Physical Science

Sourcebook

Nuffield **A**dvanced **S**cience

Copy 2

Sourcebook Physical science

14 082758 7

Nuffield Advanced Science

The physical science group

J. E. Spice, Winchester College (organizer)

A. W. B. Aylmer-Kelly (1968–9), Royal Grammar School, Worcester, and Nuffield Advanced Chemistry Project

M. G. Brown (1965–7), University of Sussex

B. E. Dawson, King's College, London

H. Derbyshire (1967–9), Wyndham School, Egremont, Cumberland

D. G. F. Eastwood (1967–8), Fettes College

A. W. Gammage (1967–9), Liverpool Institute High School for Boys

A. Grant (1967–8), no previous appointment

A. D. C. Grassie, University of Sussex

S. W. Hockey, Marlborough College

I. B. Hopley (1965–7), Clifton College

D. E. P. Hughes (from 1967), Shrewsbury School

F. R. McKim, Marlborough College

D. J. Smith (1965–7), no previous appointment

Sourcebook
Physical Science

Edited by B. E. Dawson

Nuffield Advanced Science

Published for the Nuffield Foundation by Penguin Books

Penguin Books Ltd, Harmondsworth, Middlesex, England

Penguin Books Inc., 7110 Ambassador Road, Baltimore, Md 21207, U.S.A.

Penguin Books Ltd, Ringwood, Victoria, Australia

First published 1974

Filmset in Monophoto 10/12 Plantin by Keyspools Ltd, Golborne, Lancs.
and made and printed in Great Britain by Butler & Tanner Ltd,
Frome and London

Design and art direction by Ivan and Robin Dodd

Illustrations designed and produced by Penguin Education

Contents

Part 3

Part 4

Part 5

Part 6

Part 7

Part 8

Part 9

Part 10

Part 11

Foreword

The publication of Nuffield physical science completes the series of four Advanced Science Projects sponsored and initiated by the Trustees of the Nuffield Foundation. If it is the last of a series, so too is it an offering to the beginnings of a new era. Sixth forms in England and Wales still prepare students of high academic attainment for courses in higher education which in terms of the time allotted to them are among the shortest in the world. The consequent demands for standards on admission make our sixth-form programmes amongst the most rigorous, and yet we also try to ensure that this intensive study in depth does not deprive our students of the more broadly based education which so many desire and which they have every right to seek. It was to satisfy these needs that the Joint Committee for the Physical Sciences established by the Trustees under the guidance of Sir Nevill Mott, FRS, of the University of Cambridge, and the late Sir Ronald Nyholm, FRS, of University College, London, decided that an experimental programme should be established in physical science to bring together some of the most important underlying and unifying concepts in the separate sciences of chemistry and physics. It was hoped that in this way teachers and schools might have the best of both worlds – an opportunity to prepare students not only adequately but indeed better for further education in science, technology, and medicine, whilst at the same time permitting the inclusion of a non-science discipline in the usual combinations of subjects at advanced level. Trials have shown that these aims are indeed practicable. Physical science comes at a time when there is increasing pressure for the broadening of sixth-form studies. We believe it offers a well-tried example pointing to the direction in which these changes could go.

The physical science project has been under the direction of Dr John Spice of Winchester College (now Staff Inspector of Science for the Inner London Education Authority) and a most able team of physics and chemistry teachers from school and university. Their task has been enormous and no greater tribute could be paid to their achievement of integration than to record that physicists and chemists are both able to find within this work those important concepts common to both and particular to each.

Furthermore, the programme will, I believe, find a special and additional use in Colleges of Education and Colleges of Technology, and indeed the materials may also have much of value in preparatory courses at university level.

To Dr Spice, Dr Dawson (the editor of this book), and their colleagues, to the Joint Committee for the Physical Sciences, to all those who helped in schools and Local Education Authorities in the development and success of the trials, to the Cambridge Local Examinations Syndicate, to William Anderson and his colleagues in Nuffield Publications, to Penguin Education for their help, kindness, and ever-willing assistance, and particularly to the Trustees of the Nuffield Foundation, we offer our sincere gratitude for making possible this new step forward in science education.

K. W. Keohane
Coordinator of the Nuffield Foundation Science Teaching Project

Acknowledgements

The Editor wishes to acknowledge the generous assistance he has received throughout the pre-publication period from the various contributors to this *Sourcebook* and to the original set of *Source Papers*. Specific acknowledgements appear in the text and a list of contributors of articles for the volume appears separately. The response to initial requests for appropriate material was so great that it has been necessary to restrict the total number of articles for publication. The compilation of materials has been no simple task and much credit rests with those who have assisted in various ways during the preparative period, especially Mr A. W. Gammage, Mrs P. Bewley, Mr G. R. Pierce, and Mrs J. F. Harris.

It proved an even more difficult task to select representative examples of the projects devised by students from the schools and colleges using the Nuffield physical science course materials. Indeed, it has been possible only to include in abstract form a small proportion of the total number of project reports submitted for examination purposes and these are acknowledged individually. The task of coordinating these entries was undertaken by Mr D. R. Browning, editor of *SATIS* – the abstracts journal associated with *Schools Council Project Technology*. (It should be noted that a detailed list of schools and colleges contributing to the trials of physical science materials appears in the *Introduction and Guide* and is not reproduced in this book).

Acknowledgement is also made to Mr W. Anderson, Mrs H. Ellis, and Mrs D. Williams of the Publications Department of the Nuffield Foundation Science Teaching Project for their advice and expert assistance during the production of this book.

Acknowledgements

The Editor wishes to acknowledge the generous assistance he has received throughout the pre-publication period from the various contributors to this *Sourcebook* and to the original set of *Source Papers*. Specific acknowledgements appear in the text and a list of contributors of articles for the volume appears separately. The response to initial requests for appropriate material was so great that it has been necessary to restrict the total number of articles for publication. The compilation of materials has been no simple task and much credit rests with those who have assisted in various ways during the preparative period, especially Mr A. W. Gammage, Mrs P. Bowley, Mr G. R. Pierce, and Mrs J. F. Harris.

It proved an even more difficult task to select representative examples of the projects devised by students from the schools and colleges using the Nuffield physical science course materials. Indeed, it has been possible only to include in abstract form a small proportion of the total number of project reports submitted for examination purposes and these are acknowledged individually. The task of coordinating these entries was undertaken by Mr D. R. Browning, editor of *SATIS* – the abstracts journal associated with *Schools Council Project Technology*. (It should be noted that a detailed list of schools and colleges contributing to the trials of physical science materials appears in the *Introduction and Guide* and is not reproduced in this book).

Acknowledgement is also made to Mr W. Anderson, Mrs H. Ellis, and Mrs D. Williams of the Publications Department of the Nuffield Foundation Science Teaching Project for their advice and expert assistance during the production of this book.

