171 700
4	4 421 275
5	92 846 775
10	4.5093338×10^{13}
20	2.8073848×10^{22}

In practice, we always deal with the numbers of states within a finite range of energies, say, from E to $E + \mathrm{d}E$. Then, the proper statement is that the density of states is $g(E)\,\mathrm{d}E$ where $g(E)$ is a continuous function with dimensions $[\text{energy}]^{-1}$.

To convert these ideas in a theory of statistical thermodynamics, we need to make some assumptions. Central to the subject is the *principle of equal equilibrium probabilities*. This is the statement that

> In statistical equilibrium, all quantum states accessible to the system are equally likely to be occupied.

Thus, if we share three quanta among 100 oscillators, there are 171 700 ways of sharing out this energy and, in equilibrium, the system is equally likely to be found in any one of them. There is no point in disguising the fact that there are a number of problems with this assumption – for example, does the system really explore all possible distributions in the age of the Universe? Likewise, we need to assume that there are equal probabilities for a transition between any pair of states to occur in either direction. Waldram refers to this as the *principle of jump-rate symmetry*, which fortunately follows from the equality of the transition probabilities between quantum states according to quantum mechanics. These principles work incredibly well and so provide their own justification.

This is the beginning of a journey which leads to the statistical definition of the law of increase of entropy and what we mean by temperature in statistical physics. Waldram gives a very pleasant example of how this works. Consider two large systems A and B, A containing $N_{\mathrm{A}} = 5000$ and B $N_{\mathrm{B}} = 10\,000$ oscillators. All the oscillators have the same energy levels and so we can work out the density of states for each of them. For A,

$$g_{\mathrm{A}} = \frac{(N_{\mathrm{A}} + Q_{\mathrm{A}} - 1)!}{Q_{\mathrm{A}}!\,(N_{\mathrm{A}} - 1)!} \approx \frac{(N_{\mathrm{A}} + Q_{\mathrm{A}})!}{Q_{\mathrm{A}}!\,N_{\mathrm{A}}!}, \tag{12.71}$$

where the last equality follows from the fact that the numbers are so large that the -1s are of no importance. We can now simplify this result using Stirling's approximation in the form $Q! \approx Q^{Q}$. Therefore,

$$g_{\mathrm{A}} = \frac{(N_{\mathrm{A}} + Q_{\mathrm{A}})^{N_{\mathrm{A}}+Q_{\mathrm{A}}}}{Q_{\mathrm{A}}^{Q_{\mathrm{A}}}\,N_{\mathrm{A}}^{N_{\mathrm{A}}}}. \tag{12.72}$$

The densities of states for A and B are therefore

$$g_{\mathrm{A}} = \frac{(5000 + Q_{\mathrm{A}})^{5000+Q_{\mathrm{A}}}}{Q_{\mathrm{A}}^{Q_{\mathrm{A}}}\;5000^{5000}}, \qquad g_{\mathrm{B}} = \frac{(10\,000 + Q_{\mathrm{B}})^{10\,000+Q_{\mathrm{B}}}}{Q_{\mathrm{B}}^{Q_{\mathrm{B}}}\;10\,000^{10\,000}}. \tag{12.73}$$

If we now place the two systems in thermal contact, the quanta are free to explore all the states of the combined system (A + B). This means that the density of states available to the combined system is $g_{\mathrm{A}}g_{\mathrm{B}}$ since we can combine every state of A with every state of B. Thus, for the combined system,

$$g = g_{\mathrm{A}}g_{\mathrm{B}} = \frac{(5000 + Q_{\mathrm{A}})^{5000+Q_{\mathrm{A}}}}{Q_{\mathrm{A}}^{Q_{\mathrm{A}}}\;5000^{5000}} \times \frac{(10\,000 + Q_{\mathrm{B}})^{10\,000+Q_{\mathrm{B}}}}{Q_{\mathrm{B}}^{Q_{\mathrm{B}}}\;10\,000^{10\,000}}. \tag{12.74}$$

The total number of quanta $Q = Q_A + Q_B$ is fixed. It is left as an exercise to the reader to find the maximum value of g subject to this constraint. It is simplest to take logarithms of both sides, throw away the constants and maximise by finding $d(\ln g)/dQ_A$. The answer is

$$Q_A = \frac{Q}{3}, \qquad Q_B = \frac{2Q}{3}. \tag{12.75}$$

Thus, the energy is shared exactly proportional to the number of oscillators in A and B. To put it another way, the equilibrium distribution is found by maximising the density of states which can be occupied by the system.

Now suppose the system were not in its most probable state. For example, suppose that the energies were the other way round with

$$Q_A = \frac{2Q}{3}, \qquad Q_B = \frac{Q}{3}. \tag{12.76}$$

Inserting these values into (10.74), we find

$$\frac{g}{g_{\max}} = 10^{-738}. \tag{12.77}$$

How does the system evolve towards equilibrium? Because of the principle of jump-rate symmetry, when the systems are placed in thermal contact, there is twice the probability of a quantum jumping from A to B as from B to A, because there are twice as many states available in B as in A. Thus, the 'natural flow of heat' is from A to B. In statistical terms, heat flows in natural processes so that $g_A g_B$ increases, that is,

$$d(g_A g_B) \geq 0. \tag{12.78}$$

We can equally well find the maximum of the logarithm of $g_A g_B$ and so

$$d\,[\ln(g_A g_B)] = d(\ln g_A) + d(\ln g_B) \geq 0. \tag{12.79}$$

But, we recall that g_A and g_B are functions of energy, $g_A(E_A)$ and $g_B(E_B)$. Therefore, differentiate (10.79) with respect to energy:

$$\frac{\partial(\ln g_A)}{\partial E_A}\, dE_A + \frac{\partial(\ln g_B)}{\partial E_B}\, dE_B \geq 0. \tag{12.80}$$

But, when energy is exchanged between A and B, $dE_A = -dE_B$. Therefore,

$$\left[\frac{\partial(\ln g_A)}{\partial E_A} - \frac{\partial(\ln g_B)}{\partial E_B}\right] dE_A \geq 0. \tag{12.81}$$

This is a fundamental equation in statistical physics. The signs are important. If dE_A is positive, A gains energy and

$$\frac{\partial(\ln g_A)}{\partial E_A} > \frac{\partial(\ln g_B)}{\partial E_B}. \tag{12.82}$$

Contrariwise, if dE_A is negative, B gains energy and

$$\frac{\partial(\ln g_B)}{\partial E_B} > \frac{\partial(\ln g_A)}{\partial E_A}. \tag{12.83}$$

Thus, $\partial(\ln g)/\partial E$ is a measure of temperature. From the direction in which the heat flows take place, according to (12.81), the greater $\partial(\ln g)/\partial E$, the lower the temperature, while the smaller $\partial(\ln g)/\partial E$, the higher the temperature. This leads to the definition of *statistical temperature* T_s, namely

$$\frac{1}{k_B T_s} = \frac{\partial(\ln g)}{\partial E}. \tag{12.84}$$

In equilibrium, $d\ln(g_A g_B) = 0$ and so

$$\frac{\partial(\ln g_A)}{\partial E_A} = \frac{\partial(\ln g_B)}{\partial E_B}, \quad \text{that is,} \quad T_s(A) = T_s(B). \tag{12.85}$$

The definition (12.84) can be compared with the corresponding equation derived from classical thermodynamics (11.57):

$$\left(\frac{\partial S}{\partial U}\right)_V = \frac{1}{T}. \tag{12.86}$$

Evidently, we have derived a statistical definition of entropy:

$$S = k_B \ln g = k_B \ln g(E)\, dE, \tag{12.87}$$

where we recall that g is the maximum density of states, corresponding to the equilibrium state.

12.10 Pedagogical Digression (2): A Numerical Approach to the Law of Increase of Entropy

We can use the techniques described in Section 12.6 to illustrate how the results described in the last section come about. We begin with two small isolated systems A and B, each of them with all the molecules having the same, but different, energies – A has 100 molecules each with five quanta of energy and B 25 molecules each with two quanta of energy (Fig. 12.11(*a*)). As in the simulations in Fig. 12.6, since A and B are isolated, they come separately to stationary equilibrium distributions after many random exchanges. These each result in exponential Boltzmann distributions, as illustrated by the much large simulations shown in Fig. 12.12(*a*) – system A has 10 000 molecules with 20 quanta per cell and B 2500 molecules with five quanta per cell. Both tend separately to Boltzmann distributions which can be fitted by distributions of the form $p(E_i) \propto \exp(-\alpha E_i)$ and $p(E_i) \propto \exp(-\beta E_i)$ for the systems A and B respectively. It is clear from Fig. 12.12(*a*) that $\alpha \ll \beta$ and so, since α and β are both proportional to T^{-1}, the temperature of B is less than that of A.

Now we place A and B into thermal contact. In the statistical model, this means that quanta can now be exchanged randomly between the two grids (Fig. 12.11(*b*)), shown schematically by the 'thermal' link between the arrays. After a large number of random exchanges within and between the grids, there is a net transfer of energy elements from A to B but entirely on a random statistical basis, as indicated by the net numbers moving

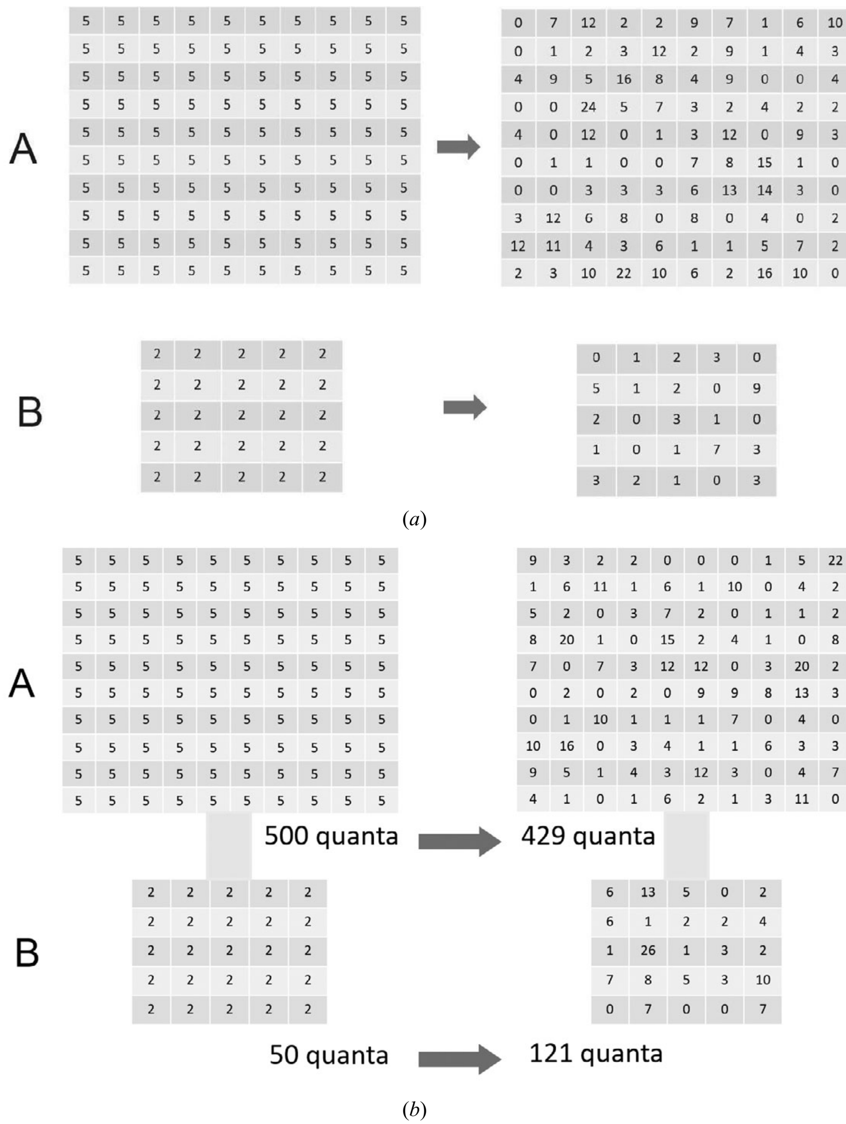
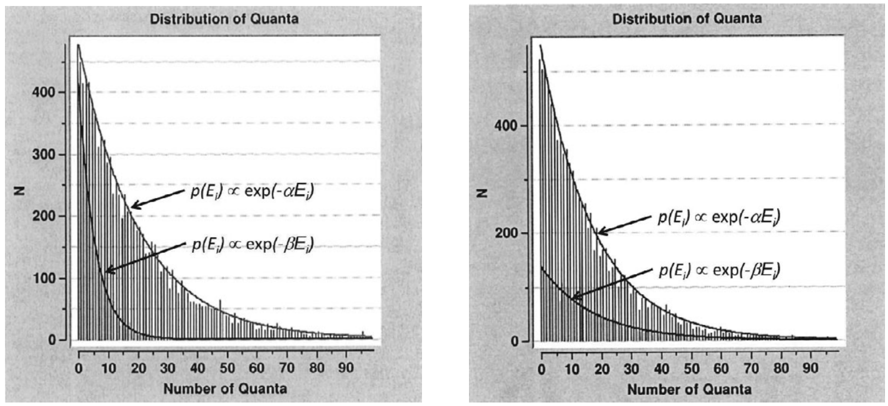


Fig. 12.11 Illustrating the equilibrium energy distribution of molecules with different numbers of cells and different energies. (*a*) System A begins 100 molecules with five quanta per cell and system B 25 molecules with two quanta per cell. In the simulation in (*a*), A and B are thermally isolated, and when the random reassignment of energy quanta is made within each array, typical examples of the energy distributions are in the arrays on the right of the diagram. (*b*) The energy distributions if A and B are in thermal contact, shown symbolically by the junction joining the arrays. Now the reassignment of energies takes place both within and between the arrays, resulting in the long term in a net transfer of quanta from A and B, entirely by random processes.

between the arrays in Fig. 12.11(*b*). In that example, there is a net transfer of 71 energy quanta from A to B but this figure fluctuates about the mean number of transfers of 60, which would result in the same average energy per molecule of 4.4 energy units. These fluctuations are an inescapable consequence of the statistical nature of the equilibrium



(*a*) (*b*)

Fig. 12.12 Illustrating the equilibrium energy distribution of molecules with different numbers of cells and different energies. System A has 10 000 molecules with 20 quanta per cell and B 2500 molecules with five quanta per cell. (*a*) The separate energy distributions if A and B are thermally isolated. (*b*) The energy distributions if A and B are in thermal contact. The distributions have been fitted by exponential distributions.

states of both arrays. In the example shown in Fig. 12.11(*b*), the net transfer of 71 quanta is about 1σ greater than the mean figure of 60.

The distribution of molecular energies for the very much larger combined array is shown in Fig. 12.12(*b*), which shows that they take up the same exponential distributions with the same exponents $\alpha = \beta$. Thus, by purely random exchange of quanta between the two arrays, thermal equilibrium is established at the same temperature.

The simulations also make it apparent why the combined system is in a state of higher entropy as compared with that of two separate systems. Since $S = k_B \ln g$, where g is the density of states, by combining the arrays into a single larger array, there is a much greater density of states because there is a larger number of boxes to be explored than for either system separately. In other words, the equilibrium state of the combined, thermally interacting system (A + B) is a state of maximum entropy which is greater than the sum of the entropies of A and B separately.

12.11 Gibbs Entropy and Information

These concepts were gradually crystallised through the last two decades of the nineteenth century, largely thanks to the heroic efforts of Boltzmann and Gibbs. Gibbs, in particular, discovered the correct procedure for defining the probabilities p_i that the system be found in state i and how these probabilities should be combined to define the entropy of the system. What Gibbs discovered turned out to be one aspect of a very general theorem in probability theory. In 1948 Claude Shannon discovered a theorem which enabled the statistical approach to be formulated in a completely general way on the basis of

Table 12.2 Examples of different sums over probabilities

p_i					$-\sum p_i \ln p_i$	$-\sum \ln p_i$	$-\sum p_i^2 \ln p_i$
0.2	0.2	0.2	0.2	0.2	1.609	0.322	0.32
0.1	0.3	0.2	0.3	0.1	1.504	8.622	0.33
0.5	0	0	0	0.5	0.693	1.386	0.35
0	0	1	0	0	0	0	0

information theory. *Shannon's theorem*[16] is one of the most fundamental theorems in the theory of probability and can be stated as follows:

> If p_i are a set of mutually exclusive probabilities, then the function
>
> $$f(p_1, p_2, p_3, \dots p_n) = -C \sum_1^n p_i \ln p_i \tag{12.88}$$
>
> is a unique function which, when maximised, gives the most likely distribution of p_i for a given set of constraints.

We will not prove this theorem, but give a simple numerical example which illustrates how it works. The important point is that the form $-\sum p_i \ln p_i$ is not at all arbitrary. Suppose there are five possible outcomes of an experiment and we assign probabilities p_i to these. In Table 12.2, the results of forming various sums of these probabilities are shown.

This simple demonstration shows that the maximum value of $-\sum p_i \ln p_i$ corresponds to the case in which all the p_is are equal. In the words of the theorem, the most likely distribution of probabilities in the absence of any constraints is that all five probabilities are the same. In the case in which the system is completely ordered, as in the fourth example, $-\sum p_i \ln p_i$ is zero. It can be seen from the table that, as the probabilities become more and more uniform, $-\sum p_i \ln p_i$ tends towards a maximum value. The correspondence with the law of increase of entropy is striking and leads directly to the definition of the *Gibbs entropy* as

$$S = -k_B \sum_i p_i \ln p_i, \tag{12.89}$$

and the maximum of this function is the most probable distribution, corresponding to the equilibrium state of the system. Notice that (12.89) provides a very powerful definition of the entropy of the system which can be used even if it has not attained an equilibrium state.

Notice how this definition quantifies the relation between entropy and disorder. The fourth example is completely ordered, has zero entropy and contains the maximum amount of information about the state of the system. The distribution in the first example is completely uniform, the entropy is a maximum and we obtain the minimum information about which state the system is actually in. Accordingly, in information theory, the information is defined to be

$$\text{Information} = k_B \sum_i p_i \ln p_i, \tag{12.90}$$

and is sometimes called the *negentropy*.

We can now complete the analysis by relating the Gibbs definition of entropy to the results obtained in the previous sections. According to the principle of equal equilibrium probabilities, the equilibrium state corresponds to that in which there is an equal probability of the system being found in any one of the states accessible to it, subject to a set of given constraints. In the language of the last section, each of the g states accessible to the system is equally probable and so

$$p_i = \frac{1}{g} = \frac{1}{g(E)\,\mathrm{d}E},$$

where the last equality indicates how we would carry out the same analysis if the density of states were described by a continuous function. Therefore, adopting the Gibbs definition of entropy,

$$S = -k_B \sum_{i=1}^{i=g} \frac{1}{g} \ln \frac{1}{g}.$$

Since the sum is over all g states,

$$S = k_B \ln g = k_B \ln(\text{density of states}), \tag{12.91}$$

exactly the definition of the equilibrium entropy derived in (12.87).

To convince ourselves of the equivalence of these procedures, we can use the example shown in Table 12.2 in which we place one quantum in any one of the five oscillators. We showed in Section 12.8 that the density of states then corresponds to $Q = 1$ and $N = 5$. Hence,

$$g = \frac{(Q + N - 1)!}{Q!\,(N - 1)!} = 5,$$

which we could have guessed without using factorials. Therefore, $\ln g = \ln 5 = 1.609$, which is exactly the same as the value of $-\sum_i p_i \ln p_i$ shown in the first line of Table 12.2.

Finally, let us consider the Joule expansion in the same light. In the case of a single volume, all the p_is are equal to unity for all N_0 particles and so $-\sum_i p_i \ln p_i = 0$. When there are two volumes, the probability is 0.5 that the molecule will be in one volume and 0.5 that it will be in the other. In equilibrium, these probabilities are the same for all molecules and so the Gibbs entropy is

$$S = -k_B \sum_{N_0} \left(\tfrac{1}{2} \ln \tfrac{1}{2} + \tfrac{1}{2} \ln \tfrac{1}{2}\right) = k_B N_0 \ln 2. \tag{12.92}$$

Thus, the Gibbs entropy is again identical to the classical definition of entropy which resulted in $\Delta S = R \ln 2 = k_B N_0 \ln 2$.

12.12 Concluding Remarks

With this introduction, the way is now clear for the development of the full statistical mechanical interpretation of classical thermodynamics. This development is traced in such standard texts as Kittel's *Thermal Physics*,[17] Mandl's *Statistical Physics*[18] and Waldram's *The Theory of Thermodynamics*.[15]

Boltzmann's great discovery was appreciated by relatively few physicists at the time. The stumbling blocks were the facts that the status of the kinetic theory of gases and the equipartition theorem was uncertain since it failed to account for the specific heat capacities of gases and it was not clear how to incorporate the internal vibrations of molecules, as exhibited by spectral lines. Towards the end of the nineteenth century, a reaction against atomic and molecular theories of the properties of matter gained currency in some quarters in continental Europe. According to this view, one should only deal with the bulk properties of systems, that is, with classical thermodynamics, and abolish atomic and molecular concepts as unnecessary. These were profoundly discouraging developments for Boltzmann. They may well have contributed to his suicide in 1906. It is a tragedy that he was driven to such an act at the very time when the correctness of his fundamental insights were appreciated by Einstein. The problem of the specific heat capacities was solved in Einstein's classic paper of 1906 and in the process opened up a completely new vision of the nature of elementary processes.

We will return to Boltzmann's procedures and the concepts developed in this chapter when we survey Planck's and Einstein's extraordinary achievements which resulted in the discovery of the concepts of quantisation and quanta.

Notes

1 Clausius, R. (1857). *Annalen der Physik*, **100**, 497. (English translation in S.G. Brush (ed.), *Kinetic Theory*, Vol. 1, p. 111. Oxford: Pergamon Press (1966).)

2 For economy of expression, I will refer to both atoms and molecules as 'molecules' in this chapter.

3 Clausius, R. (1858). *Annalen der Physik*, **105**, 239. (English translation in S.G. Brush (ed.), *Kinetic Theory*, Vol. 1, p. 135. Oxford: Pergamon Press (1966).)

4 Waterston, J.J. (1843). See article by S.G. Brush, *Dictionary of Scientific Biography*, Vol. 14, p. 184. New York: Charles Scribner's Sons. See also *The Collected Scientific Papers of John James Waterston*, ed. J.G. Haldane. Edinburgh: Oliver and Boyd.

5 Rayleigh, Lord (1892). *Philosophical Transactions of the Royal Society*, **183**, 1–78.

6 Rayleigh, Lord (1892). *op. cit.*, 2.

7 Rayleigh, Lord (1892). *op. cit.*, 3.

8 Maxwell, J.C. (1860). *Philosophical Magazine, Series 4*, **19**, 19 and **20**, 21. See also *The Scientific Papers of James Clerk Maxwell*, ed. W.D. Niven, p. 377. Cambridge: Cambridge University Press (1890).

9 Everitt, C.W.F. (1970). *Dictionary of Scientific Biography*, Vol. 9, p. 218. New York: Charles Scribner's Sons.

10 Maxwell, J.C. (1860). *Report of the British Association for the Advancement of Science*, **28**, Part 2, 16.
11 Maxwell, J.C. (1867). *Philosophical Transactions of the Royal Society*, **157**, 49.
12 Maxwell fully appreciated the significance of van der Waals' brilliant dissertation, remarking in an article in *Nature*, 'that there can be no doubt that the name of Van der Waals will soon be among the foremost in molecular science.' Maxwell, J.C. (1973). *Nature*, **10**, 477–480.
13 This simulation was first shown to me by Peter Duffett-Smith in his earlier first year course on statistical physics.
14 Maxwell, J.C. (1867). Communication to P.G. Tait, quoted by P.M. Harman, *Energy, Force and Matter: The Conceptual Development of Nineteenth-Century Physics*, p. 140. Cambridge: Cambridge University Press.
15 Waldram, J.R. (1985). *The Theory of Thermodynamics*. Cambridge: Cambridge University Press.
16 Shannon, C.E. (1948). *The Bell System Technical Journal*, **27**, 379 and 623. (See also *The Collected Papers of Claude Elween Shannon*, (eds. N.J.A. Sloane and A.D. Wyner. Piscataway, New Jersey: IEEE Press.)
17 Kittel, C. (1969). *Thermal Physics*. New York: John Wiley & Sons.
18 Mandl, F. (1971). *Statistical Physics*. Chichester: John Wiley & Sons.

Case Study V

THE ORIGINS OF THE CONCEPTS OF QUANTISATION AND QUANTA

Quanta and relativity are phenomena quite outside our everyday experience – they are undoubtedly the greatest discoveries of twentieth century physics. In the debate between the Copernican and geocentric models of the Universe discussed in Chapter 3, one of the issues raised by opponents of the Copernican picture concerned 'The deception of the senses' – if these phenomena are of such fundamental importance, why are we not aware of them in our everyday lives?' Quanta and relativity were discovered by very careful experiment, following James Clerk Maxwell's dictum clearly enunciated in his inaugural lecture of 1871:

> But the history of science shews that even during that phase of her progress in which she devotes herself to improving the accuracy of the numerical measurement of quantities with which she has long been familiar, she is preparing the materials for the subjugation of new regions, which would have remained unknown if she had contented with the rough guide of the earlier pioneers.[1]

Later he writes:

> I might bring forward instances gathered from every branch of science, showing how the labour of careful measurement has been rewarded by the discovery of new fields of research, and by the development of new scientific ideas.

As precise experiment developed during the later nineteenth and early twentieth centuries, phenomena were discovered which were seriously at variance with Newtonian physics. Out of these the foundations of modern physics were created.

In this case study, we will investigate in some detail the origins of the concepts of quantisation and quanta. For me, this is one of the most dramatic stories in scientific history and catches the flavour of an epoch when, within 25 years, physicists' view of nature changed totally and completely new perspectives were opened up. The story illustrates many important aspects of how physics and theoretical physics work in practice. We find the greatest physicists making mistakes, individuals having to struggle against the accepted views of virtually all physicists, and, most of all, a level of inspiration and scientific creativity which I find dazzling. If only everyone, and not only those who have

[1] Maxwell, J.C. (1890). Introductory Lecture on Experimental Physics, *Scientific Papers*, **2**, 241–255.

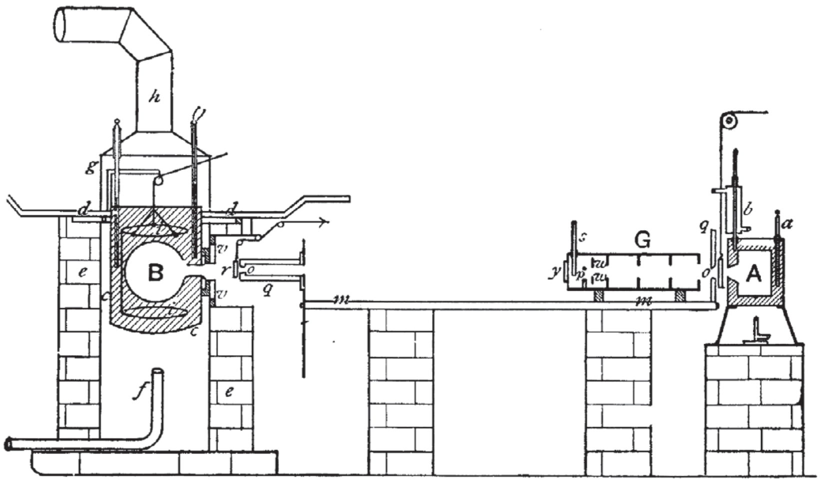


Fig. V.1 Lummer and Pringsheim's apparatus of 1897 for determining experimentally the Stefan–Boltzmann law with high precision. (From H.S. Allen and R.S. Maxwell, 1939. *A Text-book of Heat*, Part II, p. 746. London: MacMillan & Co.)

had a number of years of training as physicists or mathematicians, could appreciate the intellectual beauty of this story.

In addition to telling a fascinating and compelling story, everything essential to it will be derived using the physics and mathematics available at the time. This provides a splendid opportunity for reviewing a number of important areas of basic physics in their historical context. We will find a striking contrast between those phenomena which can be explained classically and those which necessarily involve quantum concepts.

This case study will centre upon the work of two very great physicists – Planck and Einstein. Planck is properly given credit for the discovery of *quantisation* and we will trace how this came about. Einstein's contribution was even greater in that, long before anyone else, he inferred that all natural phenomena are *quantum in nature* and he was the first to put the subject on a firm theoretical basis.

The inspiration for this case study was a set of lectures published by the late Martin J. Klein entitled *The Beginnings of the Quantum Theory* in the proceedings of the 57th Varenna Summer School.[2] When I first read the lectures, they were a revelation. I acknowledge fully my indebtedness to Professor Klein for inspiring what I consider in many ways to be the core of the present volume. Kindly, he sent me further materials after the publication of the first edition of this book. His article *Einstein and the Wave–Particle Duality*[3] is particularly valuable.

Another powerful influence was the important volume by Thomas S. Kuhn *Black-Body Theory and the Quantum Discontinuity, 1894–1912* (1978),[4] which is, in my view, a

[2] Klein, M.J. (1977). *History of Twentieth Century Physics*, Proceedings of the the International School of Physics, 'Enrico Fermi', Course 57, pp. 1–39. New York and London: Academic Press.
[3] Klein, M.J. (1964). Einstein and the Wave–Particle Duality. *The New Philosopher*, **3**, 1–49.
[4] Kuhn, T.S. (1978). *Black-Body Theory and the Quantum Discontinuity, 1894–1912*. Oxford: Clarendon Press.

masterpiece in which he digs deeply into much of the material discussed in this case study. For those who wish to pursue these topics in more detail, his book is essential reading.

The story will cover the years from about 1890 to the mid-1920s, by which date all physicists were forced to come to terms with a new view of the whole of physics in which all the fundamental entities have to be quantised. This case study ends just at the time when quantum mechanics was being codified by Werner Heisenberg, Max Born, Pascual Jordan, Paul Dirac, Wolfgang Pauli, Erwin Schrödinger and many others.

The problem I faced was that to cover that material would have required an enormous and rather different type of effort. I subsequently tackled these more mathematically elaborate developments in a book which I think of as a sequel, or companion, to the present volume, with the title *Quantum Concepts in Physics: An Alternative Approach to the Understanding of Quantum Mechanics*.[5] I regard the two books as a continuum, the transition from the classical physics of Newton and Maxwell to the contemporary physics of relativity and quantum mechanics, two of the most dramatic paradigm shifts in the history of science.

[5] Longair, M.S. (2013). *Quantum Concepts in Physics: An Alternative Approach to the Understanding of Quantum Mechanics*. Cambridge: Cambridge University Press.

13 Black-Body Radiation up to 1895

13.1 Physics and Theoretical Physics in 1890

In the course of the four case studies treated so far, we have been building up a picture of the state of physics and theoretical physics towards the end of the nineteenth century. The achievement had been immense. In mechanics and dynamics, the various approaches described in Chapter 8 were well understood. In thermodynamics, the first and second laws were firmly established, largely through the efforts of Clausius and Lord Kelvin, and the full ramifications of the concept of entropy in classical thermodynamics had been elaborated. In Chapters 5, 6 and 7, we described how Maxwell derived the equations of electromagnetism, which were completely validated by Hertz's experiments of 1887–89 – it was demonstrated beyond any shadow of doubt that light is a form of electromagnetic wave. This discovery provided a firm theoretical foundation for the wave theory of light which could account for virtually all the known phenomena of optics.

The impression is sometimes given that most physicists of the 1890s believed that the combination of thermodynamics, electromagnetism and classical mechanics could account for all known physical phenomena and that all that remained to be done was to work out the consequences of these recent hard-won achievements. In fact, it was a period of ferment when there were still many fundamental unresolved problems which exercised the greatest minds of the period.

We have discussed the ambiguous status of the kinetic theory of gases and the equipartition theorem as expounded by Clausius, Maxwell and Boltzmann. The fact that these theoretical analyses could not account satisfactorily for all the known properties of gases was a major barrier to their general acceptance. The status of atomic and molecular theories of the structure of matter came under attack both for the technical reasons outlined above and also because of a movement away from mechanistic atomic models for physical phenomena in favour of empirical or phenomenological theories. The origin of the 'resonances' within molecules, which were presumed to be the cause of spectral lines, had no clear interpretation and were an embarrassment to supporters of the kinetic theory. Maxwell and Boltzmann had discovered the statistical basis of thermodynamics, but the theory had won little support, particularly in the face of a movement which denied that kinetic theories had any value, even as hypotheses. The negative result of the Michelson–Morley experiment was announced in 1887 – we will take up that story in Case Study VI on Special and General Relativity. A useful portrait of the state of physics in the last decades of the nineteenth century is provided by the David Lindley in his book *Boltzmann's Atom*.[1]

Among these problems was the origin of the spectrum of *black-body radiation*, which was to prove to be the key, not only to the discovery of quantisation, but also to the resolution of many of the other problems listed above. The discovery of quantisation and quanta provided the groundwork for the modern quantum theory of matter and radiation, which only came into being in the years 1925 to 1930.

13.2 Kirchhoff's Law of Emission and Absorption of Radiation

The first step along the path to the understanding of the spectrum of black-body radiation was the discovery of spectral lines. In 1802, spectroscopic measurements of the solar spectrum were made by William Wollaston, who observed five strong dark lines. This observation was followed up in much greater detail by Joseph Fraunhofer, the son of a glazier and an experimenter of genius. Early in his career, Fraunhofer became a director of the Benediktbreuern Glass Factory in Bavaria, one of its main functions being the production of high quality glass components for surveying and military purposes. His objective was to improve the quality of the glasses and lenses used in these instruments and for this he needed much improved wavelength standards. The breakthrough came in 1814 with his discovery of the vast array of dark absorption features in the solar spectrum (Fig. 13.1). Ten prominent lines were labelled A, a, B, C, D, E, b, F, G and H and, in addition, there were 574 fainter lines between B and H. One of his great contributions was the invention of the *spectrograph* with which the deflection of light passing through the prism could be measured precisely. To achieve this, he placed a theodolite on its side and observed the spectrum through a telescope mounted on the rotating ring. Fraunhofer also measured the wavelengths of the lines precisely using a diffraction grating, a device invented by Thomas Young just over a decade earlier.

Fraunhofer noted that the strong absorption line which he labelled D consisted of two lines and had the same wavelength as the strong double emission line observed in

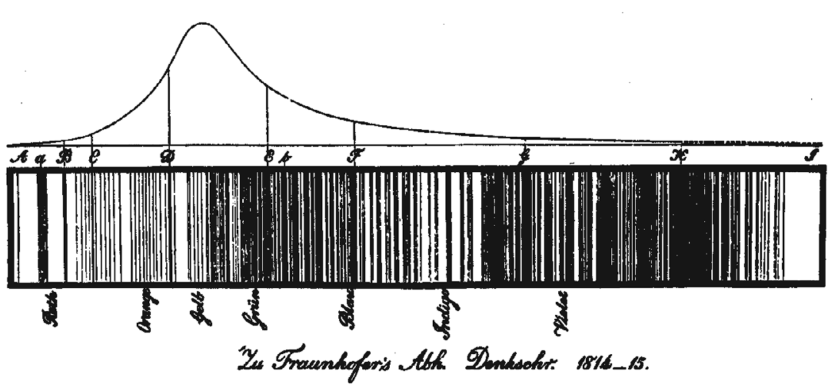


Fig. 13.1 Fraunhofer's solar spectrum of 1814 showing the vast number of absorption lines. The colours of the various regions of the spectrum are shown, as well as the labels of the prominent absorption lines.

lamp-light. Foucault was able to reproduce the D-lines in the laboratory by passing the radiation from a carbon arc through a sodium flame in front of the slit of the spectrograph. It was realised that the absorption lines are the characteristic signatures of different elements, the most important experiments being carried out by Robert Bunsen and Gustav Kirchhoff. The Bunsen burner, as it came to be called, had the great advantage of enabling the spectral lines of different elements to be determined without contamination by the burning gas. In Kirchhoff's monumental series of papers of 1861–63, entitled *Investigations of the Solar Spectrum and the Spectra of the Chemical Elements*,[2] the solar spectrum was compared with the spectra of 30 elements using a four prism arrangement, designed so that the spectrum of the element and the solar spectrum could be observed simultaneously. He concluded that the cool, outer regions of the solar atmosphere contained iron, calcium, magnesium, sodium, nickel and chromium and probably cobalt, barium, copper and zinc as well.[3]

In the course of these studies, in 1859, Kirchhoff formulated his law concerning the relation between the coefficients of emission and absorption of radiation and how these can be related to the spectrum of thermal equilibrium radiation, what came to be known as *black-body radiation*.

13.2.1 Radiation Intensity and Energy Density

We begin by clarifying the definitions we will use throughout this case study.

Consider a region of space in which an element of area $\mathrm{d}A$ is exposed to a radiation field. In the time $\mathrm{d}t$, the total amount of energy passing through $\mathrm{d}A$ is $\mathrm{d}E = S\,\mathrm{d}A\,\mathrm{d}t$, where S is the *total flux density* of radiation and has units W m^{-2}. We also need the *spectral energy distribution* of the radiation, the energy arriving *per unit frequency interval* S_ν, $S = \int S_\nu\,\mathrm{d}\nu$. We will refer to S_ν as the *flux density* and its units are W m^{-2} Hz^{-1}. In the case of a point source of radiation of luminosity L_ν at distance r, placing the elementary area normal to the direction of the source, the flux density of radiation is, by conservation of energy,

$$S_\nu = \frac{L_\nu}{4\pi r^2}. \tag{13.1}$$

The units of luminosity are W Hz^{-1}.

The flux density depends upon the orientation of the area $\mathrm{d}\boldsymbol{A}$ in the radiation field. It is meaningless to think of the radiation approaching the area along a particular ray. Rather, we consider the flux of radiation arriving within an element of solid angle $\mathrm{d}\Omega$ in the direction of the unit vector $\boldsymbol{i}_\theta$ with respect to the normal to the elementary area $\mathrm{d}\boldsymbol{A}$. This leads to one of the key quantities needed in our study, the *intensity of radiation*. If we orient the area $\mathrm{d}\boldsymbol{A}$ normal to the direction $\boldsymbol{i}_\theta$ of the incoming rays, the *intensity*, or *brightness*, of the radiation I_ν is defined to be the quotient of the flux density S_ν and the solid angle $\mathrm{d}\Omega$ from within which this flux density originates,

$$I_\nu = \frac{S_\nu}{\mathrm{d}\Omega}. \tag{13.2}$$

Thus, the intensity is the flux density per unit solid angle – its units are W m^{-2} Hz^{-1} sr^{-1}.

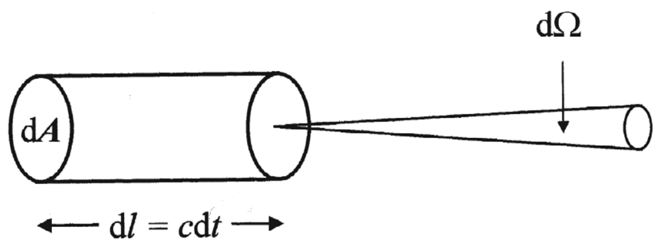


Fig. 13.2 Illustrating how to relate intensity I_ν to the energy density u_ν of radiation.

Just as in the case of fluid flow, we can write the energy flux through the area d$\boldsymbol{A}$ as

$$\mathrm{d}E = \int_\Omega I_\nu\,(\boldsymbol{i}_\theta \cdot \mathrm{d}\boldsymbol{A})\,\mathrm{d}t\,\mathrm{d}\Omega, \tag{13.3}$$

where the direction of arrival of the radiation is that of the unit vector $\boldsymbol{i}_\theta$. In the case of an isotropic radiation field, I_ν is independent of direction and so $I_\nu(\boldsymbol{\theta}) = I_0 =$ constant. Therefore,

$$\mathrm{d}E = I_0\,\mathrm{d}t \int_\Omega (\boldsymbol{i}_\theta \cdot\ \mathrm{d}\boldsymbol{A})\,\mathrm{d}\Omega = I_0\,\mathrm{d}t\,|\mathrm{d}\boldsymbol{A}| \int_0^\pi \cos\theta \times \tfrac{1}{2}\sin\theta\,\mathrm{d}\theta = 0. \tag{13.4}$$

As expected, there is no net flux density through any elementary area placed in an isotropic radiation field.

Let us now relate the intensity I_ν to the energy density of radiation u_ν within an elementary volume dV. Consider the energy arriving normal to the elementary area d$\boldsymbol{A}$ within the element of solid angle dΩ, which we can make as small as we like. Then, the radiation which passes through this area in this direction in time dt is contained in the volume of a cylinder of area dA and length $c\,\mathrm{d}t$ (Fig. 13.2). By conservation of energy, as much energy flows into the far end of the cylinder as flows out the other end at the speed of light. Therefore, the energy $U_\nu(\boldsymbol{\theta})$ contained in the volume at any time is

$$U_\nu(\boldsymbol{\theta})\,\mathrm{d}\Omega = I_\nu\,\mathrm{d}A\,\mathrm{d}t\,\mathrm{d}\Omega. \tag{13.5}$$

The energy density of radiation $u_\nu(\boldsymbol{\theta})\,\mathrm{d}\Omega$ is found by dividing by the volume of the cylinder $\mathrm{d}V = \mathrm{d}A\,c\,\mathrm{d}t$, that is

$$u_\nu(\boldsymbol{\theta})\,\mathrm{d}\Omega = \frac{\mathrm{d}U_\nu(\boldsymbol{\theta})}{\mathrm{d}V}\mathrm{d}\Omega = \frac{I_\nu\,\mathrm{d}\Omega}{c}. \tag{13.6}$$

To find the total energy density of radiation, now integrate over solid angle:

$$u_\nu = \int_\Omega u_\nu(\boldsymbol{\theta})\,\mathrm{d}\Omega = \frac{1}{c}\int I_\nu\,\mathrm{d}\Omega = \frac{4\pi}{c}J_\nu, \tag{13.7}$$

where we have defined the *mean intensity* J_ν to be

$$J_\nu = \frac{1}{4\pi}\int I_\nu\,\mathrm{d}\Omega. \tag{13.8}$$

The total energy density is found by integrating over all frequencies:

$$u = \int u_\nu \, \mathrm{d}\nu = \frac{4\pi}{c} \int J_\nu \, \mathrm{d}\nu. \tag{13.9}$$

There is one interesting feature of the intensity of a source of radiation. Suppose the source is a sphere of luminosity L_ν and radius a at distance r. Then, the intensity of radiation is

$$I_\nu = \frac{S_\nu}{\Omega} = \frac{L_\nu/4\pi r^2}{\pi a^2/r^2} = \frac{L_\nu}{4\pi} \frac{1}{\pi a^2}, \tag{13.10}$$

which is independent of the distance of the source. The intensity is just the luminosity of the source per steradian divided by its projected physical surface area.

13.2.2 Kirchhoff's Law of the Emission and Absorption of Radiation

Radiation in thermodynamic equilibrium with its surroundings was first described theoretically by Gustav Kirchhoff in 1859, while he was conducting the brilliant experiments described at the beginning of this section. They are fundamental to the understanding of the physics of radiation processes.[4]

When radiation propagates through a medium, its intensity can decrease because of absorption by the material of the medium, or increase because of its radiative properties. To simplify the argument, we assume that the emission is *isotropic* and that we can neglect scattering of the radiation. We can then define a *monochromatic emission coefficient* j_ν such that the increment in intensity into the solid angle $\mathrm{d}\Omega$ from the cylindrical volume shown in Fig. 13.2 is

$$\mathrm{d}I_\nu \, \mathrm{d}A \, \mathrm{d}\Omega = j_\nu \, \mathrm{d}V \, \mathrm{d}\Omega, \tag{13.11}$$

where j_ν has units W m^{-3} Hz^{-1} sr^{-1}. Since the emission is assumed to be isotropic, the *volume emissivity* of the medium is $\varepsilon_\nu = 4\pi j_\nu$. The volume of the cylinder is $\mathrm{d}V = \mathrm{d}A \, \mathrm{d}l$ and so

$$\mathrm{d}I_\nu = j_\nu \, \mathrm{d}l. \tag{13.12}$$

We define the *monochromatic absorption coefficient* α_ν, which results in the loss of intensity from the beam, by the relation

$$\mathrm{d}I_\nu \, \mathrm{d}A \, \mathrm{d}\Omega = -\alpha_\nu \, I_\nu \, \mathrm{d}A \, \mathrm{d}\Omega \, \mathrm{d}l. \tag{13.13}$$

This can be regarded as a phenomenological relation based upon experiment. It can be interpreted in terms of the absorption cross-section of the atoms or molecules of the medium, but this was not necessary for Kirchhoff's argument. Equation (13.13) can be simplified:

$$\mathrm{d}I_\nu = -\alpha_\nu \, I_\nu \, \mathrm{d}l. \tag{13.14}$$

Therefore, the transfer equation for radiation can be written, including both emission and absorption terms, as

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}l} = -\alpha_\nu\, I_\nu + j_\nu. \tag{13.15}$$

This equation enabled Kirchhoff to understand the relation between the emission and absorption properties of the spectral lines observed in laboratory experiments and in solar spectroscopy. If there is no absorption, $\alpha_\nu = 0$, the solution of the transfer equation is

$$I_\nu(l) = I_\nu(l_0) + \int_{l_0}^{l} j_\nu(l')\,\mathrm{d}l', \tag{13.16}$$

the first term representing the background intensity in the absence of the emission region and the second the emission from the medium located between l_0 and l. If there is absorption and no emission, $j_\nu = 0$, the solution of (13.15) is

$$I_\nu(l) = I_\nu(l_0)\exp\left[-\int_{l_0}^{l} \alpha_\nu(l')\,\mathrm{d}l'\right], \tag{13.17}$$

where the exponential term describes the absorption of radiation by the medium. The term in square brackets is often written in terms of the *optical depth* of the medium, τ_ν, where

$$\tau_\nu = \int_{l_0}^{l} \alpha_\nu(l')\,\mathrm{d}l', \quad \text{and so} \quad I_\nu(l) = I_\nu(l_0)\,\mathrm{e}^{-\tau_\nu}. \tag{13.18}$$

Kirchhoff then determined the relation between the coefficients of emission and absorption by considering the processes of emission and absorption within an enclosure which has reached thermodynamic equilibrium. The first part of the argument concerns the general properties of the equilibrium spectrum. That a unique spectrum must exist can be deduced from the second law of thermodynamics. Suppose we have two enclosures of arbitrary shape, both in thermal equilibrium at temperature T, and a filter is placed between them which allows only radiation in the frequency interval ν to $\nu + \mathrm{d}\nu$ to pass between them. Then, if $I_\nu(1)\,\mathrm{d}\nu \neq I_\nu(2)\,\mathrm{d}\nu$, energy could flow spontaneously between them without any work being done, violating the second law. The same type of argument can be used to show that the radiation must be isotropic. Therefore, the intensity spectrum of equilibrium radiation must be a *unique function of only temperature and frequency*, which can be written

$$\underbrace{I_\nu}_{\text{thermal equilibrium}} = \text{universal function of } T \text{ and } \nu \equiv B_\nu(T). \tag{13.19}$$

Now, suppose a volume of the emitting medium is maintained at temperature T. The emission and absorption coefficients have some dependence upon temperature, which we do not know. Suppose we place the emitting and absorbing volume in an enclosure in thermodynamic equilibrium at temperature T. After a very long time, the emission and absorption processes must be in balance so that there is no change in the intensity of the radiation throughout the volume. In other words, in the transfer equation (13.15), $\mathrm{d}I_\nu/\mathrm{d}l = 0$ and the intensity spectrum is the universal equilibrium spectrum $B_\nu(T)$, that is,

$$\alpha_\nu B_\nu(T) = j_\nu. \tag{13.20}$$

This is *Kirchhoff's law of emission and absorption of radiation*, showing that the emission and absorption coefficients for any physical process are related by the unknown spectrum of equilibrium radiation. This expression enabled Kirchhoff to understand the relation between the emission and absorption properties of flames, arcs, sparks and the solar atmosphere. In 1859, however, very little was known about the form of $B_\nu(T)$. As Kirchhoff remarked,

> It is a highly important task to find this function.[5]

This was one of the great experimental challenges for the remaining decades of the nineteenth century.

13.3 The Stefan–Boltzmann Law

In 1879, Josef Stefan, Director of the Institute of Experimental Physics in Vienna, deduced the law which bears his name, primarily from experiments carried out by John Tyndall at the Royal Institution of Great Britain on the radiation from a platinum strip heated to different known temperatures. Stefan also reanalysed cooling experiments carried out by Dulong and Petit in the early years of the nineteenth century.[6] He found that the total energy emitted over all wavelengths, or frequencies, is proportional to the fourth power of the absolute temperature T:

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right) = \text{total radiant energy per second} \propto T^4. \tag{13.21}$$

In 1884, his pupil Ludwig Boltzmann, by then a Professor at Graz, deduced this law from classical thermodynamics.[7] This analysis was entirely classical.

Consider a volume filled only with electromagnetic radiation and suppose that the volume is closed by a piston so that the 'gas' of radiation can be compressed or expanded. Now add some heat $đQ$ to the 'gas' reversibly. As a result, the total internal energy increases by $\mathrm{d}U$ and work is done on the piston, so that the volume increases by $\mathrm{d}V$. Then, by conservation of energy,

$$đQ = \mathrm{d}U + p\,\mathrm{d}V. \tag{13.22}$$

Writing this expression in terms of the increase in entropy $\mathrm{d}S = đQ/T$,

$$T\,\mathrm{d}S = \mathrm{d}U + p\,\mathrm{d}V, \tag{13.23}$$

as derived in Section 11.8. Equation (13.23) is now converted into a partial differential equation by taking the partial derivative of (13.23) with respect to V at constant T:

$$T\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial U}{\partial V}\right)_T + p. \tag{13.24}$$

Finally, we use one of Maxwell's relations (A11.12),

$$\left(\frac{\partial p}{\partial T}\right)_V = \left(\frac{\partial S}{\partial V}\right)_T,$$

to recast this relation as

$$T\left(\frac{\partial p}{\partial T}\right)_V = \left(\frac{\partial U}{\partial V}\right)_T + p. \tag{13.25}$$

This is the relation we were seeking. If we can work out the relation between pressure and internal energy density for a gas of radiation, we can determine the dependence of its energy density upon temperature. The relation we need was derived by Maxwell in his *Treatise on Electricity and Magnetism* of 1873.[8] Let us derive the radiation pressure of a 'gas' of electromagnetic radiation by essentially the same route followed by Maxwell. This takes a little effort, but it emphasises the important point that the equation of state of a gas of electromagnetic radiation can be derived entirely classically.

13.3.1 The Reflection of Electromagnetic Waves by a Conducting Plane

The expression for radiation pressure can be derived from the observation that, when electromagnetic waves are reflected from a conductor, they exert a force on it. This calculation needs some care and we follow the derivation by Bleaney and Bleaney.[9] We need to work out the currents which flow in the conductor under the influence of the incident waves, as well as the stresses in the materials on either side of the interface in the presence of the waves. Let us work out first the field strengths on either side of the interface for the case in which the waves are normally incident on a conducting medium which has large but finite conductivity (Fig. 13.3). This is useful revision material of some of the tools developed in our study of electromagnetism (Chapters 5 and 7).

Medium 1 is a vacuum and medium 2 a highly conducting dielectric with conductivity σ. The incident and reflected waves in a vacuum and the wave transmitted through the conductor are shown in Fig. 13.3. Using the results of Chapters 5 and 7, we can derive the

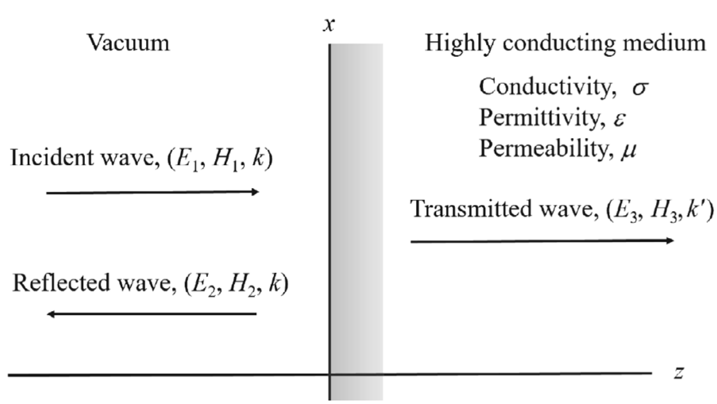


Fig. 13.3 Illustrating the boundary conditions at the interface between a vacuum and a highly conducting medium for a normally incident electromagnetic wave.

dispersion relations for waves propagating in the vacuum and in the conducting medium:

$$\textit{Medium 1} \qquad k^2 = \omega^2/c^2; \tag{13.26}$$

$$\textit{Medium 2} \qquad \nabla\times\boldsymbol{H} = \frac{\partial\boldsymbol{D}}{\partial t} + \boldsymbol{J}, \qquad \nabla\times\boldsymbol{E} = -\frac{\partial\boldsymbol{B}}{\partial t}, \tag{13.27}$$

$$\boldsymbol{J} = \sigma\boldsymbol{E}, \qquad \boldsymbol{D} = \epsilon\epsilon_0\boldsymbol{E}, \qquad \boldsymbol{B} = \mu\mu_0\boldsymbol{H}.$$

Using the relations developed in Appendix A5.6 for a wave of the form $\exp(\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t))$,

$$\nabla\times \to \mathrm{i}\boldsymbol{k}\times \quad \text{and} \quad \frac{\partial}{\partial t} \to -\mathrm{i}\omega,$$

we find

$$\begin{cases} \boldsymbol{k}\times\boldsymbol{H} = -(\omega\epsilon\epsilon_0 + \mathrm{i}\sigma)\boldsymbol{E}, \\ \boldsymbol{k}\times\boldsymbol{E} = \omega\boldsymbol{B}. \end{cases} \tag{13.28}$$

The result is similar to (13.26), but with $\omega\epsilon\epsilon_0$ replaced by $(\omega\epsilon\epsilon_0 + \mathrm{i}\sigma)$ and the inclusion of $\epsilon\mu$, that is,

$$k^2 = \epsilon\mu\frac{\omega^2}{c^2}\left(1 + \frac{\mathrm{i}\sigma}{\omega\epsilon\epsilon_0}\right). \tag{13.29}$$

Let us consider the case in which the conductivity is very high, $\sigma/\omega\epsilon\epsilon_0 \gg 1$. Then

$$k^2 = \mathrm{i}\,\frac{\mu\omega\sigma}{\epsilon_0 c^2}.$$

Since $\mathrm{i}^{1/2} = 2^{-1/2}(1 + i)$, the solution for k is

$$k = \pm(\mu\omega\sigma/2\epsilon_0 c^2)^{1/2}(1 + \mathrm{i}), \tag{13.30}$$

or

$$k = \pm(\mu\omega\sigma/\epsilon_0 c^2)^{1/2}\mathrm{e}^{\mathrm{i}\pi/4}. \tag{13.31}$$

The phase relations between $\boldsymbol{E}$ and $\boldsymbol{B}$ in the wave can now be found from the second relation of (13.28), $\boldsymbol{k}\times\boldsymbol{E} = \omega\boldsymbol{B}$.

In *free space*, $\boldsymbol{k}$ is real and $\boldsymbol{E}$ and $\boldsymbol{B}$ oscillate in phase with constant amplitude. In the conductor, however, (13.31) shows that there is a difference of $\pi/4$ between the phases of $\boldsymbol{E}$ and $\boldsymbol{B}$, and both fields decrease exponentially into the conductor, that is, from (13.30),

$$\begin{aligned} E &\propto \exp\left[\mathrm{i}(kz - \omega t)\right] \\ &= \exp\left(-\frac{z}{l}\right)\exp\left[\mathrm{i}\left(\frac{z}{l} - \omega t\right)\right], \end{aligned} \tag{13.32}$$

where $l = (2\epsilon_0 c^2/\mu\omega\sigma)^{1/2}$. The amplitude of the wave decreases by a factor 1/e in this length l, known as the *skin depth* of the conductor. This is the typical depth to which electromagnetic fields can penetrate into the conductor.

There is an important general feature of the solution represented by (13.32). If we trace back the steps involved in arriving at the solution, the assumption that the conductivity is very high corresponds to neglecting the displacement current $\partial \boldsymbol{D}/\partial t$ in comparison with $\boldsymbol{J}$. Then, the equation we have solved has the form of a diffusion equation:

$$\nabla^2 \boldsymbol{H} = \sigma \epsilon \epsilon_0 \frac{\partial \boldsymbol{H}}{\partial t}. \tag{13.33}$$

In general, wave solutions of diffusion equations have dispersion relations of the form $k = A(1 + \mathrm{i})$, that is, the real and imaginary parts of the wave vector are equal, corresponding to waves which decay in amplitude by $\mathrm{e}^{-2\pi}$ in each complete cycle. This is true for equations such as:

(a) *The diffusion equation*

$$D\nabla^2 N - \frac{\partial N}{\partial t} = 0, \tag{13.34}$$

where D is the diffusion coefficient and N the number density of particles.

(b) *The heat diffusion equation*

$$\kappa \nabla^2 T - \frac{\partial T}{\partial t} = 0, \qquad \kappa = \frac{K}{\rho C}, \tag{13.35}$$

where K is the thermal conductivity of the medium, ρ its density, C the specific heat capacity and T its temperature;

(c) *Viscous waves*

$$\frac{\mu}{\rho}\nabla^2 \omega - \frac{\partial \boldsymbol{\omega}}{\partial t} = 0, \tag{13.36}$$

where μ is the viscosity, ρ the density of the fluid and ω the vorticity. This equation is derived from the Navier–Stokes equation for fluid flow in a viscous medium (see Section 9.5.2).

(d) *Deep water waves*

$$g\,\nabla^2 \phi - \frac{\partial \phi}{\partial t} = 0, \tag{13.37}$$

where g is the acceleration due to gravity and ϕ is the velocity potential (see Section 9.4).

Returning to our story, we now match the $\boldsymbol{E}$ and $\boldsymbol{H}$ vectors of the waves at the interface between the two media. Taking z to be the direction normal to the interface, we find:

Incident wave

$$\begin{cases} E_x = E_1 \exp[\mathrm{i}(kz - \omega t)], \\ H_y = \dfrac{E_1}{Z_0} \exp[\mathrm{i}(kz - \omega t)], \end{cases} \tag{13.38}$$

where $Z_0 = (\mu_0/\epsilon_0)^{1/2}$ is the impedance of free space.

Reflected wave

$$\begin{cases} E_x = E_2 \exp[-\mathrm{i}(kz + \omega t)], \\ H_y = -\dfrac{E_2}{Z_0} \exp[-\mathrm{i}(kz + \omega t)]. \end{cases} \tag{13.39}$$

Transmitted wave

$$\begin{cases} E_x = E_3 \exp[\mathrm{i}(k'z - \omega t)], \\ H_y = E_3 \dfrac{(\mu\omega\sigma/2\epsilon_0 c^2)^{1/2}}{\omega\mu\mu_0}(1 + \mathrm{i}) \exp[\mathrm{i}(k'z - \omega t)], \end{cases} \tag{13.40}$$

where k' is given by the value of k in the relation (13.30). The values of H_y are found by substituting for k in the relation between $\boldsymbol{E}$ and $\boldsymbol{B}$, $\boldsymbol{k} \times \boldsymbol{E} = \omega \boldsymbol{B}$.

Let us write $q = [(\mu\omega\sigma/2\epsilon_0 c^2)^{1/2}/\omega\mu\mu_0](1 + \mathrm{i})$. Then,

$$H_y = qE_3 \exp[\mathrm{i}(k'z - \omega t)]. \tag{13.41}$$

The boundary conditions require E_x and H_y to be continuous at the interface (see Section 7.5), that is, at $z = 0$,

$$\begin{cases} E_1 + E_2 = E_3, \\ \dfrac{E_1}{Z_0} - \dfrac{E_2}{Z_0} = qE_3. \end{cases} \tag{13.42}$$

Therefore,

$$\frac{E_1}{(1 + qZ_0)} = \frac{E_2}{(1 - qZ_0)} = \frac{E_3}{2}. \tag{13.43}$$

In general, q is a complex number and hence there are phase differences between E_1, E_2 and E_3. We are, however, interested in the case in which the conductivity is very large, $|q|Z_0 \gg 1$ and hence

$$\frac{E_1}{qZ_0} = -\frac{E_2}{qZ_0} = \frac{E_3}{2}, \tag{13.44}$$

that is,

$$E_1 = -E_2 = \frac{qZ_0}{2} E_3.$$

Therefore, on the vacuum side of the interface, the total electric field strength $E_1 + E_2$ is zero and the magnetic field strength $H_1 + H_2 = 2H_1$.

It may seem as though we have strayed rather far from the thermodynamics of radiation, but we are now able to work out the pressure exerted by the incident wave upon the surface.

13.3.2 The Formula for Radiation Pressure

Suppose the radiation is confined within a box with rectangular sides and that the waves bounce back and forth between the walls at $z = \pm z_1$. Assuming that the walls of the box are highly conducting, we now know the values of the electric and magnetic field strengths of the waves in the vicinity of the walls for normal incidence.

Part of the origin of the phenomenon of radiation pressure may be understood as follows. Referring to Fig. 13.3, the electric field in the conductor E_x causes a current to flow in the $+x$-direction with current density

$$J_x = \sigma E_x. \tag{13.45}$$

But the force per unit volume acting on this electric current in the presence of a magnetic field of the wave is

$$\boldsymbol{F} = N_q q(\boldsymbol{v} \times \boldsymbol{B}) = \boldsymbol{J} \times \boldsymbol{B}, \tag{13.46}$$

where N_q is the number of conduction electrons per unit volume and q the electronic charge. Since $\boldsymbol{B}$ is in the $+y$-direction, this force acts in the $\boldsymbol{i}_x \times \boldsymbol{i}_y$ direction, that is, in the positive x-direction parallel to the wave vector $\boldsymbol{k}$. Therefore, the pressure acting on a layer of thickness $\mathrm{d}z$ in the conductor is

$$\mathrm{d}p = J_x B_y \, \mathrm{d}z. \tag{13.47}$$

We also know that, because the conductivity is very high in the conductor, curl $\boldsymbol{H} = \boldsymbol{J}$ and so we can relate J_x and B_y by

$$\left(\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z}\right) = J_x.$$

Since $H_z = 0$, $-\partial H_y/\partial z = J_x$. Substituting into (13.46),

$$\mathrm{d}p = -B_y \frac{\partial H_y}{\partial z} \, \mathrm{d}z. \tag{13.48}$$

Therefore,

$$p = -\int_0^\infty B_y \frac{\partial H_y}{\partial z} \, \mathrm{d}z = \int_0^{H_0} B_y \, \mathrm{d}H_y, \tag{13.49}$$

where H_0 is the value of the magnetic field strength in the wave at the interface and, according to the analysis of Section 13.3.1, $H \to 0$ as $z \to \infty$. For a linear medium, $B_0 = \mu\mu_0 H_0$ and hence

$$p = \tfrac{1}{2}\mu\mu_0 H_0^2.$$

Notice that this pressure is associated with currents flowing in the conducting medium under the influence of the wave.

The other forces acting at the vacuum–conductor interface are associated with the stresses in the electromagnetic fields themselves and are given by the appropriate components of the Maxwell stress tensor. We can derive the magnitude of these stresses by an argument based on Faraday's concept of lines of force. Suppose a uniform longitudinal magnetic field is confined within a long rectangular perfectly conducting tube. If the medium is linear, $B = \mu\mu_0 H$, the energy per unit length of tube is

$$E = \tfrac{1}{2} BHzl,$$

where z is the width and l the breadth of the tube (Fig. 13.4). Now squash the rectangle by $\mathrm{d}z$ in the z-direction so that the width of the tube becomes $(z - \mathrm{d}x)$ whilst maintaining the same number of lines of force through it, which occurs in the case of a very highly conducting medium because of the phenomenon of magnetic flux freezing (see Section 9.7.3).

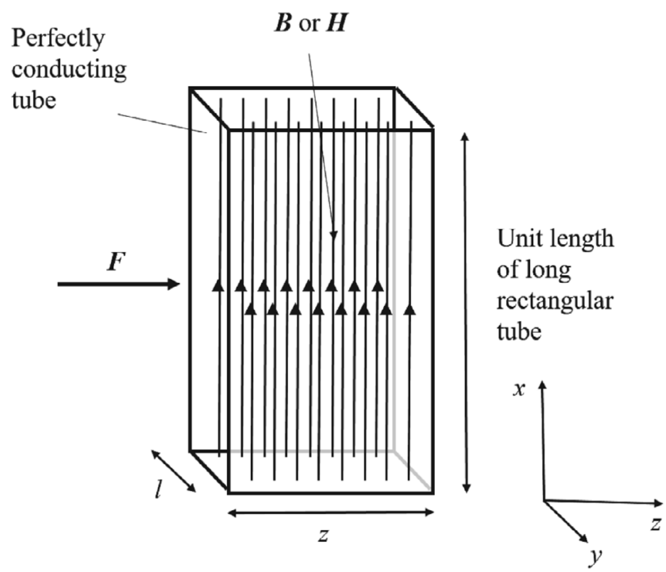


Fig. 13.4 A long perfectly conducting rectangular tube enclosing a longitudinal magnetic field. When the tube is compressed, the magnetic flux enclosed by the tube is conserved. For more details of magnetic flux freezing, see, for example, M.S. Longair (2011). *High Energy Astrophysics*, pp. 304–311. Cambridge: Cambridge University Press.

Then the magnetic flux density increases to $Bz/(z - \mathrm{d}z)$ and correspondingly, because the field is linear, H becomes $Hz/(z - \mathrm{d}z)$. Thus, the energy in the volume becomes

$$\begin{aligned} E + \mathrm{d}E &= \tfrac{1}{2}BHlz^2(z - \mathrm{d}z)^{-1}\,, \\ \mathrm{d}E &= \tfrac{1}{2}BHlz\left(1 + \frac{\mathrm{d}z}{z}\right) - \tfrac{1}{2}BHlz \\ &= \tfrac{1}{2}BHl\,\mathrm{d}z. \end{aligned}$$

This increase in energy must be the result of work done on the field,

$$F\,\mathrm{d}z = \mathrm{d}E = \tfrac{1}{2}BHl\,\mathrm{d}z.$$

Thus, the force per unit area is $F/l = \frac{1}{2}BH$ and so the pressure associated with the magnetic field within the tube is

$$p = \tfrac{1}{2}\mu\mu_0 H^2, \qquad (13.50)$$

perpendicular to the field direction. We can apply exactly the same argument to electrostatic fields, in which case the total pressure associated with the electric and magnetic fields is

$$p = \tfrac{1}{2}\epsilon\epsilon_0 E^2 + \tfrac{1}{2}\mu\mu_0 H^2. \qquad (13.51)$$

Because the values of ϵ and μ are different on either side of the interface, there is a pressure difference across it in the opposite direction to that associated with the induced electric currents.

We have shown that, at the interface, $E_x = 0$ and hence in the vacuum $p = \frac{1}{2}\mu_0 H_0^2$. Inside the conductor at $z = 0$, the stress is $\frac{1}{2}\mu\mu_0 H_0^2$. Therefore, the total pressure on the conductor is

$$p = \underbrace{\tfrac{1}{2}\mu_0 H_0^2}_{\text{stress in vacuum}} - \underbrace{\tfrac{1}{2}\mu\mu_0 H_0^2}_{\text{stress in conductor}} + \underbrace{\tfrac{1}{2}\mu\mu_0 H_0^2}_{\text{force on conduction current}},$$

or,

$$p = \tfrac{1}{2}\mu_0 H_0^2. \tag{13.52}$$

When the conductivity is very high, we have shown that the magnetic field strength $H_0 = 2H_1$, where H_1 is the field strength of the wave propagating in the positive z-direction in the vacuum and the energy density in this wave is $\frac{1}{2}(\epsilon_0 E_1^2 + \mu_0 H_1^2) = \mu_0 H_1^2$. Therefore,

$$p = \tfrac{1}{2}\mu_0 H_0^2 = 2\mu_0 H_1^2 = 2\varepsilon_1 = \varepsilon_0, \tag{13.53}$$

where ε_0 is the total energy density of radiation in the vacuum, being the sum of the energy densities in the incident and reflected waves.

The expression (13.53) is the relation between the pressure and energy density for a 'one-dimensional' gas of electromagnetic radiation, confined between two reflecting walls. In an isotropic three-dimensional case, there are equal energy densities associated with radiation propagating in the three orthogonal directions, that is,

$$\varepsilon_x = \varepsilon_y = \varepsilon_z = \varepsilon_0,$$

and hence

$$p = \tfrac{1}{3}\varepsilon, \tag{13.54}$$

where ε is the total energy density of radiation.

This somewhat lengthy demonstration shows how it is possible to derive the pressure of a gas of electromagnetic radiation entirely by classical electromagnetic arguments.

13.3.3 The Derivation of the Stefan–Boltzmann Law

The relation (13.54) between p, V and U for a gas of electromagnetic radiation solves the problem of deriving its equation of state, since $U = \varepsilon V$ and $p = \frac{1}{3}\varepsilon$. Hence, from (13.25),

$$\begin{aligned}
T\left(\frac{\partial p}{\partial T}\right)_V &= \left(\frac{\partial U}{\partial V}\right)_T + p,\\
\tfrac{1}{3}T\left(\frac{\partial \varepsilon}{\partial T}\right)_V &= \left(\frac{\partial (\varepsilon V)}{\partial V}\right)_T + \tfrac{1}{3}\varepsilon,\\
\tfrac{1}{3}T\left(\frac{\partial \varepsilon}{\partial T}\right)_V &= \varepsilon + \tfrac{1}{3}\varepsilon = \tfrac{4}{3}\varepsilon.
\end{aligned}$$

This relation between ε and T can be integrated to give

$$\frac{\mathrm{d}\varepsilon}{\varepsilon} = 4\,\frac{\mathrm{d}T}{T}, \qquad \ln\varepsilon = 4\ln T, \qquad \varepsilon = aT^4. \tag{13.55}$$

This is the calculation which Boltzmann carried out in 1884 and his name is justly associated with the *Stefan–Boltzmann law* and the Stefan–Boltzmann constant a, which has value $a = 7.566 \times 10^{-16}$ J m^{-3} K^{-4}. The expression for the Stefan–Boltzmann constant will be derived theoretically in Section 15.3.

The expression (13.55) for the energy density of radiation can be related to the energy emitted per second from the surface of a black-body at temperature T. If the black-body is placed in an enclosure maintained in thermodynamic equilibrium at temperature T, the expression for the rate of arrival of waves at unit area of its surface is given by (12.13), namely, the flux of waves is $\frac{1}{4}N\overline{c} = \frac{1}{4}Nc$, since all the waves travel at the speed of light; N is the number density of waves. If the energy of each wave is E, the rate at which energy arrives at the wall per unit area per second, and consequently the rate at which the energy must be reradiated from it, is $\frac{1}{4}N(E)Ec = \frac{1}{4}N\overline{E}c = \frac{1}{4}\varepsilon c$. Therefore,

$$I = \tfrac{1}{4}\varepsilon c = \frac{ac}{4}T^4 = \sigma T^4 = 5.67 \times 10^{-8}\, T^4 \text{ W m}^{-2}. \tag{13.56}$$

The σ in (13.56) is conventionally used – it is *not* the conductivity.

The experimental evidence for the Stefan–Boltzmann law was not particularly convincing in 1884 and it was not until 1897 that Lummer and Pringsheim undertook very careful experiments which showed that the law was indeed correct with high precision. Their apparatus is shown in Fig. V.1 of the introduction to this case study.

13.4 Wien's Displacement Law and the Spectrum of Black-Body Radiation

The spectrum of black-body radiation was not particularly well known in 1895, but there had already been important theoretical work carried out by Wilhelm Wien on the form which the radiation law should have. *Wien's displacement law*[10] was derived using a combination of electromagnetism, thermodynamics and dimensional analysis. This work, published in 1893, was to be of central importance in the subsequent story.

First of all, Wien worked out how the properties of a 'gas' of radiation change when it undergoes a reversible adiabatic expansion. The analysis starts with the basic thermodynamic relation

$$đQ = \mathrm{d}U + p\,\mathrm{d}V.$$

In an adiabatic expansion $đQ = 0$ and, for a 'gas' of radiation $U = \varepsilon V$ and $p = \frac{1}{3}\varepsilon$. Therefore,

$$\begin{aligned} \mathrm{d}(\varepsilon V) + \tfrac{1}{3}\varepsilon\,\mathrm{d}V &= 0, \\ V\,\mathrm{d}\varepsilon + \varepsilon\,\mathrm{d}V + \tfrac{1}{3}\varepsilon\,\mathrm{d}V &= 0, \\ \frac{\mathrm{d}\varepsilon}{\varepsilon} &= -\tfrac{4}{3}\frac{\mathrm{d}V}{V}. \end{aligned}$$

Integrating,

$$\varepsilon = \text{constant} \times V^{-4/3}. \tag{13.57}$$

But $\varepsilon = aT^4$ and hence

$$TV^{1/3} = \text{constant}. \tag{13.58}$$

Since V is proportional to the cube of the radius r of a spherical volume,

$$T \propto r^{-1}. \tag{13.59}$$

The next step is to work out the relation between the wavelength of the radiation and the volume of the enclosure in a reversible adiabatic expansion. Two straightforward calculations provide the answer. First of all, we determine the change in wavelength if a wave is reflected from a slowly moving mirror (Fig. 13.5). The initial position of the mirror is at X and we suppose that at that time one of the maxima of the incident waves is at A. It is then reflected along the path AC. By the time the next wavecrest arrives at the mirror, the latter has moved to X' and hence the maximum has to travel an extra distance ABN as compared with the first maximum, that is, the wavelength has increased by $\mathrm{d}\lambda = AB + BN$. By symmetry,

$$AB + BN = A'N = AA' \cos\theta.$$

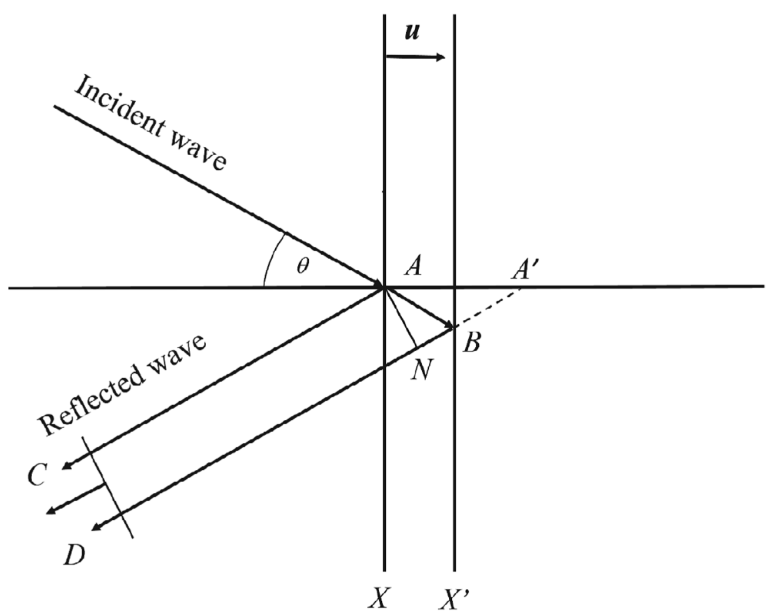


Fig. 13.5 Illustrating the change in wavelength of an electromagnetic wave reflected from a moving perfectly conducting plane.

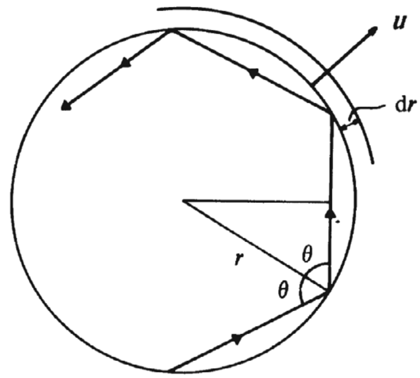


Fig. 13.6 Illustrating the differential change in wavelength of an electromagnetic wave as it is reflected inside a spherical volume expanding at speed $\boldsymbol{u}$.

But

$$AA' = 2 \times \text{distance moved by mirror} = 2XX' = 2uT,$$

where u is the velocity of the mirror and T the period of the wave. Hence,

$$\begin{aligned} \mathrm{d}\lambda &= 2uT\cos\theta, \\ &= 2\lambda\frac{u}{c}\cos\theta, \quad \text{since} \quad \mathrm{d}\lambda \ll \lambda. \end{aligned} \tag{13.60}$$

This result is correct to first order in small quantities, but that is all we need since we are only interested in differential changes.

Now suppose the wave is confined within a spherical reflecting cavity, which is expanding slowly, and that its angle of incidence with respect to the normal to the sphere is θ (Fig. 13.6). We can work out the number of reflections which the wave makes with the sphere, and hence the total change in wavelength, as the cavity expands from r to $r + \mathrm{d}r$ at velocity $\boldsymbol{u}$, assuming $|\boldsymbol{u}| \ll c$.

If the velocity of expansion is small, θ remains the same for all reflections as the sphere expands by an infinitesimal distance $\mathrm{d}r$. The time it takes the sphere to expand a distance $\mathrm{d}r$ is $\mathrm{d}t = \mathrm{d}r/u$ and, from the geometry of Fig. 13.6, the time between reflections for a wave propagating at velocity c is $2r\cos\theta/c$. The number of reflections in time $\mathrm{d}t$ is therefore $c\,\mathrm{d}t/2r\cos\theta$ and the change of wavelength $\mathrm{d}\lambda$ is

$$\mathrm{d}\lambda = \left(\frac{2u\lambda}{c}\cos\theta\right)\left(\frac{c\,\mathrm{d}t}{2r\cos\theta}\right),$$

that is,

$$\frac{\mathrm{d}\lambda}{\lambda} = \frac{u\,\mathrm{d}t}{r} = \frac{\mathrm{d}r}{r}.$$

Integrating, we find

$$\lambda \propto r. \tag{13.61}$$

Thus, the wavelength of the radiation increases linearly proportionally to the radius of the spherical volume.

We can now combine this result with (13.59), $T \propto r^{-1}$, with the result

$$T \propto \lambda^{-1}. \tag{13.62}$$

This is one aspect of *Wien's displacement law*. If radiation undergoes a reversible adiabatic expansion, the wavelength of a particular set of waves changes inversely with temperature. In other words, the wavelength of the radiation is 'displaced' as the temperature changes. In particular, if we follow the maximum of the radiation spectrum, it should follow the law $\lambda_{\max} \propto T^{-1}$. This was found to be in good agreement with experiment.

Wien now went significantly further and combined the two laws, the Stefan–Boltzmann law and the law $T \propto \lambda^{-1}$, to set constraints upon the form of the equilibrium radiation spectrum. The first step is to note that, if any system of bodies is enclosed in a perfectly reflecting enclosure, eventually all of them come to thermodynamic equilibrium at the same temperature through the emission and absorption of radiation. In this state, as much energy is radiated as is absorbed by the bodies per unit time. If we wait long enough, the radiation attains a black-body radiation spectrum at a certain temperature T. According to the arguments presented in Section 13.2.2, first enunciated by Kirchhoff, the radiation must be isotropic and the only parameters which can characterise the radiation spectrum are the temperature, T, and the wavelength of the radiation, λ.

The next step is to note that, if black-body radiation is initially at temperature T_1, and the enclosure undergoes a reversible adiabatic expansion, by definition, the expansion proceeds infinitely slowly so that the radiation takes up an equilibrium spectrum at all stages of the expansion to the lower temperature T_2. The crucial point is that the radiation spectrum must be of equilibrium black-body form at the beginning and end of the expansion. The unknown radiation law must therefore scale appropriately with temperature.

Consider the radiation in the wavelength interval λ_1 to $\lambda_1 + \mathrm{d}\lambda_1$, and let its energy density be $\varepsilon = u(\lambda_1)\,\mathrm{d}\lambda_1$, where $u(\lambda)$ is the energy density per unit wavelength interval. Then, according to Boltzmann's analysis, the energy associated with any particular set of waves changes as T^4 and hence

$$\frac{u(\lambda_1)\,\mathrm{d}\lambda_1}{u(\lambda_2)\,\mathrm{d}\lambda_2} = \left(\frac{T_1}{T_2}\right)^4. \tag{13.63}$$

But $\lambda_1 T_1 = \lambda_2 T_2$, and hence $\mathrm{d}\lambda_1 = (T_2/T_1)\,\mathrm{d}\lambda_2$. Therefore,

$$\frac{u(\lambda_1)}{T_1^5} = \frac{u(\lambda_2)}{T_2^5}, \quad \text{that is,} \quad \frac{u(\lambda)}{T^5} = \text{constant}. \tag{13.64}$$

Since λT = constant, (11.64) can be rewritten

$$u(\lambda)\,\lambda^5 = \text{constant}. \tag{13.65}$$

Now, the only combination of T and λ which is a constant throughout the expansion is the product λT, and hence we can write that, in general, the constant in (13.65) can only be constructed out of functions involving λT. Therefore, the radiation law must have the form

$$u(\lambda)\,\lambda^5 = f(\lambda T), \tag{13.66}$$

or

$$u(\lambda)\,\mathrm{d}\lambda = \lambda^{-5} f(\lambda T)\,\mathrm{d}\lambda. \tag{13.67}$$

This is the complete form of *Wien's displacement law* and it sets important constraints upon the form of the black-body radiation spectrum.

We will find it more convenient to work in terms of frequencies rather than wavelengths, and so we convert Wien's displacement law into frequency form:

$$u(\lambda)\,\mathrm{d}\lambda = u(\nu)\,\mathrm{d}\nu,$$
$$\lambda = c/\nu, \quad \mathrm{d}\lambda = -\frac{c}{\nu^2}\,\mathrm{d}\nu.$$

Hence

$$u(\nu)\,\mathrm{d}\nu = \left(\frac{c}{\nu}\right)^{-5} f\left(\frac{\nu}{T}\right)\left(-\frac{c}{\nu^2}\,\mathrm{d}\nu\right), \tag{13.68}$$

that is,

$$u(\nu)\,\mathrm{d}\nu = \nu^3 f\left(\frac{\nu}{T}\right)\mathrm{d}\nu. \tag{13.69}$$

This is really rather clever. Notice how much Wien was able to deduce using only rather general thermodynamic arguments. This argument proved to be crucial in establishing the correct formula for black-body radiation.

This was all new when Max Planck first became interested in the problem of the spectrum of equilibrium radiation in 1895.

Notes

1 Lindley, D. (2001). *Boltzmann's Atom*. New York: The Free Press.
2 Kirchhoff, G. (1861–63). *Abhandlungen der königlich Preussischen Akademie der Wissenschaften zu Berlin*, Part 1 (1861) pp. 62–95. (1862) pp. 227–240; Part 2 (1863) pp. 225–240.
3 An excellent description of these early spectroscopic studies is given by J.B. Hearnshaw in *The Analysis of Starlight*. Cambridge: Cambridge University Press (1986).
4 I particularly like the careful discussion of the fundamentals of radiative processes by G.B. Rybicki and A.P. Lightman in *Radiative Processes in Astrophysics*. New York: John Wiley and Sons (1979).
5 Kirchhoff, G. (1859). *Berlin Monatsberichte*, pp. 783–787 (trans. *Philosophical Magazine*, **19**, 193, 1860).
6 Stefan, J. (1879). *Sitzungberichte der Kaiserlichen Akademie der Wissenschaften (Wien)*, **79**, 391–428.

7 Boltzmann, L. (1884). *Annalen der Physik*, **22**, 291–294.
8 Maxwell, J.C. (1873). *A Treatise on Electricity and Magnetism*, Vol. 2, articles 792 and 793. Oxford: Clarendon Press. (Reprint of the 2nd edition of 1891 published by Oxford University Press (1998) in two volumes.)
9 Bleaney, B.I. and Bleaney, B. (1965). *Electricity and Magnetism*, 2nd edition, pp. 278–281. Oxford, Clarendon Press.
10 Wien, W. (1893). *Berlin Monatsberichte*, pp. 55–62.

14 1895–1900: Planck and the Spectrum of Black-Body Radiation

14.1 Planck's Early Career

Max Planck was typically straightforward and honest about his early career, as described in his short scientific autobiography.[1] He studied under Helmholtz and Kirchhoff in Berlin, but in his own words:

> I must confess that the lectures of these men netted me no perceptible gain. It was obvious that Helmholtz never prepared his lectures properly … Kirchhoff was the very opposite … but it would sound like a memorised text, dry and monotonous.

Planck's research interests were inspired by his reading of the works of Clausius and he set about investigating how the second law of thermodynamics could be applied to a wide variety of different physical problems, as well as elaborating as clearly as possible the basic tenets of the subject. He defended his dissertation *On the Second Law of Thermodynamics* in 1879. In his words,

> the effect of my dissertation on the physicists of those days was nil. None of my professors at the University had any understanding for its contents, as I learned for a fact in my conversations with them … Helmholtz probably did not even read my paper at all. Kirchhoff expressly disapproved of its contents, with the comment that … the concept of entropy … must not be applied to irreversible processes. I did not succeed in reaching Clausius. He did not answer my letters and I did not find him at home when I tried to see him in person in Bonn.[2]

It is interesting that Kirchhoff was incorrect in stating that the concept of entropy cannot be applied to irreversible processes. Entropy is a function of state, and therefore can be determined for any given state of the system, independent of whether or not reversible or irreversible processes were involved in attaining that state. In 1880, Planck obtained first a doctorate and then presented his thesis *Equilibrium states of isotropic bodies at different temperatures*, for which he was awarded his habilitation. Following the death of Kirchhoff in 1889, he succeeded to Kirchhoff's chair at the University of Berlin. This was a propitious advance since the University was one of the most active centres of physics research in the world and German physics was at the very forefront of the discipline.

In 1894, he turned his attention to the problem of the spectrum of black-body radiation, the topic which was to dominate his subsequent career and which led to the discovery of quantisation. It is likely that his interest in this problem was stimulated by Wien's important paper which had been published in the previous year (Section 13.4). Wien's analysis had a strong thermodynamic flavour, which must have appealed to Planck.

In 1895, he published the first results of his work on the resonant scattering of plane electromagnetic waves by an oscillating dipole. This was the first of Planck's papers in which he diverged from his previous areas of interest in that it appeared to be about electromagnetism rather than thermodynamics. In the last words of the paper, however, Planck made it clear that he regarded this as a first step towards tackling the problem of the spectrum of black-body radiation. His aim was to set up a system of oscillators in an enclosed cavity which would radiate and interact with the material and contents of the cavity so that after a very long time the system would come into thermodynamic equilibrium. He could then apply the laws of thermodynamics to black-body radiation with a view to understanding the origin of its spectrum.

He explained why this approach offered the prospect of providing insights into basic thermodynamic processes. When energy is lost by an oscillator by radiation, it is not dissipated as heat, but emitted as electromagnetic waves. The process can be considered 'conservative' because, if the radiation is enclosed in a box with perfectly reflecting walls, it can then react back on the oscillators and the energy returned to them by the corresponding absorption process according to Kirchhoff's law. Furthermore, the process is independent of the nature of the oscillator. In Planck's words:

> The study of conservative damping seems to me to be of great importance, since it opens up the prospect of a possible general explanation of irreversible processes by means of conservative forces – a problem that confronts research in theoretical physics more urgently every day.[3]

Thermodynamics was Planck's area of expertise and he was a follower of Clausius as an exponent of classical thermodynamics. He looked upon the second law of thermodynamics as having absolute validity – processes in which the total entropy decreased, he believed, should be strictly excluded. This is a very different point of view from that set forth by Boltzmann in his memoir of 1877 in which the second law of thermodynamics is only statistical in nature. As demonstrated in Chapter 12, according to Boltzmann, there is a very high probability indeed that entropy increases in all natural processes, but there remains an extremely small but finite probability that the system will evolve to a state of lower entropy. As is clearly indicated in Kuhn's history of these years,[4] there were many technical concerns about the Boltzmann's approach and Planck and his students had published papers criticising some of the steps in Boltzmann's statistical arguments.

Planck believed that, by studying classically the interaction of oscillators with electromagnetic radiation, he would be able to show that entropy increases absolutely for a system consisting of matter and radiation. These ideas were elaborated in a series of five papers. The idea did not work, as was pointed out by Boltzmann. One cannot obtain a monotonic approach to equilibrium without some statistical assumptions about the way in which the system approaches the equilibrium state. This can be understood from Maxwell's simple but compelling argument concerning the time reversibility of the laws of mechanics, dynamics and electromagnetism.

Finally, Planck conceded that statistical assumptions were necessary and introduced the concept of 'natural radiation', which corresponded to Boltzmann's assumption of 'molecular chaos'. It must have been a painful experience for Planck to have the assumptions upon which his researches had been based for 20 years to be undermined.[5] Much worse was, however, to come.

Once the assumption was made that there exists a state of 'natural radiation', Planck was able to make considerable progress with his programme. The first thing he did was to relate the energy density of radiation in an enclosure to the average energy of the oscillators within it. This is a crucial result and can be derived entirely from classical arguments. Let us show how it can be done – it is a beautiful piece of theoretical physics.

14.2 Oscillators and Their Radiation in Thermal Equilibrium

Why do we treat oscillators rather than, say, atoms, molecules, lumps of rock, and so on? The reason is that, in thermal equilibrium, everything is in equilibrium with everything else at the same temperature – rocks are in equilibrium with atoms and oscillators and therefore there is no advantage in treating complicated objects. The advantage of considering simple harmonic oscillators is that the radiation and absorption laws can be calculated exactly. To express this important point in another way, Kirchhoff's law (13.20), which relates the emission and absorption coefficients, j_ν and α_ν respectively,

$$\alpha_\nu B_\nu(T) = j_\nu,$$

shows that, if we can determine these coefficients for any process, we can find the universal black-body equilibrium spectrum $B_\nu(T)$. Let us build up the essential classical tools we need.

14.2.1 The Rate of Radiation of an Accelerated Charged Particle

Charged particles emit electromagnetic radiation when they are accelerated. First of all, we need the rather beautiful formula for the rate of emission of radiation of an accelerated charged particle. The normal derivation of this formula proceeds from Maxwell's equations through the use of retarded potentials for the field components at a large distance. We can give a much simpler derivation of the exact results of such calculations from a remarkable argument given by J.J. Thomson in his book *Electricity and Matter*,[6] the published version of his Silliman Memorial Lectures delivered in Harvard in 1903. He used this formula to derive the cross-section for the scattering of X-rays by the electrons in atoms in his subsequent book *Conduction of Electricity through Gases*.[7] This was the origin of the *Thomson cross-section* for the scattering of electromagnetic radiation by free electrons. Thomson's argument indicates very clearly why an accelerated charge radiates electromagnetic radiation and also clarifies the origin of the polar diagram and polarisation properties of the radiation.

Consider a charge q stationary at the origin O of some frame of reference S at time $t = 0$. The charge then suffers a small acceleration to velocity Δv in a short interval of time Δt. Thomson visualised the resulting electric field distribution in terms of the electric field lines attached to the accelerated charge. After a time $t \gg \Delta t$, we can distinguish between the field configuration inside and outside a sphere of radius ct centred on the origin of S. Outside this sphere, the field lines do not yet know that the charge has moved away from the origin because information cannot travel faster than the speed of light and therefore they are radial, centred on O. Inside this sphere, the field lines are radial about the origin of the frame of reference centred on the moving charge at O' at distance $\Delta v\, t$ from O. Between these two regions, there is a thin shell of thickness $c\Delta t$ in which we join up corresponding electric field lines, indicated schematically in Fig. 14.1(a). Geometrically, it is clear that there must be a component of the electric field in the circumferential direction in the

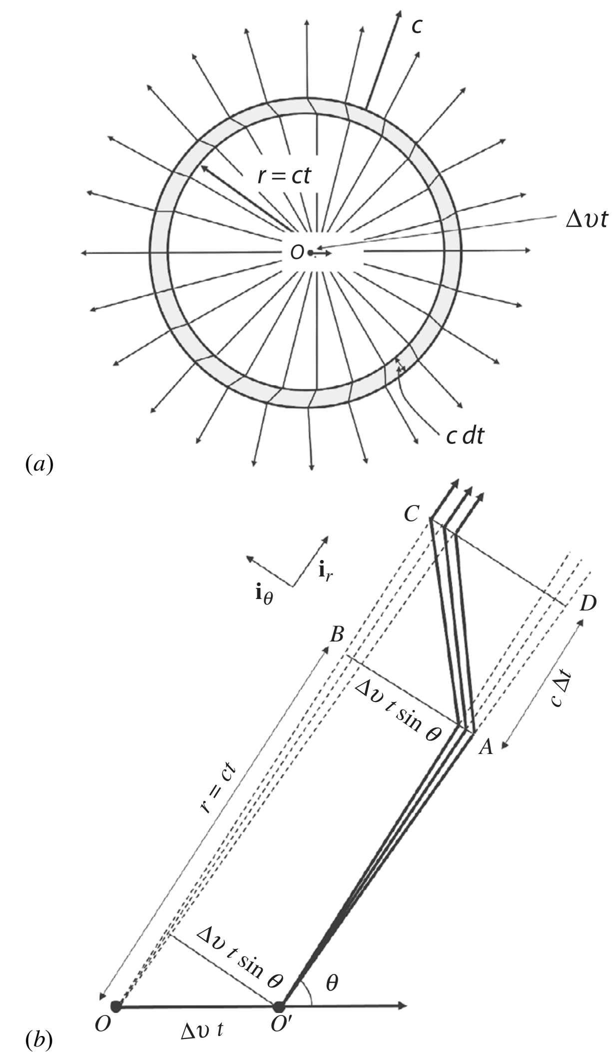


Fig. 14.1 Illustrating J.J. Thomson's method of evaluating the radiation of an accelerated charged particle. The diagram shows schematically the configuration of electric field lines at time t due to a charge accelerated to a velocity Δv in time Δt at $t = 0$. (b) An expanded version of (a) used to evaluate the strength of the E_θ component of the electric field due to the acceleration of the electron. (From M.S. Longair, 2011, *High Energy Astrophysics*, 3rd edition. Cambridge: Cambridge University Press.)

shell, that is, in the $\mathbf{i}_\theta$ direction. This 'pulse' of electromagnetic field is propagated away from the charge at the speed of light and represents an energy loss from the accelerated charged particle.

Let us work out the strength of the electric field in the pulse. We assume that the increment in velocity Δv is very small, that is, $\Delta v \ll c$, and therefore it is safe to assume that the field lines are radial at $t = 0$ and also at time t in the frame of reference S. There will, in fact, be small aberration effects associated with the velocity Δv, but these are second-order compared with the gross effects we are discussing. We may therefore consider a small cone of field lines at angle θ with respect to the acceleration vector of the charge at $t = 0$ and at some later time t when the charge is moving at a constant velocity Δv (Fig. 12.1(b)). We join up electric field lines between the two cones through the thin shell of thickness $c\Delta t$, as shown in the diagram. The strength of the E_θ component of the field is given by the number of field lines per unit area in the $\mathbf{i}_\theta$ direction. From the geometry of Fig. 12.1(b), which exaggerates the discontinuities in the field lines, the ratio of the E_θ and E_r field components is given by the relative sizes of the sides of the rectangle $ABCD$, that is,

$$\frac{E_\theta}{E_r} = \frac{\Delta v\, t \sin\theta}{c\Delta t}. \tag{14.1}$$

But, according to the inverse square law,

$$E_r = \frac{q}{4\pi\epsilon_0 r^2}, \qquad \text{where} \qquad r = ct.$$

Therefore,

$$E_\theta = \frac{q(\Delta v/\Delta t)\sin\theta}{4\pi\epsilon_0 c^2 r}.$$

$\Delta v/\Delta t$ is the acceleration $\ddot{r}$ of the charge and hence

$$E_\theta = \frac{q\ddot{r}\sin\theta}{4\pi\epsilon_0 c^2 r}. \tag{14.2}$$

Notice that the radial component of the field decreases as r^{-2}, according to Coulomb's law, but the field in the pulse decreases only as r^{-1} because the field lines become more and more stretched in the E_θ direction with time, as can be seen from (14.1). Alternatively, we can write $qr = p$, where p is the dipole moment of the charge with respect to some origin, and hence

$$E_\theta = \frac{\ddot{p}\sin\theta}{4\pi\epsilon_0 c^2 r}. \tag{14.3}$$

This is a pulse of electromagnetic radiation and hence the energy flow per unit area per second at distance r is given by the Poynting vector $|\boldsymbol{E} \times \boldsymbol{H}| = E_0^2/Z_0$, where $Z_0 = (\mu_0/\epsilon_0)^{1/2}$ is the impedance of free space (see Section 7.11). The rate of loss of energy through solid angle $\mathrm{d}\Omega$ at distance r from the charge is therefore

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right)_{\text{rad}} \mathrm{d}\Omega = \frac{|\ddot{\boldsymbol{p}}|^2 \sin^2\theta}{16\pi^2 Z_0 \epsilon_0^2 c^4 r^2}\, r^2\, \mathrm{d}\Omega = \frac{|\ddot{\boldsymbol{p}}|^2 \sin^2\theta}{16\pi^2 \epsilon_0 c^3}\, \mathrm{d}\Omega. \tag{14.4}$$

To find the total radiation rate, we integrate over solid angle, that is, we integrate over θ with respect to the direction of the acceleration. Integrating over solid angle means integrating over $d\Omega = 2\pi \sin\theta \, d\theta$ and so

$$-\left(\frac{dE}{dt}\right)_{\text{rad}} = \int_0^{\pi} \frac{|\ddot{\boldsymbol{p}}|^2 \sin^2\theta}{16\pi^2 \epsilon_0 c^3} 2\pi \sin\theta \, d\theta.$$

We find the key result

$$-\left(\frac{dE}{dt}\right)_{\text{rad}} = \frac{|\ddot{\boldsymbol{p}}|^2}{6\pi\epsilon_0 c^3} = \frac{q^2|\ddot{\boldsymbol{r}}|^2}{6\pi\epsilon_0 c^3}. \tag{14.5}$$

This result is sometimes referred to as *Larmor's formula* – precisely the same result comes out of the full theory. This formula embodies the three essential properties of the radiation of an accelerated charged particle.

(a) The total radiation rate is given by Larmor's formula (14.5). Notice that, in this formula, the acceleration is the *proper acceleration* of the charged particle and that the radiation loss rate is measured in the *instantaneous rest frame* of the particle.

(b) The *polar diagram* of the radiation is of *dipolar* form, that is, the electric field strength varies as $\sin\theta$ and the power radiated per unit solid angle varies as $\sin^2\theta$, where θ is the angle with respect to the acceleration vector of the particle (Fig. 14.2). Notice that there is no radiation along the acceleration vector and the field strength is greatest at right angles to it.

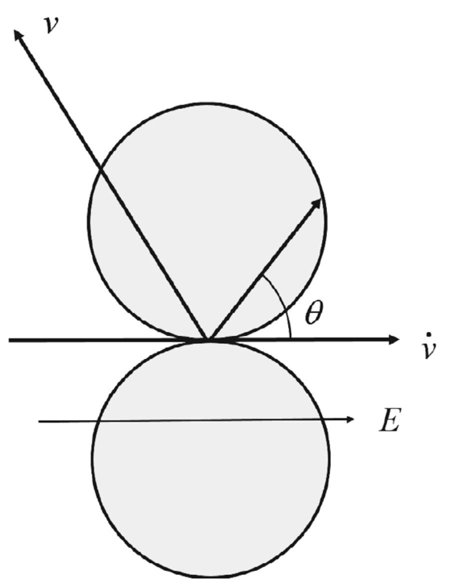


Fig. 14.2 The polar diagram of the radiation emitted by an accelerated electron. The polar diagram shows the magnitude of the electric field strength as a function of polar angle θ with respect to the instantaneous acceleration vector $\dot{\boldsymbol{v}} = \boldsymbol{a}$. The polar diagram $E \propto \sin\theta$ corresponds to circular lobes with respect to the acceleration vector. (From M.S. Longair, 2011, *High Energy Astrophysics*, 3rd edition. Cambridge: Cambridge University Press.)

(c) The radiation is *polarised* with the electric field vector lying in the direction of the acceleration vector of the particle, as projected onto a sphere at distance r from the origin (see Fig. 14.2).