Contributors to the *Sourcebook*

Professor D. R. Arthur, Department of Zoology, University of London King's College

Dr J. R. Atkinson, Department of Metallurgy, Houldsworth School of Applied Science, The University of Leeds

A. W. B. Aylmer-Kelly, Department of Chemistry, University of Southampton

The late Professor F. P. Bowden FRS, formerly Department of Surface Physics, University of Cambridge

D. R. Browning, Department of Science, Bristol Polytechnic

C. D. Curling, Faculty of Natural Science, University of London King's College

Dr A. Darbre, Faculty of Medical Science, University of London King's College

Dr B. E. Dawson, Faculty of Education, University of London King's College

Dr H. Derbyshire, Reed International Limited

D. G. F. Eastwood, Gordonstoun School, Moray

Dr John Emsley, Department of Chemistry, University of London King's College

Dr W. V. Farrar, Department of History of Science and Technology, University of Manchester Institute of Science and Technology

J. M. Faulkner, formerly William Hulme Grammar School, Manchester

R. Gallear, Malmesbury School, Wiltshire

A. W. Gammage, Minchenden School, Hertfordshire

Mrs H. Gammage

Dr P. S. H. Henry, formerly The Shirley Institute, Manchester

Professor S. E. Hunt, Department of Physics, The University of Aston in Birmingham

Dr J. H. E. Jeffes, Department of Metallurgy, Imperial College of Science and Technology, University of London

Dr A. C. Legon, Department of Chemistry, University College London

Dr F. R. McKim, Marlborough College, Wiltshire

T. S. Moss, Royal Aircraft Establishment, Farnborough

Dr R. B. Moyes, Department of Chemistry, The University, Hull

Dr J. E. Nicholas, Department of Chemistry, University of London King's College

Dr A. E. Owen, School of Engineering Science, University of Edinburgh

M. W. Poole, Faculty of Education, University of London King's College

Dr V. V. Raman, Department of Physics, Rochester Institute of Technology, New York, U.S.A.

Professor J. P. Roberts, Department of Ceramics, Houldsworth School of Applied Science, The University of Leeds

Professor A. J. B. Robertson, Department of Chemistry, University of London King's College

Dr M. A. Rothman, Department of Physics, Trenton State College, New Jersey, U.S.A.

Dr A. G. Sharpe, Department of Organic and Inorganic Chemistry, University of Cambridge

Dr A. R. Stokes, Department of Physics, University of London King's College

Dr A. G. Sykes, Department of Inorganic and Structural Chemistry, The University of Leeds

Dr D. Tabor FRS, Department of Surface Physics, University of Cambridge

Dr E. Wichers, Division of Chemistry and Chemical Technology, National Research Council, National Academy of Sciences, Washington D.C., U.S.A.

Dr R. V. Williams, Physics Department, British Iron and Steel Research Association

Introduction

This book is not a textbook in the generally accepted sense. Its purpose is to provide the means for students to become familiar with the activities and thoughts of some practising scientists and technologists on specific topics. The *Sourcebook* is intended to complement the various volumes of the *Students' workbook* for the Nuffield physical science course.

Experience gained during the school trials of this course has shown that students require something more than formal classwork and formal laboratory exercises on which to base their ideas and experience. Indeed, discussion and a selective use of relevant literature can be significant for students. To stimulate such activities, a small number of articles were prepared for use during the trials of course materials. The present collection has been derived from experience gained in these trials. Obviously, many different selections of articles relating to the physical science course could have been made and this collection does not seek to provide a complete coverage: no one book can replace the use of a well-stocked library equipped to demonstrate the dependence of our civilized society upon the use of scientific knowledge and on the adoption of technological innovations.

In this collection, some attempt has been made to relate separate aspects of the common themes and topics of the physical science course, and to reflect current concern for a right use of all scientific knowledge and its subsequent evaluation in relation to specific needs. Utilitarian applications of scientific knowledge and technological innovations bring not only benefits but also responsibilities to the professional scientist and engineer as well as to the ordinary citizen. Indeed, the issue of responsibility can be considered from several points of view and it forms a special feature of that part of the collection where the impact of our civilization on the environment is reviewed. It should be noted that not all of the articles are concerned with scientific advances or with specialized applications of basic concepts. Some direct attention to patterns of development and ideas, others provide commentaries on our present knowledge, whereas yet others can be classified as original contributions which have been reprinted for the general convenience of students following the Nuffield physical science course.

In addition, each part of the *Sourcebook* contains a section written by students about their own projects – a necessary requirement of the physical science course. These contributions are in the now familiar form of abstracts used in *SATIS*, a publication issued by the Schools Council Project Technology. (Additional information relating to the development of project work as an integral part of the Nuffield physical science course appears, for example, in the *Introduction and guide*).

To the student who uses this collection as an aid to study in the physical sciences, we hope that your course work will enable you to make appropriate references to various articles. You should read to extend your general knowledge of the topic under discussion – even if you do not fully understand some of the theoretical parts at first reading. As you progress through your course, theoretical matters will receive a more detailed treatment and you will be able to re-read appropriate articles to gain additional understanding and insights. One very important function of this collection is to provide the reader with the ideas and views of other scientists. No one will expect you to sit down and learn the contents of this book: it has been compiled to encourage you to explore and test matters for yourself in the hope that you will enjoy such inquiries.

Part 1

Forces and motion

Contents

Introduction

A great deal of our knowledge of forces and motion is summarized concisely by Newton's laws. It is therefore appropriate for this part of the Sourcebook *to open with a statement of and a commentary on these laws in Newton's own words. Newton submitted his* Philosophiae Naturalis Principia Mathematica *to the Royal Society in three instalments during 1686 and 1687. The English translation was completed about two years after Newton's death by Andrew Motte and this is also reproduced. A logical survey of these deceptively 'simple' laws of nature is provided in the next article and underlines for the reader just how difficult it is to be aware of basic assumptions. Subsequent articles deal with the implications of these principles and ideas in daily life. To relate principles formulated some three centuries ago to current theoretical ideas is equally significant and so the concluding article considers the four types of interaction which account for all changes in the chemico-physical world, governed by various principles of conservation.*

Related sections of the physical science course

Section 1	Forces, motion, and energy
Section 12	Simple harmonic motion and wave motion
Option G3	Rotational motion

Related sections of other Nuffield Advanced Science courses

Physics

Unit 1	Materials and structure
Unit 2	Electricity, electrons, and energy levels
Unit 3	Field and potential
Unit 4	Waves and oscillations
Unit 5	Atomic structure
Unit 10	Waves, particles, and atoms

1.1 Newton's laws of motion

Advances in scientific knowledge and understanding often take a leap forward after the formulation of a general principle. Among such formulations are Mendeleev's Periodic Table (1869), Einstein's Laws of Relativity (1905), Heisenberg's Uncertainty Principle (1925), and Darwin's contributions to the Theory of Evolution through his *Origin of Species by Means of Natural Selection* (1859). A fundamental and even earlier contribution was made by Newton. Newton's laws of motion first appeared in his now famous book *Philosophiae Naturalis Principia Mathematica* in 1686–7. They were translated into English by Andrew Motte and published in *Mathematical principles of natural philosophy*, 1729. The reproductions of relevant parts of these two books – given on the pages which now follow – appear through the courtesy of the Royal Society.