In fact, most radiation problems involving accelerated charged particles can be understood in terms of these simple rules. Several important examples, including the derivation of the Thomson cross-section, are given in my book *High Energy Astrophysics*.[8] Although Thomson's is a rather pleasant argument, you should not regard it as a substitute for the full theory which comes out of strict application of Maxwell's equations.

14.2.2 Radiation Damping of an Oscillator

We now apply the result (14.5) to the case of an oscillator performing simple harmonic oscillations with amplitude x_0 at angular frequency ω_0, $x = x_0 \exp(\mathrm{i}\omega t)$. Therefore, $\ddot{x} = -\omega_0^2 x_0 \exp(\mathrm{i}\omega t)$, or, taking the real part, $\ddot{x} = -\omega_0^2 x_0 \cos \omega t$. The instantaneous rate of loss of energy by radiation is therefore

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right) = \frac{\omega_0^4 e^2 x_0^2}{6\pi\epsilon_0 c^3} \cos^2 \omega_0 t. \tag{14.6}$$

The average value of $\cos^2 \omega t$ is $\frac{1}{2}$ and hence the average rate of loss of energy in the form of electromagnetic radiation is

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right)_{\text{average}} = \frac{\omega_0^4 e^2 x_0^2}{12\pi\epsilon_0 c^3}. \tag{14.7}$$

We now introduce the equation of damped simple harmonic motion, since the oscillator loses energy by virtue of emitting electromagnetic waves. It can be written in the form

$$m\ddot{x} + a\dot{x} + kx = 0,$$

where m is the (reduced) mass, k the spring constant and $a\dot{x}$ the damping force. The energies associated with each of these terms is found by multiplying through by $\dot{x}$ and integrating with respect to time:

$$\int_0^t m\ddot{x}\dot{x}\,\mathrm{d}t + \int_0^t a\dot{x}^2\,\mathrm{d}t + \int_0^t kx\dot{x}\,\mathrm{d}t = 0,$$

$$\tfrac{1}{2}m\int_0^t \mathrm{d}(\dot{x}^2) + \int_0^t a\dot{x}^2\,\mathrm{d}t + \tfrac{1}{2}k\int_0^t \mathrm{d}(x^2) = 0. \tag{14.8}$$

We identify the terms in (14.8) with the kinetic energy, the damping energy loss and the potential energy of the oscillator respectively. Now evaluate each term for simple harmonic motion of the form $x = x_0 \cos \omega t$. The average values of the kinetic and potential energies are

$$\text{average kinetic energy} = \tfrac{1}{4}m x_0^2 \omega_0^2,$$

$$\text{average potential energy} = \tfrac{1}{4}k x_0^2.$$

If the damping is very small, the natural frequency of oscillation of the oscillator is $\omega_0^2 = k/m$. Thus, as we already know, the average kinetic and potential energies of the oscillator are equal and the total energy is the sum of the two,

$$E = \tfrac{1}{2} m x_0^2 \omega_0^2. \tag{14.9}$$

Inspection of (14.8) shows that the average rate of loss of energy by radiation from the oscillator is

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right)_{\text{rad}} = \tfrac{1}{2} a x_0^2 \omega_0^2. \tag{14.10}$$

Therefore, taking the ratio of (14.9) and (14.10), we find

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right)_{\text{rad}} = \frac{a}{m} E. \tag{14.11}$$

We can now compare this relation with the expression for the loss rate by radiation of the oscillator. Substituting (14.9) into (14.7), we obtain

$$-\frac{\mathrm{d}E}{\mathrm{d}t} = \gamma E, \tag{14.12}$$

where $\gamma = \omega_0^2 e^2/6\pi\epsilon_0 c^3 m$. This expression becomes somewhat simpler if we introduce the *classical electron radius* $r_\mathrm{e} = e^2/4\pi\epsilon_0 m_\mathrm{e} c^2$. Thus $\gamma = 2r_\mathrm{e}\omega_0^2/3c$, where m_e is the mass of the electron. We can therefore obtain the correct expression for the decay in amplitude of the oscillator by identifying γ with a/m in (14.11).

We can now appreciate why Planck believed this was a fruitful way of attacking the problem. The radiation loss does not go into heat, but into electromagnetic radiation and the constant γ depends only upon fundamental constants, if we take the oscillator to be an oscillating electron. In contrast, in frictional damping, the energy goes into heat and the loss rate formula contains constants appropriate to the structure of the material. Furthermore, in the case of electromagnetic waves, if the oscillators and waves are confined within an enclosure with perfectly reflecting walls, energy is not lost from the system and the waves react back on the oscillators, returning energy to them. This is why Planck called the damping 'conservative damping'. The more common name for this phenomenon nowadays is *radiation damping*.

14.3 The Equilibrium Radiation Spectrum of a Harmonic Oscillator

We have now obtained the expression for the dynamics of an oscillator undergoing natural damping:

$$m\ddot{x} + a\dot{x} + kx = 0,$$
$$\ddot{x} + \gamma\dot{x} + \omega_0^2 x = 0. \tag{14.13}$$

If an electromagnetic wave is incident on the oscillator, energy is transferred to it and a forcing term is added to (14.13),

$$\ddot{x} + \gamma \dot{x} + \omega_0^2 x = \frac{F}{m}. \tag{14.14}$$

The oscillator is accelerated by the E field of an incident wave and so we can write $F = eE_0 \exp(\mathrm{i}\omega t)$. To find the response of the oscillator, we adopt a trial solution for x of the form $x = x_0 \exp(\mathrm{i}\omega t)$. Then

$$x_0 = \frac{eE_0}{m\left(\omega_0^2 - \omega^2 + \mathrm{i}\gamma\omega\right)}. \tag{14.15}$$

There is a complex term in the denominator and so the oscillator does not vibrate in phase with the incident wave. This does not matter for our calculation since we are only interested in the modulus of the amplitude, as we will demonstrate.

We are already quite far along the track to finding the amount of energy transferred to the oscillator, but let us not do that. Rather, let us work out the rate of radiation of the oscillator under the influence of the incident radiation field. If we set this equal to the 'natural' radiation of the oscillator, we will have found the equilibrium spectrum. This provides a physical picture for what we mean by the oscillator being in equilibrium with the radiation field – the work done by the incident radiation field is just enough to supply the energy loss per second by the oscillator.

From (14.6), the rate of radiation of the oscillator is

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right) = \frac{\omega^4 e^2 x_0^2}{6\pi\epsilon_0 c^3}\cos^2\omega_0 t,$$

where we use the value of x_0 derived in (14.15). We need the square of the modulus of x_0 and this is found by multiplying x_0 by its complex conjugate, that is,

$$x_0^2 = \frac{e^2 E_0^2}{m^2\left[\left(\omega_0^2 - \omega^2\right)^2 + \gamma^2\omega^2\right]}.$$

Therefore, the radiation rate is

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right) = \frac{\omega^4 e^4 E_0^2}{6\pi\epsilon_0 c^3 m^2\left[\left(\omega_0^2 - \omega^2\right)^2 + \gamma^2\omega^2\right]}\cos^2\omega_0 t. \tag{14.16}$$

It is now convenient to take averages over time, that is

$$\langle\cos^2\omega t\rangle = \tfrac{1}{2}.$$

We notice that the term in E_0^2 is closely related to the energy in the incident wave. The incident energy per unit area per second is given by the Poynting vector $\boldsymbol{E} \times \boldsymbol{H}$. For electromagnetic waves, $\boldsymbol{B} = \boldsymbol{k} \times \boldsymbol{E}/\omega$, $|\boldsymbol{H}| = E/\mu_0 c$ and $|\boldsymbol{E} \times \boldsymbol{H}| = (E_0^2/\mu_0 c)\cos^2\omega t$, and so the average incident energy per unit area per second is

$$\frac{1}{2\mu_0 c}E_0^2 = \tfrac{1}{2}\epsilon_0 c E_0^2.$$

Furthermore, when we deal with a superposition of waves of random phase, as is the case for equilibrium radiation, the total incident energy is found by adding the energies of the individual waves (see Section 17.3). Therefore, we can replace the value of E^2 in our formula by the sum over all the waves of angular frequency ω incident upon the oscillator to find the total *average* reradiated power. From (14.16),

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right) = \frac{\omega^4 e^4 \frac{1}{2} \sum_i E_{0i}^2}{6\pi\epsilon_0 c^3 m^2 \left[\left(\omega_0^2 - \omega^2\right)^2 + \gamma^2\omega^2\right]}. \tag{14.17}$$

The next step is to note that the loss rate (14.17) is part of a continuum intensity distribution[a] which can be written as a sum of energies in the frequency band ω to $\omega + \mathrm{d}\omega$, that is,

$$I(\omega)\,\mathrm{d}\omega = \tfrac{1}{2}\epsilon_0 c \sum_i E_{0i}^2. \tag{14.18}$$

Therefore, the total average radiation loss rate is

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right) = \frac{\omega^4 e^4}{6\pi\epsilon_0^2 c^4 m^2} \frac{I(\omega)\,\mathrm{d}\omega}{\left[\left(\omega_0^2 - \omega^2\right)^2 + \gamma^2\omega^2\right]}.$$

Introducing the classical electron radius, $r_\mathrm{e} = e^2/4\pi\epsilon_0 m_\mathrm{e} c^2$,

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right) = \frac{8\pi r_\mathrm{e}^2}{3} \frac{\omega^4 I(\omega)\,\mathrm{d}\omega}{\left[\left(\omega_0^2 - \omega^2\right)^2 + \gamma^2\omega^2\right]}. \tag{14.19}$$

Now the response curve of the oscillator, as described by the factor in the denominator, is very sharply peaked about the value ω_0, because the radiation rate is very small in comparison with the total energy of the oscillator, that is, $\gamma \ll 1$ (Fig. 14.3). We can therefore make some simplifying approximations. If ω appears on its own, we can let $\omega \to \omega_0$ and $(\omega_0^2 - \omega^2) = (\omega_0 + \omega)(\omega_0 - \omega) \approx 2\omega_0(\omega_0 - \omega)$. Therefore,

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right) = \frac{2\pi r_\mathrm{e}^2}{3} \frac{\omega_0^2\, I(\omega)\,\mathrm{d}\omega}{\left[(\omega - \omega_0)^2 + (\gamma^2/4)\right]}. \tag{14.20}$$

Finally, we expect $I(\omega)$ to be a slowly varying function in comparison with the sharpness of the response curve of the oscillator and so we can set $I(\omega)$ equal to a constant value $I(\omega_0)$ over the range of values of ω of interest:

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right) = \frac{2\pi r_\mathrm{e}^2}{3} I(\omega_0) \int_0^\infty \frac{\omega_0^2\,\mathrm{d}\omega}{\left[(\omega - \omega_0)^2 + (\gamma^2/4)\right]}. \tag{14.21}$$

[a] Notice that this intensity is the total power per unit area from 4π steradians. The usual definition is in terms of W m^{-2} Hz^{-1} sr^{-1}. Correspondingly, in (14.23), the relation between $I(\omega)$ and $u(\omega)$ is $I(\omega) = u(\omega)c$ rather than the usual relation $I(\omega) = u(\omega)c/4\pi$.

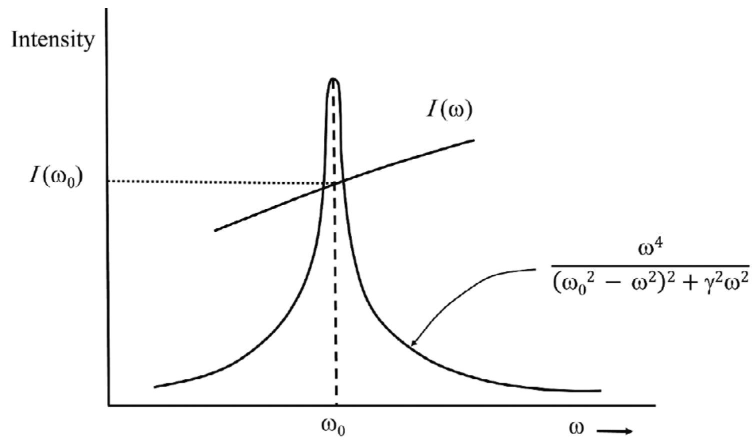


Fig. 14.3 Illustrating the response curve of the oscillator to waves of different frequencies.

The remaining integral is easy if we set the lower limit equal to minus infinity, and this is perfectly permissible since the function tends very rapidly to zero at frequencies away from that of the peak intensity. Using

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}x}{x^2 + a^2} = \frac{\pi}{a},$$

the integral (14.21) becomes $2\pi/\gamma$ and so

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right) = \frac{2\pi\omega_0^2 r_\mathrm{e}^2}{3}\frac{2\pi}{\gamma} I(\omega_0) = \frac{4\pi^2\omega_0^2 r_\mathrm{e}^2}{3\gamma} I(\omega_0). \tag{14.22}$$

According to the prescription set out at the beginning of this section, this rate of radiation should now be set equal to the spontaneous radiation rate of the oscillator. There is only one complication. We have assumed that the oscillator can respond to incident radiation arriving from any direction. For a single oscillator with axis in, say, the x-direction, there are directions of incidence in which the oscillator does not respond to the incident wave – this occurs if a component of the incident electric field is perpendicular to the dipole axis of the oscillator. We can get round this problem by a clever argument used by Richard Feynman in his analysis of this problem.[9] Suppose three oscillators are mutually at right angles to each other. Then this system can respond like a completely free oscillator and follow any incident electric field. Therefore, we obtain the correct result if we suppose that (14.22) is the radiation which would be emitted by *three mutually perpendicular oscillators*, each oscillating at frequency ω_0. Equating the radiation rates, we find the rather spectacular result:

$$I(\omega_0)\frac{4\pi^2\omega_0^2 r_\mathrm{e}^2}{3\gamma} = 3\gamma E \qquad \text{where} \qquad \gamma = \frac{2r_\mathrm{e}\omega_0^2}{3c},$$

and hence

$$I(\omega_0) = \frac{\omega_0^2}{\pi^2 c^2} E. \tag{14.23}$$

Writing (12.23) in terms of the spectral energy density $u(\omega_0)$,[b]

$$u(\omega_0) = \frac{I(\omega_0)}{c} = \frac{\omega_0^2}{\pi^2 c^3} E. \tag{14.24}$$

We now drop the subscript zero since this result applies to all frequencies in equilibrium. In terms of frequencies,

$$u(\omega)\,\mathrm{d}\omega = u(\nu)\,\mathrm{d}\nu = \frac{\omega^2}{\pi^2 c^3} E\,\mathrm{d}\omega,$$

that is,

$$u(\nu) = \frac{8\pi\nu^2}{c^3} E. \tag{14.25}$$

This is the result which Planck derived in a paper published in June 1899.[10] It is a remarkable formula. All information about the nature of the oscillator has completely disappeared from the problem. There is no mention of its charge or mass. All that remains is the average energy of the oscillator. The meaning behind the relation is obviously very profound and fundamental in a thermodynamic sense. It is an intriguing calculation – the whole analysis has proceeded through a study of the electrodynamics of oscillators, and yet the final result contains no trace of the means by which we arrived at the answer. One can imagine how excited Planck must have been when he discovered this basic result.

Even more important, if we work out the average energy of an oscillator of frequency ν in an enclosure of temperature T, we can find immediately the spectrum of black-body radiation. The next step is therefore to find the relation between E, T and ν.

14.4 Towards the Spectrum of Black-Body Radiation

The big surprise is what Planck *did not do next*. We know the mean energy of an oscillator at temperature T according to classical statistical mechanics. Classically, in thermodynamic equilibrium, each degree of freedom is allotted $\frac{1}{2}k_B T$ of energy and hence the mean energy of a harmonic oscillator is $k_B T$, because it has two degrees of freedom, those associated with the squared terms $\dot{x}^2$ and x^2 in the expression for its total energy. Setting $E = k_B T$, we find

$$u(\nu) = \frac{8\pi\nu^2}{c^3} k_B T. \tag{14.26}$$

This turns out to be the correct expression for the black-body radiation law at low frequencies, the Rayleigh–Jeans law, which will be described in Section 14.6.

Why did Planck not do this? First of all, the equipartition theorem of Maxwell and Boltzmann is a result of statistical thermodynamics and this was the point of view which he had until recently rejected. At least, he was certainly not as familiar with statistical mechanics as he was with classical thermodynamics. In addition, as described in

[b] See footnote a of this section.

Chapter 12, it was far from clear in 1899 how secure the equipartition theorem really was. Maxwell's kinetic theory could not account for the ratio of the specific heats of diatomic gases and Planck still had reservations about Boltzmann's statistical approach.

Planck had already had his fingers burned by Boltzmann when he failed to note the necessity of statistical assumptions in order to derive the equilibrium state of black-body radiation. He therefore adopted a rather different approach. To quote his words:

> I had no alternative than to tackle the problem once again – this time from the opposite side – namely from the side of thermodynamics, my own home territory where I felt myself to be on safer ground. In fact, my previous studies of the Second Law of Thermodynamics came to stand me in good stead now, for at the very outset I hit upon the idea of correlating not the temperature of the oscillator but its entropy with its energy ... While a host of outstanding physicists worked on the problem of the spectral energy distribution both from the experimental and theoretical aspect, every one of them directed his efforts solely towards exhibiting the dependence of the intensity of radiation on the temperature. On the other hand, I suspected that the fundamental connection lies in the dependence of entropy upon energy ... Nobody paid any attention to the method which I adopted and I could work out my calculations completely at my leisure, with absolute thoroughness, without fear of interference or competition.[11]

In March 1900, Planck[12] worked out the following relation for the change in entropy of a system which is not in equilibrium – we will not derive this result but it was important to Planck at the time:

$$\Delta S = \tfrac{3}{5} \frac{\partial^2 S}{\partial E^2} \, \Delta E \, \mathrm{d}E. \tag{14.27}$$

This equation applies to an oscillator with entropy which deviates from the maximum value by an amount ΔE from the equilibrium energy E. The entropy change occurs when the energy of the oscillator is changed by $\mathrm{d}E$. Thus, if ΔE and $\mathrm{d}E$ have opposite signs, so that the system tends to return towards equilibrium, then the entropy change must be positive and hence the function $\partial^2 S/\partial E^2$ must necessarily have a negative value. By inspection, we see that a negative value of $\partial^2 S/\partial E^2$ means that there is an entropy maximum and thus if ΔE and $\mathrm{d}E$ are of opposite signs, the system must approach equilibrium.

To appreciate what Planck did next, we need to review the state of experimental determinations of the spectrum of black-body radiation at this time. Wien had followed up his studies of the spectrum of thermal radiation by attempting to derive the radiation law from theory.[13] We need not go into his ideas, but he derived an expression for the radiation law which was consistent with his displacement law and which provided a good fit to all the data available in 1896. The displacement law is

$$u(\nu) = \nu^3 f(\nu/T),$$

and Wien's theory suggested that

$$u(\nu) = \frac{8\pi\alpha}{c^3} \nu^3 \mathrm{e}^{-\beta\nu/T}. \tag{14.28}$$

This is *Wien's law*, written in our notation, rather than Wien's. There are two unknown constants in the formula, α and β; the constant $8\pi/c^3$ has been included on the right-hand

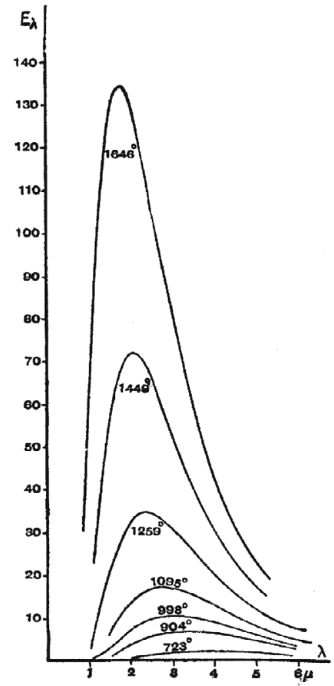


Fig. 14.4 The spectrum of black-body radiation plotted on linear scales of intensity and wavelength as determined by Lummer and Pringsheim in 1899 for temperatures between 700 and 1600 °C. (From H.S. Allen and R.S. Maxwell, 1939, *A Text-book of Heat*, Part II, p. 748, London: MacMillan & Co.)

side for reasons which will become apparent in a moment. Rayleigh's comment on Wien's paper was terse:

> Viewed from the theoretical side, the result appears to me to be little more than a conjecture.[14]

The importance of the formula was that it could account for all the experimental data available at the time and therefore could be adopted in theoretical studies. Typical black-body spectra are shown in Fig. 14.4. What is needed is a function which rises steeply to a maximum and then cuts off abruptly at short wavelengths. As might be expected, the region of the maximum was best defined by the experimental data – the uncertainties in the wings of the energy distribution were much greater.

The next step in Planck's paper was the introduction of a definition of the entropy of an oscillator:

$$S = -\frac{E}{\beta\nu} \ln \frac{E}{\alpha\nu \mathrm{e}}. \tag{14.29}$$

S is its entropy, E its energy and α and β are constants; e is the base of natural logarithms. In fact, we know from his writings that he derived this formula by working backwards from Wien's law. Let us show how this can be done.

Wien's law (14.28) can be inserted into (14.25) relating $u(\nu)$ and E:

$$u(\nu) = \frac{8\pi\nu^2}{c^3}E \quad \text{and so} \quad E = \alpha\nu\,\mathrm{e}^{-\beta\nu/T}. \tag{14.30}$$

Now suppose the oscillator (or a set of oscillators) is placed in a fixed volume V. The thermodynamic equality (11.55) states that

$$T\,\mathrm{d}S = \mathrm{d}U + p\,\mathrm{d}V,$$

and hence

$$\left(\frac{\partial S}{\partial U}\right)_V = \frac{1}{T}. \tag{14.31}$$

U and S are additive functions of state and hence (14.31) refers to the properties of an individual oscillator, as well as to an ensemble of them. Inverting (14.30), we can relate $(\partial S/\partial E)_V$ to E, that is,

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_V = -\frac{1}{\beta\nu}\ln\left(\frac{E}{\alpha\nu}\right). \tag{14.32}$$

Now adopting Planck's definition of the entropy of the oscillator (14.29) and differentiating with respect to E,

$$\begin{aligned}\frac{\mathrm{d}S}{\mathrm{d}E} &= -\frac{1}{\beta\nu}\ln\left(\frac{E}{\alpha\nu\mathrm{e}}\right) - \frac{E}{\beta\nu}\frac{1}{E} = -\frac{1}{\beta\nu}\left[\ln\left(\frac{E}{\alpha\nu\mathrm{e}}\right) + 1\right] \\ &= -\frac{1}{\beta\nu}\left[\ln\left(\frac{E}{\alpha\nu}\right)\right].\end{aligned} \tag{14.33}$$

The identity of (14.32) and (14.33) demonstrates how Wien's law combined with (14.25) leads to Planck's definition of entropy (14.29).

Now we arrive at the step which was crucial for Planck. Take the next derivative of (14.33) with respect to E:

$$\frac{\mathrm{d}^2 S}{\mathrm{d}E^2} = -\frac{1}{\beta\nu}\frac{1}{E}. \tag{14.34}$$

β, ν and E are all necessarily positive quantities and therefore $\partial^2 S/\partial E^2$ is necessarily negative. Hence, Wien's law (14.28) is entirely consistent with the second law of thermodynamics. Notice the remarkable simplicity of the expression for the second derivative of the entropy with respect to energy – it is proportional to the inverse of the energy. This profoundly impressed Planck who remarked:

> I have repeatedly tried to change or generalise the equation for the electromagnetic entropy of an oscillator in such a way that it satisfies all theoretically sound electromagnetic and thermodynamic laws but I was unsuccessful in this endeavour.[15]

In his paper presented to the Prussian Academy of Sciences in May 1899 he stated:

> I believe that this must lead me to conclude that the definition of radiation entropy and therefore Wien's energy distribution law necessarily result from the application of the principle of the increase of entropy to the electromagnetic radiation theory and therefore

> the limits of validity of this law, in so far as they exist at all, coincide with those of the second law of thermodynamics.[16]

This is a dramatic conclusion and we now know that it is not correct – he had, however, made significant progress towards the correct formulation of the theory.

14.5 The Primitive Form of Planck's Radiation Law

These calculations were presented to the Prussian Academy of Sciences in June 1900. By October 1900, Rubens and Kurlbaum had shown beyond any doubt that Wien's law was inadequate to explain the spectrum of black-body radiation at low frequencies and high temperatures. Their experiments were carried out with the greatest care over a wide range of temperatures. They showed that, at low frequencies and high temperatures, the intensity of radiation was proportional to temperature. This is clearly inconsistent with Wien's law because, if $u(\nu) \propto \nu^3\,\mathrm{e}^{-\beta\nu/T}$, then for $\beta\nu/T \ll 1$, $u(\nu) \propto \nu^3$ and is independent of temperature. Wien's displacement law requires the form of the spectrum to depend only upon ν/T and so the functional dependence must be more complex than (14.28) for small values of ν/T.

Rubens and Kurlbaum showed Planck their new results before they were presented in October 1900 and he was given the opportunity to make some remarks about their implications. The result was his paper entitled *An Improvement of the Wien Distribution*[17] – it was the first time that the Planck formula appeared in its primitive form. Here is what Planck did.

He now knew that he had to find a law which would result in the relation $u \propto T$ in the limit $\nu/T \to 0$. Let us run through the relations he had already derived. Planck knew from (14.25) that

$$u(\nu) = \frac{8\pi\nu^2}{c^3}E.$$

At low frequencies, the experiments of Rubens and Kurlbaum showed that $u(\nu) \propto T$. Hence,

$$E \propto T, \qquad \frac{\mathrm{d}S}{\mathrm{d}E} = \frac{1}{T},$$

and so

$$\frac{\mathrm{d}S}{\mathrm{d}E} \propto \frac{1}{E}, \quad \frac{\mathrm{d}^2S}{\mathrm{d}E^2} \propto \frac{1}{E^2}. \tag{14.35}$$

Therefore, $\mathrm{d}^2S/\mathrm{d}E^2$ must change its functional dependence upon E between large and small values of ν/T. Wien's law remains good for large values of ν/T and leads to

$$\frac{\mathrm{d}^2S}{\mathrm{d}E^2} \propto \frac{1}{E}. \tag{14.36}$$

The standard technique for combining functions of the form (14.35) and (14.36) is to try an expression of the form

$$\frac{d^2S}{dE^2} = -\frac{a}{E(b+E)}, \tag{14.37}$$

which has exactly the required properties for large and small values of E, $E \gg b$ and $E \ll b$ respectively. The rest of the analysis is straightforward. Integrating,

$$\frac{dS}{dE} = -\int \frac{a}{E(b+E)}\, dE = -\frac{a}{b}[\ln E - \ln(b+E)]. \tag{14.38}$$

But $dS/dE = 1/T$ and hence

$$\frac{1}{T} = -\frac{a}{b}\ln\left(\frac{E}{b+E}\right), \qquad e^{b/aT} = \frac{b+E}{E},$$

$$E = \frac{b}{e^{b/aT}-1}. \tag{14.39}$$

Now we can find the radiation spectrum:

$$u(\nu) = \frac{8\pi\nu^2}{c^3}E = \frac{8\pi\nu^2}{c^3}\frac{b}{e^{b/aT}-1}. \tag{14.40}$$

From the high frequency, low temperature limit, we can compare the constants with those which appear in Wien's law:

$$u(\nu) = \frac{8\pi\nu^2 b}{c^3 e^{b/aT}} = \frac{8\pi\alpha}{c^3}\frac{\nu^3}{e^{\beta\nu/T}}.$$

Thus b must be proportional to the frequency ν. We can therefore write Planck's formula in its primitive form:

$$u(\nu) = \frac{A\nu^3}{(e^{\beta\nu/T}-1)}. \tag{14.41}$$

Notice that this formula satisfies Wien's displacement law.

Of equal importance is that Planck was also able to find an expression for the entropy of an oscillator by integrating dS/dE. Integrating (14.41),

$$\frac{dS}{dE} = -\frac{a}{b}[\ln E - \ln(b+E)] = -\frac{a}{b}\left[\ln\frac{E}{b} - \ln\left(1+\frac{E}{b}\right)\right].$$

Integrating again,

$$S = -\frac{a}{b}\left\{\left(E\ln\frac{E}{b} - E\right) - \left[E\ln\left(1+\frac{E}{b}\right) - E + b\ln\left(1+\frac{E}{b}\right)\right]\right\},$$

$$= -a\left[\frac{E}{b}\ln\frac{E}{b} - \left(1+\frac{E}{b}\right)\ln\left(1+\frac{E}{b}\right)\right], \tag{14.42}$$

with $b \propto \nu$.

Planck's new formula was remarkably elegant and could now be confronted with the experimental evidence. Before doing that, let us study the origins of another radiation formula which had been proposed at about the same time, the *Rayleigh–Jeans law*.

14.6 Rayleigh and the Spectrum of Black-Body Radiation

Lord Rayleigh was the author of the famous text *The Theory of Sound*[18] and a leading exponent of the theory of waves in general. His interest in the theory of black-body radiation was stimulated by the inadequacies of Wien's law in accounting for the low frequency behaviour of black-body radiation as a function of temperature. His original paper is a short and elegant exposition of the application of the theory of waves to black-body radiation and is reproduced in full as an appendix to this chapter.[19] The reason for the inclusion of Jeans' name in what became known as the *Rayleigh–Jeans law* is that there was a numerical error in Rayleigh's analysis which was corrected by Jeans in a paper published in *Nature* in 1906. Let us look first at the structure of the paper.

The first paragraph is a description of the state of knowledge of the black-body spectrum in 1900. In the second, Rayleigh acknowledges the success of Wien's formula in accounting for the maximum in the black-body spectrum, but then expresses his worries about the long wavelength behaviour of the formula. He sets out his proposal in the third and fourth paragraphs and in the fifth describes his preferred form for the radiation spectrum, which, in the sixth paragraph, he hopes will be compared with experiment by 'the distinguished experimenters who have been occupied with this subject'.

Let us rework the essence of paragraphs three, four and five in modern notation with the correct numerical factors. We begin with the problem of waves in a box. Suppose the box is a cube with sides of length L. Inside the box, all possible wave modes consistent with the boundary conditions are allowed to come into thermodynamic equilibrium at temperature T. The general wave equation for the waves in the box is

$$\nabla^2\psi = \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} = \frac{1}{c_s^2}\frac{\partial^2\psi}{\partial t^2}, \tag{14.43}$$

where c_s is the speed of the waves. The walls are fixed and so the waves must have zero amplitude $\psi = 0$ at $x, y, z = 0$ and $x, y, z = L$. The solution of this problem is well known:

$$\psi = C\mathrm{e}^{-\mathrm{i}\omega t}\, \sin\frac{l\pi x}{L}\, \sin\frac{l\pi y}{L}\, \sin\frac{l\pi z}{L}, \tag{14.44}$$

corresponding to standing waves which fit perfectly into the box in three dimensions, provided l, m and n are integers. Each combination of l, m and n is called a *mode of oscillation* of the waves in the box. What this means in physical terms is that, for any mode, the medium within the box through which the waves propagate oscillates in phase at a particular frequency. These modes are also *orthogonal* and so represent independent ways in which the medium can oscillate. In addition, the set of orthogonal modes with $0 \le l, m, n \le \infty$ is a complete set and so any pressure distribution in the medium can be described as a sum over this complete set of the orthogonal functions.

We now substitute (14.44) into the wave equation (14.43) and find the relation between the values of l, m, n and the angular frequency of the waves, ω:

$$\frac{\omega^2}{c^2} = \frac{\pi^2}{L^2}(l^2 + m^2 + n^2). \tag{14.45}$$

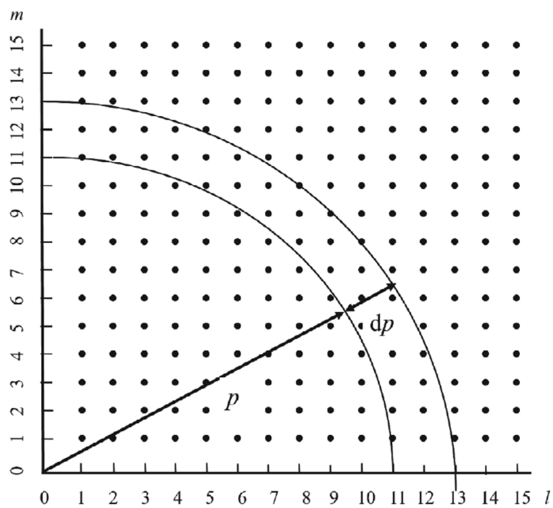


Fig. 14.5 Illustrating how the number of modes in the interval dl, dm, dn can be replaced by an increment in the phase space volume $\frac{1}{2}\pi p^2$ dp where $p^2 = l^2 + m^2 + n^2$. The diagram shows a two-dimensional cross-section through the three-dimensional volume in l, m and n.

Writing $p^2 = l^2 + m^2 + n^2$,

$$\frac{\omega^2}{c^2} = \frac{\pi^2 p^2}{L^2}. \tag{14.46}$$

We thus find a relation between the modes as parameterised by $p^2 = l^2 + m^2 + n^2$ and their angular frequency ω. According to the Maxwell–Boltzmann equipartition theorem, energy is shared equally among each independent mode and therefore we need to know how many modes there are in the range p to $p+\mathrm{d}p$. We find this by drawing a three-dimensional lattice in l, m, n space and evaluating the number of modes in the octant of a sphere illustrated in Fig. 14.5. If p is large, the number of modes is given by an octant of the spherical shell of radius p and width dp,

$$n(p)\,\mathrm{d}p = \tfrac{1}{8}4\pi p^2\,\mathrm{d}p. \tag{14.47}$$

Expressing this result in terms of ω rather than p,

$$p = \frac{L\omega}{\pi c}, \quad \mathrm{d}p = \frac{L}{\pi c}\,\mathrm{d}\omega, \tag{14.48}$$

and hence

$$n(p)\,\mathrm{d}p = \frac{L^3\omega^2}{2\pi^2 c^3}\,\mathrm{d}\omega. \tag{14.49}$$

In the present case, the waves are electromagnetic waves and so there are two independent linear polarisations for any given mode. Therefore, there are twice as many modes as given by (14.49).

All we have to do now is to apply the Maxwell–Boltzmann doctrine of the equipartition of energy and give each mode of oscillation an energy $E = k_B T$ according to the classical prescription. Then, the energy density of electromagnetic radiation in the box is

$$u(\nu)\,\mathrm{d}\nu\,L^3 = En(p)\,\mathrm{d}p = \frac{L^3\omega^2 E}{\pi^2 c^3}\,\mathrm{d}\omega, \tag{14.50}$$

that is,

$$u(\nu) = \frac{8\pi\nu^2}{c^3}E = \frac{8\pi\nu^2}{c^3}k_B T. \tag{14.51}$$

This is exactly the same result as that derived by Planck from electrodynamics (14.25) but, unlike Planck, Rayleigh did not hesitate in setting $E = k_B T$, according to the principle of equipartition of energy. It is intriguing that two such different approaches should result in exactly the same answer.

A number of features of Rayleigh's analysis are worth noting.

(i) Rayleigh deals directly with the electromagnetic waves themselves, rather than with the oscillators which are the source of the waves and which are in equilibrium with them.

(ii) Central to the result is the doctrine of the equipartition of energy. I had great difficulty with the principle when I was first taught it as a student. It is only when one realises that one is dealing with independent, or orthogonal, modes of oscillation that the idea becomes clear physically. We can decompose any motion of the gas in the box into its normal modes of oscillation and these modes will be independent of one another. However, they *cannot be completely independent* or there would be no way of exchanging energy between them so that they can come into equipartition. There must be processes by which energy can be exchanged between the so-called independent modes of oscillation. In fact what happens is that the modes are very weakly coupled together when we study higher order processes in the gas. Thus, if one mode gets more energy than another, there are physical processes by which energy can be redistributed among the modes. What the Maxwell–Boltzmann doctrine states is that, if you leave the system long enough, irregularities in the energy distribution among the oscillations are smoothed out by these energy interchange mechanisms. In many natural phenomena, the equilibrium distributions are set up very quickly and so there is no need to worry, but it is important to remember that there is a definite assumption about the ability of the interaction processes to bring about the equipartition distribution of energy – in practice, this is not always the case.

(iii) Rayleigh was aware of 'the difficulties which attend the Boltzmann–Maxwell doctrine of the partition of energy'. Among these is the fact that (14.51) must break down at high frequencies because the spectrum of black-body radiation does not increase as ν^2 to infinite frequency. He proposes, however, that the analysis 'may apply to the graver modes', that is, to long wavelengths or low frequencies.

(iv) Then, right out of the blue, we read in the fifth paragraph, 'If we introduce the exponential factor, the complete expression is', in our notation,

$$u(\nu) = \frac{8\pi\nu^2}{c^3} k_B T\, e^{-\beta\nu/T} = \frac{8\pi\nu^3}{c^3} k_B \frac{T}{\nu} e^{-\beta\nu/T}. \tag{14.52}$$

Rayleigh included this factor empirically so that the radiation spectrum would converge at high frequencies and because the exponential in Wien's law provided a good fit to the data. Notice that the last equality in (14.52) shows that the formula is consistent with Wien's displacement law (13.69).

Rayleigh's analysis is brilliant but, as we will see, it did not get its due credit in 1900.

14.7 Comparison of the Laws for Black-Body Radiation with Experiment

On 25 October 1900, Rubens and Kurlbaum compared their new precise measurements of the black-body spectrum with five different predictions. These were (i) Planck's formula (14.41), (ii) Wien's relation (14.28), (iii) Rayleigh's result (14.52) and (iv) and (v) two empirical relations which had been proposed by Thiesen and by Lummer and Jahnke (Fig. 14.6). Rubens and Kurlbaum concluded that Planck's formula was superior to all the others and that it gave precise agreement with experiment. Rayleigh's proposal was found to be a poor representation of the experimental data. Rayleigh was justifiably upset by the tone of voice in which they discussed his result. When his scientific papers were republished two years later he remarked on his important conclusion that the intensity of

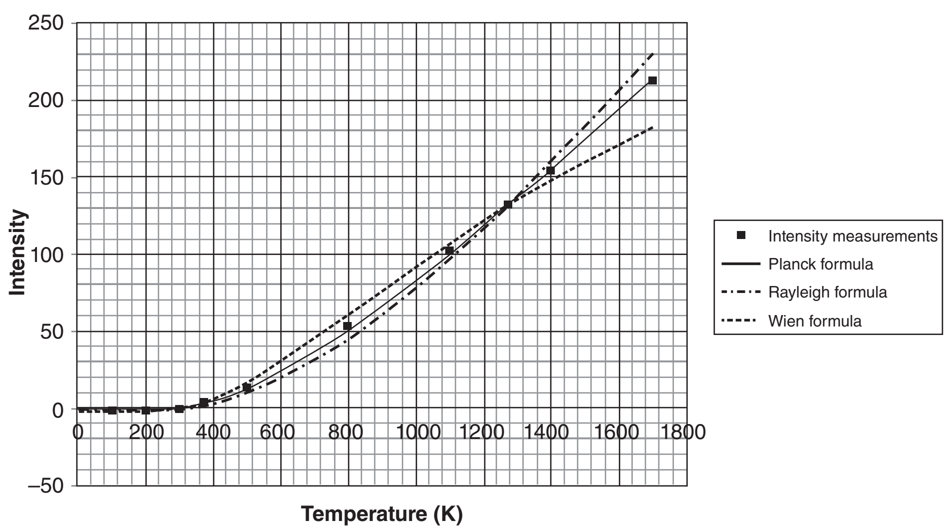


Fig. 14.6 Comparison of the radiation formulae of Planck (14.41, solid line), Wien (14.28, dotted line) and Rayleigh (14.52, dot-dashed line) with the intensity of black-body radiation at 8.85 μm as a function of temperature as measured by Rubens and Kurlbaum (filled boxes). Similar experiments were carried out at longer wavelengths, (24 + 31.6) μm and 51.2 μm, at which there was little difference between the predictions of the Planck and Rayleigh functions. (After H. Rubens and F. Kurlbaum, 1901, *Annalen der Physik*, **4**, 649–666.)

radiation should be proportional to temperature at low frequencies (see the final footnote in the Appendix). He pointed out,

> This is what I intended to emphasise. Very shortly afterwards the anticipation above expressed was confirmed by the important researches of Rubens and Kurlbaum who operated with exceptionally long waves.[19]

The essential theoretical point of Rayleigh's paper that the radiation spectrum can be described accurately by the Rayleigh–Jeans formula in the wavelength region in which it is applicable had been missed by the experimenters, who had simply compared the formulae with their experimental measurements.

It seems inconceivable that Planck was not aware of Rayleigh's work. Rubens had told Planck about his experiments and he must have seen the comparison of the various curves with the experiments. We can only assume that it was the statistical basis of the equipartition theorem which put Planck off, together with the fact that Rayleigh's law did not give a good account of the experimental data. Even so, he ought to have been impressed by the ease with which Rayleigh found the correct low frequency, high temperature relation and also by the method used to obtain a result identical to Planck's relation between the energy density of radiation and the mean energy of each mode (14.51). It has to be said that Rayleigh's paper does not exactly make this point clear.

Planck had not, however, explained anything. All he had was a formula and it had no sound theoretical basis. He immediately embarked upon this problem. The formula (14.41) was presented to the German Physical Society on 19 October 1900 and on 14 December 1900, he presented another paper entitled *On the theory of the Energy Distribution law in the Normal Spectrum*.[20] In his memoirs, he wrote,

> After a few weeks of the most strenuous work of my life, the darkness lifted and an unexpected vista began to appear.[21]

Appendix to Chapter 14: Rayleigh's Paper of 1900

REMARKS UPON THE LAW OF COMPLETE RADIATION.[19]

[*Philosophical Magazine*, XLIX. pp. 539, 540, 1900.]

By complete radiation I mean the radiation from an ideally black-body, which according to Stewart[a] and Kirchhoff is a definite function of the absolute temperature θ and the wavelength λ. Arguments of (in my opinion[b]) considerable weight have been brought forward by Boltzmann and W. Wien leading to the conclusion that the function is of the form

$$\theta^5\,\phi(\theta\lambda)\,d\lambda, \tag{1}$$

[a] Stewart's work appears to be insufficiently recognized upon the Continent. [See *Phil. Mag.* I. p. 98, 1901; p. 494 below.]

[b] *Phil. Mag.* Vol. XLV. p. 522 (1898).

expressive of the energy in that part of the spectrum which lies between λ and $\lambda + \mathrm{d}\lambda$. A further specialization by determining the form of the function ϕ was attempted later[c]. Wien concludes that the actual law is

$$c_1 \lambda^{-5} \, \mathrm{e}^{-c_2/\lambda\theta} \, \mathrm{d}\lambda, \tag{2}$$

in which c_1 and c_2 are constants, but viewed from the theoretical side the result appears to me to be little more than a conjecture. It is, however, supported upon general thermodynamic grounds by Planck[d].

Upon the experimental side, Wien's law (2) has met with important confirmation. Paschen finds that his observations are well represented, if he takes

$$c_2 = 14,455,$$

θ being measured in centigrade degrees and λ in thousandths of a millimetre (μ). Nevertheless, the law seems rather difficult of acceptance, especially the implication that as the temperature is raised, the radiation of given wavelength approaches a limit. It is true that for visible rays the limit is out of range. But if we take $\lambda = 60\mu$, as (according to the remarkable researches of Rubens) for the rays selected by reflexion at surfaces of Sylvin, we see that for temperatures over 1000° (absolute) there would be but little further increase of radiation.

The question is one to be settled by experiment; but in the meantime I venture to suggest a modification of (2), which appears to me more probable *à priori*. Speculation upon this subject is hampered by the difficulties which attend the Boltzmann-Maxwell doctrine of the partition of energy. According to this doctrine every mode of vibration should be alike favoured; and although for some reason not yet explained the doctrine fails in general, it seems possible that it may apply to the graver modes. Let us consider in illustration the case of a stretched string vibrating transversely. According to the Boltzmann-Maxwell law, the energy should be equally divided among all the modes, whose frequencies are as $1, 2, 3, \ldots$ Hence if k be the reciprocal of λ, representing the frequency, the energy between the limits k and $k + \mathrm{d}k$ is (when k is large enough) represented by $\mathrm{d}k$ simply.

When we pass from one dimension to three dimensions, and consider for example the vibrations of a cubical mass of air, we have (*Theory of Sound*, §267) as the equation for k^2,

$$k^2 = p^2 + q^2 + r^2$$

where p, q, r are integers representing the number of subdivisions in the three directions. If we regard p, q, r as the coordinates of points forming a cubic array, k is the distance of any point from the origin. Accordingly the number of points for which k lies between k and $k + \mathrm{d}k$, proportional to the volume of the corresponding spherical shell, may be represented by $k^2 \, \mathrm{d}k$, and this expresses the distribution of energy according to the Boltzmann-Maxwell law, so far as regards the wave-length or frequency. If we apply this result to radiation, we will have, since the energy in each mode is proportional to θ,

$$\theta k^2 \, \mathrm{d}k \tag{3}$$

[c] *Wied. Ann.* Vol. LVIII. p. 662 (1896).
[d] *Wied. Ann.* Vol. I. p. 74 (1900).

or, if we prefer it,

$$\theta\lambda^{-4}\,\mathrm{d}\lambda. \tag{4}$$

It may he regarded as some confirmation of the suitability of (4) that it is of the prescribed form (1).

The suggestion is that (4) rather than, as according to (2),

$$\lambda^{-5}\,\mathrm{d}\lambda \tag{5}$$

may be the proper form when $\lambda\theta$ is great[e]. If we introduce the exponential factor, the complete expression will be

$$c_1\theta\lambda^{-4}\,\mathrm{e}^{-c_2/\lambda\theta}\,\mathrm{d}\lambda. \tag{6}$$

If, as is probably to be preferred, we make k the independent variable, (6) becomes

$$c_1\theta k^2\,\mathrm{e}^{-c_2 k/\theta}\,\mathrm{d}k. \tag{7}$$

Whether (6) represents the facts of observation as well as (2) I am not in a position to say. It is to be hoped that the question may soon receive an answer at the hands of the distinguished experimenters who have been occupied with this subject.

Notes

1 Planck, M. (1950). *Scientific Autobiography and Other Papers*, p. 15. London: Williams and Norgate.
2 Planck, M. (1950). *op. cit.*, pp. 18–19.
3 Planck, M. (1896). Quoted by M.J. Klein (1977) in *History of Twentieth Century Physics*, Proceedings of the International School of Physics 'Enrico Fermi', Course 57, p. 3. New York & London: Academic Press.
4 Kuhn, T.S. (1978). *Black-Body Theory and the Quantum Discontinuity, 1894–1912*. Oxford: Clarendon Press.
5 See the remarks by Douglas Gough in Section 1.2, paragraph (ii).
6 Thomson, J.J. (1906). *Electricity and Matter*. London: Archibald Constable and Co.
7 Thomson, J.J. (1907). *Conduction of Electricity through Gases*. Cambridge: Cambridge University Press.
8 Longair, M.S. (2011). *High Energy Astrophysics*, 3rd edition. Cambridge: Cambridge University Press.
9 Feynman, R.P. (1963). *The Feynman Lectures on Physics*, Vol. 1, eds R.P. Feynman, R.B. Leighton and M. Sands, pp. 41–45. Redwood City, California: Addison-Wesley Publishing Co.
10 Planck, M. (1899). *Berl. Ber.*, 440–486.
11 Planck, M. (1950). *op. cit.*, pp. 37–38.

[e] [1902. This is what I intended to emphasize. Very shortly afterwards the anticipation above expressed was confirmed by the important researches of Rubens and Kurlbaum (*Drude Ann.* IV. p. 649, 1901), who operated with exceptionally long waves. The formula of Planck, given about the same time, seems best to fit the observations. According to this modification of Wien's formula, $\mathrm{e}^{-c_2/\lambda\theta}$ is replaced by $1/(\mathrm{e}^{c_2/\lambda\theta} - 1)$. When $\lambda\theta$ is great, this becomes $\lambda\theta/c_2$, and the complete expression reduces to (4).]

12 Planck, M. (1900). *Annalen der Physik*, **1**, 719–737.
13 Wien, W. (1896). *Annalen der Physik*, **581**, 662–669.
14 Rayleigh, Lord (1900). *Philosophical Magazine*, **49**, 539. See also *Scientific Papers by John William Strutt, Baron Rayleigh. Volume 4, 1892–1901*, p. 483. Cambridge: Cambridge University Press.
15 Planck, M. (1958). *Physikalische Abhandlungen und Vortrage (Collected Scientific Papers)*, Vol. 1, p. 596. Braunschweig: Friedr. Vieweg und Sohn. (English translation: see Hermann, A. (1971). *The Genesis of Quantum Theory (1899–1913)*, p. 10. Cambridge, Massachusetts: MIT Press.)
16 Planck, M. (1958). *op. cit.*, Vol. 1, p. 597 (English translation: Hermann, A. (1971). *op. cit.*, p. 10.)
17 Planck, M. (1900). *Verhandlungen der Deutschen Physikalischen Gessellschaft*, **2**, 202. See also *Collected Scientific Works* (1958), *op. cit.*, Vol. 1, p. 698. (English Translation: *Planck's Original Papers in Quantum Physics* (1972), annotated by H. Kangro, pp. 38–45. London: Taylor and Francis.)
18 Rayleigh, Lord (1894). *The Theory of Sound*, 2 volumes. London: MacMillan.
19 Rayleigh, Lord (1902). *Scientific Papers by John William Strutt, Baron Rayleigh. Volume 4, 1892–1901*, p. 483. Cambridge: Cambridge University Press. This version of Rayleigh's paper includes his commentary of 1902 on its reception as footnote e.
20 Planck, M. (1900). *Verhandlungen der Deutschen Physikalischen Gessellschaft*, **2**, 237–245. See also *Collected Scientific Works* (1958), *op. cit.*, Vol. 1, p. 698–706. (English Translation: *Planck's Original Papers in Quantum Physics* (1972), *op. cit.*, pp. 38–45.)
21 Planck. M. (1925). *A Survey of Physics*, p. 166. London: Methuen and Co.

15 Planck's Theory of Black-Body Radiation

15.1 Introduction

> On the very day when I formulated this law, I began to devote myself to the task of investing it with a true physical meaning. This quest automatically led me to study the interrelation of entropy and probability – in other words, to pursue the line of thought inaugurated by Boltzmann.[1]

Planck recognised that the way forward involved adopting a point of view which he had rejected in essentially all his previous work. As is apparent from his words at the end of Chapter 14, Planck was working at white-hot intensity. He was not a specialist in statistical physics and it will turn out that his analysis did not, in fact, follow the precepts of classical statistical mechanics as expounded by Boltzmann. Despite the flaws in his argument, he discovered the essential role which *quantisation* plays in accounting for (14.41) for the spectrum of black-body radiation – we recall that Planck's derivation of October 1900 was essentially a thermodynamic argument.

We have already discussed Boltzmann's expression for the relation between entropy and probability, $S \propto \ln W$, in Sections 12.7 and 12.8. The constant of proportionality was not known at the time and so we write $S = C \ln W$, where C is an unknown universal constant. First of all, let us describe how Planck should have proceeded according to classical statistical mechanics.

15.2 Boltzmann's Procedure in Statistical Mechanics

Planck knew that he had to work out the average energy $\overline{E}$ of an oscillator in thermal equilibrium in an enclosure at temperature T. Suppose there are N oscillators in such an enclosure so that their total energy is $E = N\overline{E}$. The entropy is also an additive function, and so, if $\overline{S}$ is the average entropy of an oscillator, the entropy of the whole system is $S = N\overline{S}$.

One of the tricks of classical statistical mechanics is to begin by considering the molecules to have discrete energies, $0, \epsilon, 2\epsilon, 3\epsilon, \ldots$ This procedure is adopted so that exact probabilities can be determined statistically. At the appropriate step in the argument, the value of ϵ is allowed to become infinitesimally small, whilst the total energy of the system remains finite. Then, the probability distribution changes from a discrete to a continuous distribution.

Table 15.1 Planck's example of one possible distribution of 100 energy elements among 10 oscillators

Oscillator number i	1	2	3	4	5	6	7	8	9	10	
Energy E_i in units of ϵ	7	38	11	0	9	2	20	4	4	5	Total energy $\sum_i E_i = 100$

This is exactly how Planck's paper of December 1900 begins.[2] Suppose there exists an energy unit ϵ and a fixed amount of energy E is to be distributed among N oscillators. There are many different ways in which the $r = E/\epsilon$ energy elements can be distributed over the N oscillators. The example presented by Planck in his paper is shown in Table 15.1 – he supposed that there were $N = 10$ oscillators and $r = 100$ energy elements to be distributed among them.

There are obviously many different ways of distributing the energy among the oscillators besides this one. If Planck had followed Boltzmann's prescription, here is how he should have proceeded.

Boltzmann noted that each such distribution of energies can be represented by a set of numbers $w_0, w_1, w_2, w_3, \ldots$, which describe the number of molecules or oscillators with energies $0, 1, 2, 3 \ldots$ in units of ϵ so that $\sum_i w_i = N$. Therefore, we should work out the number of ways in which the energy elements can be distributed over the oscillators and result in the same distribution of energies, $w_0, w_1, w_2, w_3, \ldots$

Let us revise some basic elements of permutation theory. The number of different ways in which n different objects can be ordered is $n!$. For example, for three objects a, b, c, these can be ordered in the following ways: $abc, acb, bac, bca, cab, cba$, that is, in $3! = 6$ ways. If m of these are identical, the number of different orderings is reduced, because we can interchange these m objects and it makes no difference to the distribution. Since the m objects can be ordered in $m!$ ways, the number of different arrangements is reduced to $n!/m!$. Thus, if $a \equiv b$ in the above example, the possible arrangements are aac, aca, caa, that is, $3!/2! = 3$ different arrangements. If a further l objects are identical, the number of different arrangements is further reduced to $n!/m!\,l!$, and so on.

Now we ask, in how many different ways can we select x from n objects? We divide the set of n objects into two groups of identical objects, the x objects which make up the set selected and the $(n - x)$ objects which are not selected. From the last paragraph, the number of different ways of making this selection is $n!/(n - x)!\,x!$. This is often written $\binom{n}{x}$ and will be recognised as the set of coefficients in the binomial expansion of $(1 + t)^n$, that is,

$$
\begin{aligned}
(1 + t)^n &= 1 + nt + \frac{n(n-1)}{2!}t^2 + \cdots + \frac{n!}{(n-x)!\,x!}t^x + \cdots + t^n \\
&= \sum_{x=0}^{n} \binom{n}{x} t^x. \qquad (15.1)
\end{aligned}
$$

Let us now return to our aim of working out the number of different ways in which we can obtain the particular distribution of oscillator energies $w_0, w_1, w_2, \ldots w_r$ over the N

oscillators, recalling that $N = \sum_{j=0}^{r} w_j$. First of all, we select w_0 oscillators from the N, which can be done in $\binom{N}{w_0}$ ways. This leaves $(N - w_0)$ from which we can select w_1 in $\binom{N-w_0}{w_1}$ different ways. Next, we select w_2 from the remaining $(N - w_0 - w_1)$ in $\binom{N-w_0-w_1}{w_2}$ ways, and so on until we have accounted for all N oscillators. Therefore, the total number of different ways of arriving at the particular distribution $w_0, w_1, w_2, \ldots w_r$ is the product of all these numbers, which is

$$\begin{aligned} W_i(w_0, w_1, w_2, \ldots w_r) &= \binom{N}{w_0}\binom{N-w_0}{w_1}\binom{N-w_0-w_1}{w_2}\cdots\binom{N-w_0-w_1-\cdots w_{r-1}}{w_r} \\ &= \frac{N!}{w_0!\, w_1!\, w_2! \cdots w_r!}. \end{aligned} \tag{15.2}$$

We have used the fact that $N = \sum_{j=0}^{r} w_j$ and $0! = 1$. It is clear from the discussion of the last two paragraphs that $W_i(w_0, w_1, w_2, \ldots w_r)$ is just the number of different ways of selecting the set of $\{w_0, w_1, w_2, \ldots w_r\}$ objects from a total of N.

According to the *principle of equal equilibrium probabilities*, each distribution of the energy elements over the system is equally likely, and so the probability of finding a particular energy distribution is

$$p_i = \frac{W_i(w_0, w_1, w_2, \ldots w_N)}{\sum_i W_i(w_0, w_1, w_2, \ldots w_N)}. \tag{15.3}$$

According to the Boltzmann procedure, the equilibrium state is that which has the greatest value of p_i. This is equivalent to maximising the entropy, if it is defined as

$$S = C \ln p, \tag{15.4}$$

that is, the state of maximum W_i also corresponds to the state of maximum entropy (see Section 12.8).

We have the choice of continuing to work with the probabilities p_i, or the density of states W_i, which was introduced in Section 12.9. These only differ by an unimportant (very large) constant. For consistency with modern usage, let us work in terms of the density of states. Therefore, using $S = C \ln W_i$ and taking the logarithm of (15.2),

$$\ln W_i = \ln r! - \sum_{j=0}^{j=N} \ln w_j! \,. \tag{15.5}$$

Using Stirling's formula,

$$n! \approx (2\pi n)^{1/2} \left(\frac{n}{\mathrm{e}}\right)^n, \qquad \ln n! \approx n \ln n - n, \tag{15.6}$$

if n is very large indeed. Substituting into (15.5),

$$\begin{aligned} \ln W_i &\approx r \ln r - r - \sum_j w_j \ln w_j + \sum_j w_j \\ &= r \ln r - r + N - \sum_j w_j \ln w_j, \end{aligned} \tag{15.7}$$

where we have used the fact that $N = \sum_j w_j$.

To find the state of maximum entropy, now maximise $\ln W_i$ subject to the constraints:

$$\begin{cases} \text{number of oscillators,} & N = \sum_j w_j = \text{constant,} \\ \text{total energy of oscillators,} & E = \sum_j \epsilon_j w_j = \text{constant,} \end{cases} \tag{15.8}$$

where ϵ_j is the energy of the ith state, $\epsilon_j = j\epsilon$. This is a classic example of a problem to be solved by the technique of *undetermined multipliers*. We need only preserve those terms in the variable w_j since the number of oscillators N and the total number of energy elements r are fixed. Therefore, we need to find the turning values of the function

$$S(w_j) = -\sum_j w_j \ln w_j - A\sum_j w_j - B\sum_j \epsilon_j w_j, \tag{15.9}$$

where A and B are constants to be found from the boundary conditions. Maximising $S(w_j)$, we find

$$\begin{aligned} \delta[S(w_j)] &= -\delta\sum_j w_j \ln w_j - A\,\delta\sum_j w_j - B\,\delta\sum_j \epsilon_j w_j \\ &= -\sum_j [(\ln w_j\,\delta w_j + \delta w_j) + A\,\delta w_j + B\,\epsilon_j\,\delta w_j] = 0 \\ &= -\sum_j \delta w_j\,[\ln w_j + \alpha + \beta\epsilon_j] = 0. \end{aligned} \tag{15.10}$$

This must be true for all j and therefore

$$w_j = \mathrm{e}^{-\alpha-\beta\epsilon_j}. \tag{15.11}$$

This is the primitive form of Boltzmann's distribution. The term $\mathrm{e}^{-\alpha}$ is a constant in front of the exponential, $w_j \propto \mathrm{e}^{-\beta\epsilon_j}$. There is nothing in this analysis so far to tell us what the constant β should be. Indeed, so far, we have simply carried out a statistical analysis. We have to appeal to analyses such as that which leads to Maxwell's velocity distribution to find the value of β (see Section 12.4). From (12.46), it can be seen that $\beta \propto T^{-1}$ and, in modern notation, we write

$$w_j \propto \mathrm{e}^{-\epsilon_j/k_\mathrm{B}T}, \tag{15.12}$$

where k_B is Boltzmann's constant. Finally, Boltzmann allows the energy elements ϵ to tend to zero in such a way that $E = \epsilon_j = j\epsilon$ remains finite and so ends up with a continuous energy distribution,

$$w(E) \propto \mathrm{e}^{-E/k_\mathrm{B}T}. \tag{15.13}$$

15.3 Planck's Analysis

Planck's analysis began by following Boltzmann's procedure. There is a fixed total energy E to be divided among the N oscillators and energy elements ϵ are introduced. Therefore, as above, there are $r = E/\epsilon$ energy elements to be shared among the oscillators. Rather (*a*) ×|×××||×××|××|×××××|×|××××||×

(*b*) ××××|××||×|×××|××|×|×××|×××|×

Fig. 15.1 Two examples of distributing 20 energy elements among 10 boxes.

than following the procedure outlined in the last section, however, Planck simply worked out the *total number of ways* in which r energy elements can be distributed over the N oscillators. We can work this number out using the permutation theory outlined in the last section.

The problem can be represented by the diagram shown in Fig. 15.1. Planck had two fixed quantities, the total number of energy elements r and the number of boxes N into which he wished to place them. In Fig. 15.1(*a*), one way of placing the 20 energy elements into 10 boxes is illustrated. It can be seen that the whole problem reduces to no more than determining the total number of ways in which the elements and the walls of the boxes can be arranged between the end stops. A second example is shown in Fig. 15.1(*b*). Thus, the problem amounts to working out the number of different ways in which r elements and $N-1$ walls can be distributed, remembering that the r elements are identical and the $N-1$ walls are identical. We deduced the answer in the last section:

$$\frac{(N+r-1)!}{r!\,(r-1)!}. \tag{15.14}$$

This elegant argument was first presented by Ehrenfest in 1914.[3] Expression (15.14) represents the *total number of ways* of distributing an energy E over N oscillators. Now Planck made the crucial step in his argument. He *defined* (15.14) to be the probability which should be used in the relation

$$S = C \ln W.$$

Let us see where this leads. N and r are very large indeed and so we can use Stirling's approximation:

$$n! = (2\pi n)^{1/2} \left(\frac{n}{\mathrm{e}}\right)^n \left(1 + \frac{1}{12n} + \cdots\right). \tag{15.15}$$

We need to take the logarithm of (15.14) and so we can use an even simpler approximation than that used in (15.6). Since n is very large, we write to a good approximation $n! \approx n^n$ and so

$$W = \frac{(N+r-1)!}{r!\,(N-1)!} \approx \frac{(N+r)!}{r!\,N!} \approx \frac{(N+r)^{N+r}}{r^r N^N}. \tag{15.16}$$

Therefore,

$$\begin{cases} S = C[(N+r)\ln(N+r) - r\ln r - N\ln N], \\ r = \dfrac{E}{\epsilon} = \dfrac{N\overline{E}}{\epsilon}, \end{cases} \tag{15.17}$$

where $\overline{E}$ is the average energy of the oscillators. Therefore,

$$S = C\left\{N\left(1+\frac{\overline{E}}{\epsilon}\right)\ln\left[N\left(1+\frac{\overline{E}}{\epsilon}\right)\right] - \frac{N\overline{E}}{\epsilon}\ln\frac{N\overline{E}}{\epsilon} - N\ln N\right\}. \tag{15.18}$$

The average entropy per oscillator $\overline{S}$ is therefore

$$\overline{S} = \frac{S_N}{N} = C\left[\left(1+\frac{\overline{E}}{\epsilon}\right)\ln\left(1+\frac{\overline{E}}{\epsilon}\right) - \frac{\overline{E}}{\epsilon}\ln\frac{\overline{E}}{\epsilon}\right]. \tag{15.19}$$

But this looks rather familiar. Equation (15.19) is exactly the expression for the entropy of an oscillator which Planck had derived to account for the spectrum of black-body radiation. From (14.42), we see that

$$S = a\left[\left(1+\frac{\overline{E}}{b}\right)\ln\left(1+\frac{\overline{E}}{b}\right) - \frac{\overline{E}}{b}\ln\frac{\overline{E}}{b}\right],$$

with the requirement that $b \propto \nu$. Thus, the energy elements ϵ must be proportional to frequency and Planck wrote this result in the form which has persisted to this day:

$$\epsilon = h\nu, \tag{15.20}$$

where h is rightly known as Planck's constant. This is the origin of the concept of *quantisation*. According to classical statistical mechanics, we ought now to allow $\epsilon \to 0$, but evidently we cannot obtain the expression for the entropy of an oscillator unless the energy elements do not disappear, but have finite magnitude $\epsilon = h\nu$. Furthermore, we have now determined the value of a in terms of the universal constant C. Therefore, we can write the complete expression for the energy density of black-body radiation:

$$u(\nu) = \frac{8\pi h\nu^3}{c^3}\frac{1}{\mathrm{e}^{h\nu/CT}-1}. \tag{15.21}$$

What about C? Planck pointed out that C is a universal constant relating the entropy to the probability and Boltzmann had implicitly worked out its value for a perfect gas. Since C is a universal constant, any suitable law such as the perfect gas law determines its value for all processes. For example, we could use the results of the Joule expansion of a perfect gas, treated classically and statistically, as was demonstrated in Section 12.7. As shown in that section, the ratio $C = k_B = R/N_A$, where R is the gas constant and N_A is Avogadro's number, the number of molecules per gram molecule. We can at last write down the Planck distribution once and for all in its final form:

$$u(\nu) = \frac{8\pi h\nu^3}{c^3}\frac{1}{\mathrm{e}^{h\nu/k_B T}-1}. \tag{15.22}$$

It is straightforward to integrate this expression to find the total energy density of radiation u in the black-body spectrum:

$$u = \int_0^\infty u(\nu)\,\mathrm{d}\nu = \frac{8\pi h}{c^3}\int_0^\infty \frac{\nu^3\,\mathrm{d}\nu}{\mathrm{e}^{h\nu/k_B T}-1}. \tag{15.23}$$

This is a standard integral, the value of which is

$$\int_0^\infty \frac{x^3\,\mathrm{d}x}{\mathrm{e}^x - 1} = \frac{\pi^4}{15}.$$

Therefore,

$$u = \left(\frac{8\pi^5 k_\mathrm{B}^4}{15c^3h^3}\right) T^4 = aT^4. \tag{15.24}$$

We have recovered the Stefan–Boltzmann law for the energy density of radiation u. Substituting the values of the constants,

$$a = 7.566 \times 10^{-16}\ \mathrm{J\ m^{-3}\ K^{-4}}.$$

We can relate this energy density to the energy emitted per second from the surface of a black-body maintained at temperature T using the same argument employed in Section 13.3.3. The rate at which energy arrives at the walls per second, and consequently the rate at which the energy must be reradiated is $\frac{1}{4}uc$. Therefore,

$$I = \tfrac{1}{4}uc = \frac{ac}{4}T^4 = \sigma T^4 = 5.67 \times 10^{-8} T^4\ \mathrm{W\ m^{-2}}.$$

This analysis provides a determination of the value of the Stefan–Boltzmann constant in terms of fundamental constants:

$$\sigma = \frac{ac}{4} = \left(\frac{2\pi^5 k_\mathrm{B}^4}{15c^2h^3}\right) = 5.67 \times 10^{-8}\ \mathrm{W\ m^{-2}\ K^{-4}}. \tag{15.25}$$

What is one to make of Planck's argument? There are two fundamental concerns:

(1) Planck certainly does not follow Boltzmann's procedure for finding the equilibrium distribution. What he defines as a probability is not really a probability of anything drawn from any parent population. Planck had no illusions about this. In his own words:

> In my opinion, this stipulation basically amounts to a definition of the probability W; for we have absolutely no point of departure, in the assumptions which underlie the electromagnetic theory of radiation, for talking about such a probability with a definite meaning.[4]

Einstein repeatedly pointed out this weakness in Planck's argument:

> The manner in which Mr Planck uses Boltzmann's equation is rather strange to me in that a probability of a state W is introduced without a physical definition of this quantity. If one accepts this, then Boltzmann's equation simply has no physical meaning.[5]

(2) A second problem concerns a logical inconsistency in Planck's analysis. On the one hand, the oscillators can only take energies $E = r\epsilon$ and yet a classical result has been used to work out the rate of radiation of the oscillator (Section 14.2.1). Implicit in that analysis was the assumption that the energies of the oscillators can vary continuously rather than take only discrete values.

These are major stumbling blocks and it is fair to say that nobody quite understood the significance of what Planck had done. The theory did not in any sense gain immediate acceptance. Nonetheless, whether one likes it or not, the concept of quantisation has been introduced by the energy elements without which it is not possible to reproduce the Planck function. It was some time before Planck fully appreciated the profound significance of what he had done. In 1906, Einstein showed that, if Planck had followed strictly Boltzmann's procedure, he would have obtained the same answer whilst maintaining the essential concept of energy quantisation. We repeat Einstein's analysis in Section 16.3.

A second point concerns the contrast between the quantum and classical parts of the derivation of Planck's formula for the black-body spectrum. Einstein was not nearly as worried as other physicists about mixing up macroscopic and microscopic physical concepts. He regarded equations such as Maxwell's equations for the electromagnetic field as statements only about the *average values* of the quantities measured. In the same way, the expression relating $u(\nu)$ and E may be another of these relations which has significance independent of electromagnetic theory, although it can be derived by that means. More important, while the relation might not be precisely true on the microscopic scale, it may still be a good representation of the average behaviour of the system as measured in laboratory experiments. This was a sophisticated point of view and characteristic of Einstein's thinking in the crucial years of the first decade of the twentieth century. We will see in the next chapter how this approach led to Einstein's spectacular discovery of light quanta.

15.4 Planck and 'Natural Units'

Why did Planck take his derivation of the black-body radiation formula so seriously? As noted by Martin Klein, Planck had a deeply held conviction that

> the search for the absolute (is) the loftiest of all scientific activity.[6]

Planck realised that there were two fundamental constants in his theory of black-body radiation, k_B and h. The fundamental nature of k_B was apparent from its appearance in the kinetic theory as the gas constant per molecule, Boltzmann's constant. Furthermore, Planck had shown that the constant C in Boltzmann's relation $S = C \ln W$ was in fact the same as Boltzmann's constant k_B. Both constants could be determined with some precision from the experimentally measured form of the spectrum of black-body radiation (15.22), and from the value of the constant in the Stefan–Boltzmann law (15.24) or (15.25). Combining the known value of the gas constant R with his new determination of k_B, Planck found a value for Avogadro's number N_A of 6.175×10^{23} molecules per mole, by far the best estimate known at that time. The present adopted value is $N_A = 6.022 \times 10^{23}$ molecules per mole.

The electric charge carried by a gram equivalent of monovalent ions was known from electrolytic theory and is known as Faraday's constant. Knowing N_A precisely, Planck was able to derive the elementary unit of charge and found it to be $e = 4.69 \times 10^{-10}$ esu, corresponding to 1.56×10^{-19} C. Again, Planck's value was by far the best available at that time, contemporary experimental values ranging between 1.3 and 6.5×10^{-10} esu. The present standard value is 1.602×10^{-19} C.

Equally compelling for Planck was the fact that h, in conjunction with the constant of gravitation G and the velocity of light c, enabled a set of 'natural' units to be defined in terms of fundamental constants. It is worthwhile quoting Klein's eloquent words on this topic:

> All systems of units previously employed owed their origins to accidents of human life on Earth, wrote Planck. The usual units of length and time derived from the size of the Earth and the period of its orbit, those of mass and temperature from the special properties of water, the Earth's most characteristic feature. Even the standardisation of length using some spectral line would be quite as arbitrary, as anthropomorphic, since the particular line, say the sodium D-line, would be chosen to suit the convenience of the physicist. The new units he was proposing would be truly 'independent of particular bodies or substances, would necessarily retain their significance for all times and for all cultures including extra-terrestrial and non-human ones,' and therefore deserve the term *natural units*.[7]

From the dimensions of these quantities, we can derive 'natural units' of time, length and mass as follows:

Time $t_{\rm Pl} = (Gh/c^5)^{1/2}$ 10^{-43} s
Length $l_{\rm Pl} = (Gh/c^3)^{1/2}$ 4×10^{-35} m
Mass $m_{\rm Pl} = (hc/G)^{1/2}$ 5.4×10^{-8} kg $\equiv 3 \times 10^{19}$ GeV

Often these *Planck units* are written in terms of $\hbar = h/2\pi$, in which case their values are 2.5 times smaller than those quoted in the table. It is evident that the natural units of time and length are very small indeed, while the mass is much greater than that of any known elementary particle. Nonetheless, Planck was absolutely correct to note that these are the fundamental dimensions which involve combinations of the 'classical' constants c and G with the fundamental *quantum of action* h, as Planck referred to it.

It is striking that, a century later, these quantities play a central role in the physics of the very early Universe. They define the times, scales and masses at which we need to take seriously the quantisation of gravity.[8] There is no theory of quantum gravity, but current thinking is that many of the large-scale features of our Universe were laid down by physical processes occurring at the Planck era $t_{\rm Pl}$ at the enormous temperatures associated with the natural unit of mass–energy $k_{\rm B}T = 3 \times 10^{19}$ GeV. We will return to some aspects of these ideas in Chapter 21.

15.5 Planck and the Physical Significance of *h*

It was a number of years before the truly revolutionary nature of what Planck had achieved in these crucial last months of 1900 was appreciated. Little interest was shown in the physical content of his paper, largely, one suspects, because no-one quite understood what he had done. Perhaps surprisingly, Planck wrote no papers on the subject of quantisation over the next five years. The next publication which casts some light on his understanding was his book *Lectures on the Theory of Thermal Radiation* of 1906.[9] Thomas Kuhn gives a detailed analysis of Planck's thoughts on quantisation though the period 1900 to 1906 on the basis of these lectures in Chapter 5 of his monograph.[10]

Kuhn's analysis is brilliant, but it is a tortured, unresolved story. What is clear from Kuhn's analysis is that Planck undoubtedly believed that the classical laws of electromagnetism were applicable to the processes of the emission and absorption of radiation, despite the introduction of the finite energy elements in his theory. In the *Lectures*, Planck describes two versions of Boltzmann's procedure in statistical physics. One is the version described in Section 15.2, in which the energies of the oscillators are supposed to take values $0, \epsilon, 2\epsilon, 3\epsilon$, and so on. But there is also a second version in which the molecules were considered to lie within the energy ranges 0 to ϵ, ϵ to 2ϵ, 2ϵ to 3ϵ, and so on. This procedure leads to exactly the same statistical probabilities as the first version. In a subsequent passage, in which the motions of the oscillators are traced in phase space, he again refers to the energies of the trajectories corresponding to certain energy ranges U to $U + \Delta U$, the ΔU eventually being identified with $h\nu$. Thus, in some sense, Planck regarded quantisation as referring to the average properties of the oscillators.

In the *Lectures*, Planck had little to say about the nature of the quantum of action h, but he was well aware of its fundamental importance. In his words,

> The thermodynamics of radiation will therefore not be brought to an entirely satisfactory state until the full and universal significance of the constant h is understood.[10]

He suggested that the solution might lie in a more detailed understanding of the microphysics of the process of radiation and this is supported by a letter to Ehrenfest of July 1905 in which he remarks

> It seems to me not impossible that this assumption (the existence of an elementary quantum of electricity e) offers a bridge to the existence of an elementary energetic quantum h, particularly since h has the same dimensions as e^2/c (in electrostatic units).[11]

Planck spent many years trying to reconcile his theory with classical physics, but he failed to find any physical significance for h beyond its appearance in the radiation formula. In his words:

> My futile attempts to fit the elementary quantum of action somehow into the classical theory continued for a number of years and they cost me a great deal of effort. Many of my colleagues saw in this something bordering on tragedy. But I feel differently about it. For the thorough enlightenment I thus received was all the more valuable. I now knew for a fact that the elementary quantum of action played a far more significant part in physics than I had originally been inclined to suspect and this recognition made me see clearly the need for the introduction of totally new methods of analysis and reasoning in the treatment of atomic problems.[12]

Indeed, it was not until after about 1908 that Planck fully appreciated the quite fundamental nature of quantisation, which has no counterpart in classical physics. His original view was that the introduction of energy elements was

> a purely formal assumption and I really did not give it much thought except that no matter what the cost, I must bring about a positive result.[13]

This quotation is from a letter by Planck to R.W. Wood written in 1931, 30 years after the events described in this chapter. I find it a rather moving letter and it is worthwhile reproducing it in full.

> October 7 1931
>
> My dear colleague,
>
> You recently expressed the wish, after our fine dinner in Trinity Hall, that I should describe from a psychological viewpoint the considerations which had led me to propose the hypothesis of energy quanta. I shall attempt herewith to respond to your wish.
>
> Briefly summarised, what I did can be described as simply an act of desperation. By nature I am peacefully inclined and reject all doubtful adventures. But by then I had been wrestling unsuccessfully for six years (since 1894) with the problem of equilibrium between radiation and matter and I knew that this problem was of fundamental importance to physics; I also knew the formula that expresses the energy distribution in the normal spectrum. A theoretical interpretation therefore had to be found at any cost, no matter how high. It was clear to me that classical physics could offer no solution to this problem and would have meant that all energy would eventually transfer from matter into radiation. In order to prevent this, a new constant is required to assure that energy does not disintegrate. But the only way to recognise how this can be done is to start from a definite point of view. This approach was opened to me by maintaining the two laws of thermodynamics. The two laws, it seems to me, must be upheld under all circumstances. For the rest, I was ready to sacrifice every one of my previous convictions about physical laws. Boltzmann had explained how thermodynamic equilibrium is established by means of a statistical equilibrium, and if such an approach is applied to the equilibrium between matter and radiation, one finds that the continuous loss of energy into radiation can be prevented by assuming that energy is forced at the outset to remain together in certain quanta. This was purely a formal assumption and I really did not give it much thought except that no matter what the cost, I must bring about a positive result.
>
> I hope that this discussion is a satisfactory response to your inquiry. In addition I am sending you as printed matter the English version of my Nobel lecture on the same topic. I cherish the memory of my pleasant days in Cambridge and the fellowship with our colleagues.
>
> With kind regards,
>
> Very truly yours
>
> M. Planck.[13]

It is intriguing that Planck was prepared to give up the whole of physics, *except the two laws of thermodynamics*, in order to understand the black-body radiation spectrum.

There is undoubtedly a sense of incompleteness about Planck's epoch-making achievements. They are very great indeed and the sense of a titanic intellectual struggle is vividly conveyed. This tortuous struggle is in sharp contrast to the virtuosity of the next giant steps taken by Albert Einstein. Before moving on to this new phase of development, let us conclude the story of Planck's statistical mechanics.

15.6 Why Planck Found the Right Answer

Why did Planck find the correct expression for the radiation spectrum, despite the fact that the statistical procedures he used were more than a little suspect? There are two answers, one methodological, the other physical. The first is that it seems quite likely that Planck worked backwards. It was suggested by Rosenfeld, and endorsed by Klein on the basis of an article by Planck of 1943, that he started with the expression for the entropy of an oscillator (14.42) and worked backwards to find W from $\exp(S/k_B)$. This results in the permutation formula on the right in (15.16), which is more or less exactly the same as (15.14) for large values of N and r. Equation (15.14) was a well-known formula in permutation theory, which appeared in the works of Boltzmann in his patient exposition of the fundamentals of statistical physics. Planck then regarded (15.14) as the definition of entropy according to statistical physics. If this is indeed what happened, it in no sense diminishes Planck's achievement, in my view.

The second answer is that Planck had stumbled by accident upon one of the correct methods of evaluating the statistics of indistinguishable particles according to quantum mechanics, a concept which was unheard of at the time. These procedures were first demonstrated by the Indian mathematical physicist Satyendra Nath Bose in a manuscript entitled *Planck's Law and the Hypothesis of Light Quanta*,[14] which he sent to Einstein in 1924. Einstein immediately appreciated its deep significance, translated it into German himself and arranged for it to be published in the journal *Zeitschrift für Physik*. Bose's paper and his collaboration with Einstein led to the establishment of the method of counting indistinguishable particles in quantum mechanics known as *Bose–Einstein statistics*, which differ radically from classical Boltzmann statistics. Equally remarkable is the fact that the new statistical procedures were formulated *before* the discovery of quantum mechanics. Einstein went on to apply these new procedures to the statistical mechanics of an ideal gas.[15]

Bose was not really aware of the profound implications of his derivation of the Planck spectrum. To paraphrase Pais's account of the originality of his paper,[16] Bose introduced three new features into statistical physics:

(i) Photon number is not conserved.
(ii) Bose divides phase space into coarse-grained cells and works in terms of the numbers of particles per cell. The counting explicitly requires that, because the photons are taken to be identical, each possible distribution of states should be counted only once. Thus, Boltzmann's axiom of the distinguishability of particles is gone.
(iii) Because of this way of counting, the statistical independence of particles has gone.

These very profound differences from the classical Boltzmann approach were to find an explanation in quantum mechanics and are associated with the symmetries of the wave functions for particles of different spins. It was only with Dirac's discovery of relativistic quantum mechanics that the significance of these deep symmetries was to be fully appreciated. As Pais remarks,

> The astonishing fact is that Bose was correct on all three counts. (In his paper, he commented on none of them.) I believe there had been no such successful shot in the dark since Planck introduced the quantum in 1900.[16]

The full explanation of these topics would take us far from the main development of this chapter, but let us show how (15.22) can derived according to Bose–Einstein statistics and why Plancks's expression for the probability and hence the entropy turned out to be correct in this case. An excellent concise treatment of these topics is given by Huang in his book *Introduction to Statistical Physics*.[17] The approach is different from the procedure discussed above in that the argument starts by dividing the volume of phase space into elementary cells. Consider one of these cells, which we label k and which has energy ϵ_k and degeneracy g_k, the latter meaning the number of available states with the same energy ϵ_k within that cell. Now suppose there are n_k particles to be distributed over these g_k states and that the particles are identical. Then, the number of different ways the n_k particles can be distributed over these states is

$$\frac{(n_k + g_k - 1)!}{n_k!\,(g_k - 1)!}, \tag{15.26}$$

using exactly the same logic which led to (15.14). This is the key step in the argument and differs markedly from the corresponding Boltzmann result (15.3). The key point is that in (15.3) all possible ways of distributing the particles over the energy states are included in the statistics, whereas in (15.26), duplications of the same distribution are eliminated because of the factorials in the denominator. At the quantum level, the explanation for this distinction is neatly summarised by Huang who remarks:

> The classical way of counting in effect accepts all wave functions regardless of their symmetry properties under the interchange of coordinates. The set of acceptable wave functions is far greater than the union of the two quantum cases [the Fermi–Dirac and Bose–Einstein cases]

The result (15.26) refers only to a single cell in phase space and we need to extend it to all the cells which make up the phase space. The total number of possible ways of distributing the particles over all the cells is the product of all numbers such as (15.26), that is,

$$W = \prod_k \frac{(n_k + g_k - 1)!}{n_k!\,(g_k - 1)!}. \tag{15.27}$$

We have not specified yet how the $N = \sum_k n_k$ particles are to be distributed among the k cells. To do so, we ask as before, 'What is the arrangement of n_k over the states which results in the maximum value of W?' At this point, we return to the recommended Boltzmann procedure. First, Stirling's theorem is used to simplify $\ln W$:

$$\ln W = \ln \prod_k \frac{(n_k + g_k - 1)!}{n_k!\,(g_k - 1)!} \approx \sum_k \ln \frac{(n_k + g_k)^{n_k+g_k}}{(n_k)^{n_k}\,(g_k)^{g_k}}. \tag{15.28}$$

Now we maximise W subject to the constraints $\sum_k n_k = N$ and $\sum n_k \epsilon_k = E$. Using the method of undetermined multipliers as before,

$$\delta(\ln W) = 0 = \sum_k \delta n_k \{[\ln(g_k + n_k) - \ln n_k] - \alpha - \beta\epsilon_k\},$$

so that

$$[\ln(g_k + n_k) - \ln n_k] - \alpha - \beta\epsilon_k = 0,$$

and finally

$$n_k = \frac{g_k}{e^{\alpha+\beta\epsilon_k} - 1}. \tag{15.29}$$

This is known as the *Bose–Einstein distribution* and is the correct statistics for counting the indistinguishable particles known as *bosons*. According to quantum mechanics, bosons are particles with even spin. For example, photons are spin-1 particles, gravitons spin-2 particles, and so on.

In the case of black-body radiation, we do not need to specify the number of photons present. We can see this from the fact that the distribution is determined solely by one parameter – the total energy, or the temperature of the system. Therefore, in the method of undetermined multipliers, we can drop the restriction on the total number of particles. The distribution will automatically readjust to the total amount of energy present, and so $\alpha = 0$. Therefore,

$$n_k = \frac{g_k}{e^{\beta\epsilon_k} - 1}. \tag{15.30}$$

By inspecting the low frequency behaviour of the Planck spectrum, we can show that $\beta = 1/k_B T$, as in the classical case.

Finally, the degeneracy of the cells in phase space g_k for radiation in the frequency interval ν to $\nu + d\nu$ has already been worked out in our discussion of Rayleigh's approach to the origin of the black-body spectrum (Section 14.6). One of the reasons for Einstein's enthusiasm for Bose's paper was that Bose had derived this factor entirely by considering the phase space available to the photons, rather than appealing to Planck's or Rayleigh's approaches, which relied upon results from classical electromagnetism. In Planck's analysis, the derivation of his expression (14.25) was entirely electromagnetic and Rayleigh's argument proceeded by fitting electromagnetic waves within a box with perfectly conducting walls. Bose considered the photons to have momenta $p = h\nu/c$ and so the volume of momentum, or phase, space for photons in the energy range $h\nu$ to $h(\nu + d\nu)$ is, using the standard procedure,

$$dV_p = V\, dp_x\, dp_y\, dp_z = V 4\pi p^2\, dp = \frac{4\pi h^3 \nu^2\, d\nu}{c^3} V, \tag{15.31}$$

where V is the volume of real space. Now Bose considered this volume of phase space to be divided into elementary cells of volume h^3, following ideas first stated by Planck in his 1906 lectures, and so the numbers of cells in phase space was

$$dN_\nu = \frac{4\pi \nu^2\, d\nu}{c^3} V. \tag{15.32}$$

He needed to take account of the two polarisation states of the radiation and so Rayleigh's result was recovered,

$$dN = \frac{8\pi\nu^2}{c^3}\,d\nu \quad \text{with} \quad \epsilon_k = h\nu. \tag{15.33}$$

We immediately find the expression for the spectral energy density of the radiation:

$$u(\nu)\,d\nu = \frac{8\pi h\nu^3}{c^3}\frac{1}{e^{h\nu/k_B T} - 1}\,d\nu. \tag{15.34}$$

This is Planck's expression for the black-body radiation spectrum and it has been derived using Bose–Einstein statistics for indistinguishable particles. The statistics are applicable not only to photons but also to integral-spin particles of all types.

We have now run far ahead of our story. None of this was known in 1900 and it was in the following years that Einstein made his revolutionary contributions to modern physics.

Notes

1 Planck, M. (1950). *Scientific Autobiography and Other Papers*, p. 41. London: Williams and Norgate.
2 Planck, M. (1900). *Verhandlungen der Deutschen Physikalischen Gessellschaft*, **2**, 237–245. See also *Collected Scientific Works* (1958), *Physikalische Abhandlungen und Vortrage (Collected Scientific Papers)*, Vol. 1, p. 698. Braunschweig: Friedr. Vieweg und Sohn. (English Translation: *Planck's Original Papers in Quantum Physics* (1972), annotated by H. Kangro, pp. 38–45. London: Taylor and Francis.)
3 Ehrenfest, P. and Kamerlingh Onnes, H. (1914). *Proc. Acad. Amsterdam*, **17**, p. 870.
4 Planck, M. (1900). Quoted by M.J. Klein (1977) *History of Twentieth Century Physics* Proceedings of the International School of Physics 'Enrico Fermi', Course 57, p. 17. New York & London: Academic Press.
5 Einstein, A. (1912). *The Theory of Radiation and Quanta*, transactions of the First Solvay Conference, eds. P. Langevin and M. de Broglie, p. 115. Paris: Gauthier-Villars. (Translation of quotation: see A. Hermann (1971) *The Genesis of Quantum Theory (1899–1913)*, p. 20. Cambridge, Massachusetts: MIT Press.)
6 Planck, M. (1950). *Scientific Autobiography and Other Papers*, p. 35. London: Williams and Norgate.
7 Klein, M.J. (1977). *History of Twentieth Century Physics*, Proceedings of the International School of Physics 'Enrico Fermi', Course 57, pp. 13–14. New York and London: Academic Press.
8 Frolov, V.I. and Novikov, I.D. (1998). *Black Hole Physics*. Dordrecht: Kluwer Academic Publishers.
9 Planck, M. (1906). *Vorlesungen über die Theorie der Wärmstrahlung*, 1st edition. Leipzig: Barth.
10 Kuhn, T.S. (1978). *Black-Body Theory and the Quantum Discontinuity, 1894–1912*. Oxford: Clarendon Press.
11 Planck, M. (1905). Quoted by Kuhn (1978), *op. cit.*, 132.
12 Planck, M. (1950). *op. cit.*, pp. 44–45.

13 Planck, M. (1931). Letter from M. Planck to R.W. Wood. See Hermann, A. (1971). *op. cit.*, pp. 23–4. Cambridge, Mass: MIT Press.
14 Bose, S.N. (1924). *Zeitschrift für Physik*, **26**, 178.
15 Einstein, A. (1924). *Sitzungsberichte, Preuss. Akad. der Wissenschaften*, p. 261; (1925). *ibid.*, p. 3.
16 Pais, A. (1982). *Subtle is the Lord … – the Science and Life of Albert Einstein*. Oxford: Clarendon Press.
17 Huang, K. (2001). *Introduction to Statistical Physics*. London: Taylor and Francis.

16 Einstein and the Quantisation of Light

16.1 1905: Einstein's *Annus Mirabilis*

By 1905, Planck's work had made little impression and he was no further forward in understanding the profound implications of what he had done. He expended a great deal of unsuccessful effort in trying to find a classical interpretation for the 'quantum of action' h, which he correctly recognised had fundamental significance for understanding the spectrum of black-body radiation. The next great steps were taken by Albert Einstein and it is no exaggeration to state that he was the first person to appreciate the full significance of quantisation and the reality of quanta. He showed that these are fundamental aspects of all physical phenomena, rather than just a 'formal device' for accounting for the Planck distribution. From 1905 onwards, he never deviated from his belief in the reality of quanta – it was some considerable time before the great figures of the day conceded that Einstein was indeed correct. He came to this conclusion in a series of brilliant papers of dazzling scientific virtuosity.

Einstein completed his undergraduate studies in August 1900. Between 1902 and 1904, he wrote three papers on the foundations of Boltzmann's statistical mechanics. Again, a deep understanding of thermodynamics and statistical physics provided the point of departure for the investigation of fundamental problems in theoretical physics.

16.1.1 Einstein's Great Papers of 1905

In 1905, Einstein was 26 yeas old and was employed as 'technical expert, third class' at the Swiss patent office in Bern. In that year, he completed his doctoral dissertation on *A New Determination of Molecular Dimensions*, which he presented to the University of Zurich on 20 July 1905. In the same year, he published three papers which are among the greatest classics of the physics literature. Any one of them would have ensured that his name remained a permanent fixture in the scientific canon. These papers were:

(1) *On the motion of small particles suspended in stationary liquids required by the molecular-kinetic theory of heat*.[1]
(2) *On the electrodynamics of moving bodies*.[2]
(3) *On a heuristic viewpoint concerning the production and transformation of light*.[3]

These papers are conveniently available in English translations in *Einstein's Miraculous Year*, edited and introduced by John Stachel.[4] These translations are taken from *The*

Collected Papers of Albert Einstein: Volume 2. The Swiss Years: Writings, 1900–1909,[5] which can also be strongly recommended.

The first paper is more familiarly known by the title of a subsequent paper published in 1906 entitled *On the Theory of Brownian Motion*[6] and is a reworking of some of the results of his doctoral dissertation. Brownian motion is the irregular motion of microscopic particles in fluids and had been studied in detail in 1828 by the botanist Robert Brown, who had noted the ubiquity of the phenomenon. The motion results from the aggregate effect of very large numbers of collisions between molecules of the fluid and the microscopic particles. Although each impact is very small, the net result of a very large number of them colliding randomly with the particle results in a 'drunken man's walk'. Einstein quantified the theory of this process by relating the diffusion of the particles to the properties of the molecules responsible for the collisions. In a beautiful analysis, he derived the formula for the mean squared distance travelled by the particle in time t:

$$\langle r^2 \rangle = \frac{k_B T t}{3\pi\eta a}, \tag{16.1}$$

where a is the radius of the particle, T the temperature, η the viscosity of the fluid and k_B is Boltzmann's constant. Crucially, Einstein had discovered the relation between the microscopic molecular properties of fluids and the observed diffusion of macroscopic particles. In his estimates of the magnitude of the effect for particles 1 μm in diameter, he needed a value for Avogadro's number N_A and he used the values which Planck and he had found from their studies of the spectrum of black-body radiation (see Section 16.2 below). He predicted that such particles would diffuse about 6 μm in one minute. In the last paragraph of the paper, Einstein states,

> Let us hope that a researcher will soon succeed in solving the problem presented here, which is so important for the theory of heat!

Remarkably, Einstein wrote this paper 'without knowing that observations concerning Brownian motion were already long familiar', as he remarked in his autobiographical notes[7] – this accounts for the absence of the term 'Brownian motion' in the title of his first paper on the subject.

Einstein's concern about the importance of this calculation for the theory of heat was well founded. The battle to establish the kinetic theory and Boltzmann's statistical procedures against the energeticists, led by Wilhelm Ostwald and Georg Helm, who denied the existence of atoms and molecules, was not yet won. In defence of their position, it has to be recognised that, at that time, there was no definite proof that heat is indeed associated with the random motions of atoms and molecules. Ernst Mach was hostile to the concept of atoms and molecules since they are not accessible directly to our senses, although he admitted that atomism was a useful conceptual tool. Many scientists simply regarded the concepts of atoms and molecules as useful working hypotheses. Einstein demonstrated that the agitational motion of the particles observed in Brownian motion *is* heat – the macroscopic particles reflect the motion of the molecules on the microscopic scale.

Precise observations of Brownian motion were difficult at that time, but in 1908, Jean Perrin[8] carried out a meticulous series of brilliant experiments which confirmed in detail all Einstein's predictions. This work convinced everyone, even the sceptics, of the reality of molecules. In Perrin's words,

> I think it is impossible that a mind free from all preconception can reflect upon the extreme diversity of the phenomena which thus converge to the same result without experiencing a strong impression, and I think it will henceforth be difficult to defend by rational arguments a hostile attitude to molecular hypotheses.[9]

The second paper is Einstein's famous paper on special relativity and its contents will be analysed Chapter 18. While listing Einstein's great papers of 1905, it is worth remarking that there is a second much shorter paper on special relativity entitled *Does the Inertia of a Body Depend on its Energy Content?*[10] This is an explicit statement that inertial mass and energy are the same thing, a forceful reminder that the expression $E = mc^2$ can be read forwards or backwards.

The third paper is often referred to as Einstein's paper on the photoelectric effect. This is a serious misrepresentation of the profundity of the paper. As Einstein wrote to his friend Conrad Habicht in May 1905:

> I promise you four papers in return, the first of which I might send you soon, since I will soon get the complimentary reprints. The paper deals with radiation and the energy properties of light and is very revolutionary, . . . [11]

Einstein was referring to the theoretical content of the paper. It is indeed revolutionary and is the subject of the Section 16.2.

16.1.2 Einstein in 1905

1905 is rightly referred to as Einstein's *annus mirabilis*. The term was first used in physics in connection with Newton's extraordinary achievements of 1665–66, which were discussed in Section 4.4. Einstein's achievement certainly ranks with Newton's, although they were very different characters with quite different outlooks on physics and mathematics. Stachel[4] gives a revealing comparison of their characters.

Newton's genius was not only in physics, but also in mathematics and, whilst the seeds of his great achievements were sown when he was only 22, it was many years before his ideas were crystallised into the forms we celebrate today. Einstein confessed that he was no mathematician and the mathematics needed to understand his papers of 1905 is no more than is taught in the first couple of years of an undergraduate physics course. His genius lay in his extraordinary physical intuition, which enabled him to see deeper into physical problems than his contemporaries. The three great papers were not the result of a sudden burst of creativity, but the product of deep pondering about fundamental problems of physics over almost a decade, which suddenly came to fruition almost simultaneously in 1905. Despite their obvious differences, the three papers have a striking commonality of approach. In each case, Einstein stands back from the specific problem at hand and studies carefully the underlying physical principles. Then, with dazzling virtuosity, he reveals the underlying physics in a new light. Nowhere is this genius more apparent than in his stunning papers on quanta and quantisation of 1905 and 1906.

16.2 Einstein (1905) *On a Heuristic Viewpoint Concerning the Production and Transformation of Light*

Let us reproduce the opening paragraphs of Einstein's paper – they are revolutionary and startling. They demand attention, like the opening of a great symphony.

> There is a profound formal difference between the theoretical ideas which physicists have formed concerning gases and other ponderable bodies and Maxwell's theory of electromagnetic processes in so-called empty space. Thus, while we consider the state of a body to be completely defined by the positions and velocities of a very large but finite number of atoms and electrons, we use continuous three-dimensional functions to determine the electromagnetic state existing within some region, so that a finite number of dimensions is not sufficient to determine the electromagnetic state of the region completely . . .
>
> The undulatory theory of light, which operates with continuous three-dimensional functions, applies extremely well to the explanation of purely optical phenomena and will probably never be replaced by any other theory. However, it should be kept in mind that optical observations refer to values averaged over time and not to instantaneous values. Despite the complete experimental verification of the theory of diffraction, reflection, refraction, dispersion and so on, it is conceivable that a theory of light operating with continuous three-dimensional functions will lead to conflicts with experience if it is applied to the phenomena of light generation and conversion.[3]

In other words, there may well be circumstances under which Maxwell's theory of the electromagnetic field cannot explain all electromagnetic phenomena and Einstein specifically gives as examples the spectrum of black-body radiation, photoluminescence and the photoelectric effect. His proposal is that, for some purposes, it may be more appropriate to consider light to be

> . . . discontinuously distributed in space. According to the assumption considered here, in the propagation of a light ray emitted from a point source, the energy is not distributed continuously over ever-increasing volumes of space, but consists of a finite number of energy quanta localised at points in space that move without dividing, and can be absorbed and generated only as complete units.

He ends by hoping that

> the approach to be presented will prove of use to some researchers in their investigations.

Einstein was really asking for trouble. The full implications of Maxwell's theory were still being worked out and here he was proposing to replace these hard-won achievements by light quanta. To physicists, this must have looked like a re-run of the controversy between the wave picture of Huygens and the particle, or corpuscular, picture of Newton – everyone knew which theory had won the day. The proposal must have seemed particularly untimely when Maxwell's discovery of the electromagnetic nature of light had been so completely validated only 15 years earlier by Heinrich Hertz (see Section 6.8).

Notice how Einstein's proposal differs from Planck's. Planck had discovered the necessity of quantising the energies of the oscillators responsible for electromagnetic radiation in terms of integral numbers of ‘energy elements’ $\epsilon = h\nu$. He had absolutely nothing to say about the radiation emitted by them. He firmly believed that the waves emitted by the oscillators were simply the classical electromagnetic waves of Maxwell. In contrast, Einstein proposed that the radiation field itself should be quantised.

Like Einstein’s other papers, the article is beautifully written and very clear. It all looks so very simple and obvious that it can be forgotten just how revolutionary its contents were. For clarity of exposition, we will continue to use the same modern notation which we have used up till now, rather than Einstein’s notation.

After the introduction, Einstein states Planck’s formula (14.25) relating the average energy of an oscillator to the energy density of black-body radiation in thermodynamic equilibrium. He does not hesitate, however, to take the average energy of the oscillator to be $\overline{E} = k_{\mathrm{B}}T$ according to the kinetic theory and then writes the total energy in the black-body spectrum in the provocative form

$$u(\nu) = \frac{8\pi\nu^2}{c^3} k_{\mathrm{B}}T, \tag{16.2}$$

$$\text{total energy} = \int_0^\infty u(\nu)\,\mathrm{d}\nu = \frac{8\pi k_{\mathrm{B}}T}{c^3}\int_0^\infty \nu^2\,\mathrm{d}\nu = \infty. \tag{16.3}$$

This is exactly the problem which had been pointed out by Rayleigh in 1900 and which led to his arbitrary introduction of the ‘standard’ exponential factor to prevent the spectrum diverging at high frequencies (see Section 14.6). This phenomenon was later called the *ultraviolet catastrophe* by Paul Ehrenfest.

Einstein next goes on to show that, despite the high frequency divergence of the expression (16.2), it is a very good description of the black-body radiation spectrum at low frequencies and high temperatures, and hence the value of Boltzmann’s constant k_{B} can be derived from that part of the spectrum alone. Einstein’s value of k_{B} agreed precisely with Planck’s estimate, which Einstein interpreted as meaning that Planck’s value was, in fact, independent of the details of the theory he had developed to account for the black-body spectrum.