AXIOMATA
SIVE
LEGES MOTUS

Lex. I.

Corpus omne perſeverare in ſtatu ſuo quieſcendi vel movendi uniformiter in directum, niſi quatenus à 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 adjicitur, & cum eo ſecundum utriuſq; determinationem componitur.

Lex. III.

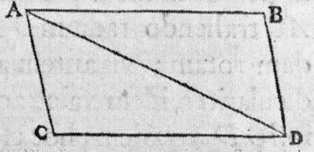
Actioni contrariam ſemper & æqualem eſſe reactionem: ſive corporum duorum actiones in ſe mutuo ſemper eſſe æquales & in partes contrarias dirigi.

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 *B D* parallelam, hæc vis nihil mutabit velocitatem accedendi ad lineam illam *B D* a vi altera genitam. Accedet igitur corpus eodem tempore ad lineam *B D* ſive vis *N* imprimatur, ſive non, atq; adeo in fine illius temporis reperietur alicubi in linea illa.



Axioms or Laws of Motion.

LAW I.

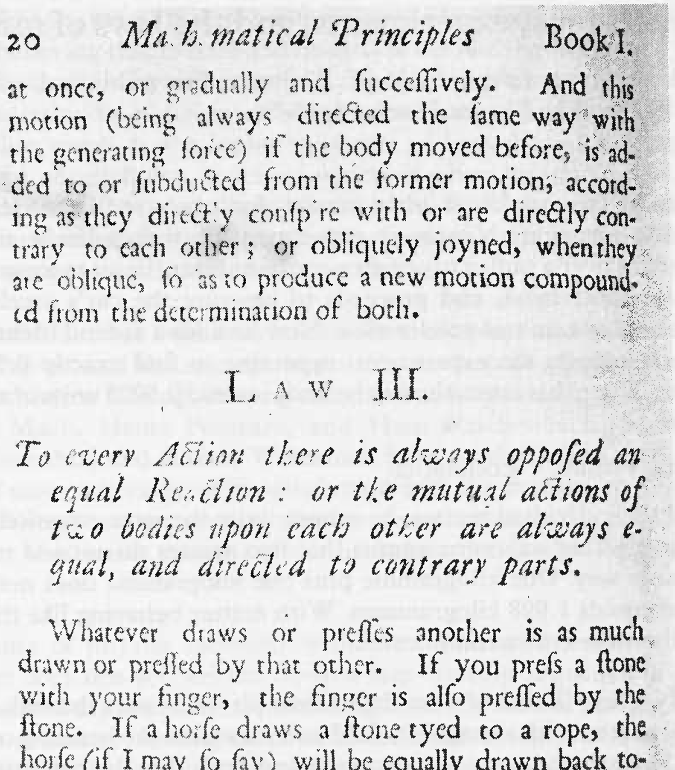
Every body perſeveres in its ſtate of reſt, or of uniform motion in a right line, unleſs it is compelled to change that ſtate by forces impreſs'd thereon.

PRojectiles perſevere in their motions, ſo far as they are not retarded by the reſiſtance of the air, or impell'd downwards by the force of gravity. A top, whoſe parts by their coheſion are perpetually drawn aſide from rectilinear motions, does not ceaſe its rotation, otherwiſe than as it is retarded by the air. The greater bodies of the Planets and Comets, meeting with leſs reſiſtance in more free ſpaces, preſerve their motions both progreſſive and circular for a much longer time.

LAW II.

The alteration of motion is ever proportional to the motive force impreſs'd; and is made in the direction of the right line in which that force is impreſs'd.

If any force generates a motion, a double force will generate double the motion, a triple force triple the motion, whether that force be impreſs'd altogether and



20 *Mathematical Principles* Book I.

at once, or gradually and ſucceſſively. And this motion (being always directed the ſame way with the generating force) if the body moved before, is added to or ſubducted from the former motion; according as they directly conſpire with or are directly contrary to each other; or obliquely joyned, when they are oblique, ſo as to produce a new motion compounded from the determination of both.

LAW III.

To every Action there is always oppoſed an equal Reaction: or the mutual actions of two bodies upon each other are always equal, and directed to contrary parts.

Whatever draws or preſſes another is as much drawn or preſſed by that other. If you preſs a ſtone with your finger, the finger is alſo preſſed by the ſtone. If a horſe draws a ſtone tyed to a rope, the horſe (if I may ſo ſay) will be equally drawn back towards the ſtone: For the diſtended rope, by the ſame endeavour to relax or unbend it ſelf, will draw the horſe as much towards the ſtone, as it does the ſtone towards the horſe; and will obſtruct the progreſs of the one as much as it advances that of the other. If a body impinge upon another, and by its force change the motion of the other; that body alſo (becauſe of the equality of the mutual preſſure) will undergo an equal change, in its own motion, towards the contrary part. The changes made by theſe actions are equal, not in the velocities, but in the motions of bodies; that is to ſay, if the bodies are not hinder'd by any other impediments. For becauſe the motions

4 are

Motte's translation and discussion of Law III continues on the next page: 'are equally changed, the changes of the velocities made towards contrary parts, are reciprocally proportional to the bodies. This Law takes place also in Attractions, as will be proved in the next Scholium.'

1.2 Definitions, experiments, and the laws of motion

Adapted from a paper by M. A. Rothman, first published in 1970 in *The Physics Teacher* **8**, 16–21.

How lucky we are that kilogramme masses don't behave like nuclear particles. Imagine this situation: Newton is experimenting with a linear air-track. He hangs a weight over a pulley to obtain a constant force, ties it to a car loaded with a one-kilogramme mass, and proceeds to measure the car's acceleration. He finds, let's say, one unit of acceleration. Now he adds a second identical mass to the first and repeats the experiment, expecting to find exactly 0.500 units of acceleration. But to his astonishment he finds instead 0.5005 units of acceleration. If he writes:

$$\text{force} = \text{mass} \times \text{acceleration}$$

for each of the individual masses, he cannot write the same equation for the two masses put together unless he admits that two masses do not add together in a simple, linear way. One kilogramme plus one kilogramme does not equal two, but instead equals 1.998 kilogrammes. With matter behaving like this, the laws of motion become rather complicated.