Now we come to the heart of the paper. We have already emphasised the central role which entropy plays in the thermodynamics of radiation. Einstein now derives a suitable expression for the entropy of black-body radiation using only thermodynamics and the observed form of the radiation spectrum. Entropies are additive, and since, in thermal equilibrium, we may consider the radiation of different wavelengths to be independent, the entropy of the radiation enclosed in volume V can be written

$$S = V\int_0^\infty \phi[u(\nu), \nu]\,\mathrm{d}\nu. \tag{16.4}$$

The function ϕ is the entropy of the radiation per unit frequency interval per unit volume. The aim of the calculation is to find an expression for the function ϕ in terms of the spectral energy density $u(\nu)$ and frequency ν. No other quantities besides the temperature T can be involved in the expression for the equilibrium spectrum, as was shown by Kirchhoff according to the argument reproduced in Section 13.2.2. The problem had already been solved by Wien but, for completeness, Einstein gave a simple proof of the result as follows.

The function ϕ must be such that, in thermal equilibrium, the entropy is a maximum for a fixed value of the total energy and so the problem can be written in terms of the calculus of variations:

$$\delta S = \delta \int_0^\infty \phi[u(\nu), \nu]\, \mathrm{d}\nu = 0, \tag{16.5}$$

subject to the constraint

$$\delta E = \delta \int_0^\infty u(\nu)\, \mathrm{d}\nu = 0, \tag{16.6}$$

where E is the total energy. Using the method of undetermined multipliers, the equation for the function $\phi(u)$ is given by

$$\int_0^\infty \left(\frac{\partial \phi}{\partial u}\, \mathrm{d}u\, \mathrm{d}\nu - \lambda\, \mathrm{d}u\, \mathrm{d}\nu \right) = 0,$$

where the constant λ is independent of frequency. The integrand must be zero to ensure that the integral is zero and so

$$\frac{\partial \phi}{\partial u} = \lambda.$$

Now suppose the temperature of unit volume of black-body radiation is increased by an infinitesimal amount $\mathrm{d}T$. The increase in entropy is then

$$\mathrm{d}S = \int_{\nu=0}^{\nu=\infty} \frac{\partial \phi}{\partial u}\, \mathrm{d}u\, \mathrm{d}\nu.$$

But $\partial\phi/\partial u$ is independent of frequency and hence

$$\mathrm{d}S = \frac{\partial \phi}{\partial u}\, \mathrm{d}E,$$

where

$$\mathrm{d}E = \int_{\nu=0}^{\nu=\infty} \mathrm{d}u\, \mathrm{d}\nu. \tag{16.7}$$

But $\mathrm{d}E$ is just the energy added to cause the infinitesimal increase in temperature and the thermodynamic relation (11.57) states that

$$\frac{\mathrm{d}S}{\mathrm{d}E} = \frac{1}{T}. \tag{16.8}$$

Therefore,

$$\frac{\partial \phi}{\partial u} = \frac{1}{T}. \tag{16.9}$$

This is the equation we have been seeking. Notice the pleasant symmetry between the relations:

$$\begin{cases} S = \displaystyle\int_0^\infty \phi\, \mathrm{d}\nu, & E = \displaystyle\int_0^\infty u(\nu)\, \mathrm{d}\nu, \\[2ex] \dfrac{\mathrm{d}S}{\mathrm{d}E} = \dfrac{1}{T}, & \dfrac{\partial \phi}{\partial u} = \dfrac{1}{T}. \end{cases} \tag{16.10}$$

Einstein now uses (16.9) to work out the entropy of black-body radiation. Rather than use Planck's formula, he uses Wien's formula since, although it is not correct for low frequencies and high temperatures, it is an excellent fit to the measured spectrum of black-body radiation in the region where the classical theory breaks down. Therefore, the analysis of the entropy associated with this part of the spectrum is likely to give insight into where the classical calculation has gone wrong.

First, Einstein writes down the form of Wien's law as derived from experiment. In the notation of (14.28),

$$u(\nu) = \frac{8\pi\alpha}{c^3}\frac{\nu^3}{e^{\beta\nu/T}}. \tag{16.11}$$

From (16.11), we immediately find an expression for $1/T$,

$$\frac{1}{T} = \frac{1}{\beta\nu}\ln\frac{8\pi\alpha\nu^3}{c^3 u(\nu)} = \frac{\partial\phi}{\partial u}. \tag{16.12}$$

The expression for ϕ is found by integration:

$$\begin{aligned}
\frac{\partial\phi}{\partial u} &= -\frac{1}{\beta\nu}\left(\ln u + \ln\frac{c^3}{8\pi\alpha\nu^3}\right), \\
\phi &= -\frac{u}{\beta\nu}\left(\ln u - 1 + \ln\frac{c^3}{8\pi\alpha\nu^3}\right) \\
&= -\frac{u}{\beta\nu}\left(\ln\frac{uc^3}{8\pi\alpha\nu^3} - 1\right).
\end{aligned} \tag{16.13}$$

Now Einstein makes clever use of this formula for ϕ. Consider the energy density of radiation in the spectral range ν to $\nu + \Delta\nu$ and suppose it has energy $\varepsilon = Vu\,\Delta\nu$, where V is the volume. Then the entropy associated with this radiation is

$$S = V\phi\,\Delta\nu = -\frac{\varepsilon}{\beta\nu}\left(\ln\frac{\varepsilon c^3}{8\pi\alpha\nu^3 V\,\Delta\nu} - 1\right). \tag{16.14}$$

Suppose the volume changes from V_0 to V, while the total energy remains constant. Then, the entropy change is

$$S - S_0 = \frac{\varepsilon}{\beta\nu}\ln(V/V_0). \tag{16.15}$$

But this formula looks familiar. Einstein shows that this entropy change is exactly the same as that found in the Joule expansion of a perfect gas according to elementary statistical mechanics. Let us repeat the analysis of Section 12.7 which led to (12.55). Boltzmann's relation can be used to work out the entropy difference $S - S_0$ between the initial and final states, $S - S_0 = k_B \ln W/W_0$, where the Ws are the probabilities of these states. In the initial state, the system has volume V_0 and the particles move randomly throughout this volume. The probability that a single particle occupies a smaller volume V is V/V_0 and hence the probability that all N end up in the volume V is $(V/V_0)^N$. Therefore, the entropy difference for a gas of N particles is

$$S - S_0 = k_B N\,\ln(V/V_0). \tag{16.16}$$

Einstein notes that (16.15) and (16.16) are formally identical. He immediately concludes that the radiation behaves thermodynamically as if it consisted of discrete particles, their number N being equal to $\varepsilon/k_B\beta\nu$. In Einstein's words,

> Monochromatic radiation of low density (within the limits of validity of Wien's radiation formula) behaves thermodynamically as though it consisted of a number of independent energy quanta of magnitude $k_B\beta\nu$.

Rewriting this result in Planck's notation, since $\beta = h/k_B$, the energy of each quantum is $h\nu$.

Einstein then works out the average energy of the quanta according to Wien's formula for a black-body spectrum. The energy in the frequency interval ν to $\nu + d\nu$ is ε and the number of quanta is $\varepsilon/k_B\beta\nu$. Therefore, the average energy is

$$\overline{E} = \frac{\int_0^\infty (8\pi\alpha/c^3)\nu^3 \, e^{-\beta\nu/T} \, d\nu}{\int_0^\infty (8\pi\alpha/c^3)(\nu^3/k_B\beta\nu) \, e^{-\beta\nu/T} \, d\nu} = k_B\beta \frac{\int_0^\infty \nu^3 \, e^{-\beta\nu/T} \, d\nu}{\int_0^\infty \nu^2 \, e^{-\beta\nu/T} \, d\nu}.$$

Integrating the denominator by parts,

$$\int_0^\infty \nu^2 \, e^{-\beta\nu/T} \, d\nu = \cancel{\left[\frac{\nu^3}{3} e^{-\beta\nu/T}\right]_0^\infty} + \frac{\beta}{3T}\int_0^\infty \nu^3 \, e^{-\beta\nu/T}, \tag{16.17}$$

and hence

$$\overline{E} = k_B\beta \times \frac{3T}{\beta} = 3k_BT. \tag{16.18}$$

The average energy of the quanta is closely related to the mean kinetic energy per particle in the black-body enclosure, $\frac{3}{2}k_BT$. This is a highly suggestive result.

So far, Einstein has stated that the radiation 'behaved as though' it consisted of a number of independent particles. Is this meant to be taken seriously, or is it just another 'formal device'? The last sentence of Section 6 of his paper leaves the reader in no doubt:

> ... the next obvious step is to investigate whether the laws of emission and transformation of light are also of such a nature that they can be interpreted or explained by considering light to consist of such energy quanta.

In other words, 'Yes, let us assume that they are real particles and see whether or not we can understand other phenomena'.

Einstein considers three phenomena which cannot be explained by classical electromagnetic theory: (i) Stokes' rule of photoluminescence, (ii) the photoelectric effect, and (iii) ionisation of gases by ultraviolet light.

Stokes' rule is the observation that the frequency of photoluminescent emission is less than the frequency of the incident light. This is explained as a consequence of the conservation of energy. If the incoming quanta have energy $h\nu_1$, the re-emitted quanta can at most have this energy. If some of the energy of the quanta is absorbed by the material before re-emission, the emitted quanta of energy $h\nu_2$ must have $h\nu_2 \leq h\nu_1$.

The photoelectric effect. This is probably the most famous result of the paper because Einstein makes a definite quantitative prediction on the basis of the theory expounded above. Ironically, the photoelectric effect had been discovered by Heinrich Hertz in 1887 in the same set of experiments which fully validated Maxwell's equations. Perhaps the most remarkable feature of the photoelectric effect was Lénard's discovery that the energies of the electrons emitted from the metal surface are independent of the intensity of the incident radiation.

Einstein's proposal provided an immediate solution to this problem. Radiation of a given frequency consists of quanta of the same energy $h\nu$. If one of these is absorbed by the material, the electron may receive sufficient energy to remove it from the surface against the forces which bind it to the material. If the intensity of the light is increased, more electrons are ejected, but their energies remain unchanged. Einstein wrote this result as follows. The maximum kinetic energy which the ejected electron can have, E_{k}, is

$$E_{\mathrm{k}} = h\nu - W, \tag{16.19}$$

where W is the amount of work necessary to remove the electron from the surface of the material, its *work function*. Experiments to estimate the magnitude of the work function involve placing the photocathode in an opposing potential so that, when the potential reaches some value V, the ejected electrons can no longer reach the anode and the photoelectric current falls to zero. This occurs at the potential at which $E_{\mathrm{k}} = eV$. Therefore,

$$V = \frac{h}{e}\nu - \frac{W}{e}. \tag{16.20}$$

In Einstein's words,

> If the formula derived is correct, then V must be a straight line function of the frequency of the incident light, when plotted in Cartesian coordinates, whose slope is independent of the nature of the substance investigated.

Thus, the quantity h/e, the ratio of Planck's constant to the electronic charge, can be found directly from the slope of this relation. These were remarkable predictions because nothing was known at that time about the dependence of the photoelectric effect upon the frequency of the incident radiation. After ten years of difficult experimentation, all aspects of Einstein's equation were fully confirmed experimentally. In 1916, Millikan was able to summarise the results of his extensive experiments:

> Einstein's photoelectric equation has been subjected to very searching tests and it appears in every case to predict exactly the observed results.[12]

Photoionisation of gases. The third piece of experimental evidence discussed in Einstein's paper was the fact that the energy of each photon has to be greater than the ionisation potential of the gas if photoionisation is to take place. He showed that the smallest energy quanta for the ionisation of air were approximately equal to the ionisation potential determined independently by Stark. Once again, the quantum hypothesis was in agreement with experiment.

At this point, the paper ends. It is one of the great papers of physics and is the work described in Einstein's Nobel Prize citation.

16.3 The Quantum Theory of Solids

In 1905, Planck and Einstein had somewhat different views about the role of quantisation and quanta. Planck had quantised the oscillators and there is no mention of this approach in Einstein's paper. Indeed, it seems that Einstein was not at all clear that they were actually describing the same phenomenon, but in 1906 he showed that the two approaches were, in fact, the same.[13] Then, in a paper submitted to *Annalen der Physik* in November 1906 and published in 1907,[14] he came to the same conclusion by a different argument, and went on to extend the idea of quantisation to solids.

In the first of these papers, Einstein asserts that he and Planck are actually describing the same phenomena of quantisation.

> At that time [1905] it seemed to me as though Planck's theory of radiation formed a contrast to my work in a certain respect. New considerations which are given in the first section of this paper demonstrate to me, however, that the theoretical foundation on which Planck's radiation theory rests differs from the foundation that would result from Maxwell's theory and the electron theory and indeed differs exactly in that Planck's theory implicitly makes use of the hypothesis of light quanta just mentioned.[13]

These arguments are developed further in the paper of 1907.[14] Einstein demonstrated that, if Planck had followed the proper Boltzmann procedure, he would have obtained the correct formula for black-body radiation, whilst maintaining the assumption that the oscillator can only take definite energies, $0, \epsilon, 2\epsilon, 3\epsilon, \ldots$

Let us repeat Einstein's argument, which appears in all the standard textbooks. As demonstrated in Section 15.2, Boltzmann's expression for the probability that a state of energy $E = r\epsilon$ is occupied is

$$p(E) \propto \mathrm{e}^{-E/k_\mathrm{B}T}.$$

The energy of the oscillator is assumed to be quantised in units of ϵ. If there are N_0 oscillators in the ground state, the number in the $r = 1$ state is $N_0\,\mathrm{e}^{-\epsilon/k_\mathrm{B}}$, in the $r = 2$ state $N_0\mathrm{e}^{-2\epsilon/k_\mathrm{B}T}$, and so on. Therefore, the average energy of the oscillator is

$$\begin{aligned}\overline{E} &= \frac{N_0 \times 0 + \epsilon N_0\,\mathrm{e}^{-\epsilon/k_\mathrm{B}T} + 2\epsilon N_0\,\mathrm{e}^{-2\epsilon/k_\mathrm{B}T} + \cdots}{N_0 + N_0\,\mathrm{e}^{-\epsilon/k_\mathrm{B}T} + N_0\,\mathrm{e}^{-2\epsilon/k_\mathrm{B}T} + \cdots} \\ &= \frac{N_0\epsilon\,\mathrm{e}^{-\epsilon/k_\mathrm{B}T}\left[1 + 2(N_0\,\mathrm{e}^{-\epsilon/k_\mathrm{B}T}) + 3(N_0\,\mathrm{e}^{-\epsilon/k_\mathrm{B}T})^2 + \cdots\right]}{N_0\left[1 + \mathrm{e}^{-\epsilon/k_\mathrm{B}T} + (\mathrm{e}^{-\epsilon/k_\mathrm{B}T})^2 + \cdots\right]}.\end{aligned} \tag{16.21}$$

We recall the following series:

$$\begin{cases} \dfrac{1}{1-x} &= 1 + x + x^2 + x^3 + \cdots, \\ \dfrac{1}{(1-x)^2} &= 1 + 2x^2 + 3x^3 + \cdots, \end{cases} \tag{16.22}$$

and hence the mean energy of the oscillator is

$$\overline{E} = \frac{\epsilon\,\mathrm{e}^{-\epsilon/k_\mathrm{B}T}}{1 - \mathrm{e}^{-\epsilon/k_\mathrm{B}T}} = \frac{\epsilon}{\mathrm{e}^{\epsilon/k_\mathrm{B}T} - 1}. \tag{16.23}$$

Thus, using the proper Boltzmann procedure, Planck's relation for the mean energy of the oscillator is recovered, provided the energy element ϵ does not vanish. Einstein's approach indicates clearly the origin of the departure from the classical result. The mean energy $\overline{E} = k_{\mathrm{B}}T$ is recovered in the classical limit $\epsilon \to 0$ from (16.23). Notice that, by allowing $\epsilon \to 0$, the averaging takes place over a continuum of energies which the oscillator might take. To put it another way, it is assumed that equal volumes of phase space are given equal weights in the averaging process, and this is the origin of the classical equipartition theorem. Einstein showed that Planck's formula requires that this assumption is wrong. Rather, only those volumes of phase space with energies $0, \epsilon, 2\epsilon, 3\epsilon, \ldots$ should have non-zero weights and these should all be equal.

Einstein then relates this result directly to his previous paper on light quanta:

> ... we must assume that for ions which can vibrate at a definite frequency and which make possible the exchange of energy between radiation and matter, the manifold of possible states must be narrower than it is for the bodies in our direct experience. We must in fact assume that the mechanism of energy transfer is such that the energy can assume only the values $0,\ \epsilon,\ 2\epsilon,\ 3\epsilon \ldots$.[15]

But this is only the beginning of the paper. Much more is to follow – Einstein puts it beautifully:

> I now believe that we should not be satisfied with the result. For the following question forces itself upon us. If the elementary oscillators that are used in the theory of the energy exchange between radiation and matter cannot be interpreted in the sense of the present kinetic molecular theory, must we not also modify the theory for the other oscillators that are used in the molecular theory of heat? There is no doubt about the answer, in my opinion. If Planck's theory of radiation strikes to the heart of the matter, then we must also expect to find contradictions between the present kinetic molecular theory and experiment in other areas of the theory of heat, contradictions that can be resolved by the route just traced. In my opinion, this is actually the case, as I try to show in what follows.[16]

This paper is often referred to as Einstein's application of the quantum theory to solids, but, just as in the case of his paper on the photoelectric effect, this paper is very much deeper than this name suggests and strikes right at the heart of the quantum nature of matter and radiation.

The problem discussed by Einstein concerns the heat capacities of solids. According to the *Dulong and Petit law*, the heat capacity per mole of a solid is about $3R$. This result can be derived simply from the equipartition theorem. The model of the solid consists of N_{A} atoms per mole and it is supposed that they can all vibrate in the three independent directions, x, y, z. According to the equipartition theorem, the internal energy per mole of the solid should therefore be $3N_{\mathrm{A}}k_{\mathrm{B}}T$, since each independent mode of vibration is awarded an energy $k_{\mathrm{B}}T$. The heat capacity per mole follows directly by differentiation: $C = \partial U/\partial T = 3N_{\mathrm{A}}k_{\mathrm{B}}T = 3R$.

It was known that some materials do not obey the Dulong and Petit law in that they have significantly smaller heat capacities than $3R$ – this was particularly true for light elements such as carbon, boron and silicon. In addition, by 1900, it was known that the heat capacities of some elements change rapidly with temperature and only attain the value $3R$ at high temperatures.

The problem is readily solved if Einstein's quantum hypothesis is adopted. For oscillators, the classical formula for the average energy of the oscillator $k_B T$ should be replaced by the quantum formula

$$\overline{E} = \frac{h\nu}{e^{h\nu/k_B T} - 1}.$$

Atoms are complicated systems, but let us suppose for simplicity that, for a particular material, they all vibrate at the same frequency, the *Einstein frequency* ν_E, and that these vibrations are independent. Since each atom has three independent modes of vibration, the internal energy is

$$U = 3N_A \frac{h\nu_E}{e^{h\nu_E/k_B T} - 1}, \tag{16.24}$$

and the heat capacity is

$$\begin{aligned}\frac{dU}{dT} &= 3N_A h\nu_E (e^{h\nu_E/k_B T} - 1)^{-2} e^{h\nu_E/k_B T} \frac{h\nu_E}{k_B T^2} \\ &= 3R \left(\frac{h\nu_E}{k_B T}\right)^2 \frac{e^{h\nu_E/k_B T}}{(e^{h\nu_E/k_B T} - 1)^2}.\end{aligned} \tag{16.25}$$

This expression turned out to provide a remarkably good fit to the measured variation of the heat capacity with temperature. In his paper, Einstein compared the experimentally determined variation of the heat capacity of diamond with his formula, with the results shown in Fig. 16.1. The decrease in the heat capacity at low temperatures is apparent, although the experimental points lie slightly above the predicted relation at low temperatures.

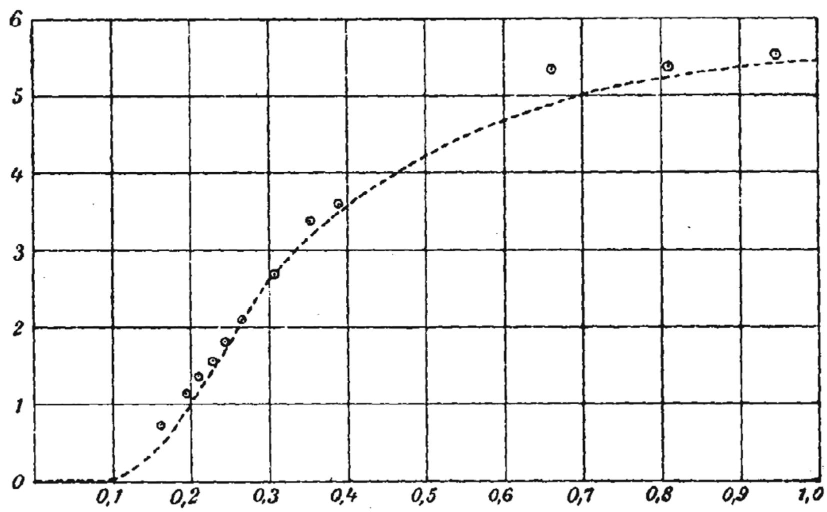


Fig. 16.1 The variation of the heat capacity of diamond with temperature compared with the prediction of Einstein's quantum theory. The abscissa is T/Θ_E, where $k_B \Theta_E = h\nu_E$, and the ordinate the molar heat capacity in calories mole^{-1}. This diagram appears in Einstein's paper of 1907[14] and uses the results of H.F. Weber which were in the tables of Landolt and Börnstein.

We can now understand why light elements have smaller heat capacities than the heavier elements. Presumably, the lighter elements have higher vibrational frequencies than heavier elements and hence, at a given temperature, ν_E/T is larger and the heat capacity is smaller. To account for the experimental data shown in Fig. 16.1, the frequency ν_E must lie in the infrared waveband. As a result, all vibrations at higher frequencies make only a vanishingly small contribution to the heat capacity. As expected, there is strong absorption at infrared wavelengths corresponding to frequencies $\nu \sim \nu_E$. Einstein compared his estimates of ν_E with the strong absorption features observed in a number of materials and found remarkable agreement, granted the simplicity of the model.

Perhaps the most important prediction of the theory was that the heat capacities of all solids should decrease to zero at low temperatures, as indicated in Fig. 16.1. This was of key importance from the point of view of furthering the acceptance of Einstein's ideas, because, at about this time, Walther Nernst began a series of experiments to measure the heat capacities of solids at low temperatures. Nernst's motivation for undertaking these experiments was to test his *heat theorem*, or the *third law of thermodynamics*, which he had developed theoretically in order to understand the nature of chemical equilibria. The heat theorem enabled the calculation of chemical equilibria to be carried out precisely and also led to the prediction that the heat capacities of all materials should tend to zero at low temperatures. As recounted by Frank Blatt,

> ... shortly after Einstein had assumed a junior faculty position at the University of Zurich [in 1909], Nernst paid the young theorist a visit so that they could discuss problems of common interest. The chemist George Hevesy ... recalls that among his colleagues it was this visit by Nernst that raised Einstein's reputation. He had come as an unknown man to Zurich. Then, Nernst came, and the people at Zurich said, 'This Einstein must be a clever fellow, if the great Nernst comes so far from Berlin to Zurich to talk to him'.[17]

16.4 Debye's Theory of Specific Heat Capacities

Although Einstein took little interest in the heat capacities of solids after 1907, the role of quantisation in solids was taken significantly further by Peter Debye in an important paper of 1912.[18] Einstein was well aware of the fact that the assumption that the atoms of a solid vibrate independently was a crude approximation. Debye took the opposite approach of returning to a continuum picture, almost identical to that developed by Rayleigh in his treatment of the spectrum of black-body radiation (Section 14.6). Debye realised that the collective modes of vibration of a solid could be represented by the complete set of normal modes which follows from fitting waves in a box, as described by Rayleigh. Each independent mode of oscillation of the solid as a whole should be awarded an energy

$$\overline{E} = \frac{\hbar\omega}{\exp(\hbar\omega/k_B T) - 1}, \tag{16.26}$$

according to Einstein's prescription, where ω is the angular frequency of vibration of the mode. The number of modes $\mathcal{N}$ had been evaluated by Rayleigh according to the procedure shown in Fig. 14.5, which resulted in the expression (14.49). In the case of the modes of vibration of a solid, the expression becomes

$$\mathrm{d}\mathcal{N} = \frac{L^3\omega^2}{2\pi^2 c_s^3}\,\mathrm{d}\omega, \tag{16.27}$$

where c_s is now the speed of propagation of the waves in the material. Just as in the case of electromagnetic radiation, we need to determine how many independent polarisation states there are for the waves. In this case, there are two transverse modes and one longitudinal mode, corresponding to the independent directions in which the material can be stressed by the wave, and so in total there are $3\,\mathrm{d}\mathcal{N}$ modes, each of which is awarded the energy (16.26). Debye made the assumption that these modes have the same speed of propagation and that it is independent of the frequency of the modes. Therefore, the total internal energy of the material is

$$\begin{aligned} U &= \int_0^{\omega_{\max}} \frac{\hbar\omega}{\exp(\hbar\omega/k_B T) - 1}\,3\,\mathrm{d}\mathcal{N} \\ &= \frac{3}{2\pi^2}\left(\frac{k_B T L}{\hbar c_s}\right)^3 \int_0^{x_{\max}} \frac{x^3}{e^x - 1}\,\mathrm{d}x, \end{aligned} \tag{16.28}$$

where $x = \hbar\omega/k_B T$.

The problem is now to determine the value of $x_{\max}$. Debye introduced the idea that there must be a limit to the total number of modes in which energy could be stored. In the high temperature limit, he argued that the total energy should not exceed that given by the classical equipartition theorem, namely $3Nk_BT$. Since each mode of oscillation has energy k_BT in this limit, there should be a maximum of $3N$ modes in which energy is stored. Therefore, Debye's condition can be found by integrating (16.27):

$$\begin{aligned} 3N = 3\int_0^{\omega_{\max}} \mathrm{d}\mathcal{N} &= 3\int_0^{\omega_{\max}} \frac{L^3\omega^3}{2\pi^2 c_s^3}\,\mathrm{d}\omega, \\ \omega_{\max}^3 &= \frac{6\pi^2 N}{L^3} c_s^3. \end{aligned} \tag{16.29}$$

It is conventional to write $x_{\max} = \hbar\omega_{\max}/k_B T = \theta_D/T$, where θ_D is known as the *Debye temperature*. Therefore, the expression for the total internal energy of the material (16.28) can be rewritten

$$U = 9RT\left(\frac{T}{\theta_D}\right)^3 \int_0^{\theta_D/T} \frac{x^3}{e^x - 1}\,\mathrm{d}x \tag{16.30}$$

for one mole of the material. This is the famous expression derived by Debye for the internal energy per mole of the solid.

To find the heat capacity, it is simplest to consider in infinitesimal increment $\mathrm{d}U$ associated with a single frequency ω and then integrate over x as before:

$$C = \frac{\mathrm{d}U}{\mathrm{d}T} = 9R\left(\frac{T}{\theta_D}\right)^3 \int_0^{\theta_D/T} \frac{x^4}{(e^x - 1)^2}\,\mathrm{d}x. \tag{16.31}$$

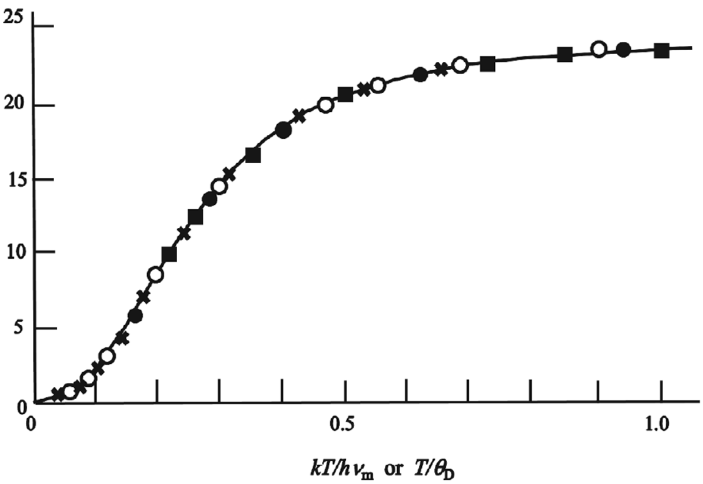


Fig. 16.2 The molar heat capacity for various solids as a function of $k_B T/\hbar\omega_D$, where ω_D is the angular Debye frequency, compared with the predictions of the Debye theory. θ_D is the Debye temperature. The substances plotted are copper (open circles: $\theta_D = 315$ K), silver (filled circles: $\theta_D = 215$ K), lead (filled squares: $\theta_D = 88$ K) and carbon (crosses: $\theta_D = 1860$ K). (After D. Tabor, 1991, *Gases, Liquids and Solids and Other States of Matter*, p. 236. Cambridge: Cambridge University Press.)

This integral cannot be written in closed form, but it provides an even better fit to the data on the heat capacity of solids than Einstein's expression (16.25), as can be seen from Fig. 16.2.

As pointed out by Tabor,[19] there are two big advantages of Debye's theory. First, from (16.29), the value of the Debye temperature can be worked out from the bulk properties of the material by measuring c_s, whereas the Einstein temperature is an arbitrary parameter to be found by fitting the data. Second, the expression (16.31) for the heat capacity gives a much improved fit to the data at low temperatures, as can be seen from Fig. 16.2. If $T \ll \theta_D$, the upper limit to the integral in (16.31) can be set to infinity, and then the integral has the value $4\pi^4/15$. Therefore, at low temperatures $T \ll \theta_D$, the heat capacity is

$$C = \frac{dU}{dT} = \frac{12\pi^4}{5} R \left(\frac{T}{\theta_D}\right)^3. \tag{16.32}$$

Rather than decreasing exponentially at low temperatures, the heat capacity varies as T^3.

Two other points are of interest. First of all, for many solids, it turns out that $\theta_E \approx 0.75\theta_D$, $\nu_E \approx 0.75\nu_D$. Secondly, there is a simple interpretation of ω_{max} in terms of wave propagation in the solid. From (16.29), the maximum frequency is

$$\nu_{max} = \left(\frac{3}{4\pi}\right)^{1/3} \left(\frac{N_A}{L^3}\right)^{1/3} c_s, \tag{16.33}$$

for one mole of the solid, where N_A is Avogadro's number. But $L/N_A^{1/3}$ is just the typical interatomic spacing, a. Therefore, (16.33) states that $\nu_{max} \approx c_s/a$, that is, the minimum wavelength of the waves $\lambda_{\rm min} = c_{\rm s}/\nu_{\rm max} \sim a$. This makes a great deal of sense physically. On scales less than the interatomic spacing a, the concept of collective vibration of the atoms of the material ceases to have any meaning.

16.5 The Specific Heat Capacities of Gases Revisited

Einstein's quantum theory solves the problem of accounting for the specific heats of diatomic gases, which had caused Maxwell such distress (Section 12.4). According to Einstein, all energies associated with atoms and molecules should be quantised and then the energy states which can be excited depend upon the temperature of the gas. Einstein's analysis of Sections 16.2 and 16.3 showed that the average energy of an oscillator of frequency ν at temperature T is

$$\overline{E} = \frac{h\nu}{{\rm e}^{h\nu/k_{\rm B}T} - 1}. \qquad (16.34)$$

If $k_{\rm B}T \gg h\nu$, the average kinetic energy is $k_{\rm B}T$, as expected from kinetic theory. If however $k_{\rm B}T \ll h\nu$, the average energy tends to $h\nu\,{\rm e}^{-h\nu/k_{\rm B}T}$, which is negligibly small.

Let us therefore look again at the various modes by which energy can be stored by the molecules of a gas.

16.5.1 Linear Motion

For atoms or molecules confined to a volume of, say, a cubic metre, the energy levels are quantised, but their spacing is absolutely tiny. To anticipate de Broglie's brilliant insight of 1924, a wavelength λ can be associated with the momentum p of a particle, such that $p = h/\lambda$, where λ is known as the *de Broglie wavelength* and h is Planck's constant. Therefore, quantum effects only become important on the metre scale at energies $\epsilon = p^2/2m \sim h^2/2m\lambda^2 = 3 \times 10^{-41}$ J. Setting ϵ equal to $\frac{3}{2}k_{\rm B}T$, we find $T = 10^{-18}$ K. Thus, for all accessible temperatures, $k_{\rm B}T \gg \epsilon$ and we need not worry about the effects of quantisation on the linear motion of molecules. Therefore, for the atoms or molecules of a gas, there are always three degrees of freedom associated with their linear motion.

16.5.2 Molecular Rotation

Energy can also be stored in the rotational motion of the molecule. The amount of energy which can be stored can be found from a basic result of quantum mechanics, that orbital angular momentum J is quantised such that only discrete values are allowed according to the relation

$$J = [j(j+1)]^{1/2}\hbar,$$

where $j = 0, 1, 2, 3, 4, \ldots$ and $\hbar = h/2\pi$. The rotational energy of the molecule is given by the analogue of the classical relation between energy and angular momentum:

$$E = \tfrac{1}{2}J^2/I,$$

where I is the moment of inertia of the molecule about the rotation axis. Therefore,

$$E = \frac{j(j+1)}{2I}\hbar^2. \tag{16.35}$$

If we write the rotational energy of the molecule as $E = \frac{1}{2}I\omega^2$, the angular frequency of rotation ω is

$$\omega = [j(j+1)]^{1/2}\hbar/I.$$

The first rotational state of the molecule corresponds to $j = 1$ and so the first excited state has energy $E = \hbar^2/I$. The condition that rotational motion contributes to the storage of energy at temperature T is

$$k_B T \geq E = \hbar^2/I,$$

which is equivalent to $k_B T \geq \hbar\omega$.

Let us consider the cases of hydrogen and chlorine molecules and evaluate the temperatures at which the energies of the first excited state, $j = 1$, is equal to $k_B T$. Table 16.1 shows the results of these calculations.

Thus, for hydrogen, at room temperature, energy can be stored in its rotational degrees of freedom, but at low temperatures, $T \ll 173$ K, the rotational degrees of freedom are 'frozen out', or 'quenched', and do not contribute to the heat capacity of the gas. For the case of chlorine, the temperature at which $E_1 = k_B T$ is so low that the rotational degrees of freedom contribute to the heat capacity at all temperatures at which it is gaseous.

There is one further key issue. For a rotating *diatomic molecule*, how many of its degrees of freedom are actually excited? Figure 16.3 shows rotational motion about the three perpendicular axes, with angular frequencies ω_x, ω_y and ω_z. The calculation performed above refers to the moments of inertia about the y and z axes. It is certainly *not* the correct calculation, however, for the moment of inertia about the x-axis. All the mass of the atoms is concentrated in their nuclei and so the moment of inertia about the x-axis is absolutely tiny compared with those about the y and z axes. In fact, it is about 10^{10} times smaller than that about the other two axes. Therefore, the temperature needed to excite rotation about the x-axis is about 10^{10} times greater than the others and so it is not excited at the temperatures at which the molecules can exist.

Thus, for diatomic molecules such as hydrogen and oxygen, there are only two rotational degrees of freedom which can be excited at room temperatures. Therefore, five degrees

Table 16.1 Properties of the diatomic molecules hydrogen and chlorine

Property	Hydrogen	Chlorine
Mass of atom (m/amu)	1	35
Bond length (d/nm)	0.0746	0.1988
Moment of inertia ($I = \frac{1}{2}md^2$/kg m^2)	4.65×10^{-48}	1.156×10^{-45}
Energy (E_1/J)	2.30×10^{-21}	9.62×10^{-24}
$E_1 = k_B T$ at	173 K	0.697 K

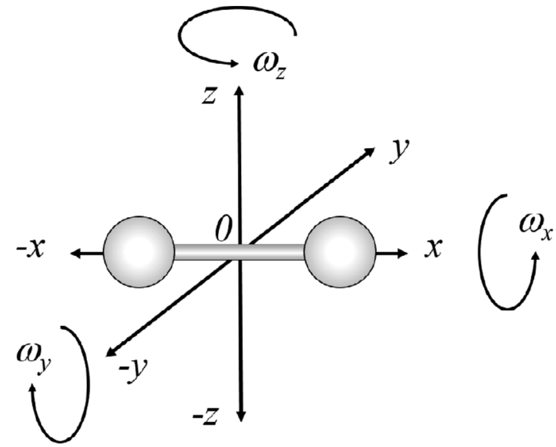


Fig. 16.3 Illustrating the rotational modes by which angular momentum can be stored in a linear diatomic molecule.

of freedom f in total are excited, three translational and two rotational, and the ratio of heat capacities is expected to be $\gamma = (f + 2)/f = 1.4$ and $C_V = 2.5R$. This solves Maxwell's problem.

This argument can be extended to molecules which have finite moments of inertia about all three axes of comparable magnitude and then there would be three rotational degrees of freedom and $\gamma = 1.33$, as found by Maxwell.

16.5.3 Molecular Vibration

To complete the analysis, consider the quantisation of the vibrational modes of molecules. A diatomic molecule behaves like a linear oscillator and so has two degrees of freedom, one associated with its kinetic energy and the other with its potential energy. As in the case of rotational motion, however, energy can only be stored in these degrees of freedom if the vibrational states of the molecule can be excited. Thus, we need to know the vibrational frequencies of the molecules in their ground states. Let us consider again the cases of hydrogen and chlorine molecules.

In the case of hydrogen, its vibrational frequency is about 2.6×10^{14} Hz and so $\epsilon = h\nu = 1.72 \times 10^{-19}$ J. Setting $h\nu = k_B T$, $T = 12\,500$ K and so, for hydrogen, the vibrational modes cannot contribute to the heat capacity at room temperature. In fact, the vibrational modes scarcely come into play at all, because hydrogen molecules are dissociated at about 2000 K.

In the case of chlorine, the atoms are much heavier and the vibrational frequency is about an order of magnitude smaller. Thus, the vibrational modes come into play at temperatures above about 600 K. At high temperatures, the degrees of freedom of the chlorine molecule which come into play are

$$3 \text{ translational} + 2 \text{ rotational} + 2 \text{ vibrational} = 7 \text{ degrees of freedom.}$$

Thus, the heat capacity of chlorine at very high temperatures tends to $\frac{7}{2}R$.

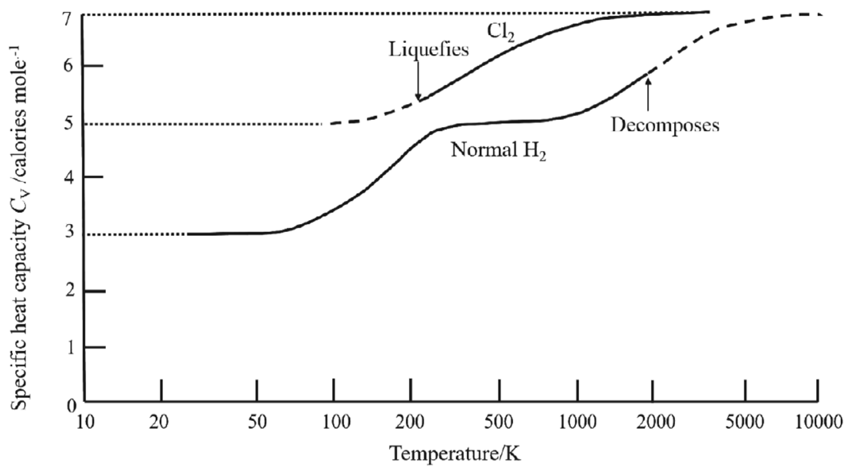


Fig. 16.4 The variation of the molar heat capacity with temperature for molecular hydrogen and chlorine. The units on the ordinate are calories per mole. One calorie = 4.184 J. Therefore, $\frac{3}{2}R_0 = \frac{3}{2} \times 8.31$ J mole^{-1} = 12.465 J mole^{-1} = 3.0 calories mole^{-1}. (After D. Tabor, 1991, *Gases, Liquids and Solids and Other States of Matter.*, p. 110. Cambridge: Cambridge University Press.)

16.5.4 The Cases of Hydrogen and Chlorine

We are now in a position to understand the details of Fig. 16.4, which shows the variation of the heat capacities of hydrogen and chlorine with temperature.

For hydrogen at low temperatures, only the translational degrees of freedom can store energy and the heat capacity is $\frac{3}{2}R_0$. Above about 170 K, the rotational modes can store energy and so five degrees of freedom come into play. Above 2000 K, the vibrational modes begin to play a role, but the molecules are dissociated at about 2000 K before these modes can come completely into play.

For chlorine, even at low temperatures, both the translational and rotational modes are excited and so the heat capacity is $\frac{5}{2}R_0$, although, as noted in the diagram, chlorine begins to liquefy between 200 and 300 K. Above 600 K, the vibrational modes begin to play a role and at the highest temperatures, the heat capacity becomes $\frac{7}{2}R_0$. The corresponding values for the ratios of heat capacities are found from the rule $\gamma = (C_V + R)/C_V$.

Thus, Einstein's quantum theory can account in detail for the properties of ideal gases and solve completely the problems which had frustrated Maxwell and Boltzmann.

16.6 Conclusion

Einstein's revolutionary ideas on quanta and quantisation predate the discovery of wave and quantum mechanics by Schrödinger and Heisenberg in the mid-1920s by about 20 years. Consequently, to today's physicist, Einstein's arguments appear schematic, rather than formally precise in the sense of being derived from a fully developed theory of quantum mechanics. New concepts and approaches were needed to provide a complete description of quantum mechanical systems. Despite these technical problems, Einstein's fundamental insights were indeed revolutionary in opening up completely new realms for physics – all physical processes are basically quantum in nature, despite the fact that classical physics is successful in explaining so much.

Notes

1 Einstein, A. (1905*a*). *Annalen der Physik*, **17**, 547.
2 Einstein, A. (1905*b*). *Annalen der Physik*, **17**, 891.
3 Einstein, A. (1905*c*). *Annalen der Physik*, **17**, 132.
4 Stachel, J. (ed.) (1998). *Einstein's Miraculous Year: Five Papers That Changed the Face of Physics*. Princeton: Princeton University Press.
5 Einstein, A. (1900–1909). *The Collected Papers of Albert Einstein: Volume 2. The Swiss Years: Writings, 1900–1909*, eds. J. Stachel and D.C. Cassidy. (English translation supplement) Princeton: Princeton University Press.
6 Einstein, A. (1906). *Annalen der Physik*, **19**, 371.
7 Einstein, A. (1979). *Autobiographical Notes*, trans. and ed. P.A. Phillips, pp. 44–45. La Salle, Illinois; Open Court.
8 Perrin, J. (1909). *Annales de Chimie et de Physique*, **18**, 1.
9 Perrin, J. (1910). *Brownian Movement and Molecular Reality*, trans. F. Soddy. London: Taylor and Francis.
10 Einstein, A. (1905*d*). *Annalen der Physik*, **17**, 639.
11 Einstein, A. (1995). *The Collected Papers of Albert Einstein: Volume 5.*, p. 20. (English translation supplement) Princeton: Princeton University Press.
12 Millikan, (1916). *Physical Review*, **7**, 18.
13 Einstein, A. (1906). *Annalen der Physik*, **20**, 199.
14 Einstein, A. (1907). *Annalen der Physik*, **22**, 180.
15 Einstein, A. (1907). *op. cit.*, pp. 183–184.
16 Einstein, A. (1907). *op. cit.*, p. 184.
17 Blatt, F.J. (1992). *Modern Physics*. New York: McGraw-Hill.
18 Debye, P. (1912). *Annalen der Physik*, **39**, 789.
19 From D. Tabor (1991). *Gases, Liquids and Solids and Other States of Matter*, p. 236. Cambridge: Cambridge University Press.

17 The Triumph of the Light Quantum Hypothesis

17.1 The Situation in 1909

In no sense were Einstein's startling new ideas on light quanta immediately accepted by the scientific community at large. Most of the major figures in physics rejected the idea that light could be considered to be made up of discrete quanta. In a letter to Einstein in 1907, Planck wrote:

> I look for the significance of the elementary quantum of action (light quantum) not in vacuo but rather at points of absorption and emission and assume that processes in vacuo are accurately described by Maxwell's equations. At least, I do not yet find a compelling reason for giving up this assumption which for the time being seems to be the simplest.[1]

Planck continued to reject the light quantum hypothesis as late as 1913. In 1909, Lorentz, who was widely regarded as the leading theoretical physicist in Europe and whom Einstein held in the highest esteem, wrote:

> While I no longer doubt that the correct radiation formula can only be reached by way of Planck's hypothesis of energy elements, I consider it highly unlikely that these energy elements should be considered as light quanta which maintain their identity during propagation.[2]

Einstein never deviated from his conviction concerning the reality of quanta and continued to find other ways in which the experimental features of black-body radiation lead inevitably to the conclusion that light consists of quanta. In one of his most inspired papers, written in 1909,[3] he showed how fluctuations in the intensity of the black-body radiation spectrum provide further evidence for the quantum nature of light. Let us first review some elementary ideas concerning the statistical fluctuations of particles and waves as a prelude to reviewing Einstein's paper.

17.2 Fluctuations of Particles and Waves

17.2.1 Particles in a Box

Consider first the statistics of particles in a box. The box is divided into N cells and a large number of particles n is distributed randomly among them. The numbers of particles in each cell are counted. If n is very large, the mean number of particles in each cell is roughly the same, but there is a real scatter about the mean value because of statistical fluctuations.

In the simplest case of tossing coins, we ask, ‘What is the probability of obtaining x heads (or successes) in n throws?’ In each throw, the probability of success $p = \frac{1}{2}$ and that of failure $q = \frac{1}{2}$, so that $p + q = 1$. If we throw two coins, the possible outcomes of the experiment are

HH, HT, TH, TT.

If we toss three coins, the possible combinations are

HHH, HHT, HTH, THH, HTT, THT, TTH, TTT.

Since the order is not important, the frequencies of occurrence of the different combinations for two throws are:

HH	H T	TT
1	2	1

or to probabilities of $\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$ for two coins. Similarly, for three throws,

HHH	HH T	H TT	TTT
1	3	3	1

or probabilities $\frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8}$. These probabilities correspond to the coefficients in t of the binomial expansions of $(p + qt)^2$ and $(p + qt)^3$ with $p = q = \frac{1}{2}$.

Let us derive this result more generally. We ask, ‘What is the probability of tossing one head and then two tails?’ The answer is $pq^2 = \frac{1}{8}$ for this particular ordering. However, if the order does not matter, we need to know the number of ways in which we could have obtained one success and two failures. This is exactly the same problem we discussed in connection with the derivation of the Boltzmann distribution (Section 15.2). The number of different ways of selecting x and y identical objects from n is $n!/x!\,y!$. In the present case, $y = n - x$ and hence the answer is $n!/x!\,(n-x)!$. Thus, the number of ways is $3!/1!\,2! = 3$, that is, the total probability is $\frac{3}{8}$ in agreement with above. Thus, in general, the probability of x successes out of n events is $n!/x!\,(n-x)!\; p^x q^{n-x}$, which is just the coefficient of the term in t^x in the expansion of $(q + pt)^n$. Hence, the probability of obtaining x successes in n attempts, $P_n(x)$, is

$$P_n(x) = \frac{n!}{(n-x)!\,x!} p^x q^{n-x}, \tag{17.1}$$

and

$$(q + pt)^n = P_n(0) + P_n(1)t + P_n(2)t^2 + \cdots P_n(x)t^x + \cdots + P_n(n)t^n. \tag{17.2}$$

Setting $t = 1$, we obtain

$$1 = P_n(0) + P_n(1) + P_n(2) + \cdots P_n(x) + \cdots + P_n(n), \qquad (17.3)$$

showing that the total probability is 1, as it must be.

Now take the derivative of (17.2) with respect to t:

$$np(q + pt)^{n-1} = P_n(1) + 2P_n(2)t + \cdots xP_n(x)t^{x-1} + \cdots + nP_n(n)t^{n-1}. \qquad (17.4)$$

Setting $t = 1$,

$$pn = \sum_{x=0}^{n} xP_n(x). \qquad (17.5)$$

The quantity on the right-hand side of (17.5) is just the average value of x, that is, $\overline{x} = pn$, which agrees with our intuitive expectation.

Let us repeat the procedure to find the variance of the distribution. Taking the derivative of (17.4) with respect to t,

$$p^2n(n-1)(q + pt)^{n-2} = 2P_n(2) + \cdots (x-1)P_n(x)t^{x-2} + n(n-1)P_n(n)t^{n-2}. \qquad (17.6)$$

Again, setting $t = 1$, we find

$$p^2n(n-1) = \sum_{x=0}^{n} x(x-1)P_n(x) = \sum_{x=0}^{n} x^2P_n(x) - \sum_{x=0}^{n} xP_n(x) = \sum_{x=0}^{n} x^2P_n(x) - np, \qquad (17.7)$$

that is,

$$\sum_{x=0}^{n} x^2P_n(x) = np + p^2n(n-1). \qquad (17.8)$$

Now $\sum_{x=0}^{n} x^2P_n(x)$ is a measure of the variance of the distribution of x measured with respect to the origin rather than the mean. Fortunately, there is a rule which tells us how to measure the variance σ^2 with respect to the mean:

$$\sigma^2 = \sum_{x=0}^{n} x^2P_n(x) - \overline{x}^2. \qquad (17.9)$$

Therefore,

$$\sigma^2 = \sum_{x=0}^{n} x^2P_n(x) - (pn)^2 = np + p^2n(n-1) - (pn)^2 = np(1-p) = npq. \qquad (17.10)$$

Finally, we convert (17.1) from a discrete to a continuous distribution by a procedure carried out in all the standard textbooks on statistics. The result is the normal, or Gaussian, distribution $p(x)\,\mathrm{d}x$, which can be written

$$p(x)\,\mathrm{d}x = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) \mathrm{d}x. \qquad (17.11)$$

σ^2 is the variance and has the value npq; x is measured with respect to the mean value np. We have already met this distribution in the context of the derivation of the Maxwell–Boltzmann velocity distribution, the expression (12.40).

This is the answer we have been seeking. If the box is divided into N cells, the probability of a particle being in a single cell in one experiment is $p = 1/N$, $q = (1 - 1/N)$. The total number of particles is n. Therefore, the average number of particles per sub-box is n/N and the variance about this mean value, that is, the mean squared statistical fluctuation about the mean, is

$$\sigma^2 = \frac{n}{N}\left(1 - \frac{1}{N}\right). \tag{17.12}$$

If N large, $\sigma^2 = n/N$ and is just the average number of particles in each cell, that is, $\sigma = (n/N)^{1/2}$. This is the well-known result that, for large values of N, the mean is equal to the variance.

This is the origin of the useful rule that the fractional fluctuation about the average value is $1/M^{1/2}$, where M is the number of discrete objects counted. This is the statistical behaviour we expect of particles in a box.

17.3 Fluctuations of Randomly Superposed Waves

The random superposition of waves is different in important ways from the statistics of particles in a box. The electric field $\boldsymbol{E}$ measured at some point in space is the random superposition of the electric fields from N sources, where N is very large. For simplicity, we consider only propagation in the z-direction and only one of the two linear polarisations of the waves, E_x or E_y. We also assume that the frequencies ν and the amplitudes ξ of all the waves are the same, the only difference being their random phases. Then, the quantity $E_x^* E_x = |\boldsymbol{E}|^2$ is proportional to the Poynting vector flux density in the z-direction for the E_x component and so is proportional to the energy density of the radiation, where E_x^* is the complex conjugate of E_x. Summing all the waves and assuming they have random phases, $E_x^* E_x$ can be written

$$E_x^* E_x = \left(\xi \sum_k \mathrm{e}^{\mathrm{i}\phi_k}\right)^* \left(\xi \sum_j \mathrm{e}^{\mathrm{i}\phi_j}\right) = \xi^2 \left(\sum_k \mathrm{e}^{-\mathrm{i}\phi_k}\right)\left(\sum_j \mathrm{e}^{\mathrm{i}\phi_j}\right), \tag{17.13}$$

$$= \xi^2 \left(N + \sum_{j \neq k} \mathrm{e}^{\mathrm{i}(\phi_j - \phi_k)}\right) = \xi^2 \left[N + 2\sum_{j>k} \cos(\phi_j - \phi_k)\right]. \tag{17.14}$$

The average of the term in $\cos(\phi_j - \phi_k)$ is zero, since the phases of the waves are random and therefore

$$\langle E_x^* E_x \rangle = N\xi^2 \propto u_x. \tag{17.15}$$

This is a familiar result. For incoherent radiation, meaning for waves with random phases, the total energy density is equal to the sum of the energies in all the waves.

Let us now work out the fluctuations in the average energy density of the waves. We need to work out the quantity $\langle (E_x^* E_x)^2 \rangle$ with respect to the mean value (17.15). We recall that $\langle \Delta n^2 \rangle = \langle n^2 \rangle - \langle \overline{n} \rangle^2$ and hence

$$\Delta u_x^2 \propto \langle (E_x^* E_x)^2 \rangle - \langle E_x^* E_x \rangle^2. \tag{17.16}$$

Now

$$\begin{aligned}\langle (E_x^* E_x)^2 \rangle &= \xi^4 \left(N + \sum_{j \neq k} \mathrm{e}^{\mathrm{i}(\phi_j - \phi_k)} \right)^2 \\ &= \xi^4 \left(N^2 + 2N \sum_{j \neq k} \mathrm{e}^{\mathrm{i}(\phi_j - \phi_k)} + \sum_{l \neq m} \mathrm{e}^{\mathrm{i}(\phi_l - \phi_m)} \sum_{j \neq k} \mathrm{e}^{\mathrm{i}(\phi_j - \phi_k)} \right). \end{aligned} \tag{17.17}$$

The second term on the right-hand side again averages to zero because the phases are random. In the double sum in (17.17), most of the terms average to zero, because the phases are random, but not all of them: those terms for which $l = k$ and $m = j$ do not vanish. Recalling that $l = m$ is excluded from the summation, the matrix of l, m values that give non-vanishing contributions $(E_x^* E_x)^2$ is

	$l \rightarrow$					
m	–	2, 1	3, 1	4, 1	...	N, 1
$\downarrow$	1, 2	–	3, 2	4, 2	...	N, 2
	1, 3	2, 3	–	4, 3	...	N, 3
	1, 4	2, 4	3, 4	–	...	N, 4
	⋮	⋮	⋮	⋮	⋮	⋮
	1, N	2, N	3, N	4, N	...	–

There are clearly $N^2 - N$ terms and therefore we find that

$$\langle (E_x^* E_x)^2 \rangle = \xi^4 [N^2 + N(N-1)] \approx 2N^2 \xi^4, \tag{17.18}$$

since N is very large. Therefore, from (17.15) and (17.16), we find

$$\Delta u_x^2 \propto \langle (E_x^* E_x)^2 \rangle = 2N^2 \xi^4 - N^2 \xi^4 = N^2 \xi^4.$$

This is the key result:

$$\Delta u_x^2 = u_x^2, \tag{17.19}$$

that is, *the fluctuations in the energy density are of the same magnitude as the energy density of the radiation field itself*. This is a remarkable property of electromagnetic radiation and is the reason why phenomena such as interference and diffraction occur for incoherent radiation. Despite the fact that the radiation measured by a detector is a superposition of a large number of waves with random phases, the fluctuations in the fields are as large as the magnitude of the total intensity.

The physical meaning of this calculation is clear. The matrix of non-vanishing contributions to $(E_x^* E_x)^2$ is no more than the sum of all pairs of waves added separately. Every pair of waves of frequency ν interferes to produce fluctuations in intensity of the radiation $\Delta u \approx u$, that is, multiplying out an arbitrary pair of terms of (17.13) with $j \neq k$,

$$\xi^2 \sin(kz - \omega t) \sin(kz - \omega t + \phi) = \frac{\xi^2}{2}\{\cos\phi - \cos[2(kz - \omega t) + \phi]\}.$$

Notice that this analysis refers to waves of random phase ϕ and of a particular angular frequency ω, that is, what we would refer to as waves corresponding to a single mode.

Let us now study Einstein's analysis of fluctuations in black-body radiation.

17.4 Fluctuations in Black-Body Radiation

17.4.1 Einstein's Analysis of 1909

Einstein begins his paper of 1909 by reversing Boltzmann's relation between entropy and probability:

$$W = \mathrm{e}^{S/k_\mathrm{B}}. \tag{17.20}$$

Consider only the radiation in the frequency interval ν to $\nu + \mathrm{d}\nu$. As before, we write $\varepsilon = Vu(\nu)\,\mathrm{d}\nu$. Now divide the volume into a large number of cells and suppose that $\Delta\varepsilon_i$ is the fluctuation in the ith cell. Then the entropy of this cell is

$$S_i = S_i(0) + \left(\frac{\partial S}{\partial U}\right)\Delta\varepsilon_i + \tfrac{1}{2}\left(\frac{\partial^2 S}{\partial U^2}\right)(\Delta\varepsilon_i)^2 + \cdots, \tag{17.21}$$

But, averaging over all cells, we know that there is no net fluctuation, $\sum_i \Delta\varepsilon_i = 0$, and therefore

$$S = \sum S_i = S(0) + \tfrac{1}{2}\left(\frac{\partial^2 S}{\partial U^2}\right)\sum(\Delta\varepsilon_i)^2. \tag{17.22}$$

Therefore, using (17.20), the probability distribution for the fluctuations is

$$W \propto \exp\left[\frac{1}{2}\left(\frac{\partial^2 S}{\partial U^2}\right)\frac{\sum(\Delta\varepsilon_i)^2}{k_\mathrm{B}}\right]. \tag{17.23}$$

This is the sum of a set of normal distributions which, for any individual cell, can be written

$$W_i \propto \exp\left[-\frac{1}{2}\frac{(\Delta\varepsilon_i)^2}{\sigma^2}\right], \tag{17.24}$$

where the variance of the distribution is

$$\sigma^2 = \frac{k_\mathrm{B}}{-\left(\dfrac{\partial^2 S}{\partial U^2}\right)}. \tag{17.25}$$

Notice that we have obtained a physical interpretation for the second derivative of the entropy with respect to energy, which had played a prominent part in Planck's original analysis.

Let us now derive σ^2 for a black-body spectrum:

$$u(\nu) = \frac{8\pi h\nu^3}{c^3}\frac{1}{e^{h\nu/k_B T} - 1}. \tag{17.26}$$

Inverting (17.26),

$$\frac{1}{T} = \frac{k_B}{h\nu}\ln\left(\frac{8\pi h\nu^3}{c^3 u} + 1\right). \tag{17.27}$$

Now express this result in terms of the total energy in the cavity in the frequency interval ν to $\nu + d\nu$, $\varepsilon = Vu\,d\nu$. As before, $dS/dU = 1/T$ and we may identify ε with U. Therefore,

$$\frac{\partial S}{\partial \varepsilon} = \frac{k_B}{h\nu}\ln\left(\frac{8\pi h\nu^3}{c^3 u} + 1\right) = \frac{k_B}{h\nu}\ln\left(\frac{8\pi h\nu^3 V\,dV}{c^3\varepsilon} + 1\right),$$

$$\frac{\partial^2 S}{\partial \varepsilon^2} = -\frac{k_B}{h\nu}\frac{1}{\left(\dfrac{8\pi h\nu^3 V\,d\nu}{c^3\varepsilon} + 1\right)} \times \frac{8\pi h\nu^3 V\,d\nu}{c^3\varepsilon^2},$$

$$\frac{k_B}{\partial^2 S/\partial\varepsilon^2} = -\left(h\nu\varepsilon + \frac{c^3}{8\pi\nu^2 V\,d\nu}\,\varepsilon^2\right) = -\sigma^2. \tag{17.28}$$

In terms of fractional fluctuations,

$$\frac{\sigma^2}{\varepsilon^2} = \left(\frac{h\nu}{\varepsilon} + \frac{c^3}{8\pi\nu^2 V\,d\nu}\right). \tag{17.29}$$

Einstein noted that the two terms on the right-hand side of (17.29) have quite specific meanings. The first originates from the Wien part of the spectrum and, if we suppose the radiation consists of photons, each of energy $h\nu$, it corresponds to the statement that the fractional fluctuation in the intensity is just $1/N^{1/2}$ where N is the number of photons, that is,

$$\Delta N/N = 1/N^{1/2}. \tag{17.30}$$

As shown in Section 17.2.1, this is exactly the result expected if light consists of discrete particles.

The second term originates from the Rayleigh–Jeans part of the spectrum. We ask, 'How many independent modes are there in the box in the frequency range ν to $\nu + d\nu$?' We have already shown in Section 14.6 that there are $8\pi\nu^2 V\,d\nu/c^3$ modes (see (14.50)). We have also shown in Section 17.2.2 that the fluctuations associated with each wave mode have magnitude $\Delta\varepsilon^2 = \varepsilon^2$. When we add together randomly all the independent modes in the frequency interval ν to $\nu + d\nu$, we add their variances and hence

$$\frac{\langle \delta E^2\rangle}{E^2} = \frac{1}{N_{\text{mode}}} = \frac{c^3}{8\pi\nu^2 V\,d\nu},$$

which is exactly the second term on the right-hand side of (17.29).

Thus, the two parts of the fluctuation spectrum correspond to particle and wave statistics, the former originating from the Wien part of the spectrum and the latter from the Rayleigh–Jeans region. The amazing aspect of this formula for the fluctuations is that we recall that we add together the variances due to independent causes and the equation

$$\frac{\sigma^2}{\varepsilon^2} = \left(\frac{h\nu}{\varepsilon} + \frac{c^3}{8\pi\nu^2 V\,\mathrm{d}\nu}\right) \tag{17.31}$$

states that *we should add independently the 'wave' and 'particle' aspects of the radiation field to find the total magnitude of the fluctuations*. I regard this as a quite miraculous piece of theoretical physics.

17.4.2 Interlude: Nyquist's Theorem, Johnson Noise and the Detection of Electromagnetic Signals in the Presence of Noise

Before continuing the story, it is helpful to show how powerful Einstein's expression for the fluctuations in black-body radiation is in its application to understanding the *noise power* inevitably present in resistors and electronic circuits. In radio receiving systems, weak signals need to be detected in the presence of the thermal noise in the receiving system.

Nyquist's Theorem and Johnson Noise

Nyquist's theorem is a beautiful application of Einstein's expression for the mean energy per mode to the case of the spontaneous electrical fluctuations of energy in a resistor R at temperature T. We aim to derive the expression for the noise power generated in a transmission line which is terminated at either end by matched resistors R, that is, the wave impedance of the transmission line Z_0 for the propagation of electromagnetic waves is equal to R (Fig. 17.1). Suppose the circuit is placed in an enclosure in thermodynamic equilibrium at temperature T. Energy is shared equally among all the modes present in the system according to Einstein's prescription (16.23) that, per mode, the average energy is

$$\overline{E} = \frac{h\nu}{\mathrm{e}^{h\nu/k_\mathrm{B}T} - 1}, \tag{17.32}$$

where ν is the frequency of the mode. Therefore, we need to know how many modes there are associated with the transmission line and award each of them energy $\overline{E}$ according to Einstein's version of the Maxwell–Boltzmann doctrine of equipartition.

We perform the same calculation as Rayleigh carried out for the normal modes of oscillation of waves in a box, but now we are dealing with a one-dimensional, rather than

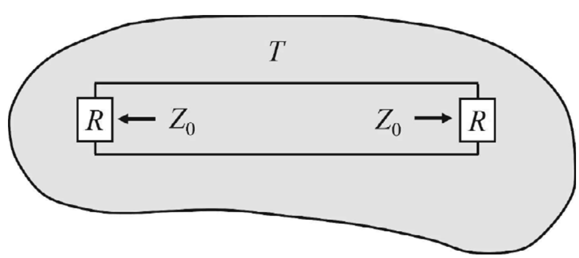


Fig. 17.1 A transmission line of wave impedance Z, terminated at each end with matched resistors R in thermodynamic equilibrium within an enclosure at temperature T.

a three-dimensional, problem. The one-dimensional case of a line of length L is easily derived from the relation (14.45):

$$\frac{\omega}{c} = \frac{n\pi}{L}, \tag{17.33}$$

where n takes only integral values, $n = 1, 2, 3 \ldots$ and c is the speed of light. Standard analysis of the nature of these modes starting from Maxwell's equations shows that there is only one polarisation associated with each value of n. In thermodynamic equilibrium, each mode is awarded energy $\overline{E}$ and hence, to work out the energy per unit frequency range, we need to know the number of modes in the frequency range $d\nu$. The number is found by differentiating (17.33):

$$dn = \frac{L}{\pi c}\, d\omega = \frac{2L}{c}\, d\nu, \tag{17.34}$$

and hence the energy per unit frequency range is $2L\overline{E}/c$.

One of the basic properties of standing waves such as those represented by (17.32) is that they correspond precisely to the superposition of waves of equal amplitude propagating at velocity c in opposite directions along the line. These travelling waves correspond to the only two permissible solutions of Maxwell's equations for propagation along the transmission line. A certain amount of power is therefore delivered to the matched resistors R at either end of the line and, since they are matched, all the energy of the waves is absorbed by them. Equal amounts of energy $L\overline{E}/c$ per unit frequency interval propagate in either direction and travel along the wire into the resistors in a travel time $t = L/c$. Therefore, the power delivered to each resistor is

$$P = \frac{L\overline{E}/c}{L/c} = \overline{E} \tag{17.35}$$

in units of W Hz^{-1}. In thermodynamic equilibrium, however, the same amount of power must be returned to the transmission line, or else the resistor would heat up above temperature T. This proves the fundamental result that the noise power delivered by a resistor at temperature T is

$$P = \frac{h\nu}{e^{h\nu/k_B T} - 1}. \tag{17.36}$$

At low frequencies, $h\nu \ll k_B T$, which is normally the case for radio and microwave receivers, (17.36) reduces to

$$P = k_B T. \tag{17.37}$$

This is often referred to as *Nyquist noise* applied to the thermal noise of a resistor at low frequencies.[4] The power available in the frequency range ν to $\nu + d\nu$ is $P\, d\nu = k_B T\, d\nu$. In the opposite limit $h\nu \gg k_B T$, the noise power decreases exponentially as $P = h\nu \exp(-h\nu/k_B T)$.

The result (17.37) was confirmed experimentally by John Johnson in 1926 and is known as *Johnson noise*.[5] For a variety of resistors, he found exactly the relation predicted by expression (17.37) and from it derived an estimate of Boltzmann's constant k_B within 8% of the standard value.

The expression (17.37) provides a convenient way of describing the performance of a radio receiver which generates a noise power P_n. We define an *equivalent noise temperature* T_n for the performance of the receiver at frequency ν by the relation

$$T_n = P_n/k_B. \tag{17.38}$$

For low frequencies, we note another key feature of the result (17.37). Per unit frequency range per second, the electrical noise power k_BT corresponds exactly to the energy of a single mode in thermodynamic equilibrium. Since this energy is in the form of electrical signals, or waves, the fluctuations in this mode correspond to

$$\Delta E/E = 1, \tag{17.39}$$

as was shown in Section 17.3. Therefore, we expect the amplitude of the noise fluctuations per unit frequency interval to be k_BT.

The Detection of Electromagnetic Waves in the Presence of Noise

In the case $h\nu \ll k_BT$, black-body radiation behaves like classical electromagnetic radiation. The signal induces a current or voltage in the detector which can be modelled as a resistor with noise power $P = k_BT$. Suppose the signal is received within a waveband ν to $\nu + d\nu$. The detected signal is the vector sum of all the electric fields of the waves as displayed on an Argand diagram (Fig. 17.2). This diagram rotates at a rate νHz. If all the waves were of exactly the same frequency, the pattern would continue to rotate in this configuration forever. However, because there is a spread in frequencies, the vectors change their phase relations so that after a time $T \sim 1/\Delta\nu$, the vector sum of the waves is different. The time T is called the *coherence time* of the waves, that is, roughly the time during which the vector sum of the waves provides an independent estimate of the resultant amplitude.

Whereas with photons we obtain an independent piece of information every time a photon arrives, in the case of waves we obtain an independent estimate once per coherence

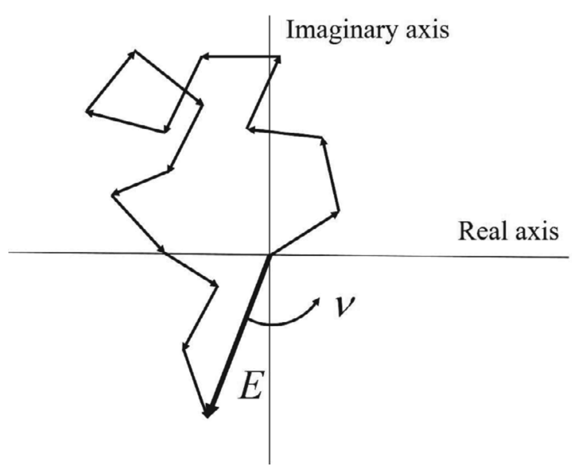


Fig. 17.2 Illustrating the random superposition of electromagnetic waves in the frequency range ν to $\nu + d\nu$ on an Argand diagram.

time $T = 1/\Delta\nu$. Thus, if a source is observed for a time t, we obtain $t/T = t\,\Delta\nu$ independent estimates of the intensity of the source. These results tell us how often we need sample the signal.

Often we are interested in measuring very weak signals in the presence of noise in the receiver. The noise power itself fluctuates with amplitude $\Delta E/E = 1$ per mode per second according to our analysis of Section 17.3. We can therefore reduce the amplitude of the variations by integrating for long periods and by increasing the bandwidth of the receiver. In both cases, the number of independent estimates of the strength of the signal is $\Delta\nu\, t$ and hence the amplitude of the power fluctuations after time t is reduced to

$$\Delta P = \frac{k_{\mathrm{B}}T}{(\Delta\nu\, t)^{1/2}}.$$

Thus, if we are prepared to integrate long enough and use large enough bandwidths, very weak signals can be detected.

17.5 The First Solvay Conference

Einstein published his remarkable analysis of fluctuations in black-body radiation in 1909, but there was still only modest support for the concept of light quanta. Among those who were persuaded of the importance of quanta was Walther Nernst, who at that time was measuring the low temperature heat capacities of various materials. As recounted in Section 16.3, Nernst visited Einstein in Zurich in March 1910 and they compared Einstein's theory with his recent experiments. These experiments showed that Einstein's predictions of the low temperature variation of the specific heat capacity with temperature (16.25) gave a good description of the experimental results. As Einstein wrote to his friend Jakob Laub after the visit,

> I consider the quantum theory certain. My predictions with respect to specific heats seem to be strikingly confirmed. Nernst, who has just been here, and Rubens are eagerly occupied with experimental tests, so that people will soon be informed about this matter.[6]

By 1911, Nernst was convinced, not only of the importance of Einstein's results, but also of the theory underlying them. The outcome of the meeting with Einstein was dramatic. The number of papers on quanta began to increase rapidly, as Nernst popularised his results on the quantum theory of solids.

Nernst was a friend of the wealthy Belgian industrialist Ernest Solvay and he persuaded him to sponsor a meeting of a select group of prominent physicists to discuss the issues of quanta and radiation. The idea was first mooted in 1910, but Planck urged that the meeting should be postponed for a year. As he wrote,

> My experience leads me to the opinion that scarcely half of those you envisage as participants have a sufficiently lively conviction of the pressing need for reform to be motivated to attend the conference ... Of the entire list you name, I believe that besides ourselves [only] Einstein, Lorentz, W. Wein and Larmor are deeply interested in the topic.[7]

Fig. 17.3 The participants in the first Solvay conference on physics, Brussels 1911. (From *La Théorie du rayonnement et les quanta*, 1912, eds. P. Langevin and M. De Broglie, Paris: Gautier-Villars.)

By 1911, the balance of opinion was shifting. This was the first, and perhaps the most significant, of the famous series of Solvay conferences. The eighteen official participants met on 29 October 1911 in the Hotel Metropole in Brussels and the meeting took place between the 30th of that month and the 3rd of November (Fig. 17.3). By this time, the majority of the participants were supporters of the quantum hypothesis. Two were definitely against quanta – Jeans and Poincaré. Five were initially neutral – Rutherford, Brillouin, Curie, Perrin and Knudsen. The eleven others were basically pro-quanta – Lorentz (chairman), Nernst, Planck, Rubens, Sommerfeld, Wien, Warburg, Langevin, Einstein, Hasenöhrl and Onnes. The secretaries were Goldschmidt, Maurice de Broglie and Lindemann; Solvay, who hosted the conference, was there, as well as his collaborators Herzen and Hostelet. Rayleigh and van der Waals were unable to attend.

As a result of the meeting, Poincaré was converted to the notion of quanta and the physicists who took a neutral position had done so because they were unfamiliar with the arguments. The conference had a profound effect in that it provided a forum at which all the arguments could be presented. In addition, all the participants wrote up their lectures for publication beforehand and these were then discussed in detail. These discussions were recorded and the full proceedings published within a year as the important volume *La Théorie du rayonnement et les quanta: Rapports et discussions de la réunion tenue à Bruxelles, du 30 octobre au 3 novembre 1911*.[8] Thus, all the important issues were made available to the scientific community in one volume. As a result, the next generation of students became fully familiar with the arguments and many of them set to work

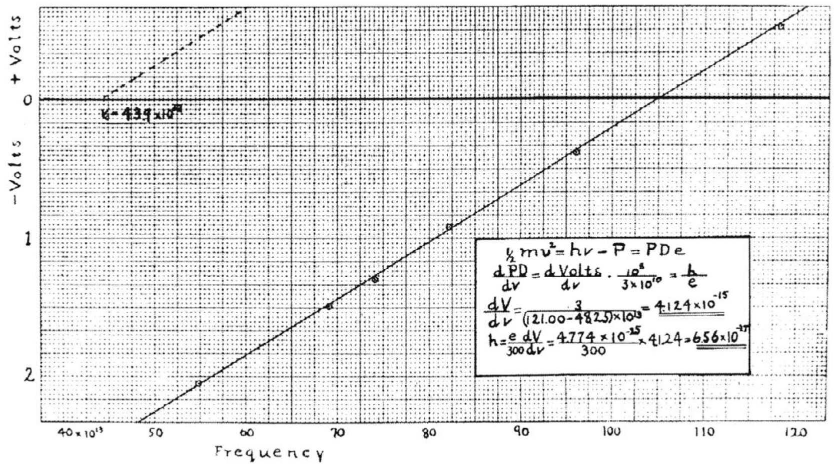


Fig. 17.4 Millikan's results on the photoelectric effect compared with the predictions of Einstein's quantum theory. (From R.A. Millikan, 1916, *Physical Review*, **7**, 355.)

immediately to tackle the problems of quanta. Furthermore, these problems began to be appreciated beyond the central European German-speaking scientific community.

It would be wrong, however, to believe that everyone was suddenly convinced of the existence of quanta. In his paper of 1916 on his great series of experiments in which he verified Einstein's prediction of the dependence of the photoelectric effect upon frequency (Fig. 17.4), Millikan stated:

> We are confronted however by the astonishing situation that these facts were correctly and exactly predicted nine years ago by a form of quantum theory which has now been generally abandoned.[9]

Millikan refers to Einstein's

> bold, not to say reckless, hypothesis of an electromagnetic light corpuscle of energy $h\nu$ which flies in the face of the thoroughly established facts of interference.[9]

17.6 Experimental and Theoretical Advances 1911 to 1925

After 1911, the light quantum hypothesis had to be taken seriously and the application of quantum concepts to atoms began a new and crucial development which would eventually lead to the definitive theory of quantum mechanics in the period 1925 to 1930, but it would prove to be a rocky road, with many blind alleys. Between 1911 and 1925, the unsuccessful efforts to create a consistent theory of quanta, what is known as the *Old Quantum Theory*, were the precursors of the truly revolutionary concepts propounded by Heisenberg, Born, Schrödinger, Pauli, Dirac and many others from 1925 onwards.

This remarkable and complex story is told in some detail in my companion volume *Quantum Concepts in Physics*.[10] For our present purposes, this convoluted story is only outlined here, except for those aspects bearing on the phenomena of light quanta which were to prove crucial in the discovery of quantum mechanics.

17.6.1 Thomson and Rutherford

Great advances were made in understanding the nature of atoms, largely as a result of the brilliant experiments associated with the names of J.J. Thomson and Ernest Rutherford.

The discovery of the electron in 1897 is attributed to J.J. Thomson on the basis of his famous series of experiments, in which he established that the charge-to-mass ratio, e/m_e, of cathode rays is about two-thousand times that of hydrogen ions. Thomson was the first to interpret the cathode ray e/m experiments in terms of a sub-atomic particle. In his words, cathode rays constituted

> . . . a new state, in which the subdivision of matter is carried very much further than in the ordinary gaseous state.[11]

In 1899, Thomson used one of C.T.R. Wilson's early condensation chambers to measure the charge of the electron. The sensitivity of the apparatus was such that a single electronic charge could act as a condensation centre for a water droplet and so they counted the total number of droplets formed and their total charge. From these, Thomson estimated the charge of the electron to be $e = 2.2 \times 10^{-19}$ C, compared with the present standard value of 1.602×10^{-19} C. This experiment was the precursor of the famous Millikan oil drop experiment, in which the water droplets were replaced by fine drops of a heavy oil, which did not evaporate during the course of the experiment.

Thomson pursued a much more sustained and detailed campaign than the other physicists in establishing the universality of what became known as *electrons*, the name coined for cathode rays by Johnstone Stoney in 1891. Thomson demonstrated that the β-particles emitted in radioactive decays and those ejected in the photoelectric effect had the same charge-to-mass ratio as the cathode rays.

It was inferred that there must be electrons within the atom, but it was not clear how many there were. There might have been as many as 2000 electrons, if they were to make up the mass of the hydrogen atom. The answer was provided by a brilliant series of experiments carried out by Thomson and his colleagues, who studied the scattering of X-rays by thin films. X-rays are scattered by the electrons in atoms by *Thomson scattering*, the theory of which was worked out by Thomson[12] using the classical expression for the radiation of an accelerated electron (see Section 14.2.1). The cross-section for the scattering of a beam of incident radiation by an electron is

$$\sigma_T = \frac{e^4}{6\pi\epsilon_0^2 m_e^2 c^4} = \frac{8\pi r_e^2}{3} = 6.653 \times 10^{-29}\ \text{m}^2, \qquad (17.40)$$

the *Thomson cross-section*, where $r_e = e^2/4\pi\epsilon_0 m_e c^2$ is the classical electron radius.[13]

Thomson found that the intensity of scattered X-rays from thin films indicated that there were not thousands of electrons in each atom, but a much smaller number. In collaboration with Charles Barkla, he showed that, except for hydrogen, the number of electrons was roughly half the atomic weight. The picture which emerged was one in which the number of electrons and, consequently, the amount of positive charge in the atom increased by units of the electronic charge. Furthermore, most of the mass of the atom had to be associated with the positive mass. A key question was, 'How are the electrons and the positive charge distributed inside atoms?' In the picture favoured by Thomson, the positive charge was distributed throughout the atom and, within this sphere, the negatively charged electrons were placed on carefully chosen orbits, such that their dipole, quadrupole and higher multipole moments cancelled out – this rather clever model became known by the somewhat pejorative name, the 'plum-pudding' model of the atom, the plums representing the electrons and the pudding the positive charge.

The discovery of the nuclear structure of atoms resulted from a brilliant series of experiments carried out by Ernest Rutherford and his colleagues Hans Geiger and Ernest Marsden in the period 1909–12 at Manchester University. In 1907 Rutherford and Royds demonstrated convincingly that α-particles are helium nuclei.[14] Rutherford had also been impressed by the fact that α-particles could pass through thin films rather easily, suggesting that much of the volume of atoms is empty space, although there was clear evidence of small-angle scattering. Rutherford persuaded Marsden, who was still an undergraduate, to investigate whether or not α-particles were deflected through large angles on being fired at a thin gold foil. To Rutherford's astonishment, a few particles were deflected by more than 90°, and a very small number almost returned along the direction of incidence. In Rutherford's words:

> It was quite the most incredible event that has ever happened to me in my life. It was almost as incredible as if you fired a 15-inch shell at a piece of tissue paper and it came back and hit you.[15]

Rutherford realised that it required a very considerable force to send the α-particle back along its track. It was only in 1911 that he hit upon the idea that, if all the positive charge were concentrated in a compact nucleus, the scattering could be attributed to the repulsive electrostatic force between the incoming α-particle and the positive nucleus. Rutherford was no theorist, but he used his knowledge of central orbits in inverse-square law fields of force to work out the properties of what became known as *Rutherford scattering*.[16] As was shown in the appendix to Chapter 4, the orbit of the α-particle is a *hyperbola*, the angle of deflection ϕ being given by

$$\cot\frac{\phi}{2} = \left[\frac{4\pi\epsilon_0 m_\alpha}{2Ze^2}\right] p_0 v_0^2, \tag{17.41}$$

where p_0 is the collision parameter, v_0 is the initial velocity of the α-particle and Z the nuclear charge. It is straightforward to work out the probability that the α-particle is scattered through an angle ϕ. The result is

$$p(\phi) \propto \frac{1}{v_0^4}\,\mathrm{cosec}^4\,\frac{\phi}{2}, \tag{17.42}$$

the famous $\mathrm{cosec}^4(\phi/2)$ law derived by Rutherford, which was found to explain precisely the observed distribution of scattering angles of the α-particles.[17]

Rutherford had, however, achieved much more. The fact that the scattering law was obeyed so precisely, even for large angles of scattering, meant that the inverse-square law of electrostatic repulsion held good to very small distances indeed. They found that the nucleus had to have size less than about 10^{-14} m, very much less than the sizes of atoms, which are typically about 10^{-10} m. It was not until 1914 that Rutherford was thoroughly convinced of the necessity of adopting the nuclear model of the atom. Before that date, however, someone else did – Niels Bohr, the first theorist to apply successfully quantum concepts to the structure of atoms.

17.6.2 The Structure of Atoms and the Bohr Model of the Hydrogen Atom

The construction of atomic models was a major industry in the early years of the twentieth century, particularly in Great Britain, and has been splendidly surveyed by Heilbron.[18] The model builders, of whom Thomson was a leading practitioner, faced two major problems. The first was the dynamical stability of the distribution of electrons within atoms and the second, and potentially an even greater problem, was the fact that an electron accelerated in the field of the positive charge radiates electromagnetic radiation, and so would lose energy and spiral very rapidly to the centre of the atom.

However the electrons are distributed in the atom, they cannot be stationary because of *Earnshaw's theorem*, which states that any static distribution of electric charges is mechanically unstable, in that they either collapse or disperse to infinity under the action of electrostatic forces. The alternative is to place the electrons in orbits, what is often called the 'Saturnian' model of the atom. The most famous of these early models was that due to the Japanese physicist Hantoro Nagaoka, who attempted to associate the spectral lines of atoms with small vibrational perturbations of the electrons about their equilibrium orbits. The problem with this model was that perturbations in the plane of the electron's orbit are unstable, leading to instability of the atom as a whole.

The radiative instability of the electron can be understood as follows. Suppose the electron is in a circular orbit of radius a. Then, equating the centripetal force to the electrostatic force of attraction between the electron and the nucleus of charge Ze,

$$\frac{Ze^2}{4\pi\epsilon_0 a^2} = \frac{m_e v^2}{a} = m_e|\ddot{\boldsymbol{r}}|, \tag{17.43}$$

where $|\ddot{\boldsymbol{r}}|$ is the centripetal acceleration. The rate at which the electron loses energy by radiation is given by (14.5). The kinetic energy of the electron is $E = \frac{1}{2}m_e v^2 = \frac{1}{2}m_e a|\ddot{\boldsymbol{r}}|$ and so the time it takes the electron to lose all its kinetic energy by radiation is

$$T = \frac{E}{|\mathrm{d}E/\mathrm{d}t|} = \frac{2\pi a^3}{\sigma_T c}. \tag{17.44}$$

Taking the radius of the atom to be $a = 10^{-10}$ m, the electron loses all its energy in about 3×10^{-10} s. Something is profoundly wrong. As the electron loses energy, it moves into an orbit of smaller radius, loses energy more rapidly and spirals into the centre.

The pioneer atom model builders were well aware of this problem. The solution was to place the electrons in orbits such that there is no net acceleration when the acceleration vectors of all the electrons in the atom are added together. This requires, however, the electrons to be very well ordered in their orbits about the nucleus. This was the basis of Thomson's 'plum-pudding' model. The 'plums', or electrons, were not randomly located within the sphere of positive charge, but were precisely located to ensure that the radiative instability did not occur. The radiative problem remained, however, particularly for hydrogen, which possesses only one electron.

Niels Bohr completed his doctorate on the electron theory of metals in 1911 and spent the following year in England, working for seven months with J.J. Thomson at the Cavendish Laboratory in Cambridge, and then four months with Ernest Rutherford in Manchester. Bohr was immediately struck by the significance of Rutherford's model of the nuclear structure of the atom and began to devote all his energies to understanding atomic structure on that basis. He quickly appreciated the distinction between the chemical properties of atoms, which are associated with the orbiting electrons, and radioactive processes, which are associated with activity in the nucleus. On this basis, he could understand the nature of the isotopes of a particular chemical species. Bohr also realised from the outset that the structure of atoms could not be understood on the basis of classical physics. The obvious way forward was to incorporate the quantum concepts of Planck and Einstein into the models of atoms. Einstein's statement, quoted in Section 16.3,

> ... for ions which can vibrate with a definite frequency, ... the manifold of possible states must be narrower than it is for bodies in our direct experience.[19]

was precisely the type of constraint which Bohr was seeking. Such a mechanical constraint was essential to understand how atoms could survive the inevitable instabilities according to classical physics. How could these ideas be incorporated into models for atoms?

Bohr was not the first physicist to introduce quantum concepts into the construction of atomic models. In 1910, a Viennese doctoral student, Arthur Erich Haas, realised that, if Thomson's sphere of positive charge were uniform, an electron would perform simple harmonic motion through the centre of the sphere, since the restoring force at radius r from the centre would be, according to Gauss's theorem in electrostatics,

$$f = m_e \ddot{r} = -\frac{eQ(\leq r)}{4\pi\epsilon_0 r^2} = -\left(\frac{eQ}{4\pi\epsilon_0 a^3}\right) r, \tag{17.45}$$

where a is the radius of the atom and Q the total positive charge. For a hydrogen atom, for which $Q = e$, the frequency of oscillation of the electron is

$$\nu = \frac{1}{2\pi}\left(\frac{e^2}{4\pi\epsilon_0 m_e a^3}\right)^{1/2}. \tag{17.46}$$

Haas argued that the energy of oscillation of the electron, $E = e^2/4\pi\epsilon_0 a$, should be quantised and set equal to $h\nu$. Therefore,

$$h^2 = \frac{\pi m_e e^2 a}{\epsilon_0}. \tag{17.47}$$

Haas used (17.47) to show how Planck's constant could be related to the properties of atoms, taking for ν the short wavelength limit of the Balmer series, that is, setting $n \to \infty$ in the Balmer formula (17.52). Haas's efforts were discussed by Lorentz at the 1911 Solvay conference, but they did not attract much attention.[20] According to Haas's approach, Planck's constant was simply a property of atoms according to (17.47), whereas those already converted to quanta preferred to believe that h had much deeper significance.

In the summer of 1912, Bohr wrote an unpublished Memorandum for Rutherford, in which he made his first attempt at quantising the energy levels of the electrons in atoms.[21] He proposed relating the kinetic energy T of the electron to the frequency $\nu' = v/2\pi a$ of its orbit about the nucleus through the relation

$$T = \tfrac{1}{2} m_e v^2 = K\nu', \tag{17.48}$$

where K is a constant, which he expected would be of the same order of magnitude as Planck's constant h. Bohr believed there must be some such non-classical constraint in order to guarantee the stability of atoms. Indeed, his criterion (17.48) fixed absolutely the kinetic energy of the electron about the nucleus. For a bound circular orbit,

$$\frac{mv^2}{a} = \frac{Ze^2}{4\pi\epsilon_0 a^2}, \tag{17.49}$$

where Z is the positive charge of the nucleus in units of the charge of the electron e. As is well known, the binding energy of the electron is

$$E = T + U = \tfrac{1}{2} m_e v^2 - \frac{Ze^2}{4\pi\epsilon_0 a} = -T = \frac{U}{2}, \tag{17.50}$$

where U is the electrostatic potential energy. The quantisation condition (17.48) enables both v and a to be eliminated from the expression for the kinetic energy of the electron. A straightforward calculation shows that

$$T = \frac{mZ^2 e^2}{32\epsilon_0^2 K^2}, \tag{17.51}$$

which was to prove to be of great significance for Bohr. His Memorandum containing these ideas was principally about issues such as the number of electrons in atoms, atomic volumes, radioactivity, the structure and binding of diatomic molecules and so on. There is no mention of spectroscopy, which he and Thomson considered too complex to provide useful information.

The next clue was provided by the work of the Cambridge physicist John William Nicholson,[22] who arrived at the concept of the quantisation of angular momentum by a quite different route. Nicholson had shown that, although the Saturnian model of the atom is unstable for perturbations in the plane of the orbit, perturbations perpendicular to the plane are stable for orbits containing up to five electrons – he assumed that the unstable modes in the plane of the orbit were suppressed by some unspecified mechanism. The frequencies of the stable oscillations were multiples of the orbital frequency, and he compared these with the frequencies of the lines observed in the spectra of bright nebulae, particularly with the unidentified 'nebulium' and 'coronium' lines. Performing the same exercise for ionised atoms with one less orbiting electron, further matches to the astronomical spectra were obtained. The frequency of the orbiting electrons remained a free parameter, but when he worked out the angular momentum associated with them, Nicholson found that they were multiples of $h/2\pi$. When Bohr returned to Copenhagen later in 1912, he was perplexed by the success of Nicholson's model, which seemed to provide a successful, quantitative model for the structure of atoms and which could account for the spectral lines observed in astronomical spectra.

The breakthrough came in early 1913, when H.M. Hansen told Bohr about the Balmer formula for the wavelengths, or frequencies, of the spectral lines in the spectrum of hydrogen,

$$\frac{1}{\lambda} = \frac{\nu}{c} = R\left(\frac{1}{2^2} - \frac{1}{n^2}\right), \tag{17.52}$$

where $R = 1.097 \times 10^7$ m^{-1} is known as the *Rydberg constant* and $n = 3, 4, 5, \ldots$. As Bohr recalled much later,

> As soon as I saw Balmer's formula, the whole thing was clear to me.[23]

He realised immediately that this formula contained within it the crucial clue for the construction of a model of the hydrogen atom, which he took to consist of a single negatively charged electron in a circular orbit about a positively charged nucleus. He went back to his Memorandum on the quantum theory of the atom, in particular, to his expression for the binding energy, or kinetic energy, of the electron (17.51). He realised that he could determine the value of the constant K from the expression for the Balmer series. The running term in $1/n^2$ can be associated with (17.51), if we write for hydrogen with $Z = 1$,

$$T = \frac{m_e e^2}{32\epsilon_0 n^2 K^2}. \tag{17.53}$$

Then, when the electron changes from an orbit with quantum number n to that with $n = 2$, the energy of the emitted radiation would be the difference in the total energies of the two states. Applying Einstein's quantum hypothesis, this energy should be equal to $h\nu$. Inserting the numerical values of the constants into (17.53), Bohr found that the constant K was exactly $h/2$. Therefore, the energy of the state with quantum number n is

$$T = \frac{m_e e^2}{8\epsilon_0 n^2 h^2}. \tag{17.54}$$

The angular momentum of the state could be found directly by writing $T = \frac{1}{2} I \omega'^2 = 8\pi^4 m_e a^2 \nu'^2$, from which it follows that

$$J = I\omega' = \frac{nh}{2\pi}. \tag{17.55}$$

This is how Bohr arrived at the quantisation of angular momentum according to the 'old' quantum theory. In the first paper of his famous trilogy of 1913,[24] Bohr acknowledged that Nicholson had discovered the quantisation of angular momentum in his papers of 1912. These results were the inspiration for what became known as the *Bohr model of the atom*.

In the same paper of 1913, Bohr noted that a similar formula to (17.52) could account for the *Pickering series*, which had been discovered in 1896 by Edward Pickering in the spectra of hot Wolf–Rayet stars, usually in absorption. In 1912, Alfred Fowler discovered the same series in laboratory experiments. Bohr argued that singly ionised helium atoms would have exactly the same spectrum as hydrogen, but the wavelengths of the corresponding lines would be four times shorter, as observed in the Pickering series. Fowler objected, however, that the ratio of the Rydberg constants for singly ionised helium and hydrogen was not 4, but 4.00163. Bohr realised that the problem arose from neglecting the contribution of the mass of the nucleus to the computation of the moments of inertia of the hydrogen atom and the helium ion. If the angular velocity of the electron and the nucleus about their centre of mass is ω, the condition for the quantisation of angular momentum becomes

$$\frac{nh}{2\pi} = \mu\omega R^2, \tag{17.56}$$

where $\mu = m_e m_N/(m_e + m_N)$ is the *reduced mass* of the atom, or ion, which takes account of the contributions of both the electron and the nucleus to the angular momentum; R is their separation. Therefore, the ratio of Rydberg constants for ionised helium and hydrogen should be

$$\frac{R_{He^+}}{R_H} = 4\left(\frac{1 + \dfrac{m_e}{M}}{1 + \dfrac{m_e}{4M}}\right) = 4.00160, \tag{17.57}$$

where M is the mass of the hydrogen atom. Thus, precise agreement was found between the theoretical and laboratory estimates of the ratio of Rydberg constants for hydrogen and ionised helium.

The Bohr theory of the hydrogen atom was a quite remarkable achievement and the first convincing application of quantum concepts to atoms. Bohr's dramatic results were persuasive evidence for many scientists that Einstein's quantum theory had to be taken really seriously for processes occurring on the atomic scale. The 'old' quantum theory was, however, seriously incomplete and constituted an uneasy mixture of classical and quantum ideas without any self-consistent theoretical underpinning.

From Einstein's point of view, the results provided further strong support for his quantum picture of elementary processes. In his biography of Bohr, Pais tells the story of Hevesy's encounter with Einstein in September 1913.[25] When Einstein heard of Bohr's analysis of the Balmer series of hydrogen, Einstein remarked cautiously that Bohr's work was very interesting, and important if right. When Hevesy told him about the helium results, Einstein responded,

> This is an enormous achievement. The theory of Bohr must then be right.

17.7 Einstein (1916) 'On the Quantum Theory of Radiation'

During the years 1911 to 1916, Einstein was preoccupied with the formulation of general relativity, which will be discussed in Chapter 19. By 1916, the pendulum of scientific opinion was beginning to swing in favour of the quantum theory, particularly following the success of the Bohr's theory of the hydrogen atom. These ideas fed back into Einstein's thinking about the problems of the emission and absorption of radiation and resulted in his famous derivation of the Planck spectrum through the introduction of what are now called *Einstein’s A and B coefficients*.

Einstein’s great paper of 1916 begins by noting the formal similarity between the Maxwell–Boltzmann distribution for the velocity distribution of the molecules in a gas and Planck’s formula for the black-body spectrum.[26] Einstein shows how these distributions can be reconciled through a derivation of the Planck spectrum, which gives insight into what he refers to as the ‘still unclear processes of emission and absorption of radiation by matter’. The paper begins with a description of a quantum system consisting of a large number of molecules which can occupy a discrete set of states $Z_1, Z_2, Z_3, \dots$ with corresponding energies $\epsilon_1, \epsilon_2, \epsilon_3, \dots$ According to classical statistical mechanics, the relative probabilities W_n of these states being occupied in thermodynamic equilibrium at temperature T are given by Boltzmann’s relation:

$$W_n = g_n \exp\left(-\frac{\epsilon_n}{k_{\mathrm{B}}T}\right), \qquad (17.58)$$

where the g_n are the *statistical weights*, or *degeneracies*, of the states Z_n. As Einstein remarks forcefully in his paper,

> [17.58] expresses the farthest-reaching generalisation of Maxwell’s velocity distribution law.

Consider two quantum states of the gas molecules, Z_m and Z_n, with energies ϵ_m and ϵ_n respectively, such that $\epsilon_m > \epsilon_n$. Following the precepts of the Bohr model, it is assumed that a quantum of radiation is emitted if the molecule changes from the state Z_m to Z_n, the energy of the quantum being $h\nu = \epsilon_m - \epsilon_n$. Similarly, when a photon of energy $h\nu$ is absorbed, the molecule changes from the state Z_n to Z_m.

The quantum description of these processes follows by analogy with the classical processes of the emission and absorption of radiation.

Spontaneous emission Einstein notes that a classical oscillator emits radiation in the absence of excitation by an external field. The corresponding process at the quantum level is called *spontaneous emission*, and the probability of it taking place in the time interval dt is

$$\mathrm{d}W = A_m^n\,\mathrm{d}t, \qquad (17.59)$$

similar to the law of radioactive decay.

Induced emission and absorption By analogy with the classical case, if the oscillator is excited by waves of the same frequency as the oscillator, it can either gain or lose energy, depending upon the phase of the wave relative to that of the oscillator, that is, the work done on the oscillator can be either positive or negative. The magnitude of the positive or negative work done is proportional to the energy density of the incident waves. The quantum mechanical equivalents of these processes are those of *induced absorption*, in which the molecule is excited from the state Z_n to Z_m, and *induced emission*, in which the molecule emits a photon under the influence of the incident radiation field. The probabilities of these processes are written

$$\textit{Induced absorption} \qquad \mathrm{d}W = B_n^m \varrho \, \mathrm{d}t,$$
$$\textit{Induced emission} \qquad \mathrm{d}W = B_m^n \varrho \, \mathrm{d}t.$$

The lower indices refer to the initial state and the upper indices the final state; ϱ is the energy density of radiation with frequency ν; B_n^m and B_m^n are constants for a particular physical processes, and are referred to as 'changes of state by induced emission and absorption'.

We now seek the spectrum of the radiation $\varrho(\nu)$ in thermal equilibrium. The relative numbers of molecules with energies ϵ_m and ϵ_n in thermal equilibrium are given by the Boltzmann relation (17.58) and so, in order to leave the equilibrium distribution unchanged under the processes of spontaneous and induced emission and induced absorption of radiation, the probabilities must balance, that is,

$$\underbrace{g_n \mathrm{e}^{-\epsilon_n/k_\mathrm{B}T} B_n^m \varrho}_{\text{absorption}} = \underbrace{g_m \mathrm{e}^{-\epsilon_m/k_\mathrm{B}T} \left(B_m^n \varrho + A_m^n\right)}_{\text{emission}}. \qquad (17.60)$$

Now, in the limit $T \to \infty$, the radiation energy density $\varrho \to \infty$, and the induced processes dominate the equilibrium. Therefore, allowing $T \to \infty$ and setting $A_m^n = 0$, (17.60) becomes

$$g_n B_n^m = g_m B_m^n. \qquad (17.61)$$

Reorganising (17.60), the equilibrium radiation spectrum ϱ can be written

$$\varrho = \frac{A_m^n / B_m^n}{\exp\left(\dfrac{\epsilon_m - \epsilon_n}{k_\mathrm{B}T}\right) - 1}. \qquad (17.62)$$

But, this is Planck's radiation law. Suppose we consider only Wien's law, which is known to be the correct expression in the frequency range in which light can be considered to consist of quanta. Then, in the limit $\epsilon_m - \epsilon_n / k_\mathrm{B}T \gg 1$,

$$\varrho = \frac{A_m^n}{B_m^n} \exp\left(-\frac{\epsilon_m - \epsilon_n}{k_\mathrm{B}T}\right) \propto \nu^3 \exp\left(-\frac{h\nu}{k_\mathrm{B}T}\right). \qquad (17.63)$$

Therefore, we find the following 'thermodynamic' relations:

$$\frac{A_m^n}{B_m^n} \propto \nu^3, \qquad (17.64)$$

$$\epsilon_m - \epsilon_n = h\nu. \qquad (17.65)$$

The value of the constant in (17.64) can be found from the Rayleigh–Jeans limit of the black-body spectrum, $\epsilon_m - \epsilon_n / k_\mathrm{B}T \ll 1$. From (14.51),

$$\varrho(\nu) = \frac{8\pi\nu^2}{c^3} k_\mathrm{B}T = \frac{A_m^n}{B_m^n} \frac{k_\mathrm{B}T}{h\nu},$$

$$\frac{A_m^n}{B_m^n} = \frac{8\pi h\nu^3}{c^3}. \qquad (17.66)$$

The importance of these relations between the A and B coefficients is that they are associated with atomic processes at the microscopic level. Once A_m^n or B_m^n or B_n^m is known, the other coefficients can be found from (17.61) and (17.66).

It is intriguing that this important analysis occupies only the first three sections of Einstein's paper. The important thrust of the paper for Einstein was that he could now use these results to determine how the motions of molecules would be affected by the emission and absorption of quanta. The analysis was similar to his earlier studies of Brownian motion, but now applied to the case of quanta interacting with molecules. The quantum nature of the processes of emission and absorption were essential features of his argument. We will not go into the details of this remarkable calculation, except to quote the key result that, when a molecule emits or absorbs a quantum of energy $h\nu$, there must be a positive or negative change in the momentum of the molecule of magnitude $|h\nu/c|$, even in the case of spontaneous emission. In Einstein's words,

> There is no radiation of spherical waves. In the spontaneous emission process, the molecule suffers a recoil of magnitude $h\nu/c$ in a direction that, in the present state of the theory, is determined only by 'chance'.

The Einstein A and B coefficients can be related to the emission and absorption coefficients appearing in Kirchhoff's law (see Section 13.2.2). Einstein's profound insight opened up new ways of understanding the interaction between matter and radiation. The transfer equation for radiation (13.16) is

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}l} = -\alpha_\nu I_\nu + j_\nu .$$

Suppose the upper and lower states are labelled 2 and 1 respectively. It is natural to associate the emission coefficient j_ν with the spontaneous rate of emission of quanta per unit solid angle, that is,

$$j_\nu = \frac{h\nu}{4\pi} n_2 A_{21}, \tag{17.67}$$

where n_2 is the number density of molecules in the upper state. In the same way, the energy absorbed by the molecules by induced absorption is

$$\frac{h\nu}{4\pi} n_1 B_{12} I_\nu, \tag{17.68}$$

so that the absorption coefficient is

$$\alpha_\nu = \frac{h\nu}{4\pi} n_1 B_{12}. \tag{17.69}$$

We still have to take account of induced emission. Although it might seem that this term should be included in the emission coefficient, it is preferable to consider the term to represent *negative absorption* since it depends upon the intensity of radiation along the direction of the ray, as does the induced absorption term. Therefore, we can write the absorption coefficient *corrected for stimulated emission* as

$$\alpha_\nu = \frac{h\nu}{4\pi} (n_1 B_{12} - n_2 B_{21}) = \frac{h\nu}{4\pi} n_1 B_{12} \left(1 - \frac{g_1 n_2}{g_2 n_1}\right). \tag{17.70}$$

The transfer equation for radiation then becomes

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}l} = -\frac{h\nu}{4\pi} n_1 B_{12} \left(1 - \frac{g_1 n_2}{g_2 n_1}\right) I_\nu + \frac{h\nu}{4\pi} n_2 A_{21}. \qquad (17.71)$$

In the case in which the populations of the states are given by the Boltzmann distribution and the system has reached thermodynamic equilibrium, $\mathrm{d}I_\nu/\mathrm{d}l = 0$, the Planck distribution is recovered. Equation (17.71) is a remarkably powerful expression. In many physical situations, the distribution of energies among the molecules is maintained in thermal equilibrium by particle collisions, but the radiation is not in thermodynamic equilibrium with the matter – the matter is said to be in a state of *local thermodynamic equilibrium*. In such a circumstance,

$$\frac{n_2}{n_1} = \frac{g_2}{g_1} \exp\left(-\frac{h\nu}{k_\mathrm{B}T}\right); \quad \frac{n_2 g_1}{n_1 g_2} = \exp\left(-\frac{h\nu}{k_\mathrm{B}T}\right) < 1. \qquad (17.72)$$

When the inequality in (17.72) is satisfied, even if the distribution of particle energies is not Maxwellian, the populations of the states are said to be *normal*. It is, however, possible to violate this condition by overpopulating the upper state, in which case the absorption coefficient can become *negative* and the intensity of radiation is amplified along the path of the rays. In this case, the populations of the states of the molecules are said to be *inverted*. This is the principle behind the operation of *masers* and *lasers*.

One aspect of the introduction of Einstein's A and B coefficients was crucial in understanding the equilibrium between matter and radiation. In his analyses of 1909, Einstein had considered the thermodynamic equilibrium between oscillators and the radiation field, but Lorentz argued that the same equilibrium spectrum should be obtained if, rather than oscillators, free electrons and radiation were maintained in thermal contact within an enclosure at temperature T.[27] The process by which energy is exchanged between photons and the electrons is the *Compton scattering* mechanism discussed below. Photons with energy $h\nu$ collide with electrons with momentum $\boldsymbol{p}$ resulting in an increase (or decrease) in the momentum of the electron to $\boldsymbol{p}'$, while the photon leaves with decreased (or increased) energy $h\nu'$.

In 1923, Wolfgang Pauli showed that the mean energy of the electrons was $\frac{1}{2}m\overline{v^2} = \frac{3}{2}k_\mathrm{B}T$ and the Planck spectrum obtained, *provided* induced processes are taken into account in determining the probabilities of scattering into the direction $h\nu'$.[28] The full analysis of the problem is non-trivial, but the essence of it is that, if the energy density of radiation at frequency ν is ϱ and the energy density at frequency ν' is ϱ', the probability of scattering from ν to ν' is given, in the notation used by Pauli, by an expression of the form $(A\varrho + B\varrho\varrho')\,\mathrm{d}t$, where the term in A is the spontaneous scattering and B the induced scattering. The remarkable part of this expression is the term in B which includes the energy density of radiation at the scattered frequency ν' and which corresponds to *induced Compton scattering*. In fact, analysing the origin of these two terms, the A term corresponds to the normal Compton scattering process, in which the radiation can be considered to consist of photons, while the B term corresponds the wave properties of the radiation. Both terms need to be included if the mean energy of the electrons is to be $\frac{3}{2}k_\mathrm{B}T$.

The relation of Pauli's calculation to Einstein's paper of 1916 was elucidated a few months later by Einstein and Paul Ehrenfest,[29] who argued that the process of Compton scattering could be considered to consist of the absorption of a photon of frequency ν and the emission of a photon of frequency ν'. Then, the concepts of induced absorption, spontaneous and induced emission can be applied to the process. The probability of scattering into a particular state with frequency ν' is proportional to the product of the probabilities of absorption and emission by an expression of the form $(B\varrho)(A' + B'\varrho')$ where the A' and B' represent the spontaneous and induced emission terms respectively.[30] It turns out that the A' term includes Planck's constant while B' does not, again betraying the quantum origin of the Wien cut-off and the classical origin of the Rayleigh–Jeans region of the spectrum.

17.8 Compton Scattering

Direct experimental evidence for the correctness of Einstein's conclusion was provided by Arthur Holly Compton's beautiful X-ray scattering experiments of 1923.[31] These showed that X-ray quanta could undergo collisions in which they behave like particles, the *Compton effect* or *Compton scattering*. It is a standard result of elementary special relativity that the increase in wavelength of an electromagnetic quantum in collision with a stationary electron is

$$\lambda' - \lambda = \frac{hc}{m_e c^2}(1 - \cos\theta), \tag{17.73}$$

where θ is the angle through which it is scattered (Fig. 17.5). [32] Implicit in this calculation is the conservation of relativistic three-momentum in which the momentum of the quantum is $h\nu/c$. In June 1929, Werner Heisenberg wrote in a review article, entitled *The Development of the Quantum Theory 1918–1928*,

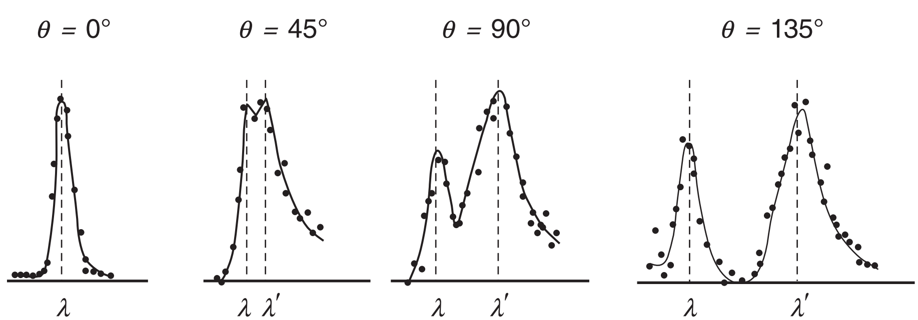


Fig. 17.5 The results of Compton's X-ray scattering experiments showing the increase in wavelength of the K_α line of molybdenum as the deflection angle θ increases. (From A.H. Compton, 1923, *Physical Review*, **21**, 483.) The unscattered wavelength λ of the K_α line is 0.07107 nm.

> At this time [1923] experiment came to the aid of theory with a discovery which would later become of great significance for the development of quantum theory. Compton found that with the scattering of X-rays from free electrons, the wavelength of the scattered rays was measurably longer than that of the incident light. This effect, according to Compton and Debye, could easily be explained by Einstein's light quantum hypothesis; the wave theory of light, on the contrary, failed to explain this experiment. With that result, the problems of radiation theory which had hardly advanced since Einstein's works of 1906, 1909 and 1917 were opened up.[33]

Compton's remark in his reminiscences in the year before he died is revealing:

> These experiments were the first to give, at least to physicists in the United States, a conviction of the fundamental validity of the quantum theory.[34]

17.9 The Threshold of Quantum Mechanics

In this case study, Albert Einstein's central role in the discovery of light quanta and the wave–particle duality have been emphasised. Over the years 1913 to 1925, many other key quantum phenomena were discovered which had to be built into a coherent theory of quantum mechanics. These discoveries resulted from brilliant experiments, most of them carried out by German experimental groups and supported by theoretical investigations, led by pioneers such as Bohr and Sommerfeld. Full details of these achievements are described in Part 2 *The Old Quantum Theory* of my book *Quantum Concepts in Physics*, Chapters 4 to 9. The time-line of discoveries and advances can be summarised as follows:

1911 First Solvay conference held at the Hotel Metropole, Brussels (Section 17.5).
1913 The Bohr model of the atom and the concept of *stationary states* which electrons can occupy in orbits about the atomic nucleus. The explanation of the Balmer series of hydrogen and ionised helium followed from the quantisation of angular momentum (Section 17.6.2).
1916 The explanation of the splitting of emission and absorption lines by magnetic and electric fields, the Zeeman and Stark effects respectively, in terms of the effects of quantisation upon atoms in electric and magnetic fields.
1916 Einstein's A and B coefficients (Section 17.7).
1918 The introduction of additional quantum numbers to account for multiplets in atomic spectra and the associated selection rules for permitted transitions.
1918 Bohr's correspondence principle, enabling classical physical concepts to inform the rules of the old quantum theory.
1919–23 Landé's vector model of the atom which could account for the splitting of spectral lines in terms of the addition of angular momenta according to the prescription given by the expression for the Landé g-factor.
1922 The experimental demonstration of space quantisation by the Stern–Gerlach experiment.

1922 The confirmation of the existence of light quanta as a result of Compton's X-ray scattering experiments (Section 17.8).
1922 The concept of electron shells in atoms and the understanding of the electronic structure of the periodic table.
1923 The introduction of de Broglie waves associated with the momentum of the electron and the extension of the wave–particle duality to particles as well as to light quanta.
1924 The discovery of Bose–Einstein statistics which differed very significantly from classical Boltzmann's statistics (Section 15.6).
1925 The discovery of Pauli's exclusion principle and the requirement of four quantum numbers to describe the properties of atoms.
1925 The discovery of electron spin as a distinctive quantum property of all particles – although inferred from angular momentum concepts, the intrinsic spin of the electron is not an 'angular momentum' in the sense of rotational motion.

The old quantum theory could not accommodate these features of atomic physics into a single coherent theory. In essence, the old quantum theory should be thought of as classical physics with the Bohr–Sommerfeld quantisation rules bolted on, namely,

$$\oint p_\phi \, \mathrm{d}\phi = nh, \tag{17.74}$$

as well as a set of selection rules to ensure that the observed optical and X-ray spectral features could be reproduced. Jammer summarises the situation admirably:

> In spite of its high-sounding name and its successful solutions of numerous problems in atomic physics, quantum theory, and especially the quantum theory of polyelectron systems, prior to 1925, was, from the methodological point of view, a lamentable hodgepodge of hypotheses, principles, theorems and computational recipes rather than a logical consistent theory. Every single quantum-theoretic problem had to be solved first in terms of classical physics; its classical solution had then to pass through the mysterious sieve of the quantum conditions or, as it happened in the majority of cases, the classical solution had to be translated into the language of quanta in conformance with the correspondence principle. Usually, the process of finding the 'correct solution' was a matter of skillful guessing and intuition, rather than of deductive or systematic reasoning.[35]

The big problem for all physicists was to discover a theory which could reconcile the apparently contradictory wave and particle properties of light. During the period 1905 to 1925, the fundamentals of classical physics were progressively undermined. The necessity of adopting quantum concepts became increasingly compelling with Planck's, Einstein's and Bohr's deep insights into the deficiencies of classical physics at the microscopic level. The 'old' quantum theory, which has been the subject of the last five chapters, was a patchwork of ideas and, as the theory developed in the hands of theorists such as Sommerfeld, became more and more *ad hoc* and unsatisfactory.

The breakthrough came with the discovery of quantum and wave mechanics by Werner Heisenberg and Erwin Schrödinger in 1925–27, leading to the relativistic wave equation for the electron by Paul Dirac in 1928. These established a new set of physical principles and a new mathematical framework for understanding the wave–particle duality which had caused such distress to physicists of the first two decades of the twentieth century. How they negotiated what Born referred to as not even ‘a straight staircase upward’ but rather a tangle of ‘interconnected alleys’ is told in Part III of my book *Quantum Concepts in Physics*.

The resolution of the problem is beautifully expressed by Stephen Hawking:

> The theory of quantum mechanics is based on an entirely new type of mathematics that no longer describes the real world in terms of particles and waves; it is only observations of the world that may be described in these terms.[36]

17.10 The Story Concluded

The story which has unfolded over the last five chapters is one of the most exciting and dramatic intellectual developments in theoretical physics, and indeed in the whole of science. The contributions of many very great physicists were crucial in reaching this new level of understanding. In the course of reviewing this text, I have read again many of Einstein’s papers and have been astounded by the imagination and fertility of his thinking. It is not only the quality, but also the sheer quantity, of papers spanning the whole of physics

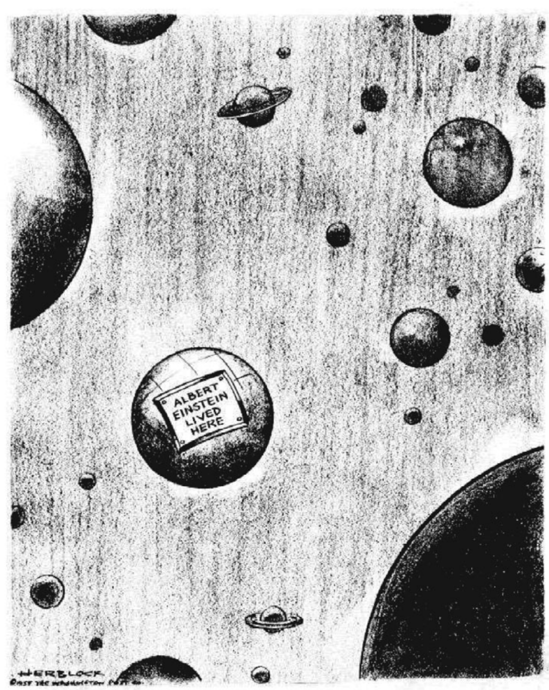


Fig. 17.6 Cartoon by Herblock in memory of Albert Einstein, who died on 15 April 1955. (From *Herblock’s Here and Now*, New York: Simon and Schuster.)

which is quite staggering. What comes through is a sustained clarity of thinking and a sureness of intuition in tackling head-on the most challenging problems facing physicists in the period 1905 to 1925. If relativity and quanta, leading to quantum mechanics, were the greatest discoveries of twentieth century physics, the person at the heart of both of these developments was undoubtedly Albert Einstein.

I have been asked by non-scientists whether or not Einstein deserves his reputation as the outstanding figure of modern science as he is portrayed in the popular press (Fig. 17.6). The case study of the last five chapters and those to come on special and general relativity should leave the reader in no doubt that Einstein's contributions transcend those of virtually all other physicists, the only physicists who can seriously be mentioned in the same breath being Newton and Maxwell. I am reminded of an occasion in the 1970s when I had dinner in Moscow with Academician V.L. Ginzburg. He recounted how Landau, the greatest Soviet theoretical physicist of the twentieth century, put physicists into leagues, some belonging to the first division, others to the second division and so on. The first division contained names such as Newton and Maxwell; he rated his own contribution rather modestly. But there was a separate division, the zero division, in which there was only one physicist – Albert Einstein.

Notes

1 Planck, M. (1907). Letter to Einstein of 6 July 1907, Einstein Archives, Princeton, New Jersey, quoted by A. Hermann (1971) in *The Genesis of Quantum Theory (1899–1913)*, p. 56. Cambridge, Massachusetts: MIT Press.
2 Lorentz, H.A. (1909). Letter to W. Wein of 12 April 1909, quoted by A. Hermann (1971) in *The Genesis of Quantum Theory (1899–1913)*, p. 56. Cambridge, Massachusetts: MIT Press.
3 Einstein, A. (1909). *Physikalische Zeitschrift*, **10**, 185.
4 Nyquist, H. (1928). Thermal Agitation of Electric Charge in Conductors. *Physical Review*, **32**, 110–113.
5 Johnson, J. (1928). Thermal Agitation of Electricity in Conductors. *Physical Review*, **32**, 97–109.
6 Einstein, A. (1910). Letter to Laub, 16 March 1910. See Einstein Archives, *op. cit.* and Kuhn, T.S. (1978). *Black-Body Theory and the Quantum Discontinuity, 1894–1912*, p. 214. Oxford: Clarendon Press.
7 Planck, M. (1910). Letter to W. Nernst of 11 June 1910. See Kuhn (1978). *op. cit.*, p. 230.
8 P. Langevin and M. De Broglie (eds.) (1912). *La Théorie du rayonnement et les quanta: Rapports et discussions de la réunion tenue à Bruxelles, du 30 octobre au 3 novembre 1911*. Paris: Gauthier-Villars.
9 Millikan, R.A. (1916). *Physical Review*, **7**, 355.
10 Longair, M.S. (2013). *Quantum Concepts in Physics: An Alternative Approach to the Understanding of Quantum Mechanics*. Cambridge: Cambridge University Press.
11 Thomson, J.J. (1897). *Philosophical Magazine*, **44**, 311.
12 Thomson, J.J. (1907). *Conduction of Electricity through Gases*. Cambridge: Cambridge University Press.
13 Longair, M.S. (2011). *High Energy Astrophysics*, 3rd edition. Cambridge: Cambridge University Press.

14 Rutherford, E. and Royds, T. (1908). *Philosophical Magazine*, **17**, 281, 1909.
15 See Andrade, E.N. da C. (1964). *Rutherford and the Nature of the Atom*, p. 111. New York: Doubleday.
16 Rutherford, E. (1911). *Philosophical Magazine*, **21**, 669.
17 Geiger, H. and Marsden, E. (1913). *Philosophical Magazine*, **25**, 604.
18 See Heilbron, J. (1978). *op. cit.*, 70.
19 Einstein, A. (1907). *Annalen der Physik*, **22**, 184.
20 Haas, A.E. (1910). *Wiener Berichte IIa*, **119**, 119; *Jahrbuch der Radioaktivität und Elektronik*, **7**, 261; *Physikalische Zeitschrift*, **11**, 537.
21 Bohr, N. (1912). Bohr's Memorandum is reproduced in *On the Constitution of Atoms and Molecules*, ed. L. Rosenfeld. Copenhagen: Munksgaard (1963).
22 Nicholson, J.W. (1911). *Philosophical Magazine*, **22**, 864. (1912) *Monthly Notices of the Royal Astronomical Society*, **72**, 677. See also, McCormmach, R. (1966). The Atomic Theory of John William Nicholson, *Archive for History of Exact Sciences*, **2**, 160.
23 Heilbron, J. (1978). *History of Twentieth Century Physics*, Proceedings of the International School of Physics 'Enrico Fermi', Course 57, p. 40. New York and London: Academic Press.
24 Bohr, N. (1913). *Philosophical Magazine*, **26**, 1, 476, 857.
25 Pais, A. (1991). *Niels Bohr's Times: In Physics, Philosophy and Polity*, p. 154. Oxford: Clarendon Press.
26 Einstein, A. (1916). *Physikalischen Gesellschaft Zürich Mitteilungen*, **16**, p. 47. See also, *Verhandlungen der Deutschen Physikalischen Gesellschaft* **18**, 318 and *Zeitschrift für Physik*, **18**, 121 (1917).
27 Lorentz, H.A. (1912). See note 6, pp. 35–39.
28 Pauli, W. (1923). *Zeitschrift für Physik*, **18**, p. 272.
29 Einstein, A. and Ehrenfest, P. (1923). *Zeitschrift für Physik*, **19**, p. 301.
30 The subject of Comptonisation is an important aspect of high energy astrophysics. I have given a simple introduction to the subject in Section 9.4 of my book *High Energy Astrophysics*, 3rd edition, 2011. Cambridge: Cambridge University Press. See also the extensive survey by Pozdnyakov, L.A., Sobol, I.M. and Sunyaey, R.A. (1983). *Soviet Scientific Reviews, Section E: Astrophysics and Space Physics Reviews*, **2**, 189–331.
31 Compton, A.H. (1923). *Physical Review*, **21**, p. 483.
32 See, for example, note 13, p. 235 *et seq.*
33 Heisenberg, W. (1929). The Development of the Quantum Theory 1918–1928, *Naturwissenschaften*, **17**, 491. (See translation of the quotation in R.H. Stewer, *The Compton Effect*, p. 287. New York: Science History Publications.)
34 Compton, A.H. (1961). *American Journal of Physics*, **29**, p. 817.
35 Jammer, M. (1989). *The Conceptual Development of Quantum Mechanics*, 2nd edition, New York: American Institute of Physics and Tomash Publishers. The first edition of Jammer's important book was published in 1966 by the McGraw-Hill Book Company in the International Series of Pure and Applied Physics. The 1989 edition is an enlarged and revised version of the first edition and I use this as the primary reference to Jammer's history.
36 Hawking, S.W. (1988). *A Brief History of Time: From the Big Bang to Black Holes*, p. 56. London: Bantam Books.

Case Study VI

SPECIAL AND GENERAL RELATIVITY

Albert Einstein's name is inextricably associated with the discoveries of the special and general theories of relativity. How they came about were, however, very different. What became known as the special theory of relativity was of the greatest interest and importance for many physicists and was tackled by many distinguished theories. Einstein's contribution was to provide a brilliant, coherent derivation of the transformations needed to replace those of Galileo, which were firmly embedded in the Newtonian picture of the world. In developing the special theory, Einstein crucially demonstrated that we live in four-dimensional space-time and not three-dimensional space and a separate time coordinate. In the case of general relativity, Einstein was ploughing a lone furrow for many years and many thought he was wasting his time on what seemed a quite intractable problem. But, as so often, Einstein's instincts and insights resulted in pure gold. The final version of the theory of 1915 has been validated with exquisite precision over the succeeding century or more.

Special Relativity

Special relativity is a subject in which intuition is a dangerous tool. Statements such as 'moving clocks go slow' or 'moving rulers are length contracted' are common currency but, in my view, they can hinder rather than help an understanding of the content of special relativity. In these non-intuitive areas, it is essential to have at hand a robust formalism which can be relied upon to give the correct answers. My own view is that the best approach is to introduce the Lorentz transformations and four-vectors as quickly as possible. These provide a set of simple rules which enable most reasonably straightforward problems in special relativity to be tackled with confidence. This approach involves constructing a four-vector algebra and calculus with rules essentially identical to those of three-vectors.

On researching the history of special relativity, it became apparent that many of my favorite tricks were well known to the pioneers, and so this case study begins with a modest historical introduction. I would emphasise that this is not intended as a course in special relativity, in which every 'i' is dotted and every 't' crossed. Rather the emphasis is upon the creation of a robust mathematical structure which enables relativistic calculations to be carried out as simply as possible.

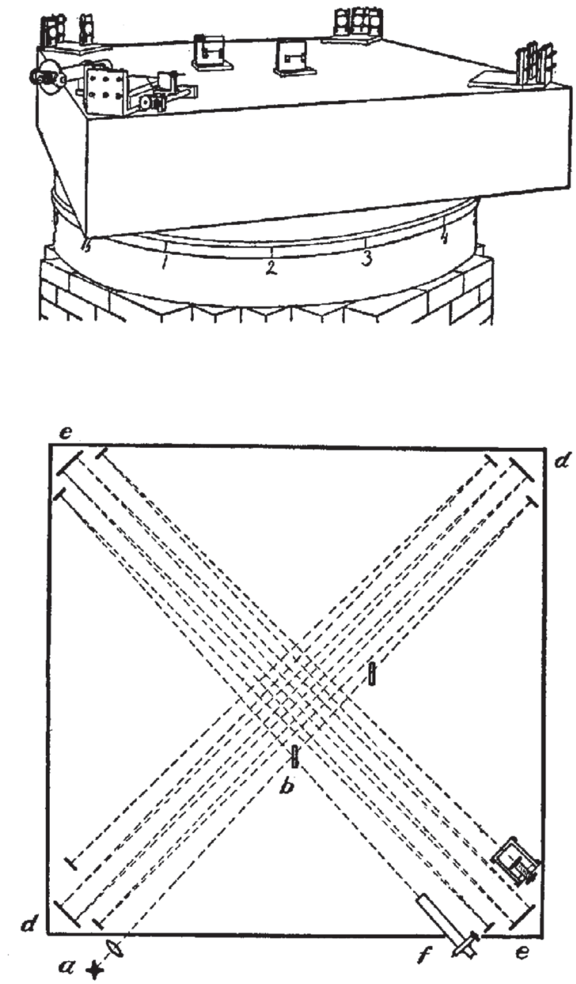


Fig. VI.1 The Michelson–Morley experiment of 1887. (From A.A. Michelson, 1927, *Studies in Optics*. Chicago: University of Chicago Press.)

Central to the development of the special theory of relativity is the great Michelson–Morley experiment of 1887 (Fig. VI.1). In a departure from my concentration upon the theoretical structure, a little space is devoted to that remarkable experiment (Section 18.1.1).

General Relativity

General relativity follows on rather naturally from special relativity and results in even more profound changes to our concepts of space and time than those which follow from special relativity. The idea that the geometry of space-time is influenced by the distribution of matter, which then moves along paths in distorted space-time, is one of the fundamental concepts of modern physics. Indeed, the bending of space-time is now regularly observed

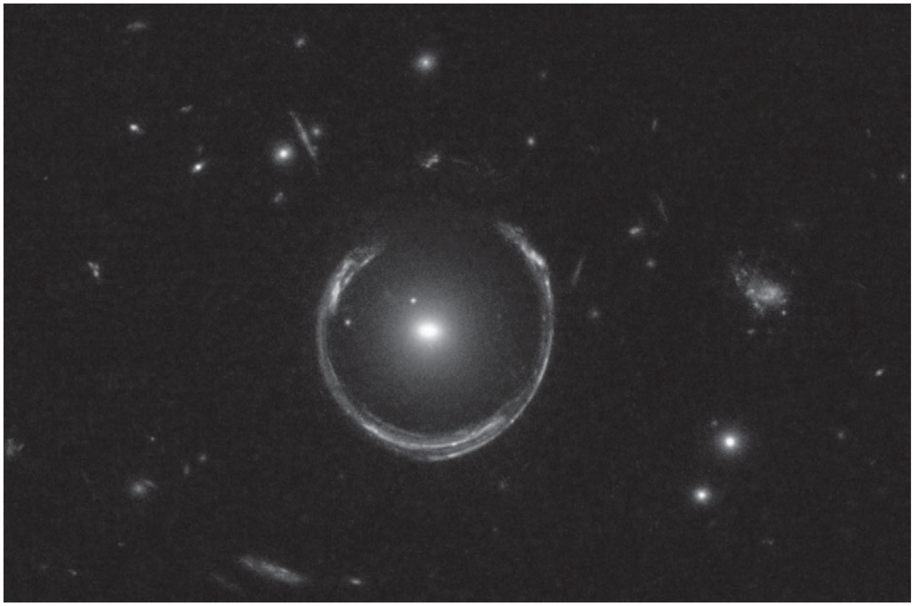


Fig. VI.2 The Horseshoe gravitational lens as observed by the Hubble Space Telescope. The giant elliptical galaxy acts as a gravitational lens, distorting the image of a distant background galaxy into an almost perfect ring. The unlensed image of the background galaxy can be reconstructed, knowing the distribution of mass in the lensing galaxy. (Courtesy of NASA and the STScI.)

as the gravitational lens effect in deep astronomical images (Fig. VI.2). Unfortunately, general relativity is technically complex, in the sense that the necessary mathematical tools go beyond what can normally be introduced in an undergraduate physics course, and it would be wrong to underestimate the technical difficulties. Nonetheless, a great deal can be achieved without the use of advanced techniques, provided the reader is prepared to accept a few results which will simply be made plausible and the exact results quoted. This seems a small price to pay for some insight into the intimate connection between space-time geometry and gravitation and for understanding some of the more remarkable phenomena which are expected to be observed in strong gravitational fields, such as those found in the vicinity of black holes and in the emission of gravitational waves.

The case for incorporating these more demanding technical developments is reinforced by the essential role general relativity plays in contemporary cosmology, as expounded in Case Study VII, and by the dramatic discovery of gravitational waves from a pair of coalescing black holes in 2016, exactly 100 years after Einstein's prediction of the existence of such waves. In these events, we observe directly the full range of the effects of general relativity in the strong field limit, something which could only be dreamed about in the twentieth century.

18 Special Relativity: A Study in Invariance

18.1 Introduction

There is an appealing logic about the standard route to the discovery of the Lorentz transformations and relativistic dynamics, which can be summarised as follows:

- Bradley's observations of stellar aberration of 1727–28, implying that the Earth moves through a stationary aether;
- The null result of the Michelson–Morley experiment of 1887, indicating that there is no detectable motion of the Earth through the aether;
- Einstein's Principle of Relativity and the second postulate that the speed of light in vacuo is the same for observers in all inertial frames of reference;
- Derivation of the Lorentz transformations and relativistic kinematics;
- Invariance of the laws of physics under Lorentz transformation;
- Relativistic dynamics;
- $E = mc^2$, Maxwell's equations, etc.

As in many of the most dramatic turning points in the history of physics, the route to the special theory of relativity was somewhat more tortuous. The origins of special relativity have been the subject of a great deal of study by historians of science, two highly accessible surveys of the key events being presented by John Stachel.[1] More details of the physical arguments which led to the special theory of relativity are contained in Chapters 6–8 of the study by Abraham Pais *'Subtle is the Lord . . .' The Science and Life of Albert Einstein*.[2]

The emphasis of this case study is on the concept of invariance, rather than on the history, but some of the events which led up to Einstein's great papers of 1905 illuminate the terminology and set them in context.

18.1.1 The Michelson–Morley Experiment

Maxwell's last brief paper was published posthumously in the *Proceedings of the Royal Society* in 1880 and concerned a means of measuring the speed of the Earth though the hypothetical aether. His idea concerned the use of accurate timing of the eclipses of the satellites of Jupiter as a means of measuring the speed of light relative to the aether in different directions on the sky, a procedure similar to that of Rømer in 1676 in his first estimate of the speed of light. The content of the paper concerned a letter from Maxwell to David P. Todd, the Director of the US Nautical Almanac Office in Washington DC, enquiring about the accuracy with which the orbits of Jupiter and its satellites were known.

In his letter to Todd, Maxwell discussed the great difficulty of performing terrestrial experiments to measure the velocity light through the aether:

> Even if we were sure of the theory of aberration, we can only get differences of position of stars, and in the terrestrial methods of determining the velocity of light, the light comes back along the same path again, so that the velocity of the earth with respect to the ether would alter the time of the double passage by a quantity depending on the square of the ratio of the earth's velocity to that of light, and this is quite too small to be observed.[3]

This remark was picked up by Albert Michelson who recognised that, although very small, a terrestrial test could be made using optical interferometry. In his book *Light Waves and their Uses*,[4] Michelson stated:

> Maxwell considered it possible, theoretically at least, to deal with the square of the ratio of two velocities; that is the square of 1/10 000 or 1/100 000 000. He further indicated that if we made two measurements of the velocity of light, one in the direction in which the earth is travelling, and one in a direction at right angles to this, then the time it takes light to pass over the same length of path is greater in the first case than in the second.

Maxwell had described the principles of optical interferometry in his *Encyclopaedia Britannica* article on the aether,[5] but he does not seem to have appreciated its potential for measuring such tiny path differences. Michelson, an experimenter of genius and the first American citizen to win a Nobel Prize in science, recognised the power of optical interferometry in measuring extremely small path differences between light rays. He made his first attempt to carry out the experiment in 1881 but the interferometer baseline was too small to obtain a significant result – no fringe shift was observed. In addition, there was an error in his calculation of the amplitude of predicted fringe shift and so the first experiments were inconclusive.

A much improved version of the experiment, shown in Fig. VI.1 was carried out with Edward Morley – the famous Michelson–Morley experiment of 1887.[6] In Michelson's words,

> ... the interferometer was mounted on a block of stone 1.5 m square and 0.25 m thick resting on an annular wooden ring which floated the whole apparatus on mercury.[7]

The lower diagram of Fig. VI shows the multiple reflections used to increase the optical path lengths of the arms of the interferometer.

The experiment was performed by observing the movement of the central fringe of the interferometer as the apparatus was rotated 'fairly uniformly and continuously' through 360° and measurements made every one-sixteenth of a revolution. In Fig. 18.1, the mean displacements of the central fringe during rotation through 360° are compared with the expected sinusoidal variation if the Earth moved through a stationary aether at 30 km s^{-1}, the dotted line corresponding to only one-eighth of the theoretical displacement. Again, in Michelson's words,

> It must be concluded that the experiment shows no evidence of a displacement greater than 0.01 fringe ... With $V/c = 1/10\,000$, this gives an expected displacement of 0.4 fringes. The actual value is certainly less than one-twentieth of this actual amount and probably less than one-fortieth.

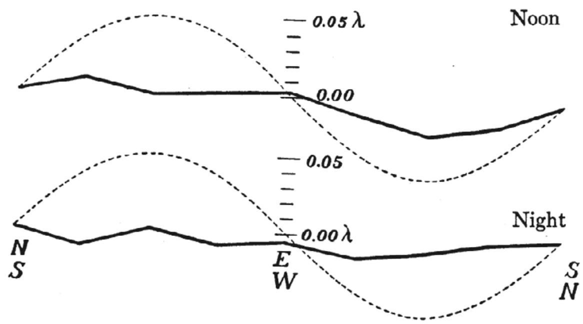


Fig. 18.1 The null result of the Michelson–Morley experiment. The solid line shows the average movement of the central fringe as the apparatus was rotated through 360°. The dotted line shows one eighth of the expected sinusoidal variation if the Earth moves through the stationary aether at 30 km s^{-1}. (From A.A. Michelson, 1927, *Studies in Optics*. Chicago: University of Chicago Press.)

The experiment was repeated at different times of the year by Morley and Miller, in case the motion of the Earth about the Sun was cancelled out by the drift of the aether, but the same null result was found.

In the 1962 reprint of Michelson's book *Studies in Optics*, Harvey B. Lemon writes in his introduction:

> To the complete astonishment and mystification of the scientific world this refined experiment also yielded absolutely negative results. Again must we note at this point the universal confidence with which any experimental fact announced by Michelson was instantly accepted. Not for twenty years did anyone have the temerity to challenge his conclusion.

This may well be the reason that there is no explicit reference to the Michelson–Morley experiment in Einstein's great paper of 1905 – Michelson's null result instantly became one of the established facts of the thorny problem of understanding the nature of the aether.

18.1.2 Maxwell's Equations and Voigt (1887)

As discussed in Section 6.8, it was some time after 1865 before Maxwell's formulation of the laws of electromagnetism, as encapsuled in the compact set of four equations (5.44), was fully accepted. Maxwell's equations were put into their familiar guise thanks to the efforts of Heaviside, Helmholtz and Hertz. Hertz's demonstration that electromagnetic phenomena are propagated at the speed of light and that electromagnetic waves have all the properties of light waves were persuasive evidence that Maxwell's equations indeed encapsulated in compact form the known phenomena of electromagnetism and unified the physics of electromagnetism and light.

This triumph came, however, at a cost. All known wave phenomena resulted from the perturbation of some material medium and so the nature of the medium through which electromagnetic waves were supposed to be propagated, the aether, became a central concern of late nineteenth century physics.[8] Bradley's observation of stellar aberration showed that the Earth must be moving through a stationary aether. To account for these observations, the Galilean expression for the addition of velocities is used, which follows from the Galilean transformations (3.3) under which Newton's laws of motion are form invariant. A problem for nineteenth century physicists was that, unlike Newton's laws of motion, Maxwell's equations for the electromagnetic field are not form invariant under Galiliean transforms.

To demonstrate this, the Galilean transformations can be used to derive expressions which describe how partial derivatives transform between inertial frames of reference. Thus, if we write

$$\begin{cases} t' = t, \\ x' = x - Vt, \\ y' = y, \\ z' = z, \end{cases} \tag{18.1}$$

using the chain rule, for example,

$$\frac{\partial}{\partial t} = \frac{\partial t'}{\partial t}\frac{\partial}{\partial t'} + \frac{\partial x'}{\partial t}\frac{\partial}{\partial x'} + \frac{\partial y'}{\partial t}\frac{\partial}{\partial y'} + \frac{\partial z'}{\partial t}\frac{\partial}{\partial z'} = \frac{\partial}{\partial t'} - V\frac{\partial}{\partial x'}, \tag{18.2}$$

the transforms of the partial derivatives become

$$\begin{cases} \dfrac{\partial}{\partial t} \to \dfrac{\partial}{\partial t'} - V\dfrac{\partial}{\partial x'}, \\ \dfrac{\partial}{\partial x} \to \dfrac{\partial}{\partial x'}, \\ \dfrac{\partial}{\partial y} \to \dfrac{\partial}{\partial y'}, \\ \dfrac{\partial}{\partial z} \to \dfrac{\partial}{\partial z'}. \end{cases} \tag{18.3}$$

The problem can be appreciated by transforming Maxwell's equations from the laboratory frame to a frame moving at velocity V along the positive x-axis. Maxwell's equations *in free space* can be written in Cartesian coordinates:

$$\begin{cases} \dfrac{\partial E_z}{\partial y} - \dfrac{\partial E_y}{\partial z} = -\dfrac{\partial B_x}{\partial t}, \\ \dfrac{\partial E_x}{\partial z} - \dfrac{\partial E_z}{\partial x} = -\dfrac{\partial B_y}{\partial t}, \\ \dfrac{\partial E_y}{\partial x} - \dfrac{\partial E_x}{\partial y} = -\dfrac{\partial B_z}{\partial t}. \end{cases} \tag{18.4}$$

$$\begin{cases} \dfrac{\partial B_z}{\partial y} - \dfrac{\partial B_y}{\partial z} = \dfrac{1}{c^2}\dfrac{\partial E_x}{\partial t}, \\ \dfrac{\partial B_x}{\partial z} - \dfrac{\partial B_z}{\partial x} = \dfrac{1}{c^2}\dfrac{\partial E_y}{\partial t}, \\ \dfrac{\partial B_y}{\partial x} - \dfrac{\partial B_x}{\partial y} = \dfrac{1}{c^2}\dfrac{\partial E_z}{\partial t}. \end{cases} \tag{18.5}$$

$$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} = 0. \tag{18.6}$$

$$\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} = 0. \tag{18.7}$$

Substituting for the partial derivatives in (18.4) and (18.5), and using (18.6) and (18.7) as transformed into the S' frame of reference, (18.4) and (18.5) become

$$\begin{cases} \dfrac{\partial(E_z + VB_y)}{\partial y'} - \dfrac{\partial(E_y - VB_z)}{\partial z'} = -\dfrac{\partial B_x}{\partial t'}, & (a) \\ \dfrac{\partial E_x}{\partial z'} - \dfrac{\partial(E_z + VB_y)}{\partial x'} = -\dfrac{\partial B_y}{\partial t'}, & (b) \\ \dfrac{\partial(E_y - VB_z)}{\partial x'} - \dfrac{\partial E_x}{\partial y'} = -\dfrac{\partial B_z}{\partial t'}, & (c) \end{cases} \tag{18.8}$$

and

$$\begin{cases} \dfrac{\partial(B_z - VE_y)}{\partial y'} - \dfrac{\partial(B_y + VE_z)}{\partial z'} = \dfrac{1}{c^2}\dfrac{\partial E_x}{\partial t'}, & (a) \\ \dfrac{\partial B_x}{\partial z'} - \dfrac{\partial(B_z - VE_y)}{\partial x'} = \dfrac{1}{c^2}\dfrac{\partial E_y}{\partial t'}, & (b) \\ \dfrac{\partial(B_y + VE_z)}{\partial x'} - \dfrac{\partial B_x}{\partial y'} = \dfrac{1}{c^2}\dfrac{\partial E_z}{\partial t'}. & (c) \end{cases} \tag{18.9}$$

We now seek new definitions for the components of $\boldsymbol{E}$ and $\boldsymbol{B}$ in the moving frame S'. For example in the first term of (18.8(a)) and the second term of (18.8(b)), we might write $E_z' = E_z + VB_y$, but then we would run into trouble with the last term in (18.9(c)) E_z. Similarly, although we might be tempted to write $B_y' = B_y + VE_z$ in the second term of (18.9(a)) and the first term of (18.9(c)), this runs into serious trouble with the last term of (18.8(b)).

It can be seen that the only frame of reference in which the Galilean transformations result in a self-consistent set of equations is that for which $V = 0$. For this reason, it was believed that Maxwell's equations would hold true only in the frame of reference of the stationary aether – they were considered to be *non-relativistic* under Galilean transformations.

Remarkably, in 1887, Woldmar Voigt[9] noted that Maxwell's wave equation for electromagnetic waves,

$$\nabla^2 \boldsymbol{H} - \frac{1}{c^2}\frac{\partial^2 \boldsymbol{H}}{\partial t^2} = 0, \tag{18.10}$$

is form invariant under the transformation

$$\begin{cases} t' = t - \dfrac{Vx}{c^2}, \\ x' = x - Vt, \\ y' = y/\gamma, \\ z' = z/\gamma, \end{cases} \tag{18.11}$$

where $\gamma = (1 - V^2/c^2)^{-1/2}$. Except for the fact that the transformations on the right-hand side have been divided by the Lorentz factor γ, the set of equations (18.11) is the Lorentz transformation. Voigt derived this expression using the invariance of the phase of a propagating electromagnetic wave, exactly the approach we will take in Section 18.2. This work was unknown to Lorentz when he set about deriving what are now known as the Lorentz transformations.

18.1.3 Fitzgerald (1889)

The Irish physicist George Francis Fitzgerald is rightly credited with the remarkable insight that the null result of the Michelson–Morley experiment could be explained by supposing that moving objects are length contracted in their direction of motion. His brief paper entitled *The Aether and the Earth's Atmosphere*[10] published in *Science* in 1889, contains the statement:

> I have read with much interest Messrs. Michelson and Morley's wonderfully delicate experiment attempting to decide the important question as to how far the aether is carried along by the Earth. Their result seems opposed to other experiments showing that the aether in the air can only be carried along only to an inappreciable extent. I would suggest that almost the only hypothesis that can reconcile this opposition is that the length of material bodies changes, according as they are moving through the aether or across it, by an amount depending on the square of the ratio of their velocity to that of light. We know that electric forces are affected by the motion of electrified bodies relative to the aether, and it seems a not improbable supposition that the molecular forces are affected by the motion, and that the size of the body alters consequently.

This quotation amounts to more than 60% of his brief note. Remarkably, this paper, by which Fitzgerald is best remembered, was not included in his complete works edited by Larmor in 1902. Lorentz knew of the paper in 1894, but Fitzgerald was uncertain whether or not it had been published when Lorentz wrote to him. The reason was that *Science* went bankrupt in 1889 and was only refounded in 1895. Notice that Fitzgerald's proposal was only qualitative and that he was proposing a real physical contraction of the body in its direction of motion because of interaction with the aether.

18.1.4 Lorentz

Hendrik Lorentz had agonised over the null result of the Michelson–Morley experiment and in 1892 came up with same suggestion as Fitzgerald, but with a quantitative expression for the length contraction. In his words,

> This experiment has been puzzling me for a long time, and in the end I have been able to think of only one means of reconciling it with Fresnel's theory. It consists in the supposition that the line joining two points of a solid body, if at first parallel to the direction of the Earth's motion, does not keep the same length when subsequently turned through 90°.[11]

Lorentz worked out that the length contraction had to amount to

$$l = l_0 \left(1 - \frac{V^2}{2c^2}\right), \tag{18.12}$$

which is just the low velocity limit of the expression $l = l_0/\gamma$, where $\gamma = (1 - V^2/c^2)^{-1/2}$. Subsequently, this phenomenon has been referred to as *Fitzgerald–Lorentz contraction*.

In 1895, Lorentz[12] tackled the problem of the transformations which would result in form invariance of Maxwell's equations and derived the following relations, which in SI notation are:

$$\begin{cases} x' = x - Vt, \quad y' = y, \quad z' = z, \\ t' = t - \dfrac{Vx}{c^2}, \\ \boldsymbol{E}' = \boldsymbol{E} + \boldsymbol{V} \times \boldsymbol{B}, \\ \boldsymbol{B}' = \boldsymbol{B} - \dfrac{\boldsymbol{V} \times \boldsymbol{E}}{c^2}, \\ \boldsymbol{P}' = \boldsymbol{P}, \end{cases} \tag{18.13}$$

where $\boldsymbol{P}$ is the polarisation. Under this set of transformations, Maxwell's equations are form invariant to *first order* in V/c. Notice that time is no longer absolute. Lorentz apparently considered this simply to be a convenient mathematical tool in order to ensure form invariance to first order in V/c. He called t the *general time* and t' the *local time*. In order to account for the null result of the Michelson–Morley experiment, he had to include an additional second-order compensation factor, the Fitzgerald–Lorentz contraction $(1 - V^2/c^2)^{-1/2}$, into the theory. One feature of this paper was the *assumption* that the force on an electron should be given by the first-order expression

$$\boldsymbol{f} = e(\boldsymbol{E} + \boldsymbol{V} \times \boldsymbol{B}). \tag{18.14}$$

This is the origin of the expression for the *Lorentz force* for the joint action of electric and magnetic fields on a charged particle, although, as we showed in Section 6.3, these terms were already present in Maxwell's paper of 1865.

Einstein knew of Lorentz's paper of 1895, but was unaware of his subsequent work. In 1899, Lorentz[13] established the invariance of the equations of electromagnetism to all orders in V/c through a new set of transformations:

$$\begin{cases} x' = \epsilon\gamma(x - Vt), \quad y' = \epsilon y, \quad z' = \epsilon z, \\ t' = \epsilon\gamma\left(t - \dfrac{Vx}{c^2}\right). \end{cases} \tag{18.15}$$

These are the Lorentz transformations, including the scale factor ϵ. By this means, he was able to incorporate length contraction into the transformations. Almost coincidentally, in 1898, Joseph Larmor wrote his prize winning essay *Aether and Matter*,[14] in which he derived the standard form of the Lorentz transformations and showed that they included the Fitzgerald–Lorentz contaction.

In his major paper of 1904, entitled *Electromagnetic Phenomena in a System Moving with Any Velocity Smaller than Light*,[15] Lorentz presented the transformations (18.15) with $\epsilon = 1$. However, as expressed by Gerald Holton, the apparatus used to arrive at this result contained a number of 'fundamental assumptions'. In Holton's words,

> (The theory) contained in fact eleven *ad hoc* hypotheses: restriction to small ratios of velocities V to light velocity c, postulation *a priori* of the transformation equations . . . , assumption of a stationary aether, assumption that the stationary electron is round, that its charge is uniformly distributed, that all mass is electromagnetic, that the moving electron changes one of its dimensions precisely in the ratio $(1 - V^2/c^2)^{1/2}$ to one, ...[16]

The complications arose because, in Lorentz's approach, the transformations were part of his theory of the electron and were postulated *a priori* as part of the aether theory of electromagnetism. Einstein was not aware of this paper in 1905.

18.1.5 Poincaré

Henri Poincaré made a number of crucial contributions to the development of special relativity. In a remarkable paper of 1898, entitled *La Mesure du Temps*,[17] he identified the problem of what is meant by simultaneity and the measurement of time intervals. In the conclusion of his paper, he writes,

> The simultaneity of two events or the order of their succession, as well as the equality of two time intervals, must be defined in such a way that the statement of the natural laws be as simple as possible. In other words, all rules and definitions are but the result of unconscious opportunism.

In 1904, Poincaré surveyed current problems in physics and included the statement:

> . . . the principle of relativity, according to which the laws of physical phenomena should be the same whether for an observer fixed or for an observer carried along by uniform movement or translation.[18]

Notice that this is no more than a restatement of Galilean relativity. He concluded by remarking,

> Perhaps likewise, we should construct a whole new mechanics, . . . where, inertia increasing with the velocity, the velocity of light would become an impassible limit.

18.1.6 Einstein before 1905

Albert Einstein had been wrestling with exactly these problems since 1898. He was certainly aware of Poincaré's writings and these were the subject of intense discussion with his scientific colleagues in Bern. According to Einstein in a letter of 25 April 1912 to Paul Ehrenfest,

> I knew that the principle of the constancy of the velocity of light was something quite independent of the relativity postulate and I weighted which was the more probable, the principle of the constancy of c, as required by Maxwell's equations, or the constancy of c exclusively for an observer located at the light source. I decided in favour of the former.[19]

In 1924, he stated:

> After seven years of reflection in vain (1898–1905), the solution came to me suddenly with the thought that our concepts of space and time can only claim validity insofar as they stand in a clear relation to our experiences; and that experience could very well lead to the alteration of these concepts and laws. By a revision of the concept of simultaneity into a more malleable form, I thus arrived at the special theory of relativity.[19]

Once he had discovered the concept of the relativity of simultaneity, it took him only five weeks to complete his great paper, *On the Electrodynamics of Moving Bodies*.[20] As we will discuss below, Einstein's great paper contains the two postulates of special relativity, in addition to which he made only four assumptions, one concerning the isotropy and homogeneity of space, the others concerning three logical properties involved in the definition of the synchronisation of clocks. Thus, Einstein's approach not only simplified greatly many of the ideas current at the time, but completely revolutionised our understanding of the nature of space and time.

Einstein did not regard what he had done as particularly revolutionary. In his own words,

> With respect to the theory of relativity, it is not at all a question of a revolutionary act, but a natural development of a line which can be pursued through the centuries.[21]

As we noted in Section 16.1.1, of the great papers published in 1905, Einstein regarded that on quanta as the really revolutionary paper of the three.

18.1.7 Reflections

Relativity is a subject in which the results are not at all intuitively obvious on the basis of our everyday experience. The number of books on the subject is immense, but, above all others, pride of place must go to Einstein's original paper of 1905 *On the Electrodynamics of Moving Bodies*.[20] It is as clear as any of the subsequent papers in expounding the basis of the theory. It is one of the miracles of theoretical physics that Einstein produced such a complete and elegant exposition with such profound implications for our understanding of the nature of space and time in this single paper.

In this case study, we investigate the special theory of relativity from the point of view of *invariance*. This is very much in the spirit of Einstein. We concentrate upon deriving the formulae of special relativity with the minimum number of assumptions and exposing clearly where we have to be guided by experiment. The approach is similar to that of Rindler in his excellent textbook *Relativity: Special, General and Cosmological*.[22] I find Rindler's exposition one of the most satisfying of all books on relativity.

18.2 Geometry and the Lorentz Transformation

As in the case of Galilean relativity, we begin with two inertial frames of reference moving with relative velocity V. We call S the laboratory frame of reference, or the local rest frame, and S' the moving frame, which has velocity V in the positive x-direction through S. We adopt rectangular Cartesian coordinate systems in the S and S' with their axes parallel – we refer to frames of reference S and S' in this orientation as 'standard configuration' (Fig. 18.2). We will not go into the formalities of showing that it is indeed possible to set up such a standard configuration – this is done by Rindler – but we will assume that it can be done.

The special theory of relativity as formulated by Einstein in his paper of 1905 contains two basic postulates. Let us quote the words of Einstein himself:

> The same laws of electrodynamics and optics are valid for all frames of reference for which the equations of mechanics hold good. We will raise this conjecture (the purport of which will hereafter be called the 'Principle of Relativity') to the status of a postulate and also introduce another postulate which is only apparently irreconcilable with the former, namely that light is always propagated in empty space with a definite velocity c which is independent of the state of motion of the emitting body.[23]

Rephrasing Einstein's postulates, the first is that the laws of physics are the same in all inertial frames of reference and, second, that the speed of light in vacuo should have the same value c in all inertial frames. The second postulate is crucial in leading to the special theory of relativity.

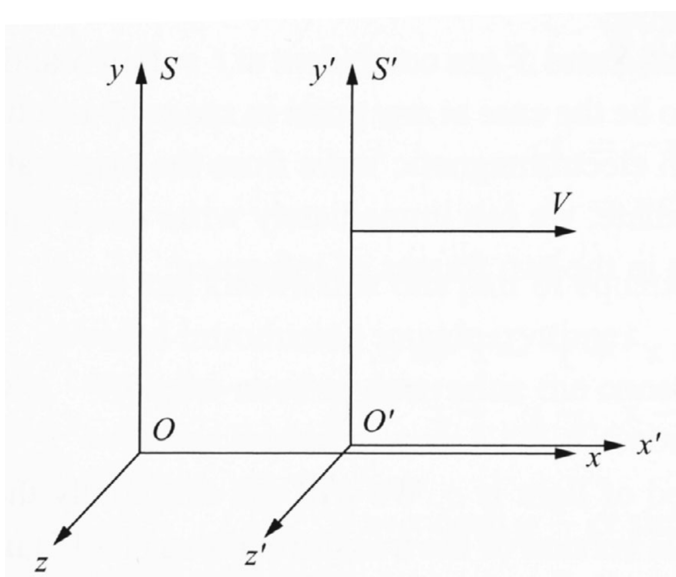


Fig. 18.2 Inertial frames of reference S and S' in standard configuration.

Suppose the origins of the inertial frames S and S' are coincident at $t = 0$ in S and $t' = 0$ in S'. It is always possible to arrange this to be the case at any point in space by resetting the clocks in S and S'. Now let us send out an electromagnetic wave from the origin at $t = 0, t' = 0$. Because of Einstein's second postulate, we can immediately write down the motion of the wavefront in the two frames of reference:

$$\begin{cases} \text{In } S, & c^2t^2 - x^2 - y^2 - z^2 = 0, \\ \text{In } S', & c^2t'^2 - x'^2 - y'^2 - z'^2 = 0, \end{cases} \tag{18.16}$$

guaranteeing that in both frames the speed of light is c. This is essentially the same argument used by Voigt to derive his version of the transformations (18.11) in 1887. Now the vectors $[ct, x, y, z]$ and $[ct', x', y', z']$ are simply coordinates in S and S' and therefore we seek a set of transformations which leave these expressions for the motion of the wavefront *form invariant*.

There is a formal analogy between (18.16) and the properties of the components of a three-vector under rotation. If two Cartesian frames of reference have the same origin and one is rotated through an angle θ with respect to the other, the norm (or the square of the magnitude) of a three-vector is the same in both frames, that is,

$$R^2 = x^2 + y^2 + z^2 = x'^2 + y'^2 + z'^2. \tag{18.17}$$

Let us consider the problem of transforming (18.16) in the same way. For simplicity, suppose the wave is propagated along the positive x-direction so that the transform we are seeking must result in

$$c^2t^2 - x^2 = c^2t'^2 - x'^2 = 0. \tag{18.18}$$

We first convert (18.18) into a form similar to (18.17) by writing $c\tau = \mathrm{i}ct$, $c\tau' = \mathrm{i}ct'$. Thus, we have introduced an *imaginary time* coordinate τ, a concept which caused so much grief to lay readers of Stephen Hawking's *A Brief History of Time*.[24] It is no more than a change of variable. Then

$$c^2\tau^2 + x^2 = c^2\tau'^2 + x'^2 = 0. \tag{18.19}$$

The necessary transformation formulae are just the rotation formulae for vectors in two-dimensions:

$$\begin{cases} c\tau' = -x\sin\theta + c\tau'\cos\theta, \\ x' = x\cos\theta + c\tau\sin\theta. \end{cases} \tag{18.20}$$

Notice that it is assumed that the 'time coordinate' is in the 'y-direction'. Now τ is imaginary and so the angle θ must be imaginary as well. To convert into real quantities, write $\theta = \mathrm{i}\phi$, where ϕ is real. Because

$$\begin{cases} \cos(\mathrm{i}\phi) = \cosh\phi, \\ \sin(\mathrm{i}\phi) = \mathrm{i}\sinh\phi, \end{cases}$$

we find

$$\begin{cases} ct' = -x \sinh \phi + ct \cosh \phi, \\ x' = x \cosh \phi - ct \sinh \phi. \end{cases} \tag{18.21}$$

If we had known that this pair of equations leaves $x^2 - c^2t^2$ form invariant, we could have avoided introducing imaginary time.

We now determine the constant ϕ. At time t in S, the origin of S' at $x' = 0$ has moved to $x = Vt$ in S. Substituting into the second relation of (18.21), we find

$$0 = x \cosh \phi - ct \sinh \phi, \quad \tanh \phi = x/ct = V/c.$$

Then, because $\cosh \phi = (1 - V^2/c^2)^{-1/2}$ and $\sinh \phi = (V/c)/(1 - V^2/c^2)^{-1/2}$, we find from (18.21)

$$\begin{cases} ct' = \gamma \left(ct - \dfrac{Vx}{c} \right), \\ x' = \gamma(x - ct), \end{cases} \tag{18.22}$$

where $\gamma = (1 - V^2/c^2)^{-1/2}$ is referred to as the *Lorentz factor* – it appears in virtually all calculations in special relativity. We also find that $y' = y$, $z' = z$ from the symmetry of the transformations between S and S'.

We have therefore derived the complete set of Lorentz transformations using the idea of form invariance and Einstein's second postulate, which is a statement about the invariance of the speed of light between inertial frames of reference. Let us write the Lorentz transformations in the standard form we will employ throughout this exposition:

$$\begin{cases} ct' = \gamma \left(ct - \dfrac{Vx}{c} \right), \\ x' = \gamma(x - ct), \\ y' = y, \\ z' = z, \end{cases} \qquad \gamma = \left(1 - \frac{V^2}{c^2} \right)^{-1/2}. \tag{18.23}$$

In the standard exposition of special relativity, the remarkable consequences of this set of transformations and the so-called 'paradoxes' which are supposed to arise are explored. Let me state immediately that there are no paradoxes – only some rather non-intuitive features of space-time which result from the Lorentz transformations. The origin of these phenomena is what Einstein recognised as the key feature of special relativity, the *relativity of simultaneity*. This concept elucidates the origin of some of the difficulties which students have in understanding relativity. As we have already stated, at any point in space, we can reset the clocks in S and S' so that they read the same time at that instant. In the above example, we can arrange $x' = 0$, $x = 0$; $t' = 0$, $t = 0$. At that point in space-time, the event is simultaneous in the two frames of reference. At all other points in space, however, events are not simultaneous.

From the first equation of (18.23), we see that $x = 0$, $t = 0$ implies $t' = 0$, that is, simultaneity. At all other values of x, the observers in S and S' disagree about the time at which an event occurs, that is,

$$ct' = \gamma\left(ct - \frac{Vx}{c}\right), \qquad (18.24)$$

and hence, if $x \neq 0, t' \neq t$. In other words, if observers in S and S' can agree on simultaneity at one point in space-time, they will disagree *at all other points*. This is the origin of the phenomena of time dilation, length contraction, twin paradoxes, etc. It is the fundamental difference between Galilean and special relativity. In Galilean relativity, observers in S and S' always agree everywhere about simultaneity. This is clear from the Newtonian limit of (18.24). If $V/c \to 0$, $\gamma \to 1$ and $t' = t$ everywhere.

18.3 Three-Vectors and Four-Vectors

Three-vectors provide a compact notation for writing the laws of physics in a form which is independent of the choice of reference frame. No matter how we rotate or translate the frame of reference, or change the system of coordinates used, the vector relations remain form invariant. For example:

(i) Vector addition is preserved:

$$\text{If} \quad \boldsymbol{a} + \boldsymbol{b} = \boldsymbol{c}, \qquad \boldsymbol{a}' + \boldsymbol{b}' = \boldsymbol{c}'.$$

(ii) Equally,

$$\text{If} \quad \boldsymbol{a} \cdot (\boldsymbol{b} + \boldsymbol{c}), \qquad \boldsymbol{a}' \cdot (\boldsymbol{b}' + \boldsymbol{c}').$$

(iii) The *norm* or *magnitude* of the three-vector is invariant with respect to rotations and displacements:

$$\boldsymbol{a}^2 = |\boldsymbol{a}|^2 = a_1^2 + a_2^2 + a_3^2 = a_1'^2 + a_2'^2 + a_3'^3.$$

(iv) Scalar products are invariant:

$$\boldsymbol{a} \cdot \boldsymbol{b} = a_1 b_1 + a_2 b_2 + a_3 b_3 = a_1', b_1' + a_2' b_2' + a_3' b_3'.$$

In other words, these vector relations express equalities which are independent of the frame of reference in which we perform our calculation.

Our objective is to find quantities similar to three-vectors which remain form invariant under Lorentz transformations. These *four-vectors* are objects which enable us to write down the laws of physics in relativistic form. The components of the four-vectors should transform like the components of the primitive four-vector $[ct, x, y, z]$ and we need to find expressions for these which relate them to physical quantities which can be measured in the laboratory.

There remains the thorny question of notation. I will use the Lorentz transformations in the form of (18.23) and define the components of the four-vectors as quantities which transform like ct and x, y, z. The components of the four-vectors are written in square brackets and I use bold, italic, capital letters to represent them. Thus,

$$\boldsymbol{R} \equiv [ct, x, y, z]. \qquad (18.25)$$

To work out the norm of any four-vector, the rule is

$$\boldsymbol{R}^2 = |\boldsymbol{R}|^2 = c^2t^2 - x^2 - y^2 - z^2, \tag{18.26}$$

that is, there is a plus sign in front of c^2t^2 and minus signs in front of x^2, y^2 and z^2 – in the language of relativity, the *signature* of the metric is $[1, -1, -1, -1]$. The time component has a different status from the spatial components, and so I will use the rule that the 'time component' has a subscript 0. Thus, if the four-vector is $\boldsymbol{A} = [A_0, A_1, A_2, A_3]$, its norm is $|\boldsymbol{A}|^2 = A_0^2 - A_1^2 - A_2^2 - A_3^2$, exactly equivalent to $R^2 = x^2 + y^2 + z^2$ for three-vectors. The transformation rules for the components of $\boldsymbol{A}$ can be found from (18.23), according to the identifications:

$$\begin{cases} ct' \to A_0' & ct \to A_0, \\ x' \to A_1' & x \to A_1, \\ y' \to A_2' & y \to A_2, \\ z' \to A_3' & z \to A_3. \end{cases} \tag{18.27}$$

In this section we are concerned with finding four-vectors for *kinematic* quantities, that is, quantities which *describe* motion. We have already introduced our primitive four-vector $\boldsymbol{R} \equiv [ct, x, y, z]$. Let us proceed to the next four-vector, the *displacement four-vector*.

18.3.1 Displacement Four-Vector

The four-vector $[ct, x, y, z]$ transforms between inertial frames of reference according to the Lorentz transformation and so does the four-vector $[ct + c\Delta t, x + \Delta x, y + \Delta y, z + \Delta z]$. From the linearity of the Lorentz transformations, it follows that $[c\Delta t, \Delta x, \Delta y, \Delta z]$ also transforms like $[ct, x, y, z]$. We therefore define the *displacement four-vector* to be

$$\Delta\boldsymbol{R} = [c\Delta t, \Delta x, \Delta y, \Delta z]. \tag{18.28}$$

To express it another way, it must be a four-vector by virtue of being the difference of two four-vectors.

Evidently, the time interval between events depends upon the interial frame of reference in which the observer is located. Of special importance is that frame of reference in which the events take place at the same spatial location, that is, $\Delta x = \Delta y = \Delta z = 0$. The *proper time* Δt_0 is the time interval between events in this special frame of reference. It is straightforward to show that it is the *shortest time interval* measured in any frame of reference. Taking the norms of the displacement four-vector in any two frames of reference,

$$c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 = c^2\Delta t'^2 - \Delta x'^2 - \Delta y'^2 - \Delta z'^2.$$

In the frame in which $\Delta x' = \Delta y' = \Delta z' = 0$, $\Delta t' = \Delta t_0$ and so

$$c^2\Delta t_0^2 = c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2. \tag{18.29}$$

Since $\Delta x^2, \Delta y^2, \Delta z^2$ are necessarily positive, Δt_0 to must be the minimum time measured in any inertial frame of reference.

An important aspect of the proper time interval Δt_0 is that it is the only invariant time interval upon which all observers can agree for a pair of events separated by $c\Delta t, \Delta x, \Delta y, \Delta z$. Observers in different inertial frames make different measurements of $c\Delta t$, Δx, Δy, Δz and, although they all measure different Δts for the time interval between the same two events, they all agree about the value of Δt_0 when they measure $\Delta x, \Delta y$ and Δz as well and insert the values into (18.29). As indicated above, the proper time is the only invariant time interval on which they can all agree.

Let us work out the relation between the proper time Δt_0 and the time measured in some other arbitrary frame of reference. The four-vectors associated with the proper time interval Δt_0 in the proper frame of reference and the other frame are

$$[c\Delta t_0, 0, 0, 0], \quad [c\Delta t, \Delta x, \Delta y, \Delta z]. \tag{18.30}$$

Equating the norms of these four-vectors:

$$c^2\Delta t_0^2 = c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 = c^2\Delta t^2 - \Delta \boldsymbol{r}^2. \tag{18.31}$$

But $\Delta\boldsymbol{r}$ is just the distance the origin of the proper frame of reference has moved in the time Δt. In other words, the velocity of the origin of the proper frame $\boldsymbol{v}$ is $\Delta\boldsymbol{r}/\Delta t$ and hence

$$\Delta t_0^2 = \Delta t^2\left(1 - \frac{V^2}{c^2}\right), \tag{18.32}$$

that is,

$$\Delta t_0 = \frac{\Delta t}{\gamma}. \tag{18.33}$$

Since γ is always greater than or equal to 1, this calculation demonstrates that the proper time interval Δt_0 is the shortest time between events. We will use (18.33) repeatedly in what follows.

The relation (18.33) explains the observation of muons at sea-level. Cosmic rays are very high energy protons and nuclei which originate in violent explosions of stars and make their way to the Solar System through the interstellar medium. Muons are created in the upper layers of the atmosphere by collisions between very high energy cosmic rays, mostly protons, and the nuclei of atoms of the atmospheric gases. A typical interaction is

$$\mathrm{p} + \mathrm{p} \to n\pi^+ + n\pi^- + n\pi^0 + \mathrm{p} + \mathrm{p} + \cdots,$$

that is, a shower of positive, negative and neutral pions (π) is created in the collision in which roughly equal numbers of the three types of pions are produced. The positive and negative pions have very short lifetimes, $\tau = 2.551 \times 10^{-8}$ s, and decay into muons, μ:

$$\pi^+ \to \mu^+ + \nu_\mu \qquad \pi^- \to \mu^- + \overline{\nu_\mu}.$$

The ν_μ and $\overline{\nu_\mu}$ are muon neutrinos and antineutrinos respectively. The muons have short lifetimes $\tau = 2.2 \times 10^{-6}$ s before decaying into electrons, positrons, neutrinos, muon neutrinos and so on. Since they are created at the top of the atmosphere at a height of about 10 km, according to Galilean relativity, they should travel a typical distance of only

$c\tau$ before decaying, that is, in a distance of $3 \times 10^5 \times 2.2 \times 10^{-6}$ km = 660 m. Therefore, we would expect very few of them to reach the surface of the Earth. However, intense fluxes of relativistic muons are observed at the surface of the Earth. These high energy muons have Lorentz factors $\gamma > 20$. Because of the effects of time dilation, the observer on the Earth measures a decay half-life $\Delta t = \gamma \Delta t_0$, an effect which is often referred to confusingly as 'moving clocks go slow'. Therefore, according to the observer on the surface of the Earth, the muon has a half-life of $\gamma\tau$, in which time a relativistic muon with $\gamma = 20$ can travel a distance of 13 km and so easily reach the surface of the Earth.

Another famous demonstration of time dilation is the experiment first carried out by Joseph Hafele and Richard Keating, who flew round the world in opposite directions on commercial airliners. As the abstract of their paper states:

> Four cesium beam clocks flown around the world on commercial jet flights during October 1971, once eastward and once westward, recorded directionally dependent time differences which are in good agreement with predictions of conventional relativity theory. Relative to the atomic time scale of the U.S. Naval Observatory, the flying clocks lost 59 ± 10 nanoseconds during the eastward trip and gained 273 ± 7 nanoseconds during the westward trip, where the errors are the corresponding standard deviations.[25]

Note that the time differences include those due to gravitational time dilation associated with the change in gravitational potential between the aircraft and the surface of the Earth.

18.3.2 The Velocity Four-Vector

To find the velocity four-vector, we need a quantity which transforms like $\Delta\boldsymbol{R}$ and has the form of a velocity. The only time which is Lorentz invariant is the proper time Δt_0 – as we discussed above, an observer in any inertial frame of reference can measure $c\Delta t, \Delta x, \Delta y, \Delta z$ for two events and compute from these $c^2\Delta t_0^2 = c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2$. An observer in any other inertial frame of reference would find exactly the same value, $c^2\Delta t_0^2$. Therefore, let us define the *velocity four-vector* by

$$\boldsymbol{U} = \frac{\Delta\boldsymbol{R}}{\Delta t_0}. \tag{18.34}$$

But, from the analysis which led to (18.33), the proper time interval is related to the time interval measured in the laboratory frame by $\Delta t_0 = \Delta t/\gamma$. We can therefore write $\boldsymbol{U}$ as

$$\begin{aligned} \boldsymbol{U} = \frac{\Delta\boldsymbol{R}}{\Delta t_0} = \gamma \left[c\frac{\Delta t}{\Delta t}, \frac{\Delta x}{\Delta t}, \frac{\Delta y}{\Delta t}, \frac{\Delta z}{\Delta t} \right] &= \left[\gamma c, \gamma u_x, \gamma u_y, \gamma u_z \right] \\ &= \left[\gamma c, \gamma \boldsymbol{u} \right]. \end{aligned} \tag{18.35}$$

In this relation $\boldsymbol{u}$ is the three-velocity of the particle in the frame S and γ the corresponding Lorentz factor; u_x, u_y, u_z, are the components of $\boldsymbol{u}$. Notice the procedure we use to work out the components of the four-velocity. In some frame of reference, we measure the three-velocity $\boldsymbol{u}$ and hence γ. We then form the quantities γc and $\gamma u_x, \gamma u_y, \gamma u_z$ and know that they will transform exactly as ct and x, y, z.

Let us use this procedure to add two velocities relativistically. The relative velocity of the two frames of reference is V in standard configuration and $\boldsymbol{u}$ is the velocity of the particle in the frame S. What is its velocity in the frame S'? The Lorentz factor associated with the relative motion of S and S' is γ_V. First, we write down the velocity four-vectors for the particle in S and S':

$$\text{In } S,\ [\gamma c, \gamma \boldsymbol{u}] \equiv [\gamma c, \gamma u_x, \gamma u_y, \gamma u_z], \qquad \gamma = (1 - u^2/c^2)^{-1/2}.$$
$$\text{In } S',\ [\gamma' c, \gamma' \boldsymbol{u}'] \equiv [\gamma' c, \gamma' u'_x, \gamma' u'_y, \gamma' u'_z], \qquad \gamma' = (1 - u'^2/c^2)^{-1/2}.$$

We can relate the components of the four-vectors through the following identities:

$$\begin{aligned} ct' &\to \gamma' c, & ct &\to \gamma c, \\ x' &\to \gamma' u'_x, & x &\to \gamma u_x, \\ y' &\to \gamma' u'_y, & y &\to \gamma u_y, \\ z' &\to \gamma' u'_z, & z &\to \gamma u_z. \end{aligned}$$

Therefore, applying the Lorentz transformation (18.23), we find

$$\begin{cases} \gamma' c = \gamma_V \left(\gamma c - \dfrac{V\gamma u_x}{c}\right), \\ \gamma' u'_x = \gamma_V(\gamma u_x - V\gamma), \\ \gamma' u'_y = \gamma u_y, \\ \gamma' u'_z = \gamma u_z. \end{cases} \tag{18.36}$$

The first relation of (18.36) gives

$$\frac{\gamma\gamma_V}{\gamma'} = \frac{1}{1 - \dfrac{Vu_x}{c^2}},$$

and therefore, from the spatial terms of (18.36), we find

$$\begin{cases} u'_x = \dfrac{u_x - V}{\left(1 - \dfrac{Vu_x}{c^2}\right)}, \\ u'_y = \dfrac{u_y}{\gamma_V\left(1 - \dfrac{Vu_x}{c^2}\right)}, \\ u'_z = \dfrac{u_z}{\gamma_V\left(1 - \dfrac{Vu_x}{c^2}\right)}. \end{cases} \tag{18.37}$$

These are the standard expressions for the addition of velocities in special relativity. They have many pleasant features. For example, if u_x is equal to the speed of light c, the speed u'_x in S' is also c, as required by Einstein's second postulate.

Note that the norm of the four-velocity is

$$|\boldsymbol{U}^2| = c^2\gamma^2 - \gamma^2 u_x^2 - \gamma^2 u_y^2 - \gamma^2 u_z^2 = \gamma^2 c^2(1 - u^2/c^2) = c^2, \tag{18.38}$$

which is an invariant as expected.

18.3.3 The Acceleration Four-Vector

We can now repeat the procedure of Section 18.3.2 to find the acceleration four-vector. First, we form the increment of the four-velocity $\Delta \boldsymbol{U} \equiv [c\Delta\gamma, \Delta(\gamma \boldsymbol{u})]$, which is necessarily a four-vector. Then we define the only invariant acceleration-like quantity we can form by dividing by the proper time interval Δt_0:

$$\boldsymbol{A} = \frac{\Delta \boldsymbol{U}}{\Delta t_0} = \left[\gamma c \frac{\Delta\gamma}{\Delta t}, \gamma \frac{\Delta}{\Delta t}(\gamma \boldsymbol{u})\right] = \left[\gamma c \frac{\mathrm{d}\gamma}{\mathrm{d}t}, \gamma \frac{\mathrm{d}}{\mathrm{d}t}(\gamma \boldsymbol{u})\right]. \tag{18.39}$$

This is the *acceleration four-vector*. Let us convert it into a more useful form for practical applications. First, notice how to differentiate $\gamma = (1 - u^2/c^2)^{-1/2}$:

$$\begin{aligned}\frac{\mathrm{d}\gamma}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t}(1 - u^2/c^2)^{-1/2} = \frac{\mathrm{d}}{\mathrm{d}t}(1 - \boldsymbol{u}\cdot\boldsymbol{u}/c^2)^{-1/2} \\ &= \frac{\boldsymbol{u}\cdot\boldsymbol{a}}{c^2}(1 - u^2/c^2)^{-3/2} = \gamma^3\left(\frac{\boldsymbol{u}\cdot\boldsymbol{a}}{c^2}\right),\end{aligned} \tag{18.40}$$

where $\boldsymbol{a}$ is the three-acceleration. Then

$$\frac{\mathrm{d}}{\mathrm{d}t}(\gamma \boldsymbol{u}) = \gamma^3\left(\frac{\boldsymbol{u}\cdot\boldsymbol{a}}{c^2}\right)\boldsymbol{u} + \gamma \boldsymbol{a},$$

and so, from (18.39),

$$\boldsymbol{A} = \frac{\mathrm{d}\boldsymbol{U}}{\mathrm{d}t} = \left[\gamma^4\left(\frac{\boldsymbol{u}\cdot\boldsymbol{a}}{c}\right), \gamma^4\left(\frac{\boldsymbol{u}\cdot\boldsymbol{a}}{c^2}\right)\boldsymbol{u} + \gamma^2\boldsymbol{a}\right]. \tag{18.41}$$

Notice what this means. In some frame of reference, say S, we measure at a particular instant the three-velocity $\boldsymbol{u}$ and the three-acceleration $\boldsymbol{a}$ of a particle. From these quantities we form the scalar quantity $\gamma^4(\boldsymbol{u}\cdot\boldsymbol{a})/c$ and the three-vector, $\gamma^4(\boldsymbol{u}\cdot\boldsymbol{a}/c^2)\boldsymbol{u} + \gamma^2\boldsymbol{a}$. We then know that these quantities transform exactly as ct and $\boldsymbol{r}$ between inertial frames of reference.

It is often convenient in relativistic calculations to transform into the frame of reference in which the particle is instantaneously at rest. The particle is accelerated, but this does not matter since there is no dependence upon acceleration in the Lorentz transformations. In the *instantaneous rest frame* $\boldsymbol{u} = 0, \gamma = 1$ and hence the four-acceleration in this frame is

$$\boldsymbol{A} = [0, \boldsymbol{a}] \equiv [0, \boldsymbol{a}_0], \tag{18.42}$$

where $\boldsymbol{a}_0$ is the *proper acceleration* of the particle. In the same frame, the four-velocity of the particle is

$$\boldsymbol{U} = [\gamma c, \gamma \boldsymbol{u}] = [c, 0].$$

Therefore, $\boldsymbol{A}\cdot\boldsymbol{U} = 0$. Since four-vector relations are true in any frame of reference, this means that, no matter which frame we care to work in, the scalar product of the velocity and acceleration four-vectors is always zero, that is, the velocity and acceleration four-vectors are orthogonal. If you are suspicious, it is a useful exercise to take the scalar product of the expressions (18.35) and (18.41) for the four-vectors $\boldsymbol{A}$ and $\boldsymbol{U}$ and show that it is indeed zero.

18.4 Relativistic Dynamics: The Momentum and Force Four-Vectors

So far, we have been dealing with *kinematics*, that is, the description of motion, but now we have to tackle the problems of *dynamics*. This means introducing the concepts of momentum and force and how they are related.

18.4.1 The Momentum Four-Vector

First of all, let us go through a purely formal exercise of defining suitable momentum and force four-vectors and then finding out if they make physical sense. Let us introduce a four-vector quantity $\boldsymbol{P}$, which has the dimensions of momentum:

$$\boldsymbol{P} = m_0 \boldsymbol{U} = [\gamma m_0 c, \gamma m_0 \boldsymbol{u}], \tag{18.43}$$

where $\boldsymbol{U}$ is the velocity four-vector and m_0 is the mass of the particle, which is taken to be a scalar invariant – it will be identified with the *rest mass* of the particle. We note the following consequences:

(i) $m_0 \boldsymbol{U}$ is certainly a four-vector since m_0 is an invariant;
(ii) the space components of $\boldsymbol{P}$ reduce to the Newtonian formula for momentum if $u \ll c$, that is, $m_0 \gamma \boldsymbol{u} \to m_0 \boldsymbol{u}$ as $u \to 0$.

Therefore, *if this agrees with experiment*, we have found a suitable form for the momentum four-vector.

We can also define a *relativistic three-momentum* $\boldsymbol{p} = \gamma m_0 \boldsymbol{u}$ from the spatial components of the four-vector. The quantity $m = \gamma m_0$ is defined to be the *relativistic inertial mass*. As will be discussed below, I try to avoid using the relativistic mass, preferring to work with the relativistic momentum. We need to show that these definitions result in a self-consistent set of dynamics.

Let us find a relation between the relativistic momentum and the relativistic mass of the particle. Equating the norms of the momentum four-vectors in the laboratory frame $\boldsymbol{P} \equiv [\gamma m_0 c, \gamma m_0 \boldsymbol{u}]$ and in the rest frame of the particle, $\boldsymbol{P} \equiv [m_0 c, 0]$,

$$\begin{aligned} m_0^2 c^2 &= \gamma^2 m_0^2 c^2 - \gamma^2 m_0^2 u^2 \\ &= m^2 c^2 - p^2, \end{aligned}$$

or

$$p^2 = m^2 c^2 - m_0^2 c^2. \tag{18.44}$$

We will rewrite this expression in terms of energies in Section 18.5.

18.4.2 The Force Four-Vector

Following the logic of the previous sections, the natural four-vector generalisation of Newton's second law of motion is

$$\boldsymbol{F} = \frac{\mathrm{d}\boldsymbol{P}}{\mathrm{d}t_0}, \tag{18.45}$$

where $\boldsymbol{F}$ is the *four-force* and t_0 proper time. We now relate the force which we measure in the laboratory to the four-force. The best approach is through the quantity we have called the relativistic three-momentum. Why did we adopt this definition? Consider the collision between two particles which initially have momentum four-vectors $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$. After the collision, these become $\boldsymbol{P}_1'$ and $\boldsymbol{P}_2'$. The conservation of four-momentum can then be written

$$\boldsymbol{P}_1 + \boldsymbol{P}_2 = \boldsymbol{P}_1' + \boldsymbol{P}_2'. \tag{18.46}$$

In terms of the components of the four-vectors, this equation implies

$$\begin{cases} m_1 + m_2 &= m_1' + m_2', \\ \boldsymbol{p}_1 + \boldsymbol{p}_2 &= \boldsymbol{p}_1' + \boldsymbol{p}_2'. \end{cases} \tag{18.47}$$

where the ms are relativistic masses. Thus, implicit in this formalism is the requirement that the relativistic three-momentum is conserved and so, for relativistic particles, $\boldsymbol{p} = \gamma m_0 \boldsymbol{u}$ plays the role of momentum. The corresponding force equation is suggested by Newton's second law,

$$\boldsymbol{f} = \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}(\gamma m_0 \boldsymbol{u}), \tag{18.48}$$

where $\boldsymbol{f}$ is the normal three-force of Newtonian dynamics.

Are these definitions self-consistent? We have to be a bit careful. On the one hand, we could just argue, 'Let's look at experiment and see if it works', and in many ways this is more or less all that the argument amounts to. We can do a little better. In point collisions of particles, the relativistic generalisation of Newton's third law should apply, that is, $\boldsymbol{f} = -\boldsymbol{f}$. We can only consider point collisions or else we would get into trouble with action at a distance in relativity – recall the relativity of simultaneity in different frames of reference. For a point collision, $\boldsymbol{f} = -\boldsymbol{f}$ is true if we adopt the definition of $\boldsymbol{f}$ given above because we have already argued that relativistic three-momentum is conserved, that is,

$$\Delta\boldsymbol{p}_1 = -\Delta\boldsymbol{p}_2, \quad \frac{\Delta\boldsymbol{p}_1}{\Delta t} = -\frac{\Delta\boldsymbol{p}_2}{\Delta t}, \quad \boldsymbol{f}_1 = -\boldsymbol{f}_2.$$

However, we cannot be absolutely sure that we have made the correct choice without appealing to experiment. We faced the same type of logical problem when we tried to understand the meaning of Newton's laws of motion (Section 8.1). They ended up being a set of definitions which give results consistent with experiment. Similarly, relativistic dynamics cannot come out of pure thought, but can be put into a logically self-consistent mathematical structure which is consistent with experiment.

We therefore adopt the definition of $\boldsymbol{f}$ as the three-force, in the same sense as in Newtonian dynamics, but now the particle may be moving relativistically and the relativistic three-momentum should be used for $\boldsymbol{p}$. Within this framework, we can derive a number of pleasant results.

18.4.3 $\boldsymbol{F} = m_0\boldsymbol{A}$

This follows directly from the definition of $\boldsymbol{F}$,

$$\boldsymbol{F} = \frac{\mathrm{d}\boldsymbol{P}}{\mathrm{d}t_0} = m_0\frac{\mathrm{d}\boldsymbol{U}}{\mathrm{d}t_0} = m_0\boldsymbol{A}. \tag{18.49}$$

In addition, because $\boldsymbol{F}\cdot\boldsymbol{U} = 0$,

$$\boldsymbol{F}\cdot\boldsymbol{U} = 0.$$

that is, the force and velocity four-vectors are orthogonal.

18.4.4 The Relativistic Generalisation of $\boldsymbol{f} = \mathrm{d}\boldsymbol{p}/\mathrm{d}t$

Let us write out the four-vector form of Newton's second law in terms of its components:

$$\boldsymbol{F} = [f_0, f_1, f_2, f_3] = \frac{\mathrm{d}\boldsymbol{P}}{\mathrm{d}t_0} = \left[\gamma\frac{\mathrm{d}(\gamma m_0 c)}{\mathrm{d}t}, \gamma\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t}\right], \tag{18.50}$$

where $\boldsymbol{p} = \gamma m_0\boldsymbol{u}$. Since we have argued that the relativistic form of Newton's second law in three-vector form is $\boldsymbol{f} = \mathrm{d}\boldsymbol{p}/\mathrm{d}t$, it follows that we must write the four-force in the form

$$\boldsymbol{F} = [f_0, \gamma\boldsymbol{f}] = \left[\gamma\frac{\mathrm{d}(\gamma m_0 c)}{\mathrm{d}t}, \gamma\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t}\right] = m_0\boldsymbol{A}. \tag{18.51}$$

Equating the spatial components of $[f_0, \gamma\boldsymbol{f}] = m_0\boldsymbol{A}$ and using (18.41), we find

$$\boldsymbol{f} = m_0\gamma^3\left(\frac{\boldsymbol{u}\cdot\boldsymbol{a}}{c^2}\right)\boldsymbol{u} + m_0\gamma\boldsymbol{a}. \tag{18.52}$$

This is the relativistic generalisation of the Newtonian expression $\boldsymbol{f} = m_0\boldsymbol{a}$.

Now let us analyse the 'time' component of the force four-vector, that is, of

$$\boldsymbol{F} = \frac{\mathrm{d}\boldsymbol{P}}{\mathrm{d}t_0} = \left[\gamma\frac{\mathrm{d}(\gamma m_0 c)}{\mathrm{d}t}, \gamma\frac{\mathrm{d}}{\mathrm{d}t}(\gamma m_0\boldsymbol{u})\right].$$

From (18.40), we can write

$$\gamma\frac{\mathrm{d}(\gamma m_0 c)}{\mathrm{d}t} = m_0\gamma^4\left(\frac{\boldsymbol{u}\cdot\boldsymbol{a}}{c}\right),$$

or, in terms of the relativistic mass m,

$$\frac{\mathrm{d}m}{\mathrm{d}t} = \gamma^3 m_0\left(\frac{\boldsymbol{u}\cdot\boldsymbol{a}}{c^2}\right). \tag{18.53}$$

Now inspect the quantity $(\boldsymbol{f}\cdot\boldsymbol{u})/c^2$. Substituting for $\boldsymbol{f}$ from (18.52),

$$\begin{aligned}\frac{\boldsymbol{f}\cdot\boldsymbol{u}}{c^2} &= m_0\gamma^3\left(\frac{\boldsymbol{u}\cdot\boldsymbol{a}}{c^2}\right)\frac{u^2}{c^2} + m_0\gamma\left(\frac{\boldsymbol{u}\cdot\boldsymbol{a}}{c^2}\right)\\ &= m_0\gamma^3\left(\frac{\boldsymbol{u}\cdot\boldsymbol{a}}{c^2}\right)\left(\frac{u^2}{c^2} + \frac{1}{\gamma^2}\right)\\ &= m_0\gamma^3\left(\frac{\boldsymbol{u}\cdot\boldsymbol{a}}{c^2}\right).\end{aligned} \tag{18.54}$$

Therefore, (18.53) becomes

$$\frac{\mathrm{d}m}{\mathrm{d}t} = \frac{\boldsymbol{f} \cdot \boldsymbol{u}}{c^2}, \tag{18.55}$$

or

$$\frac{\mathrm{d}(mc^2)}{\mathrm{d}t} = \frac{\mathrm{d}(\gamma m_0 c^2)}{\mathrm{d}t} = \boldsymbol{f} \cdot \boldsymbol{u}. \tag{18.56}$$

The quantity $\boldsymbol{f} \cdot \boldsymbol{u}$ is just the rate at which work is done on the particle, that is, its rate of increase of energy. Thus mc^2 is identified with the total energy of the particle. This is the formal proof of perhaps the most famous equation in physics,

$$E = mc^2. \tag{18.57}$$

Notice the profound implication of equation (18.57). There is a certain amount of inertial mass associated with the energy produced when work is done. It does not matter what the form of the energy is – electrostatic, magnetic, kinetic, elastic, etc. All energies are the same thing as inertial mass. Likewise, reading the equation backwards, inertial mass is energy. Nuclear power stations and nuclear explosions are vivid demonstrations of the identity of inertial mass and energy.

These key realisations were published in the same year 1905 as the major paper *On the Electrodynamics of Moving Bodies*. This second much shorter paper of 1905 is entitled *Does the Inertia of a Body Depend on Its Energy Content?*[26] and answers the question with a resounding 'Yes'. Intriguingly, direct experimental evidence for the relation $E = mc^2$ was only discovered in the famous experiment by John Cockcroft and Ernest Walton of 1932. With their intense beam of protons colliding with a ^{7}Li target, they could determine the energy balance for the interaction

$$\mathrm{p} + {}^7\mathrm{Li} \rightarrow {}^4\mathrm{He} + {}^4\mathrm{He},$$

including the kinetic energies of the incoming proton and the outgoing helium nuclei. Specifically, the mass decrement in the above interaction results in an energy release of 14.3 ± 2.7 Mev, while, from the observed ranges of the α-particles, the total liberated energy was 17.2 MeV.[27]

18.4.5 A Mild Polemic

Sometimes references are made to *longitudinal* and *transverse* masses in textbooks. These arise from an unhappy use of the relativistic form of Newton's second law of motion. If $\boldsymbol{u} \parallel \boldsymbol{a}$, the force law (18.52) reduces to

$$\begin{aligned} \boldsymbol{f} &= m_0\gamma^3 \left(\frac{\boldsymbol{u} \cdot \boldsymbol{a}}{c^2}\right) \boldsymbol{u} + m_0\gamma \boldsymbol{a} = m_0\gamma^3 \left(\frac{u^2}{c^2} + \frac{1}{\gamma^2}\right) \boldsymbol{a} \\ &= \gamma^3 m_0 \boldsymbol{a}. \end{aligned}$$

The quantity $\gamma^3 m_0$ is called the *longitudinal* mass. On the other hand, if $\boldsymbol{u} \perp \boldsymbol{a}$, we find that

$$\boldsymbol{f} = \gamma m_0 \boldsymbol{a},$$

and γm_0 is called the *transverse* mass. I dislike this introduction of different sorts of masses, which are meant to be 'relativistic' forms of Newton's second law of motion in the form $\boldsymbol{f} = \text{‘}m\text{’}\boldsymbol{a}$. It is much preferable to stick with the correct generalisation of Newton's second law in the form $\boldsymbol{f} = \mathrm{d}\boldsymbol{p}/\mathrm{d}t$, where $\boldsymbol{p} = \gamma m_0 \boldsymbol{u}$ and abolish all masses except the rest mass m_0 for the reasons discussed in the next section.

18.5 The Relativistic Equations of Motion

Let us gather together the results we have obtained so far. Consider a particle of rest mass m_0 travelling at velocity $\boldsymbol{u}$ in some inertial frame of reference S. The particle has Lorentz factor γ given by the standard expression $\gamma = \left(1 - u^2/c^2\right)^{-1/2}$. Then:

- The *relativistic three-momentum* of the particle is $\gamma m_0 \boldsymbol{u}$.
- Its *total energy* is $E = mc^2 = \gamma m_0 c^2$.
- Its *rest mass energy* is $E_0 = m_0 c^2$.
- Its *kinetic energy* is the difference between its total and rest mass energies,

$$E_{\text{kin}} = E - E_0 = \gamma mc^2 - m_0 c^2 = (\gamma - 1) m_0 c^2. \tag{18.58}$$

The last definition follows from the calculation which led to (18.56) in which we showed that, when we do work on a particle, we increase the quantity $\gamma m_0 c^2$. For $u \ll c$, (18.58) reduces to the Newtonian non-relativistic expression for the kinetic energy:

$$E_{\text{kin}} = (\gamma - 1) m_0 c^2 = \left[\left(1 - \frac{u^2}{c^2}\right)^{-1/2} - 1\right] m_0 c^2 \approx \tfrac{1}{2} m_0 u^2. \tag{18.59}$$

- The relativistic form of Newton's second law of motion is

$$\boldsymbol{f} = \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t}. \tag{18.60}$$

- The conservation laws of energy and momentum are subsumed into the single law of conservation of four-momentum:

$$\boldsymbol{P}_1 + \boldsymbol{P}_2 = \boldsymbol{P}_1' + \boldsymbol{P}_2', \tag{18.61}$$

where the components of the four vectors are given by

$$\boldsymbol{P} \equiv [p_0,\ p_x,\ p_y,\ p_z] = [\gamma m_0 c, \gamma m_0 u_x,\ \gamma m_0 u_y,\ \gamma m_0 u_z]. \tag{18.62}$$

- Equating the spatial components of the equation of conservation of four-momentum, we find the law of *conservation of relativistic three-momentum*:

$$\boldsymbol{p}_1 + \boldsymbol{p}_2 = \boldsymbol{p}_1' + \boldsymbol{p}_2'. \tag{18.63}$$

- From the 'time' component of the equation of conservation of four-momentum, we find the *relativistic law of conservation of energy*:

$$\begin{cases} p_0(1) + p_0(2) &= p_0'(1) + p_0'(2), \\ E_1 + E_2 &= E_1' + E_2', \\ \gamma_1 m_0(1) + \gamma_2 m_0(2) &= \gamma_1' m_0(1) + \gamma_2' m_0(2), \end{cases} \tag{18.64}$$

where the Es are the total energies of the particles.

18.5.1 Massless Particles

There are advantages in rewriting the relations we have just derived in terms of the various energies defined above. Let us rewrite the list entirely in terms of energies:

- *Total energy* $E = \gamma m_0 c^2$.
- *Rest mass energy* $E_0 = m_0 c^2$.
- *Relativistic momentum* $\boldsymbol{p} = \gamma m_0 \boldsymbol{u} = (E/c^2)\boldsymbol{u}$.
- *Kinetic energy* $E_{\text{kin}} = E - E_0$.

Thus, we can write the four-momentum of the particle as follows:

$$\boldsymbol{P} = m_0 \boldsymbol{U} = m_0[\gamma c,\ \gamma u_x,\ \gamma u_y,\ \gamma u_z] = \left[\frac{E}{c},\ \frac{E}{c^2}\boldsymbol{u}\right]. \tag{18.65}$$

In Case Study V, we demonstrated in some detail that light consists of massless particles, *photons*, which travel at the speed of light. Their energies are related to their frequencies ν by $E = h\nu$, where h is Planck's constant. For these particles, the rest mass energy $E_0 = 0$ and $\boldsymbol{u} = \boldsymbol{c}$. Thus, for massless particles, such as the photon, we find directly

$$\boldsymbol{p} = \frac{E}{c^2}\boldsymbol{c} \quad \text{and} \quad E_{\text{kin}} = E, \tag{18.66}$$

and the momentum four-vector for photons is

$$\boldsymbol{P} = \left[\frac{E}{c}, \frac{E}{c^2}\boldsymbol{c}\right]. \tag{18.67}$$

Notice that we have derived directly an expression for the momentum of a photon,

$$\boldsymbol{p} = \frac{E}{c^2}\boldsymbol{c} = \frac{h\nu}{c^2}\boldsymbol{c}. \tag{18.68}$$

The advantage of writing the four-momenta in terms of energies is that we can treat particles and photons on exactly the same basis. Notice also that all masses have been eliminated from the expressions.

To complete this rewrite of the expressions, let us perform the same change of variable for the expression relating the masses and momenta of the particles in their rest and moving frames (18.44). Equating the norms of the momentum four-vector (18.65) in the laboratory and rest frames of the particle,

$$\left(\frac{E}{c}\right)^2 - \left(\frac{E}{c^2}\boldsymbol{u}\right)^2 = \left(\frac{E_0}{c}\right)^2,$$

$$E^2 - \boldsymbol{p}^2 c^2 = E_0^2 = \text{constant}. \tag{18.69}$$

The quantity $E^2 - \boldsymbol{p}^2 c^2 = E_0^2$ is thus an invariant in all inertial frames of reference. This is a great help in solving problems in relativity. We can work out the total energy $E = \sum_i E_i$ and total momentum $\boldsymbol{p} = \sum_i \boldsymbol{p}_i$ in some frame of reference and then we know that the quantity $E^2 - \boldsymbol{p}^2 c^2 = E_0^2$ is an invariant in all inertial frames of reference. This is particularly useful in working out threshold energies for the creation of particles in high energy collisions. One final note is that, in the extreme high energy limit $E \gg E_0$, (18.69) becomes

$$E = |\boldsymbol{p}|c. \tag{18.70}$$

Thus, in the ultrarelativistic limit, the particles behave like photons.

18.5.2 Relativistic Particle Dynamics in a Magnetic Field

Consider next the dynamics of a relativistic particle in a magnetic field. The expression for the *Lorentz force* is

$$\boldsymbol{f} = e(\boldsymbol{E} + \boldsymbol{u} \times \boldsymbol{B}),$$

and hence, if $\boldsymbol{E} = 0$,

$$\boldsymbol{f} = e(\boldsymbol{u} \times \boldsymbol{B}).$$

Therefore, from (18.52),

$$e(\boldsymbol{u} \times \boldsymbol{B}) = m_0\gamma^3 \left(\frac{\boldsymbol{u} \cdot \boldsymbol{a}}{c^2}\right) \boldsymbol{u} + \gamma m_0 \boldsymbol{a}. \tag{18.71}$$

The left-hand side of this expression and the first term on the right are perpendicular vectors because $\boldsymbol{u} \perp (\boldsymbol{u} \times \boldsymbol{B})$ and hence we require simultaneously that

$$e(\boldsymbol{u} \times \boldsymbol{B}) = \gamma m_0 \boldsymbol{a} \quad \text{and} \quad \boldsymbol{u} \cdot \boldsymbol{a} = 0, \tag{18.72}$$

that is, the acceleration impressed by the magnetic field is perpendicular to both $\boldsymbol{B}$ and the velocity $\boldsymbol{u}$. This is the origin of the circular or spiral motion of charged particles in magnetic fields.

18.6 The Frequency Four-Vector

Finally, let us derive the *frequency four-vector* from the scalar product rule: if $\boldsymbol{A}$ is a four-vector and $\boldsymbol{A} \cdot \boldsymbol{B}$ is an invariant, $\boldsymbol{B}$ must be a four-vector. The simplest approach is to consider the phase of a wave. If we write the expression for the wave in the form $\exp[\mathrm{i}(\boldsymbol{k} \cdot \boldsymbol{r} - \omega t)]$, the quantity $(\boldsymbol{k} \cdot \boldsymbol{r} - \omega t)$ is the phase of the wave, in other words, how far through the cycle of the wave from 0 to 360° one happens to be at coordinates $[ct, \boldsymbol{r}]$. This is an invariant scalar quantity, whatever frame of reference one happens to be in. Thus, in any inertial frame of reference,

$$(\boldsymbol{k} \cdot \boldsymbol{r} - \omega t) = \text{invariant}.$$

But $[ct, \boldsymbol{r}]$ is a four-vector and therefore the quantity

$$\boldsymbol{K} = [\omega/c, \boldsymbol{k}] \tag{18.73}$$

must also be a four-vector, which is called the *frequency four-vector*.