An unlikely story? Think of what the nuclear physicist goes through. He applies a force to a proton with a magnetic field and measures the bending of its path in that field. He does the same to a deuteron and finds that the bending of its path is about a tenth of a per cent greater than he would expect by adding the mass of the proton to the mass of the neutron. Does he say that Newton's laws of motion don't apply? Does he say that the magnetic field acts differently on the proton when it is attached to a neutron? No, he simply says that the deuteron has less mass than expected, and when he looks closely enough he sees that this is because some mass left the scene in the form of a gamma ray photon emitted when the neutron and proton combined.

If kilogramme masses behaved in the same manner as the neutrons and protons, Newton would have found it difficult to express the laws of motion in a simple way. But why do we assume that kilogramme masses *don't* behave like neutrons and protons? After all, they attract each other with a gravitational force, and if the gravitational binding energy is allowed to escape from the system, then there will be a loss of mass. What makes Newton so lucky is the fact that the gravitational binding energy is so infinitesimally small that there is no possibility of noticing it in any laboratory situation. Therefore we can get away with the assumption that two equal masses put together add up to twice the individual masses.

This story has the moral that no assumption can be taken for granted. The only reason for believing that masses add linearly is the fact that this is the way we see them behave under most circumstances. The moment we try to reach more than a casual understanding of Newton's laws of motion we find that hidden behind those familiar words is a complex structure of hypotheses, assumptions, and experimental observations, much of which is never brought into the light of awareness, even in the most careful textbooks. Furthermore, there is a great amount of confusion as to what parts of the laws are simply definitions and what parts are based on experiments and observations.

This state of affairs has been well known for many years and an entire literature has grown up around the attempt to put Newton's laws of motion on a logical and self-consistent footing. Among the classic writers dealing with the subject are Ernst Mach, Henri Poincaré, and Hans Reichenbach. More recently, Leonard Eisenbud and Robert Weinstock have published detailed analyses of the laws of motion because of dissatisfaction with the customary treatments. As readers will be aware, the general trend among elementary physics textbooks is to present a reasonable amount of the background and logic behind Newton's laws.

The teaching of physics contrasts with the teaching of mathematics in that hardly ever does one see written down a step-by-step sequence of definitions, hypotheses, and experimental tests leading up to the laws of motion in the spirit of Euclid's geometry, leaving no holes in the chain of logic. One reason for this state of affairs is the fact that there is no general agreement on one best logical sequence. Indeed, there are some violent differences of opinion on the best way to approach the subject.

A fundamental reason for these differences of opinion is the fact that the laws of motion are not absolute laws handed down by higher authority, but rather they represent man's attempt to build in his mind a conceptual structure than enables him to interpret what goes on out in the world. It is the oldest problem of philosophy: how do we know what the universe is like when all that comes into our brains are coded sequences of pulses along a number of nerve channels from our sensory organs? We do this by building up in our minds a set of concepts such as acceleration, force, and mass. The organization of these concepts into a unified structure is the business of science.

Trouble arises when different people invent different structures all claiming to describe the same things existing 'out there'. Newton's statement that $F = ma$ is a simple thing, but there are at least three ways of interpreting it. The traditional engineering approach stems from the muscular feeling you have when you push or pull something. As a result, the engineer begins by defining the unit of force to be the pound force (represented by the weight of a standard pound). The unit of mass is then defined as that mass which responds to the pound force with an acceleration of 1 ft s^{-2}, and we end up with that ungainly unit of mass, the slug.

The school of physicists following the philosophy of Ernst Mach notices that forces are not really the results of direct contact between hard objects, but that they are transmitted through empty space by means of 'fields'. As a result, the point of view is changed, and we say: here we see an object undergoing acceleration, so there must be a force acting on it, even though we can't see anything touching it. If we know the mass of the object, we can define the amount of force by the equation $F = ma$. The initial definition of the object's mass is then the problem and Mach solved it by saying: we compare an unknown mass with a standard mass by using an inertia balance (see figure 1.2.1). When we push the two masses apart by a spring or other force and measure their velocities (or accelerations) the ratio of the masses is given by the ratio of the velocities.

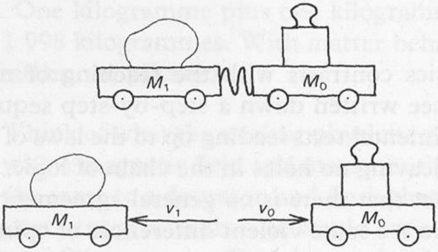


Figure 1.2.1
An inertia balance. The two cars are pushed apart by a spring, and their velocities are measured simultaneously. The masses are so defined that $M_1/M_0 = |v_0/ v_1|$.

A third approach is a compromise between the two. A standard kilogramme is defined, and then a spring is calibrated by measuring the acceleration it gives to that standard mass while the elongation of the spring is measured. We then use that calibrated spring as a known force to find the value of unknown masses by measuring their acceleration. (See figure 1.2.2.)

In this third approach (which we may call the Newtonian method) as well as in the Machian approach, we compare unknown masses with a standard mass by means of a spring. In the Machian method we use the spring to accelerate both

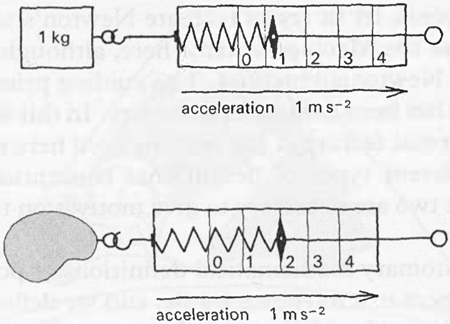


Figure 1.2.2
A spring is calibrated by using it to accelerate a standard kilogramme. If that same spring is used to accelerate an unknown mass, that mass is then proportional to the force required to give unit acceleration.

masses simultaneously, while in the Newtonian method we pull the objects one at a time, trusting that the properties of the spring do not change between pulls. However, the psychology between the two methods is quite different. The Newtonian method says, in effect: I have a *known force* that I apply to a mass. If I measure the acceleration this tells me the amount of mass. If I know the mass, then I can find out what the acceleration is going to be. This is indeed the way we go about solving practical problems when the source of the force is a known quantity.

The Machian method, on the other hand, has a very interesting connotation. It says: if I know the mass of an object and I measure its acceleration, then this operation tells me how much force is causing this acceleration. It must be given by $F = ma$, by definition, regardless of the source. This is the method used to discover new forces, and in fact was the method used by Newton when he realized that the acceleration of the Moon toward the Earth was caused by the gravitational force.

Now, however, we run into a serious logical difficulty. If $F = ma$ is true by definition, then what is the point of all the experiments we perform to demonstrate that the laws are 'true'? When we try to answer this question by close analysis of Newton's laws, we find ourselves going through a thorny philosophical thicket, but when we come out on the other side we find that we begin to understand what a 'law of nature' really is, and what experiments must be performed to verify each law.

To illustrate this process, let us try to recreate Newton's second law of motion step by step. I will use the Machian scheme here, although the same procedure can be applied to the Newtonian method. The guiding principle is that no word may be used before it has been defined in some way. In this way we avoid circular reasoning. One important feature of the analysis used here is that I find it useful to employ three different types of definitions: conceptual, behavioural, and operational. The first two are necessary to give motivation for the third.

We start with the customary mathematical definitions of position, velocity, and acceleration with respect to a reference frame, and we define a physical straight line to be the path followed by a light ray in free space. We also bring with us the primitive conceptual definition of a force as a push or a pull.

When we try to make a more sophisticated definition of force, we ask ourselves what it is that a force does, and we see that force is invariably associated with acceleration. When a force is suddenly applied or removed, the position and velocity of the object remain continuous, but it is the *acceleration* that is discontinuous. This observation encourages us to make a behavioural definition of force: *a force is anything which can cause the acceleration of an object.*

Now, just because we have made a definition, this does not ensure that the defined object exists. Hidden behind every definition is the unconscious hypothesis that the thing being defined does, indeed, exist. Let us bring this hypothesis out into the open and say: if an object is seen to undergo acceleration, then there exists a force that is acting on it. In other words, an object changes its state of motion if, and only if, there is a force acting on it.

The only thing to do with a hypothesis like this is to test it. If, every time we see an object being accelerated, we can identify the origin of the force responsible for that acceleration, then the hypothesis is a good one. When we make this test, we find that there appear to be two kinds of forces operating in the universe. First there are forces that come from an identifiable source: gravitational, electromagnetic, or stretched springs. Then there are forces whose source is not so easily identifiable, felt only when the observer happens to be in a reference frame that is rotating or otherwise being accelerated. These are the *inertial forces* – the centrifugal, Coriolis, and so-called '*g*-force' felt by a rocketing astronaut (figure 1.2.3).

There are three possible things we can do with this information. First, we can say that our hypothesis about the nature of force is all wrong, since not all forces have identifiable sources. Second, we can say that our force hypothesis is correct, but only in certain special frames of reference in which inertial forces do not

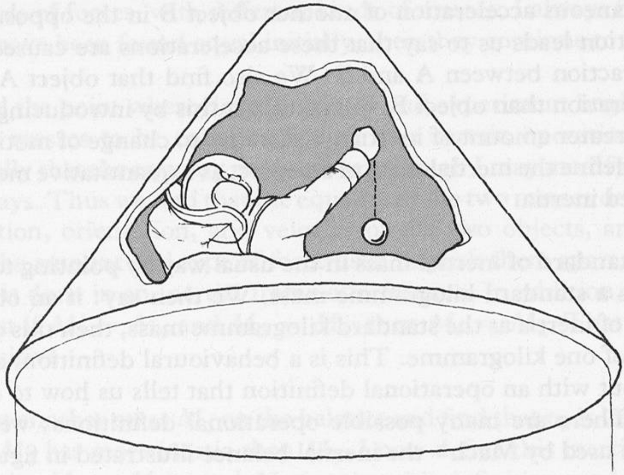


Figure 1.2.3
An astronaut in a rocket accelerating feels an inertial force pushing him down to his seat. If he drops a ball, he sees that ball accelerating downward as if a force is acting on it.

make themselves felt. In these 'inertial reference frames' we can always identify any local source of force causing an object to accelerate. In other – non-inertial – reference frames an object appears to accelerate without apparent cause, and we say that there is a 'fictitious' force acting on that object. We soothe our outraged sense of causality by saying that this action is merely the result of the inertia of the object, and sweep under the rug the question of why objects have inertia in the first place.

The third thing we can do is to keep open the possibility that the inertial forces actually can be identified as real interactions with objects elsewhere in the universe. If we follow this path we find ourselves following in the footsteps of Albert Einstein towards the General Theory of Relativity. If we want to remain within the framework of Newtonian theory, we use the second option, but keep in the back of our minds the fact that this is a man-made choice, and that the third option may lead us closer to understanding the way nature actually behaves.

Having made at least a partial definition of force, we now continue in our scheme to find a definition of mass. We start with the observation that in an inertial reference frame we always find that the acceleration of object A is associated with a simultaneous acceleration of another object B in the opposite direction. This observation leads us to say that these accelerations are caused by a force-pair, an interaction between A and B. We also find that object A may have a greater acceleration than object B. We interpret this by introducing the concept that B has a greater amount of inertia – resistance to change of motion – than A, leading us to define the inertial mass of an object as a quantitative measure of that property called inertia.

We define a standard of inertial mass in the usual way by pointing to a thing and saying: this is a standard kilogramme mass. We then say: if an object has the same amount of inertia as the standard kilogramme mass, then this object has an inertial mass of one kilogramme. This is a behavioural definition, and we must now fill this out with an operational definition that tells us how to compare the two masses. There are many possible operational definitions; we choose one similar to that used by Mach – the inertial balance illustrated in figure 1.2.1.

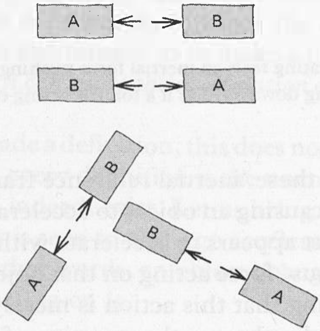


Figure 1.2.4
A test for a symmetrical inertial frame of reference. Two cars are pushed apart by a spring in a number of different orientations. If the cars all travel with the same speed regardless of orientation, then the masses A and B are equal, and the reference frame is symmetrical. Notice that it takes at least two observations to prove both points.

To make sure that there are no misleading extraneous influences, I have the spring push the two cars apart a number of times, each time with a different orientation (see figure 1.2.4). If the two cars end up travelling with the same speed regardless of orientation, then the two masses are equal. This procedure establishes a frame of reference with a certain degree of symmetry. Without this precaution we could find ourselves measuring masses that were a function of the direction of motion.

Having established the symmetry of our frame of reference (which we define as an inertial frame), I proceed with the crucial part of the experiment. Now I continue the comparison of the two masses under many different conditions: with different amounts of spring compression, with different springs, with different kinds of forces, with different kinds of masses. I always find that when two masses have been found equal initially, then they continue to be equal.

Here we find the point where definition ends and experiment begins: initially, I defined two masses to be equal according to a certain operation. I then find experimentally that the equality persists even though I may modify the operation in various ways. Thus we find that the equality of the two masses does not depend on the position, orientation, and velocity of the two objects, and it does not depend on the amount and type of force used to push them apart, as long as the experiment is done in an inertial frame of reference. In addition, I find experimentally that if $M_1 = M_0$ and $M_2 = M_1$, then $M_0 = M_2$. That is, equality of mass is transitive.