We can also derive this four-vector by considering light to consist of photons and deriving the appropriate momentum four-vector. Thus, from (18.67),

$$\boldsymbol{P} \equiv \left[\frac{E}{c}, \frac{E}{c^2}\boldsymbol{u}\right] \equiv \left[\frac{h\nu}{c}, \frac{h\nu}{c}\mathbf{i}_k\right],$$

where $h\nu/c$ is the momentum of the photon and $h\nu$ its energy; $\boldsymbol{i}_k$ is the unit vector in the direction of propagation of the photon. Writing

$$\frac{h\nu}{c^2} = \frac{\hbar\omega}{c^2}, \qquad \frac{h\nu}{c} = \frac{\hbar\omega}{c} = \hbar|\boldsymbol{k}|,$$

the momentum four-vector of the photon becomes

$$\boldsymbol{P} \equiv \hbar\left[\frac{\omega}{c}, \boldsymbol{k}\right], \tag{18.74}$$

that is,

$$\boldsymbol{P} = \hbar\boldsymbol{K}.$$

18.7 Lorentz Contraction and the Origin of Magnetic Fields

We have stated that special relativistic phenomena are outside our normal experience, but there is one remarkable example in which second-order effects in V/c are omnipresent in everyday life – the origin of the magnetic fields associated with the flow of electric currents. We noted in Section 18.1.1 the interchangeability of electric and magnetic fields, depending upon the inertial frame of reference in which the measurements are made. We need to repeat the analyses of Section 18.1.1 to work out the transformation laws for the electric field strength $\boldsymbol{E}$ and the magnetic flux density $\boldsymbol{B}$ according to special relativity. This involves replacing (18.3) by the Lorentz transformations for partial derivatives, which are

$$\begin{cases} \dfrac{\partial}{\partial t} \to \gamma\left(\dfrac{\partial}{\partial t'} - V\dfrac{\partial}{\partial x'}\right), \\ \dfrac{\partial}{\partial x} \to \gamma\left(\dfrac{\partial}{\partial x'} - \dfrac{V}{c^2}\dfrac{\partial}{\partial t'}\right), \\ \dfrac{\partial}{\partial y} \to \dfrac{\partial}{\partial y'}, \\ \dfrac{\partial}{\partial z} \to \dfrac{\partial}{\partial z'}. \end{cases} \tag{18.75}$$

It is left to the reader to show that Maxwell's equations are form invariant under these transformations and that the transformed components of $\boldsymbol{E}$ and $\boldsymbol{B}$ are

$$\begin{cases} E'_x = E_x, \\ E'_y = \gamma(E_y - VB_z), \\ E'_z = \gamma(E_z - VB_y). \end{cases} \tag{18.76}$$

$$\begin{cases} B'_x = B_x, \\ B'_y = \gamma\left(B_y - \dfrac{V}{c^2}E_z\right), \\ B'_z = \gamma\left(B_z - \dfrac{V}{c^2}E_y\right). \end{cases} \tag{18.77}$$

If, for example, there is a uniform magnetic field in some frame of reference, it can be transformed to zero in some inertial frame in which the effects of the magnetic field would be attributed to the action of the electric field $\boldsymbol{E}'$. Let us consider the case of the magnetic field of a current-carrying wire.

Suppose the current I flows in the wire as observed in the laboratory frame of reference S, the electrons drifting at speed v while the ions remain stationary. The current is $I = en_\mathrm{e}v$, where n_e is the number of electrons per unit length and is equal to n_i, the number of ions per unit length, so that charge neutrality is preserved in S; e is the charge of the electron. Applying Ampère's law (5.10), the magnetic flux density at radial distance r from the wire is

$$\oint \boldsymbol{B}\cdot \mathrm{d}\boldsymbol{s} = \mu_0 I, \qquad B = \frac{\mu_0 e n_\mathrm{e} v}{2\pi r}. \tag{18.78}$$

If a charge q moves at velocity $\boldsymbol{u}$ parallel to the wire in the positive x-direction, as shown in Fig. 18.3(a), the Lorentz force is

$$\boldsymbol{f} = q(\boldsymbol{u}\times\boldsymbol{B}) = q\frac{\mu_0 e n_\mathrm{e} u v}{2\pi r} \tag{18.79}$$

towards the wire, if $q > 0$. Notice that, in the arrangement shown in Fig. 18.3(a), this is exactly the same as the rule that parallel currents attract each other.

We now repeat the calculation in the frame of reference of the moving charge q (Fig. 18.3(b)). The ions now move in the negative x'-direction at speed u and the electrons have speed v', which is the relativistic sum of v and u also in the negative x-direction. The corresponding Lorentz factors are

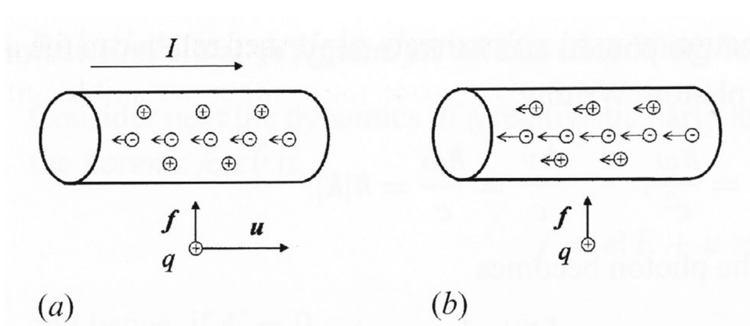


Fig. 18.3 Illustrating the origin of the force on a moving charge in the vicinity of a current-carrying wire. There is no charge on the wire in S. Transforming to the frame moving with the charge q, there is a net negative charge on the wire.

$$\gamma'_{\mathrm{i}} = \frac{1}{\left(1 + \dfrac{u^2}{c^2}\right)^{1/2}}, \qquad \gamma'_{\mathrm{e}} = \frac{1 + \dfrac{uv}{c^2}}{\left(1 - \dfrac{u^2}{c^2}\right)^{1/2}\left(1 - \dfrac{v^2}{c^2}\right)^{1/2}},$$

where the first relation of (18.36) has been used to find the Lorentz factor for the electrons – notice that we add together the speeds relativistically although the speeds of the ions and electrons are very small, $u \ll c$, $v \ll c$. The charge densities per unit length in S' are given by the length-contracted values for the ions and electrons. For the ions, which are stationary in S, the number per unit length increases to

$$n'_{\mathrm{i}} = n_{\mathrm{i}}\gamma'_{\mathrm{i}}.$$

For the electrons, the number per unit length is already increased relative to the number density in their own rest frame by the factor γ_{e}. Therefore, the number per unit length when moving at speed v' relative to its value in S is

$$n'_{\mathrm{e}} = n_{\mathrm{e}}\frac{\gamma'_{\mathrm{e}}}{\gamma_{\mathrm{e}}}.$$

Therefore, the net charge on the wire per unit length in S' is

$$en' = e(n'_{\mathrm{i}} - n'_{\mathrm{e}}). \tag{18.80}$$

Preserving quantities to second order in v/c, we find

$$en' = -\frac{n_{\mathrm{e}}euv}{c^2}. \tag{18.81}$$

There is therefore a net negative charge on the wire which results in an *electrostatic attractive force* on the stationary charge q in S'. The force acting on the charge q due to a line charge ρ_{e} per unit length is found from Gauss's law for electrostatics, $E_r = \rho_{\mathrm{e}}/2\pi\epsilon_0 r$, and so the force acting on the charge q in S' is

$$f_{\mathrm{E}} = q\frac{en_{\mathrm{e}}uv}{2\pi\epsilon_0 c^2 r}\gamma = q\frac{\mu_0 en_{\mathrm{e}}uv}{2\pi r}, \tag{18.82}$$

which is exactly the same as (18.79). Thus, we can understand the origin of magnetic fields as being due to the *second-order effects* of length contraction upon the charges flowing in the wire.

It is remarkable just how small the effect is. Using the figures adopted by French,[28] suppose a current of 10 A flows in a copper wire of cross-section $\sigma = 1$ mm^2. In solid copper, there are $n_{\mathrm{e}} \sim 10^{20}$ conduction electrons mm^{-3}. Therefore, the current is $I = n_{\mathrm{e}}ev\sigma$, from which we find $v \sim 0.6$ mm s^{-1}. This corresponds to $v/c \sim 2 \times 10^{-12}$ – the electrons are scarcely moving at all!

The reason the effect is so tiny is that, without the almost perfect cancellation of the charges of the ions, the electrostatic forces would be enormous. To see this, we can compare the electrostatic force due to the ions alone, which would be $f_{\mathrm{i}} = q(n_{\mathrm{i}}e/2\pi\epsilon_0 r)$, with the expression (18.81):

$$\frac{f_{\mathrm{E}}}{f_{\mathrm{i}}} = \frac{q\mu_0 en_{\mathrm{e}}uv/2\pi r}{qn_{\mathrm{i}}e/2\pi\epsilon_0 r} = \mu_0\epsilon_0 uv = \frac{uv}{c^2}. \tag{18.83}$$

Notice that $\mu_0\epsilon_0$ is just the ratio of the strengths of the magnetostatic and electrostatic forces for unit poles and charges (see Section 5.1).

18.8 Reflections

This case study is a beautiful example of how an appropriate mathematical structure can clarify the underlying symmetries of the theory. I could have gone through the whole exercise as a piece of mathematics and it would hardly have been apparent when I had to look at the real world. In fact, there are assumptions present as we have shown, just as there are in Newtonian theory. It so happens that in this case the mathematical structure is very elegant indeed – it might not have turned out that way.

From the practical point of view, the use of four-vectors greatly simplifies all calculations in relativity. It is only necessary to remember a single set of transformations and then understand how to derive the relevant four-vectors. There are great advantages in having a prescription such as this which can be trusted in carrying out relativistic calculations. I am very suspicious of arguments which try to provide intuitive ways of looking at relativistic problems. Phenomena such as length contraction and time dilation should be treated cautiously and the simplest way of ensuring that confusion does not reign is to write down the four-vectors associated with events and then use the Lorentz transforms to relate coordinates.

Notes

1 Stachel, J. (1995). *The History of Relativity*, in *Twentieth Century Physics*, Vol. 1, eds. L.M. Brown, A. Pais and A.B. Pippard, pp. 249–356. Bristol: Institute of Physic Publishing and New York: American Institute of Physics Press. Stachel, J. (1998). *Einstein's Miraculous Year: Five Papers that Changed the Face of Physics*, pp. 123–160. Princeton: Princeton University Press.

2 Pais, A. (1982). *Subtle Is The Lord . . . – The Science and Life of Albert Einstein*. Oxford: Clarendon Press.

3 Maxwell, J.C. (1880). On a Possible Mode of Detecting a Motion of the Solar System through the Luminiferous Ether, *Proceedings of the Royal Society*, **30**, 108–110. See also, *Nature*, **21**, 314–315 (1880).

4 Michelson, A.A. (1903). *Light Waves and their Uses*, pp. 156–157. Chicago: Chicago University Press.

5 Maxwell, J.C. (1878). Aether, *Encyclopaedia Britannica*, 9th edition, **8**, 568–572.

6 Michelson, A.A and Morley, E.W. (1887). On the Relative Motion of the Earth and of the Luminiferous Ether, *Sidereal Messenger*, **6**, 306–310.

7 Michelson, A.A. (1927). *Studies in Optics*, p. 151. Chicago: University of Chicago Press (first Phoenix edition 1962).

8 Whittaker, E.T. *History of the Theories of the Aether and Electricity*, Vol. 1 (1910). London: Longmans, Green and Co (revised and enlarged, New York: Nelson and Sons (1951)). Vol. 2

(1953). New York: Nelson and Sons. This history is regarded as the standard work on the subject, with the one caveat that for some reason Whittaker gives the credit for special relativity to Lorentz and Poincaré, rather than Einstein. As will be apparent from this case study, this scarcely matches the facts as must have been known at the time. Holton's critique of Whittaker's opinions is contained in his essay *On the origins of the special theory of relativity* in his book *Thematic Origins of Scientific Thought: Kepler to Einstein*, pp. 191–236. Cambridge, Massachusetts: Harvard University Press.

9 Voigt, W. (1887). *Goett. Nachr*. p. 41.

10 G.F. Fitzgerald, (1889). *Science*, **13**, 390.

11 Lorentz, H.A. (1892). *Versl. K. Ak. Amsterdam*, **1**, 74 (see note 3, pp. 123–4).

12 Lorentz, H.A. (1895). *Versuch einer Theorie der Elektischen und Optischen Erschweinungen in bewgten Körpen*. Leiden: Brill (see note 3, pp. 124–125).

13 Lorentz, H.A. (1899). *Versl. K. Ak. Amsterdam*, **10**, 793 (see note 3, pp. 125–6).

14 Larmor, J. (1900). *Aether and Matter*. Cambridge: Cambridge University Press.

15 Lorentz, H.A. (1904). *Proc. K. Ak. Amsterdam*, **6**, 809 (see note 3, pp. 126).

16 Holten, G. (1988). *Thematic Origins of Scientific Thought: Kepler to Einstein*, p. 195. Cambridge, Massachusetts: Harvard University Press.

17 Poincaré, H. (1898). *Revue de Métaphysique et de Morale*, **6**, 1.

18 Poincaré, H. (1904) L'Etat actuel et l'avenir de la physique mathématique, *Bulletin des Sciences Mathématiques*.

19 See note 1, Stachel (1998), p. 112.

20 Einstein, A. (1905). *Annalen der Physik*, **17**, 891. (See translation in note 1, Stachel (1998), pp. 123–160.)

21 See note 13, p. 197.

22 Rindler, W. (2001). *Relativity: Special, General and Cosmological*. Oxford: Oxford University Press.

23 See note 1, Stachel (1998), p. 124.

24 Hawking, S.W. (1988). *A Brief History of Time*. London: Bantam Press.

25 Hafele, J.C. and Keating, R.E. (1972). Around-the-World Atomic Clocks: Observed Relativistic Time Gains, *Science*, **177**, 168–170.

26 Einstein, A. (1905). *Annalen der Physik*, **17**, 891–892. For translation, see note 1, Stachel (1998), pp. 161–164.

27 Cockcroft, J.D. and Walton, E.T.S. (1932). Experiments with High Velocity Positive Ions, II. The Disintegration of Elements by High Velocity Protons, *Proceedings of the Royal Society of London*, **A137**, 229–242. See also, my book, Longair, M.S. (2016). *Maxwell's Enduring Legacy: A Scientific History of the Cavendish Laboratory*, Section 9.6, 213–218.

28 French, A.P. (1968). *Special Relativity*. London: Chapman and Hall.

19 An Introduction to General Relativity

19.1 Introduction

Einstein's paper on the special theory of relativity of 1905 appears effortless, and, as discussed in Chapter 18, he did not regard the formulation of the theory as a particularly 'revolutionary act'. The route to the discovery of the general theory of relativity was very different. Whereas others had come close to elucidating the features of special relativity, Einstein was on his own and went far beyond all his contemporaries in his discovery of the general theory. How he arrived at the theory is one of the great stories of theoretical physics and involved the very deepest physical insight, imagination, intuition and sheer doggedness. It led to concepts which could barely have been conceived of in 1915 – the phenomenon of black holes and the possibility of testing theories of gravity in the strong field limit through the observation of relativistic stars. It also provided for the first time a relativistic theory of gravity which could be used to construct fully self-consistent models of the Universe as a whole (Chapter 20).

The history of the discovery of general relativity is admirably told by Abraham Pais in his scientific biography of Einstein, *'Subtle is the Lord . . .' The Science and Life of Albert Einstein,*[1] which discusses many of the technical details of the papers published in the period 1907 to 1915. Equally recommendable is the survey by John Stachel of the history of the discovery of both theories of relativity.[2] The thrust of this case study is much more upon the content of the theory than upon its history, but a brief chronology of the discovery is illuminating.

19.1.1 Einstein and Relativistic Gravity in 1907

In his Kyoto address of December 1922,[3] Einstein stated:

> In 1907, while I was writing a review of the consequences of special relativity, . . . I realised that all the natural phenomena could be discussed in terms of special relativity except for the law of gravitation. I felt a deep desire to understand the reason behind this . . . It was most unsatisfactory to me that, although the relation between inertia and energy is so beautifully derived [in special relativity], there is no relation between inertia and weight. I suspected that this relationship was inexplicable by means of special relativity.

In the same lecture, he remarks:

> I was sitting in a chair in the patent office in Bern when all of a sudden a thought occurred to me: "If a person falls freely he will not feel his own weight." I was startled. This simple thought made a deep impression upon me. It impelled me towards a theory of gravitation.

In his comprehensive review of relativity published in 1907,[4] Einstein devoted the whole of the last section, Section V, to 'The Principle of Relativity and Gravitation'. In the very first paragraph, he raised the question,

> Is it conceivable that the principle of relativity also applies to systems that are accelerated relative to one another?

He had no doubt about the answer and stated the *Principle of Equivalence* explicitly for the first time:

> … in the discussion that follows, we shall therefore assume the complete physical equivalence of a gravitational field and a corresponding acceleration of the reference system.

From this postulate, he derived the time-dilation formula in a gravitational field:

$$\mathrm{d}t = \mathrm{d}\tau \left(1 + \frac{\Phi}{c^2}\right), \tag{19.1}$$

where Φ is the gravitational potential, τ is proper time and t is the time measured at zero potential. Then, applying Maxwell's equations to the propagation of light in the gravitational potential $\Phi(\boldsymbol{r})$, he found that the equations are form invariant, provided the speed of light varies as

$$c(\boldsymbol{r}) = c\left(1 + \frac{\Phi(\boldsymbol{r})}{c^2}\right), \tag{19.2}$$

recalling that Φ is always negative. Einstein realised that, as a result of Huygens' principle, or equivalently Fermat's principle of least time, light rays are bent in a non-uniform gravitational field. He was disappointed to conclude that the effect was too small to be detected in any terrestrial experiment.

19.1.2 Bending of Light Rays 1911

Einstein published nothing further on gravity and relativity until 1911. In his paper of that year,[5] he reviewed his earlier ideas, but noted that the gravitational dependence of the speed of light would result in the deflection of the light of background stars by the Sun. Applying Huygens' principle to the propagation of light rays with a variable speed of light, he found the 'Newtonian' result that the angular deflection of light by a mass M would amount to

$$\Delta\theta = \frac{2GM}{pc^2}, \tag{19.3}$$

where p is the collision, or impact, parameter (see Fig. A4.5).[6] For the Sun, this deflection amounted to 0.87 arcsec, although Einstein estimated 0.83 arcsec. Einstein urged astronomers to attempt to measure this deflection. Intriguingly, (19.3) had been derived by Johann Soldner in 1801 on the basis of the Newtonian corpuscular theory of light.[7]

19.1.3 The Years of Toil 1912–15

Following the Solvay conference of 1911, from 1912 to 1915, Einstein's efforts were principally devoted to formulating the relativistic theory of gravity. It was to prove to be a titanic struggle.

During 1912, he realised that he needed more general space-time transformations than those of special relatively. Two quotations illustrate the evolution of his thought:

> The simple physical interpretation of the space-time coordinates will have to be forfeited, and it cannot yet be grasped what form the general space-time transformations could have.[8]

> If all accelerated systems are equivalent, then Euclidean geometry cannot hold in all of them.[9]

Towards the end of 1912, he realised that what was needed was non-Euclidean geometry. From his student days, he vaguely remembered Gauss's theory of surfaces, which had been taught to him by Carl Friedrich Geiser. Einstein consulted his old school friend, the mathematician Marcel Grossmann, about the most general forms of transformation between frames of reference for metrics of the form

$$ds^2 = g_{\mu\nu}\,dx^\mu\,dx^\nu. \tag{19.4}$$

Although outside Grossmann's field of expertise, he soon came back with the answer that the most general transformation formulae were the Riemannian geometries, but that they had the 'bad feature' that they are non-linear. Einstein instantly recognised that, on the contrary, this was a great advantage since any self-consistent theory of relativistic gravity must be non-linear, as discussed in Section 19.2.4.

The collaboration between Einstein and Grossmann was crucial in elucidating the features of Riemannian geometry essential for the development of the theory, Einstein fully acknowledging the central role which Grossmann had played. At the end of the introduction to his first monograph on general relativity, Einstein wrote:

> Finally, grateful thoughts go at this place to my friend the mathematician Grossmann, who by his help not only saved me the study of the relevant mathematical literature but also supported me in the search for the field equations of gravitation.[10]

The Einstein–Grossmann paper[11] of 1913 was the first exposition of the role of Riemannian geometry in the search for a relativistic theory of gravity. The details of Einstein's struggles over the next three years are fully recounted by Pais. It was a huge and exhausting intellectual endeavour which culminated in the presentation of the theory in its full glory in November 1915. In that month, Einstein discovered that he could account precisely for the perihelion shift of Mercury, discovered by Le Verrier in 1859, as a natural consequence of his general theory of relativity. He knew he must be right.

In the next section, we will put flesh on the various arguments which led Einstein to the conclusion that the relativistic theory of gravity had to involve the complexities of Riemannian geometry and the tools needed to relate that structure to the properties of matter and radiation.

19.1.4 The Plan of Attack

The aim of this chapter is to make plausible the basic concepts of space, time and gravity, as embodied in the general theory of relativity. Then, we study one specific solution of the theory, the Schwarzschild solution, to illustrate how Einstein's general relativity differs from Newtonian gravity.

My recommended approach is to begin with Berry's *Principles of Cosmology and Gravitation*[12] and Rindler's *Relativity: Special, General and Cosmological*,[13] both of which are excellent introductory texts. To deepen appreciation of the theory, Weinberg's *Gravitation and Cosmology*[14] shows clearly why general relativity has to be as complex as it is. What I particularly like about his approach is that the physical content of the theory is clearly described at each stage of its development. Another recommendation is d'Inverno's *Introducing Einstein's Relativity*,[15] which is particularly clear on the geometric aspects of the theory. Finally, the real enthusiast should not be without the classic, mammoth volume *Gravitation*[16] by Misner, Thorne and Wheeler.

19.2 Essential Features of the Relativistic Theory of Gravity

19.2.1 The Principle of Equivalence

Einstein's insight that 'a person in free fall will not feel his own weight' is embodied in the *Principle of Equivalence*. The starting point is the comparison of Newton's second law of motion,

$$\boldsymbol{f} = \frac{\mathrm{d}}{\mathrm{d}t}(m\boldsymbol{v}) = m\ddot{\boldsymbol{r}}, \tag{19.5}$$

for the case of point masses with Newton's law of gravity,

$$\boldsymbol{f} = -\frac{Gm_1m_2}{r^2}\,\boldsymbol{i}_r. \tag{19.6}$$

Three different types of mass appear in these formulae. The mass which appears in (19.5) is the *inertial mass* of the body m_{i}, meaning the constant which appears in the relation between force and acceleration. It describes the resistance, or inertia, of the body to changes in its state of motion, quite independent of the nature of the impressed force $\boldsymbol{f}$.

The masses appearing in (19.6) are *gravitational masses*, but again care is needed. Suppose we ask, 'What is the gravitational force acting on the mass m_2 due to the presence of the mass m_1?' The answer is in two parts. First, the mass m_1 is the source of a gravitational field $\boldsymbol{g}_2$ at the location of the mass m_2,

$$\boldsymbol{g}_2 = -\frac{Gm_1}{r^2}\,\boldsymbol{i}_r.$$

The mass m_1, the source of the field, is called the *active gravitational mass*. When the mass m_2 is placed in this field, it experiences a force $m_2\boldsymbol{g}_2$, describing the response of the body to the pre-existing field, and so m_2 is called the *passive gravitational mass*. The proportionality of the active and passive gravitational masses follows directly from Newton's third law of motion,

$$\boldsymbol{f}_1 = -\boldsymbol{f}_2,$$

that is,

$$Gm_{1\mathrm{a}}m_{2\mathrm{p}} = Gm_{2\mathrm{a}}m_{1\mathrm{p}} \qquad \text{and so} \qquad \frac{m_{1\mathrm{a}}}{m_{2\mathrm{a}}} = \frac{m_{1\mathrm{p}}}{m_{2\mathrm{p}}},$$

where the subscripts a and p refer to the active and passive gravitational masses respectively. Since the ratios of the masses of the two bodies are the same, the active and passive gravitational masses can be taken to be identical.

The relation between gravitational and inertial mass has to be found from experiment. The gravitational mass which appears in Newton's law of gravity is akin to the electric charge in electrostatics – they both describe the forces associated with a particular physical property of the bodies involved. In the case of electrostatics, there is a clear distinction between the electric charge of a body and its inertial mass. In fact, it would be more helpful to refer to the gravitational mass as a 'gravitational charge', to emphasise the physical distinction between it and inertial mass.

In deriving his law of gravity, Newton was well aware of the fact that he had assumed that gravitational and inertial mass are the same thing – this was implicit in his analysis which demonstrated that the planets move in elliptical orbits about the Sun. Newton tested this assumption by comparing the periods of pendulums of identical dimensions, but made of different materials. Preserving the distinction between gravitational and inertial mass, the equation of simple harmonic motion for a simple pendulum of length l is

$$\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} = -\left(\frac{m_\mathrm{g}}{m_\mathrm{i}}\right)\left(\frac{g}{l}\right)\theta. \tag{19.7}$$

Newton showed that the inertial and gravitational masses are proportional to each other to better than one part in 1000.

The expression for centripetal force also involves the inertial mass and so the same type of test can be carried out for bodies in which the centripetal force is balanced by the gravitational force. This approach was subject to increasingly accurate measurements by Loránd Eötvös in the 1880s, by Dicke in 1964 and by Braginskii in 1971. In the version of Dicke's experiment carried out by Braginskii and Panov, the linearity of the relation between gravitational and inertial mass was established to about one part in 10^{12}. These experiments and many others in gravitational physics are carefully reviewed by Clifford Will in his excellent book *Theory and Experiment in Gravitational Physics*.[17] His article in *Living Reviews* entitled *The Confrontation between General Relativity and Experiment* also provides an in depth assessment of the various ways in which the equivalence principle might need to be revised in the light of high precision tests of the principle.[18]

The precise linearity of the relation between the two masses enables rather subtle tests to be carried out. For example, the inertial mass associated with the kinetic energy $T = (\gamma - 1)mc^2$ of electrons moving at relativistic speeds in atoms, $m = T/c^2$, behaves exactly as a gravitational mass of the same magnitude. There is a need to continue to improve the precision with which the linearity of this relation is established. The French MICROSCOPE project[19] has already improved the limits by an order of magnitude with the ultimate objective of reaching a sensitivity level of one part in 10^{-15}. Any deviation from exact proportionality would have a profound impact upon fundamental physics.

The null result of these experiments is the experimental basis for Einstein's *Principle of Equivalence*. Rindler quote's Einstein's statement of the principle as follows:

> All local, freely falling, non-rotating laboratories are fully equivalent for the performance of all physical experiments.[20]

Einstein expressed this somewhat more *in extenso* in an unpublished review of the general theory written in January 1920, now in the Pierpoint Morgan Library in New York City:

> ... *for an observer falling freely from the roof of a house there exists* – at least in his immediate surroundings – *no gravitational field* [Einstein's italics]. Indeed, if the observer drops some bodies, then these remain relative to him in the same state of rest or of uniform motion, independent of their particular chemical or physical nature. ... The observer therefore has the right to interpret his state as "at rest".[21]

An equivalent statement of the principle is as follows:

> At any point in a gravitational field, in a frame of reference moving with the free-fall acceleration at that point, all the laws of physics have their usual special relativistic form, except for the force of gravity, which disappears locally.[22]

By *free fall*, we mean a frame of reference which is accelerated at the gravitational acceleration at that point in space, $\boldsymbol{a} = \boldsymbol{g}$. These statements formally identify inertial and gravitational mass, since the force acting on a particle in a gravitational field depends upon the particle's *gravitational mass*, whereas the acceleration depends upon its *inertial mass*.

The first statement of the principle is a natural extension of the postulates which form the basis of special relativity. In particular, any two inertial frames in relative motion are 'freely falling' frames in zero gravity. Hence, the interval between events must be given by the metric of special relativity, the Minkowski metric,

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \frac{1}{c^2}\,\mathrm{d}\boldsymbol{r}^2. \tag{19.8}$$

But the principle of equivalence goes far beyond this. Inside a 'freely falling laboratory', there is no gravitational field. By transforming into such a frame of reference, gravity is replaced at that point in space by an accelerated frame of reference. The principle of equivalence asserts that the laws of physics should be identical in all such accelerated frames of reference. Einstein's great insight was to 'abolish' gravity and replace it by appropriate transformations between accelerated frames of reference.

To reinforce the idea of the equivalence of gravity and accelerations locally, it is useful to consider pictures of laboratories in free fall, in gravitational fields and accelerated so as to mimic the effects of gravity. In Figs. 19.1(*a*) and (*b*), the equivalence of a static gravitational field and an accelerated laboratory is demonstrated by a spring balance. For

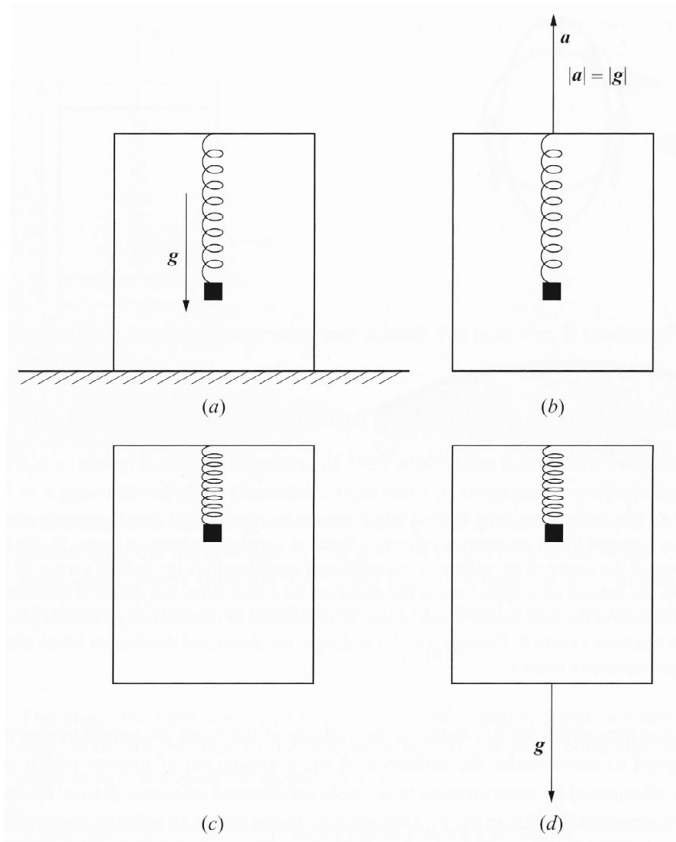


Fig. 19.1 Illustrating the equivalence between gravitational fields and accelerated reference frames. In each drawing the laboratory has a spring suspended from the ceiling with a weight attached. In (*a*), the lift is stationary in a gravitational field of strength $\boldsymbol{g}$. In (*b*), the lift is accelerated with acceleration $|\boldsymbol{a}| = |\boldsymbol{g}|$ in the absence of a gravitational field. In (*a*) and (*b*), the springs are stretched. In (*c*), the lift is located in a region of zero gravitational acceleration. In (*d*), the lift is in free fall in a gravitational field of strength $\boldsymbol{g}$. In (*c*) and (*d*), there is no extension of the spring.

comparison, a laboratory in a 'region of zero gravitational field' and one in free fall in a gravitational field of field strength $\boldsymbol{g}$ are illustrated in Figs. 19.1(c) and (d) respectively.

Although the *acceleration* due to gravity can be eliminated at a particular point in space, the effects of gravity cannot be eliminated in the vicinity of that point. This is most easily seen by considering the gravitational field at distance r in the vicinity of a point mass M, or equivalently, a spherically symmetric mass, such as the Earth (Fig. 19.2). It is apparent that different freely falling laboratories are needed at different points in space in order to eliminate gravity everywhere.

If very precise measurements are made in the vicinity of the mass M, neighbouring particles are observed to move under the influence of the components of gravity which were not precisely eliminated by transforming to a single accelerated reference frame.

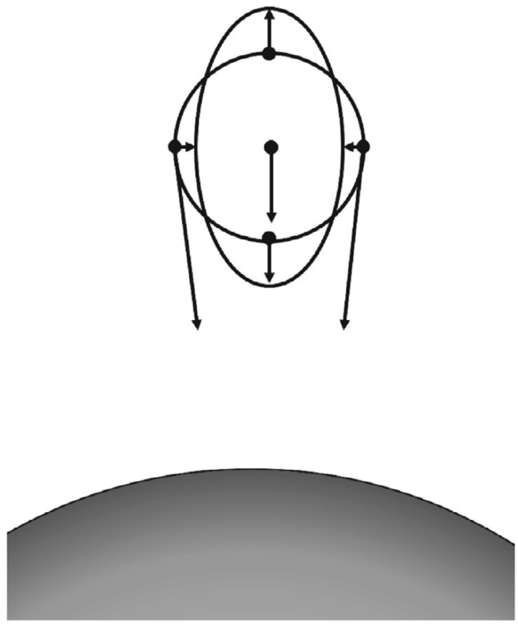


Fig. 19.2 Illustrating the 'tidal forces' which cannot be eliminated when the acceleration due to gravity ***g*** is replaced by an accelerated reference frame at a particular point in space. In this example, the observer is in centrifugal equilibrium in the field of the spherical Earth. Initially, test masses are located on a sphere about the observer. At a later time, the sphere is distorted into an ellipsoid because of the tidal forces, which cannot be eliminated by transforming away the gravitational field at the observer. (After R. Penrose, 1997, *The Large, the Small and the Human Mind*. Cambridge: Cambridge University Press.)

For example, consider a standard Euclidean (x, y, z) coordinate frame inside an orbiting space station, the z-coordinate being taken to lie in the radial direction. It is a useful exercise to show that, if two test particles are released from rest, with an initial separation vector ξ, the separation vector varies with time as

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\begin{bmatrix}\xi^x \\ \xi^y \\ \xi^z\end{bmatrix} = \begin{bmatrix} -\dfrac{GM}{r^3} & 0 & 0 \\ 0 & -\dfrac{GM}{r^3} & 0 \\ 0 & 0 & +\dfrac{2GM}{r^3}\end{bmatrix}\begin{bmatrix}\xi^x \\ \xi^y \\ \xi^z\end{bmatrix}. \tag{19.9}$$

The uncompensated forces vary as r^{-3}, exactly the same dependence as the 'tidal forces', which cause Earth–Moon and Earth–Sun tides. We therefore need a theory which reduces locally to Einstein's special relativity in a freely falling reference frame, and which transforms correctly into another freely falling frame when we move to a slightly different point in space. There is no such thing as a global Lorentz frame in the presence of a non-uniform gravitational field.

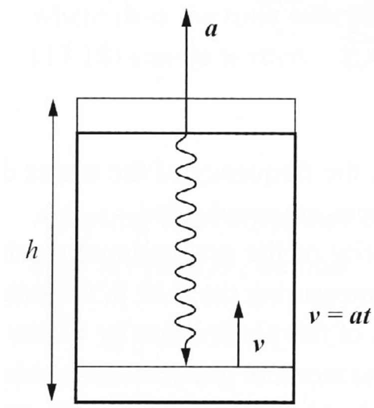


Fig. 19.3 Illustrating the gravitational redshift. The laboratory is accelerated to velocity $|\boldsymbol{a}t|$ after time t.

19.2.2 The Gravitational Redshift and Time Dilation in a Gravitational Field

Let us repeat Einstein's argument of 1907 concerning the redshift of electromagnetic waves in a gravitational field. Consider a light wave of frequency ν propagating from the ceiling to the floor of a lift in a gravitational field $\boldsymbol{g}$. The gravitational force can be mimicked by accelerating the lift upwards so that $\boldsymbol{g} = -\boldsymbol{a}$ (Fig. 19.3). It is assumed that the acceleration is small. If the height of the lift is h, the light signal travels from the roof to the floor in a time $t = h/c$. In this time, the floor is accelerated to a speed $u = at = |\boldsymbol{g}|t$ and hence,

$$u = \frac{|\boldsymbol{g}|h}{c}. \tag{19.10}$$

Therefore, the light wave is observed with a higher frequency when it arrives at the floor of the lift because of the Doppler effect. To first order in u/c, the observed frequency ν' is

$$\nu' = \nu\left(1 + \frac{u}{c}\right) = \nu\left(1 + \frac{|\boldsymbol{g}|h}{c^2}\right). \tag{19.11}$$

Let us now express this result in terms of the change in gravitational potential between the ceiling and floor of the lift. Since $\boldsymbol{g} = -\text{grad}\,\phi$,

$$|\boldsymbol{g}| = -\frac{\Delta\phi}{h}. \tag{19.12}$$

Notice that, because of the attractive nature of the gravitational force, ϕ is more negative at $h = 0$ than at the ceiling. Therefore,

$$\nu' = \nu\left(1 - \frac{\Delta\phi}{c^2}\right). \tag{19.13}$$

This is the formula for the *gravitational redshift* z_g in the 'Newtonian' limit, the redshift being defined to be

$$z = \frac{\lambda_\mathrm{obs} - \lambda_\mathrm{em}}{\lambda_\mathrm{em}} = \frac{\nu' - \nu}{\nu'}, \tag{19.14}$$

and so

$$|z_g| \approx \frac{\Delta\phi}{c^2}, \tag{19.15}$$

for very small changes $|\Delta\phi|/c^2 \ll 1$. Thus, the frequency of the waves depends upon the *gravitational potential* in which the light waves are propagated.

Laboratory experiments to measure the gravitational redshift were carried out by Pound and Rebka[23] in 1959 and by Pound and Snider in 1964.[24] They measured the difference in redshift of γ-ray photons propagating up and then down a tower 22.5 m high at Harvard University using the Mössbauer effect. In this effect, the recoil effects of the emission and absorption of the γ-ray photons are zero, since the momentum is absorbed by the atomic lattice as a whole. The γ-ray resonance is very sharp indeed and only tiny Doppler shifts are needed to move off resonance absorption. In the Harvard experiment, the predicted difference in redshifts for γ-ray photons travelling up and down the tower were

$$z_{\rm up} - z_{\rm down} = \frac{2gh}{c^2} = 4.905 \times 10^{-15}.$$

The measured value was $(4.900 \pm 0.037) \times 10^{-15}$, providing excellent agreement within about 1%.

Notice that the gravitational redshift is incompatible with special relativity. According to that theory, observers at the top and bottom of the tower are considered to be at rest in the same inertial frame of reference.

Suppose we now write (19.13) in terms of the period of the waves T. Then,

$$T' = T\left(1 + \frac{\Delta\phi}{c^2}\right). \tag{19.16}$$

This expression is exactly the same as the time dilation formula between inertial frames of reference in special relativity, only now the expression refers to different locations in the gravitational field. This expression for the time dilation is exactly what would be evaluated for any time interval, and so we can write in general

$$\mathrm{d}t' = \mathrm{d}t\left(1 + \frac{\Delta\phi}{c^2}\right). \tag{19.17}$$

Let us take the gravitational potential to be zero at infinity and measure the gravitational potential at any point in the field relative to that value. Assuming the weak field limit in which changes in the gravitational potential are small, at any point in the gravitational field, we can write

$$\mathrm{d}t'^2 = \mathrm{d}t^2\left[1 + \frac{\phi(r)}{c^2}\right]^2, \tag{19.18}$$

where $\mathrm{d}t$ is the time interval measured at $\phi = 0$, that is, at $r = \infty$. Since $\phi(r)/c^2$ is small, (19.18) can be written

$$\mathrm{d}t'^2 = \mathrm{d}t^2\left[1 + \frac{2\phi(r)}{c^2}\right]. \tag{19.19}$$

Adopting the Newtonian expression for the gravitational potential of a point mass M, $\phi(r) = -GM/r$, we find

$$\mathrm{d}t'^2 = \mathrm{d}t^2 \left(1 - \frac{2GM}{rc^2}\right). \tag{19.20}$$

Let us now introduce (19.20) into the Minkowski metric of special relativity,

$$\mathrm{d}s^2 = \mathrm{d}t'^2 - \frac{1}{c^2}\,\mathrm{d}l^2, \tag{19.21}$$

where dl is the differential element of proper distance. The metric of space-time about the point mass can therefore be written

$$\mathrm{d}s^2 = \mathrm{d}t^2 \left(1 - \frac{2GM}{rc^2}\right) - \frac{1}{c^2}\,\mathrm{d}l^2. \tag{19.22}$$

This calculation suggests how the metric coefficients become more complicated than those of Minkowski space when we attempt to construct a relativistic theory of gravity. It is important to appreciate that, in the above analysis, the times dt and dt' are *proper times*. Thus, if we consider the observer to be located at zero potential at $r = \infty$, dt is the proper time measured on that clock at a fixed point in space, that is, the interval ds^2 is a pure time interval and hence proper time. In exactly the same way, the observer at rest at gravitational potential $\Delta\phi$ measures a proper time interval dt'. Since $\Delta t \neq \Delta t'$, there is a synchronisation problem – the rate at which proper time runs depends upon the gravitational potential.

The result of the last paragraph is the 'general relativistic' version of the twin paradox. If one twin goes on a cosmic journey through a deeper gravitational potential than the twin who stays at home, less proper time passes on the clock of the travelling twin than on the clock of the less adventurous twin. This is no more than time dilation in a gravitational field. The clock synchronisation problem is resolved if we decide to refer all time intervals to a *coordinate time* which takes the same value at all points in space. The natural time interval is dt, the proper time measured at infinity, which can then be related to that measured in a gravitational potential $\phi(r)$ by (19.19). Notice that both dt and dt' are different from the time difference dτ measured by an observer in free fall in the gravitational field at the same point in space.

The difference between dt and dt' as expressed by (19.17) enables us to understand Einstein's result of 1907 that the speed of light can be considered to depend upon the gravitational potential ϕ. At some point in the field, light propagates at the speed of light, that is, dl/d$t' = c$. When we measure the speed of light as observed at infinity, however, the observer measures dt and not dt'. Hence, the observer at infinity measures the speed of light in the radial direction to be less than c by the factor appearing in (19.17), that is

$$c(r) = c\left(1 + \frac{\Delta\phi}{c^2}\right).$$

Note that this is not the complete story, because we have not yet dealt with the space component of the metric. As Einstein realised, this is a convenient expression for estimating time-lag effects when an electromagnetic wave passes through a varying gravitational potential and also for working out the deflection of electromagnetic waves in a varying gravitational potential.

One final comment about the gravitational redshift is worth making. Notice that it is no more than an expression of the conservation of energy in a gravitational field. In the final form of the metric, this property will be built into the metric coefficients.

19.2.3 Space Curvature

Let us illustrate how the expression for the spatial component of the metric (19.22), $\mathrm{d}l$, has to change as well. Consider the propagation of light rays in our lift, but now travelling perpendicular to the gravitational acceleration. Again, we use the principle of equivalence to replace the gravitational field in a stationary laboratory by an accelerated lift in free space (Fig. 19.4).

In the time the light ray propagates across the lift, a distance l, the lift moves upwards a distance $\frac{1}{2}|\boldsymbol{g}|t^2$. Therefore, in the frame of reference of the accelerated lift, and also in the stationary frame in the gravitational field, the light ray follows a parabolic path as illustrated in the diagram. Let us approximate the light path by a circular arc of radius R. The length of the chord d across the circular arc is then

$$d^2 = \tfrac{1}{4}|\boldsymbol{g}|^2 t^4 + l^2. \tag{19.23}$$

From the geometry of the diagram, it can be seen that the angle $\phi = |\boldsymbol{g}|t^2/l$. Hence, since $R\phi = d$,

$$R^2 = \frac{d^2}{\phi^2} = \frac{l^2}{4} + \frac{l^4}{|\boldsymbol{g}|^2 t^4}. \tag{19.24}$$

Now, $\frac{1}{2}|\boldsymbol{g}|t^2 \ll l$, $l = ct$ and so

$$R = \frac{l^2}{|\boldsymbol{g}|t^2} = \frac{c^2}{|\boldsymbol{g}|}. \tag{19.25}$$

Thus, the radius of curvature of the path of the light rays depends only upon the local gravitational acceleration $|\boldsymbol{g}|$. Since $\boldsymbol{g}$ is determined by the gradient of the gravitational potential, it follows that the curvature of the paths of light rays depends upon the mass distribution.

This simple argument indicates that, according to the principle of equivalence, the geometry along which light rays propagate in an accelerated laboratory, which we recall is the frame in which gravity has been abolished, is not flat but curved space, the spatial curvature depending upon the local value of the gravitational acceleration $\boldsymbol{g}$. It was arguments of this nature which convinced Einstein that he had to take non-Euclidean geometries seriously in order to derive a relativistic theory of gravity. In other words, the appropriate geometry for '$\mathrm{d}\boldsymbol{l}^2$' in the metric (19.22) must, in general, refer to curved space. But, we have just shown in Section 19.2.2 that the time component of the metric also depends upon the gravitational potential and so, in the full theory, the metric must refer to four-dimensional curved space-time.

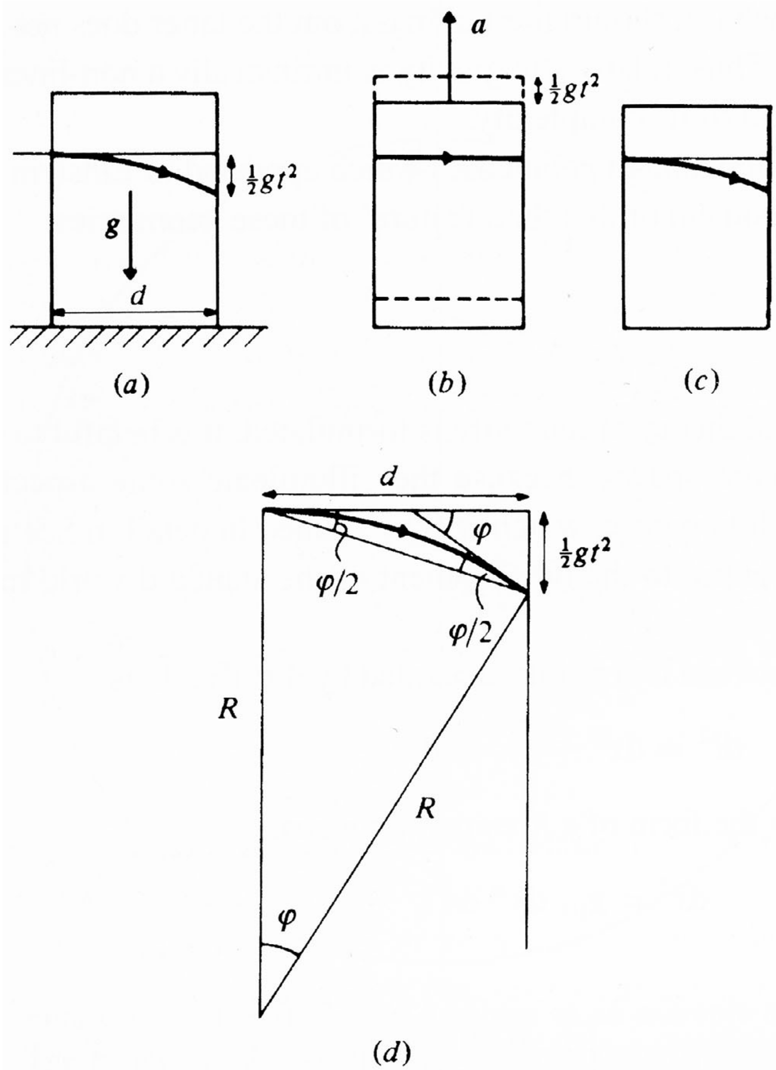


Fig. 19.4 Illustrating the gravitational deflection of light according to the principle of equivalence. (*a*) The path of a light ray in a gravitational field $\boldsymbol{g}$. (*b*) The path of the light ray in the equivalent accelerated frame of reference, as viewed by a stationary external observer. The dashed line shows the position of the laboratory after time t. (*c*) The path of the light ray as observed in the accelerated frame of reference. (*d*) The geometry of light deflection in an accelerated frame of reference.

19.2.4 The Non-Linearity of Relativistic Gravity

To complicate matters further, any relativistic theory of gravity must be non-linear. This follows from Einstein's mass–energy relation $E = mc^2$ as applied to the gravitational field. The gravitational field of some mass distribution has a certain local energy density at each point in the field. Since $E = mc^2$, it follows that there is a certain inertial mass density in the gravitational field which is itself a source of gravitational field. This property contrasts with, say, the case of an electric field which possesses a certain amount of electromagnetic field energy and which consequently has a certain inertial mass, but this does not generate additional electrostatic charge. Thus, relativistic gravity is intrinsically a non-linear theory and this accounts for a great deal of its complexity. This was the aspect of the Riemannian geometries which appealed to Einstein in 1912, when he was informed by Grossmann of this 'bad feature' of these geometries.

19.3 Isotropic Curved Spaces

Before outlining how the general theory is formulated, it is helpful to consider the properties of isotropic curved spaces because they illuminate some aspects of the interpretation of the Schwarzschild metric and because these spaces are central to the development of the standard world models in Chapter 20.

19.3.1 A Brief History of Non-Euclidean Geometries

During the late eighteenth century, non-Euclidean geometries began to be taken seriously by mathematicians who realised that the fifth postulate of Euclid, that parallel lines meet only at infinity, might not be essential for the construction of self-consistent geometries.[25] The first proposals that the global geometry of space might not be Euclidean were discussed by Girolamo Saccheri and Johann Lambert. In 1766, Lambert noted that, if space were hyperbolic rather than flat, the radius of curvature of the space could be used as an absolute measure of distance.[26] In 1816, Carl Friedrich Gauss repeated this proposal in a letter to Christian Gerling and was aware of the fact that a test of the local geometry of space could be carried out by measuring the sum of the angles of a triangle between three high peaks in the Harz mountains, the Brocken, Hoherhagen and Inselberg. In 1818, Gauss was asked to carry out a geodetic survey of the state of Hanover and he devoted a large effort to carrying out and reducing the data himself. He was certainly aware of the fact that the sum of the angles of the triangle formed by the three Harz Mountains was 180° within the limits of geodetic measurements.[27]

The fathers of non-Euclidean geometry were Nicolai Ivanovich Lobachevsky, who became rector of Kazan University in Russia in 1827, and János Bolyai in Transylvania, then part of Hungary. In the 1820s, they independently solved the problem of the existence of non-Euclidean geometries and showed that Euclid's fifth postulate could not be deduced from the other postulates.[28] In his papers entitled *On the Principles of Geometry*, Lobachevsky also proposed an astronomical test of the geometry of space. If the geometry were hyperbolic, the minimum parallax of any object would be

$$\theta = \arctan\left(\frac{a}{\mathcal{R}}\right),$$

where a is the radius of the Earth's orbit and $\mathcal{R}$ the radius of curvature of the geometry. He found a minimum value of $\mathcal{R} \geq 1.66\times10^5$ AU = 2.6 light years. It is intriguing that this estimate was made 8 years before Bessel's announcement of the first successful parallax measurement of 61 Cygni. In making this estimate, Lobachevsky used the observational upper limit of 1 arcsec for the parallax of bright stars. In a statement which will warm the hearts of observational astronomers, he remarked,

> There is no means other than astronomical observations for judging the exactness which attaches to the calculations of ordinary geometry.

Non-Euclidean geometry was placed on a firm theoretical basis by Bernhard Riemann[29] and the English-speaking world was introduced to these ideas through the works of William Kingdon Clifford and Arthur Cayley. In 1900, Karl Schwarzschild returned to the problem of the geometry of space and was able to set more stringent limits to its radius of curvature. Repeating Lobachevsky's argument, he found $\mathcal{R} \geq 60$ light years if space were hyperbolic.[30] If space were closed, he could set limits to the radius of curvature of the closed geometry because the total volume of the closed space is $V = 2\pi^2\mathcal{R}^3$. Since there were only 100 stars with measurable parallaxes and at least 10^8 for which no parallax could be measured, he concluded that $\mathcal{R} \geq 2500$ light years. He also noted that, if space were spherical, it should be possible to observe an image of the Sun in the direction precisely 180° away from its direction on the sky at any time.

Until Einstein's discovery of the general theory of relativity, considerations of the geometry of space and the role of gravity in defining the large-scale structure of the Universe were separate questions. After 1915, they were inextricably linked.

It turns out that it is not necessary to become embroiled in the details of Riemannian geometry to appreciate the geometrical properties of isotropic curved spaces. We can demonstrate simply why the only isotropic curved spaces are those in which the two-dimensional curvature of any space section κ is constant throughout the space and can only take positive, zero or negative values.[31]

19.3.2 Parallel Transport and Isotropic Curved Spaces

Consider first of all the simplest two-dimensional curved geometry, the surface of a sphere (Fig. 19.5). In the diagram, a triangle is shown consisting of two lines drawn from the north pole down to the equator, the triangle being completed by the line drawn along the equator. For simplicity, we consider first the case in which the angle between the lines AB and AC at the 'north pole' is 90°. The three sides of the triangle are all segments of great circles on the sphere and so are the shortest distances between the three corners of the triangle. The three lines are *geodesics* in the curved geometry.

We need a procedure for working out how non-Euclidean the curved geometry is. The way this is done in general is by the procedure known as the *parallel displacement* or *parallel transport* of a vector on making a complete circuit around a closed figure such as the triangle in Fig. 19.5. If the surface were flat Euclidean space, we know that the answer would be 180°, but for the spherical surface shown in Fig. 19.5 the answer is different. Suppose we start with a little vector perpendicular to AC at the pole and lying in the surface of the sphere. We then transport that vector from A to C, keeping it perpendicular to AC. At C, we rotate the vector through 90° so that it is now perpendicular to CB. We then transport

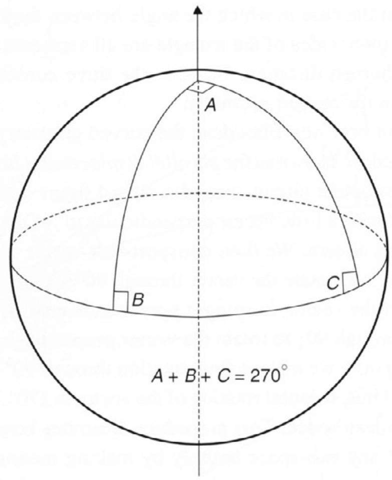


Fig. 19.5 Illustrating the sum of the angles of a triangle on the surface of a sphere.

the vector, keeping it perpendicular to BC to the corner B. We make a further rotation through 90° to rotate the vector perpendicular to BA and then transport it back to A. At that point, we make a final rotation through 90° to bring the vector back to its original direction. Thus, the total rotation of the vector is 270°. Clearly, the surface of the sphere is a non-Euclidean space. This procedure illustrates how the geometrical properties of any two-space can be worked out, entirely by making measurements within the two-space.

Another simple calculation illustrates an important feature of parallel transport on the surface of a sphere. Suppose the angle at A is not 90°, but some arbitrary angle θ. Then, if the radius of the sphere is R_c, the surface area of the triangle ABC is $A = \theta R_c^2$. Thus, if $\theta = 90°$, the area is $\pi R_c^2/2$ and the sum of the angles of the triangle is 270°; if $\theta = 0°$, the area is zero and the sum of the angles of the triangle is 180°. Evidently, the difference of the sum of the angles of the triangle from 180° is proportional to the area of the triangle, that is

$$(\text{Sum of angles of triangle} - 180°) \propto (\text{Area of triangle}). \tag{19.26}$$

This result is a general property of isotropic curved spaces.

Let us now take a vector around a loop in some general two-space (Fig. 19.6). Define a small area $ABCD$ by light rays which travel along geodesics in the two-space. Consider two light rays, OAD and OBC, originating from O, which are crossed by the light rays AB and DC as shown in Fig. 19.6(a). The light rays AB and DC are chosen so that they cross the light ray OAD at right angles. We start at A with the vector to be transported round the loop parallel to AD as shown. We then parallelly transport the vector along AB until it reaches B. At this point, it has to be rotated through an angle $90° - \beta$ in order to lie perpendicular to BC at B. It is then transported along BC to C, where it is rotated through an angle $90° + (\beta + \mathrm{d}\beta)$, as shown in the diagram, in order to lie perpendicular to CD at C. It can be seen by inspection that the extra rotations at B and C are in opposite directions.

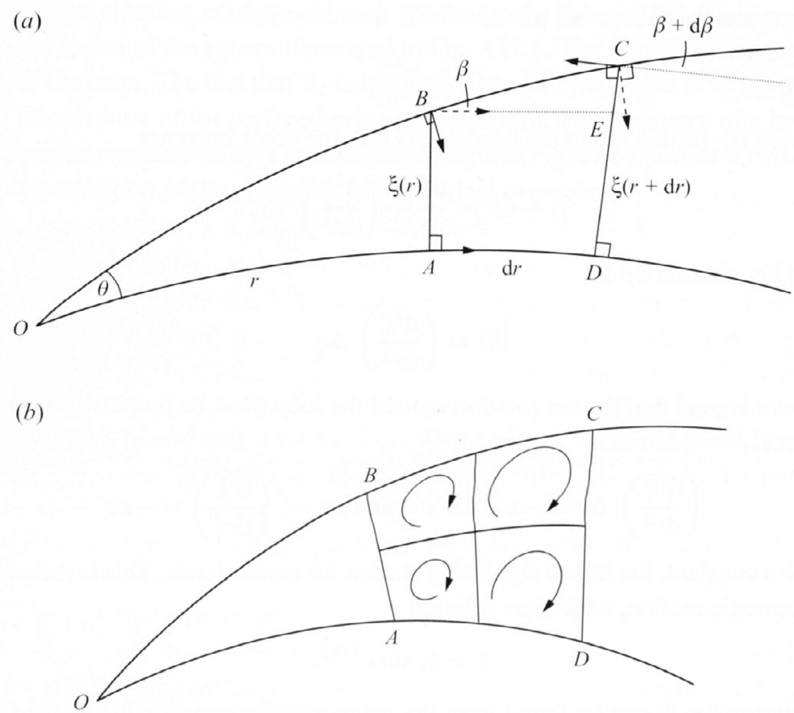


Fig. 19.6 (*a*) Illustrating the parallel transport of a vector around the small loop *ABCD* formed by the light rays *OBC*, *OAD*, *AB* and *DC*. (*b*) Illustrating how the sum of the rotations around the sub-loops must add up linearly to a total rotation dβ.

Since CD and AB are perpendicular to the light rays, the subsequent rotations of the vector on parallel transport amount to 180° and so the total rotation round the loop is 360° + dβ. Thus, dβ is a measure of the departure of the two-space from Euclidean geometry, for which dβ = 0.

Now, the rotation of the vector dβ must depend upon the area of the loop. In the case of an isotropic space, we should obtain the same rotation wherever we place the loop in the two-space. Furthermore, if we were to split the loop into a number of sub-loops, the rotations around the separate sub-loops must add up linearly to the total rotation dβ (Fig. 19.6(*b*)). Thus, in an isotropic two-space, the rotation dβ should be proportional to the area of the loop $ABCD$. We conclude that the constant relating the rotation dβ to the area of the loop must be a constant everywhere in the two-space, just as was found in the particular case of a spherical surface in Fig. 19.5.

The complication is that, since the space is non-Euclidean, we do not know how to relate the length AB to the geodesic distance $OA = r$ along the light path and the angle θ subtended by AB at the origin. Therefore, we have to write the distance AB as an unknown function of r, $\xi(r)$, such that

$$\theta = \frac{\xi(r)}{r}. \tag{19.27}$$

The angle of rotation β can now be found by determining how ξ changes when we move a distance dr along the geodesic:

$$\xi(r + \mathrm{d}r) = \xi(r) + \frac{\mathrm{d}\xi}{\mathrm{d}r}\,\mathrm{d}r. \tag{19.28}$$

Thus, the rotation at B is

$$\beta = \frac{\mathrm{d}\xi}{\mathrm{d}r}. \tag{19.29}$$

At a distance Δr further along the geodesic, the rotation becomes

$$\beta + \mathrm{d}\beta = \frac{\mathrm{d}\xi}{\mathrm{d}r} + \left(\frac{\mathrm{d}^2\xi}{\mathrm{d}r^2}\right)\Delta r, \tag{19.30}$$

and so the net rotation dβ is

$$\mathrm{d}\beta = \left(\frac{\mathrm{d}^2\xi}{\mathrm{d}r^2}\right)\Delta r. \tag{19.31}$$

But, we have argued that the net rotation around the loop must be proportional to the area of the loop, $\mathrm{d}A = \xi\Delta r$,

$$\left(\frac{\mathrm{d}^2\xi}{\mathrm{d}r^2}\right)\Delta r = -\kappa\,\Delta r\,\xi \quad \text{and so} \quad \left(\frac{\mathrm{d}^2\xi}{\mathrm{d}r^2}\right) = -\kappa\xi, \tag{19.32}$$

where κ is a constant, the minus sign being chosen for convenience. This is the equation of simple harmonic motion which has solution

$$\xi = \xi_0 \sin \kappa^{1/2} r. \tag{19.33}$$

ξ_0 can be found from the value of ξ for very small values of r which must reduce to the Euclidean expression $\theta = \xi/r$. Therefore, $\xi_0 = \theta/\kappa^{1/2}$ and

$$\xi = \frac{\theta}{\kappa^{1/2}} \sin \kappa^{1/2} r. \tag{19.34}$$

κ is defined to be the *curvature* of the two-space and can be either positive, negative or zero. If it is negative, we can write $\kappa = -\kappa'$, where κ' is positive and then the circular functions become hyperbolic functions:

$$\xi = \frac{\theta}{\kappa'^{1/2}} \sinh \kappa'^{1/2} r. \tag{19.35}$$

In the case $\kappa = 0$, we find the Euclidean result:

$$\xi = \theta r. \tag{19.36}$$

The results include all possible isotropic curved two-spaces. The constant κ can be positive, negative or zero corresponding to spherical, hyperbolic and flat spaces respectively. In geometric terms, $R_\mathrm{c} = \kappa^{-1/2}$ is the radius of curvature of a two-dimensional section through the isotropic curved space and has the same value at all points and in all orientations within the plane. It is often convenient to write the expression for ξ in the form

$$\xi = \theta R_\mathrm{c} \sin \frac{r}{R_\mathrm{c}}, \tag{19.37}$$

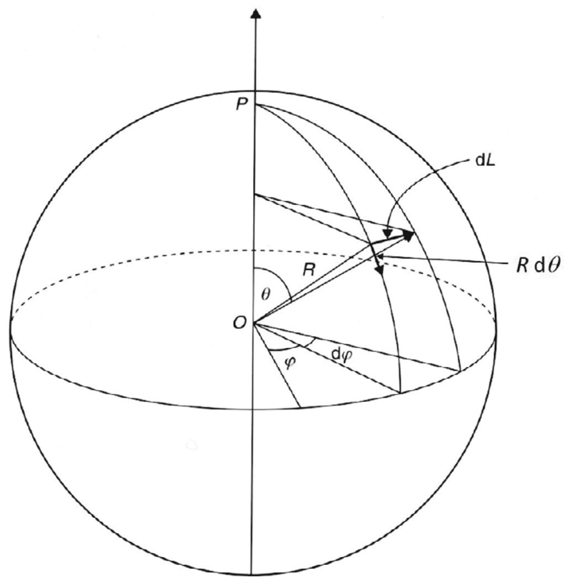


Fig. 19.7 The surface of a sphere as an example of a two-dimensional isotropic curved space.

where R_c is real for closed spherical geometries, imaginary for open hyperbolic geometries and infinite for the case of Euclidean geometry.

19.3.3 The Metric of an Isotropic Curved Space

The simplest examples of such spaces are the spherical geometries in which R_c is just the radius of the sphere as illustrated in Fig. 19.7. The hyperbolic spaces are more difficult to envisage. The fact that R_c is imaginary can be interpreted in terms of the principal radii of curvature of the surface having opposite sign. The geometry of a hyperbolic two-sphere can be represented by a saddle-shaped figure (Fig. 19.8), just as a two-sphere provides a visualisation of the properties of a spherical two-space.

In flat space, the distance between two points separated by $\mathrm{d}x, \mathrm{d}y, \mathrm{d}z$ is

$$\mathrm{d}l^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2.$$

If we write

$$\mathrm{d}s^2 = g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu, \qquad \text{then} \qquad g_{\mu\nu} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Now consider the simplest two-dimensional isotropic curved space, the surface of a sphere. This 'two-space' is isotropic because the radius R of the sphere is the same at all points on the surface. An orthogonal frame of reference can be set up at each point locally on the surface of the sphere. Let us choose spherical polar coordinates to describe positions of

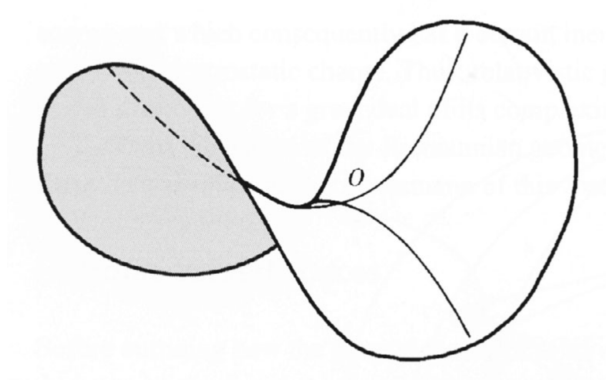


Fig. 19.8 Illustrating the geometry of an isotropic hyperbolic two-space. The principal radii of curvature of the surface are equal in magnitude, but have opposite signs in orthogonal directions.

points on the surface of the two-sphere as indicated in Fig. 19.7. In this case, the natural orthogonal coordinates are the angular coordinates θ and ϕ, and hence

$$\mathrm{d}x^1 = \mathrm{d}\theta, \qquad \mathrm{d}x^2 = \mathrm{d}\phi.$$

From the geometry of the sphere, we know that

$$\mathrm{d}l^2 = R^2\,\mathrm{d}\theta^2 + \mathrm{d}L^2 = R^2\,\mathrm{d}\theta^2 + R^2\sin^2\theta\,\mathrm{d}\phi^2. \tag{19.38}$$

Therefore, the elements of the metric tensor are

$$g_{\mu\nu} = \begin{bmatrix} R^2 & 0 \\ 0 & R^2\sin^2\theta \end{bmatrix}. \tag{19.39}$$

The key point is that the metric tensor $g_{\mu\nu}$ contains information about the intrinsic geometry of the two-space. We need a prescription which enables us to find the intrinsic geometry of the surface from the components of the metric tensor. In the case of a sphere, this may appear hardly necessary, but we could easily have chosen some strange set of coordinates, which would have obscured the intrinsic geometry of the space.

This is where the memory of Gauss's theory of surfaces played a key role in Einstein's thinking in 1912. For the case of two-dimensional metric tensors, which can be reduced to diagonal form, that is, with $g_{12} = g_{21} = 0$, Gauss showed that the curvature of the surface is given by the formula

$$\kappa = \frac{1}{2g_{11}g_{22}}\left\{-\frac{\partial^2 g_{11}}{\partial x_2^2} - \frac{\partial^2 g_{22}}{\partial x_1^2} + \frac{1}{2g_{11}}\left[\frac{\partial g_{11}}{\partial x_1}\frac{\partial g_{22}}{\partial x_1} + \left(\frac{\partial g_{11}}{\partial x_2}\right)^2\right] + \frac{1}{2g_{22}}\left[\frac{\partial g_{11}}{\partial x_2}\frac{\partial g_{22}}{\partial x_2} + \left(\frac{\partial g_{22}}{\partial x_1}\right)^2\right]\right\}. \tag{19.40}$$

A proof of this theorem is outlined in Berry's book[32] and the general case for two-spaces is quoted by Weinberg.[33] With $g_{11} = R^2$, $g_{22} = R^2\sin^2\theta$ and $x^1 = \theta$, $x^2 = \phi$, it is straightforward to show that $\kappa = 1/R^2$, that is, the space is one of constant curvature, meaning that it is independent of θ and ϕ. In Section 19.3.2, we demonstrated that the only isotropic two-spaces correspond to the curvature κ being a constant which is positive, negative or zero. The value $\kappa = 0$ corresponds to flat, Euclidean space. If κ is negative, the two-space is hyperbolic, the radii of curvature being in opposite senses at all points of the two-space, as illustrated schematically in Fig. 19.8. For these spaces, trigonometric functions such as $\sin\theta$ and $\cos\theta$ are replaced by their hyperbolic counterparts, $\sinh\theta$ and $\cosh\theta$. In isotropic curved spaces, κ is a constant everywhere, but in the case of a general two-space, the curvature is a function of the spatial coordinates.

The extension to isotropic three-spaces is straightforward, but it is no longer possible to envisage the three-space geometrically since it would have to be embedded in a isotropic four-space. We can, however, proceed in a straightforward manner since a two-dimensional section through an isotropic three-space must itself be an isotropic two-space, for which we have already worked out the metric tensor.

Suppose we want to work out the angle $d\phi$ subtended by a rod of length L at polar angle θ, as illustrated in Fig. 19.7. In the notation of that figure,

$$L = R \sin\theta \, d\phi. \tag{19.41}$$

If x is the distance round the arc from P to L, that is, a 'geodesic distance' in the surface of the sphere, $\theta = x/R$ and hence

$$L = R \sin(x/R) \, d\phi. \tag{19.42}$$

By *geodesic distance*, we mean the shortest distance between two points in the two-space and, in the isotropic case, this is the great circle joining P and L.

To extend the argument to isotropic three-spaces, suppose we want to work out the solid angle subtended by a circular area perpendicular to the 'line of sight' from P in the isotropic three-space. We take a two-sphere section through the curved space perpendicular to the previous direction, as illustrated schematically in Fig. 19.9. Because of the isotropy of the space, the relation between L and $d\phi$ is the same in the perpendicular direction and hence the area of the circle of diameter L is

$$A = \frac{\pi L^2}{4} = \frac{\pi R^2}{4} \sin^2(x/R) \, d\phi^2 = R^2 \sin^2(x/R) \, d\Omega, \tag{19.43}$$

where $d\Omega$ is the solid angle subtended by the area dA in the isotropic three-space. Notice that, if $R \gg x$, (19.43) reduces to $A = x^2 \, d\Omega$, the Euclidean result.

This example illustrates an important feature of the idea of 'distance' in non-Euclidean geometries. If x is taken to be the geodesic distance, the formula for surface area involves

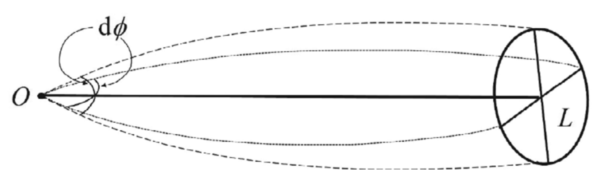


Fig. 19.9 Illustrating how surface areas can be measured in an isotropic three-space by taking orthogonal sections through the space, each of which is an isotropic two-space.

something more complex than the Euclidean formula. We can, however, make the formulae for areas look like the Euclidean relation by writing $A = r^2\,\mathrm{d}\Omega$, defining a new distance coordinate r. In this example, we would write $r = R\sin(x/R)$.

Now the formula for the spatial interval in a spherical two-space is

$$\mathrm{d}l^2 = \mathrm{d}x^2 + R^2\sin^2(x/R)\,\mathrm{d}\phi^2. \tag{19.44}$$

Let us rewrite (19.44) in terms of r rather than x. Differentiating $r = R\sin(x/R)$,

$$\mathrm{d}r = \cos(x/R)\,\mathrm{d}x, \tag{19.45}$$

and so

$$\mathrm{d}r^2 = \cos^2(x/R)\,\mathrm{d}x^2 = [1 - \sin^2(x/R)]\,\mathrm{d}x^2 = \left(1 - \frac{r^2}{R^2}\right)\mathrm{d}x^2, \tag{19.46}$$

that is,

$$\mathrm{d}x^2 = \frac{\mathrm{d}r^2}{1 - \kappa r^2}, \tag{19.47}$$

where $\kappa = 1/R^2$ is the curvature of the two-space. The metric can therefore be written

$$\mathrm{d}l^2 = \frac{\mathrm{d}r^2}{1 - \kappa r^2} + r^2\,\mathrm{d}\phi^2. \tag{19.48}$$

The metrics (19.44) and (19.48) are exactly equivalent but notice the different meanings of x and r: x is a *geodesic distance* in the three-space, whereas r is a distance coordinate which gives the correct answer for distances normal to the geodesic according to the relation $L = r\,\mathrm{d}\phi$. Notice how conveniently the metric (19.48) takes care of spherical, flat and hyperbolic spaces depending on the sign of κ.

19.4 The Route to General Relativity

The considerations of Sections 19.2 and 19.3 make it clear why general relativity is a theory of considerable technical complexity. Space-time has to be defined by a general curved metric tensor $g_{\mu\nu}$, which is a function of space-time coordinates. It is apparent from the calculations carried out in the preceding sections that the $g_{\mu\nu}$ are analogous to gravitational potentials, their variation from point to point in space-time defining the local curvature of space. From the purely geometrical point of view, we have to give mathematical substance to the principle of equivalence, that is, to devise transformations in Riemannian space which can relate the values of $g_{\mu\nu}$ at different points in space-time. From the dynamical point of view, we have to be able to relate the $g_{\mu\nu}$ to the distribution of matter in the Universe.

We are led to four-dimensional curved spaces which locally reduce to a Minkowski metric of the form (19.8). The appropriate choice of the metric, in general, is a *Riemannian metric*, which can be written

$$ds^2 = \sum_{\mu,\nu} g_{\mu\nu}\, dx^\mu\, dx^\nu = g_{\mu\nu}\, dx^\mu\, dx^\nu, \tag{19.49}$$

where x^μ and x^ν are coordinates defining points in the four-dimensional space and ds^2 is given by a homogeneous quadratic differential form in these coordinates. The convention of summing over identical indices is used in (19.49). The indices μ and ν in dx are written as superscripts for the technical reason that in the full theory of four-dimensional spaces we need to distinguish between covariant and contravariant quantities. We will not need this distinction in our approach. Locally, these spaces reduce to the Minkowski metric (19.8), for which

$$g_{\mu\nu} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{1}{c^2} & 0 & 0 \\ 0 & 0 & -\frac{1}{c^2} & 0 \\ 0 & 0 & 0 & -\frac{1}{c^2} \end{bmatrix}, \tag{19.50}$$

where $[x^0, x^1, x^2, x^3] = [t, x, y, z]$. The *signature* of the metric is $[+, -, -, -]$ and this property is preserved under general Riemannian transformations.

An important conceptual point is that the coordinates should be considered as 'labels' for the points in the space rather than 'distances', as is conventional in Euclidean space. It so happens that x, y and z are real distances in Euclidean space. In general, however, the coordinates may not have this intuitively obvious meaning.

19.4.1 Four-Tensors in Relativity

All the laws of physics, with the exception of gravity, can be written in Lorentz invariant form, that is, they are *form-invariant* under Lorentz transformations. The *four-vectors*, introduced in Chapter 18, are designed to be objects which are form invariant under Lorentz transformations. Just for this section, we will use the notation used by professional relativists according to which the velocity v is measured in units of the speed of light, so that we set $c = 1$. Then, the time coordinate $x^0 = t$ and the spatial components are $x = x^1$, $y = x^2$ and $z = x^3$. The transformation of a four-vector V^α between two inertial frames of reference in standard configuration is then written

$$V^\alpha \to V'^\alpha = \Lambda^\alpha_\beta V^\beta, \tag{19.51}$$

where the matrix Λ^α_β is the Lorentz transformation,

$$\Lambda^\alpha_\beta = \begin{bmatrix} \gamma & -\gamma v & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{19.52}$$

with $\gamma = (1 - v^2/c^2)^{-1/2}$. The convention of summing over identical indices is adopted in (19.51).

Many physical quantities are naturally described in terms of *tensors* rather than vectors. Therefore, the natural extension of the idea of four-vectors in relativity is to *four-tensors*, which are objects which transform according to the rule

$$T^{\alpha\gamma} \rightarrow T'^{\alpha\gamma} = \Lambda^{\alpha}_{\beta}\Lambda^{\gamma}_{\delta}T^{\beta\delta}. \tag{19.53}$$

As an example of how this works, consider what relativists call ‘dust’, by which they mean matter without any internal pressure. The *energy–momentum tensor* for dust is $T^{\alpha\beta} = \varrho_0 u^{\alpha}u^{\beta}$, where ϱ_0 is the proper mass density of the dust, meaning the density measured by an observer moving with the flow, what is called a *comoving observer*; u^{α} is the velocity four-vector. $T^{\alpha\beta}$ can therefore be written

$$T^{\alpha\beta} = \varrho_0 \begin{bmatrix} 1 & u_{x_1} & u_{x_2} & u_{x_3} \\ u_{x_1} & u_{x_1}^2 & u_{x_1}u_{x_2} & u_{x_1}u_{x_3} \\ u_{x_2} & u_{x_1}u_{x_2} & u_{x_2}^2 & u_{x_2}u_{x_3} \\ u_{x_3} & u_{x_1}u_{x_3} & u_{x_2}u_{x_3} & u_{x_3}^2 \end{bmatrix}. \tag{19.54}$$

The simplest case is the transformation of the T^{00} component of this four-tensor, $T'^{00} = \Lambda^0_{\beta}\Lambda^0_{\delta}T^{\beta\delta}$. Performing this double sum, the result is $\varrho' = \gamma^2\varrho_0$, which has a natural interpretation in special relativity. The observed density of the dust ϱ' is increased by two powers of the Lorentz factor γ over the proper value ϱ_0. One of these is associated with the formula for the momentum of the dust $p = \gamma m\boldsymbol{v}$ and the other with length contraction in the direction of motion of the flow, $l = l_0/\gamma$.

For a perfect fluid, the pressure p cannot be neglected and then the energy–momentum tensor becomes $T^{\alpha\beta} = (\varrho_0 + p)u^{\alpha}u^{\beta} - pg^{\alpha\beta}$, where $g^{\alpha\beta}$ is the metric tensor. It is a useful exercise to show that the equation

$$\partial_{\beta}T^{\alpha\beta} = 0 \tag{19.55}$$

expresses both the laws of conservation of momentum and energy in relativity, where ∂_{β} means partial differentiation of the tensor components with respect to β, the operator ∂_{β} taking the form

$$[\partial/\partial x_0, \partial/\partial x_1, \partial/\partial x_2, \partial/\partial x_3]. \tag{19.56}$$

These aspects of the energy–momentum tensor $T^{\mu\nu}$ are elegantly described by Rindler.[34]

Maxwell’s equations in a vacuum can be written in compact form in terms of the antisymmetric electromagnetic field tensor $F^{\alpha\beta}$:

$$F^{\alpha\beta} = \begin{bmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & -B_z & 0 & B_x \\ -E_z & B_y & -B_x & 0 \end{bmatrix}, \tag{19.57}$$

and the current density four-vector $j^{\alpha} = [\varrho_e, \boldsymbol{j}]$. This form of Maxwell’s equations is written in Heaviside–Lorentz units with $c = 1$. The equation of continuity becomes

$$\partial_{\alpha}j^{\alpha} = 0. \tag{19.58}$$

Maxwell's equations for the relations between electric and magnetic fields and their sources become

$$\partial_\beta F^{\alpha\beta} = j^\alpha. \tag{19.59}$$

Thus, four-tensors provide the natural language for expressing the laws of physics in a form which guarantees that they transform correctly according to the Lorentz transformations.

19.4.2 What Einstein Did

Einstein's objective was to find the relation between the $g_{\mu\nu}$ and the mass–energy distribution, that is, the analogue of Poisson's equation in Newtonian gravity, which involves second-order partial differential equations. For example, in the metric (19.22), we rationalised that the g_{00} component should have the form

$$g_{00} = \left(1 - \frac{2GM}{rc^2}\right) = \left(1 + \frac{2\phi}{c^2}\right). \tag{19.60}$$

Poisson's equation for gravity is

$$\nabla^2\phi = 4\pi G\varrho_0, \tag{19.61}$$

and hence, using (19.60) and the rest-frame value for $T^{00} = \rho_0$, we find

$$\nabla^2 g_{00} = \frac{8\pi G}{c^2}T^{00}. \tag{19.62}$$

This simple, and somewhat crude, calculation shows why it is reasonable to expect a close relation between the derivatives of $g_{\mu\nu}$ and the energy–momentum tensor $T^{\mu\nu}$. In fact, the constant in front of T^{00} is the same as in the full theory.

The tensor equivalent of Poisson's equation involves the differentiation of tensors and this is where the complications begin – partial differentiation of tensors does not generally yield other tensors. As a result, the operations corresponding to grad, div and curl are more complicated for tensors than for vectors.

What is needed is a tensor which involves the metric tensor $g_{\mu\nu}$ and its first and second derivatives and which is linear in its second derivatives. It turns out that there is a unique answer to this problem, the fourth-rank tensor known as the *Riemann–Christoffel tensor* $R^{\lambda}_{\mu\nu\kappa}$. Other tensors can be formed from this tensor by contraction, the most important of these being the *Ricci tensor*:

$$R_{\mu\kappa} = R^{\lambda}_{\mu\lambda\kappa}, \tag{19.63}$$

and the *curvature scalar*:

$$R = g^{\mu\kappa}R_{\mu\kappa}. \tag{19.64}$$

Einstein's stroke of genius was to relate these tensors to the energy–momentum tensor in the following way:

$$R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R = -\frac{8\pi G}{c^2}T_{\mu\nu}. \tag{19.65}$$

This is the key relation which shows how the components of the metric tensor $g_{\mu\nu}$ are related to the mass–energy distribution in the Universe.

We will go no further along this route, except to note that Einstein realised that he could add an additional term to the left-hand side of (19.65). This is the origin of the famous *cosmological constant* Λ and was introduced in order to construct a static closed model for the Universe. Equation (19.65) then becomes

$$R_{\mu\nu} - \tfrac{1}{2} g_{\mu\nu} R + \Lambda g_{\mu\nu} = -\frac{8\pi G}{c^2} T_{\mu\nu}. \tag{19.66}$$

The cosmological constant will reappear in Chapters 20 and 21.

We still need a rule which tells us how to find the path of a particle or light ray in space-time from the metric (19.49). In Euclidean three-space, the answer is that we seek that path which minimises the distance ds between the points A and B. We need to find the corresponding result for Riemannian space-time. If we join the points A and B by a large number of possible routes through space-time, the path corresponding to free-fall between A and B must be the shortest path. From consideration of the twin paradox in general relativity, the free-fall path between A and B must correspond to the maximum proper time between A and B, that is, we require $\int_A^B \mathrm{d}s$ to be a *maximum* for the shortest path. In terms of the calculus of variations, this can be written

$$\delta \int_A^B \mathrm{d}s = 0. \tag{19.67}$$

For our 'Newtonian' metric (19.22), this condition is exactly equivalent to Hamilton's principle in mechanics and dynamics. The 'action integral' derived from (19.22) is

$$\begin{aligned} \int_A^B \mathrm{d}s &= \int_{t_1}^{t_2} \frac{\mathrm{d}s}{\mathrm{d}t}\,\mathrm{d}t \\ &= \int_{t_1}^{t_2} \left[\left(1 + \frac{2\phi}{c^2}\right) - \frac{v^2}{c^2}\right]^{1/2} \mathrm{d}t, \end{aligned} \tag{19.68}$$

since $\mathrm{d}l/\mathrm{d}t = v$. For weak fields, $(2\phi/c^2) - (v^2/c^2) \ll 1$, (19.68) reduces to

$$\int_A^B \mathrm{d}s = \int_{t_1}^{t_2} (U - T)\,\mathrm{d}t, \tag{19.69}$$

where U is the potential energy of the particle and T its kinetic energy. Thus, the constraint that ds must be a maximum corresponds to Hamilton's principle in dynamics (see Section 8.3).

19.5 The Schwarzschild Metric

The solution of Einstein's field equations for the case of a point mass was discovered by Karl Schwarzschild[35] in 1916, only two months after the publication of Einstein's theory in its definitive form. Schwarzschild was a brilliant astronomer and Director of the Potsdam Astrophysical Observatory, who had volunteered for military service in 1914 at the beginning of the First World War. In 1916, while serving on the Russian front, he wrote two papers on his exact solution of Einstein's field equations for a stationary point mass. Tragically, he contracted a rare skin disease at the front and died in May 1916 after a short illness. The *Schwarzschild metric* for a point mass M has the form

$$\mathrm{d}s^2 = \left(1 - \frac{2GM}{rc^2}\right)\mathrm{d}t^2 - \frac{1}{c^2}\left[\frac{\mathrm{d}r^2}{\left(1 - \dfrac{2GM}{rc^2}\right)} + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right)\right]. \tag{19.70}$$

From the analyses of Sections 19.2 and 19.3, we can understand the meanings of the various terms in the metric. We recall that (19.70) reduces to the local Minkowski metric at any point in space.

(a) The time coordinate is written in a form such that clocks can be synchronised everywhere, that is, the interval of proper time for a stationary observer at radial distance r is

$$\mathrm{d}\tau = \left(1 - \frac{2GM}{rc^2}\right)^{1/2}\mathrm{d}t, \tag{19.71}$$

whereas the corresponding interval of coordinate time measured at infinity is $\mathrm{d}t$. Thus, $\mathrm{d}t$ keeps track of time at infinity and proper time at any point in the field is found from (19.71). Note that, whereas the factor $(1 - 2GM/rc^2)^{1/2}$ in the metric (19.22) was only approximate, (19.71) is *exact* in general relativity for all values of the parameter $2GM/rc^2$.

(b) The angular coordinates have been written in terms of spherical polar coordinates centred on the point mass. The radial coordinate r is such that proper distances $\mathrm{d}l_\perp$ perpendicular to the radial coordinate are given by

$$\mathrm{d}l_\perp^2 = r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2). \tag{19.72}$$

Hence, the distance measure r can be thought of as an *angular diameter distance* and is distinct from the *geodesic distance* x from the mass to the point labelled r. The increment of proper, or geodesic, distance $\mathrm{d}x$ in the radial direction is the first component of the spatial part of the metric (19.70) and so

$$x = \int_0^r \mathrm{d}x = \int_0^r \frac{\mathrm{d}r}{\left(1 - \dfrac{2GM}{rc^2}\right)^{1/2}}. \tag{19.73}$$

The combination of (19.71) and (19.73) enables us to derive the correct version of Einstein's argument concerning the apparent variability of the speed of light according to an observer at infinity. The local speed of light is $c = \mathrm{d}x/\mathrm{d}\tau$ and hence, for a distant observer, the speed of light in terms of coordinate time and the distance measure r is

$$c(r) = \frac{\mathrm{d}r}{\mathrm{d}t} = c\left(1 - \frac{2GM}{rc^2}\right). \tag{19.74}$$

This is a convenient expression for working out time delays due to time dilation in the gravitational field of a point mass.

(c) The metric describes how the gravity of the point mass bends space-time. The effects of time dilation were described in (a) and, in addition, the space sections are curved. For isotropic curved spaces, we showed that the spatial component of the metric could be written

$$\frac{\mathrm{d}r^2}{(1-\kappa r^2)} + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right). \tag{19.75}$$

Comparing (19.70) with (19.75), it can be seen that the local space curvature is found by equating κr^2 and $2GM/rc^2$,

$$\kappa = \frac{2GM}{r^3 c^2}. \tag{19.76}$$

This provides a measure of the curvature of space produced by a point mass M at distance coordinate r. Notice that the space is isotropic at all points, but the curvature changes as r^{-3}. The expression (19.76) has the attractive features that κ tends to zero as r tends to infinity, and that κ is proportional to M. Writing the curvature κ as $1/R^2$,

$$\frac{r^2}{R^2} = \frac{2GM}{rc^2}. \tag{19.77}$$

This is a measure of the curvature of space relative to the distance from the point mass. Notice that $2GM/rc^2$ is the 'relativistic factor' appearing in the time and radial components of the metric. Some typical values of this parameter are:

At the surface of the Earth,	$2GM_{\mathrm{E}}/rc^2 = 1.4\times10^{-8}$,
At the surface of the Sun,	$2GM_{\odot}/rc^2 = 4\times10^{-6}$,
At the surface of a neutron star,	$2GM_{\mathrm{ns}}/rc^2 \approx 0.3$.

Thus, within the Solar System, the effects of space curvature are very small and require measurements of extreme precision in order to be detectable at all. The first example indicates the magnitude of the factor by which the sum of the angles of a triangle differ from 180° in measurements made at the surface of the Earth, indicating the enormous challenge which faced Gauss in trying to measure the curvature of space.

(d) Note that something 'funny' must happen at the radial coordinate $2GM/rc^2 = 1$. The radius defined by this relation, $r_{\mathrm{g}} = 2GM/c^2$, is known as the *Schwarzschild radius* of a stationary black hole and plays a particularly important role in their study, which will be discussed later in this chapter.

19.6 Particle Orbits about a Point Mass

Let us now analyse the motion of a test particle about a point mass according to the Schwarzschild metric. This is the simplest solution of Einstein's field equations – the point mass has no angular momentum and is uncharged. To simplify the analysis, we consider orbits in the equatorial plane $\theta = \pi/2$. There is no loss of generality in making this simplification, since we can always rotate the coordinate system so that $\mathrm{d}\theta = 0$. The Schwarzschild metric can therefore be written

$$\mathrm{d}s^2 = \alpha\,\mathrm{d}t^2 - \frac{1}{c^2}(\alpha^{-1}\,\mathrm{d}r^2 + r^2\,\mathrm{d}\phi^2). \tag{19.78}$$

In this notation, ϕ is the angle in the equatorial plane and $\alpha = (1 - 2GM/rc^2)$. We now form the 'action integral':

$$s_{AB} = \int_A^B \mathrm{d}s = \int_A^B G(x^\mu, \dot{x}^\mu)\,\mathrm{d}s, \tag{19.79}$$

where the function $G(x^\mu, \dot{x}^\mu)$ is found by dividing (19.78) by $\mathrm{d}s^2$:

$$G(x^\mu, \dot{x}^\mu) = \left[\alpha\dot{t}^2 - \frac{1}{c^2}\alpha^{-1}\dot{r}^2 - \frac{r^2}{c^2}\dot{\phi}^2\right]^{1/2}. \tag{19.80}$$

This is the function to be inserted into the Euler–Lagrange equation (8.24) for the independent coordinates $x^\mu = t, r, \phi$.

Notice what we are doing in following this procedure. The dots refer to differentiation with respect to $\mathrm{d}s$, which is identical to $\mathrm{d}\tau$ in the case of the motion of particles. Therefore, in the analysis which follows, the dots refer to differentiation with respect to proper time. Note also that, by inspection, $G(x^\mu, \dot{x}^\mu) = 1$. The important point is that, when we insert $G(x^\mu, \dot{x}^\mu)$ into the Euler–Lagrange equation, it is only the functional dependence of G upon x^μ and $\dot{x}^\mu$ which is important. Exactly the same procedure occurs in Hamiltonian mechanics in which the Hamiltonian is a constant, but it is the dependence upon the coordinates which determines the dynamics. To reinforce this point, consider the case of flat space in which $G(x^\mu, \dot{x}^\mu)$ has the form

$$G(x^\mu, \dot{x}^\mu) = \left[\dot{t}^2 - \frac{1}{c^2}\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right)\right]^{1/2}, \tag{19.81}$$

where again the dots mean differentiation with respect to proper time. It is apparent that (19.81) is the magnitude of the velocity four-vector

$$\boldsymbol{U} = \left[\dot{t}, \frac{\dot{x}}{c}, \frac{\dot{y}}{c}, \frac{\dot{z}}{c}\right] = \left[\gamma, \gamma\frac{\boldsymbol{u}}{c}\right], \tag{19.82}$$

where $\boldsymbol{u}$ is the three-velocity and so $|G| = 1$.

We notice that the expression (19.80) for G does not depend upon either t or ϕ and so the Euler–Lagrange equation,

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial G}{\partial \dot{x}^\mu}\right) - \frac{\partial G}{\partial x^\mu} = 0, \tag{19.83}$$

is greatly simplified for these coordinates. Considering first the t coordinate,

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial G}{\partial \dot{t}}\right) = 0, \quad \frac{\partial G}{\partial \dot{t}} = \text{constant}. \tag{19.84}$$

Taking the partial derivative of (19.80) with respect to $\dot{t}$ and recalling that $G = 1$,

$$\alpha\dot{t} = k, \tag{19.85}$$

where k is a constant. This is a very important result. We have derived an expression for a constant of the motion and we guess that it is equivalent to the *conservation of energy* in the gravitational field. What the variational procedure has done is to find a property of the motion which is stationary with respect to proper time – in classical mechanics, this is the total energy for a conservative field of force. We will return to the interpretation of (19.85) in a moment.

We now carry out the same analysis for the ϕ-coordinate. Then,

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial G}{\partial \dot{\phi}}\right) = 0, \qquad \frac{\partial G}{\partial \dot{\phi}} = \text{constant}. \tag{19.86}$$

Taking the partial derivative of (19.80) with respect to $\dot{\phi}$ and recalling that $G = 1$,

$$r^2\dot{\phi} = h, \tag{19.87}$$

where h is a constant. This is another important result. This time, the variational procedure has found a constant of the motion for variations in the ϕ-direction. The Newtonian analogue of this result is the *conservation of angular momentum*, where h is the *specific angular momentum*, that is, the angular momentum per unit mass.

The radial equation is somewhat more complicated. It is simplest to substitute for $\dot{t}$ and $\dot{\phi}$ into (19.80) in two stages. First substituting for $\dot{t}$ into expression (19.80),

$$\begin{aligned} G^2 &= \alpha\dot{t}^2 - \frac{1}{c^2}\alpha^{-1}\dot{r}^2 - \frac{r^2}{c^2}\dot{\phi}^2 \\ &= \frac{k^2}{\alpha} - \frac{1}{\alpha c^2}\dot{r}^2 - \frac{r^2}{c^2}\dot{\phi}^2 = 1. \end{aligned} \tag{19.88}$$

The expression (19.88) can now be reorganised as follows:

$$\dot{r}^2 + \alpha(r\dot{\phi})^2 - \frac{2GM}{r} = c^2(k^2 - 1). \tag{19.89}$$

Finally, multiplying through by the mass of a test particle m,

$$\tfrac{1}{2}m\dot{r}^2 + \tfrac{1}{2}\alpha m(r\dot{\phi})^2 - \frac{GMm}{r} = \frac{mc^2}{2}(k^2 - 1). \tag{19.90}$$

This calculation shows that the constant k is related to conservation of energy. All the terms on the left-hand side look superficially like those appearing in the Newtonian expression for the conservation of energy in bound orbits, but there are *important differences*:

- The three terms on the left-hand side look formally the same as the Newtonian expression for conservation of energy, except for the inclusion of the multiplying factor α in front of the term $\frac{1}{2}m(r\dot{\phi})^2$, which corresponds to motion in the circumferential direction.
- The dots refer to differentiation with respect to *proper time* $\mathrm{d}\tau$, not to coordinate time.
- The r coordinate is an angular diameter distance rather than a geodesic distance, but we are now familiar with the care needed in the use of distance measures in curved space-time.

19.6.1 Energy in General Relativity

The definition of energy in general relativity is discussed by Shapiro and Teukolsky.[36] The simplest approach is to adopt the Lagrangian formulation of mechanics in which the total energy is defined as the momentum conjugate to the time coordinate through the relation

$$E = \frac{\partial \mathcal{L}}{\partial \dot{t}}, \tag{19.91}$$

(see Section 8.4.1). Shapiro and Teukolsky show that $\mathcal{L} \propto G^2$ and so the definition of E can be found from (19.80),

$$E \propto \alpha \dot{t} = k. \tag{19.92}$$

In fact, the reasoning which led to (19.85) performed precisely the same operation. To find the constant relating E to k, it is simplest to consider the case of a stationary particle at $r = \infty$, in which case (19.90) shows that $k = 1$. The total energy is the rest mass energy mc^2 and hence, in general, the total energy must take the form

$$E = kmc^2 \qquad \text{or} \qquad k = \frac{E}{mc^2}. \tag{19.93}$$

Notice that, by taking the partial derivative with respect to $\dot{t}$, we find the energy according to coordinate time t, that is, the energy available at infinity, rather than at a point in the field.

Let us illustrate this in another way. The four-vector velocity of the point mass m can be written down from (19.80), by analogy with the flat space case (19.82). The four-momentum is

$$\boldsymbol{P} = \left[mc\alpha^{1/2}\dot{t},\ m\alpha^{-1/2}\dot{r},\ 0,\ mr\dot{\phi}\right]. \tag{19.94}$$

for the case of motion in the (r, ϕ) plane. The second and fourth terms correspond to the radial and tangential components of the particle's momentum, and the first to its total energy at that point in the field. Therefore, we can write that the *local* energy of the particle in the field is

$$E_{\text{local}} = mc^2\alpha^{1/2}\dot{t} = k\alpha^{-1/2}mc^2. \tag{19.95}$$

The energy available at infinity is, however, subject to the same time-dilation factor as the time component of the metric, since it is the 'time' component of a parallel four-vector. Therefore, using (19.92), the available energy according to the observer at infinity is

$$E = \alpha^{1/2}E_{\text{local}} = \alpha^{1/2}k\alpha^{-1/2}mc^2 = kmc^2. \tag{19.96}$$

(see also Shapiro and Teukolsky). This result will prove to be important in working out the maximum amount of energy which can be extracted from matter which forms an accretion disc about a black hole.

19.6.2 Particle Orbits in Newtonian Gravity and General Relativity

Let us now contrast the orbits of particles about a point mass M according to Newtonian gravity and general relativity. Substituting for $\dot{\phi}$ in (19.90) using (19.87),

$$m\dot{r}^2 + \frac{\alpha m h^2}{r^2} - \frac{2GMm}{r} = mc^2(k^2 - 1),$$

$$m\dot{r}^2 + \frac{mh^2}{r^2} - \frac{2GMm}{r} - \frac{2GMmh^2}{r^3 c^2} = mc^2(k^2 - 1). \tag{19.97}$$

This substitution incorporates the law of conservation of angular momentum into the law of conservation of energy in Newtonian gravity. For Newtonian gravity, the equivalent expression is

$$m\dot{r}^2 + \frac{mh^2}{r^2} - \frac{2GMm}{r} = m\dot{r}_\infty, \tag{19.98}$$

where $\dot{r}_\infty$ is the radial velocity of the test particle at infinity. If the particle does not reach infinity, $\dot{r}_\infty$ is imaginary.

Although (19.97) and (19.98) look similar, the meanings of the coordinates are quite different, as noted in Section 19.6. The most important difference is the presence of the term $-2GMmh^2/r^3c^2$ on the left-hand side of (19.97). The origin of this term can be traced to the α-factor in front of the term representing rotational motion about the point mass.

This term has profound implications for the dynamics of particles about the central point mass. The radial component of the velocity of a test particle in Newtonian theory can be represented in terms of the variation of the different terms in (19.98) with radius. Rewriting (19.98) for unit mass in terms of the Schwarzschild radius r_g and a dimensionless specific angular momentum $\eta = h^2/r_\mathrm{g}^2 c^2$,

$$\text{for the Newton case} \qquad \dot{r}^2 = -c^2\left[\frac{\eta}{(r/r_\mathrm{g})^2} - \frac{1}{(r/r_\mathrm{g})}\right] + \dot{r}_\infty^2. \tag{19.99}$$

The term in large square brackets acts as a potential Φ such that

$$\Phi = \frac{\eta}{(r/r_\mathrm{g})^2} - \frac{1}{(r/r_\mathrm{g})}, \tag{19.100}$$

where the first term is a 'centrifugal' potential and the second the Newtonian gravitational potential. The variation of $-\Phi$ with r/r_g is shown in Fig. 19.10, in which the contributions of the two terms are displayed separately.

Then:

- If the particle has zero radial velocity at infinity, that is, it starts from A', the radial velocity at different radii can be found from the bold curve. When the particle reaches the radial coordinate at A, the radial component of its velocity is zero and this is the closest the particle can approach to $r = 0$: $\dot{r}$ becomes imaginary for smaller values of r. At A, all the kinetic energy of the particle is in its tangential motion.
- Bound orbits are found if $\dot{r}_\infty^2$ is negative. Then, the particle moves in an elliptical orbit between B and B'.