Now, if I put another mass M_3 on the balance and find that v_3 is less than v_0, this means that M_3 has more inertia than M_0. I now define this inertial mass M_3 to be the quantity $M_3 = M_0|v_0/v_3|$. Notice that this definition used one particular spring under one set of conditions. Now I find that if I repeat the experiment with the variations described above (different springs, etc.) then I always find M_3 to be the same.

This discovery leads me to rearrange the above equation so that it reads

$$M_0v_0+M_3v_3 = 0$$

and since it is always true, I call it a law of nature and name it conservation of momentum. The criticism is often made that if we define mass by the above recipe, then we have made conservation of momentum true by definition, and we wonder what is the experimental content of the law. We now see what it is: when we define the mass of an object by one measurement under one set of conditions, then the above relationship continues to be true for all time regardless of the kind of force used to push the masses apart, and regardless of the materials they are made of. *The invariance of the relationship is the essence of the law.*

Other important relationships remain to be shown by experiment. For example, if $M_1 = M_2$, then $M_1+M_2 = 2M_1$ (linearity of addition of mass). This property allows us to set up a simple system of mass units, and as described at the beginning of this article, it is a property that cannot be taken for granted or defined into existence.

Having defined mass, we are now ready for the quantitative definition of force, which we do by saying that whenever we see an object being accelerated, the amount of force acting on that object is equal numerically to the rate of change of momentum (the momentum being defined as Mv). For constant mass this becomes our familiar friend $F = Ma$.

Once more the question is raised: what right do we have to call this a law of nature? If $F = Ma$ is true by definition, what is the experimental content of the law? The experimental content is considerable, but not always obvious. First of all we observe that if we do these experiments in an inertial frame of reference, a force acting on mass M always arises from an interaction with another object – that is, a stretched spring, an electric charge at a given distance, another mass, etc. Second, we observe that if a particular interaction (for example, a particular spring stretched a certain distance) causes a force of magnitude $F = Ma$, then the same interaction under the same conditions will always give the same mass the same acceleration. It is this constancy of the interaction that allows us to calibrate a spring to measure force. Third, if a particular interaction acts on one object with a force $F_1 = M_1a_1$, and if the same interaction acts on another object with a force $F_2 = M_2a_2$, then it is found that $F_1 = F_2$. That is, the same spring produces the same force (for a given stretch) regardless of the object it is accelerating. This is another property that makes springs useful. In general, if a given arrangement of objects produces a certain force, then this same arrangement will always produce the same amount of force.

Fourth, and most important, if a number of forces (in terms of stretched springs) act on an object, then we observe that the forces add linearly according to the rules of vector addition.

We see that the essence of the law is the invariance of the relationship $F = Ma$; once a force is defined by this relationship, the relationship applies consistently under all circumstances. The experimental tests of the law must deal with the invariance properties and the summation properties, not with the original definition of force.

We can see how this logic works in practice by looking at some of the experiments done in the teaching laboratory with springs and masses. Suppose we take a kilogramme mass, accelerate it by pulling it with a spring, and measure both the acceleration and the spring's stretch. This operation tells us how much force was being applied, and so we are able to calibrate the spring. If we now use the same spring on another mass, we can come to two different conclusions, depending on whether we are using the Newtonian approach or the Machian approach. According to the first, once the spring is calibrated we know how much force it applies, and so we can use it on a number of other masses to show experimentally that the acceleration is proportional to the force and inversely proportional to the mass. But according to the Machian view, any time we accelerate a known mass (previously measured by an inertia-balance experiment) the amount of force is given by the mass times the acceleration, so that all the experiment can show is that the calibrated spring retains its calibration unchanged every time it is used.

Suppose we add a second spring in parallel with the first. The Newtonian method would say this experiment shows that doubling the force doubles the acceleration. According to the Machian point of view we note that the acceleration doubled, and therefore the force must have doubled. The experiment thus verifies the fact that parallel forces exerted by two springs add linearly.

We see from these examples that the conclusions drawn from an experiment depend very much on the point of view of the experimenter. We also see that in order to design an experiment to verify a law of nature we must be very clear in our minds what parts of the law are definitions and what parts remain free to be verified.

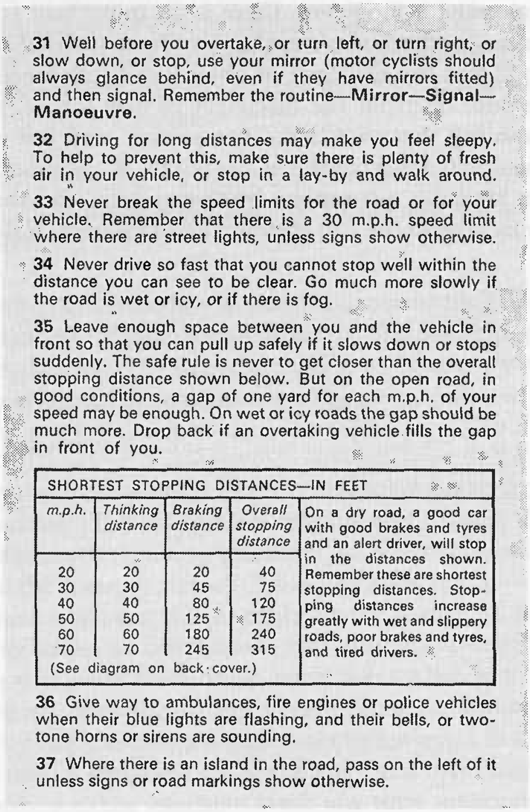
Both the Newtonian and Machian points of view are equivalent as far as practical applications are concerned. However, there is an important conceptual difference. The Newtonian method assumes that you have defined all the forces and all you want to do is find the response of objects to these forces. The Machian method leaves the door open for the discovery of new forces, and in particular allows for the possibility that 'fictitious' forces may turn out to be real forces. The unresolved question of inertial forces has led D. W. Sciama to say 'Newton's laws of motion are logically incomplete by themselves, and the problems they raise lead one step by step to the full complexity of General Relativity'.

1.3 Forces and cars

by A. W. Gammage

Avoiding crashes

The extract from the current issue of the Highway Code (figure 1.3.1) gives figures for stopping distances in terms of 'thinking distance' and 'braking distance', and tells drivers not to get too close to vehicles in front of them. Forces and motion are very much involved in what happens when the car ahead makes a sudden emergency stop.



31 Well before you overtake, or turn left, or turn right, or slow down, or stop, use your mirror (motor cyclists should always glance behind, even if they have mirrors fitted) and then signal. Remember the routine—**Mirror—Signal—Manoeuvre.**

32 Driving for long distances may make you feel sleepy. To help to prevent this, make sure there is plenty of fresh air in your vehicle, or stop in a lay-by and walk around.

33 Never break the speed limits for the road or for your vehicle. Remember that there is a 30 m.p.h. speed limit where there are street lights, unless signs show otherwise.

34 Never drive so fast that you cannot stop well within the distance you can see to be clear. Go much more slowly if the road is wet or icy, or if there is fog.

35 Leave enough space between you and the vehicle in front so that you can pull up safely if it slows down or stops suddenly. The safe rule is never to get closer than the overall stopping distance shown below. But on the open road, in good conditions, a gap of one yard for each m.p.h. of your speed may be enough. On wet or icy roads the gap should be much more. Drop back if an overtaking vehicle fills the gap in front of you.

SHORTEST STOPPING DISTANCES—IN FEET

m.p.h.	*Thinking distance*	*Braking distance*	*Overall stopping distance*	
20	20	20	40	On a dry road, a good car with good brakes and tyres and an alert driver, will stop in the distances shown. Remember these are shortest stopping distances. Stopping distances increase greatly with wet and slippery roads, poor brakes and tyres, and tired drivers.
30	30	45	75	
40	40	80	120	
50	50	125	175	
60	60	180	240	
70	70	245	315	

(See diagram on back cover.)

36 Give way to ambulances, fire engines or police vehicles when their blue lights are flashing, and their bells, or two-tone horns or sirens are sounding.

37 Where there is an island in the road, pass on the left of it unless signs or road markings show otherwise.

Figure 1.3.1
Extract from the Highway Code (1968).
Department of the Environment (1968) The Highway Code. *Crown copyright, reproduced by permission of the Controller of Her Majesty's Stationery Office.*

Imagine you are driving the car behind. Assume both cars are moving initially at the same velocity, the stopping is smooth and without skidding, and the deceleration is the same for both cars. We might take this to be $0.3g$* (or, preferably, set up an experiment to measure it). How do you decide the separation that will ensure that the two cars do not touch?

When the brake lights of the car come on, a short interval elapses before you apply your brakes – in this time you travel the 'thinking distance'. Measure your reaction time and remember it must be your eye-to-foot reaction time. This can be done by stopping the fall of some object with your right foot. Before the object is unexpectedly allowed to fall, your right foot must be pressing against something that simulates the accelerator pedal. You must move your foot the appropriate distance left to stop the falling object.

Alternatively, an electrical timing system can be used with a red light coupled, for example, to a sewing-machine foot-switch. Do not practice to shorten your reaction time; you only get one chance on the road! Test several people and use the worst result.

Now compare this 'thinking time' with that in the Highway Code, and work out what the *minimum* separation between cars should be for several different velocities. Here are a few questions you might try to answer:

1 Does the 'thinking time' vary with age, sex, or tiredness?

2 Is it *significantly* different between experienced drivers and non-drivers?

3 If the front car's brakes are poorer than yours, does that change the result?

4 Is the safe separation proportional to the velocity as the Highway Code suggests (Section 35)?

5 Does the value of the deceleration influence the result if it is the same for both cars? Or if it is not the same?

6 Can you present your results so as to satisfy a 'jury' of non-scientists?

Protection in car crashes

This part of the article is about protecting the occupants of a colliding car from fatal or serious injuries. The measures suggested may not always be best for protection against minor injuries sustained in a low speed collision. In collisions at 16 km h^{-1} (10 mile h^{-1}), soft padding on the dashboard, steering wheel, etc., may be effective but such precautions are virtually useless for higher speeds.

---

* g is the acceleration due to gravity.

The fundamental problem is to disperse the kinetic energy of the occupant, $\frac{1}{2}mu^2$ (where m = mass of occupant and u = velocity of car just before collision), in the most satisfactory way, and the following points are worth considering:

1 the retardation should be spread over the maximum possible time so that the force on the body at any instant is minimal;

2 the retarding force should be spread over a large area so that the pressure on a given part of the body is reduced;

3 the retarding force should be applied to the strongest parts of the body.

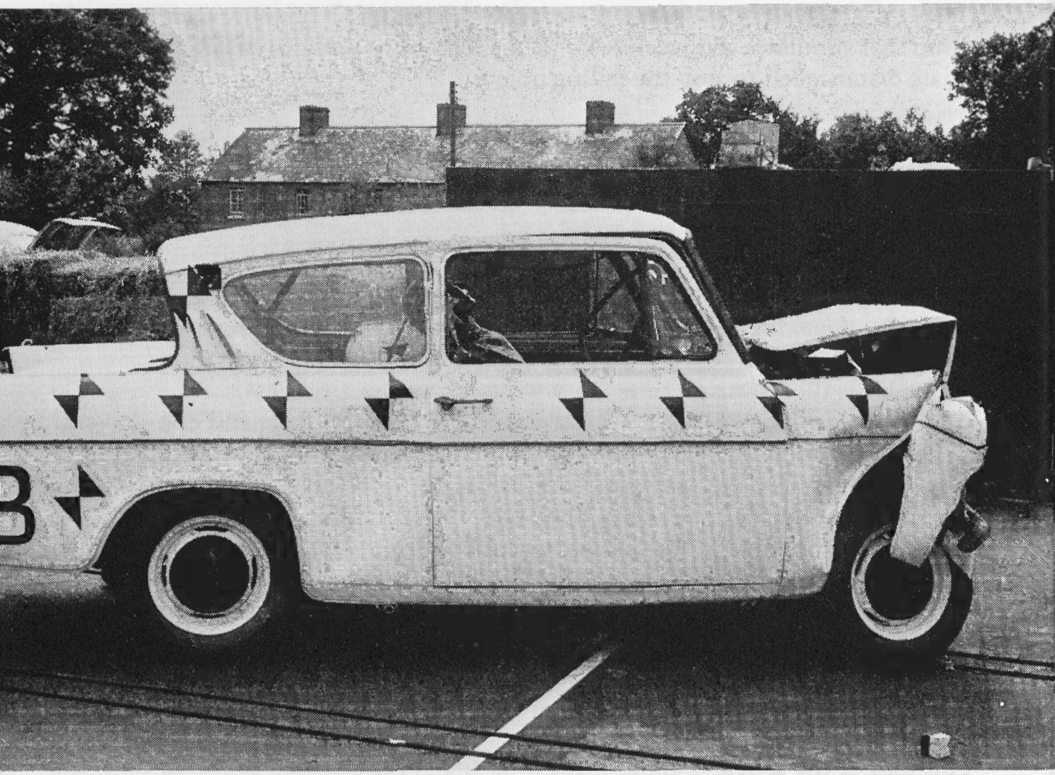


Figure 1.3.2
Deformation of the front of a medium-sized car after a head-on collision.

Frontal collisions are by far the most frequent and serious types that occur on our roads. They can be staged either by the use of two cars travelling in opposite directions or by one car running into a rigid obstacle. The two methods give similar results and for ease of treatment the second alternative will be considered.

About seventy per cent of all serious injuries are due to the car occupant colliding with objects in front of him. Also, about sixty per cent of all fatal injuries occur when the head or neck collides with the windscreen and frame or the projecting features over the windscreen, and most of the other fatalities occur when the body impacts with the steering column or dashboard. These figures indicate where protection is required, and safety devices must be suitably designed to prevent this type of injury.