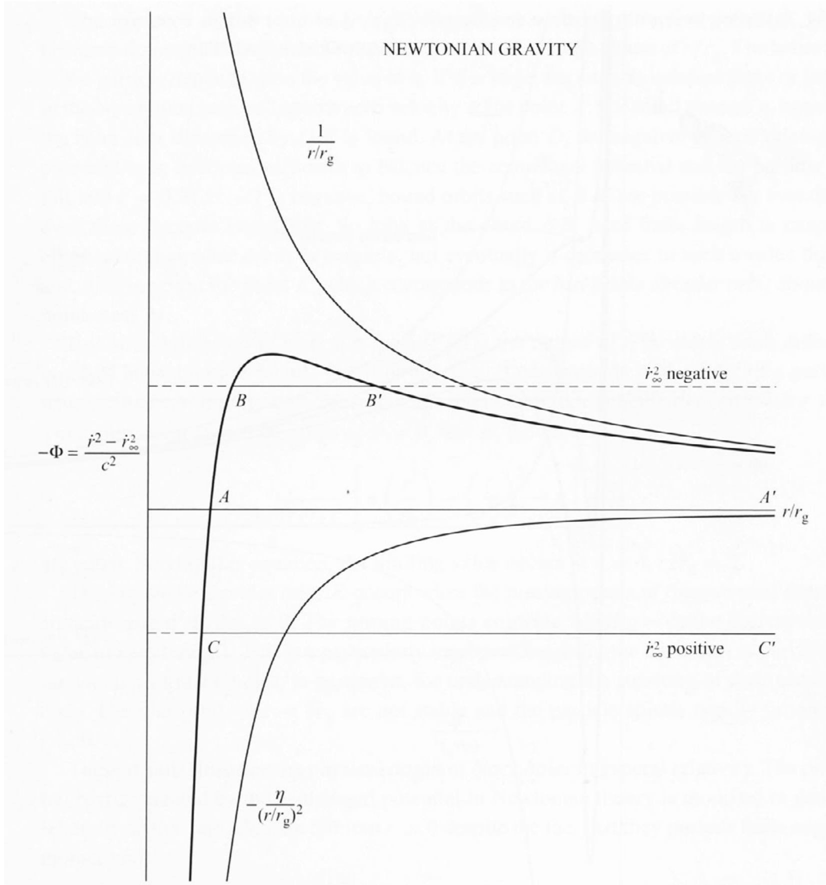


Fig. 19.10 The gravitational potential $-\Phi$ in Newtonian gravity, in which there are two contributions to the potential, plotted against r/r_g.

- Hyperbolic orbits of unbound particles correspond to positive values of $\dot{r}^2_\infty$ and are represented by loci such as CC'. C is the distance of closest approach and is found from the solution of (19.100) at which $\dot{r} = 0$. Notice that the particle can only reach $r = 0$ if $\eta = 0$. This is an important result in that, even if η is tiny, it is always sufficient to prevent the test mass reaching the origin, according to classical dynamics.

In the corresponding analysis for general relativity, the expression for $\dot{r}^2$ becomes,

$$\text{for the GR case,} \qquad \dot{r}^2 - c^2(k^2 - 1) = -c^2\left[\frac{\eta}{(r/r_g)^2} - \frac{1}{(r/r_g)} - \frac{\eta}{(r/r_g)^3}\right], \qquad (19.101)$$

so that now

$$\Phi = \frac{\eta}{(r/r_g)^2} - \frac{1}{(r/r_g)} - \frac{\eta}{(r/r_g)^3}. \qquad (19.102)$$

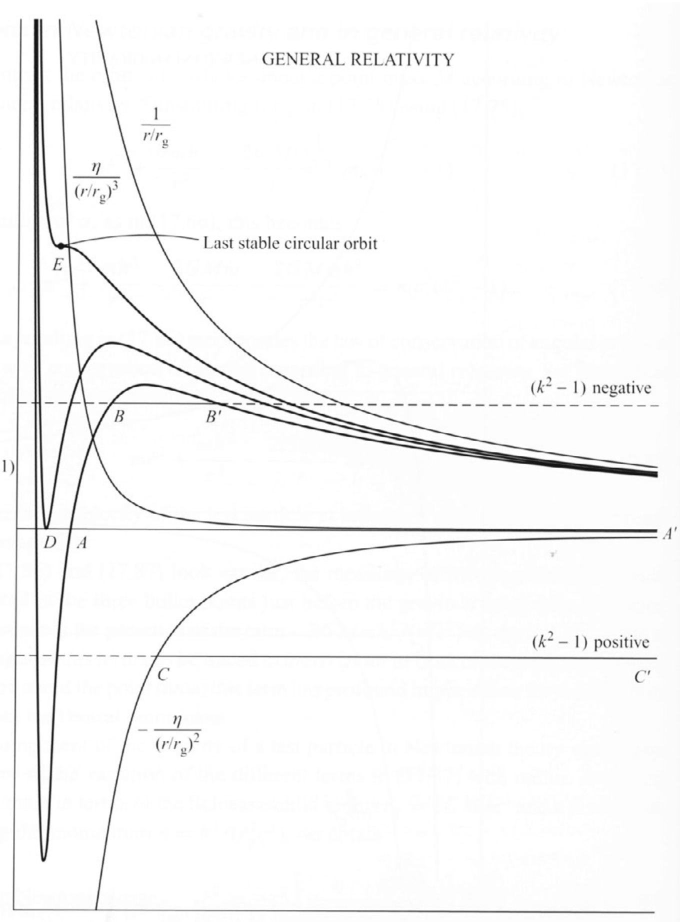


Fig. 19.11 The variation of the potential $-\Phi$ according to general relativity. There is an attractive potential associated with the term $\eta/(r/r_g)^3$, in addition to the two found in Newtonian gravity. This term dominates the dynamics close to the origin.

Much of the analysis is similar to the Newtonian case, but now we have to include the specifically general relativistic term $\eta/(r/r_g)^3$, which acts as a negative potential. The key point is that its magnitude increases even more rapidly with decreasing r than does the centrifugal term. The corresponding energy diagram for general relativity is shown in Fig. 19.11.

The presence of the term in $(r/r_g)^{-3}$ introduces a strong attractive potential which becomes dominant if the particle can penetrate to small enough values of r/r_g. The behaviour of the particle depends upon the value of η. If η is large, the particle behaves more or less as in the Newtonian case and attains zero velocity at the point A. For small enough η, however, the behaviour illustrated by DA' in Fig. 19.11 is found. At the point D, the negative general relativistic potential term becomes sufficient to balance the centrifugal potential and the particle can fall into $r = 0$. If $c^2(k^2 - 1)$ is negative, bound orbits such as BB' are possible but eventually even these become impossible. So long as the chord BB' is of finite length, a range of elliptical and circular orbits is possible, but eventually η decreases to such a value that B and B' coincide at the point E, which corresponds to the *last stable circular orbit* about the point mass M.

It is straightforward to work out the ranges of η in which these different types of behaviour are found. The limiting case D corresponds to that in which any particle which falls in from infinity can reach $r = 0$. This condition occurs when the roots of Φ are equal to zero:

$$\frac{1}{(r/r_g)^3}\left[\eta\left(\frac{r}{r_g}\right) - \left(\frac{r}{r_g}\right)^2 - \eta\right] = 0. \tag{19.103}$$

Solving this equation, the limiting value occurs at $\eta = 4$, $r/r_g = 2$.

The last stable circular orbit E occurs when the turning points of the potential function coincide and $\mathrm{d}^2\Phi/\mathrm{d}r^2 = 0$. The roots coincide when $\eta = 3$, the corresponding value of r being $3r_g$. This is a particularly important result for the dynamics of particles in the vicinity of black holes, in particular for understanding the structure of their accretion discs. Circular orbits at $r < 3r_g$ are not stable and the particle spirals rapidly in to $r = 0$.

These results illustrate the physical origin of *black holes* in general relativity. The potential barrier created by the centrifugal potential in Newtonian theory is modified in general relativity so that particles can fall into $r = 0$, despite the fact that they possess finite angular momentum.

19.7 Advance of the Perihelia of Planetary Orbits

The detailed analysis of the orbit of a test mass m proceeds from the energy equation (19.97),

$$m\dot{r}^2 + \frac{mh^2}{r^2} - \frac{2GMm}{r} - \frac{2GMmh^2}{r^3c^2} = mc^2(k^2 - 1), \tag{19.104}$$

where $\dot{r}$ means $\mathrm{d}r/\mathrm{d}\tau$ and the dependence upon τ can be eliminated by replacing it with the dependence upon ϕ, thus obtaining an expression for the orbit in the polar coordinates r and ϕ:

$$\dot{r} = \frac{\mathrm{d}r}{\mathrm{d}\tau} = \frac{\mathrm{d}r}{\mathrm{d}\phi}\frac{\mathrm{d}\phi}{\mathrm{d}\tau} = \frac{\mathrm{d}r}{\mathrm{d}\phi}\dot{\phi} = \frac{h}{r^2}\frac{\mathrm{d}r}{\mathrm{d}\phi}, \tag{19.105}$$

using (19.87). Therefore,

$$\left(\frac{h}{r^2}\frac{\mathrm{d}r}{\mathrm{d}\phi}\right)^2 + \frac{h^2}{r^2} - \frac{2GM}{r} - \frac{2GMh^2}{r^3c^2} = c^2(k^2 - 1). \tag{19.106}$$

There is a standard substitution $u = 1/r$, which simplifies the analysis. Then, $\mathrm{d}r = -(1/u^2)\,\mathrm{d}u$ and, dividing by h^2,

$$\left(\frac{\mathrm{d}u}{\mathrm{d}\phi}\right)^2 + u^2 - \frac{2GM}{h^2}u - \frac{2GM}{c^2}u^3 = \frac{c^2}{h^2}(k^2 - 1). \tag{19.107}$$

Now differentiating with respect to ϕ,

$$\frac{\mathrm{d}^2u}{\mathrm{d}\phi^2} + u = \frac{GM}{h^2} + \frac{3GM}{c^2}u^2. \tag{19.108}$$

It is illuminating to compare this equation with the corresponding Newtonian result. The Newtonian law of conservation of energy is

$$\tfrac{1}{2}mv^2 - \frac{GMm}{r} = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\phi}^2) - \frac{GMm}{r} = \text{constant}. \tag{19.109}$$

Because of conservation of angular momentum, $v_\phi r = \dot{\phi}r^2 = h = \text{constant}$, where h is the specific angular momentum, and hence

$$m\dot{r}^2 + \frac{mh^2}{r^2} - \frac{2GMm}{r} = \text{constant}. \tag{19.110}$$

Performing the same types of substitution involved in deriving (19.108), namely, dividing through by m, setting $u = 1/r$ and then converting the time derivative into a derivative with respect to ϕ by writing

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \frac{\mathrm{d}r}{\mathrm{d}\phi}\frac{\mathrm{d}\phi}{\mathrm{d}t} = \frac{h}{r^2}\frac{\mathrm{d}r}{\mathrm{d}\phi}, \tag{19.111}$$

we find the result

$$\left(\frac{\mathrm{d}u}{\mathrm{d}\phi}\right)^2 + u^2 - \frac{2GM}{h^2}u = \text{constant}. \tag{19.112}$$

Differentiating with respect to ϕ, we find

$$\frac{\mathrm{d}^2u}{\mathrm{d}\phi^2} + u = \frac{GM}{h^2}. \tag{19.113}$$

The expressions (19.108) and (19.113) are similar, the important difference being the presence of the term $(3GM/c^2)u^2$ in (19.108). We already know the solutions of (19.113) – for bound orbits, these are *ellipses*. Notice that, in the limit $c \to \infty$, the term $(3GM/c^2)u^2$ in (19.108) vanishes, enabling us to recover the Newtonian expression (19.113). This term is a very small correction to the Newtonian solution within the Solar System. To see this, we can compare the magnitudes of the two 'gravitational' terms on the right-hand side of (19.108) for a circular orbit, $u = \text{constant}$,

$$\frac{(3GM/c^2)u^2}{(GM/h^2)} = \frac{3u^2h^2}{c^2} = 3\frac{v^2}{c^2}. \tag{19.114}$$

Thus, the correction to the orbit is second order in v/c. Since the speed of the Earth about the Sun is 30 km s^{-1}, the correction amounts to only about three parts in 10^8.

Although the effect is very small, it results in measurable changes in the orbits of the planets. The largest effect within the Solar System is found for Mercury's orbit about the Sun, which has ellipticity $\epsilon = 0.2$. The position of closest approach of a planet to the Sun is known as its *perihelion* and the effect of the small general relativistic correction is to cause the perihelion of Mercury's orbit to advance by a small amount each orbit. Let us determine the magnitude of the advance for a *circular orbit*.

For a circular orbit, u = constant, $\mathrm{d}^2u/\mathrm{d}\phi^2 = 0$ and so, in the Newtonian approximation,

$$u = \frac{GM}{h^2}. \tag{19.115}$$

We can insert this solution into equation (19.108) and solve for a small perturbation $g(\phi)$ to the orbital phase,

$$u = \frac{GM}{h^2} + g(\phi). \tag{19.116}$$

Substituting into (19.108) and preserving terms to first order in $g(\phi)$,

$$\frac{\mathrm{d}^2 g}{\mathrm{d}\phi^2} + g\left[1 - \left(\frac{3GM}{c^2}\right)\left(\frac{2GM}{h^2}\right)\right] = \frac{3GM}{c^2}\left(\frac{GM}{h^2}\right)^2. \tag{19.117}$$

This is a harmonic equation for g and, in the non-relativistic limit in which $(3GM/c^2) \to 0$, becomes

$$\frac{\mathrm{d}^2 g}{\mathrm{d}\phi^2} + g = 0. \tag{19.118}$$

The harmonic solutions have period 2π, in other words, a perfect circular orbit.

Because of the relativistic perturbation term, the phase of the planet in its orbit is slightly different from 2π per revolution. From the expression (19.117),

$$\omega^2 = \left[1 - \left(\frac{3GM}{c^2}\right)\left(\frac{2GM}{h^2}\right)\right], \tag{19.119}$$

and the period T of the orbit is

$$T = \frac{2\pi}{\left[1 - \left(\frac{3GM}{c^2}\right)\left(\frac{2GM}{h^2}\right)\right]^{1/2}} \approx 2\pi\left(1 + \frac{3G^2M^2}{c^2h^2}\right). \tag{19.120}$$

Thus, the orbit closes up slightly later each orbit. The fractional change of phase per orbit is

$$\frac{\mathrm{d}\phi}{2\pi} = \frac{3G^2M^2}{c^2h^2} = \frac{3}{4}\left(\frac{2GM}{hc}\right)^2 = \frac{3}{4}\left(\frac{c}{v}\right)^2\left(\frac{r_\mathrm{g}}{r}\right)^2, \tag{19.121}$$

since $h = rv$. For elliptical orbits, the exact answer is

$$\frac{\mathrm{d}\phi}{2\pi} = \frac{3}{4}\left(\frac{c}{v}\right)^2\left(\frac{r_\mathrm{g}}{r}\right)^2\frac{1}{(1-\epsilon^2)}. \tag{19.122}$$

For Mercury, $r = 5.8 \times 10^{10}$ m, $T = 88$ days, and $\epsilon = 0.2$, and for the Sun $r_\mathrm{g} = 3$ km. The advance of the perihelion of Mercury is therefore predicted to amount to 43 arcsec per century. This was the extraordinary resolution of an unsolved problem in nineteenth century celestial mechanics. Once the effects of the perturbations of the other planets upon Mercury's orbit about the Sun had been taken into account, Le Verrier found in 1859 that there remained an unaccounted advance of its perihelion of exactly this value. In his paper on general relativity of November 1915, Einstein triumphantly showed that this was precisely accounted for by the general theory of relativity.

The effect is very small within the Solar System, but it is very much larger in the case of compact close binary star systems. The most important example is the binary pulsar PSR 1913+16, which consists of two neutron stars orbiting their common centre of mass. The orbital period is only 7.75 hours, and the ellipticity of the orbits of the neutron stars is $\epsilon = 0.617$. In this case, the rate of advance of the perihelia of their elliptical orbits provides important information about the masses of the neutron stars, as well as providing very sensitive tests of general relativity itself (See Section 19.11).

19.8 Light Rays in Schwarzschild Space-Time

Einstein realised as early as 1907 that light rays are deflected by the presence of matter and in 1911 urged astronomers to attempt to measure the deflection of light rays by the Sun, four years before the general theory was presented in its definitive form. The predicted deflection can be determined from the Schwarzschild metric of a point mass. Light rays propagate along null geodesics $\mathrm{d}s = \mathrm{d}\tau = 0$. This has consequences for the energy and angular momentum conservation relations, (19.85) and (19.87) respectively, which can be written

$$\alpha\dot{t} = \left(1 - \frac{2GM}{rc^2}\right)\frac{\mathrm{d}t}{\mathrm{d}s} = k, \quad r^2\dot{\phi} = r^2\frac{\mathrm{d}\phi}{\mathrm{d}s} = h. \tag{19.123}$$

Since $\mathrm{d}s = \mathrm{d}\tau = 0$, both k and h are infinite, but their ratio remains finite. The propagation equation for light rays can be found by inserting the value $h = \infty$ into (19.108),

$$\frac{\mathrm{d}^2u}{\mathrm{d}\phi^2} + u = \frac{3GM}{c^2}u^2. \tag{19.124}$$

The term on the right-hand side represents the influence of curved space-time upon the propagation of the light ray. Notice that there is an interesting limiting case in which photons propagate in a circular orbit about a black hole. In this case, the first term on the left-hand side of (19.124) is zero, and the radius of the photon orbit is

$$r = \frac{3GM}{c^2} = \frac{3}{2}r_{\mathrm{g}}. \tag{19.125}$$

Returning to the much less extreme case of the deflection of light rays by the Sun, let us first work out the path of the light ray in the absence of the term $(3GM/c^2)u^2$. The propagation equation becomes

$$\frac{\mathrm{d}^2u}{\mathrm{d}\phi^2} + u = 0. \tag{19.126}$$

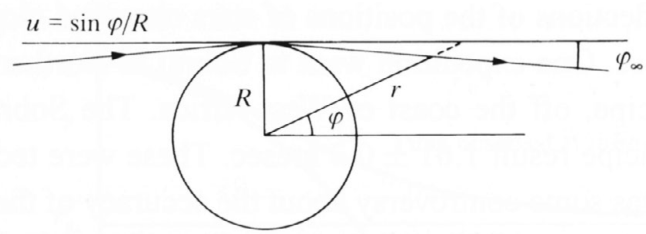


Fig. 19.12 The coordinate system used to work out the deflection of light rays by the Sun.

A suitable solution of this equation is $u_0 = \sin\phi/R$ where R is the distance of closest approach of the light ray, as illustrated in Fig. 19.12. This solution corresponds to a straight line, tangent to the (empty) sphere at radius R – if there is no mass, there is no deviation of the light ray.

To find the path of the light ray in the next order of approximation, we seek solutions of (19.124) with $u = u_0 + u_1$. Then, we find

$$\frac{\mathrm{d}^2 u_1}{\mathrm{d}\phi^2} + u_1 = \frac{3GM}{c^2R^2}\sin^2\phi. \tag{19.127}$$

By inspection, a suitable trial solution is $u_1 = A + B\cos 2\phi$. Differentiating and equating coefficients results in the following solution for u:

$$u = u_0 + u_1 = \frac{\sin\phi}{R} + \frac{3}{2}\frac{GM}{c^2R^2}\left(1 + \frac{1}{3}\cos 2\phi\right). \tag{19.128}$$

Since the angle ϕ is very small, we need only study the asymptotic solutions at large values of r and so $\sin\phi \approx \phi$ and $\cos 2\phi \approx 1$. In the limit $r \to \infty,\ u \to 0$, we find

$$u = \frac{\phi_\infty}{R} + \frac{3}{2}\frac{GM}{c^2R^2}\left(1 + \frac{1}{3}\right) = 0,$$

$$\phi_\infty = -\frac{2GM}{Rc^2}. \tag{19.129}$$

From the geometry of Fig. 19.12, the total deflection is twice ϕ_∞,

$$\Delta\phi = \frac{4GM}{Rc^2}. \tag{19.130}$$

For light rays just grazing the limb of the Sun, the deflection amounts to 1.75 arcsec.

Historically, this was a very important result. As described in Section 19.1.2 and endnote 6 of this chapter, Einstein derived the 'Newtonian' value of the deflection and found it to be half the value (19.130). These predictions led to the famous eclipse expeditions of 1919 led by Eddington and Crommelin to measure precisely the angular deflections of the positions of stars observed close to the limb of the Sun during a solar eclipse.[37] One expedition went to Sobral in Northern Brazil and the other to the island of Principe, off the coast of West Africa. The Sobral result was 1.98 ± 0.12 arcsec and the Principe result 1.61 ± 0.30 arcsec. These were technically demanding observations and there was some controversy about the accuracy of the results. The results were, however, clearly consistent with Einstein's prediction and from that time on Einstein's name became synonymous with scientific genius in the public imagination. The matter was laid to rest in the 1970s when radio interferometry found excellent agreement with Einstein's prediction. The precision of radio astronomical position measurements is now so high that the gravitational deflections of rays even at very large angles to the solar direction need to be corrected for the effects of the bending of space-time. The same is also true of the ultrahigh precision astrometry satellites of the European Space Agency, the Hipparcos mission of the 1990s and the Gaia mission of the 2010s.

The deflection of the images of background stars and galaxies by intervening masses has developed dramatically over the last 40 years, and the subject of *gravitational lensing* has become a major and exciting astronomical industry. Figure VI.2 in the introduction to this case study shows the 'horseshoe' arc observed about the centre of an elliptical galaxy, the mass of the galaxy acting as a gravitational lens which magnifies and distorts the image of a very distant background galaxy. Images such as these enable the mass distribution in the galaxy to be determined in some detail, producing convincing evidence for dark matter in the haloes of the galaxies and within clusters of galaxies.

19.9 Particles and Light Rays near Black Holes

Let us now consider a particle falling radially from rest at infinity to $r = 0$. We begin with the energy equation (19.97) with $h = 0$,

$$m\dot{r}^2 - \frac{2GMm}{r} = mc^2(k^2 - 1). \tag{19.131}$$

If the particle is at rest at infinity, the total energy of the particle is its rest mass energy $E = mc^2$ and hence, according to (19.93), $k = 1$ and

$$\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 = \frac{2GM}{r}. \tag{19.132}$$

This equation describes the dynamics of the particle in terms of the *proper time* measured at that point in space. Performing this integral,

$$\int_{\tau_1}^{\tau_2} \mathrm{d}\tau = -\int_{r_1}^{0} \frac{r^{1/2}}{(2GM)^{1/2}}\,\mathrm{d}r,$$

$$(\tau_2 - \tau_1) = \left(\frac{2}{9GM}\right)^{1/2} r_1^{3/2}. \tag{19.133}$$

Thus, the particle falls from r_1 to $r = 0$ in a *finite proper time*, and nothing strange takes place at the Schwarzschild radius r_g. The observer at infinity has, however, a different view of the world. That observer measures $\mathrm{d}t$ and not $\mathrm{d}\tau$. The metric,

$$ds^2 = d\tau^2 = \alpha\, dt^2 - \frac{1}{c^2}\alpha^{-1}\, dr^2, \tag{19.134}$$

can be reorganised so that

$$dt^2 = \alpha^{-1} d\tau^2 + \frac{1}{(\alpha c)^2}\, dr^2. \tag{19.135}$$

We now substitute for $d\tau^2$ using (19.71) to find the relation between dt and dr. After some reorganisation, we find

$$dt = -\frac{dr}{c\left(\dfrac{2GM}{rc^2}\right)^{1/2}\left(1-\dfrac{2GM}{rc^2}\right)} = -\frac{dr}{c\left(\dfrac{r_g}{r}\right)^{1/2}\left(1-\dfrac{r_g}{r}\right)}. \tag{19.136}$$

We can now integrate to find the coordinate time t for collapse from r_1 to $r = 0$,

$$t = \int_{t_1}^{t_2} dt = -\frac{1}{c r_g^{1/2}} \int_{r_1}^{0} \frac{r^{3/2}}{(r - r_g)}\, dr. \tag{19.137}$$

Now, the coordinate time t as measured by the distant observer diverges as r tends to r_g. Thus, although the particle takes a finite proper time to collapse to $r = 0$, it takes an infinite time to reach the Schwarzschild radius r_g according to the external observer. These different types of behaviour are illustrated in Fig. 19.13.

Correspondingly, light signals detected by a distant observer are progressively redshifted as $r \to r_g$. The proper time interval for signals emitted from some point in the field is

$$ds = d\tau = \left(1 - \frac{2GM}{rc^2}\right)^{1/2} dt. \tag{19.138}$$

Hence, in terms of frequencies,

$$\nu_{obs} = \left(1 - \frac{2GM}{rc^2}\right)^{1/2} \nu_{em} = \left(1 - \frac{r_g}{r}\right)^{1/2} \nu_{em}, \tag{19.139}$$

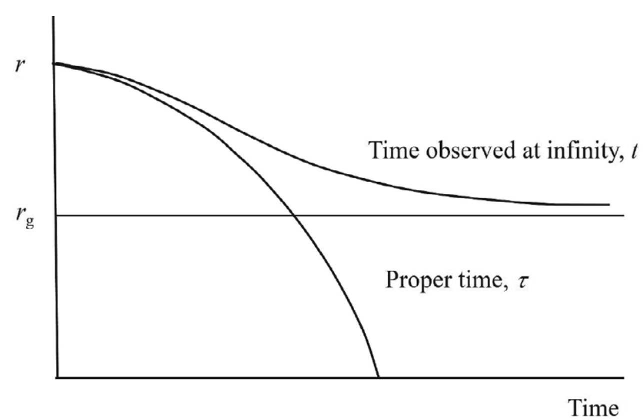


Fig. 19.13 Comparison of the proper time τ and the time observed at infinity t for a test particle falling radially from rest at infinity to $r = 0$.

where ν_{em} and ν_{obs} are the emitted and observed frequencies respectively. This is the correct expression for the *gravitational redshift* z_g according to general relativity, which can be written

$$z_g = \frac{\Delta\lambda}{\lambda} = \left(1 - \frac{r_g}{r}\right)^{-1/2} - 1. \tag{19.140}$$

Thus, the gravitational redshift tends to infinity as $r \to r_g$. As a result, we cannot observe light signals originating from radii $r < r_g$. Light signals can certainly travel inwards to $r = 0$, but those within r_g cannot propagate outwards from within r_g.

We have uncovered a number of the key properties of *black holes*. They are *black* because no radiation can travel outwards through r_g; they are *holes* because matter inevitably falls into a singularity at $r = 0$ if it approaches too close to the Schwarzschild radius r_g. Similar properties are found in the more general case of rotating, or Kerr, black holes.

Various singularities appear in the standard form of the Schwarzschild metric. The singularity at $r = r_g$ is now understood not to be a real physical singularity, but to be associated with the choice of coordinate system. In other coordinate systems, for example, in Kruskal or Finkelstein coordinates, this singularity disappears. More details of these features of coordinate systems in general relativity are discussed by Rindler in his book *Relativity: Special, General and Cosmological*.[13] In contrast, the singularity at $r = 0$ is a real physical singularity and so there is something missing from the theory. The obvious guess is that we need a quantum theory of gravity to eliminate the singularity, but no such theory yet exists.

19.10 Circular Orbits about Schwarzschild Black Holes

Let us use the Schwarzschild metric to work out some of the properties of particles in circular orbits about black holes. These studies are of central importance for high energy astrophysics, in particular, in the study of accretion discs about black holes. Matter falling towards a black hole is most unlikely to possess zero angular momentum and, consequently, collapse occurs along the rotation axis of the infalling material, resulting in the formation of an *accretion disc*. I have discussed some elementary features of thin accretion discs in my book *High Energy Astrophysics*.[38] Matter in the accretion disc moves in essentially Keplerian orbits about the black hole. In order that matter can arrive at the last stable orbit, it has to lose angular momentum and this takes place through the action of viscous forces in the differentially rotating accretion disc. Viscosity results in the transport of angular momentum outwards, enabling the material to spiral slowly in towards the last stable orbit at $r = 3r_g$ and at the same time heating up the disc to a high temperature. Let us use the tools we have developed to understand how much energy can be released as the matter drifts in towards the last stable orbit.

We showed in Section 19.6.1 that the total energy of a test mass orbiting a black hole, as observed at infinity, is $E = kmc^2$. From (19.97),

$$m\dot{r}^2 + \alpha m(r\dot{\phi})^2 - \frac{2GMm}{r} = mc^2(k^2 - 1). \tag{19.141}$$

For circular orbits, $\dot{r} = 0$ and $r^2\dot{\phi} = h$ is an invariant of the motion. Therefore,

$$\frac{\alpha mh^2}{r^2} - \frac{2GMm}{r} = mc^2(k^2 - 1). \tag{19.142}$$

The value of h can be found from (19.108). For a circular orbit, $\mathrm{d}^2u/\mathrm{d}\phi^2 = 0$, and so

$$u = \frac{GM}{h^2} + \frac{3GM}{c^2}u^2,$$
$$h^2 = \frac{GMr}{1 - \dfrac{3GM}{rc^2}}. \tag{19.143}$$

In the non-relativistic case, (19.143) reduces to the Newtonian result for the specific angular momentum, $h^2 = GMr$. In the case of the last stable orbit about the black hole at $r = 3r_g$, $h = \sqrt{12}GM/c$. The component of the velocity four-vector in the circumferential direction is

$$\gamma v = r\frac{\mathrm{d}\phi}{\mathrm{d}\tau} = r\dot{\phi} = \frac{h}{r}, \tag{19.144}$$

and hence for the last stable orbit $\gamma v = \sqrt{3}$ and $v = 0.5c$ – particles on the last stable orbit move at half the speed of light.

Substituting (19.143) into (19.142), we obtain the important result

$$k = \frac{1 - \dfrac{2GM}{rc^2}}{\left(1 - \dfrac{3GM}{rc^2}\right)^{1/2}}. \tag{19.145}$$

In the Newtonian limit, this result reduces to

$$E = kmc^2 = mc^2 - \frac{1}{2}\frac{GMm}{r}. \tag{19.146}$$

This represents the sum of the particle's rest mass energy mc^2 and the *binding energy* of its circular orbit about the point mass. We recall that, for a bound circular orbit, the kinetic energy T is one-half the magnitude of the gravitational potential energy $|U|$, what is known as the *virial theorem* in stellar dynamics,

$$T = \tfrac{1}{2}|U|. \tag{19.147}$$

Recall, however, that the gravitational potential energy is negative and so the total energy of the particle is

$$E_{\mathrm{orbit}} = T + U = \frac{1}{2}\frac{GM}{r} - \frac{GM}{r} = -\frac{1}{2}\frac{GM}{r}. \tag{19.148}$$

According to Newtonian gravity, if the particle is released from rest at infinity, as it falls towards the point mass, $T = |U|$ and so, if it does not lose kinetic energy, it can return back to infinity. In order to attain a bound orbit, half the particle's kinetic energy has to be lost, either dissipated as heat or by some other physical process. Thus, the expression for the *total energy* tells us how much energy has to be released in order that the particle can attain a bound orbit at that radius. This is why it is often referred to as the *binding energy* of the particle's orbit – this energy would have to be supplied to return the particle to $r = \infty$.

The corresponding result in general relativity follows directly from (19.145). The expression for the binding energy becomes

$$E_{\text{orbit}} = E - mc^2 = (k-1)mc^2 = \left[\frac{1 - \dfrac{2GM}{rc^2}}{\left(1 - \dfrac{3GM}{rc^2}\right)^{1/2}} - 1 \right] mc^2, \tag{19.149}$$

which describes how much energy is released by infall of matter from infinity to a circular orbit at radius r about the black hole. According to the Newtonian version of this argument (19.148), the amount of energy available is unbounded – as $r \to 0$, the binding energy tends to infinity. This does not take place in general relativity, as shown by (19.149). The plot of the binding energy as a function of radius in Fig. 19.14 shows that only a finite amount of binding energy can be made available for powering binary X-ray sources, active galactic nuclei and quasars.

The maximum binding energy occurs at $r = 3r_g$, the last stable orbit, and is

$$E_{\text{orbit}} = -\left[1 - \left(\frac{8}{9}\right)^{1/2}\right] mc^2 = -0.0572mc^2. \tag{19.150}$$

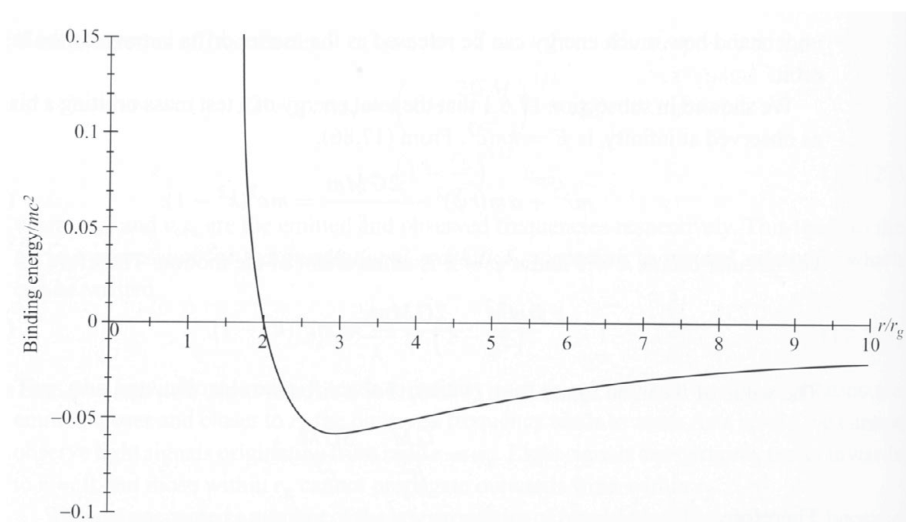


Fig. 19.14 The variation of the binding energy of a circular orbit in units of mc^2 with distance coordinate r in units of GM/c^2 for the Schwarzschild metric. The last stable orbit occurs at $r/(GM/c^2) = 6$ and the Schwarzschild radius at $r/(GM/c^2) = 2$.

This is a very important result for at least two reasons:

- To reach the last stable orbit, 5.72% of the rest mass energy of the particle has to be released. This is almost an order of magnitude greater than can be liberated by nuclear reactions – if four hydrogen nuclei are combined to form a helium nucleus, only 0.7% of the rest mass energy of the hydrogen atoms is released.
- This is the most compact volume from which an object of mass M can liberate this energy. If time variations on a time-scale T are observed from any source, causality requires the size of the source to be less than $r = cT$. The last stable orbit is the smallest dimension from which energy can be emitted from an object of mass M.
- Even greater efficiencies of energy release can be obtained by accretion onto rotating black holes. Solutions of Einstein's field equations for a rotating black hole were discovered by Roy Kerr in 1962. For a maximally rotating black hole, up to 42% of the rest mass energy of the infalling matter can be released, providing an enormous source of energy for powering extreme astrophysical objects such as quasars and radio galaxies.

Notice an interesting feature of (19.149). The binding energy is zero for an orbit with $r = 2r_g$, although it is unstable. This corresponds precisely to the locus passing though the point D in Fig. 19.11, which shows the binding energy as a function of radius r. Recall that a particle which reaches the point D along the locus from A can fall in from infinity to $r = 0$ despite the fact that it has a finite specific angular momentum – the orbit has zero binding energy.

19.11 Gravitational Waves

Einstein predicted the existence of gravitational waves in his paper of 1916, the year after the general theory was presented in its final form.[39] There was considerable debate about the reality of gravitational waves since they correspond to ripples in the fabric of space-time itself. The issue was resolved in favour of their existence.[40] The challenge facing the experimenters was enormous and it was to be 100 years after Einstein's prediction that it was fully confirmed by the remarkable detection of gravitational waves by the LIGO experiment.

19.11.1 The Search for Gravitational Waves

The search was begun by Joseph Weber in a pioneering set of experiments carried out in the 1960s. Weber was an electrical engineer by training and, having become fascinated by general relativity in the late 1950s, devised laboratory experiments to test the theory. He published his ideas in his book *General Relativity and Gravitational Radiation*[41] and set about constructing an aluminium bar detector to attempt to measure gravitational radiation from cosmic sources.[42] He estimated that with this detector he could measure strains as small as 10^{-16}, the fractional change in the dimensions of the detector caused by the passage of gravitational waves.

His first published results caused a sensation when he claimed to have found a positive detection of gravitational waves by correlating the signals from two gravitational wave detectors separated by a distance of 1000 km from the University of Maryland to the Argonne National Laboratory.[43] In a subsequent paper, he reported that the signal originated from the general direction of the Galactic Centre.[44] These results were received with considerable scepticism by the astronomical community since the reported fluxes far exceeded what even the most optimistic relativists would have predicted for the flux of gravitational waves originating anywhere in the Galaxy. The positive effect of Weber's experiments was that a major effort was made by experimentalists to check his results and, in the end, they could not be reproduced.

19.11.2 The Rate of Emission of Gravitational Waves

Let us give a simple heuristic estimate of the expected rate of radiation of gravitational waves. The following elementary arguments give an impression of how the result comes about. The essay by Bernard Schutz can be recommended as a gentle introduction to the physics of gravitational radiation.[45]

We use the analogy between the inverse square laws of electrostatics and gravity to work out the rate of radiation of quadrupole radiation from a system of charges or masses. The expression for the rate of loss of energy by electromagnetic radiation of an accelerated charged particle is given by (14.5):

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right) = \frac{|\ddot{\boldsymbol{p}}|^2}{6\pi\epsilon_0 c^3}, \tag{19.151}$$

where $\boldsymbol{p}$ is the dipole moment of the accelerated charge relative to some origin. In the case of an oscillating dipole, we can write $\boldsymbol{p} = \boldsymbol{p}_0\, \mathrm{e}^{-\mathrm{i}\Omega t} = q\boldsymbol{r}_0\, \mathrm{e}^{-\mathrm{i}\Omega t}$ where Ω is the angular frequency of oscillation and $\boldsymbol{p} = q\boldsymbol{r}$. Therefore, the expression for the time-averaged rate of radiation loss is

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right) = \frac{q^2\Omega^4|\boldsymbol{r}_0|^2}{12\pi\epsilon_0 c^3} = \frac{\Omega^4|\boldsymbol{p}_0|^2}{12\pi\epsilon_0 c^3}. \tag{19.152}$$

The equivalent expression for gravity results in no energy loss because a gravitational dipole has zero dipole moment, both 'gravitational charges' having the same sign. Hence, there is no dipole emission of gravitational waves.

A system of gravitating masses can, however, have a finite quadrupole moment and, if this is time varying, there is an energy loss by gravitational radiation. The corresponding equation to (19.152) for the time-averaged rate of loss of energy by electric quadrupole radiation is

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right) = \frac{\Omega^6}{1440\,\pi\epsilon_0 c^5}\sum_{j,k}|Q_{jk}|^2, \tag{19.153}$$

where the elements of the electric quadrupole moment tensor Q_{jk} are

$$Q_{jk} = \int (3x_j x_k - r^2\,\delta_{jk})\,\rho(x)\,\mathrm{d}^3 x, \tag{19.154}$$

and $\rho(x)$ is the electric charge density distribution.[46]

We obtain an approximate expression for quadrupole gravitational radiation in this 'Coulomb approximation' if we use the equivalence $e^2/4\pi\epsilon_0 \to Gm^2$. Then

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right)_{\mathrm{gr}} = \frac{G\Omega^6}{360\,c^5}\sum_{j,k}|Q_{jk}|^2, \tag{19.155}$$

where the Q_{jk} are now the components of the gravitational quadrupole moment tensor. The exact expression quoted by Schutz is

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right)_{\mathrm{gr}} = \frac{G}{5c^5}\left(\sum_{j,k}\dddot{Q}_{jk}\dddot{Q}_{jk} - \tfrac{1}{3}\dddot{Q}^2\right), \tag{19.156}$$

where the Q_{jk} are the components of spatial quadrupole tensor, or matrix, the second moment of the mass distribution ρ,[47]

$$Q_{jk} = \int \rho x_j x_k \,\mathrm{d}^3x. \tag{19.157}$$

Following Schutz's order of magnitude estimates, let us evaluate the rate of radiation of a binary star system, each star having mass M and moving in a circular orbit of radius r at speed v about their common centre of mass. To order of magnitude, the third derivative of the quadrupole moment is then

$$\frac{\mathrm{d}^3Q}{\mathrm{d}t^3} \sim \frac{Q}{t^3} \sim \frac{Mr^2}{t^3} \sim \frac{Mv^3}{r}, \tag{19.158}$$

where r and t are the typical spatial and time scales over which the quadrupole moment varies. Inserting this value into (19.156) for the term within large round brackets, we find the approximate gravitational radiation loss rate

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right)_{\mathrm{gr}} \sim \frac{G}{5c^5}\left(\frac{Mv^3}{r}\right)^2. \tag{19.159}$$

Since $v = \Omega r$, we find the result

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right)_{\mathrm{gr}} \sim \frac{G}{5c^5}M^2r^4\Omega^6. \tag{19.160}$$

Notice that (19.156) and (19.160) have exactly the same dependence upon the quantities which appear in the loss rate for gravitational radiation since $Q_{jk} \sim Mr^2$. The angular velocity of the binary star is $\Omega^2 = GM/4r^3$ and so (19.160) can be rewritten

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right)_{\mathrm{gr}} \sim \frac{c^5}{80G}\left(\frac{GM}{rc^2}\right)^5, \tag{19.161}$$

where, following Schutz, it is assumed that four comparable terms contribute to the total gravitational loss rate. As Schutz points out, the term c^5/G corresponds to a luminosity of 3.6×10^{52} W, an enormous value, which is modified by the 'relativistic factor' $(GM/rc^2)^5$. Thus, really close binary neutron stars and black holes are expected to be extraordinarily powerful sources of gravitational waves. In particular, the very last phases of inspiralling

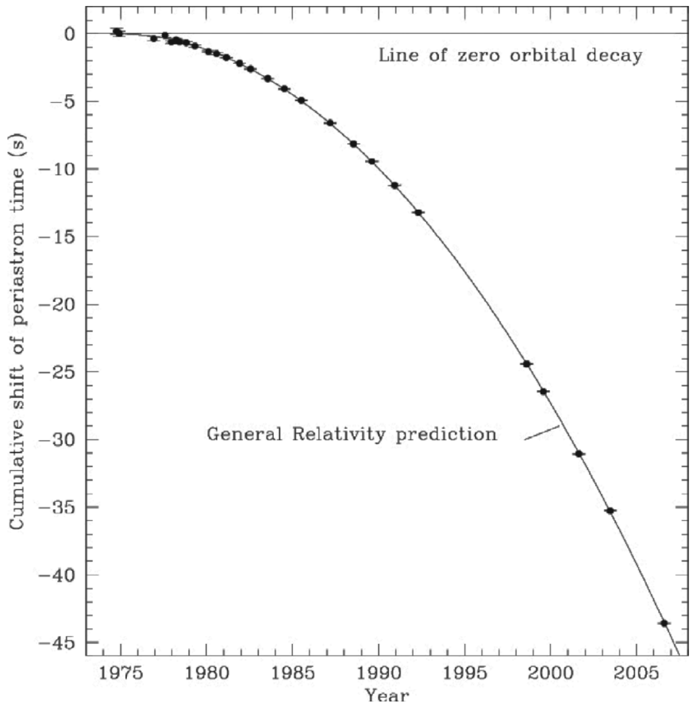


Fig. 19.15 The change of orbital phase as a function of time for the binary neutron star system PSR 1913+16 compared with the expected change due to gravitational radiation energy loss by the binary system (solid line). (see endnote[18].)

of a pair of neutron stars or black holes in a very close binary system, before they coalesce to a single black hole, were expected to be promising targets for observation by the present generation of gravitational wave detectors.

A key relation from our present perspective is that (19.160) shows that the rate of energy loss by gravitational waves is proportional to Ω^6. This energy is extracted from the binding energy of the binary system, $E = -\frac{1}{2}I\Omega^2$, and so we can write

$$-\left(\frac{\mathrm{d}E}{\mathrm{d}t}\right)_{\mathrm{gr}} = \left[\frac{\mathrm{d}\left(\frac{1}{2}I\Omega^2\right)}{\mathrm{d}t}\right] \propto \Omega^6, \tag{19.162}$$

$$\dot{\Omega} \propto \Omega^5. \tag{19.163}$$

Exactly this relation was found by Russell Hulse and Joseph Taylor for the spin-up of the binary pulsar PSR 1913+16 (Fig. 19.15). They achieved much more than this, however, because the orbits of the two neutron stars and their masses were determined by very precise timing observations and so the exact radiation loss rate for their highly elliptical orbits, $e = 0.615$, could be estimated and found to be in excellent agreement with the observed spin-up rate of the binary system. The resulting tests of general relativity were among the best in constraining possible deviations from general relativity.

19.11.3 Discovery of Gravitational Waves

Weber's endeavours set a challenge to the experimental community to devise enhanced methods for detecting the extremely tiny strains expected from strong sources of gravitational waves. The outcome was the funding of a number of major projects aimed at detecting the elusive gravitational waves. Many new technologies had to be developed and the programmes carried out in the 1980s can best be regarded as development programmes. For example, the LIGO project, an acronym for Laser Interferometer Gravitational-Wave Observatory, was initiated by Kip Thorne and Rainer Weiss in 1984, but approval by the US National Science Foundation was only given in 1994 at a revised cost for development, construction and early operations of $365 million.[48] The project consists of two essentially identical interferometers each with a 4-km baseline located at Livingston, Louisiana, and Hanford near Richland, Washington. The VIRGO project is a French–Italian collaboration to construct an interferometer with a 3-km baseline at a site near Pisa, Italy. The GEO600 experiment is a German–British interferometer project with a 600-m baseline, while the Japanese TAMA project is a 300-m baseline interferometer located at Mitaka, near Tokyo. For all these projects, there was a long development programme to reach the sensitivities at which there is a good chance of detecting gravitational waves from celestial sources. The importance of having a number of gravitational wave detectors operating simultaneously is that they form a global network and any significant events should be detectable by all of them. By precise timing, the direction of arrival of the gravitational waves can be estimated.

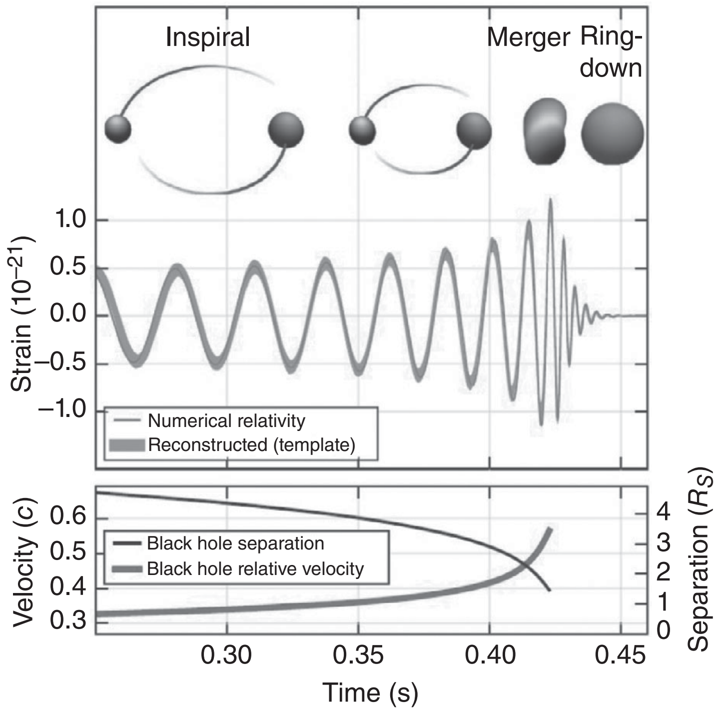
The goal was achieved in 2015 with the detection of a pulse, or chirp, of gravitational waves from a pair of inspiralling black holes and their coalescence into a single more massive black hole. This process has a characteristic signature which can be calculated precisely because the black holes are such simple objects. The minute strains expected from such a stupendous event are shown in Fig. 19.16(a), along with the signals observed at the two LIGO detectors located at Hanford, Washington, and Livingston, Louisiana, on 14 September 2015 (Fig. 19.16(b)). The observation of the black hole coalescence signature is unambiguous.

The properties of the coalescing black holes involved in this event were as follows:

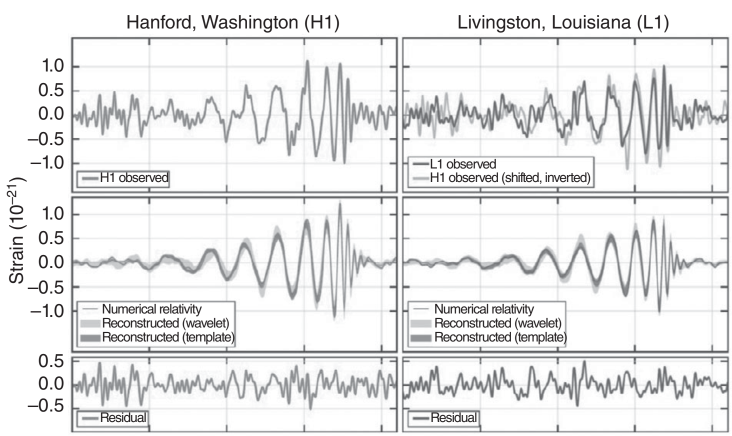
- The masses of the two black holes were both about 30 times the mass of the Sun before coalescence.
- They moved about each other at about half the speed of light just before coalescence.
- Three solar masses of the material of the black holes were converted into the energy of gravitational waves in less than a tenth of a second.

This is really serious high energy astrophysics. Since that date several other wholly convincing examples of coalescing binary black holes have been discovered.

One of the most dramatic detections was the discovery on 17 August 2017 of a pair of coalescing neutron stars.[49] In this case, a γ-ray burst was observed 2 seconds after the gravitational wave event. The resulting afterglows were followed up by a huge number of telescopes throughout the electromagnetic spectrum and the associated explosion, a



(*a*)



(*b*)

Fig. 19.16 (*a*)The predicted variation of the strain in the LIGO detector system resulting from the coalescence of a pair of black holes. (*b*)The signatures and observed signals from the two LIGO detectors of the event of 14 September 2015. (From B.P. Abbott, *et al.*, the LIGO Scientific Collaboration and the Virgo Collaboration, 2016, 'Observation of Gravitational Waves from a Binary Black Hole Merger', *Physical Review Letters*, **116**, 061102.)

'kilonova', was detected in the relatively nearby elliptical galaxy NGC 4993. The vision of multi-wavelength astronomy supported by outstanding technology, computation and theory had come about with a vengeance. And a new era of gravitational wave and black hole physics was born.

Notes

1 Pais, A. (1982). *Subtle is the Lord . . . the Science and Life of Albert Einstein*. Oxford: Clarendon Press.
2 Stachel, J. (1995). The History of Relativity, in *Twentieth Century Physics.* Vol. 1, eds. L.M. Brown, A. Pais and A.B. Pippard, pp. 249–356. *Bristol: Institute of Physics Publishing and New York*: American Institute of Physics Press.
3 Einstein, A. (1922). Kyoto address of December 1922. J. Ishiwara (1977). *Einstein Köen-Roku*. Tokyo: Tokyo-Tosho.
4 Einstein, A. (1907). *Jahrbuch Radioaktivität Elektronik*, **4**, 411–462.
5 Einstein, A. (1911). *Annalen der Physik*, **35**, 898–908.
6 This result can be readily obtained from the Rutherford scattering formula (A4.22) by converting from the electrostatic to the gravitational inverse square laws. The speed of the particle is the speed of light c and $zZe^2/4\pi\epsilon_0$ is replaced by GMm.
7 Soldner, J.G. von (1804). *Berliner Astronomisches Jahrbuch*, p. 161.
8 Einstein, A. (1912). *Annalen der Physik*, **38**, 1059.
9 Einstein, A. (1922). In Ishiwara (1977), *op. cit*.
10 Einstein, A. (1916). *Die Grundlage der Allgemeinen Relativitätstheorie*, p. 6. Leipzig: J.A. Barth.
11 Einstein, A. and Grossmann, M. (1913). *Zeitschrift für Mathematik und Physik*, **62**, 225.
12 Berry, M. (1995). *Principles of Cosmology and Gravitation*. Cambridge: Cambridge University Press.
13 Rindler, W. (2001). *Relativity: Special, General and Cosmological*. Oxford: Oxford University Press.
14 Weinberg, S. (1972). *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. New York: John Wiley and Sons.
15 d'Inverno, R. (1992). *Introducing Einstein's Relativity*. Oxford: Clarendon Press.
16 Misner, C.W., Thorne, K.S. and Wheeler, J.A. (1973). *Gravitation*. San Francisco: Freeman and Co.
17 Will, C.M. (2018). *Theory and Experiment in Gravitational Physics*, 2nd edition. Cambridge: Cambridge University Press.
18 Will, C.W. (2014). www.livingreviews.org/lrr-2014-4.
19 Touboul, P. *et al.* (2017). MICROSCOPE Mission: First Results of a Space Test of the Equivalence Principle, *Physical Review Letters*, **119**, 231101.
20 Rindler, W. (2001). *op. cit.*, p. 19.
21 Einstein, A. (1920). Unpublished manuscript, now in the Pierpoint Morgan Library, New York. See Pais (1982). *op. cit.*, p. 178.
22 This statement is due to Professor Anthony Lasenby in his lectures on Gravitation and Cosmology.
23 Pound, R.V. and Rebka, G.A. (1959). Gravitational Red-Shift in Nuclear Resonance. *Physical Review Letters*, **3**, 439–441.
24 Pound, R.V. and Snider, J.L. (1964). Effect of Gravity on Nuclear Resonance. *Physical Review Letters*, **13**, 539–540.

25 An excellent history of the discovery of non-Euclidean geometry is provided by R. Bonola in his book *Non-Euclidean Geometry, and the Theory of Parallels by Nikolas Lobachevsky, with a Supplement Containing the Science of Absolute Space by John Bolyai*, trans. H.S. Carslaw. New York: Dover (1955).
26 See North, J.D. (1965). *The Measure of the Universe*, pp. 74–75. Oxford: Clarendon Press.
27 It is often stated that Gauss carried out the measurements himself, but this is probably a myth. In his biography of Gauss, Bühler writes as follows (W.K. Bühler, *Gauss: A Biographical Study*, p. 100. Berlin: Springer-Verlag, 1981): 'The often-told story according to which Gauss wanted to decide the question [of whether space is non-Euclidean] by measuring a particularly large triangle is, as far as we know, a myth. The great triangle Hohenhagen-Inselberg-Brocken was a useful control for the smaller triangles which it contains. Gauss was certainly aware of the fact that the error of measurement was well within the possible deviation from 180 degrees from which, under strict conditions, one could have derived the non-Euclidean nature of space.'
28 Lobachevsky, I.N. (1829, 1830). On the Principles of Geometry. *Kazanski Vesnik (Kazan Messenger)*; Bolyai, J. (1832). Appendix: Scientiam spatii absolute veritam exhibens (Appendix explaining the absolutely true science of space). Published as an appendix to the essay by his father F. Bolyai, *An Attempt to Introduce Studious Youth to the Elements of Pure Mathematics*. Maros Vásárhely, Transylvania.
29 Riemann, B. (1854). *Uber die Hypothesen welche der Geometrie zu Grunde liegen (On the Hypotheses that lie at the Foundations of Geometry)*. Göttingen: Habilitationschrift.
30 Schwarzschild, K. (1900). Über das zulgässige Vergleichung der Helligkeit Verscheidenfarbiger Sterne (On an upper limit to the curvature of space), *Vierteljahrsschrift des Astronomischen Gesellschaft*, **35**, 337–347.
31 The essence of the following arguments was first shown to me by Peter Scheuer.
32 Berry, M. (1995). *Principles of Cosmology and Gravitation*. Cambridge: Cambridge University Press.
33 Weinberg, S. (1972). *Gravitation and Cosmology*.
34 Rindler, W. (2001). Note 13, *op. cit.*
35 Schwarzschild, K. (1916). *Sitzungsberichte der Preussischen Akademie der Wissenschaften*, 189.
36 Shapiro, S.L. and Teukolsky, S.A. (1983). *Black Holes, White Dwarfs and Neutron Stars: The Physics of Compact Objects*. New York: John Wiley and Sons.
37 See my essay on the 1919 eclipse expedition. Longair, M.S. (2016). Bending spacetime: a commentary on Dyson, Eddington and Davidson (1920) 'A determination of the deflection of light by the Sun's gravitational field', *Philosophical Transactions of the Royal Society*, **A373**, 20140287.
38 Longair, M.S. (2011). *High Energy Astrophysics*, 3rd edition. Cambridge: Cambridge University Press.
39 Einstein, A. (1916). Näherungsweise Integration der Feldgleichungen der Gravitation. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften Berlin*. Part 1: 688–696.
40 A brief history of some of the problems involved in understanding the existence of gravitational waves in general relativity is given by John Stachel in his review 'History of Relativity', in *Twentieth Century Physics*, eds. L.M. Brown, A. Pais and A.B. Pippard, Vol. 1, pp. 249–356. Bristol and Philadelphia: Institute of Physics Publishing and New York: American Institute of Physics Press (1995).
41 Weber, J. (1961). *General Relativity and Gravitational Waves*. Interscience Tracts on Physics and Astronomy, New York: Interscience.
42 Weber, J. (1966). Observation of the Thermal Fluctuations of a Gravitational-Wave Detector, *Physical Review Letters*, **17**, 1228–1230.

43 Weber, J. (1969). Evidence for Discovery of Gravitational Radiation, *Physical Review Letters*, **22**, 1320–1324.

44 Weber, J. (1970). Anisotropy and Polarization in the Gravitational-Radiation Experiments, *Physical Review Letters*, **25**, 180–184.

45 *Encyclopaedia of Astronomy and Astrophysics*, ed. P. Murdin, Vol. 2, pp. 1030–1042. Bristol and Philadelphia: Institute of Physics Publishing and London: Nature Publishing Group, 2001.

46 The expression (19.154) is derived by J.D. Jackson in his book *Classical Electrodynamics*, 3rd edition, Section 9.3, New York: John Wiley and Sons, Inc., 1999.

47 Notice the difference of a factor three between the definitions of the quadrupole moment tensor for electromagnetism (19.154) and gravity (19.156). There is a compensatory factor of three in the term in large brackets in (19.156). The difference in the gravitational case is because of technicalities about the choice of gauge to be used in the full development of the theory of gravitational radiation.

48 A very readable account of black holes, gravitational waves and their significance for physics and astronomy is contained in Kip Thorne's book *Black Holes and Time Warps: Einstein's Outrageous Legacy*, New York: W.W. Norton and Company, 1994.

49 B.P. Abbott *et al.* The LIGO Scientific Collaboration and Virgo Collaboration (2017). GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral. *Physical Review Letters*, **119**, 161101.

Case Study VII

COSMOLOGY AND PHYSICS

Famously, Ernest Rutherford, Cavendish Professor of Experimental Physics and Head of the Cavendish Laboratory, stated:

> Don't let me catch anyone talking about the Universe in my department.[1]

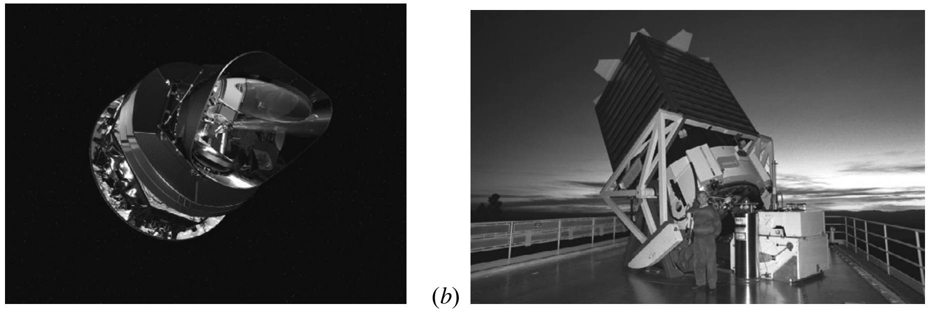
He made this remark despite the fact that already in 1904 he had made the first unambiguous estimate of the age of the Earth, and consequently the Universe, from the study of radioactive decay chains. Understandably, in the early twentieth century, there were neither the observations nor the theoretical tools which could have made the large-scale structure of the Universe and its evolution an area of serious study for physicists.

The many observational discoveries since the 1960s throughout the electromagnetic spectrum have completely changed this perspective. The physics of the large-scale structure and geometry of the Universe are now just as much integral parts of physics as is its structure on the microscopic scales of atoms, nuclei, elementary particles and so on.

We now have a clear physical picture of the processes involved in the evolution of the Universe from the time it was of the order a microsecond old to the present era – this story is full of splendid physics and illustrates how the laws established in the laboratory can be applied successfully on the scale of the very large, as well as the very small. This new understanding has also posed a number of fundamental problems for physicists and cosmologists and these will undoubtedly become part of the new physics of the twenty-first century. Fortunately, the essential physics necessary to understand what is involved in reaching the present level of understanding can be developed rather precisely using the tools of an undergraduate physics course, combined with ideas from elementary general relativity discussed in Chapter 19.

The two chapters of this case study concern the standard model of cosmology, often referred to colloquially as the Big Bang. The first of these, Chapter 20, is designed to be one which would be agreed by essentially all cosmologists, taking into account the various caveats along the way. What is compelling about this story is that, at face value and giving due regard to the uncertainties, the standard model – the cold dark matter model with a finite cosmological constant Λ, or the ΛCDM model – can account for a vast amount of independent observational data, now of the very highest quality. Even one convincingly

[1] As quoted by John Kendrew in 'J.D. Bernal and the Origin of Life', BBC Radio Talk 26 July 1968, and in Daintith, J. (2008), *Biographical Encyclopedia of Scientists, Third Edition*, p. 662. Boca Raton, FL: CRC Press.



(*a*) (*b*)

Fig. VII.1 (*a*) The Planck satellite of the European Space Agency which made the definitive whole-sky map of the Cosmic Microwave Background Radiation (Courtesy of ESA and the Planck project). (*b*) The 2.5-metre Sloan Telescope which carried out the Sloan Digital Sky Survey (SDSS).

divergent piece of observational data would bring down the whole house of cards. The fact that this has not happened is a very striking achievement for the standard model. At the same time, the search is on for observational phenomena which would act as further critical tests of the standard picture.

Sadly, this is not the place to recount the extraordinary advances in observational technique which have made these confrontations of theory with observation possible. Just as a couple of examples, Fig. VII.1 (*a*) and (*b*) show the Planck satellite of ESA, which has set the benchmark for precise observations of the Cosmic Microwave Background Radiation, and the Sloan Telescope which has delineated with exquisite precision the large-scale distribution of galaxies at the present epoch.

The second chapter, Chapter 21, concerns the deep problems which arise from the success of the ΛCDM model. These are exercising the imaginations of the world-wide community of theorists and cosmologists and all the possible routes ahead involve innovative physics beyond what has been established in the laboratory. The most popular solution to many of the basic problems is the *inflationary scenario* for the early Universe. There is no proper physical realisation of the inflationary picture but it has undoubted attractions. As in the past, this may well be another example where astrophysics and cosmology provide clues to as yet unknown physics, particularly when combined with contemporaneous advances in particle physics.

The reader interested in many more details of the historical development of modern cosmology may find the volume edited by Helge Kragh and me, *The Oxford Handbook of the History of Modern Cosmology*,[2] of some help in understanding the complex history of this development. The authors range from observational and theoretical cosmologists to professional historians of physics and cosmology, as well as philosophers of science. This broad perspective will provide some appreciation of the richness and fecundity of the discipline. Consequently, rather than history, the exposition in this final case study will concentrate more upon the logical development of the construction of cosmological models and comparison with the key observations.

[2] Kragh, H., and Longair, M.S. (eds.) (2019). *The Oxford Handbook of the History of Modern Cosmology*. Oxford: Oxford University Press.

20 Cosmology

20.1 Cosmology and Physics

By cosmology, I mean the application of the laws of physics to the Universe as a whole. The validation of theory thus depends upon precise observation rather than experiment, placing us at one stage further removed from our 'apparatus', than is in the case in laboratory physics. Yet, throughout history, astronomical observations have played their role in establishing new physics, which was rapidly assimilated into the mainstream of established science. In Chapter 2, we discussed how Tycho Brahe's observations of the motions of the planets led to Newton's law of gravity. From observations of the eclipses of the satellites of Jupiter, Ole Rømer showed conclusively that the speed of light is finite, and in 1676 estimated its value from the time it took light to travel across the Earth's orbit about the Sun. Most tests of general relativity have featured astronomical measurements, for example, the use of the binary pulsar PSR 1913+16 to constrain deviations from general relativity and, perhaps most spectacularly, the discovery of gravitational waves (Section 19.11.2). We will find other examples of constraining physical parameters as the physics of the standard ΛCDM model became securely established.

Thus, astronomy has played, and continues to play, a major role in fundamental physics. Equally impressive is how the very best of laboratory physics enables us to understand the properties of celestial objects in a convincing way. Conversely, astronomical observations enable us to test and extend the laws of physics in physical circumstances which are unattainable in the laboratory.

There is, however, a major issue which besets the application of the laws of physics to the Universe as a whole. We have only one Universe to observe, and physicists are rightly cautious about accepting the results of any one-off experiment. In cosmology, however, there is no prospect of doing any better. Despite this limitation, the application of the laws of physics to cosmological problems turns out to be a remarkably successful programme and leads to the standard ΛCDM model. But, this picture also leads to a number of fundamental problems, which are addressed in Chapter 21. The general consensus is that these problems can only be resolved by new types of physics which are applicable at the ultra-high temperatures and densities encountered in the very early Universe. These conditions far exceed those at which the laws of physics have been tested in terrestrial laboratories. Nonetheless, the numerous successes of the standard picture have led theorists to consider the very earliest phases of the Big Bang as a laboratory for physics at extreme energies.

20.2 Basic Cosmological Data

The models which describe the large-scale structure and evolution of the Universe are based upon two key observations, the isotropy and homogeneity of the Universe on the large scale and Hubble's law. These observations on their own enable us to set up the theoretical infrastructure needed to undertake cosmological studies.[1]

20.2.1 The Isotropy and Homogeneity of the Universe

The Universe is highly anisotropic and inhomogeneous on small scales but, as we observe larger and larger scales, it becomes more and more isotropic and homogeneous. Galaxies are the building blocks which define the large-scale distribution of visible matter in the Universe. On the small scale, they are clustered into groups and clusters of galaxies, but, when viewed on a large enough scale, these irregularities are smoothed out. Figure 20.1 shows the distribution of over two million galaxies derived from the Cambridge APM galaxy survey in a region of sky covering about one-tenth of the celestial sphere.

It is apparent from Fig. 20.1 that, looked at on a large enough scale, one region of the Universe looks very much like another, which is *prime facie* evidence that, in the simplest

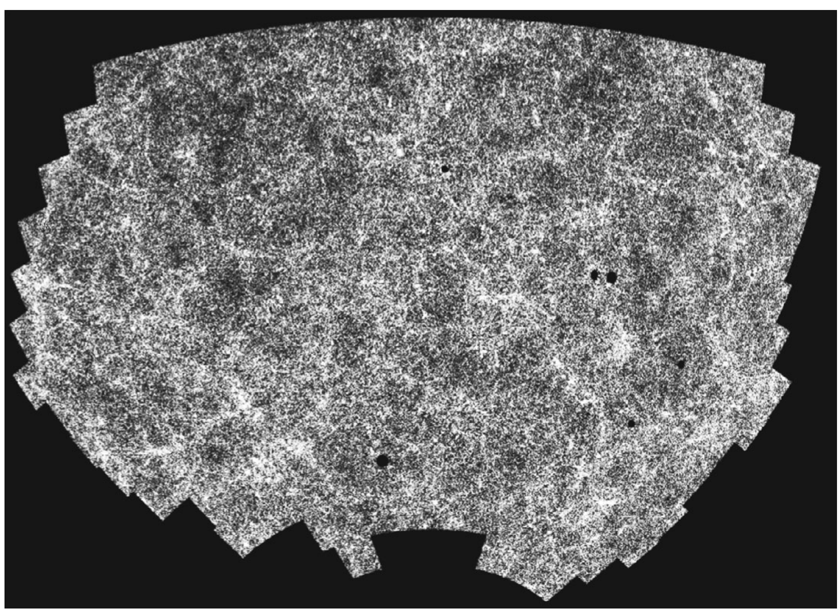


Fig. 20.1 The distribution of galaxies with apparent astronomical magnitudes $17 \leq b_j \leq 20.5$ shown in an equal area projection centred on the Southern Galactic Pole. This image was reconstructed from machine scans of 185 UK Schmidt plates by the Cambridge APM machine. There are over two million galaxies in this image. The small empty patches are regions which have been excluded about bright stars, nearby dwarf galaxies, globular clusters and step wedges. (From Maddox, S.J., Efstathiou, G., Sutherland, W.G. and Loveday, J. (1990), *Monthly Notices of the Royal Astronomical Society*, **242**, 43P.)

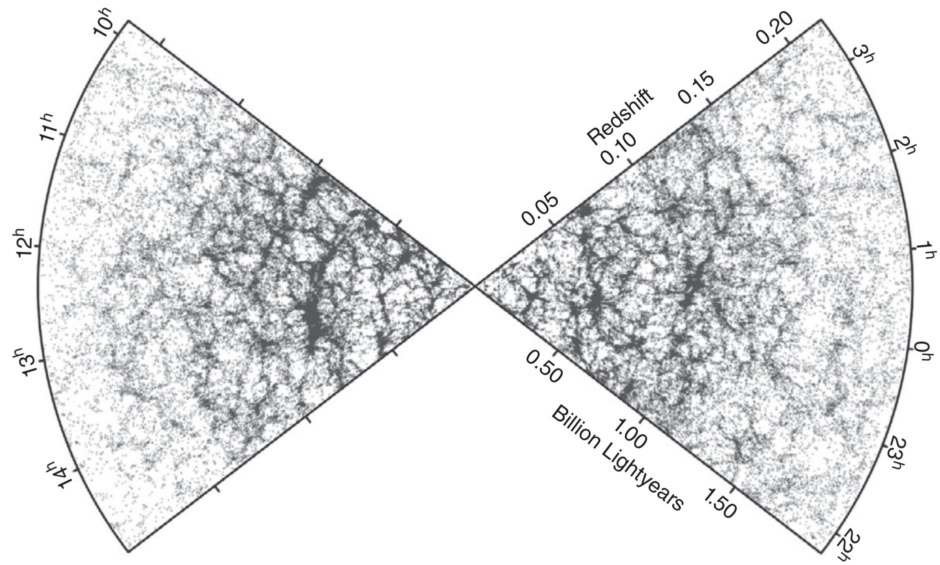


Fig. 20.2 The 2dF Galaxy Redshift Survey used the two-degree field spectroscopic facility of the Anglo-Australian Telescope to measure the redshifts of approximately 220 000 galaxies. (From Colless *et al.* (the 2dF-GRS team) (2001). *Monthly Notices of the Royal Astronomical Society*, **328**, 1039–63. doi: 10.1046/j.1365-8711.2001.04902.x.)

approximation, the Universe is isotropic on a large enough scale. Nonetheless, there are non-random features in the galaxy distribution. The nature of these non-random features is best appreciated from the three-dimensional mapping of the spatial distribution of large samples of galaxies for which distances have been measured. A slice through the galaxy distribution out to redshifts of about 0.2 is shown in Fig. 20.2 from the Anglo-Australian Telescope two-degree field Galaxy Redshift Survey (2dF-GRS).

There appear to be ‘walls’ of galaxies, as well as regions of significant under-density, showing that the distribution of galaxies exhibits a ‘cellular’ structure. Statistical analyses have shown that topologically the distribution of galaxies resembles a sponge. Where the galaxies are located corresponds to the material of the sponge, while the ‘holes’ or ‘voids’ in the galaxy distribution correspond to the holes in the sponge. Both the material of the sponge and the holes are continuously connected. The sizes of the holes can be very large indeed, some of them being as large as about 50 times the typical size of a rich cluster of galaxies. Analyses of the distribution of galaxies, such as that shown in Fig. 20.2, have demonstrated that the same ‘spongy’ distribution persists from the local Universe out to the limit of the APM survey. This is direct evidence, not only for the isotropy of the Universe, but also for its homogeneity on the large scale – although the distribution of galaxies is irregular, it displays the same degree of irregularity as observations are extended to larger and larger distances.

The evidence of the distribution of galaxies is compelling enough, but even more spectacular are the observations of the Cosmic Microwave Background Radiation (CMB). Soon after its discovery by Arno Penzias and Robert Wilson in 1965, it was established that

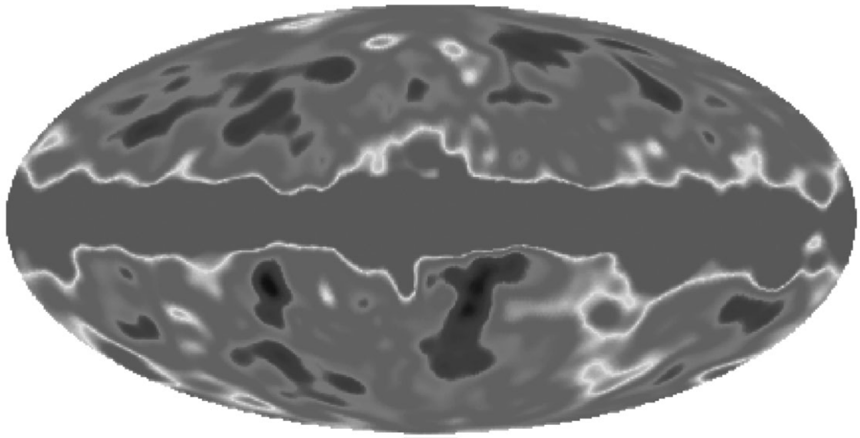


Fig. 20.3 A map of the whole sky in galactic coordinates as observed by the COBE satellite at a wavelength of 5.7 mm (53 GHz). The centre of our Galaxy is in the centre of the image and the plane of our Galaxy is the broad band across the centre of the image. The intensity fluctuations away from the Galactic plane are real fluctuations in the intensity of the background radiation and correspond to an rms temperature fluctuation of only $35 \pm 2\,\mu$K, compared with a total brightness temperature of 2.728 K. (From Bennett, *et. al.* (1996), *Astrophysical Journal*, **464**, L1.)

the spectrum of the radiation is of black-body form and is remarkably isotropic over the sky. These features of the CMB were studied in detail by the Cosmic Background Explorer, COBE, which was launched by NASA in September 1989. The spectrum of the radiation is that of a perfect black body at a radiation temperature $T = 2.728$ K, deviations from this spectrum amounting to less than 0.03% of its maximum intensity in the wavelength interval $2.5 > \lambda > 0.5$ mm. The distribution of the radiation away from the plane of our Galaxy is quite extraordinarily isotropic, there being no evidence for deviations from isotropy at the level of about one part in 10^5 on angular scales greater than $10°$ (Fig. 20.3), but at about this level real cosmological fluctuations are observed at high galactic latitudes – a much higher resolution and sensitivity map of the CMB is shown in Fig. 20.14. This radiation is the cooled remnant of the very hot early phases of the expanding Universe.

We can be certain that the isotropy refers to the Universe as a whole because ‘shadows’ of very distant clusters of galaxies, containing hot gas, are observed in the background radiation as a result of the Sunyaev–Zeldovich effect.[2] Thus, we can be certain that, on large scales, the Universe is isotropic to better than one part in 100 000 – this is a quite astounding cosmological result which simplifies greatly the construction of cosmological models.

20.2.2 Hubble’s Law

Edwin Hubble discovered the velocity–distance relation for galaxies in 1929,[3] using the radial velocities of 24 galaxies, most of which had been measured painstakingly by Vesto M. Slipher. Hubble estimated the distances of these galaxies using a variety of different astronomical techniques and found that the recession velocity v of a galaxy from our

Galaxy is proportional to its distance r, $v = H_0 r$, where H_0 is known as Hubble's constant – this relation is known as *Hubble's law*. The extension of the velocity–distance relation to very much greater distances was convincingly demonstrated by Hubble and Humason a few years later.[4]

The velocities were inferred from the *redshifts* of the galaxies, meaning the shifts of spectral lines to longer wavelengths because of the recessional velocities of the galaxies. If λ_e is the wavelength of the line as emitted and λ_0 the observed wavelength, the redshift z is defined by

$$z = \frac{\lambda_0 - \lambda_e}{\lambda_e}. \tag{20.1}$$

If the galaxy is moving radially away from the observer at velocity v, in the limit of small redshifts, $v/c \ll 1$,

$$v = cz. \tag{20.2}$$

This is the type of velocity used in Hubble's law. It is strongly recommended that redshifts rather than velocities are used since the former are dimensionless and have a very much deeper significance for cosmology than the velocities, as will be discussed below.

Traditionally, the distances of galaxies have been estimated by finding some class of astronomical object, all the members of which have the same intrinsic luminosity L. Then, the observed flux density S of the object is given by the inverse square law, $S = L/4\pi r^2$. Astronomers traditionally use a logarithmic scale of flux density known as the *apparent magnitude m* such that $m = -2.5 \log S + \text{constant}$. Therefore, if the intrinsic luminosity of the distance indicator is independent of distance, we expect

$$m = -2.5 \log S + \text{constant} = 5 \log r + \text{constant}. \tag{20.3}$$

Figure 20.4 shows the redshift–distance relation for the brightest galaxies in rich clusters of galaxies as determined by Allan Sandage in the 1960s. There is an excellent linear correlation between apparent magnitude and the logarithm of the redshift, with exactly the slope expected if redshift were precisely proportional to distance. The same relation is found for all classes of extragalactic object.

A major programme for observational cosmologists has been the precise determination of the value of Hubble's constant H_0. Thanks to the efforts of many astronomers and a large programme of observations using the Hubble Space Telescope and the analysis of the high resolution maps of the CMB, there is now general agreement that the value of H_0 is close to $H_0 = 70$ km s^{-1} Mpc^{-1}, with an uncertainty of a few percent (Section 20.9).

20.2.3 The Local Expansion of the Distribution of Galaxies

The combination of the isotropy of the Universe and Hubble's law shows that the Universe as a whole is expanding uniformly at the present time. In a uniform expansion, the ratio of the distances between any two points increases by the same factor in a given time interval t_1 to t_2, that is,

$$\frac{r_1(t_2)}{r_1(t_1)} = \frac{r_2(t_2)}{r_2(t_1)} = \cdots \frac{r_n(t_2)}{r_n(t_1)} = \cdots = \alpha = \text{constant}, \tag{20.4}$$

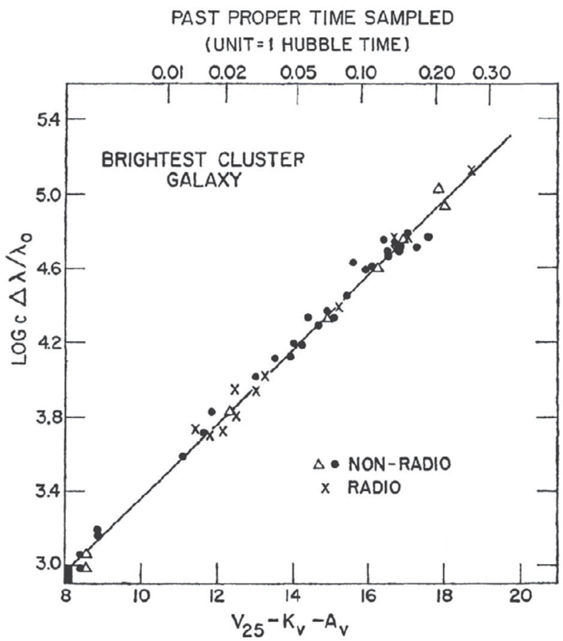


Fig. 20.4 The velocity–distance relation for the brightest galaxies in rich clusters. The abscissa is the corrected apparent magnitude in the *V* waveband and the ordinate the logarithm of the recessional velocity. This correlation indicates that the brightest galaxies in clusters have remarkably standard properties and that their velocities of recession from our Galaxy are proportional to their distances. (From Sandage, A.R. (1968), *Observatory*, **88**, 91.)

where the distances r_n are measured from some arbitrary point. The recession velocity of galaxy 1 from the origin is therefore

$$v_1 = \frac{r_1(t_2) - r_1(t_1)}{t_2 - t_1} = \frac{r_1(t_1)}{t_2 - t_1}\left[\frac{r_1(t_2)}{r_1(t_1)} - 1\right]$$
$$= \frac{(\alpha - 1)}{t_2 - t_1} r_1(t_1) = H_0 r_1(t_1). \quad (20.5)$$

Hence, for any galaxy,

$$v_{\rm n} = \frac{(\alpha - 1)}{t_2 - t_1} r_{\rm n}(t_1) = H_0 r_{\rm n}(t_1). \quad (20.6)$$

Thus, a uniformly expanding distribution of galaxies automatically results in a velocity–distance relation of the form $v = H_0 r$. Notice that it does not matter which location we choose as the origin. All stationary observers moving with the uniform expansion find that they are at the centre of an expanding Universe with the same Hubble relation, if the observations are made at the same cosmic epoch.

20.3 The Robertson–Walker Metric

From the considerations of the last section, the natural starting point for the construction of cosmological models is the assumption that the Universe is isotropic, homogeneous and expanding uniformly at the present epoch. One of the problems facing the pioneers of relativistic cosmology was the definition of the space and time coordinates to be used in building the simplest cosmological models. The problem was solved independently by Howard Robertson and Arthur Walker in 1935–36, who derived the form of the metric of space-time for *all* isotropic, homogeneous, uniformly expanding models of the Universe.[5] The form of the metric is independent of the assumption that the large-scale dynamics of the Universe are described by general relativity – whatever the physics of the expansion, the space-time metric must be of Robertson–Walker form, because of the assumptions of isotropy and homogeneity.

A key step in the development of these models was the introduction by Hermann Weyl in 1923 of what is known as *Weyl's postulate*. To eliminate the arbitrariness in the choice of coordinate frames, Weyl introduced the idea that, in the words of Hermann Bondi,

> The particles of the substratum (representing the nebulae) lie in space-time on a bundle of geodesics diverging from a point in the (finite or infinite) past.[6]

This postulate asserts that the geodesics, which represent the world lines of galaxies, do not intersect, except at a singular point in the finite, or infinite, past. Intriguingly, Weyl introduced this postulate *before* Hubble's discovery of the recession of the nebulae. By the term 'substratum', Bondi meant an imaginary medium which can be thought of as a fluid which defines the overall kinematics of the system of galaxies. A consequence of Weyl's postulate is that there is only one geodesic passing through each point in space-time, except at the origin. Once this postulate is adopted, a notional observer, known as a *fundamental observer*, can be assigned to each world line. Each fundamental observer carries a standard clock and time measured on that clock from the singular point is called *cosmic time*. This procedure resolves the problem of synchronising the clocks of fundamental observers.

One further assumption is needed, which is known as the *cosmological principle*:

> We are not located at any special location in the Universe.

In other words, we are located at a *typical* position in the Universe and any other fundamental observer located anywhere in the Universe at the same cosmic epoch would observe the same large-scale features. Thus, we assert that every fundamental observer at the same cosmic epoch observes the same Hubble expansion of the distribution of galaxies, the same isotropic CMB, the same large-scale spongy structure in the distribution of galaxies and voids, and so on.

The *Robertson–Walker metric* is the metric, in the sense of special relativity, of all isotropically expanding universes. Because of the observed large-scale isotropy of the Universe, the starting point is to consider models in which we smooth out all structure, in other words, a uniformly expanding isotropic, homogeneous Universe. Since the curvature of space is an intrinsic property of space-time, it follows that the curvature of space must also be isotropic at any epoch and so the geometry must be one of the isotropic curved spaces discussed in Section 19.3.

20.3.1 Isotropic Curved Spaces Again

The key result derived in Section 19.3 was that all possible isotropic curved spaces are described by a spatial increment of the form

$$dl^2 = dx^2 + R_c^2 \sin^2\left(\frac{x}{R_c}\right)\left[d\theta^2 + \sin^2\theta \, d\phi^2\right], \tag{20.7}$$

in terms of the spherical polar coordinates (x, θ, ϕ), where R_c is real in closed spherical geometries, imaginary in open, hyperbolic geometries and infinite in flat Euclidean geometry. If instead the spatial increment is written in terms of the angular diameter distance ρ,

$$dl^2 = \frac{d\rho^2}{1 - \kappa\rho^2} + \rho^2\left[d\theta^2 + \sin^2\theta \, d\phi^2\right], \tag{20.8}$$

where $\kappa = R_c^2$ is the curvature of the three-space. The notation for the 'angular diameter distance' has been changed from r to ρ for reasons which will become apparent in a moment – ρ will only have this meaning in this section. Recall that the distance coordinate x describes the *metric* or *geodesic distance* in the radial direction, while the ρ coordinate ensures that distances perpendicular to the radial direction are given by $dl = \rho \, d\theta$.

20.3.2 The Robertson–Walker Metric

Cosmic time is defined to be *proper time* as measured by a fundamental observer with a standard clock. According to the Weyl hypothesis, all the clocks can be synchronised by assuming that the world lines of all fundamental observers diverged from some point in the distant past and we set $t = 0$ at that origin.

The *Minkowski metric* for any isotropic three-space is then given by

$$ds^2 = dt^2 - \frac{1}{c^2}dl^2, \tag{20.9}$$

where dl is then given by either of the above forms of the spatial increment and t is cosmic time.

There is a complication in using the metric (20.9) with dl^2 given by (20.7) or (20.8), which is illustrated by the space-time diagram shown in Fig. 20.5. Since light travels at a finite speed, all astronomical objects are observed along the *past light cone* of a fundamental observer, which in our case is centred on the Earth at the present epoch t_0. Therefore, distant objects are observed at earlier epochs than the present, when the Universe was homogeneous and isotropic, but the distances between fundamental observers were smaller and the spatial curvature different. The problem is that we can only apply the metric (20.9) to an isotropic curved space *at a single epoch*.

To derive a distance measure which can be included in the metric, we carry out a thought experiment in which we line up a set of fundamental observers between the Earth and a

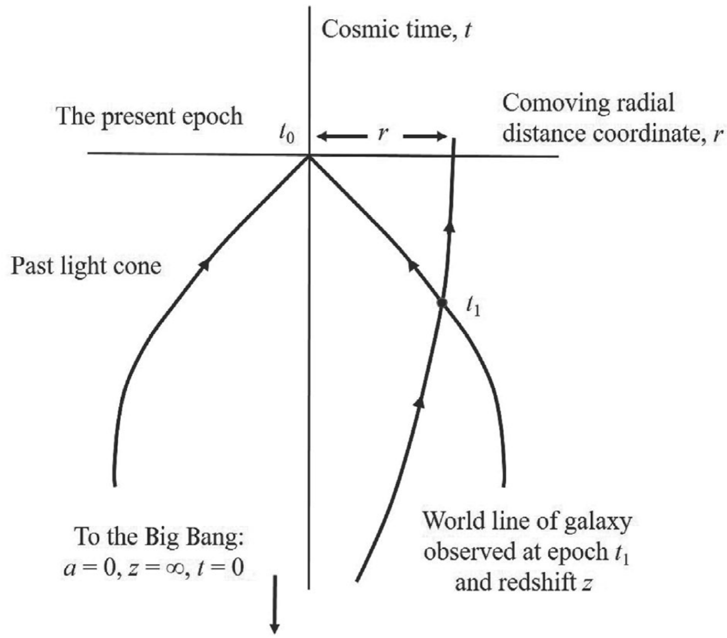


Fig. 20.5 A space-time diagram showing the past light cone along which a fundamental observer observes all objects in the Universe.

distant galaxy. The observers are all instructed to measure the geodesic distance dx to the next observer *at a fixed cosmic time t*, which they read on their own clocks. By adding together all the dxs, a geodesic distance x at a single epoch is obtained which can be used in the metric (20.9). Notice that x is a *fictitious distance*. Distant galaxies can only be observed as they were at some epoch earlier than the present and we do *not* know how to project their positions relative to us forward to the present epoch without a knowledge of the kinematics of the expanding Universe. In other words, *the distance measure x depends upon the choice of cosmological model*.

As discussed in Section 20.2.3, the definition of a uniform expansion is that between two cosmic epochs, t_1 and t_2, the distances of any two fundamental observers, i and j, change such that

$$\frac{x_i(t_1)}{x_j(t_1)} = \frac{x_i(t_2)}{x_j(t_2)} = \text{constant}, \tag{20.10}$$

that is,

$$\frac{x_i(t_1)}{x_i(t_2)} = \frac{x_j(t_1)}{x_j(t_2)} = \cdots = \text{constant} = \frac{a(t_1)}{a(t_2)}. \tag{20.11}$$

The function $a(t)$ is defined to be the *scale factor* which describes how the relative distances between any two fundmental observers change with cosmic time t. We set $a(t) = 1$ at the present epoch t_0 and let the value of x at t_0 be r. Therefore,

$$x(t) = a(t)\, r. \tag{20.12}$$

r therefore becomes a *distance label* attached to a galaxy for all time and the variation of proper distance between fundamental observers in the expanding Universe is taken care of by the scale factor $a(t)$; r is called the *comoving radial distance coordinate*.

Proper distances perpendicular to the line of sight must also change by a factor a between the epochs t and t_0, because of the isotropy of the world model,

$$\frac{\Delta l(t)}{\Delta l(t_0)} = a(t), \tag{20.13}$$

and hence, from the metric (20.7),

$$a(t) = \frac{R_c(t)\sin\left[\dfrac{x}{R_c(t)}\right]\mathrm{d}\theta}{R_c(t_0)\sin\left[\dfrac{r}{R_c(t_0)}\right]\mathrm{d}\theta} \quad \text{or} \quad \frac{R_c(t)}{a(t)}\sin\left[\frac{a(t)r}{R_c(t)}\right] = R_c(t_0)\sin\left[\frac{r}{R_c(t_0)}\right].$$

This is only true if

$$R_c(t) = R_c(t_0)\,a(t), \tag{20.14}$$

that is, the radius of curvature of the spatial sections is proportional to $a(t)$. In order to preserve isotropy and homogeneity, *the curvature of space changes as the Universe expands as* $\kappa = R_c^{-2}(t) \propto a^{-2}$.

To simplify the notation, we define the radius of curvature of the spatial geometry at the present epoch, $R_c(t_0) \equiv \mathcal{R}$ and then

$$R_c(t) = \mathcal{R}\,a(t). \tag{20.15}$$

Substituting (20.12) and (20.15) into the metric (20.9), we obtain the result we have been seeking:

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \frac{a^2(t)}{c^2}\left[\mathrm{d}r^2 + \mathcal{R}^2\sin^2\left(\frac{r}{\mathcal{R}}\right)\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right)\right]. \tag{20.16}$$

This is the *Robertson–Walker metric* in the form we will use in analysing observations in cosmology. It contains one unknown function $a(t)$, the scale factor, and an unknown constant, the curvature of space at the present epoch $\kappa = \mathcal{R}^{-2}$.

The metric can be written in different ways. If we use a *comoving angular diameter distance* $r_1 = \mathcal{R}\sin(r/\mathcal{R})$, the metric becomes

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \frac{a^2(t)}{c^2}\left[\frac{\mathrm{d}r_1^2}{1-\kappa r_1^2} + r_1^2\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right)\right], \tag{20.17}$$

where $\kappa = 1/\mathcal{R}^2$. By a suitable rescaling of the r_1 coordinate $\kappa r_1^2 = r_2^2$, the metric can equally well be written

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \frac{a_1^2(t)}{c^2}\left[\frac{\mathrm{d}r_2^2}{1-kr_2^2} + r_2^2\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right)\right], \tag{20.18}$$

with $k = +1, 0$ and -1 for universes with spherical, flat and hyperbolic geometries respectively. In this rescaling, the value of $a_1(t_0)$ at the present epoch is $\mathcal{R}$ and not unity.

The Robertson–Walker metric enables us to define the invariant interval $\mathrm{d}s^2$ between events at any epoch or location in the expanding Universe, noting:

- t is cosmic time;
- $a(t)\,\mathrm{d}r$ is the element of proper (or geodesic) distance in the radial direction;
- $a(t)[\mathcal{R}\sin(r/\mathcal{R})]\,\mathrm{d}\theta = a(t)r_1\,\mathrm{d}\theta$ is the element of proper distance perpendicular to the radial direction subtended by the angle $\mathrm{d}\theta$ at the origin;
- Similarly, $a(t)[\mathcal{R}\sin(r/\mathcal{R})]\sin\theta\,\mathrm{d}\phi = a(t)r_1\sin\theta\,\mathrm{d}\phi$ is the element of proper distance in the ϕ-direction.

Note the following features of the above arguments:

- The metrics (20.16) to (20.18) have been derived using only special relativity and the postulates of isotropy and homogeneity – there is nothing explicitly general relativistic about the argument.
- We have specified nothing about the physics which determines the dynamics of the expanding Universe – this has all been absorbed into the function $a(t)$.

20.4 Observations in Cosmology

Many useful results relating the intrinsic properties of distant objects to their observed properties are independent of the specific cosmological model. It is useful to provide a list of results, which can be applied to the cosmology of the reader's choice.

20.4.1 Redshift

Consider a wave packet of frequency ν_1 emitted between cosmic times t_1 and $t_1 + \Delta t_1$ from a distant galaxy. This wave packet is received by the observer at the present epoch in the cosmic time interval t_0 to $t_0 + \Delta t_0$. The signal propagates along null cones, $\mathrm{d}s^2 = 0$ and so, considering radial propagation from source to observer, $\mathrm{d}\theta = 0, \mathrm{d}\phi = 0$, the metric (20.16) reduces to

$$\mathrm{d}t = -\frac{a(t)}{c}\mathrm{d}r, \qquad \frac{c\,\mathrm{d}t}{a(t)} = -\mathrm{d}r. \tag{20.19}$$

The minus sign appears because the origin of the r-coordinate is at the observer. Considering first the leading edge of the wave packet, integrate (20.19) from the source to observer:

$$\int_{t_1}^{t_0} \frac{c\,\mathrm{d}t}{a(t)} = -\int_r^0 \mathrm{d}r. \tag{20.20}$$

The end of the wave packet must travel the same radial comoving coordinate distance, since the r-coordinate is fixed to the galaxy for all time. Therefore,

$$\int_{t_1+\Delta t_1}^{t_0+\Delta t_0} \frac{c\,\mathrm{d}t}{a(t)} = -\int_r^0 \mathrm{d}r.$$

Hence,

$$\int_{t_1}^{t_0} \frac{c\,\mathrm{d}t}{a(t)} + \frac{c\,\Delta t_0}{a(t_0)} - \frac{c\,\Delta t_1}{a(t_1)} = \int_{t_1}^{t_0} \frac{c\,\mathrm{d}t}{a(t)}. \tag{20.21}$$

Since $a(t_0) = 1$,

$$\Delta t_0 = \frac{\Delta t_1}{a(t_1)}. \tag{20.22}$$

This is the cosmological expression for the phenomenon of *time dilation*. When distant galaxies are observed, $a(t_1) < 1$, and so time intervals are observed to be longer in our frame of reference than they are in that of the source.

This result also provides directly an expression for the *redshift*. If $\Delta t_1 = \nu_1^{-1}$ is the period of the emitted waves and $\Delta t_0 = \nu_0^{-1}$ the observed period,

$$\nu_0 = \nu_1\, a(t_1). \tag{20.23}$$

Rewriting (20.23) in terms of the redshift z,

$$z = \frac{\lambda_0 - \lambda_e}{\lambda_e} = \frac{\lambda_0}{\lambda_e} - 1 = \frac{\nu_1}{\nu_0} - 1,$$

that is,

$$1 + z = \frac{1}{a(t_1)}. \tag{20.24}$$

Thus, *redshift is a direct measure of the scale factor of the Universe when the source emitted the radiation*. For example, when a galaxy is observed with redshift $z = 1$, the scale factor of the Universe when the light was emitted was $a(t) = 0.5$, that is, the separation between fundamental observers (or galaxies) was half its present value. We obtain no information, however, about *when* the light was emitted. This explains why it is best to dispense with velocities in Hubble's law.

We can now find an expression for the comoving radial distance coordinate r:

$$r = \int_{t_1}^{t_0} \frac{c\,\mathrm{d}t}{a(t)} = \int_{t_1}^{t_0} (1 + z)\, c\, \mathrm{d}t. \tag{20.25}$$

Thus, once we know how $a(t)$ varies with cosmic time t, we can find r by integration. This relation emphasises the somewhat artificial nature of r – we need to know the kinematics of the Universe as described by $a(t)$ between the times when the light was emitted and received before we can work out r.

20.4.2 Hubble's Law

In terms of proper distances, Hubble's law at any epoch can be written $v = H\,x$ and so

$$\frac{\mathrm{d}x}{\mathrm{d}t} = H\,x, \tag{20.26}$$

where H is Hubble's constant at cosmic time t. Substituting $x = a(t)r$,

$$r\frac{\mathrm{d}a(t)}{\mathrm{d}t} = H\,a(t)\,r, \tag{20.27}$$

that is,

$$H = \dot{a}/a. \tag{20.28}$$

At the present epoch, $t = t_0$, $a = 1$ and hence

$$H_0 = (\dot{a})_{t_0}. \tag{20.29}$$

Hubble's constant H_0 describes the present expansion rate of the Universe. Its value changes with cosmic epoch and can be defined at any cosmic time through the more general relation

$$H(t) = \dot{a}/a. \tag{20.30}$$

20.4.3 Angular Diameters

The spatial element of the metric in the θ-direction is a proper length d and is related to the observed angular diameter $\Delta\theta$ by

$$d = a(t) D\,\Delta\theta = \frac{D\Delta\theta}{(1+z)}, \quad \Delta\theta = \frac{d(1+z)}{D}, \tag{20.31}$$

where we have introduced a *distance measure* $D = \mathcal{R}\sin(r/\mathcal{R})$. For small redshifts $z \ll 1$, $r \ll \mathcal{R}$, (20.31) reduces to the Euclidean relation $d = r\,\Delta\theta$. Equation (20.31) can also be written

$$\Delta\theta = \frac{d}{D_{\rm A}}, \tag{20.32}$$

so that the relation between d and $\Delta\theta$ looks like the standard Euclidean relation. To do this, we introduce the distance measure $D_{\rm A} = D/(1+z)$, which is known as an *angular diameter distance* and which is often used in the literature.

20.4.4 Apparent Intensities

Suppose a source at redshift z has luminosity $L(\nu_1)$ measured in W Hz^{-1}, that is, the total energy emitted over 4π steradians per unit time per unit frequency interval. What is the flux density $S(\nu_0)$ of the source at the observing frequency ν_0, by which we mean the energy received per unit time, per unit area and per unit bandwidth (W m^{-2} Hz^{-1}), where $\nu_0 = a(t_1)\nu_1 = \nu_1/(1+z)$?

Suppose the source emits $N(\nu_1)$ photons of energy $h\nu_1$ in the bandwidth $\Delta\nu_1$ in the proper time interval Δt_1. Then, the luminosity of the source is

$$L(\nu_1) = \frac{N(\nu_1)\,h\nu_1}{\Delta\nu_1\,\Delta t_1}. \tag{20.33}$$

These photons are distributed over a 'sphere' centred on the source at epoch t_1, and when this 'shell' of photons arrives at the observer at the present epoch t_0, a certain fraction of them is intercepted by the telescope. The photons are observed at the present epoch t_0 with frequency $\nu_0 = a(t_1)\nu_1$ in a proper time interval $\Delta t_0 = \Delta t_1/a(t_1)$ and in the waveband $\Delta\nu_0 = a(t_1)\,\Delta\nu_1$.

We also need to relate the diameter of our telescope Δl to the angular diameter which it subtends at the source at epoch t_1. The metric (20.16) provides the answers. The proper distance Δl refers to the present epoch at which $a(t_0) = 1$ and hence

$$\Delta l = D\,\Delta\theta, \tag{20.34}$$

where $\Delta\theta$ is the angle measured by a fundamental observer located at the source. Notice the difference between the results (20.31) and (20.34). They are angular diameters measured in opposite directions along the light cone. The factor of $(1+z)$ difference between them is part of a more general relation concerning angular diameter measures along light cones, which is known as the *reciprocity theorem*.

Therefore, the surface area of the telescope is $\pi\Delta l^2/4$ and the solid angle subtended by this area at the source is $\Delta\Omega = \pi\Delta\theta^2/4$. The number of photons incident upon the telescope in time Δt_0 is therefore $N(\nu_1)\,\Delta\Omega/4\pi$, but they are observed at frequency ν_0 in the waveband $\Delta\nu_0$. Therefore, the flux density of the source, that is, the energy received per unit time, per unit area and per unit bandwidth, is

$$S(\nu_0) = \frac{N(\nu_1)h\nu_0\,\Delta\Omega}{4\pi\,\Delta t_0\,\Delta\nu_0(\pi/4)\Delta l^2}. \tag{20.35}$$

We can now relate the quantities in (20.35) to the properties of the source using the relations (20.22), (20.23) and (20.33),

$$S(\nu_0) = \frac{L(\nu_1)\,a(t_1)}{4\pi D^2} = \frac{L(\nu_1)}{4\pi D^2(1+z)}, \tag{20.36}$$

where the distance measure $D = \mathcal{R}\sin(r/\mathcal{R})$. If the spectra of sources are of power-law form, $L(\nu) \propto \nu^{-\alpha}$, this relation becomes

$$S(\nu_0) = \frac{L(\nu_0)}{4\pi D^2(1+z)^{1+\alpha}}. \tag{20.37}$$

We can repeat the analysis for *bolometric luminosities* and *bolometric flux densities*. In this case, we consider the total energy emitted in a finite bandwidth $\Delta\nu_1$ which is received in the bandwidth $\Delta\nu_0$, that is

$$L_{\rm bol} = L(\nu_1)\,\Delta\nu_1 = 4\pi D^2(1+z)^2 S_{\rm bol}, \tag{20.38}$$

where the bolometric flux density is $S_{\rm bol} = S(\nu_0)\Delta\nu_0$. Therefore,

$$S_{\rm bol} = \frac{L_{\rm bol}}{4\pi D^2(1+z)^2} = \frac{L_{\rm bol}}{4\pi D_{\rm L}^2}. \tag{20.39}$$

The quantity $D_{\rm L} = D(1+z)$ is called the *luminosity distance*, since this definition makes the relation between $S_{\rm bol}$ and $L_{\rm bol}$ look like an inverse square law.

The bolometric luminosity can be integrated over any suitable bandwidth, so long as the corresponding redshifted bandwidth is used to measure the bolometric flux density at the present epoch, that is,

$$\sum_{\nu_0} S(\nu_0)\,\Delta\nu_0 = \frac{\sum_{\nu_1} L(\nu_1)\,\Delta\nu_1}{4\pi D^2(1+z)^2} = \frac{\sum_{\nu_1} L(\nu_1)\,\Delta\nu_1}{4\pi D_{\mathrm{L}}^2}. \tag{20.40}$$

We can rewrite the expression for the flux density of the source $S(\nu_0)$ in terms of the luminosity of the source at the observing frequency ν_0 as follows:

$$S(\nu_0) = \frac{L(\nu_0)}{4\pi D_{\mathrm{L}}^2}\left[\frac{L(\nu_1)}{L(\nu_0)}(1+z)\right]. \tag{20.41}$$

The last term in square brackets is known as the *K-correction*. It was introduced by the pioneer optical cosmologists in the 1930s in order to 'correct' the apparent magnitude of distant galaxies for the effects of redshifting the spectrum, when observations are made at a fixed observing frequency ν_0 and in a fixed bandwidth $\Delta\nu_0$. Taking logarithms and multiplying by -2.5, the terms in square brackets can be converted into a correction to the apparent magnitude of the galaxy:

$$K(z) = -2.5\log_{10}\left[\frac{L(\nu_1)}{L(\nu_0)}(1+z)\right]. \tag{20.42}$$

Notice that this form of the K-correction is correct for *monochromatic* flux densities and luminosities. In the case of observations in the optical waveband, in which magnitudes are measured through broad standard filters, averages have to be taken over the spectral energy distributions of the objects within the appropriate spectral windows in the emitted and observed wavebands. Evidently, the K-correction depends upon knowing the spectra of the sources.

20.4.5 Number Densities

The numbers of sources in the redshift range, z to $z + \mathrm{d}z$ can be readily found because, by definition, r is a comoving radial distance coordinate defined at the present epoch, and hence the number we need is just the number of objects in the interval r to $r + \mathrm{d}r$ at the present day. The volume of a spherical shell of thickness $\mathrm{d}r$ at comoving distance coordinate r is

$$\mathrm{d}V = 4\pi\mathcal{R}^2\sin^2(r/\mathcal{R})\,\mathrm{d}r = 4\pi D^2\,\mathrm{d}r. \tag{20.43}$$

Therefore, if N_0 is the present space density of objects and their number is conserved as the Universe expands,

$$\mathrm{d}N = 4\pi N_0 D^2\,\mathrm{d}r. \tag{20.44}$$

The definition of comoving radial distance coordinate automatically takes account of the expansion of the system of galaxies.

20.4.6 The Age of the Universe

The differential relation (20.19) is needed to work out the age of the Universe T_0,

$$-\frac{c\,\mathrm{d}t}{a(t)} = \mathrm{d}r. \tag{20.45}$$

Therefore,

$$T_0 = \int_0^{t_0} \mathrm{d}t = \int_0^{r_{\max}} \frac{a(t)\,\mathrm{d}r}{c}, \tag{20.46}$$

where $r_{\max}$ is the comoving distance coordinate corresponding to $a = 0, z = \infty$.

20.4.7 Summary

The above results can be used to work out the relations between the intrinsic properties of galaxies and observables for any isotropic, homogeneous world model. Let us summarise the procedure:

(1) First work out from theory the function $a(t)$ and the curvature of the space at the present epoch $\kappa = \mathcal{R}^{-2}$.

(2) Next, work out the *comoving radial distance coordinate* r as a function of redshift z using the integral

$$r = \int_{t_1}^{t_0} \frac{c\,\mathrm{d}t}{a(t)}. \tag{20.47}$$

Notice what this expression means – the proper distance interval $c\,\mathrm{d}t$ at epoch t is projected forward to the present epoch t_0 by the scale factor $a(t)$.

(3) Now, find the *distance measure* D as a function of redshift z,

$$D = \mathcal{R} \sin \frac{r}{\mathcal{R}}. \tag{20.48}$$

(4) If so desired, the *luminosity distance* $D_{\mathrm{L}} = D(1+z)$ and *angular diameter distance* $D_{\mathrm{A}} = D/(1+z)$ can be introduced.

(5) The number of objects $\mathrm{d}N$ in the redshift interval $\mathrm{d}z$ and solid angle Ω can be found from the expression

$$\mathrm{d}N = \Omega N_0 D^2\,\mathrm{d}r, \tag{20.49}$$

where N_0 is the number density of objects *at the present epoch*.

20.5 The Standard World Models

In 1917, Einstein realised that, in general relativity, he had discovered a theory which enabled fully self-consistent models for the Universe as a whole to be constructed for the first time.[7] The simplest models are those in which it is assumed that the Universe is isotropic, homogeneous and uniformly expanding. These assumptions result in an enormous simplification of Einstein's field equations. The standard textbooks show that they reduce to

$$\ddot{a} = -\frac{4\pi G a}{3}\left(\rho + \frac{3p}{c^2}\right) + \tfrac{1}{3}\Lambda a, \tag{20.50}$$

$$\dot{a}^2 = \frac{8\pi G\rho}{3}a^2 - \frac{c^2}{\mathcal{R}^2} + \tfrac{1}{3}\Lambda a^2, \tag{20.51}$$

where a is the scale factor, ρ is the inertial mass density of the matter and radiation content of the Universe and p their pressure. Notice that the pressure term in (20.50) is a relativistic correction to the inertial mass density. Unlike normal pressure forces, which depend upon the gradient of the pressure, this term depends linearly on the pressure. $\mathcal{R}$ is the radius of curvature of the geometry of the world model at the present epoch, and so the term $-c^2/\mathcal{R}^2$ in (20.51) is a constant of integration.

The *cosmological constant* Λ has been included in (20.50) and (20.51). This term had a chequered history. It was introduced by Einstein in 1917 in order to produce static solutions of the field equations, more than 10 years before Hubble showed that, in fact, the Universe is not static. We first discuss the models with $\Lambda = 0$.

20.5.1 The Standard Dust Models: The Friedman World Models with $\Lambda = 0$

By *dust*, cosmologists mean a pressureless fluid, $p = 0$. It is convenient to refer the density of matter to its value at the present epoch ρ_0. Because of conservation of mass, $\rho = \rho_0\, a^{-3}$ and so (20.50) and (20.51) reduce to

$$\ddot{a} = -\frac{4\pi G\rho_0}{3}a^{-2}, \tag{20.52}$$

$$\dot{a}^2 = \frac{8\pi G\rho_0}{3}a^{-1} - \frac{c^2}{\mathcal{R}^2}. \tag{20.53}$$

In 1934, Edward Arthur Milne and William McCrea showed how these relations can be derived using Newtonian dynamics. Consider a galaxy at distance x from the Earth (Fig. 20.6). Because of the isotropy of the Universe, its deceleration is due to the gravitational attraction of the matter inside the uniform sphere of radius x and density ρ centred on the Earth. According to Gauss's theorem for gravity, we can replace the mass $M = (4\pi/3)\rho x^3$ by a point mass at the centre of the sphere. Therefore, the deceleration of the galaxy is

$$m\ddot{x} = -\frac{GMm}{x^2} = -\frac{4\pi x\rho m}{3}. \tag{20.54}$$

The mass of the galaxy m cancels out on either side of the equation, showing that the deceleration refers to the dynamics of the Universe as a whole, rather than to any particular galaxy. Now, replace x by the comoving radial distance coordinate r, $x = ar$, and express the density in terms of its value at the present epoch, $\rho = \rho_0\, a^{-3}$. Therefore,

$$\ddot{a} = -\frac{4\pi G\rho_0}{3}\frac{1}{a^2}, \tag{20.55}$$

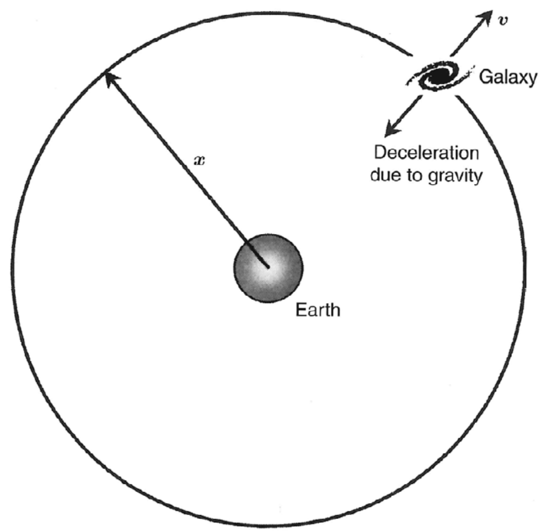


Fig. 20.6 Illustrating the dynamics of isotropic world models. A galaxy at distance *x* is decelerated by the gravitational attraction of the matter within distance *x* of our Galaxy. Because of the assumption of isotropy, a fundamental observer on any galaxy participating in the uniform expansion would carry out the same calculation.

which is identical to equation (20.52). Multiplying equation (20.55) by $\dot{a}$ and integrating,

$$\dot{a}^2 = \frac{8\pi G\rho_0}{3}a^{-1} + \text{constant}. \qquad (20.56)$$

This result is identical to (20.53), if we identify the constant with $-c^2/\mathcal{R}^2$.

This analysis illustrates a number of important features of the world models of general relativity. First of all, note that, because of the assumption of isotropy, local physics is also global physics, which is why the simple Newtonian argument works. The physics which defines the local behaviour of matter also defines its behaviour on the largest scales. For example, the curvature of space within one cubic metre is exactly the same as that on the scale of the Universe itself. A second point is that, although we might appear to have placed the Earth in a rather special position, a fundamental observer located on any galaxy would perform exactly the same calculation.

Third, notice that at no point in the argument did we ask over what physical scale the calculation was to be valid. In fact, this calculation describes correctly the dynamics of the Universe on scales which are greater than the *horizon scale*, which we take to be $r_\mathrm{H} = ct$, that is, the maximum distance between points which can be causally connected at the epoch t. The reason is the same as above – local physics is also global physics and so, if the Universe were set up in such a way that it had uniform density on scales far exceeding the horizon scale, the dynamics on these very large scales would be exactly the same as the local dynamics.

The solutions of Einstein's field equations were presented by Aleksander Friedman in two papers published in 1922 and 1924.[8] Georges Lemaître rediscovered Friedman's solutions in 1927,[9] and brought them to the notice of astronomers and cosmologists during the 1930s. The solutions of the field equations are appropriately referred to as the *Friedman models* of the Universe.

It is convenient to express the mass densities of the world models relative to the *critical density* ρ_c, which is defined to be $\rho_c = (3H_0^2/8\pi G)$, and then to refer the density of any specific model at the present epoch ρ_0 to this value through a *density parameter* $\Omega_0 = \rho_0/\rho_c$. Thus,

$$\Omega_0 = \frac{8\pi G\rho_0}{3H_0^2}. \tag{20.57}$$

The subscript 0 has been attached to Ω because the critical density ρ_c changes with cosmic epoch, as does ρ. (20.53) therefore becomes

$$\dot{a}^2 = \frac{\Omega_0 H_0^2}{a} - \frac{c^2}{\mathcal{R}^2}. \tag{20.58}$$

Several important results can be deduced from (20.58). At the present epoch, $t = t_0, a = 1$,

$$\mathcal{R} = \frac{c/H_0}{(\Omega_0 - 1)^{1/2}} \qquad \text{and so} \qquad \kappa = \frac{(\Omega_0 - 1)}{(c/H_0)^2}. \tag{20.59}$$

The expressions (20.59) show that there is a one-to-one relation between the density of the Universe and its spatial curvature for the standard world models with $\Lambda = 0$. The solutions of (20.58) shown in Fig. 20.7 are the well-known relations between the dynamics and geometry of the Friedman world models, namely:

(1) The models with $\Omega_0 > 1$ have closed, spherical geometry and they collapse to an infinite density in a finite time, an event sometimes referred to as the 'Big Crunch';
(2) The models having $\Omega_0 < 1$ have open, hyperbolic geometries and expand forever. They would reach infinity with a finite recession velocity.

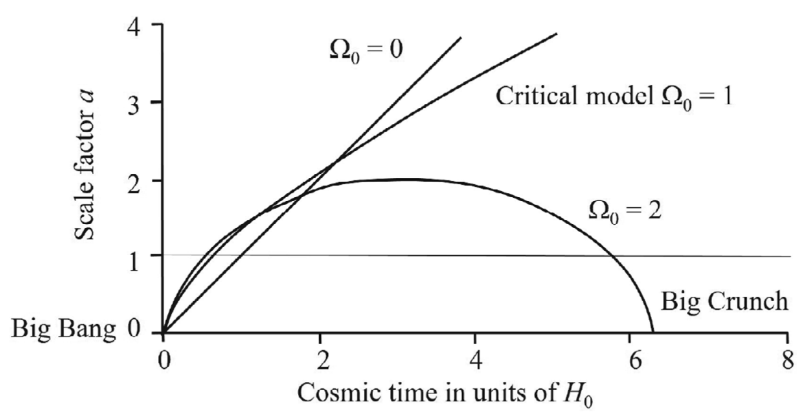


Fig. 20.7 Comparison of the dynamics of different Friedman models with $\Lambda = 0$ characterised by the density parameter $\Omega_0 = \rho/\rho_0$. Cosmic time is measured in units of $H_0 t$, where H_0 is Hubble's constant and the scale factor $a = 1$ at the present epoch. In this format, all the models have the same value of Hubble's constant at the present day.

(3) The model with $\Omega_0 = 1$ separates the open from the closed models and the collapsing models from those which expand forever. This model is often referred to as the *Einstein–de Sitter model* or the *critical model*. The velocity of expansion tends to zero as a tends to infinity. It has a particularly simple variation of $a(t)$ with cosmic epoch,

$$a = \left(\frac{3H_0 t}{2}\right)^{2/3}, \quad \kappa = 0. \tag{20.60}$$

Another useful result is the function $a(t)$ for the empty world model, $\Omega_0 = 0$, $a(t) = H_0 t$, $\kappa = -(H_0/c)^2$. This model is sometimes referred to as the *Milne model*. It is an interesting exercise to show why it is that, in an empty world model, the global geometry of the Universe is hyperbolic (Section 20.5.2).

The introduction of the critical density ρ_c and the density parameter Ω_0 is conveniently extended to describe the relative importance of different contributions to the overall mass–energy density of the Universe. In due course we will introduce density parameters to describe the average mass density in baryons Ω_b, in the dark matter Ω_d and the dark energy Ω_v, as well as a curvature density parameter Ω_K. In the present instance, the value of $\Omega_0 = \Omega_b + \Omega_d$.

The general solutions of (20.53) can be written in parametric form:

$$\text{For } \Omega_0 > 1, \quad a = a(1 - \cos\theta), \qquad t = b(\theta - \sin\theta), \tag{20.61}$$

$$\text{For } \Omega_0 < 1, \quad a = a(\cosh\phi - 1), \qquad t = b(\sinh\phi - \phi), \tag{20.62}$$

where

$$\text{for } \Omega_0 > 1, \quad a = \frac{\Omega_0}{2(\Omega_0 - 1)} \quad \text{and} \quad b = \frac{\Omega_0}{2H_0(\Omega_0 - 1)^{3/2}}, \tag{20.63}$$

$$\text{for } \Omega_0 < 1, \quad a = \frac{\Omega_0}{2(1 - \Omega_0)} \quad \text{and} \quad b = \frac{\Omega_0}{2H_0(1 - \Omega_0)^{3/2}}. \tag{20.64}$$

All the models tend towards the dynamics of the critical model at early times, but with a different constant, that is, for $\theta \ll 1, \phi \ll 1$,

$$a = \Omega_0^{1/3}\left(\frac{3H_0 t}{2}\right)^{2/3}. \tag{20.65}$$

Just as Hubble's constant H_0 measures the local expansion rate of the distribution of galaxies, so we can define the local deceleration of the Universe at the present epoch, $\ddot{a}(t_0)$. It is conventional to define the *deceleration parameter* q_0 to be the dimensionless deceleration at the present epoch through the expression

$$q_0 = -\left(\frac{\ddot{a}}{\dot{a}^2}\right)_{t_0}. \tag{20.66}$$

Substituting the values $a = 1$ and $\dot{a}(t_0) = H_0$ at the present epoch into (20.52), we find that the deceleration parameter q_0 is directly proportional to the density parameter Ω_0,

$$q_0 = \Omega_0/2, \tag{20.67}$$

provided the cosmological constant Λ is zero.

An important result for many aspects of cosmology is the relation between redshift z and cosmic time t. From (20.58) and (20.59), and recalling that $a = (1+z)^{-1}$, we find

$$\frac{dz}{dt} = -H_0(1+z)^2(\Omega_0 z + 1)^{1/2}. \quad (20.68)$$

Cosmic time t measured from the Big Bang follows immediately by integration:

$$t = \int_0^t dt = -\frac{1}{H_0}\int_\infty^z \frac{dz}{(1+z)^2(\Omega_0 z + 1)^{1/2}}. \quad (20.69)$$

It is a straightforward calculation to evaluate this integral to find the present age of the Universe T_0:

$$T_0 = \frac{\Omega_0}{H_0(\Omega - 1)^{3/2}}\left[\sin^{-1}\left(\frac{\Omega_0 - 1}{\Omega_0}\right)^{1/2} - \frac{(\Omega_0 - 1)^{1/2}}{\Omega_0}\right], \qquad \text{if } \Omega_0 > 1.$$

$$T_0 = \frac{2}{3H_0}, \qquad \text{if } \Omega_0 = 1.$$

$$T_0 = \frac{\Omega_0}{H_0(1 - \Omega_0)^{3/2}}\left[\frac{(1 - \Omega_0)^{1/2}}{\Omega_0} - \sinh^{-1}\left(\frac{1 - \Omega_0}{\Omega_0}\right)^{1/2}\right], \qquad \text{if } \Omega_0 < 1.$$

For the critical model $\Omega_0 = 1$, the present age of the Universe is $T_0 = (2/3)H_0^{-1}$ and for the empty model, $\Omega_0 = 0$, $T_0 = H_0^{-1}$. Hence, for all world models with $\Lambda = 0$, their ages are $T_0 \le H_0^{-1}$.

Just as it is possible to define Hubble's constant at any epoch by $H = \dot{a}/a$, so we can define a density parameter Ω at any epoch through the definition $\Omega = 8\pi G\rho/3H^2$. Since $\rho = \rho_0(1+z)^3$, it follows that

$$\Omega H^2 = \frac{8\pi G}{3}\rho_0(1+z)^3. \quad (20.70)$$

It is a straightforward exercise to show that this relation can be rewritten

$$\left(1 - \frac{1}{\Omega}\right) = (1+z)^{-1}\left(1 - \frac{1}{\Omega_0}\right). \quad (20.71)$$

This important result shows that, whatever the value of Ω_0 now, Ω tends to the value 1 in the distant past, because $(1+z)^{-1}$ becomes very small at large redshifts. If the value of Ω were significantly different from 1 in the very distant past, (20.71) shows that it would be very different indeed from 1 now. Our Universe must have been *very finely tuned indeed* to the value $\Omega = 1$ in the distant past, if it is to end up close to $\Omega_0 = 1$ at the present epoch. This *flatness problem* is taken up in Chapter 21.

We can now find expressions for the comoving radial coordinate distance r and the 'distance measure' D. We recall that

$$dr = \frac{c\,dt}{a(t)} = -c\,dt(1+z) = \frac{c\,dz}{H_0(1+z)(\Omega_0 z + 1)^{1/2}}. \quad (20.72)$$

Integrating from redshifts 0 to z, the expression for r is

$$r = \frac{2c}{H_0(\Omega_0 - 1)^{1/2}}\left[\tan^{-1}\left(\frac{\Omega_0 z + 1}{\Omega_0 - 1}\right)^{1/2} - \tan^{-1}(\Omega_0 - 1)^{-1/2}\right]. \quad (20.73)$$

Finally, D is found by evaluating $D = \mathcal{R} \sin(r/\mathcal{R})$, where $\mathcal{R}$ is given by the expression (20.59). It is a straightforward calculation to show that

$$D = \frac{2c}{H_0 \Omega_0^2 (1+z)} \left\{ \Omega_0 z + (\Omega_0 - 2) \left[(\Omega_0 z + 1)^{1/2} - 1 \right] \right\} . \qquad (20.74)$$

This formula was first derived by Mattig in 1958.[10] Although it has been derived for spherical geometry, it has the advantage of being correct for all values of Ω_0. It can be used in all the expressions listed in Section 20.4 for relating intrinsic properties to observables.

20.5.2 Pedagogical Interlude: The Robertson–Walker Metric for an Empty Universe

The empty world model, $\Omega_0 = 0, \Lambda = 0$, often referred to as the *Milne model*, is an interesting case since it illustrates why care has to be taken in setting up the mathematical structure of the cosmological models. There is no gravity present and so the space-time metric should be derivable from special relativity.

The origin of the uniform expansion is taken to be $[0, 0, 0, 0]$ and the world lines of particles diverge from this point, each point maintaining constant velocity with respect to the others. The space-time diagram for this universe is shown in Fig. 20.8. Our world line is the t-axis, and that of particle P, which moves with constant velocity v with respect to us, is shown.

The problem becomes apparent as soon as we attempt to define a suitable *cosmic time* for ourselves and for the fundamental observer moving with the particle P. At time t, the distance of P from us is r and, since P's velocity is constant, $r = vt$ in our reference frame. Because of the relativity of simultaneity, the observer at P measures a different time τ.

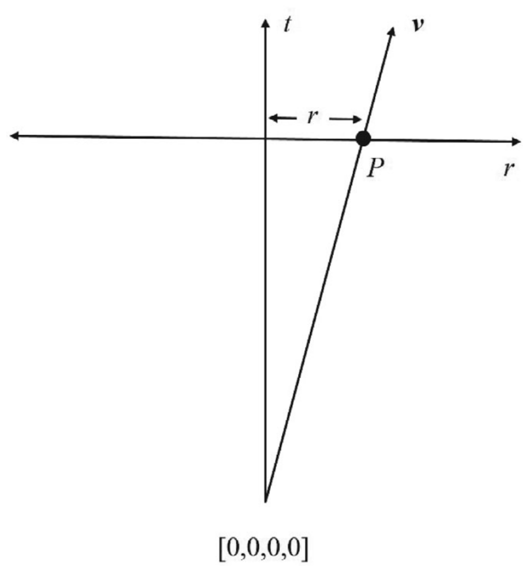


Fig. 20.8 A space-time diagram for an empty universe.

From the Lorentz transformation,

$$\tau = \gamma\left(t - \frac{vr}{c^2}\right), \qquad \gamma = \left(1 - \frac{v^2}{c^2}\right)^{-1/2}.$$

Since $r = vt$,

$$\tau = t\left(1 - \frac{r^2}{c^2 t^2}\right)^{1/2}. \tag{20.75}$$

The problem is that t is proper time for the observer at O but for nobody else. We need to be able to define surfaces of constant cosmic time τ, because it is on these surfaces that we can impose the conditions of isotropy and homogeneity at a given cosmic time, in accordance with the cosmological principle. The appropriate surface for which τ = constant is given by those points which satisfy

$$\tau = t\left(1 - \frac{r^2}{c^2 t^2}\right)^{1/2} = \text{constant}. \tag{20.76}$$

At each point in the space, this surface is normal to the world line of a fundamental observer.

The next requirement is to define the local element of radial distance $\mathrm{d}l$ at the point P on the surface τ = constant. The interval $\mathrm{d}s^2 = \mathrm{d}t^2 - (1/c^2)\,\mathrm{d}r^2$ is an invariant. Since τ = constant over the surface, $\mathrm{d}s^2 = -(1/c^2)\,\mathrm{d}l^2$ and hence

$$\mathrm{d}l^2 = \mathrm{d}r^2 - c^2\,\mathrm{d}t^2. \tag{20.77}$$

τ and $\mathrm{d}l$ define locally the proper time and proper distance of events at P.

Now let us transform from the frame S to that at P by moving at radial velocity v. Distances perpendicular to the radial coordinate remain unaltered under Lorentz transformation and therefore, if in S,

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \frac{1}{c^2}\left(\mathrm{d}r^2 + r^2\,\mathrm{d}\theta^2\right), \tag{20.78}$$

then in S',

$$\mathrm{d}s^2 = \mathrm{d}\tau^2 - \frac{1}{c^2}\left(\mathrm{d}l^2 + r^2\,\mathrm{d}\theta^2\right). \tag{20.79}$$

Now all we need do is to express r in terms of l and τ to complete the transformation to (τ, l) coordinates.

Along the surface of constant τ, $\mathrm{d}l^2 = \mathrm{d}r^2 - c^2\,\mathrm{d}t^2$. In addition, since $\mathrm{d}l$ is measured at fixed time τ in S', the relation between $\mathrm{d}r$ and $\mathrm{d}t$ can be found from the Lorentz transform of $\mathrm{d}\tau$,

$$\mathrm{d}\tau = \gamma\left(\mathrm{d}t - \frac{v}{c^2}\mathrm{d}r\right) = 0.$$

Therefore,

$$\mathrm{d}t^2 = \frac{v^2}{c^4}\,\mathrm{d}r^2, \tag{20.80}$$

and, from (20.77),

$$\mathrm{d}l^2 = \mathrm{d}r^2\left(1 - \frac{v^2}{c^2}\right) = \mathrm{d}r^2\left(1 - \frac{r^2}{c^2t^2}\right) = \mathrm{d}r^2 - \frac{r^2\,\mathrm{d}r^2}{c^2t^2}. \tag{20.81}$$

Notice that this makes sense since, in the rest frame at the point P, the recession speed is the same as observed in S', $v = \mathrm{d}l/\mathrm{d}\tau = \mathrm{d}r/\mathrm{d}t$.

Finally, we substitute for $\mathrm{d}r/t$ in the last term of (20.81) using $\mathrm{d}l/\tau = \mathrm{d}r/t$, so that

$$\mathrm{d}l^2 = \frac{\mathrm{d}r^2}{\left(1 + \dfrac{r^2}{c^2\tau^2}\right)}. \tag{20.82}$$

Integrating, using the substitution $r = c\tau \sinh x$, the solution is

$$r = c\tau\sinh(l/c\tau). \tag{20.83}$$

The metric (20.79) can therefore be written

$$\mathrm{d}s^2 = \mathrm{d}\tau^2 - \frac{1}{c^2}\left[\mathrm{d}l^2 + c^2\tau^2\sinh^2(l/c\tau)\,\mathrm{d}\theta^2\right]. \tag{20.84}$$

This corresponds precisely to the expression (20.9) with $\mathrm{d}l^2$ given by (20.7) for an isotropic curved space with hyperbolic geometry, the radius of curvature of the geometry $\mathcal{R}$ being $c\tau$. This explains why an empty universe has hyperbolic spatial sections. The conditions (20.75) and (20.82) indicate why we can only define a consistent cosmic time and radial distance coordinate in hyperbolic rather than flat space.

20.5.3 Models with Non-zero Cosmological Constant

Einstein's field equation (20.50), including the cosmological constant Λ, for dust-filled universes is

$$\ddot{a} = -\frac{4\pi Ga\rho}{3} + \tfrac{1}{3}\Lambda a = -\frac{4\pi G\rho_0}{3a^2} + \tfrac{1}{3}\Lambda a. \tag{20.85}$$

Inspection of (20.85) gives insight into the meaning of the cosmological constant. Even in an empty universe, $\rho = p = 0$, there is a net force acting on a test particle. In the words of Yakov Zeldovich, if Λ is positive, the term may be thought of as the 'repulsive effect of a vacuum', the repulsion being relative to an absolute geometrical frame of reference. There is no physical interpretation of this term in classical physics.

There is, however, a natural interpretation in the context of quantum field theory. In the modern picture of a vacuum, there are zero-point fluctuations associated with the zero-point energies of all quantum fields. The stress-energy tensor of a vacuum has a negative-pressure equation of state, $p = -\rho c^2$. This pressure may be thought of as a *tension*, rather than a pressure. When such a vacuum expands, the work done $p\,\mathrm{d}V$ in expanding from V to $V + \mathrm{d}V$ is $-\rho c^2\,\mathrm{d}V$ so that, during the expansion, the mass density of the negative pressure field remains constant. Carroll, Press and Turner[11] show how the theoretical value of Λ can be evaluated using simple concepts from quantum field theory. They find the mass density of the repulsive field to be $\rho_{\mathrm{v}} = 10^{95}$ kg m^{-3}, the subscript 'v' referring to the vacuum energy density. This is a serious problem. This mass density is about 10^{120} times greater than that associated with the cosmological constant derived from observation, which correspond to $\rho_v \approx 6 \times 10^{-27}$ kg m^{-3} (see Section 20.9).

This enormous discrepancy should be not be passed over lightly. In the inflationary model of the very early Universe, forces of exactly this type are invoked to cause a rapid exponential expansion. If the inflationary picture is adopted, we have to explain why ρ_v decreased by a factor of about 10^{120} at the end of the inflationary era.

It is encouraging that the Higgs fields associated with the Higgs bosons have negative pressure equations of state, although they cannot be identified with the cosmological constant. The upshot is that it is now quite natural to believe that there may well be forces in nature which provide Zeldovich's 'repulsion of the vacuum', and to associate a certain mass density ρ_v with the energy density of the vacuum at the present epoch. It is convenient to rewrite the dynamical equations in terms of a density parameter Ω_v associated with ρ_v as follows. We begin with (20.50), but include the density ρ_v and pressure p_v of the vacuum fields in place of the Λ term:

$$\ddot{a} = -\frac{4\pi G a}{3}\left(\rho_m + \rho_v + \frac{3p_v}{c^2}\right), \tag{20.86}$$

where the matter density is ρ_m. Since $p_v = -\rho_v c^2$,

$$\ddot{a} = -\frac{4\pi G a}{3}\left(\rho_m - 2\rho_v\right). \tag{20.87}$$

But, as the Universe expands, $\rho_m = \rho_0/a^3$, ρ_v = constant, and so

$$\ddot{a} = -\frac{4\pi G \rho_0}{3a^2} + \frac{8\pi G \rho_v a}{3}. \tag{20.88}$$

Equations (20.85) and (20.88) have exactly the same dependence of the 'cosmological term' upon the scale factor a, and so we can formally identify the cosmological constant with the vacuum mass density by

$$\Lambda = 8\pi G \rho_v. \tag{20.89}$$

At the present epoch, $a = 1$ and so

$$\ddot{a}(t_0) = -\frac{4\pi G \rho_0}{3} + \frac{8\pi G \rho_v}{3}. \tag{20.90}$$

Introducing a density parameter Ω_v associated with ρ_v, we obtain

$$\Omega_v = \frac{8\pi G \rho_v}{3H_0^2} \quad \text{and so} \quad \Lambda = 3H_0^2\Omega_v. \tag{20.91}$$

We can therefore find a new relation between the deceleration parameter q_0, Ω_0 and Ω_v from (20.90) and (20.91):

$$q_0 = \frac{\Omega_0}{2} - \Omega_v. \tag{20.92}$$

The dynamical equations (20.50) and (20.51) can be rewritten

$$\ddot{a} = -\frac{\Omega_0 H_0^2}{2}\frac{1}{a^2} + \Omega_v H_0^2 a, \tag{20.93}$$

$$\dot{a}^2 = \frac{\Omega_0 H_0^2}{a} - \frac{c^2}{\mathcal{R}^2} + \Omega_v H_0^2 a^2. \tag{20.94}$$

Substituting the values $a = 1$ and $\dot{a} = H_0$ at the present epoch into (20.94), the curvature of space is related to Ω_0 and Ω_v by

$$\frac{c^2}{\mathcal{R}^2} = H_0^2[(\Omega_0 + \Omega_v) - 1], \tag{20.95}$$

or

$$\kappa = \frac{1}{\mathcal{R}^2} = \frac{[(\Omega_0 + \Omega_v) - 1]}{(c^2/H_0^2)}. \tag{20.96}$$

Thus, the condition that the spatial sections are flat Euclidean space is

$$(\Omega_0 + \Omega_v) = 1. \tag{20.97}$$

The radius of curvature R_c of the model changes with scale factor as $R_c = a\mathcal{R}$, and so if the space curvature is zero now, it must have been zero at all times in the past. This argument leads to the flatness problem discussed in Section 20.5.1, but now for models with $\Lambda \neq 0$ as well. Since $\kappa(z) \propto R_c^{-2}(z) = (1 + z)^2/\mathcal{R}^2$, if κ were different from zero by even a very small amount at very large redshifts, it would be very different indeed from zero at the present epoch.

Let us investigate the dynamics of the models with $\Lambda \neq 0$.

Models with $\Lambda < 0$; $\Omega_v < 0$. Models with negative cosmological constant are not of great interest, because the net effect is to incorporate an attractive force, in addition to gravity, which slows down the expansion of the Universe. The one difference from the models with $\Lambda = 0$ is that, no matter how small the values of Λ and Ω_0, the universal expansion is eventually reversed, as may be seen from inspection of (20.93).

Models with $\Lambda > 0$; $\Omega_v > 0$. These models are of much greater interest, because the positive cosmological constant leads to a repulsion which opposes the attractive force of gravity. There is a minimum rate of expansion which occurs at scale factor

$$a_{\min} = (4\pi G\rho_0/\Lambda)^{1/3} = (\Omega_0/2\Omega_v)^{1/3}. \tag{20.98}$$

The minimum rate of expansion is

$$\dot{a}_{\min}^2 = \Lambda^{1/3}\left(\frac{3\Omega_0 H_0^2}{2}\right)^{2/3} - \frac{c^2}{\mathcal{R}^2} = 3H_0^2\left(\frac{\Omega_v\Omega_0^2}{4}\right)^{1/3} - \frac{c^2}{\mathcal{R}^2}. \tag{20.99}$$

If the right-hand side of (20.99) is greater than zero, the dynamical behaviour shown in Fig. 20.9(*a*) is found. For large values of a, the dynamics become those of the de Sitter Universe:

$$a(t) \propto \exp\left[(\Lambda/3)^{1/2}\, t\right] = \exp\left[\Omega_v^{1/2} H_0 t\right]. \tag{20.100}$$

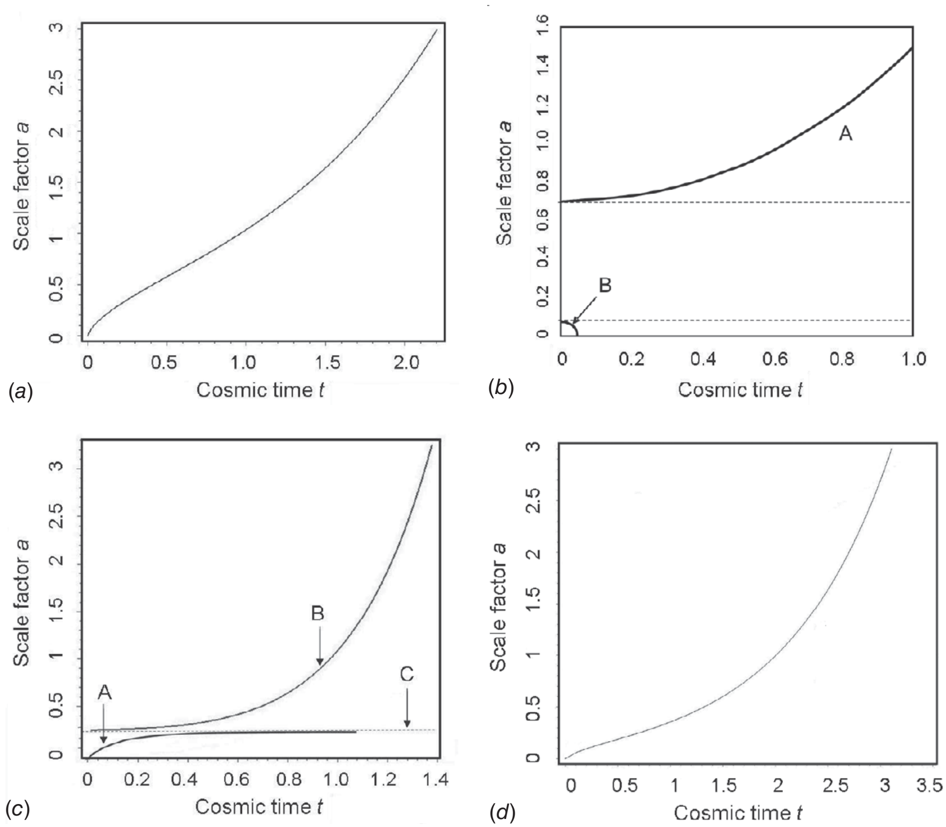


Fig. 20.9 Examples of the dynamics of world models in which $\Lambda \neq 0$. Models (*a*) and (*d*) are referred to as Lemaître models. In (*a*), the model parameters are $\Omega_0 = 0.3$ and $\Omega_v = 0.7$, a favoured model according to current best estimates of these parameters. (*b*) This 'bouncing' model has $\Omega_0 = 0.05$ and $\Omega_v = 2$. The zero of cosmic time has been set to the value when $\dot{a} = 0$. The loci are symmetrical in cosmic time with respect to this origin. (*c*) This is an Eddington–Lemaître model which is stationary at redshift $z_c = 3$, corresponding to scale factor $a = 0.25$. In (*d*), the model parameters are $\Omega_0 = 0.01$ and $\Omega_v = 0.99$ and the age of the Universe can far exceed H_0^{-1}.

If the right-hand side of (20.99) is less than zero, there is a range of scale factors for which no solution exists, and the function $a(t)$ has two branches, as illustrated in Fig. 20.9(*b*). For the branch labelled A, the dynamics are dominated by the Λ term and the asymptotic solutions are described by

$$a(t) \propto \exp\left[\pm (\Lambda/3)^{1/2}\, t\right]. \tag{20.101}$$

For the branch B, the Universe never expanded to sufficiently large values of a for the repulsive effect of the Λ term to prevent the Universe collapsing.

The most interesting cases are those for which $\dot{a}_{\min} \approx 0$. The case $\dot{a}_{\min} = 0$ is known as the *Eddington–Lemaître model* and is illustrated in Fig. 20.9(*c*). The interpretation of these models is either: A, the Universe expanded from an origin at some finite time in the past and will eventually attain a stationary state in the infinite future; B, the Universe is expanding away from a stationary solution in the infinite past. The stationary state C is unstable because, if it is perturbed, the Universe moves either onto branch B or onto the collapsing variant of branch A. In Einstein's static Universe of 1917, the stationary phase occurs at the present day. In general, the value of Λ corresponding to $\dot{a}_{\min} = 0$ is

$$\Lambda = \tfrac{3}{2}\Omega_0 H_0^2 (1 + z_c)^3 \quad \text{or} \quad \Omega_v = \frac{\Omega_0}{2}(1 + z_c)^3, \tag{20.102}$$

where z_c is the redshift of the stationary state.

The Eddington–Lemaître models can have ages much greater than H_0^{-1}; in the extreme stationary Eddington–Lemaître model with $\dot{a}_{\min} = 0$, the Universe is infinitely old. A closely related set of models, also with ages greater than H_0^{-1}, are the *Lemaître models*, which have Λ terms such that the value of $\dot{a}_{\min}$ is just greater than zero. An example of this type of model is shown in Fig. 20.9(*d*). There is a long 'coasting phase' when the velocity of expansion of the Universe was very small.

20.5.4 Λ Models with Zero Curvature

As we will find, Λ models with zero curvature are of special importance for cosmology. Conveniently, there are analytic solutions in this case. We need the expression for dz/dt to replace (20.68). From (20.94),

$$\frac{dz}{dt} = -H_0(1 + z)\left[(1 + z)^2(\Omega_0 z + 1) - \Omega_v z(z + 2)\right]^{1/2}. \tag{20.103}$$

Cosmic time t measured from the Big Bang follows immediately by integration:

$$t = \int_0^t dt = -\frac{1}{H_0}\int_\infty^z \frac{dz}{(1 + z)\left[(1 + z)^2(\Omega_0 z + 1) - \Omega_v z(z + 2)\right]^{1/2}}. \tag{20.104}$$

For models with zero curvature, $(\Omega_0 + \Omega_v) = 1$ and hence (20.103) becomes

$$\frac{dz}{dt} = -H_0(1 + z)\left[\Omega_0(1 + z)^3 + (1 - \Omega_0)\right]^{1/2}, \tag{20.105}$$

which can be substituted into the integral (20.25) to find the radial comoving distance coordinate. The cosmic time–redshift relation is found by integration:

$$t = \int_0^t dt = -\frac{1}{H_0}\int_\infty^z \frac{dz}{(1 + z)\left[\Omega_0(1 + z)^3 + (1 - \Omega_0)\right]^{1/2}}. \tag{20.106}$$

The parametric solution is

$$t = \frac{2}{3H_0\Omega_v^{1/2}} \ln\left(\frac{1 + \cos\theta}{\sin\theta}\right) \quad \text{where} \quad \tan\theta = \left(\frac{\Omega_0}{\Omega_v}\right)^{1/2}(1 + z)^{3/2}. \tag{20.107}$$

The present age of the Universe is found by setting $z = 0$,

$$t_0 = \frac{2}{3H_0\Omega_v^{1/2}} \ln\left[\frac{1 + \Omega_v^{1/2}}{(1 - \Omega_v)^{1/2}}\right]. \tag{20.108}$$

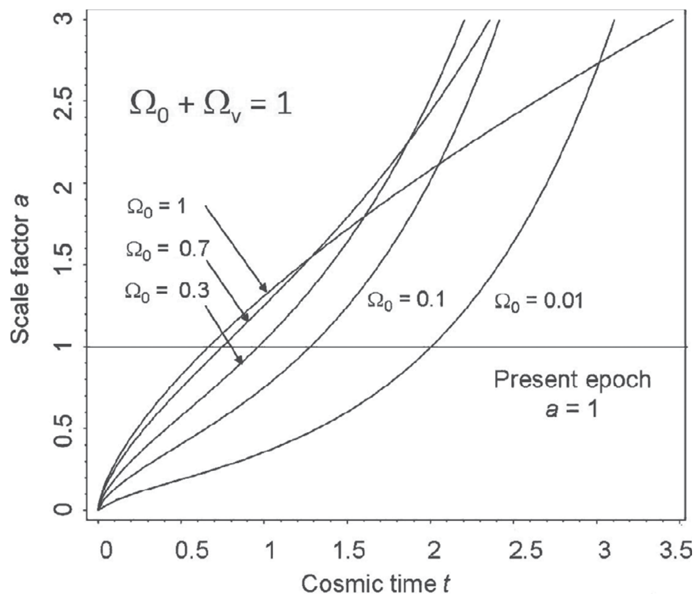


Fig. 20.10 The dynamics of spatially flat world models, $\Omega_0 + \Omega_v = 1$, with different combinations of Ω_0 and Ω_v. The abscissa is plotted in units of H_0^{-1}. The dynamics of these models can be compared with those shown in Fig. 20.7 for which $\Omega_v = 0$.

This relation illustrates how it is possible to find Friedman models with ages greater than H_0^{-1} with flat spatial sections. For example, if $\Omega_v = 0.9$ and $\Omega_0 = 0.1$, the age of the world model would be $1.28H_0^{-1}$. Examples of the dynamics of flat cosmological models with finite values of the cosmological constant are shown in Fig. 20.10.

20.6 Historical Interlude: Steady State Theory

In the 1940s and 1950s, an attempt was made by Hermann Bondi, Thomas Gold and Fred Hoyle to derive the function $a(t)$ from very general considerations in the light of what was known about the Universe at that time – the result was *Steady State Theory*. The original motivation was the fact that the value of Hubble's constant H_0 was estimated to be $H_0 \approx 500$ km s^{-1} Mpc^{-1} causing a significant time-scale problem. As has been shown, all the Friedman world models with zero cosmological constant have ages less than $T_0 = H_0^{-1}$ and so the Universe would have to be less than 2×10^9 years old. This value was less than the age of the Earth, which was known to be about 4.6×10^9 years from radioactive dating.

Bondi, Gold and Hoyle came up with an ingenious solution to the time-scale problem by replacing the cosmological principle with what they called the *perfect cosmological principle*.[12] This is the statement that the large-scale properties of the Universe as observed today should be *the same for all fundamental observers at all epochs*. In other words, the Universe should present an unchanging appearance for all fundamental observers for all time. This postulate immediately determines the kinematics of the Universe, since Hubble's constant becomes a fundamental constant of nature. Thus, from (20.30),

$$\dot{a}/a = H_0, \quad \text{and hence} \quad a \propto \exp(H_0 t). \tag{20.109}$$

Normalising to the value $a(t_0) = 1$ at the present epoch t_0,

$$a(t) = \exp[H_0(t - t_0)], \tag{20.110}$$

where t is cosmic time.

The immediate consequence of this model is that, since the matter is continually dispersing, it must be created continuously out of nothing so that the Universe preserves the same appearance at all epochs. Hence, the theory is sometimes referred to as the theory of *continuous creation*. The proponents of the theory showed that the rate at which matter would be created in order to replace the dispersing matter was unobservably small, amounting to only one particle per cubic metre every 300 000 years.

Other features of the model follow immediately from the formalism we have already derived. Equation (20.15) states that $R_c(t) = \mathcal{R}a(t)$ and so the curvature $\kappa = \mathcal{R}^{-2}$ must be zero, the only value which does not change with time. Equally straightforward is the relation between the comoving radial distance coordinate r and redshift z. Recalling that $a(t) = (1 + z)^{-1}$, the expression for r becomes

$$r = \int_t^{t_0} \frac{c\,dt}{a(t)} = \int_t^{t_0} \frac{c\,dt}{\exp[H_0(t - t_0)]} = \frac{c}{H_0}\left[\frac{1}{a(t)} - 1\right] = \frac{cz}{H_0}. \tag{20.111}$$

Remarkably, this relation is precisely Hubble's law, but now it holds good for all redshifts. Therefore, the metric for Steady State cosmology is

$$ds^2 = dt^2 - \frac{a^2(t)}{c^2}[dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)], \tag{20.112}$$

with $a(t) = \exp[H_0(t - t_0)]$ and $r = cz/H_0$. Finally, the Universe is of infinite age, resolving the time-scale problem – there is no initial singularity, unlike the standard Big Bang.

In the mid-1950s, Walter Baade showed that Hubble's constant had been significantly overestimated, and Humason, Mayall and Sandage derived a revised value of $H_0 = 180$ km s^{-1} Mpc^{-1}, giving $T_0 = H_0^{-1} = 5.6 \times 10^9$ years, which is comfortably greater than the age of the Earth. Subsequent studies reduced the value even further, as discussed in Section 20.9.

Despite the fact that one of the motivations for the theory was undermined, the simplicity and elegance of the theory was a profound attraction for a number of theorists, who saw in it a unique cosmological model, as opposed to the standard models which depend upon the mass density of the Universe and the cosmological constant, and possess a singularity at $t = 0$.

The uniqueness of the theory made it highly testable. One of the first pieces of evidence against the theory was derived from the number counts of faint radio sources, which indicated that there were many more radio sources per unit comoving volume in the past than there are at the present day, violating the basic premise of the Steady State theory. The fatal blow to the theory was, however, the CMB. There is no natural origin for this radiation in the Steady State picture – there do not exist sources which could mimic its black-body spectrum and account for its enormous energy density. On the other hand, these properties find a natural explanation in the standard Big Bang picture. More details of the various controversies involved in these cosmological debates are contained in the book by Helge Kragh, *Cosmology and Controversy: The Historical Development of Two Theories of the Universe*.[13]

The story of the Steady State theory is intriguing from the point of view of methodology in cosmology. The elegance of the model was achieved at a heavy price – the introduction of the continuous creation of matter. For some scientists, this step was so objectionable that they ruled the theory out immediately. Hoyle subsequently investigated models in which the concept of the continuous creation of matter was preserved, but the Universe oscillated so that it mimicked an evolving Universe with a hot early phase. In Hoyle's original version of the Steady State theory of 1948, matter creation was associated with what he called a creation or C-field. It turns out that this term is not so different from the ψ scalar fields which appear in the inflationary model of the very early Universe, but the interpretation is quite different.

20.7 The Thermal History of the Universe

Once the dynamical framework of the world models is defined, the thermal history of the Universe can be found by the application of the laws of physics to uniformly expanding cosmological models, leading to further important tests of the standard picture. Let us first consider the dynamics of radiation-dominated universes.

20.7.1 Radiation-Dominated Universes

For a gas of photons, massless particles or a relativistic gas in the ultrarelativistic limit $E \gg mc^2$, the pressure p is related to its energy density ε by $p = \frac{1}{3}\varepsilon$, and the inertial mass density of the radiation $\rho_{\rm rad}$ is related to its energy density ε by $\varepsilon = \rho_{\rm rad}c^2$.

If $N(\nu)$ is the number density of photons of energy $h\nu$, then the energy density of radiation is found by summing over all frequencies,

$$\varepsilon = \sum_{\nu} h\nu\, N(\nu). \tag{20.113}$$

If photons are neither created nor destroyed, their number density varies as $N = N_0\, a^{-3} = N_0(1+z)^3$, and the energy of each photon changes by the redshift factor $h\nu = h\nu_0(1+z)$. Therefore, the variation of the energy density of radiation with cosmic epoch is

$$\varepsilon = \sum_{\nu_0} h\nu_0\, N_0(\nu_0)\,(1+z)^4 = \varepsilon_0(1+z)^4 = \varepsilon_0 a^{-4}. \tag{20.114}$$

For black-body radiation, the energy density is given by the Stefan–Boltzmann law, $\varepsilon = aT^4$, and its spectral energy density by the Planck distribution:

$$\varepsilon(\nu)\,{\rm d}\nu = \frac{8\pi h\nu^3}{c^3}\frac{1}{{\rm e}^{h\nu/k_{\rm B}T}-1}\,{\rm d}\nu. \tag{20.115}$$

It immediately follows that, for black-body radiation, the temperature T varies with redshift as $T = T_0(1+z)$ and its spectrum as

$$\varepsilon(\nu_1)\,\mathrm{d}\nu_1 = \frac{8\pi h\nu_1^3}{c^3}[(\mathrm{e}^{h\nu_1/k_\mathrm{B}T_1} - 1)]^{-1}\,\mathrm{d}\nu_1 = \frac{8\pi h\nu_0^3}{c^3}[\mathrm{e}^{h\nu_0/k_\mathrm{B}T_0} - 1)^{-1}](1+z)^4\,\mathrm{d}\nu_0$$
$$= (1+z)^4\,\varepsilon(\nu_0)\,\mathrm{d}\nu_0. \tag{20.116}$$

Thus, a black-body spectrum preserves its form as the Universe expands, but the radiation temperature changes as $T = T_0(1+z)$ and the frequency of each photon as $\nu = \nu_0(1+z)$. Another way of thinking about these results is in terms of the adiabatic expansion of a gas of photons. The ratio of specific heat capacities for radiation, and a relativistic gas in the ultrarelativistic limit, is $\gamma = 4/3$. As shown in Section 11.3.3, in an adiabatic expansion, $T \propto V^{-(\gamma-1)} = V^{-1/3} \propto a^{-1} = (1+z)$, which is exactly the same as the above result.

The variations of p and ρ with scale factor a can now be substituted into Einstein's field equations (20.50) and (20.51). Setting the cosmological constant $\Lambda = 0$,

$$\ddot{a} = -\frac{8\pi G\varepsilon_0}{3c^2}\frac{1}{a^3}, \qquad \dot{a}^2 = \frac{8\pi G\varepsilon_0}{3c^2}\frac{1}{a^2} - \frac{c^2}{\mathcal{R}^2}. \tag{20.117}$$

At early epochs, the constant term $c^2/\mathcal{R}^2$ can be neglected, and so integrating,

$$a = \left(\frac{32\pi G\varepsilon_0}{3c^2}\right)^{1/4} t^{1/2} \quad \text{or} \quad \varepsilon = \varepsilon_0\, a^{-4} = \left(\frac{3c^2}{32\pi G}\right) t^{-2}. \tag{20.118}$$

The dynamics of the radiation-dominated models, $a \propto t^{1/2}$, depend only upon the total inertial mass density in relativistic and massless forms. Notice that we need to add together the contributions of all forms of relativistic matter and radiation to ε at the appropriate epochs.

20.7.2 The Matter and Radiation Content of the Universe

The CMB provides by far the largest contribution to the energy density of radiation in intergalactic space. Comparing the inertial mass density in the radiation and the matter,

$$\frac{\rho_\mathrm{rad}}{\rho_\mathrm{matter}} = \frac{aT^4(z)}{\Omega_0\rho_\mathrm{c}(1+z)^3c^2} = \frac{2.6\times10^{-5}(1+z)}{\Omega_0 h^2}, \tag{20.119}$$

where we have written $h = H_0/100$ km s^{-1} Mpc^{-1} to take account of uncertainties in the precise value of Hubble's constant H_0. The Universe was therefore *matter-dominated* at redshifts $z \le 4\times10^4\,\Omega_0 h^2$ and the dynamics described by $a \propto t^{2/3}$, provided $\Omega_0 z \gg 1$.[14] At redshifts $z \ge 4\times10^4\,\Omega_0 h^2$, the Universe was *radiation dominated* and the dynamics described by (20.118), $a \propto t^{1/2}$.

The present *photon-to-baryon ratio* is another key cosmological parameter. Adopting $T = 2.728$ K,

$$\frac{N_\gamma}{N_\mathrm{B}} = \frac{3.6\times10^7}{\Omega_\mathrm{B} h^2}, \tag{20.120}$$

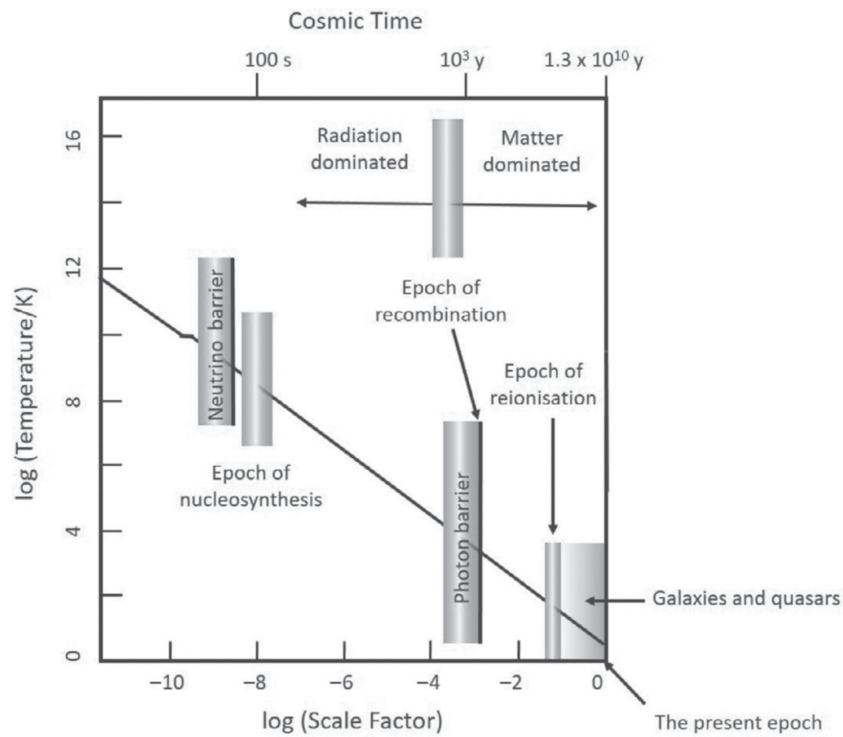


Fig. 20.11 A schematic diagram showing the thermal history of the Universe. The radiation temperature decreases as $T_r \propto a^{-1}$ as indicated by the solid diagonal line, except for small jumps as different particle–antiparticle pairs annihilate. Various important epochs in the standard Big Bang model are indicated. An approximate time-scale is shown along the top of the diagram. The photon and neutrino barriers are shown. In the standard picture, the Universe is opaque for electromagnetic radiation and neutrinos prior to these epochs.

where N_B and Ω_B refer to the number density and density parameter of baryons respectively. If photons were neither created nor destroyed during the expansion of the Universe, this ratio is an invariant and is the factor by which the photons outnumber the baryons in the Universe at the present epoch. It is also proportional to the specific entropy per baryon during the radiation-dominated phases of the expansion.

The resulting thermal history of the Universe is shown schematically in Fig. 20.11 on logarithmic scales. Certain epochs are of special significance.

20.7.3 The Epoch of Recombination

At a redshift $z \approx 1500$, the temperature of the CMB was $T \approx 4000$ K and there were then sufficient photons with energies $h\nu \geq 13.6$ eV in the Wien region of the Planck distribution to ionise all the neutral hydrogen in the intergalactic medium. The number density of photons in the Wien region of the Planck distribution, $h\nu \gg k_B T$, with energies $h\nu \geq E$ is

$$n(\geq E) = \int_{E/h}^{\infty} \frac{8\pi\nu^2}{c^3}\frac{\mathrm{d}\nu}{\mathrm{e}^{h\nu/k_{\mathrm{B}}T}} = \frac{1}{\pi^2}\left(\frac{2\pi k_{\mathrm{B}}T}{hc}\right)^3 \mathrm{e}^{-x}(x^2+2x+2), \tag{20.121}$$

where $x = E/k_{\mathrm{B}}T$. The total number density of photons in a black-body spectrum at temperature T is

$$N = 0.244\left(\frac{2\pi k_{\mathrm{B}}T}{hc}\right)^3 \mathrm{m}^{-3}, \tag{20.122}$$

and so the fraction of the photons with energies greater than E is

$$\frac{n(\geq E)}{N} = \frac{\mathrm{e}^{-x}(x^2+2x+2)}{0.244\pi^2}. \tag{20.123}$$

In the simplest approximation, the intergalactic gas is ionised if there are as many ionising photons as hydrogen atoms, that is, from (20.120), there need only be one photon for every $3.6 \times 10^7/\Omega_{\mathrm{B}}h^2$ of those in the CMB with energy greater than 13.6 eV to ionise the gas.

For illustrative purposes, let us take this ratio to be one part in 10^9. Then, the solution of the expression

$$\frac{1}{10^9} = \frac{\mathrm{e}^{-x}(x^2+2x+2)}{0.244\pi^2} \tag{20.124}$$

is $x = E/k_{\mathrm{B}}T \approx 26.5$. This is an important result. There are so many photons relative to hydrogen atoms that the temperature of the radiation can be 26.5 times less than that found by setting $E = k_{\mathrm{B}}T$ and there are still sufficient photons with energy $E \geq 13.6$ eV to ionise the gas. Since $k_{\mathrm{B}}T = 13.6$ eV at 150 000 K, the intergalactic gas is expected to be ionised at temperature $T \approx 150\,000/26.5 \sim 5600$ K. The present temperature of the CMB is 2.728 K, and so the hydrogen in the pregalactic gas was fully ionised at a scale factor $a \sim 2.728/5600 = 5 \times 10^{-4}$, or a redshift $z \sim 2000$. More detailed calculations show that the intergalactic neutral hydrogen was 50% ionised at $z \approx 1500$, corresponding to a temperature $T_{\mathrm{r}} \approx 4000$ K. This epoch is referred to as the *epoch of recombination* – the pregalactic neutral hydrogen was fully ionised prior to this epoch and so, when the clocks are run forward, the universal plasma recombined at this epoch.

The same type of calculation appears in a number of different guises in astrophysics – the nuclear reactions which power the Sun take place at a much lower temperature than expected, the temperature at which regions of ionised hydrogen become fully ionised is only about 10 000 K and light nuclei are destroyed in the early Universe at much lower temperatures than would be expected. In all these cases, the high energy tails of the Planck and Maxwell distributions contain huge numbers of particles or photons with energies very much greater than the mean.

The most important consequence of this calculation is that the Universe was opaque to Thomson scattering by free electrons at early epochs. Detailed calculations show that the optical depth of the intergalactic gas for Thomson scattering was unity at a redshift close to 1000, and then became very large as soon as the intergalactic hydrogen was fully ionised at redshift z_{r}. As a result, the Universe beyond $z \approx 1000$ is unobservable, because any photons originating from larger redshifts are scattered many times before they propagate to the Earth – there is a *photon barrier* at a redshift of 1000 beyond which we cannot obtain information directly using photons. If there is no further scattering of the photons between the epoch of recombination and the present epoch, the redshift of about 1000 is the *last scattering surface* for the CMB.[15] The spatial fluctuations at high galactic latitudes in the images of the microwave sky in Figs. 20.3 and 20.14 were imprinted on this last scattering surface at $z \approx 1000$.

20.7.4 The Epoch of Equality of Matter and Radiation Inertial Mass Densities

It follows from (20.119) that the matter and radiation made equal contributions to the inertial mass density at redshift $z \approx 4 \times 10^4\,\Omega_0 h^2$, and at larger redshifts the Universe was radiation dominated. After the intergalactic gas recombined at redshifts $z \leq 1000$, there was negligible coupling between the neutral matter and the photons of the microwave background radiation. If the matter and radiation were not thermally coupled, they would cool independently, the gas having ratio of specific heat capacities $\gamma = 5/3$ and the radiation $\gamma = 4/3$. These result in adiabatic cooling which depends upon the scale factor a as $T_{\rm m} \propto a^{-2}$ and $T_{\rm rad} \propto a^{-1}$ for the matter and radiation respectively. We therefore expect the matter to cool more rapidly than the radiation during the post-recombination epochs, once the matter and radiation are fully decoupled at a redshift $z \approx 400$.[16] This is not the case prior to recombination, however, because the matter and radiation were strongly coupled by *Compton scattering*. Before recombination, the optical depth of the intergalactic plasma for Thomson scattering was so large that the small energy transfers between the photons and the electrons by Compton scattering were sufficient to maintain the matter at the same temperature as the radiation.

Once the thermal history of the intergalactic gas has been determined, the variation of the speed of sound with cosmic epoch can be found. The adiabatic speed of sound $c_{\rm s}$ is

$$c_{\rm s}^2 = \left(\frac{\partial p}{\partial \rho}\right)_S ,$$

where the subscript S means ‘at constant entropy’. A complication is that, from the epoch when the energy densities of matter and radiation were equal to beyond the epoch of recombination and the subsequent neutral phase, the dominant contributions to p and ρ changed dramatically as the Universe changed from being radiation to matter dominated. As long as the matter and radiation are closely coupled, the square of the sound speed can be written

$$c_{\rm s}^2 = \frac{(\partial p/\partial T)_{\rm rad}}{(\partial \rho/\partial T)_{\rm rad} + (\partial \rho/\partial T)_{\rm m}}, \tag{20.125}$$

where the partial derivatives are taken at constant entropy. This expression reduces to the following expression:

$$c_{\rm s}^2 = \frac{c^2}{3}\frac{4\rho_{\rm rad}}{4\rho_{\rm rad} + 3\rho_{\rm m}}. \tag{20.126}$$

Thus, in the radiation-dominated phases, $z \geq 4 \times 10^4\, \Omega_0 h^2$, the speed of sound tends to the relativistic sound speed, $c_s = c/\sqrt{3}$. At smaller redshifts, the sound speed decreases as the contribution of the inertial mass density of the thermal matter becomes more important. Finally, after the decoupling of the matter and radiation are complete, the sound speed becomes the thermal sound speed of the cooling primordial gas. These considerations are crucial for understanding the physics of the growth of small density perturbations in the primordial plasma.

20.7.5 Early Epochs

To complete this brief thermal history of the Universe, let us summarise the physical processes which are important at even earlier epochs.

- Extrapolating back to redshifts $z \approx 10^8$, the radiation temperature was $T \approx 3 \times 10^8$ K and the background photons attained γ-ray energies, $\epsilon = k_B T = 25$ keV. At this high temperature, the photons in the Wien region of the Planck spectrum are energetic enough to dissociate light nuclei such as helium and deuterium. At earlier epochs, all nuclei are dissociated.
- At redshifts $z \approx 10^9$, electron–positron pair production from the thermal background radiation took place and the Universe was flooded with electron–positron pairs, one pair for every pair of photons present in the Universe today. When the clocks are run forward from an earlier epoch, the electrons and positrons annihilate at about this epoch and their energy is transferred to the photon field – this accounts for the little discontinuity at $a = 10^{-9}$ in the temperature history shown in Fig. 20.11.
- At a slightly earlier epoch, the opacity of the Universe for weak interactions was unity. This results is a *neutrino barrier*, similar to the photon barrier at $z \sim 1000$.
- We can extrapolate even further back in time to $z \approx 10^{12}$, when the temperature of the background radiation was sufficiently high for baryon–antibaryon pair production to take place from the thermal background. Just as in the case of the epoch of electron–positron pair production, the Universe was flooded with baryons and antibaryons, one pair for every pair of photons present in the Universe now. Again, there is a little discontinuity in the temperature history at this epoch.

The process of extrapolation further and further back into the mists of the early Universe continues until we run out of physics which has been established in the laboratory. Most particle physicists would probably agree that the standard model of elementary particles has been tried and tested to energies of about 200 GeV, and so we can probably trust laboratory physics back to epochs as early as 10^{-6} s, although the more conservative among us might be happier to accept 10^{-3} s. How far back one is prepared to extrapolate beyond these times is largely a matter of taste. The most ambitious theorists have no hesitation in extrapolating back to the very earliest Planck eras, $t_P \sim (Gh/c^5)^{1/2} = 10^{-43}$ s, when the relevant physics was certainly very different from the physics of the Universe from redshifts of about 10^{12} to the present day. Aspects of the physics of the very early Universe are discussed in Chapter 21.

20.8 Nucleosynthesis in the Early Universe

One of the reasons why the standard Big Bang model is taken so seriously is its remarkable success in accounting for the observed abundances of the light elements by *primordial nucleosynthesis*.[17] Consider a particle of mass m at very high temperatures, $k_B T \gg mc^2$. If the time-scales of the interactions which maintain this species in thermal equilibrium with all the other species present at temperature T are shorter than the age of the Universe at that epoch, the equilibrium number densities of the particle and its antiparticle are given by the statistical mechanical formulae:

$$N = \overline{N} = \frac{4\pi g}{h^3} \int_0^\infty \frac{p^2\,\mathrm{d}p}{e^{E/k_B T} \pm 1}, \tag{20.127}$$

where g is the statistical weight of the particle, p its momentum and the $\pm$ sign depends upon whether the particles are fermions (+) or bosons (−). For (i) photons, (ii) nucleons and electrons and (iii) neutrinos respectively, these are

$$\begin{aligned}
&\text{(i)} \quad g = 2 \quad N = 0.244\left(\frac{2\pi k_B T}{hc}\right)^3 \text{m}^{-3} \quad \varepsilon = aT^4,\\
&\text{(ii)} \quad g = 2 \quad N = 0.183\left(\frac{2\pi k_B T}{hc}\right)^3 \text{m}^{-3} \quad \varepsilon = \tfrac{7}{8}aT^4,\\
&\text{(iii)} \quad g = 1 \quad N = 0.091\left(\frac{2\pi k_B T}{hc}\right)^3 \text{m}^{-3} \quad \varepsilon = \tfrac{7}{16}aT^4.
\end{aligned}$$

To find the total energy density ε, we add all the equilibrium energy densities,

$$\varepsilon = \chi(T)\, aT^4. \tag{20.128}$$

This is the expression for the energy density which should be included in (20.118) to determine the dynamics of the early Universe.

When the particles become non-relativistic, $k_B T \ll mc^2$, and the abundances of the different species are maintained by interactions between the particles, the non-relativistic limit of (20.127) results in an equilibrium number density

$$N = g\left(\frac{mk_B T}{h^2}\right)^{3/2} \exp\left(-\frac{mc^2}{k_B T}\right). \tag{20.129}$$

Thus, once the particles become non-relativistic, they no longer contribute to the inertial mass density which determines the rate of expansion of the Universe.

At redshifts $z < 10^{12}$, the neutrons and protons were non-relativistic, $k_B T \ll mc^2$, and their equilibrium abundances were maintained by the electron–neutrino weak interactions

$$e^+ + n \rightarrow p + \overline{\nu}_e \quad , \quad \nu_e + n \rightarrow p + e^-. \tag{20.130}$$

The values of g are the same for the neutrons and protons and so their relative abundances are

$$\left[\frac{\mathrm{n}}{\mathrm{p}}\right] = \exp\left(-\frac{\Delta mc^2}{k_\mathrm{B}T}\right), \tag{20.131}$$

where Δmc^2 is the mass difference between the neutron and the proton.

This abundance ratio froze out when the neutrino interactions could no longer maintain the equilibrium abundances of neutrons and protons. The condition for 'freezing out' is that the time-scale of the weak interactions becomes greater than the age of the Universe. Detailed calculations show that the time-scales for the expansion of the Universe and the decoupling of the neutrinos are the same when the Universe was almost exactly one second old at a temperature of 10^{10} K. At that time, the neutron fraction, as determined by (20.131), was

$$\left[\frac{\mathrm{n}}{\mathrm{n}+\mathrm{p}}\right] = 0.21. \tag{20.132}$$

The neutron fraction decreased very slowly after this time. Detailed calculations show that after 300 s the neutron fraction had fallen to 0.123.[18]

At this epoch, the bulk of the formation of the light elements took place. Almost all the neutrons were combined with protons to form ^{4}He nuclei so that for every pair of neutrons a helium nucleus is formed. The reactions involved are

$$\begin{cases} \mathrm{p}+\mathrm{n} \to \mathrm{D}+\gamma & \mathrm{p}+\mathrm{D} \to {}^3\mathrm{He}+\gamma \\ \mathrm{n}+\mathrm{D} \to {}^3\mathrm{H}+\gamma & \mathrm{p}+{}^3\mathrm{H} \to {}^4\mathrm{He}+\gamma \\ \mathrm{n}+{}^3\mathrm{He} \to {}^4\mathrm{He}+\gamma & \mathrm{d}+\mathrm{d} \to {}^4\mathrm{He}+\gamma \\ {}^3\mathrm{He}+{}^3\mathrm{He} \to {}^4\mathrm{He}+2\mathrm{p} & \end{cases} \tag{20.133}$$

Most of the nucleosynthesis took place at a temperature less than about 1.2×10^9 K since, at higher temperatures, the deuterons are destroyed by the γ-rays of the background radiation. The binding energy of deuterium is $E_\mathrm{B} = 2.23$ MeV and this energy is equal to $k_\mathrm{B}T$ at $T = 2.6 \times 10^{10}$ K. However, just as in the case of the recombination of the intergalactic gas, the photons far outnumber the nucleons and it is only when the temperature of the expanding gas had decreased to about 26 times less than this temperature that the number of dissociating photons is less than the number of nucleons. Although some of the neutrons decay spontaneously by this time, the bulk of them survive and so, according to the above calculation, the predicted helium to hydrogen mass ratio is just twice the neutron fraction:

$$\left[\frac{^4\mathrm{He}}{\mathrm{H}}\right] \approx 0.25. \tag{20.134}$$

Detailed studies show that, in addition to ^{4}He, there are traces of the light elements deuterium (D), helium-3 (^{3}He) and lithium-7 (^{7}Li) (Fig. 20.12).

It had been a mystery why the abundance of helium is so high wherever it can be observed in the Universe. Its chemical abundance was always estimated to be greater than about 23%. In addition, there had always been a problem in understanding how deuterium could have been synthesised in stars. It is a very fragile nucleus and is destroyed rather than created in stellar interiors. Yet the observed deuterium abundance relative to hydrogen in

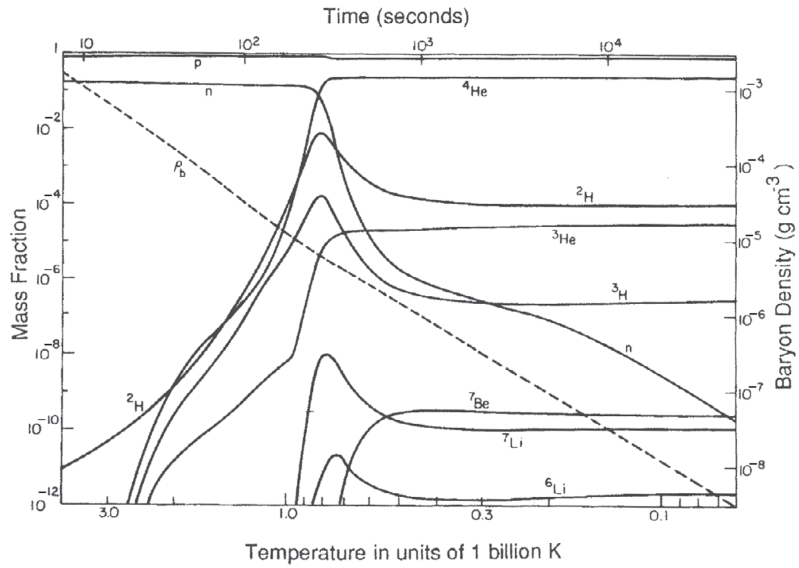


Fig. 20.12 An example of the time and temperature evolution of the abundances of the light elements in the early evolution of the Big Bang model of the Universe.[5] Before 10 seconds, no significant synthesis of light elements took place because deuterium, ^{2}H, is destroyed by hard γ-rays in the Wien region of the black-body radiation spectrum. As the temperature decreases, more and more deuterium survives and synthesis of the light elements becomes possible through the reactions (20.133). The synthesis of elements such as D(^{2}H), ^{3}He, ^{4}He, ^{7}Li and ^{7}Be was completed after about 15 minutes.

the interstellar gas is always [D/H] $\approx 1.5 \times 10^{-5}$. The type of same argument applies to the lighter isotope of helium, ^{3}He, and to ^{7}Li. Precisely these elements are synthesised in the early stages of the Big Bang. In stellar interiors, nucleosynthesis takes place in roughly thermodynamic equilibrium conditions over very long time-scales, whereas in the early stages of the Big Bang the 'explosive' nucleosynthesis is all over in a few minutes. The deuterium and ^{3}He abundances provide strong constraints on the present baryon density of the Universe.

Therefore, since we only know of ways of destroying deuterium rather than creating it, this figure provides a firm lower limit to the amount of deuterium which could have been produced by primordial nucleosynthesis. Wagoner's calculations show that the predicted abundance of deuterium is a strong function of the present density of the Universe, whereas the abundance of helium is remarkably constant (Fig. 20.13).

The reasons for this can be understood as follows. The amount of helium synthesised is determined by the equilibrium abundance of protons and neutrons as the Universe cools down, which is primarily determined by the thermodynamics of the expanding radiation-dominated Universe. On the other hand, the abundance of deuterium depends upon the number density of nucleons. If the Universe had high baryon number density, essentially all the deuterons would be converted into helium, whereas, if the Universe were of low density, not all the deuterium would be converted into ^{4}He. The same argument applies

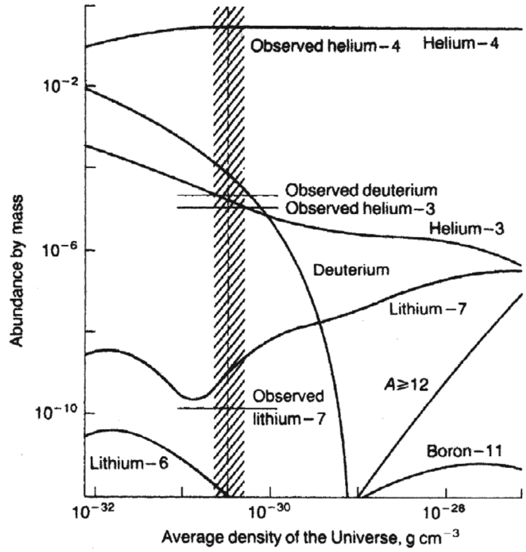


Fig. 20.13 The predicted primordial abundances of the light elements compared with their observed values. The present density of baryonic matter in the world model is plotted along the abscissa. The observed abundances of the elements are in good agreement with a model having $\Omega_B = 0.015h^{-2}$.

to ^{3}He. Thus, the deuterium and ^{3}He abundances set an upper limit to the present baryon density of the Universe. A typical figure which results from recent analyses[19] is

$$\Omega_B h^2 = 0.022^{+0.003}_{-0.022}. \tag{20.135}$$

Since $h \approx 0.7$, *baryonic matter cannot close the Universe*. Indeed, as will be discussed in Chapter 21, it cannot even account for the amount of dark matter present in the Universe.

One of the more remarkable results of these studies is that they enabled limits to be set to the number of neutrino species which could have been present during the epochs when the light elements were synthesised. If there had been more than three species of neutrino present, they would have contributed to the inertial mass density of massless particles and so would have speeded up the early expansion of the Universe (see (20.128)). The decoupling of the neutrinos would have taken place at a higher temperature resulting in the overproduction of helium. From this type of cosmological argument, it was shown that there could not be more than three families of neutrinos, a result subsequently confirmed from the width of the decay spectrum of $W^{\pm}$ and Z^0 bosons measured by LEP at CERN.

20.9 The Values of the Cosmological Parameters

With the development of powerful observational tools for cosmology across the electromagnetic spectrum, the values of the cosmological parameters have now been determined

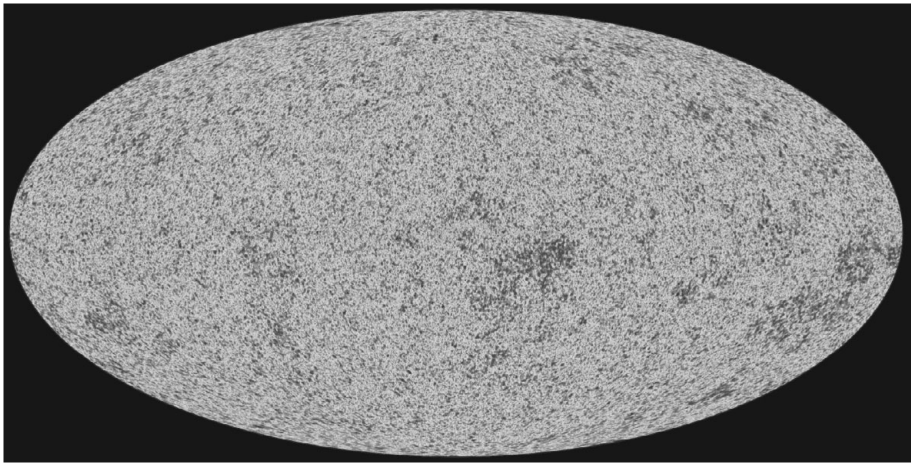


Fig. 20.14 The whole-sky map of the CMB as observed by the Planck satellite of the European Space Agency (ESA). The image shows large-scale structures in the Universe only 370 000 years after the Big Bang. It encodes a huge amount of information about the cosmological parameters which describe our Universe. (Courtesy of ESA and the Planck Collaboration.)

with unprecedented precision, opening up a whole range of fundamental problems which are addressed in Chapter 21. Here, we summarise briefly the results of a wide range of observations and experiments which bear upon the determination of the parameters. Unfortunately, this is not the place to go into the details of the observations, which are crucial in judging the weight to be attached to the various estimates. Many more details are provided in the chapters which my colleagues and I contributed to *The Oxford Handbook of the History of Modern Cosmology*, which contains references to the original papers.[20]

My preferred approach to reviewing the estimates of the comological parameters is to begin with observations of the CMB and then compare these with the many other independent routes to their estimation. The discovery of the CMB by Penzias and Wilson in 1965 was a key event of modern cosmology. We need not rehearse the subsequent 50 years of increasingly sensitive and challenging experiment, but leap directly to the remarkable results of the NASA Wilkinson Microwave Anisotropy Probe (WMAP) and the ESA Planck satellite. These space missions have measured the properties of the CMB with unprecedented precision and set a benchmark for contemporary *precision cosmology*.

The final all-sky image of the fluctuations in the CMB as measured by the Planck satellite is shown in Fig. 20.14, once the foregrounds due to Galactic radio and dust emission, as well as discrete radio sources, have been removed. The fluctuations typically have amplitude $\Delta T/T \sim 10^{-5}$; in other words, the fluctuations represent very small perturbations indeed in comparison with the total intensity of the isotropic background. The radiation originates from the last scattering surface at the epoch of recombination at $z \approx 1000$ and presents an image of the pre-galactic Universe when it was only 370 000 years old and out of which the structures we observe in the Universe today developed.

The theory of the origin of these ripples in both the total intensity and polarisation of the radiation has been the subject of a vast amount of investigation.[21] The beauty of this route to the determination of cosmological parameters is that the corresponding perturbations in the matter content of the universe from which large-scale structures of the Universe are ultimately formed are in the linear regime of their development, that is, $\delta\rho/\rho \ll 1$. Therefore, the predictions of the models can be made with considerable precision because the physics of the pre-galactic primordial plasma is well understood.

Observations of the power spectrum of the distribution of galaxies and large-scale structures show that it is continuous without any preferred scales. Harrison and Zeldovich[22] argued on general grounds that the initial power spectrum of perturbations should be of power-law form in the early Universe and this is adopted in the numerical simulations of the growth of perturbations in the matter and radiation. In order to create models for the predicted power spectrum of the CMB, remarkably few input parameters are required. The models have the following ingredients:

- The dominant forms of mass–energy are the dark matter with density parameter Ω_d and the dark energy with density parameter Ω_v.
- The baryonic matter has density parameter Ω_B.
- The spectrum of the initial perturbations from which the large-scale structure of the Universe formed is given by the Harrison–Zeldovich power spectrum $p(k) = Ak^n$, where k is the wavenumber of the perturbation, $n \approx 1$ and the phases of the perturbations are random.
- The optical depth to Thomson scattering back to the last scattering surface, which provides information about the epoch of reionisation.

In Fig. 20.15, the points show the observed power spectrum of fluctuations in the CMB for multipoles in the range $2 \leq l \leq 2500$ from observations with the Planck satellite. The solid line shows the remarkable agreement between the observed fluctuation power spectrum and the best-fit six-parameter model according to the standard ΛCDM model of structure formation.

One of the more dramatic results of the Planck mission is that gravitational lensing of fluctuations in the CMB on the last scattering surface at redshifts $z \sim 1000$ enables the power spectrum of dark matter perturbations on large scales to be determined as well. A map has been created of the distribution of dark matter on large scales and the features in it are correlated with the observed distribution of large-scale systems. As a result the six-parameter family of cosmological parameters can be determined entirely from observations of the CMB.

A direct consistency check is provided by the power spectrum of galaxies on very large scales at the present epoch as determined by the Sloan Digital Sky Survey. The first peak in the CMB power spectrum is reflected in the correlation function of galaxies on very large physical scales (Fig. 20.16) and the corresponding cosmological parameters are in excellent agreement with the value of $\Omega_b h^2$ from the Planck power spectrum.

The resulting large-scale properties of the Universe are summarised in Table 20.1. These show that, by adding together the mass–energy densities in the forms of baryonic matter, dark matter and dark energy, the Universe is geometrically flat, $\Omega_K = 0$, at the 0.002% level. [23]

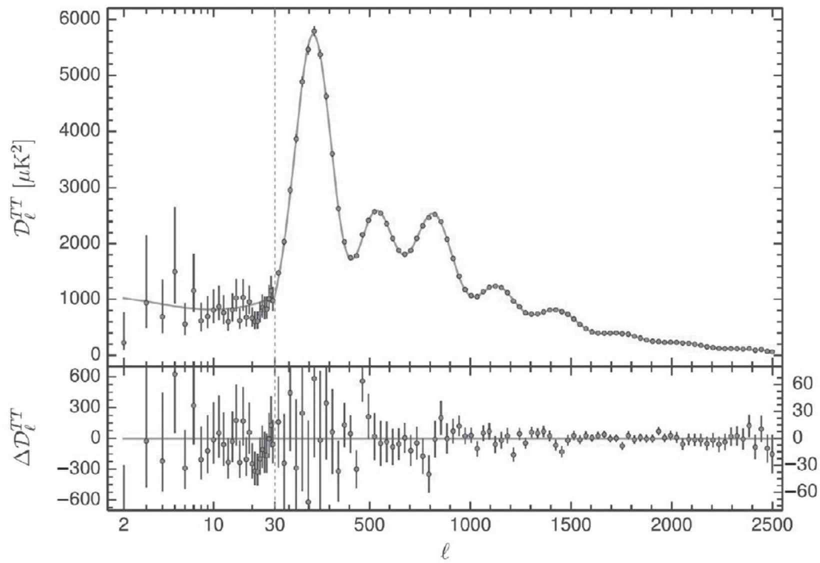


Fig. 20.15 The scalar power spectrum of fluctuations in the CMB as observed by the Planck satellite. The abscissa is logarithmic for wavenumbers $l < 30$ and linear for $l > 30$ (Courtesy of ESA and the Planck Science Team. The Planck Collaboration, ArXiv e-prints 1807.06209 (2018)).

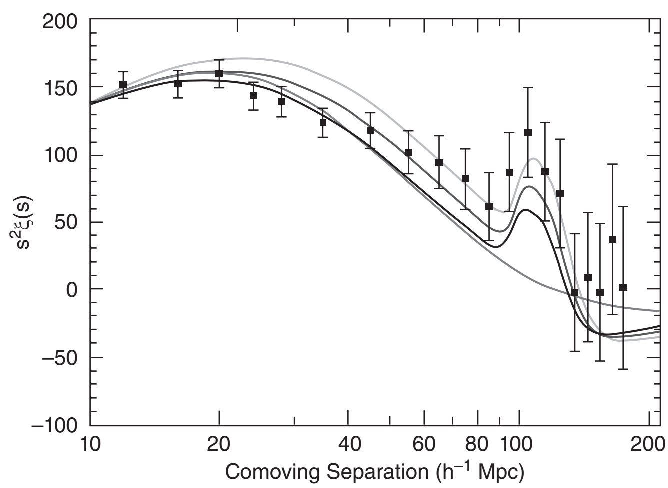


Fig. 20.16 The power spectrum of the distribution of galaxies on very large physical scales determined from observations of 17 million galaxies observed in the Sloan Digital Sky Survey showing the peak corresponding to the first maximum in the Planck power spectrum of fluctuations in the CMB seen in Fig. 20.15. The solid lines show various fits to the observations for the standard ΛCDM model, the top example having similar parameters to those used to account for the power spectrum of the CMB. (From Eisenstein, D.J. *et al.* (2005). Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies, *Astrophysical Journal*, **633**, 560–574.)

Table 20.1 The final (2018) values of cosmological parameters from the Planck mission

Baryonic matter	$\Omega_b h^2 = 0.02233 \pm 0.00015$
Cold dark matter	$\Omega_d h^2 = 0.1198 \pm 0.0012$
Dark energy	$\Omega_v = 0.6889 \pm 0.0056$
Spatial curvature	$\Omega_K = \Omega_b + \Omega_c + \Omega_v = 0.001 \pm 0.002$
Epoch of reionisation	$z_{reion} = 7.64 \pm 0.74$
Hubble's constant	$H_0 = 100h = 67.37 \pm 0.54$ km s^{-1} Mpc^{-1}

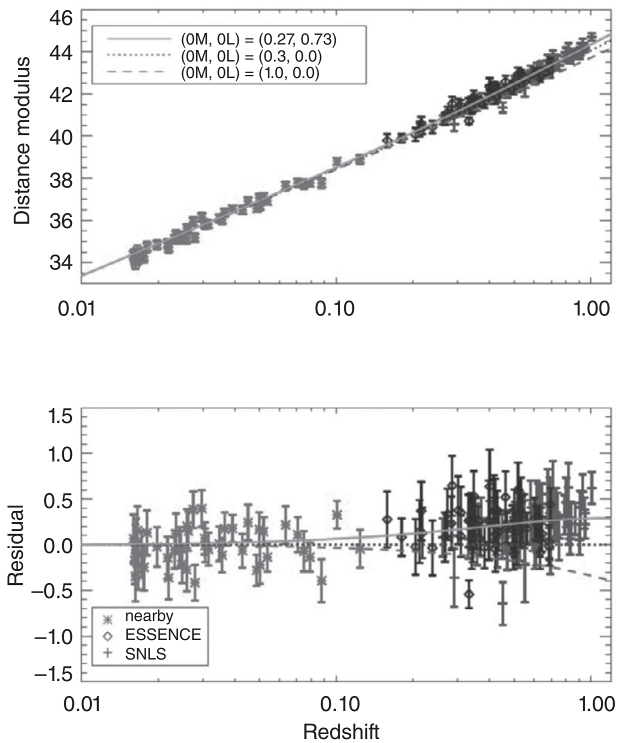


Fig. 20.17 The redshift–magnitude relation for Type Ia supernovae from the ESSENCE project. (From Krisciunas, K., arXiv astro-ph 0809.2612 (2008).)

These estimates of the cosmological parameters are in encouraging agreement with independent estimates. These include:

- *Acceleration of the Universe – Type Ia supernovae*. The demonstration of the acceleration of the Universe through the observation of Type Ia supernovae has been a triumph for the classical route to the determination of the kinematics of the Universe. This class of supernova explosion has been found to be remarkably uniform in their properties, the correlation becoming even better when account is taken of the correlation between their luminosities at maximum light and the duration of the outburst. In Fig. 20.17, the data are taken from the results of the ESSENCE project, which had the objective of measuring the redshifts and distances of about 200 Type Ia supernovae. The observations are consistent with a finite, positive value of the cosmological constant. The over-plotted lines show the expected relations for various cosmological models, including the preferred ΛCDM model with $\Omega_m = 0.27$ and $\Omega_v = 0.73$, which provides the best fit to the data.

- *Value of Hubble's constant by independent routes*. During the 1980s and 1990s, a major effort was made to resolve the differences between the various estimates of Hubble's constant, much of it stimulated by the capability of the Hubble Space Telescope (HST) to measure Cepheid variable stars out to the Virgo cluster of galaxies. The team led by Wendy Freedman used not only the HST data, but also all the other distance measurement techniques to ensure internal self-consistency of the distance estimates. The final result of the project published in 2001 was $H_0 = 70 \pm 7$ km s^{-1} Mpc^{-1}, where the errors are 1-σ errors.[24] More recent estimates indicate a value of 73.8 ± 2.4 km s^{-1} Mpc^{-1}. There is an ongoing debate about the discrepancy with the Planck estimate given in Table 20.1. A value of about 70 km s^{-1} Mpc^{-1} would probably be consistent with both estimates, that is, agreement within about 5%.
- *Mass density of the Universe from infall into large-scale structures.* The average mass density in the Universe at the present epoch is dominated by the dark matter. Studies of the average amount of dark matter estimated from the infall of galaxies into large-scale structures have found values of $\Omega_m \approx 0.3$, consistent with studies of individual clusters of galaxies.[25]
- *The abundances of the light elements by primordial nucleosynthesis*. The value of the density parameter in baryonic matter found from studies of primordial nucleosynthesis, $\Omega_B h^2 = 0.022^{+0.003}_{-0.022}$, discussed in Section 20.8, is in encouraging agreement with the values found from studies of the fluctuation spectrum of the CMB.
- *Age of the Galaxy, the ages of the oldest stars and nucleocosmochronology*. Bolte[26] argued that the ages of the oldest globular clusters were

$$T_0 = 15 \pm 2.4 \text{ (stat)}^{+4}_{-1} \text{ (syst) Gy.} \tag{20.136}$$

  The ultra-metal-poor K giant star CS 22892-052 has a thorium abundance significantly less than its scaled Solar System abundance and so the star must have been formed much earlier than the Solar System. A lower limit to the age of CS 22892-052 of $(15.2 \pm 3.7) \times 10^9$ years was found.[27] These values are consistent with the Planck estimate of 13.799 ± 0.038 Gy.
- *The statistics of gravitational lenses.* The volume enclosed by a given redshift z increases as Ω_Λ increases. The statistics of gravitationally lensed images by intervening galaxies therefore provides a test of models with finite Ω_Λ.[28] The CLASS collaboration has reported the point-source lensing rate to be one per 690 ± 190 targets.[29] They found that the observed fraction of multiply lensed sources was consistent with flat world models, $\Omega_0 + \Omega_\Lambda = 1$, in which the density parameter in the matter Ω_0 was

$$\Omega_0 = 0.31^{+0.27}_{-0.14}(68\%)^{+0.12}_{-0.10}(\text{syst}). \tag{20.137}$$

Thus, there is very encouraging agreement between a variety of independent routes to a self-consistent set of cosmological parameters. A convenient summary of the results would be to state that:

- The large scale geometry of the Universe is flat.
- The baryonic matter is about 5% of the critical density.

- The dark matter is 26% of the critical density.
- The dark energy is 69% of the critical density.
- The epoch of reionisation occurred at a redshift of about 7 to 8.
- Hubble's constant is 70 km s^{-1} Mpc^{-1} within a few percent.

These are undoubtedly rather surprising values. We live in a Universe dominated by dark energy and dark matter. The dark matter is a difficult problem and it is perhaps surprising that, despite extraordinarily impressive cryogenic experiments to detect the particles directly and searches for new types of particles with the Large Hadron Collider, such particles have evaded detection. But, if the dark matter is a difficult problem, the origin of the dark energy is *very* difficult – it only makes its effects observable on large cosmological scales.

Now the debate has shifted to the issue of how well the parameters derived from the different routes are in fact in agreement. But we are now talking about discrepancies at the few percent level at most. This may be where the next generation of cosmological problems may lie.

These results have persuaded cosmologists that the standard Big Bang model provides the most convincing framework for studies of the origin and evolution of the Universe and its contents but the elaboration of that story requires a whole book in itself. At the same time, this new consensus model brings with it a host of problems which are stretching the imaginations of cosmologists and astrophysicists – these issues are explored in Chapter 21.

Notes

1 Many more details of these observations and the theoretical background to the standard cosmological models are contained in my book *Galaxy Formation* (2008). Berlin: Springer-Verlag. These topics are brought up-to-date in the volume *The Oxford Handbook of the History of Modern Cosmology*, eds. H. Kragh and Longair, M.S., Oxford, Oxford University Press, 2019.

2 The physics of the Sunyaev–Zeldovich effect and its applications in cosmology are described in Sunyaev, R.A. and Zeldovich, Ya.B. (1980). *Annual Review of Astronomy and Astrophysics*, **18**, 537–560.

3 Hubble, E.P. (1929). A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae, *Proceedings of the National Academy of Sciences*, **15**, 168–173.

4 Hubble, E.P. and Humason, M. (1934). The Velocity–Distance Relation among Extra-Galactic Nebulae, *Astrophysical Journal*, **74**, 43–80.

5 Robertson, H.P. (1935). Kinematics and World Structure, *Astrophysical Journal*, **82**, 284–301; Walker, A.G. (1936). On Milne's Theory of World Structure, *Proceedings of the London Mathematical Society, Series 2*, **42**, 90–127.

6 Weyl, H. (1923). Zur allgemeinen Relativitätstheorie (On the theory of general relativity), *Physikalische Zeitschrift*, **29**, 230–232. Bondi, H. (1960). *Cosmology*. Cambridge: Cambridge University Press.

7 Einstein, A. (1917). Kosmologische Betrachtungen zur Allgemeinen Relativitätstheorie (Cosmological considerations in the general theory of relativity), *Sitzungsberichte, Königlich Preussische Akademie der Wissenschaften (Berlin)*, I, 142–152.

8 Friedman, A.A. (1922). On the Curvature of Space, *Zeitschrift für Physik*, **10**, 377–386; (1924). On the Possibility of a World with Constant Negative Curvature, *Zeitschrift für Physik*, **12**, 326–332.
9 Lemaître, G. (1927). A Homogeneous Universe of Constant Mass and Increasing Radius, Accounting for the Radial Velocity of Extra-Galactic Nebulae, *Annales de la Société Scientifique de Bruxelles*, **A47**, 29–39. Translation in *Monthly Notices of the Royal Astronomical Society*, **91**, 483–490 (1931).
10 Mattig, W. (1958), Über den Zusammenhang zwischen Rotverschiebung und scheinbarer Helligkeit, *Astronomische Nachrichten*, **284** (3), 109. doi:10.1002/asna.19572840303.
11 Carroll, S.M., Press, W.H. and Turner, E.L. (1992). *Annual Review of Astronomy and Astrophysics*, **30**, 499–542.
12 Bondi, H. and Gold, T. (1948). The Steady-State Theory of the Expanding Universe, *Monthly Notices of the Royal Astronomical Society*, **108**, 252–270; Hoyle, F. (1948). A New Model for the Expanding Universe, *Monthly Notices of the Royal Astronomical Society*, **108**, 372–382.
13 Kragh, H. (1999). *Cosmology and Controversy: The Historical Development of Two Theories of the Universe*. Princeton: Princeton University Press. See also Kragh's Chapter 5 of *The Oxford Handbook of the History of Modern Cosmology*, *op. cit.*
14 Note that the more complete analysis should include the inertial mass density of all types of neutrinos and any other types of hypothetical relativistic material.
15 In fact, there is further scattering of the photons of the CMB during the epoch of reionisation of the intergalactic gas at redshifts about 7. Analysis of the Planck power spectrum of fluctuations in the CMB shows that the optical depth to Thomson scattering through these epochs was $\tau_T \approx 0.05$ and so the impact on the temperature fluctuations in the CMB is small, but significant cosmologically.
16 For details of the coupling of matter and radiation during the epochs about the recombination era, see my book *Galaxy Formation* (2008), Berlin: Springer Verlag.
17 Details of the arguments given below are described in more detail in my book *Galaxy Formation*, *op. cit*.
18 Wagoner, R.V. (1973). *Astrophys. J.*, **148**, 3.
19 Steigman, G. (2007). Primordial Nucleosynthesis in the Precision Cosmology Era, *Annual Review of Nuclear and Particle Science*, **57**, 463–491. doi: 10.1146/annurev.nucl.56.080805.140437.
20 Kragh, H. and Longair, M.S. (2019). *The Oxford Handbook of the History of Modern Cosmology*. Oxford, Oxford University Press.
21 I have given a detailed account of the theory of the origin of structure in the Universe at an accessible level in my book *Galaxy Formation* (2008), Berlin: Springer Verlag.
22 Harrison, E.R. (1970). Fluctuations at the Threshold of Classical Cosmology, *Physical Review*, **D1**, 2726–2730. Zeldovich, Ya.B. (1972). A Hypothesis, Unifying the Structure and the Entropy of the Universe, *Monthly Notices of the Royal Astronomical Society*, **160**, 1P–3P.
23 The Planck Collaboration, ArXiv e-prints 1807.06209 (2018).
24 Freedman, W.L. *et al.* (2001). Final Results from the Hubble Space Telescope Key Project to Measure the Hubble Constant, *Astrophysical Journal*, **533**, 47–72.
25 Peacock, J.A. *et al.* (2001). A Measurement of the Cosmological Mass Density from Clustering in the 2dF Galaxy Redshift Survey, *Nature*, **410**, 169–173.
26 Bolte, M. (1997). Globular Clusters: Old, in *Critical Dialogues in Cosmology*, ed. Turok, N., 156–168. Singapore: World Scientific.
27 Cowan, J.J. *et al.* (1997). The Thorium Chronometer in CS 22892-052: Estimates of the Age of the Galaxy. *Astrophysical Journal*, **480**, 246–254.
28 Fukugita, M. *et al.* (1992). Statistical Properties of Gravitational Lenses with a Nonzero Cosmological Constant, *Astrophysical Journal*, **393**, 3–21.
29 Mitchell, J.L. *et al.* (2005). Improved Cosmological Constraints from Gravitational Lens Statistics, *Astrophysical Journal*, **622**, 81–98.

21 Dark Matter, Dark Energy and the Inflationary Paradigm

21.1 Introduction

The observations and reasoning supporting the standard ΛCDM model of contemporary cosmology were described in Chapter 20. In the course of that exposition, we foreshadowed a number of fundamental problems with the model. This chapter is devoted to analysing these issues and suggesting possible solutions to them.[1]

Much of the discussion concerns the *inflationary paradigm* for the early Universe and, as we will show, this model has some striking successes. But it involves physics far beyond what has been established in the laboratory, at energies which far exceed those available at the most powerful terrestrial accelerators such as the Large Hadron Collider (LHC) at CERN. This raises concerns about how far we can trust the model and indeed whether this is physics at all. Readers should make up their own minds about how seriously they wish to adopt the inflationary model, particularly when theorists have failed to come up with a physical realisation of it. Let us begin by analysing in more detail the basic problems with the preferred ΛCDM model.

21.2 Dark Matter and Dark Energy

21.2.1 Dark Matter

In the early days of astrophysical cosmology, there were many reasons why there should be dark matter in the Universe made out of ordinary baryonic material – low mass stars, dead stars, interstellar and intergalactic gas, dust and so on. If the matter did not radiate in the optical waveband, it was invisible. The subsequent story breaks naturally into two parts – first establishing the amount of dark matter present in the Universe and then determining whether or not it is baryonic. This endeavour was to require the full power of the information-gathering capacities of the new post-War astronomies, the advent of new technologies and associated astronomical facilities, and advances in interpretation and theory. Among key astrophysical landmarks along the way were the following.

- In 1933, Fritz Zwicky made pioneering observations of the velocity dispersion among the galaxies in the nearby rich cluster of galaxies, the Coma cluster.[2] To determine the total mass of the cluster, he used the *virial theorem*, according to which the kinetic energy of the galaxies in the cluster should be half the gravitational potential energy of the cluster.

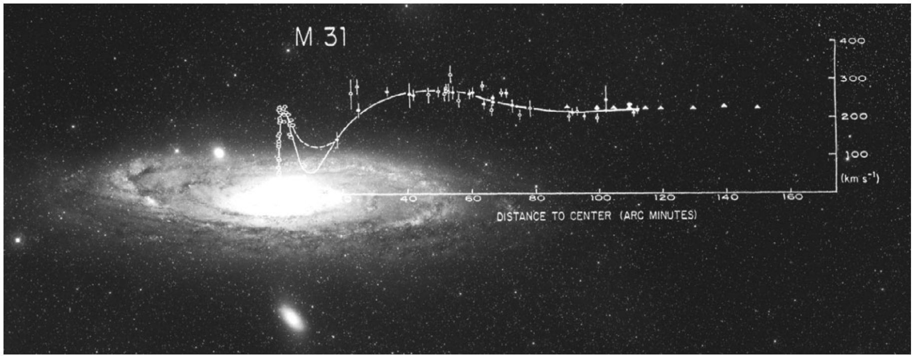


Fig. 21.1 The rotation curve for the nearby giant spiral galaxy M31, showing the flat rotation curve extending well beyond the optical image of the galaxy thanks to observations of the velocities of interstellar neutral hydrogen by Morton Roberts and his colleagues. (Courtesy of the late Vera Rubin.)

This leads to an expression for the total mass of the cluster, $M \approx 2R_{\rm cl}\langle v^2\rangle/G$, where $R_{\rm cl}$ is the cluster radius and $\langle v^2\rangle$ is the velocity dispersion of the galaxies in the cluster. In solar units, the mass-to-luminosity ratio was found to be about 500, about 100 times greater than that of a galaxy such as our own. Interpreted literally, there must be about 50 to 100 times more dark as compared with visible matter. This important result has been confirmed by all subsequent studies of the masses of rich clusters of galaxies.

- Once powerful long-slit optical spectroscopic facilities became available, the rotation curves of galaxies could be traced well beyond their central regions and flat rotation curves were observed by Vera Rubin and her colleagues. At the same time, studies of spiral galaxies using the 21-cm line of neutral hydrogen determined the rotation curves to much greater distances from their centres than optical observations, and showed that flat rotation curves are common among giant spiral galaxies (Fig. 21.1). The significance of these flat rotation curves can be appreciated from the following argument. Assuming that the distribution of mass in the galaxy is spherically symmetric, the mass within radius r is $M(\leq r)$. According to Gauss's law for gravity, we can find the radial acceleration at radius r by placing the mass within radius r, $M(\leq r)$, at the centre of the galaxy. Then, equating the centripetal acceleration at radius r to the gravitational acceleration, we find $M(\leq r) = v_{\rm rot}^2(r)\, r/G$. If the rotation curve of the spiral galaxy is flat, $v_{\rm rot}$ = constant, $M(\leq r) \propto r$ and so the mass within radius r increases linearly with distance from the centre. This contrasts with the distributions of light in the discs, bulges and haloes of spiral galaxies, which decrease exponentially with increasing distance from the centre. Consequently, the mass-to-luminosity ratio must increase dramatically with distance from the centres of the galaxies.
- Theoretical studies of the stability of the mass distributions in disc galaxies by Miller, Prendergast and Hohl found that these were unstable.[3,4] A solution to this problem was found by Ostriker and Peebles,[5] who showed that, if the galaxies had dark matter haloes, the discs of spiral galaxies would be stabilized.
- X-ray imaging of the hot intergalactic gas in rich clusters of galaxies enabled the total mass of the cluster and the mass of hot intracluster gas to be determined. The

extremely hot X-ray emitting gas forms a roughly isothermal atmosphere within the cluster gravitational potential from which the masses of the hot gas and the total gravitating mass of the cluster can be found. Typically, there is much more dark mass than the mass associated either with the intergalactic gas or the visible parts of galaxies in the cluster.[6] These results confirmed Zwicky's pioneering measurements.

- The values of the density parameters in baryonic matter estimated from primordial nucleosynthesis arguments and from the form of the power spectrum of the CMB correspond to $\Omega_B \sim 0.05$ and $\Omega_d \sim 0.26$ respectively
- Finally, the low limits to the fluctuations in the CMB forced cosmologists to take non-baryonic dark matter really seriously in the early 1980s in order to account for the formation of structure in the Universe by the present epoch, while depressing the predicted level of intensity fluctuations in the CMB below the observational upper limits.

The upshot of all these considerations was that, from the early 1980s onwards, purely baryonic dark matter theories of structure formation became untenable and non-baryonic dark matter had to be taken really seriously.

21.2.2 Baryonic Dark Matter and Black Holes

The evidence of Section 21.2.1 indicates unambiguously that ordinary baryonic dark matter cannot account for the total amount of dark matter inferred to be present in the Universe by a factor of about ten. By 'ordinary' baryonic dark matter, we mean matter composed of protons, neutrons and electrons. Thanks to advances in technology, many of the 'invisible' forms of baryonic dark matter have now been detected. For example, brown dwarf stars, which have masses less than about 0.08 $M_\odot$ and are not hot enough to burn hydrogen into helium in their cores, have been discovered in abundance in the 2MASS 2-μm infrared sky survey, as well as by the NICMOS infrared imager of the Hubble Space Telescope. The dramatic discovery of thousands of extra-Solar System planets, or exoplanets, does not make a significant contribution to the baryonic dark matter. These dark objects cannot account for the observed amount of dark matter.

Black holes are potential candidates for the dark matter. The supermassive black holes in the nuclei of galaxies have masses which are typically only about 0.1% of the mass of the bulges of their host galaxies and so they contribute negligibly to the mass density of the Universe.

There might, however, be a population of intergalactic massive black holes. Limits to their number density can be set in certain mass ranges from studies of the numbers of gravitationally lensed galaxies observed in large samples of extragalactic radio sources. Jacqueline Hewitt and her colleagues set limits to the number density of massive black holes with masses in the range $10^{10} \leq M \leq 10^{12} M_\odot$ in a major dedicated radio survey with the Very Large Array (VLA). The numbers of gravitationally lensed images corresponded to $\Omega_{BH} \ll 1$.[7]

The same technique using very long baseline interferometry (VLBI) can be used to study the mass density of lower mass black holes by searching for the gravitationally

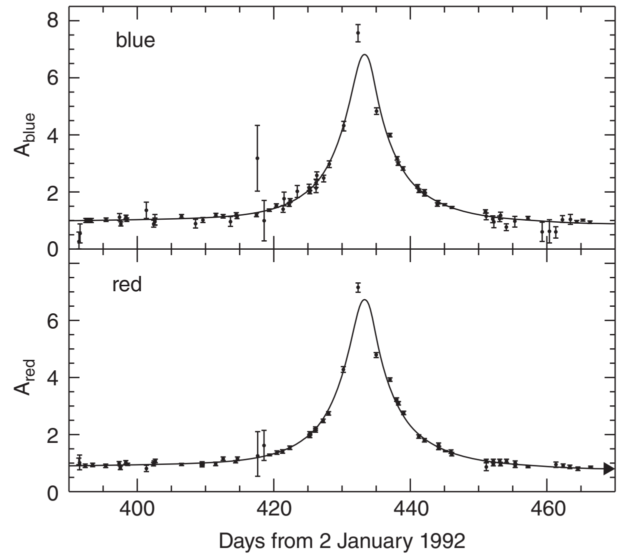


Fig. 21.2 The gravitational microlensing event recorded by the MACHO project in February and March 1993. The horizontal axis shows the date in days measured from day zero on 2 January 1992. The vertical axis shows the amplification of the brightness of the lensed star relative to the unlensed intensity in blue and red wavebands. The solid lines show the expected variations of brightness of a lensed star with time. The same characteristic light curve is observed in both wavebands, as expected for a gravitational microlensing event. (From Alcock, C. *et al.* 1993, Possible Gravitational Microlensing of a Star in the Large Magellanic Cloud, *Nature*, **365**, 621–623.)

lensed images on an angular scale of milliarcseconds, corresponding to masses in the range $10^6 \leq M \leq 10^8 M_\odot$.[8] Wilkinson and his colleagues searched a sample of 300 compact radio sources for examples of multiple gravitationally lensed images but none was found. The upper limit to the cosmological mass density of intergalactic supermassive compact objects in the mass range $10^6 \leq M \leq 10^8 M_\odot$ corresponded to less than 1% of the critical cosmological density.[9]

An impressive approach to setting limits to the contribution which discrete low mass objects, collectively known as MAssive Compact Halo Objects, or MACHOs, could make to the dark matter in the halo of our own Galaxy has been the search for gravitational microlensing signatures of such objects as they pass in front of background stars. The MACHOs include low mass stars, white dwarfs, brown dwarfs, planets and black holes. These events are very rare and so very large numbers of background stars have to be monitored. The beauty of this technique is that it is sensitive to MACHOs with a very wide range of masses, from 10^{-7} to 100 $M_\odot$, and so the contributions of a very wide range of candidates for the dark matter can be constrained. In addition, the expected light curves of such gravitational lensing events have a characteristic form which is independent of wavelength. Two large projects, the MACHO and the EROS projects, have made systematic surveys over a number of years to search for these events. The first example of a microlensing event was discovered in October 1993 (Fig. 21.2), the mass of the invisible lensing object being estimated to lie in the range $0.03 < M < 0.5\ M_\odot$.

By the end of the MACHO project, many lensing events had been observed, including over 100 in the direction towards the Galactic bulge, about three times more than expected. In addition, 13 definite and 4 possible events were observed in the direction of the Large Magellanic Cloud.[10] The numbers are significantly greater than the 2–4 detections expected from known types of star. The statistics are consistent with MACHOs making up about 20% of the necessary halo mass, the 95% confidence limits being 8–50%. Somewhat fewer microlensing events were detected in the EROS project, which found than less than 25% of the mass of the standard dark matter halo could consist of dark objects with masses in the range 2×10^{-7} to 1 $M_\odot$ at the 95% confidence level.[11] The most likely candidates for the detected MACHOs would appear to be white dwarfs, produced in large numbers in the early evolution of the Galaxy, but other more exotic possibilities cannot be excluded. The consensus view is that MACHOs alone cannot account for all the dark matter in the halo of our Galaxy and so some form of non-baryonic matter must make up the difference.

21.2.3 Non-Baryonic Dark Matter

The dark matter might consist of the types of particle predicted by theories of elementary particles but not yet detected experimentally, for example, the axions or Weakly Interacting Massive Particles (WIMPs). In addition, neutrinos with finite rest masses might make a significant contribution.

Axions

The smallest mass candidates are the *axions*, which were invented by particle theorists in order to 'save quantum chromodynamics from strong CP violation'. If the axions were produced in thermal equilibrium, they would have unacceptably large masses, which would result in conflict with observations of the Sun and the supernova SN1987A. Specifically, if the mass of the axion were greater than 1 eV, the rate of loss of energy by the emission of axions would exceed the rate at which energy is generated by nuclear reactions in the Sun and so its centre would need to be hotter, resulting in a shorter age than is acceptable and greater emission of high energy neutrinos. There is, however, another non-equilibrium route by which the axions could be created in the early Universe. If they exist, they must have been created when the thermal temperature of the Universe was about 10^{12} K, but they were out of equilibrium and never acquired thermal velocities – they remained 'cold'. Their rest mass energies are expected to lie in the range 10^{-2} to 10^{-5} eV. The roles of such particles in cosmology and galaxy formation are discussed by Efstathiou and by Kolb and Turner.[12]

WIMPs

Examples of *Weakly Interacting Massive Particles*, or WIMPs, might be the gravitino, the supersymmetric partner of the graviton, or the photino, the supersymmetric partner of the photon, or some form of as yet unknown massive neutrino-like particle. In particular, the dark matter might be in the form of the lightest supersymmetric particle which is expected to be stable.

There must, however, be a suppression mechanism to avoid the problem that, if the WIMPs were as common as the photons and neutrinos, their masses cannot be greater than about 30 eV. The physics of this process is described by Kolb and Turner.[13] According to particle theorists, almost all theories of physics beyond the standard model involve the existence of new particles at the TeV scale because of the symmetries which have to be introduced to avoid proton decay and violations of the precision tests of the electroweak theory. These considerations lead to the expectation of new particles appearing at the weak energy scale.

An example of the type of experiment which could demonstrate the presence of new particles has been carried out at the LHCb and CMS experiments at CERN. The cross-section of the extremely rare decay of the B_s meson into two muons has been measured. The observed branching fraction for this process compared with the predictions of the standard model provides a means of searching for physics beyond the standard model. The measurements were statistically compatible with standard model predictions and so allow stringent constraints to be placed on theories beyond the standard model. This experiment, involving the simplest of the routes to the detection of supersymmetric particles, gave a null result, but this does not rule out the importance of this type of search for supersymmetric particles since there are other ways in which they could be involved in particle decays.[14]

Neutrinos with Finite Rest Mass

The three known types of neutrino have finite rest masses. Laboratory tritium β-decay experiments[15] have provided an upper limit to the rest mass of the electron antineutrino of $m_\nu \le 2$ eV. This measurement does not exclude the possibility that the two other types of neutrino, the μ and τ neutrinos, could have greater masses. However, the discovery of neutrino oscillations has provided a measurement[16] of the mass difference between the μ and τ neutrinos of $\Delta m_\nu^2 \sim 3 \times 10^{-3}$. Thus, although their masses are not measured directly, they probably have masses of the order of 0.1 eV. This can be compared with the typical neutrino rest mass needed to attain the critical cosmological density of about 10–20 eV. Hence, known types of neutrino with finite rest masses cannot constitute the dark matter.

21.2.4 Astrophysical and Experimental Limits

Useful astrophysical limits can be set to the number densities of different types of neutrino-like particles in the outer regions of giant galaxies and in clusters of galaxies. The WIMPs and massive neutrinos are collisionless fermions and therefore there are constraints on the phase space density of these particles, which translate into a lower limit to their masses.[17]

Being fermions, neutrino-like particles are subject to the Pauli exclusion principle according to which there is a maximum number of particle states in phase space for a given momentum $p_{\rm max}$. It is a straightforwad calculation to show that the resulting lower bound to the mass of the neutrino is

$$m_\nu \ge \frac{1.5}{(N_\nu \sigma R^2)^{1/4}}\ \text{eV}, \tag{21.1}$$

where N_ν is the number of neutrino species, the velocity dispersion σ is measured in units of 10^3 km s^{-1} and R is measured in Mpc.

In rich clusters of galaxies, typical values are $\sigma = 1000$ km s^{-1} and $R = 1$ Mpc. If there is only one neutrino species, $N_\nu = 1$, we find $m_\nu \geq 1.5$ eV. If there were six neutrino species, namely, electron, muon, tau neutrinos and their antiparticles, $N_\nu = 6$ and then $m_\nu \geq 0.9$ eV. For giant galaxies, for which $\sigma = 300$ km s^{-1} and $R = 10$ kpc, $m_\nu \geq 20$ eV if $N_\nu = 1$ and $m_\nu \geq 13$ eV if $N_\nu = 6$. For small galaxies, for which $\sigma = 100$ km s^{-1} and $R = 1$ kpc, the corresponding figures are $m_\nu \geq 80$ eV and $m_\nu \geq 50$ eV respectively. Thus, particles with rest masses $m_\nu \sim 1$ eV could bind clusters of galaxies but they could not bind the haloes of giant or small galaxies.

Searches for different types of dark matter particles have developed into one of the major areas of *astroparticle physics*. An important class of experiment involves the search for weakly interacting particles with masses $m \geq 1$ GeV, which could make up the dark halo of our Galaxy. In order to form a bound dark halo about our Galaxy, the particles would have to have velocity dispersion $\langle v^2 \rangle^{1/2} \sim 230$ km s^{-1} and their total mass is known. Therefore, the number of WIMPs passing through a terrestrial laboratory each day is a straightforward calculation. When these massive particles interact with the sensitive volume of the detector, the collision results in the transfer of momentum to the nuclei of the atoms of the material of the detector and this recoil can be measured in various ways. The challenge is to detect the very small number of events expected because of the very small cross-section for the interaction of WIMPs with the nuclei of atoms. A typical estimate is that less than one WIMP per day would be detectable by 1 kilogram of detector material. There should be an annual modulation of the dark matter signal as a result of the Earth's motion through the Galactic halo population of dark matter particles.

A good example of the quality of the data now available is provided by the results of the Super Cryogenic Dark Matter Search (SuperCDMS) at the Soudan Laboratory.[18] With an exposure of 1690 kg days, only a single candidate event was observed, consistent with the expected background in the detector. The upper limit to the spin-independent WIMP–nucleon cross-section is $(1.4 \pm 1.0) \times 10^{-44}$ cm^2 at 46 GeV c^{-2}. These results are the strongest limits to date for WIMP–germanium-nucleus interactions for masses greater than 12 GeV c^{-2}.

21.2.5 The Dark Energy

The evidence for an accelerating universe from the redshift–magnitude relation for Type Ia supernovae and from the many different aspects of analyses of the properties of the power spectrum of the fluctuations in the CMB is unambiguous (Section 20.9). It is particularly impressive that, using a scalar power spectrum for the primordial fluctuations, the polarisation power spectra and the large-scale power spectrum of the dark matter derived from the Planck observations, the six-parameter family of the best-fit model can be derived without recourse to any other observation.[19]

But there is more to it than that. The ΛCDM model solves the problem of creating the large-scale structure of the distribution of dark matter in a simple and elegant manner without the need to patch it up essentially arbitrarily with astrophysical phenomena, which is necessary in other conceivable models.[20]

The contrast between the dark matter and the dark energy problems is striking. The estimates of the amount of dark matter depend on Newtonian gravity in domains in which we can have a great deal of confidence that it is the appropriate theory of gravity. The dark matter is acted upon by gravity in the usual way, whereas the dark energy term in Einstein's equations does not depend upon the mass distribution, as can be seen from the expression for the variation of the scale factor a with cosmic time, in which there is no density dependence on the Λ term:

$$\ddot{a} = -\frac{4\pi G}{3} a \left(\varrho + \frac{3p}{c^2} \right) + \tfrac{1}{3} \Lambda a. \tag{21.2}$$

The Λ term provides a uniform background against which the evolution of the contents of the Universe unfold. It only makes its presence known on the largest scales observable at the present epoch and becomes of decreasing importance at large redshifts.

There is also the issue of on which side of the Einstein field equation the cosmological constant term should appear. This equation can be written with the geometrical term $G_{\mu\nu}$ on the left-hand side and the stress-energy tensor $T_{\mu\nu}$ on the right:

$$G_{\mu\nu} - \tfrac{1}{2}\, g_{\mu\nu}\, G - \Lambda g_{\mu\nu} = -\kappa T_{\mu\nu}. \tag{21.3}$$

Is the Λ-term part of the intrinsic geometry, in which case it should appear on the left-hand side, or is it a source term driving the acceleration, in which case it should be included on the right-hand side of (21.3)? These are hard questions to answer observationally, but some aspects of them are feasible. Thus, if the dark energy terms were to change with cosmic epoch, that would suggest that it is a physical field. Experiments such as the Euclid project of the European Space Agency and the WFIRST mission of NASA aim to tackle this issue.

21.3 The Big Problems

Let us review the Big Problems which have been lurking in the background during the development of the standard ΛCDM picture (Section 20.9).

21.3.1 The Horizon Problem

The horizon problem, clearly recognised by Robert Dicke in 1961, is the question 'Why is the Universe so isotropic?'[21] At earlier cosmological epochs, the particle horizon $r \sim ct$ encompassed less and less mass and so the scale over which particles could be causally connected was smaller and smaller. For illustration, how far could light have travelled along the last scattering surface at $z \sim 1000$ since the Big Bang? In matter-dominated models, this distance is $r = 3ct$, corresponding to an angle $\theta_{\rm H} \approx 2°$ on the sky. Thus, regions of the sky separated by angular distances greater than this could not have been in causal communication. Why then is the CMB so isotropic? How did causally separated regions 'know' that they had to have the same temperature to better than one part in 10^5?

21.3.2 The Flatness Problem

Why is the Universe geometrically flat, $\Omega_K = 1$? The flatness problem was also recognised by Dicke in 1961 and reiterated by Dicke and Peebles in 1979 for standard world models with $\Omega_v = 0$.[22] In its original version, the problem arises from the fact that, according to the standard world models, if the Universe were set up with a value of the density parameter differing even slightly from the critical value $\Omega = 1$, it would diverge very rapidly from this value at later epochs. As shown in Section 20.5.1, if the Universe has density parameter Ω_0 today, at redshift z, $\Omega(z)$ was

$$\left[1 - \frac{1}{\Omega(z)}\right] = f(z)\left[1 - \frac{1}{\Omega_0}\right], \tag{21.4}$$

where $f(z) = (1+z)^{-1}$ for the matter-dominated era and $f(z) \propto (1+z)^{-2}$ during the radiation-dominated era. Thus, since $\Omega_0 \sim 0.3$ at the present epoch, it must have been extremely close to the critical value in the remote past. Alternatively, if $\Omega(z)$ had departed from $\Omega(z) = 1$ at a very large redshift, Ω_0 would be very far from $\Omega_0 = 1$ today. The only 'stable' value is $\Omega_0 = 1$. There is nothing in the standard world models that would lead us to prefer any particular value of Ω_0. This is referred to as the *fine-tuning problem*.

When Dicke described the horizon problem, the value of the overall density parameter was poorly known, but his argument was still compelling. Now we know that the value of the overall density parameter is $\Omega_K = \Omega_d + \Omega_b + \Omega_v = 1.001 \pm 0.002$ (Section 20.9) – why is the Universe geometrically flat?

21.3.3 The Baryon-Asymmetry Problem

The baryon-asymmetry problem arises from the fact that the photon-to-baryon ratio today is

$$\frac{N_\gamma}{N_b} = \frac{4 \times 10^7}{\Omega_b h^2} = 1.6 \times 10^9, \tag{21.5}$$

where Ω_b is the density parameter in baryons and the values of Ω_b and h have been taken from Table 20.1. If photons are neither created nor destroyed, this ratio is conserved as the Universe expands. At temperature $T \approx 10^{10}$ K, electron–positron pair production takes place from the photon field. At a correspondingly higher temperature, baryon–antibaryon pair production takes place with the result that there must have been a very small asymmetry in the baryon–antibaryon ratio in the very early Universe if we are to end up with the observed photon-to-baryon ratio at the present day. At these very early epochs, there must have been roughly $10^9 + 1$ baryons for every 10^9 antibaryons to guarantee the observed ratio of photons to baryons at the present epoch. If the Universe had been symmetric with respect to matter and antimatter, the photon-to-baryon ratio would now be about 10^{18}, in gross contradiction with the observed value.[23] Therefore, there must be some mechanism in the early Universe which results in a slight asymmetry between matter and antimatter. Fortunately, we know that spontaneous symmetry breaking results in a slight imbalance between various classes of mesons and so there is hope that this can be explained by 'standard' particle physics, but the precise mechanism has not been identified.

21.3.4 The Primordial Fluctuation Problem

What was the origin of the density fluctuations from which galaxies and large-scale structures formed? One of the key results of studies of the development of small density perturbations in the early Universe is that, when these came through their cosmological horizons, meaning when the scale of the perturbations was of the order $r = ct$ where t is cosmic time, the amplitudes of the density perturbations had to be of finite amplitude, $\delta\varrho/\varrho \sim 10^{-4}$, on a very wide range of mass scales. These could not have originated as statistical fluctuations in the numbers of particles on, say, the scales of superclusters of galaxies. There must have been some physical mechanism which generated finite amplitude perturbations with power spectrum close to $P(k) \propto k$ in the early Universe. This line of reasoning led Lemaître, Tolman and Lifshitz to conclude that galaxies could not have formed by gravitational collapse. Others, such as Zeldovich, Novikov, Peebles and their colleagues adopted the more pragmatic approach that the fluctuations were a relic of the early Universe and then worked out their consequences.[24]

21.3.5 The Values of the Cosmological Parameters

The horizon and flatness problems were recognised before compelling evidence was found for the finite value of the cosmological constant. The values for the cosmological parameters create their own problems. The density parameters in the dark matter and the dark energy are of the same order of magnitude at the present epoch but the matter density evolves with redshift as $(1 + z)^3$, while the dark energy density is unchanging with cosmic epoch. Why then do we live at an epoch when they have more or less the same values?

As discussed in Section 20.5.3, although the vacuum energy density and the discovery of the Higgs fields provided realisations of negative pressure fields of the types needed to account for the cosmological constant, the theoretical value of Ω_v is about 10^{120} times greater than its observed value. How did this come about?

As if these problems were not serious enough, they are compounded by the fact that the nature of the dark matter and the dark energy are unknown. Thus, one of the consequences of precision cosmology is the unsettling result that we do not understand the nature of about 95% of the material which drives the large-scale dynamics of the Universe. The measured values of the cosmological parameters listed in Table 20.1 really are extraordinary – wherever did they come from?

Rather than being causes for despair, however, these problems should be seen as the great challenges for the astrophysicists and cosmologists of the twenty-first century. It is not too far-fetched to see an analogy with Bohr's theory of the hydrogen atom, which was an unsatisfactory mix of classical and primitive quantum ideas, but which ultimately led to completely new and deep insights with the development of quantum mechanics.[25]

21.3.6 The Way Ahead

In the standard Friedman models, the big problems are solved by assuming that the Universe was endowed with appropriate initial conditions in its very early phases. To put it crudely, we get out at the end what we put in at the beginning. In a truly physical picture of our Universe, we should aim to do better than this.

I have suggested five possible approaches to solving these problems.[26]

- That is just how the Universe is – the initial conditions were set up that way.
- There are only certain classes of universe in which 'intelligent' life could have evolved. The Universe has to have the appropriate initial conditions and the fundamental constants of nature should not be too different from their measured values or else there would be no chance of life forming as we know it. This approach involves the *Anthropic Cosmological Principle* according to which, in a provocative version, it is asserted that the Universe is as it is because we are here to observe it.
- The inflationary scenario for the early Universe. This topic is taken up in Section 21.5 and subsequent sections.
- Seek clues from particle physics and extrapolate that understanding beyond what has been confirmed by experiment to the earliest phases of the Universe.
- Something else we have not yet thought of. We can think of this in terms of what Donald Rumsfeld called the 'unknown unknowns – the ones we don't know we don't know'.[27] This would certainly involve new physical concepts.

Let us consider some aspects of these approaches.

The Limits of Observation

Even the first, somewhat defeatist, approach might be the only way forward if it turned out to be just too difficult to disentangle convincingly the physics responsible for setting up the initial conditions from which our Universe evolved. In 1970, William McCrea considered the fundamental limitations involved in asking questions about the very early Universe, his conclusion being that we can obtain less and less information the further back in time one asks questions about the early Universe.[28] A modern version of this argument would be framed in terms of the limitations imposed by the existence of a last scattering surface for electromagnetic radiation at $z \approx 1000$ and those imposed on the accuracy of observations of the CMB and the large-scale structure of the Universe because of their cosmic variances.

In the case of the CMB, the observations made by the Planck experiment are already cosmic variance limited for multipoles $l \leq 2000$ – we will never be able to learn much more than we know already about the form of the scalar power spectrum on these angular scales. In these studies, the search for new physics will depend upon discovering discrepancies between the standard model and future observations. The optimists would argue that the advances will come through extending our technological capabilities so that new classes of observation become cosmic variance limited. For example, the detection of primordial gravitational waves through their polarisation signature at small multipoles in the CMB, the nature of dark matter particles and the nature of the vacuum energy are at the cutting edge of fundamental issues for astrophysical cosmology. These approaches will be accompanied by discoveries in particle physics with future generations of ultra-high energy particle experiments.

It is also salutary to recall that the range of particle energies which have been explored by the most powerful particle accelerators is up to about 200 GeV, corresponding to a cosmological epoch about 1 microsecond after the Big Bang. This seems very modest compared with the Planck era which occurred at $t \sim 10^{-43}$ s. Is there really no new physics to be discovered between these epochs?

It is folly to attempt to predict what will be discovered over the coming years, but we might run out of luck. How would we then be able to check that the theoretical ideas proposed to account for the properties of the very early Universe are correct? Can we do better than boot-strapped self-consistency? The great achievement of modern cosmology has been that we have made enormous strides in defining a convincing framework for astrophysical cosmology through precise observation combined with robust application of the best theoretical physics has to offer. As a result, the basic problems identified above can now be addressed as areas of genuine scientific enquiry.

The Anthropic Cosmological Principle

There is certainly some truth in the fact that our ability to ask questions about the origin of the Universe says something about the sort of universe we live in. The Cosmological Principle asserts that we do not live at any special location in the Universe, and yet we are certainly privileged in even being able to ask the questions. In this line of reasoning, there are only certain types of universe in which life as we know it could have formed. For example, the stars must live long enough for there to be time for biological life to form and evolve into sentient beings. This line of reasoning is embodied in the *Anthropic Cosmological Principle*, first expounded by Carter in 1974,[29] and dealt with *in extenso* in the books by Barrow and Tipler and by Gribbin and Rees.[30] Part of the problem stems from the fact that we have only one universe to study – we cannot go out and investigate other universes to see if they have evolved in the same way as ours. There are a number of versions of the Anthropic Principle, some of them stronger than others. In extreme interpretations, it leads to statements such as the strong form enunciated by Wheeler,[31]

> Observers are necessary to bring the Universe into being.

It is a matter of taste how seriously one takes this line of reasoning. To many cosmologists, it is not particularly appealing because it suggests that it will never be possible to find physical reasons for the initial conditions from which the Universe evolved, or for the values of the fundamental constants of nature. But some of these problems are really hard. Weinberg, for example, found it such a puzzle that the vacuum energy density Ω_v is so very much smaller than the values expected according to current theories of elementary particles, that he invoked anthropic reasoning to account for its smallness.[32] Another manifestation of this type of reasoning is to invoke the range of possible initial conditions which might come out of the picture of chaotic or eternal inflation[33] and argue that, if there were at least 10^{120} of them, then we live in one of the few which has the right conditions for life to develop as we know it. It is left to the reader how seriously these ideas should be taken.[34] My personal view is that we might regard the Anthropic Cosmological Principle as the very last resort if all other physical approaches fail.

21.4 A Pedagogical Interlude: Distances and Times in Cosmology

Let us summarise the various times and distances used in the study of the early Universe.[35] This short pedagogical interlude may help understand the nature of the problems discussed above and potential routes to their resolution.

Comoving radial distance coordinate In order to define a self-consistent distance at a specific cosmic epoch t, we projected the proper distances along our past light cone to a reference epoch which we took to be the present epoch t_0. In terms of cosmic time and scale factor a, the comoving radial distance coordinate r is defined to be

$$r = \int_t^{t_0} \frac{c\,\mathrm{d}t}{a} = \int_a^1 \frac{c\,\mathrm{d}a}{a\dot{a}}. \tag{21.6}$$

Recall that this is a 'fictitious distance' introduced to provide a suitable distance measure to include in the Robertson–Walker metric.

Proper radial distance coordinate The same problem arises in defining a proper distance at an earlier cosmological epoch. We *define* the proper radial distance x to be the comoving radial distance coordinate projected back to the epoch t (see 20.26). From (20.12), we find

$$x = a \int_t^{t_0} \frac{c\,\mathrm{d}t}{a} = a \int_a^1 \frac{c\,\mathrm{d}a}{a\dot{a}}. \tag{21.7}$$

Particle horizon The particle horizon r_H is defined as the maximum proper distance over which there can be causal communication at the epoch t,

$$r_\mathrm{H} = a \int_0^{t} \frac{c\,\mathrm{d}t}{a} = a \int_0^a \frac{c\,\mathrm{d}a}{a\dot{a}}. \tag{21.8}$$

Radius of the Hubble sphere or the Hubble radius The Hubble radius is the proper radial distance of causal contact *at a particular epoch*. It is the distance at which the velocity in the velocity–distance relation at that epoch is equal to the speed of light. This Hubble sphere has proper radius

$$r_\mathrm{HS} = \frac{c}{H(z)} = \frac{ac}{\dot{a}}. \tag{21.9}$$

For example, (21.9) can be used to estimate the wavenumber of the first maximum in the power spectrum of fluctuations in the CMB at the epoch of recombination.

Event horizon The event horizon r_E is defined as the greatest proper radial distance an object can have if it is ever to be observable by an observer who observes the Universe at cosmic time t_1:

$$r_\mathrm{E} = a \int_{t_1}^{t_\mathrm{max}} \frac{c\,\mathrm{d}t}{a(t)} = a \int_{a_1}^{a_\mathrm{max}} \frac{c\,\mathrm{d}a}{a\dot{a}}. \tag{21.10}$$

Cosmic time Cosmic time t is defined to be time measured by a fundamental observer who reads time on a standard clock:

$$t = \int_0^t \mathrm{d}t = \int_0^a \frac{\mathrm{d}a}{\dot{a}}. \tag{21.11}$$

Conformal time The conformal time is found by projecting time intervals along the past light cone to the present epoch, using the cosmological time dilation relation. There are similarities to the definition of comoving radial distance coordinate:

$$\mathrm{d}t_{\mathrm{conf}} = \mathrm{d}\tau = \frac{\mathrm{d}t}{a}. \tag{21.12}$$

Thus, according to the cosmological time dilation formula, the interval of conformal time is what would be measured by a fundamental observer observing distant events at the present epoch t_0. At any epoch, the conformal time has value

$$\tau = \int_0^t \frac{\mathrm{d}t}{a} = \int_0^a \frac{\mathrm{d}a}{a\dot{a}}. \tag{21.13}$$

The past light cone This topic requires a little care because of the way in which the standard models are set up in order to satisfy the requirements of isotropy and homogeneity. For example, Hubble's law in the form $v = H_0 r$ can be applied at the present epoch to recessions speeds which can exceed the speed of light because r is the 'artificial' comoving radial distance coordinate. Consider the proper distance between two fundamental observers at some epoch t:

$$x = a(t)r, \tag{21.14}$$

where r is the comoving radial distance coordinate. Differentiating with respect to cosmic time,

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \dot{a}r + a\frac{\mathrm{d}r}{\mathrm{d}t}. \tag{21.15}$$

The first term on the right-hand side represents the motion of the substratum and corresponds to a recession velocity $v(t)$ at any epoch since

$$v(t) = H(t)\,x = \frac{\dot{a}}{a}\,ar = \dot{a}r.$$

The second term on the right-hand side of (21.15) corresponds to the velocity of peculiar motions in the local rest frame at r. The element of proper radial distance is $a\,\mathrm{d}r$ and so, if we consider a light ray travelling towards the observer at the origin along the past light cone, we find

$$v_{\mathrm{tot}} = \dot{a}r - c. \tag{21.16}$$

This result defines the propagation of light from the source to the observer in space-time diagrams for the expanding Universe.

We can now plot the trajectories of light rays from their source towards the observer at t_0. The proper radial distance to the proper time axis at $r = 0$, r_{PLC} is

$$r_{\mathrm{PLC}} = \int_0^t v_{\mathrm{tot}}\,\mathrm{d}t = \int_0^a \frac{v_{\mathrm{tot}}\,\mathrm{d}a}{\dot{a}}. \tag{21.17}$$

Inspection of (21.16) shows that, at large enough values of r, the comoving radial distance coordinate, the light rays can have positive values of v_{tot}, meaning that they are propagating away from the observer. The light waves propagate to the observer at the present epoch through local inertial frames, which expand with progressively smaller velocities until they cross the *Hubble sphere* at which the recession velocity of the local frame of reference is the speed of light. From this epoch onwards, propagation is towards the observer until, as $t \to t_0$, the speed of propagation towards the observer is the speed of light.

These phenomema are best illustrated by specific examples of the standard cosmological models. We consider first the critical world model and then the reference ΛCDM model. It is convenient to present these space-time diagrams with time measured in units of H_0^{-1} and distance in units of c/H_0. The diagrams shown in Figs. 21.3 and 21.4 follow the attractive presentation by Davis and Lineweaver, but the time axis has been truncated at the present cosmological epoch.[36]

The Critical World Model $\Omega_0 = 1, \Omega_v = 0$

It is simplest to begin with the critical world, $\Omega_0 = 1, \Omega_v = 0$, which has a particularly simple space-time diagram. Figure 21.3(a) shows the more 'intuitive' presentation in terms of cosmic time t plotted against proper distance $x = ar$. The world lines of galaxies with redshifts 0.5, 1, 2 and 3 are shown and these all have $x \propto t^{2/3}$ because $a \propto t^{2/3}$ and r, the comoving radial distance coordinate, is a constant for any given galaxy. Also shown are the past light cone, which converges to the origin as $t \to 0$, the Hubble sphere and the particle horizon. There is no event horizon in this model.

In Fig. 21.3(b), the coordinate system is changed to conformal time τ plotted against comoving radial distance coordinate r. Now the trajectories of the galaxies at $z = 0.5$, 1, 2 and 3 are vertical lines since they have the same value of the distance 'label' r for all time. Consequently, the initial singularity at $r = 0$ becomes a singular line corresponding to the abscissa of the diagram. The Hubble sphere, the particle horizon and the past light cone become straight lines in this presentation, as can be appreciated from their definitions. Thus, comparison of (21.6) and (21.13) shows that $r = c\tau$; comparison of (21.8) with (21.13) shows that the particle horizon is proportional to τ and the past light come has slope the speed of light, propagating to the observer at the origin. Note that, in the units of Fig. 21.3(b), the numbers on the axes are exact.

The Reference ΛCDM World Model $\Omega_0 = 0.3, \Omega_v = 0.7$

The same pair of diagrams corresponding to Fig. 21.3 for the reference ΛCDM model with $\Omega_0 = 0.3$ and $\Omega_v = 0.7$ is shown in Fig. 21.4. Now the rate of change of the scale factor with cosmic time in units in which $c = 1$ and $H_0 = 1$ is

$$\dot{a} = \left[\frac{0.3}{a} + 0.7(a^2 - 1)\right]^{1/2}. \tag{21.18}$$

The diagrams shown in Fig. 21.4(a) and (b) have many of the same general features as Fig. 21.3(a) and (b), but there are key differences, the most significant being associated with the dominance of the dark energy term at late epochs.

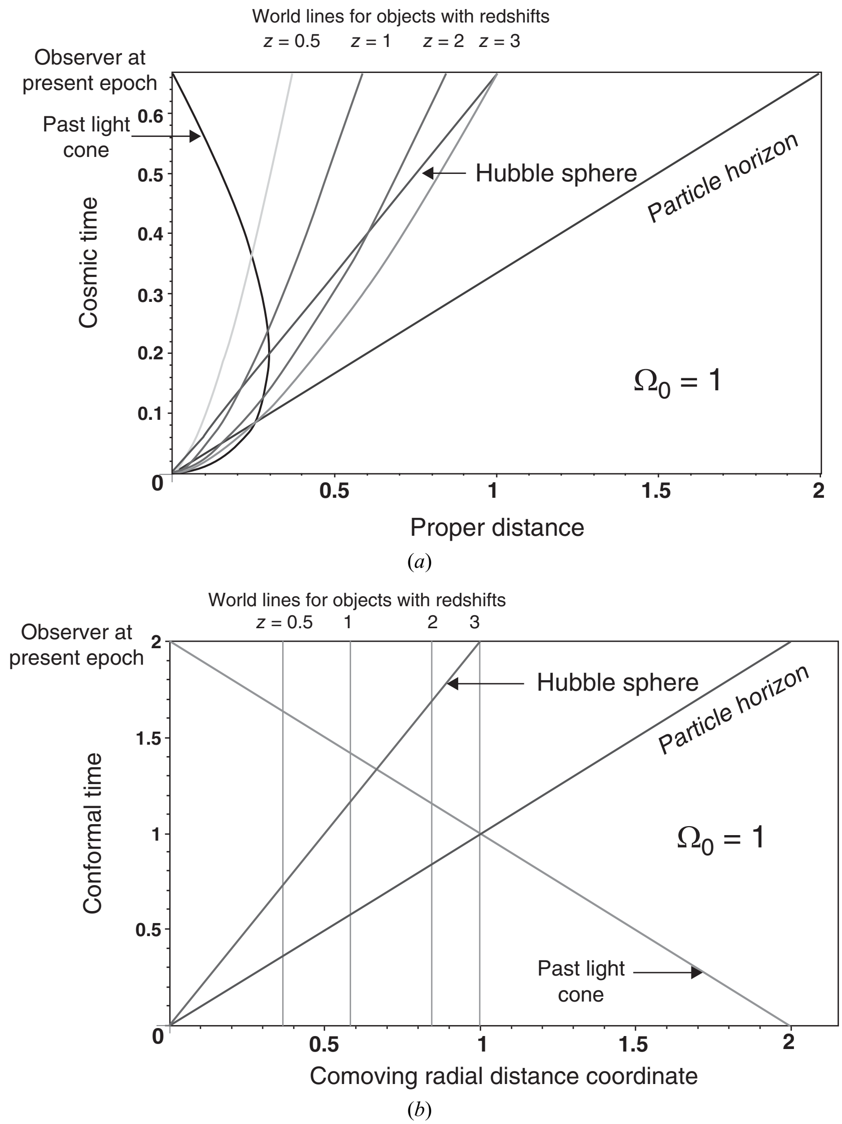


Fig. 21.3 Space-time diagrams for the critical cosmological model, $\Omega_0 = 1, \Omega_v = 0$. The times and distances are measured in units of H_0^{-1} and c/H_0 respectively. (*a*) This diagram is plotted in terms of cosmic time and proper distance. (*b*) The same space-time diagram plotted in terms of conformal time and comoving radial distance coordinate.

- In both diagrams, note that the cosmic time-scale is stretched out relative to the critical model.
- In Fig. 21.4(a), the world lines of galaxies begin to diverge at the present epoch as the repulsive effect of the dark energy dominates over the attractive force of gravity.
- In Fig. 21.4(a), the Hubble sphere begins to converge to a proper distance of 1.12 in units of c/H_0. The reason is that the expansion rate $\dot{a}$ becomes exponential in the future while Hubble's constant tends to a constant value of $\Omega_v^{1/2}$.

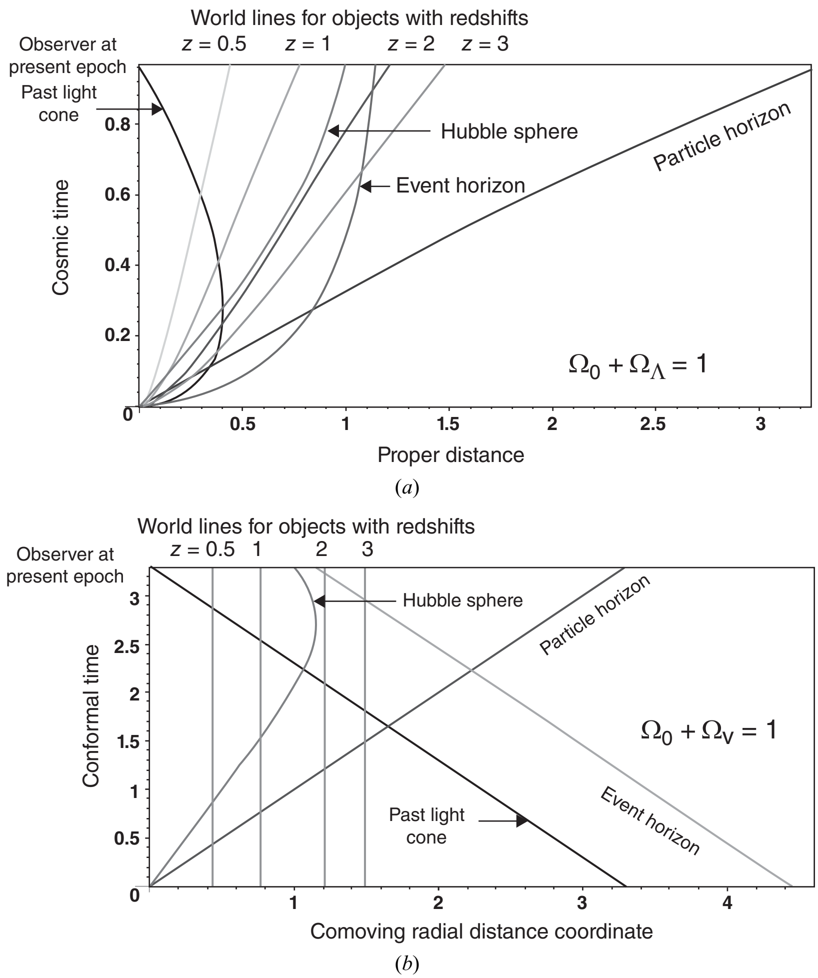


Fig. 21.4 Space-time diagrams for the reference cosmological model, $\Omega_0 = 0.3, \Omega_v = 0.7$. The times and distances are measured in units of H_0^{-1} and c/H_0 respectively.

- Unlike the critical model, there is an event horizon in the reference model. The reason is that, although the geometry is flat, the exponential expansion drives galaxies beyond distances at which there could be causal communication with an observer at epoch t. It can be seen from Fig. 21.4(a) that the event horizon tends towards the same asymptotic value of 1.12 in proper distance units as the Hubble sphere. To demonstrate this, we evaluate the integral

$$r_E = a \int_a^\infty \frac{da}{[0.3a + 0.7(a^4 - a^2)]^{1/2}}. \tag{21.19}$$

For large values of a, terms other than that in a^4 under the square root in the denominator can be neglected and the integral becomes $1/0.7^{1/2} = 1.12$, as found above for the Hubble sphere. In Fig. 21.4(b), the comoving distance coordinates of the Hubble sphere and the event horizon tend to zero as $t \to \infty$ because, for example, (21.8) has to be divided by a to convert it to a comoving distance and $a \to \infty$. This shrinking of the Hubble sphere is the origin of the statement that ultimately we will end up 'alone in the Universe'.

The paper by Davis and Lineweaver repays close study.[37] Their remarkable Appendix B indicates how even some of the most distinguished cosmologists and astrophysicists can lead the unwary newcomer to the subject astray. For our future purposes, the diagram shown in Fig. 21.4(*b*) in extended form will provide illumination on how the inflationary paradigm can help resolve some of the Big Problems.

21.5 The Inflationary Universe: Historical Background

The most important conceptual development for studies of the very early Universe can be dated to about 1980 and the proposal by Alan Guth of the *inflationary model* for the very early Universe.[38] Guth fully acknowledged that there had been earlier suggestions foreshadowing his proposal.[39] Zeldovich had noted in 1968 that there is a physical interpretation of the cosmological constant Λ in terms of the zero-point fluctuations in a vacuum.[40] Linde in 1974 and Bludman and Ruderman in 1977 had shown that the scalar Higgs fields of particle physics have similar properties to those which would result in a positive cosmological constant.[41]

In 1975, Gurevich noted that an early initial vacuum-dominated phase would provide a 'cause of cosmological expansion', this solution having later to be joined onto the standard Friedman–Lemaître solutions. Starobinsky, a member of Zeldovich's group of astrophysicists/cosmologists, found a class of cosmological solutions which indeed did just that, starting with a de Sitter phase and ultimately ending up as Freidman–Lemaître models – he noted that the exponential de Sitter expansion could lead to a solution of the singularity problem by extrapolating the de Sitter solutions back to $t \to -\infty$. He also predicted that gravitational waves would be generated during the de Sitter phase at potentially measureable levels. Commenting on this work, Zeldovich also noted that the exponential expansion would eliminate the horizon problem.

Guth realised that, if there were an early phase of exponential expansion of the Universe, this could solve the horizon problem and drive the Universe towards a flat spatial geometry, thus solving the flatness problem simultaneously. The great merit of Guth's insights was that they made the issues of the physics of the early Universe accessible to the community of cosmologists and spurred an explosion of interest in developing genuine physical theories of the very early Universe by particle theorists.

Suppose the scale factor a increased exponentially with time as $a \propto e^{t/T}$. Such exponentially expanding models were found in some of the earliest solutions of the Friedman equations, in the guise of empty de Sitter models driven by what is now termed the vacuum energy density Ω_v.[42] Consider a tiny region of the early Universe expanding under the influence of the exponential expansion. Particles within the region were initially very close together and in causal communication with each other. Before the inflationary expansion began, the region had physical scale less than the particle horizon, and so there was time for it to attain a uniform, homogeneous state. The region then expanded exponentially so that neighouring reference frames were driven to such large distances that they could no longer communicate by light signals – the causally connected reference frames were swept beyond their particle horizons by the inflationary expansion. At the end of the inflationary epoch, the Universe transformed into the standard radiation-dominated Universe and the inflated region continued to expand as $a \propto t^{1/2}$.

Let us demonstrate to order of magnitude how the argument runs in concrete terms. The time-scale 10^{-34} s is taken to be the characteristic e-folding time for the exponential expansion. Over the interval from 10^{-34} s to 10^{-32} s, the radius of curvature of the Universe increased exponentially by a factor of about $e^{100} \approx 10^{43}$. The horizon scale at the beginning of this period was only $r \approx ct \approx 3 \times 10^{-26}$ m and this was inflated to a dimension of 3×10^{17} m by the end of the inflationary era. This dimension then scaled as $t^{1/2}$, as in the standard radiation-dominated Universe so that the region would have expanded to a size of $\sim 3 \times 10^{42}$ m by the present day – this dimension far exceeds the present particle horizon $r \approx cT_0$ of the Universe, which is about 10^{26} m. Thus, our present Universe would have arisen from a tiny region in the very early Universe which was much smaller than the horizon scale at that time. This guaranteed that our present Universe would be isotropic on the large scale, resolving the horizon problem. At the end of the inflationary era, there was an enormous release of energy associated with the ‘latent heat’ of the phase transition and this reheated the Universe to a very high temperature.

The exponential expansion also had the effect of straightening out the geometry of the early Universe, however complicated it may have been to begin with. Suppose the tiny region of the early Universe had some complex geometry. The radius of curvature of the geometry $R_c(t)$ scales as $R_c(t) \propto a(t)$, and so it is inflated to dimensions vastly greater than the present size of the Universe, driving the geometry of the inflated region towards flat Euclidean geometry, $\Omega_K = 0$, and consequently the Universe must have $\Omega_0 + \Omega_v = 1$. It is important that these two aspects of the case for the inflationary picture can be made independently of a detailed understanding of the physics of inflation. There is also considerable freedom about the exact time when the inflationary expansion could have occurred, provided there are sufficient e-folding times to isotropise our observable Universe and flatten its geometry.

In Guth’s original inflationary scenario, the exponential expansion was associated with the spontaneous symmetry breaking of Grand Unified Theories of elementary particles at very high energies through a first-order phase transition, only about 10^{-34} s after the Big Bang, commonly referred to as the GUT era. The Universe was initially in a symmetric state, referred to as a false vacuum state, at a very high temperature before the inflationary phase took place. As the temperature fell, spontaneous symmetry breaking took place through the process of barrier penetration from the false vacuum state and the Universe attained a lower energy state, the true vacuum. At the end of this period of exponential expansion, the phase transition took place, releasing a huge amount of energy.

The problem with this realisation was that it predicted ‘bubbles’ of true vacuum embedded in the false vacuum, with the result that huge inhomogeneities were predicted. Another concern was that an excessive number of monopoles were created during the GUT phase transition. Kibble showed that, when this phase transition took place, topological defects are expected to be created, including point defects (monopoles), line defects (cosmic strings) and sheet defects (domain walls).[43] Kibble also showed that one monopole is created for each correlation scale at that epoch. Since that scale cannot be greater than the particle horizon at the GUT phase transition, it is expected that huge numbers of monopoles are created. According to the simplest picture of the GUT phase transition, the mass density in these monopoles in the standard Big Bang picture would vastly exceed $\Omega_0 = 1$ at the present epoch.[44]

The model was revised in 1982 by Linde and by Albrecht and Steinhardt who proposed instead that, rather than through the process of barrier penetration, the transition took place through a second-order phase transition which did not result in the formation of 'bubbles' and so excessive inhomogeneities.[45] This picture, often referred to as *new inflation*, also eliminated the monopole problem since the likelihood of even one being present in the observable Universe was very small.

By the spring of 1982 several groups were at work fleshing out the details of the new inflationary scenario. With notable exceptions, such as Stephen Hawking and John Barrow, nearly everyone in this research community came from a background in particle physics. They all met in Cambridge at a workshop sponsored by the Nuffield Foundation to hammer out the developing issues in the physics of the early Universe.[46]

An important focus of the conference was the calculation of density perturbations produced during an inflationary era. It was appreciated that this was a 'calculable problem', the answer being an estimate of the magnitude of the density perturbations, measured by the dimensionless density contrast $\Delta = \delta\rho/\rho$, produced during inflation. In 1981, Mukhanov and Chibisov had argued that a de Sitter phase could generate perturbations by 'stretching' the zero-point fluctuations of quantum fields to significant scales.[47]

At the heart of the debate was the 'gauge problem', reflecting the fact that a 'perturbed space-time' cannot be uniquely decomposed into a background space-time plus perturbations. Slicing the space-time along different surfaces of constant time leads to different magnitudes for the density perturbations. The perturbations 'disappear', for example, by slicing along surfaces of constant density. In practice, almost all studies of structure formation used a particular choice of gauge, generally the synchronous gauge, but this leads to difficulties in interpreting perturbations with length scales greater than the Hubble radius.

Hawking and Guth pursued refinements of Hawking's approach during the Nuffield workshop, the centerpiece of these calculations being the 'time delay' function characterizing the start of the scalar field's slow roll down the effective potential. This 'time delay' function can be related to the two-point correlation function characterizing fluctuations in the field ϕ prior to inflation, and it is also related to the spectrum of density perturbations, since these are assumed to arise as a result of the differences in the time at which inflation ends.

Steinhardt and Turner then enlisted James Bardeen's assistance in developing a third approach; he had recently formulated a fully *gauge invariant formulation* for the study of density perturbations on all scales.[48] Using Bardeen's formalism, the three aimed to give a full account of the behavior of different modes of the field ϕ as these evolved through the inflationary phase and up to recombination. The physical origin of the spectrum was traced to the qualitative change in behaviour as perturbation modes expand past the Hubble radius: they 'freeze out' as they cross the horizon, and leave an imprint that depends on the details of the model under consideration.[49]

The key results were that inflation leads naturally to an almost Harrison–Zeldovich spectrum of density fluctuations and these have Gaussian phases. But reducing the magnitude of these perturbations to satisfy observational constraints required an unnatural choice of coupling constants within the context of Higgs fields at the GUT era. The upshot was the introduction of an 'inflaton field', a scalar field custom-made to produce an inflationary stage.

Following the Nuffield workshop, inflation turned into a 'paradigm without a theory', as expressed by Turner, as cosmologists developed a wide variety of models bearing a loose family resemblance. The models share the basic idea that the early Universe passed through an inflationary phase, but differ on the nature of the 'inflaton' field (or fields) and the form of the effective potential $V(\phi)$. Olive's review of the first decade of inflation ended by bemoaning the ongoing failure of any of these models to renew the strong connection with particle physics achieved in old and new inflation:

> A glaring problem, in my opinion, is our lack of being able to fully integrate inflation into a unification scheme or any scheme having to do with our fundamental understanding of particle physics and gravity. … An inflaton as an inflaton and nothing else can only be viewed as a toy, not a theory.[50]

As a result, there is not a genuine physical theory of the inflationary Universe, but its basic concepts resolve some of the problems listed in Section 21.3. What it also does, and which gives it considerable appeal, is to suggest an origin for the spectrum of initial density perturbations as quantum fluctuations on the scale of the particle horizon.

21.6 The Origin of the Spectrum of Primordial Perturbations

As Andrew Liddle and David Lyth have written,[51]

> Although introduced to resolve problems associated with the initial conditions needed for the Big Bang cosmology, inflation's lasting prominence is owed to a property discovered soon after its introduction: it provides a possible explanation for the initial inhomogeneities in the Universe that are believed to have led to all the structures we see, from the earliest objects formed to the clustering of galaxies to the observed irregularities in the microwave background.

The theory also makes predictions about the spectrum of primordial gravitational waves which are accessible to experimental validation.[52] The enormous impact of particle theorists taking these cosmological problems really seriously has enlarged, yet again, the domain of astrophysical cosmology. For the 'cosmologist in the street', the theory of inflation does not make for particularly easy reading, because the reader should be comfortable with many aspects of theoretical physics which lie outside the standard tools of the observational cosmologist – ladder operators, quantum field theory, zero point fluctuations in quantum fields, all of these applied within the context of curved space-time. Developing the theory of the quantum origin of density perturbations in detail cannot be carried out with modest effort, but we can give some flavour of the new types of physics involved.

Let us list some of the clues about the formulation of a successful theory.[53]

The Equation of State

We know from analyses of the physical significance of the cosmological constant Λ that exponential growth of the scale factor is found if the dark energy has a negative pressure equation of state $p = -\varrho c^2$. More generally, exponential growth of the scale factor is found provided the strong energy condition is violated, that is, if $p < -\frac{1}{3}\varrho c^2$. To be effective in the very early Universe, the mass density of the scalar field has to be vastly greater than the value of Ω_v we measure today.

The Duration of the Inflationary Phase

In the example of the inflationary expansion given above, we arbitrarily assumed that 100 e-folding times would take place during the inflationary expansion. A more careful calculation shows that there must have been at least 60 e-folding times and these took place in the very early Universe. It is customary to assume that inflation began not long after the Planck era, but there is quite a bit of room for manoeuvre.

The Shrinking Hubble Sphere

There is a natural way of understanding how fluctuations can be generated from processes in the very early Universe. It is helpful to revisit the conformal diagrams for world models discussed in Section 21.4, in particular, Fig. 21.4(*b*) in which the comoving radial distance coordinate and conformal time were worked out for the reference model with $\Omega_0 = 0.3$ and $\Omega_v = 0.7$. The effect of using conformal time is to stretch out time in the past and shrink it into the future. Because of the use of linear scales in the ordinate, the radiation-dominated phase of the standard Big Bang is scarcely visible.

In Fig. 21.5(*a*), there are two additions to Fig. 21.4(*b*). The redshift of 1000 is shown corresponding to the last scattering surface of the CMB. The intersection with our past light cone is shown and then a past light cone from the last scattering surface to the singularity at conformal time $\tau = 0$, which is now a singular line – the abscissa of Fig. 21.5(*a*) – is shown as a shaded triangle. This is another way of demonstrating the *horizon problem* – the region of causal contact is very small compared with moving an angle of 180° over the sky which would correspond to twice the distance between the origin and the comoving radial distance coordinate at 3.09.

Let us now add the inflationary era to Fig. 21.5(*a*). It is useful to regard the end of the inflation era as the zero of time for the standard Big Bang and then to extend the diagram back to negative conformal times. In other words, we shift the zero of conformal time very

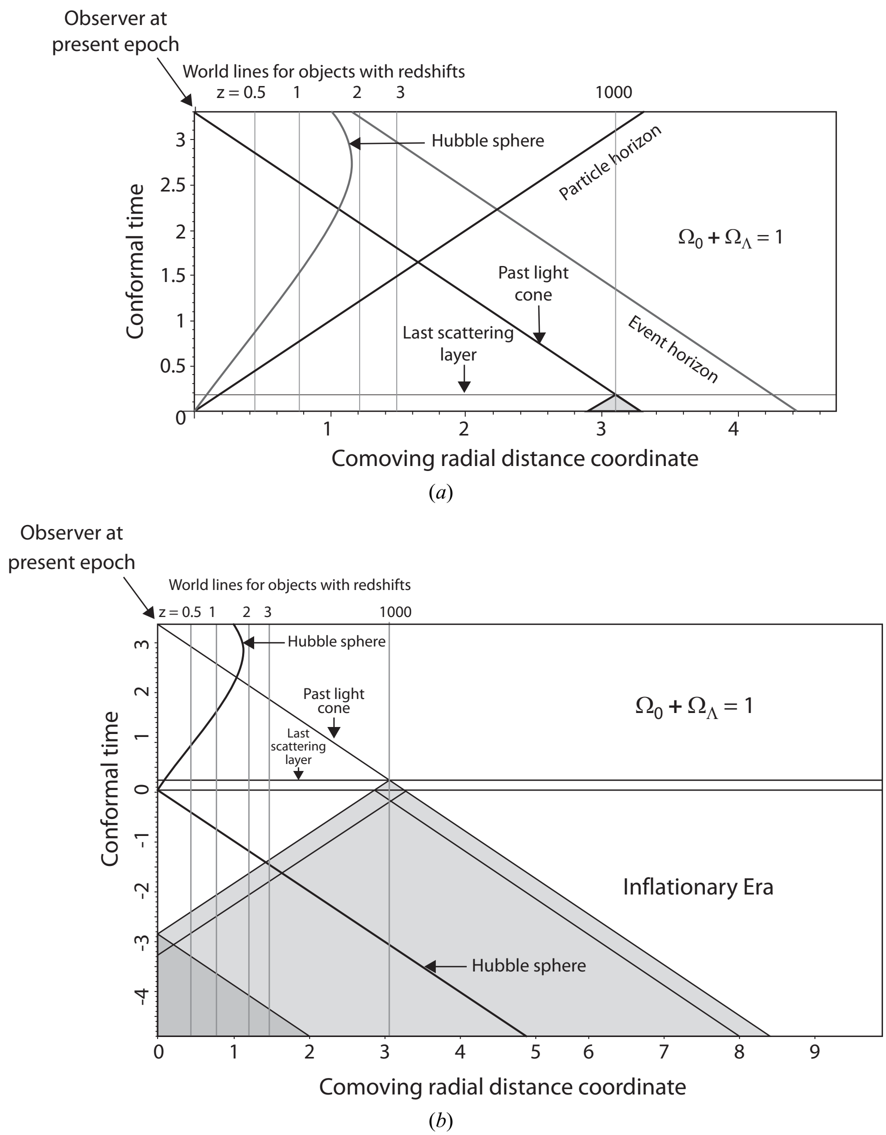


Fig. 21.5 (*a*) A repeat of conformal diagram Fig. 21.4(*b*) in which conformal time is plotted against comoving radial distance coordinate. The last scattering surface at the epoch of recombination is shown as well as the past light cone from the point at which it intersects the last scattering surface. (*b*) An extended conformal diagram now showing the inflationary era. The time coordinate is set to zero at the end of the inflationary era and evolution of the Hubble sphere and the past light cone at recombination extrapolated back to the inflationary era.

slightly to, say, 10^{-32} s and then we can extend the light cones back through the entire inflationary era (Fig. 21.5(*b*)). This construction provides another way of understanding how the inflationary picture resolves the causality problem. The light cones have unit slope in the conformal diagram and so we draw light cones from the ends of the element of comoving radial distance at $\tau = 0$ from the last scattering surface. Projecting far enough back in time, the light cones from opposite directions on the sky overlap, represented by the dark grey shaded area in Fig. 21.5(*b*). This is the region of causal contact in the very early Universe.

There is, however, an even better way of understanding what is going on. We distinguished between the Hubble sphere and the particle horizon in Section 21.4 – now this distinction becomes important. The particle horizon is defined as the maximum distance over which causal contact could have been made from the time of the singularity to a given epoch. In other words, it is not just what happened at a particular epoch which is important, but the history along the past light cone.

In contrast, writing the exponential inflationary expansion of the scale factor as $a = a_0 \exp[H(t - t_i)]$, where a_0 is the scale factor when the inflationary expansion began at τ_i, the Hubble sphere has radius $r_{HS} = c/H$ and the comoving Hubble sphere is $r_{HS}(\text{com}) = c/(Ha)$. Since H is a constant throughout most of the inflationary era, it follows that the comoving Hubble sphere *decreases* as the inflationary expansion proceeds.

We now need to join this evolution of the comoving Hubble sphere onto its behaviour after the end of inflation, that is, join it onto Fig. 21.5(*a*). The expression for conformal time during the inflationary era is

$$\tau = \int \frac{\mathrm{d}a}{a\dot{a}}, \tag{21.20}$$

and so, integrating and using the expression for $r_{HS}(\text{com})$, we find

$$\tau = \text{constant} - \frac{r_{HS}(\text{com})}{c}. \tag{21.21}$$

This solution for $r_{HS}(\text{com})$ is joined on to the standard result at the end of the inflationary epoch, as illustrated in Fig. 11.5(*b*). The complete evolution of the Hubble sphere is indicated by the heavy line labelled ‘Hubble sphere’ in that diagram.

Figure 21.5(*b*) illustrates very beautifully how the inflationary paradigm solves the horizon problem. It will be noticed that the point at which the Hubble sphere crosses the comoving radial distance coordinate of the last scattering surface corresponds exactly to the time when the past light cones from opposite directions on the sky touch at conformal time −3. This is not a coincidence – they are different ways of stating that opposite regions of the CMB were in causal contact at conformal time $t = -3$.

But we learn a lot more. Because any object preserves its comoving radial distance coordinate for all time, as represented by the vertical lines in Fig. 21.5, it can be seen that, in the early Universe, objects lie within the Hubble sphere, but during the inflationary expansion, they pass through it and remain outside it for the rest of the inflationary expansion. Only when the Universe transforms back into the standard Friedman model does the Hubble sphere begin to expand again and objects can then ‘re-enter the horizon’. Consider, for example, the region of the Universe out to redshift $z = 0.5$, which corresponds to one of the comoving coordinate lines in Fig. 21.5(*b*). It remained within the Hubble sphere during the inflationary era until conformal time $\tau = -0.4$, after which it was outside the horizon. It then re-entered the Hubble sphere at conformal time $\tau = 0.8$. This behaviour occurs for all scales and masses of interest in understanding the origin of structure in the present Universe.

Since causal connection is no longer possible on scales greater than the Hubble sphere, it follows that objects 'freeze out' when they pass through the Hubble sphere during the inflationary era, but they come back in again and regain causal contact when they re-cross it in the standard matter–radiation dominated era. This is one of the key ideas behind the idea that the perturbations from which galaxies formed were created in the early Universe, froze out on crossing the Hubble sphere and then grew again on re-entering it at conformal times $\tau > 0$.

Notice that, at the present epoch, we are entering a phase of evolution of the Universe when the comoving Hubble sphere about us has begun to shrink again. This can be seen in the upper part of Fig. 21.5(*b*) and is entirely due to the fact that the dark energy is now dominating the expansion and its dynamics are precisely those of another exponential expansion. In fact, the Hubble sphere tends asymptotically to the line labelled 'event horizon' in Fig. 21.5(*a*).

21.6.1 Scalar Fields

As Baumann noted,[54] there are three equivalent conditions necessary to produce an inflationary expansion in the early Universe:

- the decreasing of the Hubble sphere during the early expansion of the Universe;
- an accelerated expansion;
- violation of the strong energy condition, $p < -\varrho c^2/3$.

How can this be achieved physically? To quote Baumann's words, written before the discovery of the Higgs boson in 2012:

> Answer: scalar field with special dynamics! Although no fundamental scalar field has yet been detected in experiments, there are fortunately plenty of such fields in theories beyond the standard model of particle physics. In fact, in string theory for example there are numerous scalar fields (moduli), but it proves very challenging to find just one with the right characteristics to serve as an inflaton candidate.

The results of calculations of the properties of the scalar field $\phi(t)$, which is assumed to be homogeneous at a given epoch, are as follows. There is a kinetic energy $\dot{\phi}^2/2$ and a potential energy, or self-interaction energy, $V(\phi)$ associated with the field. Putting these through the machinery of field theory results in expressions for the density and pressure of the scalar field:

$$\varrho_\phi = \frac{1}{2}\dot{\phi}^2 + V(\phi); \qquad p_\phi = \frac{1}{2}\dot{\phi}^2 - V(\phi). \tag{21.22}$$

The scalar field can result in a negative pressure equation of state, provided the potential energy of the field is very much greater than its kinetic energy. In the limit in which the kinetic energy is neglected, we obtain the equation of state $p = -\varrho c^2$, where the c^2, which is set equal to one by professional field theorists, has been restored.

To find the time evolution of the scalar field, we combine (21.22) with the Einstein field equations with the results:

$$H^2 = \frac{1}{3}\left(\frac{1}{2}\dot{\phi}^2 + V(\phi)\right); \qquad \ddot{\phi} + 3H\dot{\phi} + V(\phi)_{,\phi} = 0, \tag{21.23}$$

where $V(\phi)_{,\phi}$ means the derivative of $V(\phi)$ with respect to ϕ. Thus, to obtain the inflationary expansion over many e-folding times, the kinetic energy term must be very small compared with the potential energy and the potential energy term must be very slowly varying with time. This is formalised by requiring the two *slow-roll parameters* $\epsilon(\phi)$ and $\eta(\phi)$ to be very small during the inflationary expansion. These parameters set constraints upon the dependence of the potential energy function upon the field ϕ and are formally written:

$$\epsilon(\phi) \equiv \frac{1}{2}\left(\frac{V_{,\phi}}{V}\right)^2; \quad \eta(\phi) \equiv \frac{V_{,\phi\phi}}{V} \quad \text{with} \quad \epsilon(\phi), |\eta(\phi)| \ll 1, \tag{21.24}$$

where $V(\phi)_{,\phi\phi}$ means the second derivative of $V(\phi)$ with respect to ϕ. Under these conditions, we obtain what we need for inflation, namely,

$$H^2 = \frac{1}{3}V(\phi) = \text{constant} \quad \text{and} \quad a(t) \propto e^{Ht}. \tag{21.25}$$

At this stage, it may appear that we have not really made much progress since we have adjusted the theory of the scalar field to produce what we know we need. The bonus comes when we consider fluctuations in the scalar field and their role in the formation of the spectrum of primordial perturbations.

21.6.2 The Quantised Harmonic Oscillator

The key result can be illustrated by the elementary quantum mechanics of a harmonic oscillator. The solutions of Schrödinger's equation for a harmonic potential have quantised energy levels and wave functions:

$$E = \left(n + \tfrac{1}{2}\right)\hbar\omega; \qquad \psi_n = H_n(\xi)\exp\left(-\tfrac{1}{2}\xi^2\right), \tag{21.26}$$

where $H_n(\xi)$ is the Hermite polynomial of order n and $\xi = \sqrt{\beta}x$. For the simple harmonic oscillator, $\beta^2 = am/\hbar^2$, where a is the constant in the expression for the harmonic potential $V = \frac{1}{2}ax^2$ and m is the reduced mass of the oscillator.

We are interested in fluctuations about the zero-point energy, that is, the stationary state with $n = 0$. The zero-point energy and Hermite polynomial of order $n = 0$ are

$$E = \tfrac{1}{2}\hbar\omega \quad \text{and} \quad H_0(\xi) = \text{constant}. \tag{21.27}$$

The first expression is the well-known result that the oscillator has to have finite kinetic energy in the ground state. It is straightforward to work out the variance of the position coordinate x of the oscillator:[55]

$$\langle x^2 \rangle = \frac{\hbar}{2\omega m}. \tag{21.28}$$

These are the fluctuations which must necessarily accompany the zero-point energy of the vacuum fields. This elementary calculation sweeps an enormous number of technical issues under the carpet. Baumann's clear presentation of the proper calculation can be recommended. It is reassuring that his final answer agrees exactly with the above results for the one-dimensional harmonic oscillator.

21.6.3 The Spectrum of Fluctuations in the Scalar Field

We need only one more equation – the expression for the evolution of the vacuum fluctuations in the inflationary expansion. The inflaton field is decomposed into a uniform homogeneous background and a perturbed component $\delta\phi$ which is the analogue of the deviation x of the zero point oscillations of the harmonic oscillator. Baumann outlines the derivation of this equation, warning of the numerous technical complexities which have to be dealt with. In Bertschinger's review of the physics of inflation, he deals with these issues and finds the following equation:

$$\delta\ddot{\phi}_k + 3\left(\frac{\dot{a}}{a}\right)\delta\dot{\phi}_k + (k_c^2 c_s^2 - 2\kappa)\,\delta\phi_k = 0, \tag{21.29}$$

where κ is the curvature of space at the present epoch.[56] This equation has a familiar form which can be understood by comparing it with the standard equation for the evolution of density perturbations $\Delta = \delta\rho/\rho$ in the Friedman models:

$$\frac{d^2\Delta}{dt^2} + 2\left(\frac{\dot{a}}{a}\right)\frac{d\Delta}{dt} = \Delta(4\pi G\varrho_0 - k^2 c_s^2), \tag{21.30}$$

where k is the proper wavenumber and c_s is the speed of sound.[57]

Since we are interested in flat space solutions, we set $\kappa = 0$. Furthermore, for matter with equation of state $p = -\varrho c^2$, the speed of sound is the speed of light, which according to Baumann's conventions is set equal to unity, and so we obtain an equation of the form (21.29). A big advantage of Baumann's proper derivation of (21.29) is that it can be applied on superhorizon scales as well as for those within the horizon, thanks to the use of Bardeen's gauge-invariant formulation of the perturbation equation.

We recognise that (21.29) is the equation of motion for a damped harmonic oscillator. If the 'damping term' $3H\delta\dot{\phi}_k$ is set equal to zero, we find harmonic oscillations. On the other hand, for scales much greater than the radius of the Hubble sphere, $\lambda \gg c/H$, an order of magnitude calculation shows that the damping term dominates and the velocity $\delta\dot{\phi}_k$ tends exponentially to zero, corresponding to the 'freezing' of the fluctuations on superhorizon scales.

Both x and $\delta\phi_k$ have zero point fluctuations in the ground state. In the case of the harmonic oscillator, we found $\langle x^2 \rangle \propto \omega^{-1}$. In exactly the same way, we expect the fluctuations in $\delta\phi_k$ to be inversely proportional to the 'angular frequency' in (21.29), that is,

$$\langle(\delta\phi_k)^2\rangle \propto \frac{1}{k/a} \propto \lambda, \tag{21.31}$$

where λ is the proper wavelength. Integrating over wavenumber, we find the important result

$$\langle(\delta\phi)^2\rangle \propto H^2. \tag{21.32}$$

At the end of the inflationary expansion, the scalar field is assumed to decay into the types of particles which dominate our Universe at the present epoch, releasing a vast amount of energy which reheats the contents of the Universe to a very high temperature. The final step in the calculation is to relate the fluctuations $\delta\phi$ to the density perturbations in the highly relativistic plasma in the post-inflation era. In the simplest picture, we can think of this transition as occurring abruptly between the era when $p = -\varrho c^2$ and the scale factor increases exponentially with time, as in the de Sitter metric, to that in which the standard relativistic equation of state $p = \frac{1}{3}\varrho c^2$ applies with associated variation of the inertial mass density with cosmic time $\varrho \propto H^2 \propto t^{-2}$ (see (20.118)). Guth and Pi used the time-delay formalism which enables the density perturbation to be related to the inflation parameters.[58] The end result is

$$\frac{\delta\varrho}{\varrho} \propto \frac{H_*^2}{\dot{\phi}_*}. \tag{21.33}$$

where H_* and ϕ_* are their values when the proper radius of the perturbation is equal to the Hubble radius.

This order of magnitude calculation illustrates how quantum fluctuations in the scalar field ϕ can result in density fluctuations in the matter, all of which have more or less the same amplitude when they passed through the horizon in the very early Universe. They then remained frozen in until they re-entered the horizon very much later in the radiation-dominated era, as illustrated in Fig. 21.5(*b*).

This schematic calculation is only intended to illustrate why the inflationary paradigm is taken so seriously by theorists. It results remarkably naturally in the Harrison–Zeldovich spectrum of primordial perturbations.

In the full theory, the values of the small parameters ϵ and η defined by (21.24) cannot be neglected and they have important consequences for the spectrum of the perturbations and the existence of primordial gravitational waves. Specifically, the spectral index of the perturbations on entering the horizon is predicted to be

$$n_S - 1 = 2\eta - 6\epsilon. \tag{21.34}$$

Furthermore, tensor perturbations, corresponding to gravitational waves, are also expected to be excited during the inflationary era. Quantum fluctuations generate quadrupole perturbations and these result in a similar almost scale-invariant power spectrum of perturbations. Their spectral index is predicted to be

$$n_T - 1 = -2\epsilon, \tag{21.35}$$

where scale invariance corresponds to $n_T = 1$. The tensor-to-scalar ratio is defined as

$$r = \frac{\Delta_T^2}{\Delta_S^2} = 16\epsilon, \tag{21.36}$$

where Δ_T^2 and Δ_S^2 are the power spectra of tensor and scalar perturbations respectively.

These results illustrate why the deviations of the spectral index of the observed perturbations from the value $n_S = 1$ are so important. The fact that the best fit value $n_S = 0.961^{+0.018}_{-0.019}$ is slightly, but significantly, less than one suggests that there may well be a background of primordial gravitational waves. These are really very great observational challenges, but they provide a remarkably direct link to processes which may have occurred during the inflationary epoch. To many cosmologists, the detection of primordial gravitational waves with the predicted spectrum would be the 'smoking gun' which sets the seal on the inflationary model of the early Universe.

Whilst the above calculation is a considerable triumph for the inflationary scenario, we should remember that there is as yet no physical realisation of the scalar field. Although the scale-invariant spectrum is a remarkable prediction, the amplitude of the perturbation spectrum is model dependent. There are literally hundreds of possible inflationary models depending upon the particular choice of the inflationary potential. We should also not neglect the possibility that there are other sources of perturbations which could have resulted from various types of topological defect, such as cosmic strings, domain walls, textures and so on.[59] Granted all these caveats, the startling success of the inflationary model in accounting for the observed spectrum of fluctuations in the CMB has made it the model of choice for studies of the early Universe for many cosmologists.

21.7 Baryogenesis

A key contribution of particle physics to studies of the early Universe concerns the baryon-asymmetry problem, a subject referred to as *baryogenesis*. In a prescient paper of 1967, Sakharov enunciated the three conditions necessary to account for the baryon-antibaryon asymmetry of the Universe.[60] *Sakharov's rules* for the creation of non-zero baryon number from an initially baryon symmetric state are:

- *Baryon number* must be violated;
- C (charge conjugation) and CP (charge conjugation combined with parity) must be violated;
- The asymmetry must be created under *non-equilibrium conditions*.

The reasons for these rules can be readily appreciated from simple arguments.[61] Concerning the first rule, it is evident that, if the baryon asymmetry developed from a symmetric high temperature state, baryon number must have been violated at some stage – otherwise, the baryon asymmetry would have to be built into the model from the very beginning. The second rule is necessary in order to ensure that a net baryon number is created, even in the presence of interactions which violate baryon conservation. The third rule is necessary because baryons and antibaryons have the same mass and so, thermodynamically, they would have the same abundances in thermodynamic equilibrium, despite the violation of baryon number and C and CP invariance.

There is evidence that all three rules can be satisfied in the early Universe from a combination of theoretical ideas and experimental evidence from particle physics. Thus, baryon number violation is a generic feature of Grand Unified Theories which unify the strong and electroweak interactions – the same process is responsible for the predicted instability of the proton. C and CP violation have been observed in the decay of the neutral K^0 and $\overline{K^0}$ mesons. The K^0 meson should decay symmetrically into equal numbers of particles and antiparticles but, in fact, there is a slight preference for matter over antimatter, at the level of 10^{-3}, very much greater than the degree of asymmetry necessary for baryogenesis, $\sim 10^{-8}$. The need for departure from thermal equilibrium follows from the same type of reasoning which led to the primordial synthesis of the light elements (Section 20.8). As in that case, so long as the time-scales of the interactions which maintained the various constituents in thermal equilibrium were less than the expansion time-scale, the number densities of particles and antiparticles of the same mass would be the same. In thermodynamic equilibrium, the number densities of different species did not depend upon the cross-sections for the interactions which maintain the equilibrium. It is only after decoupling, when non-equilibrium abundances were established, that the number densities depended upon the specific values of the cross-sections for the production of different species.

In a typical baryogenesis scenario, the asymmetry is associated with some very massive boson and its antiparticle, $X, \overline{X}$, which are involved in the unification of the strong and electroweak forces and which can decay into final states which have different baryon numbers. Kolb and Turner provide a clear description of the principles by which the observed baryon asymmetry can be generated at about the epoch of grand unification or soon afterwards, when the very massive bosons can no longer be maintained in equilibrium.[62] Although the principles of the calculations are well defined, the details are not understood, partly because the energies at which they are likely to be important are not attainable in laboratory experiments, and partly because predicted effects, such as the decay of the proton, have not been observed. Thus, although there is no definitive evidence that this line of reasoning is secure, well-understood physical processes of the type necessary for the creation of the baryon-antibaryon asymmetry exist. The importance of these studies goes well beyond their immediate significance for astrophysical cosmology. As Kolb and Turner remark,

> … in the absence of direct evidence for proton decay, baryogenesis may provide the strongest, albeit indirect, evidence for some kind of unification of the quarks and the leptons.

21.8 The Planck Era

Enormous progress has been made in understanding the types of physical process necessary to resolve the basic problems of cosmology, but it is not clear how independent evidence for them can be found. The methodological problem with these ideas is that they are based upon extrapolations to energies vastly exceeding those which can be tested in terrestrial laboratories. Cosmology and particle physics come together in the early Universe and they boot-strap their way to a self-consistent solution. This may be the best that we can hope for but it would be preferable to have independent constraints upon the theories.

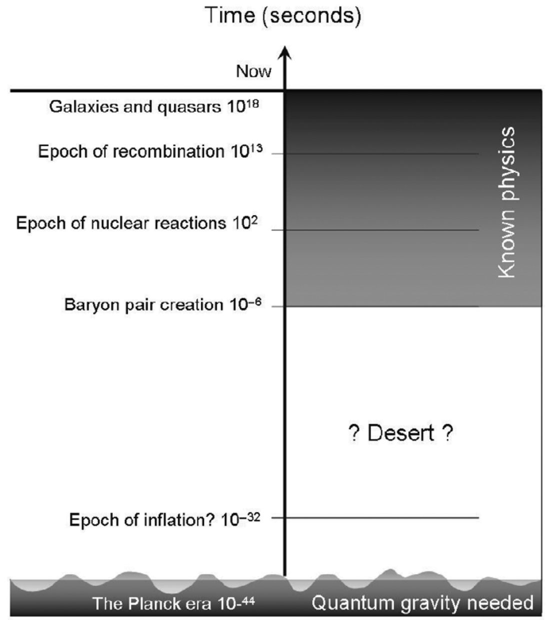


Fig. 21.6 A schematic diagram illustrating the evolution of the Universe from the Planck era to the present time. The shaded area on the right of the diagram indicates the regions of known physics.

A representation of the evolution of the Universe from the Planck era to the present day is shown in Fig. 21.6. The *Planck era* is that time in the very remote past when the energy densities were so great that a quantum theory of gravity is needed. On dimensional grounds, this era must have occurred when the Universe was only about $t_{\mathrm{Pl}} \sim (hG/c^5)^{1/2} \sim 10^{-43}$ s old. Despite enormous efforts on the part of theorists, there is no quantum theory of gravity and so we can only speculate about the physics of these extraordinary eras.

Being drawn on a logarithmic scale, Fig. 21.6 encompasses the evolution of the whole of the Universe, from the Planck area at $t \sim 10^{-43}$ s to the present age of the Universe which is about 4×10^{17} s or 13.6×10^9 years old. Halfway up the diagram, from the time when the Universe was only about a millisecond old, to the present epoch, we can be confident that the Big Bang scenario is the most convincing framework for astrophysical cosmology.

At times earlier than about 1 millisecond, we quickly run out of known physics. This has not discouraged theorists from making bold extrapolations across the huge gap from 10^{-3} s to 10^{-43} s using current understanding of particle physics and concepts from string theories, for example. Some impression of the types of thinking involved in these studies can be found in the ideas expounded in the volume *The Future of Theoretical Physics*, celebrating the 60th birthday of Stephen Hawking.[63] Maybe many of these ideas will turn out to be correct, but there must be some concern that some fundamentally new physics will emerge at higher and higher energies before we reach the GUT era at $t \sim 10^{-36}$ s and the Planck era at $t \sim 10^{-43}$ s. This is why the particle physics experiments to be carried with the facilities such as the Large Hadron Collider at CERN are of such importance for astrophysics and cosmology. The discovery of the Higgs boson was a real triumph, providing essential support for our understanding of standard model of particle physics. In addition, there is the possibility of discovering new types of particles, such as the lightest supersymmetric particle or new massive ultra-weakly interacting particles, as the accessible range of particle energies increases from about 100 GeV to 1 TeV. These experiments should provide clues to the nature of physics beyond the standard model of particle physics and will undoubtedly feed back into understanding of the physics of the early Universe.

It is certain that at some stage a quantum theory of gravity is needed which may help resolve the problems of singularities in the early Universe. The singularity theorems of Penrose and Hawking show that, according to classical theories of gravity under very general conditions, there is inevitably a physical singularity at the origin of the Big Bang, that is, as $t \rightarrow 0$ and the energy density of the Universe tends to infinity. However, it is not clear that the actual Universe satisfies the various energy conditions required by the singularity theorems, particularly if the negative pressure equation of state $p = -\varrho c^2$ holds true in the very early Universe. All these considerations show that new physics is needed if we are to develop a convincing physical picture of the very early Universe.

Notes

1 This chapter is a new version of Chapter 20 of my book *Galaxy Formation* (2008). A revised and expanded version of that chapter was published with Chris Smeenk as Chapter 11 in *The Oxford Handbook of the History of Modern Cosmology*. I am most grateful to Chris for his collaboration in writing that essay. The present chapter is a further revision of these chapters.

2 Zwicky, F. (1933). Rotverschiebung von Extragalaktischen Nebeln (The Redshift of Extragalactic Nebulae), *Helvetica Physica Acta*, **6**, 110–118. (1937) On the Masses of Nebulae and of Clusters of Nebulae, *Astrophysical Journal*, **86**, 217–246.

3 Hohl, F. (1971). Numerical Experiments with a Disk of Stars *Astrophysical Journal*, **168**, 343–359.

4 Miller, R.H., Prendergast, K.H. and Quirk, W.J. (1970). Numerical Experiments on Spiral Structure *Astrophysical Journal*, **161**, 903–916.

5 Ostriker, J.P. and Peebles, P.J.E. (1973). A Numerical study of the Stability of Flattened Galaxies: Or, Can Cold Galaxies Survive?, *Astrophysical Journal*, **186**, 467–480. doi: 10.1086/152513.

6 Fabricant, D.G., Lecar, M. and Gorenstein, P. (1980). X-ray Measurements of the Mass of M87, *Astrophysical Journal*, **241**, 552–560. Böhringer, H. (1994). Clusters of Galaxies, in *Frontiers of Space and Ground-based Astronomy*, eds. W. Wamsteker, M.S. Longair and Y. Kondo, pp. 359–368. Dordrecht: Kluwer Academic Publishers.

7 Hewitt, A. *et al.* (1987). A VLA gravitational lens survey, *Observational Cosmology: IAU Symposium No. 124*, eds. A. Hewitt, G. Burbidge and L.Z. Fang, pp. 747–750. Dordrecht: D. Reidel Publishing Company.

8 Kassiola, A. *et al.* (1991). Bounds on Intergalactic Compact Objects from Observations of Compact Radio Sources, *Astrophysical Journal*, **381**, 6–13.

9 Wilkinson, P. *et al.* (2001). Limits on the Cosmological Abundance of Supermassive Compact Objects from a Search for Multiple Imaging in Compact Radio Sources, *Physical Review Letters*, **86**, 584–587.

10 Alcock, C. *et al.* (2000). The MACHO Project: Microlensing Results from 5.7 Years of Large Magellanic Cloud Observations, *Astrophysical Journal*, **542**, 281–307.
11 Afonso, C. *et al.* (2003). Limits on Galactic Dark Matter with 5 years of EROS SMC Data, *Astronomy and Astrophysics*, **400**, 951–956.
12 Efstathiou, G. (1990). Cosmological Perturbations, in *Physics of the Early Universe*, eds. J.A. Peacock, A.F. Heavens and A.T. Davies, pp. 361–463. Edinburgh: SUSSP Publications. Kolb, E.W. and Turner, M.S. (1990). *The Early Universe*. Redwood City, California: Addison-Wesley Publishing Co.
13 Kolb, E.W. and Turner, M.S. (1990). *op. cit.* An outline of the physics involved is given in Section 13.3 of my book *Galaxy Formation* (2008). *op. cit.*
14 Aaji, R. *et al.* (2017). Observation of the Annihilation Decay Mode $B_0 \to K^+K^-$, *Physical Review Letters*, **118**(8). doi:10.1103/PhysRevLett.118.081801.
15 Weinheimer, C. 2001. Neutrino Mass from Tritium β-Decay, *Dark Matter in Astro- and Particle Physics, Proceedings of the International Conference DARK 2000*, ed. H.V. Klapdor-Kleingrothaus, pp. 513–519. Berlin: Springer-Verlag.
16 Eguchi, K. *et al.* (2003). First Results from Kamland: Evidence for Reactor Anti-neutrino Disappearance, *Physical Review Letters*, **90**, 021802(1–6). Aliu, E. *et al.* (2005). Evidence for Muon Neutrino Oscillation in an Accelerator-Based Experiment, *Physical Review Letters*, **94**, 081802.
17 Tremaine, S. and Gunn, J.E. (1979). Dynamical Role of Light Neutral Leptons in Cosmology, *Physical Review Letters*, **42**, 407–410.
18 Agnese, R. *et al.* (2017). Results from the Super Cryogenic Dark Matter Search Experiment at Soudan, *Physical Review Letters*, **120**, 061802. doi:10.1103/PhysRevLett.120.0618-2.
19 See Section 20.9
20 The reasons for this are illustrated in Section 11.4.2 of my book *Galaxy Formation*, *op. cit.* See also Fig. 14.10 of that text.
21 Dicke, R.H. (1961). Dirac's Cosmology and Mach's Principle, *Nature*, **192**, 440–441.
22 Dicke, R.H. (1961). *op. cit*. Dicke, R.H. and Peebles, P.J.E. (1979). The Big Bang Cosmology – Enigmas and Nostrums, *General Relativity: An Einstein Centenary Survey*, eds. S.W. Hawking and W. Israel, pp. 504–517. Cambridge: Cambridge University Press.
23 Zeldovich, Ya.B. (1965). Survey of Modern Cosmology, *Advances in Astronomy and Astrophysics*, **3**, 241–379.
24 Bardeen and Peebles provide clear expositions of this puzzling feature of the then standard picture of structure formation. Bardeen, J. (1980). Gauge-Invariant Cosmological Perturbations, *Physical Review D*, **22**, 1882–1905. Peebles (1980). *The Large-Scale Structure of the Universe*. Princeton: Princeton University Press.
25 I have given a history of how quantum mechanics came about in my book *Quantum Concepts in Physics* (2013). Cambridge: Cambridge University Press.
26 Longair, M.S. (1997). The Friedman Robertson–Walker Models: On Bias, Errors and Acts of Faith, in *Critical Dialogues in Cosmology*, ed. N. Turok, pp. 285–308. Singapore: World Scientific.
27 Donald Rumsfeld was President George W. Bush's United States Secretary of Defense and played a central role in the planning of the United States' response to the September 11 attacks, which included two wars, one in Afghanistan and one in Iraq.
28 McCrea, W.H. (1970). A Philosophy for Big Bang Cosmology, *Nature*, **228**, 21–24.
29 Carter, B. (1974). Large Number Coincidences and the Anthropic Principle in Cosmology, *Confrontation of Cosmological Theories with Observational Data; Proceedings of the Symposium, Krakow, Poland, September 10–12, 1973*, ed. M.S. Longair, pp. 291–298. Dordrecht: D. Reidel Publishing Company.
30 Barrow, J. and Tipler, F. (1986). *The Anthropic Cosmological Principle*. Oxford: Oxford University Press. Gribben, J. and Rees, M.J. (1989). *Dark Matter, Mankind and Anthropic Cosmology*. New York: Bantam Books.

31 Wheeler, J.A. (1977). Genesis and Observership, *Foundational Problems in the Special Science*, eds. R.E. Butts and J. Hintikka, pp. 3–33. Dordrecht: D. Reidel Publishing Company.
32 Weinberg, S. (1989). The Cosmological Constant Problem, *Reviews of Modern Physics*, **61**, 1–23. (1997). Theories of the Cosmological Constant, in *Critical Dialogues in Cosmology*, ed. N. Turok, pp. 195–203, Singapore: World Scientific.
33 Linde, A. (1983). Chaotic Inflation, *Physics Letters B*, **129**, 177–181.
34 The reader may be interested in the discussion of some of these issues in Chaps. 12 and 13 of *The Oxford Handbook of the History of Modern Cosmology*, *op. cit.*
35 For a more detailed discussion of these topics, see Section 12.2 of *Galaxy Formation*, *op. cit*.
36 Davis, T.M. and Lineweaver, C.H. (2004). Expanding Confusion: Common Misconceptions of Cosmological Horizons and the Superluminal Expansion of the Universe, *Publications of the Astronomical Society of Australia*, **21**, 97–109. doi:10.1071/AS03040.
37 Davis, T.M. and Lineweaver, C.H. (2004). *op. cit.*
38 Guth, A. (1981). Inflationary Universe: A Possible Solution for the Horizon and Flatness Problems, *Physical Review D*, **23**, 347–356.
39 These earlier works, particularly the work of the Soviet theorists Sakharov, Zeldovich, Starobinsky and their colleagues are surveyed in detail by C. Smeenk (2005). False Vacuum: Early Universe Cosmology and the Development of Inflation, in *The Universe of General Relativity, Einstein Studies*, eds. A.J. Kox and J. Eisenstaedt, **11**, 223–257. Boston: Birkhäuser.
40 Zeldovich, Ya.B. (1968). The Cosmological Constant and the Theory of Elementary Particles, *Uspekhi Fizisheskikh Nauk*, **95**, 209–230.
41 Linde, A. (1974). Is the Lee Constant a Cosmological Constant? *Zhurnal Experimentalnoi i Teoretichseskikh Fizica (JETP) Letters*, **19**, 183–184. Bludman, S. and Ruderman, M. (1977). Induced Cosmological Constant Expected above the Phase Transition Restoring the Broken Symmetry, *Physical Review Letters*, **38**, 255–257. A popular account of the history of the development of ideas about the inflation picture of the early Universe is contained in Guth's book *The Inflationary Universe: The Quest for a New Theory of Cosmic Origins* (1997). Reading, Massachusetts: Addison-Wesley. The pedagogical review by Lineweaver can also be recommended. He adopts a somewhat sceptical attitude to the concept of inflation and our ability to test inflationary models through confrontation with observations. (Lineweaver, C.H. (2005). Inflation and the Cosmic Microwave Background, in *The New Cosmology*, ed. M. Colless, 31–65. Singapore: World Scientific. doi:10.1142/97898127023570002).
42 Lanczos, K. (1922). Bemerkung sur de Sitterschen Welt, *Physikalische Zeitschrift*, **23**, 539–543.
43 Kibble, T.W.B. (1976). Topology of Cosmic Domains and Strings, *Journal of Physics A: Mathematical and General*, **9**, 1387–1398.
44 Kolb, E.W. and Turner, M.S. (1990). *op. cit.*
45 Linde, A. (1982). A New Inflationary Universe Scenario: A Possible Solution of the Horizon, Flatness, Homogeneity, Isotropy and Primordial Monopole Problems, *Physics Letters*, **108B**, 389–393. Linde, A. (1983). Chaotic Inflation, *Physics Letters B*, **129**, 177–181. Albrecht, A. and Steinhardt, P. (1982). Reheating an Inflationary Universe, *Physical Review Letters*, **48**, 1437–1440.
46 Parts of the text of this section use material from Smeenk, C. (2005). *op. cit.* and Smeenk, C. (2018). Inflation and the Origins of Structure, *Beyond Einstein, Einstein Studies*, eds. D. Rowe, S. Walter and T. Sauer, **14**. Berlin: Springer.
47 Mukhanov, V. and Chibisov, G. (1981). Quantum Fluctuations and a Nonsingular Universe, *JETP Letters*, **33**, 532–535.
48 Bardeen, J. (1980). *op. cit.*
49 Bardeen, J. *et al.* (1983). Spontaneous Creation of Almost Scale-Free Density Perturbations in an Inflationary Universe, *Physical Review D*, **28**, 679–693. doi:10.1103/PhysRevD.28.679.
50 Olive, K. (1990). Inflation, *Physics Reports*, **190**, 307–403, doi: 10.1016/0370-1573(90)90144-Q.

51 Liddle, A. and Lyth, D. (2000). *Cosmological Inflation and Large-Scale Structure.* Cambridge: Cambridge University Press.

52 There are now several recommendable books on this subject: Liddle, A. and Lyth, D. (2000). *op. cit.*; Dodelson, S. (2003). *Modern Cosmology*. Amsterdam: Academic Press; Mukhanov, V. (2005). *op. cit.*

53 I have found the pedagogical exposition by Daniel Baumann a very helpful guide. Baumann, D. (2007). On the Quantum Origin of Structure in the Inflationary Universe. arXiv:0710.3187 [hep-th].

54 Baumann, D. (2007). *op. cit.*

55 See Longair, M.S. (2008). *op. cit.*, Section 20.5.5.

56 Bertschinger, E. (1996). Cosmological Dynamics, in *Cosmology and Large Scale Structure*, eds. R. Schaeffer *et al.*, 273–346. Amsterdam: Elsevier.

57 The linearised perturbation equations (21.29) and (21.30) for the evolution of the perturbations on sub-horizon and super-horizon scales are often referred to as the Mukhanov–Sasaki equations.

58 Guth, A.H. and Pi, S.-Y. (1982). Fluctuations in the New Inflationary Universe, *Physical Review Letters*, **49**, 1110–1113. doi:10.1103/PhysRevLett.49.1110.

59 Shellard, P. (2003). The Future of Cosmology: Observational and Computational Prospects, in *The Future of Theoretical Physics and Cosmology*, eds. G.W. Gibbons *et al.*, 755–780. Cambridge: Cambridge University Press.

60 Sakharov, A. (1967). Violation of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe, *Zhurnal Experimentalnoi i Teoretichseskikh Fizica (JETP) Letters*, **5**, 32–35.

61 Kolb, E.W. and Turner, M.S. (1990). *op. cit.*

62 Kolb, E.W. and Turner, M.S. (1990). *op. cit.*

63 Gibbons, G. *et al.* (eds.) (2003). *The Future of Theoretical Physics and Cosmology*. Cambridge: Cambridge University Press.

Author Index

Subject Index

Index entries in **bold font** are the titles of sections of this book, as well as topics of central importance to the narrative. Roman fonts are used for subsections, subsubsections and topics of interest and importance. *Italic font* is used for the titles of major papers and books which are referred to in the main text. Generally, titles of references in the endnotes, which give more details of the citations, are not included in this index.