Safety devices

Safety devices suggested in recent years include the elimination of sharp projections, the provision of crash-proof door-locks, collapsible steering columns, padded dashboards, inflatable air-bags, use of seat belts, energy absorbing plastics on the bumpers, and crushable bonnets. Some of these devices are far more important than others and, added to this, the only satisfactory way to treat the problem is to consider 'car + protection devices' as one unit and to work out the interplay of one safety device on another. For example, figure 1.3.2 shows that most of the crumpling of a standard saloon car involved in a head-on collision occurs in the front of the car and the passenger compartment often escapes serious deformation at speeds under 56 km h^{-1} (35 mile h^{-1}). This is true even in the case of a Mini which has a short bonnet due to cross-mounting of the engine (figure 1.3.3).

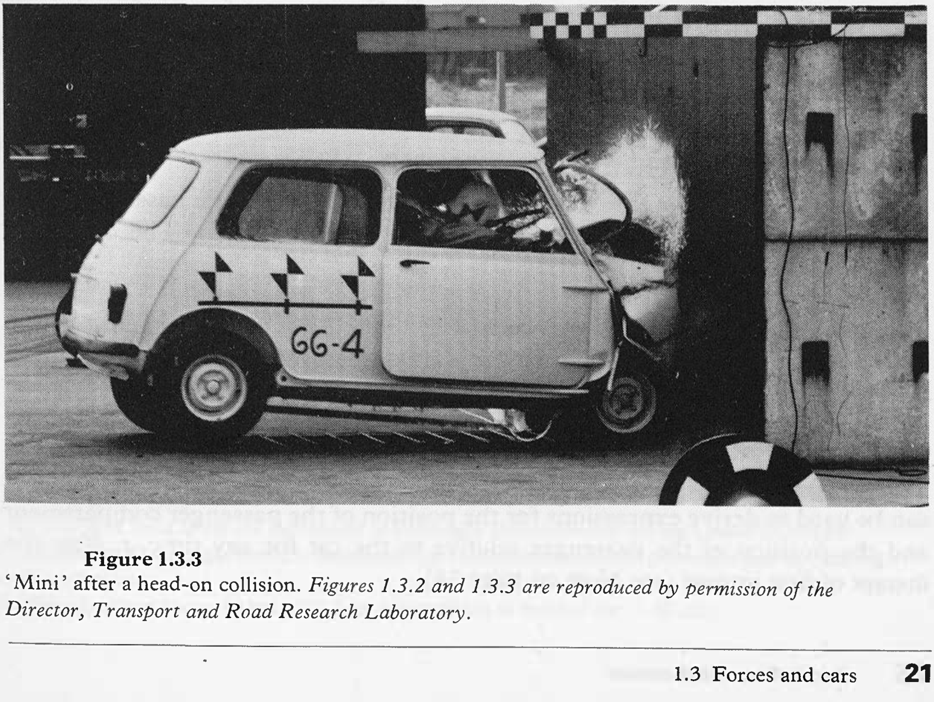


Figure 1.3.3
'Mini' after a head-on collision. *Figures 1.3.2 and 1.3.3 are reproduced by permission of the Director, Transport and Road Research Laboratory.*

The maximum deformation that occurs in the front of a saloon car is about 1 m in a 64 km h^{-1} (40 mile h^{-1}) collision with a rigid obstacle. Because the front of the car crumples, it takes about 0.1 s after the initial impact for the passenger compartment to come to rest (see figure 1.3.4). The symbol a represents the fore and aft deceleration of the passenger compartment and t the time measured from the instant of first impact.

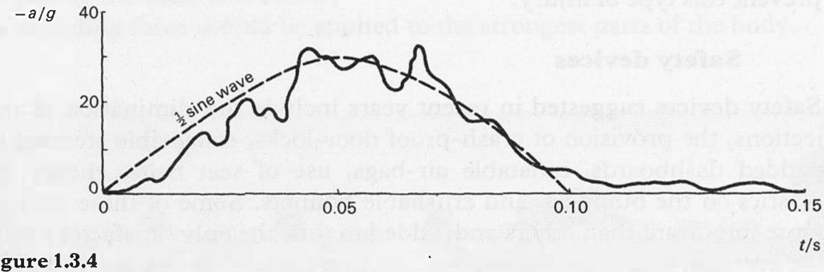


Figure 1.3.4

Deceleration of British 4-door saloon car in collision with a massive concrete barrier. (Mass = 1132 kg, speed = 59.7 km h^{-1}.)

If seat belts are worn, the crushing of the car front is advantageous to the occupant (*provided* the passenger compartment remains intact) because it increases the time available for the dissipation of the passenger's energy. The manner in which the front crushes is also important, as is the deceleration against time curve for the passenger compartment. However, if seat belts are not worn, it makes little difference how the car halts because the passenger compartment is almost stationary by the time the passenger, travelling at the original velocity of the car, hits it.

Analysis of a crash

In a head-on crash, the deceleration against time curve for the passenger compartment often resembles half a sine wave (see figure 1.3.4). The crash can therefore be approximately analysed.

Suppose a car is travelling at 50 km h^{-1} (31 mile h^{-1}) when it collides with a rigid obstacle and the front crumbles so that the passenger compartment comes to rest over a distance of 60 cm. The above approximation, that is,

$$\frac{d^2x}{dt^2} = -A \sin \omega t$$

can be used to derive expressions for the position of the passenger compartment and the position of the passenger relative to the car for any time, t, after the instant of first impact (see Note on page 34).

passenger compartment position $x = \frac{u}{2}\left[t + \frac{\sin \omega t}{\omega}\right]$

position of passenger relative to car $(s - x) = \frac{u}{2}\left[t - \frac{\sin \omega t}{\omega}\right]$

where:

u = initial velocity of car
x = coordinate of a particular point in the passenger compartment
s = coordinate of a particular point of the passenger
t = time measured from the instant of first impact
$-A$ = maximum value of the deceleration of the passenger compartment
ω = $2\pi/T$ where T is the period of oscillation for a whole sine wave

The derivation assumes that the passenger is completely free to move in the compartment; that is, the seat is horizontal and friction between the passenger and the seat is negligible.

Now, $x = 60$ cm when $t = T/2 = \pi/\omega$. Thus, the first expression becomes:

$$\frac{60}{10^2} = \frac{50 \times 10^3}{3600 \times 2}\left[\frac{\pi}{\omega} + \frac{\sin \pi}{\omega}\right]$$

therefore $\omega = 36.4 \text{ rad s}^{-1}$ or $T/2 = 0.086$ s

The values for u and ω can be substituted in these equations to give the graphs shown in figures 1.3.5*a* and *b*.

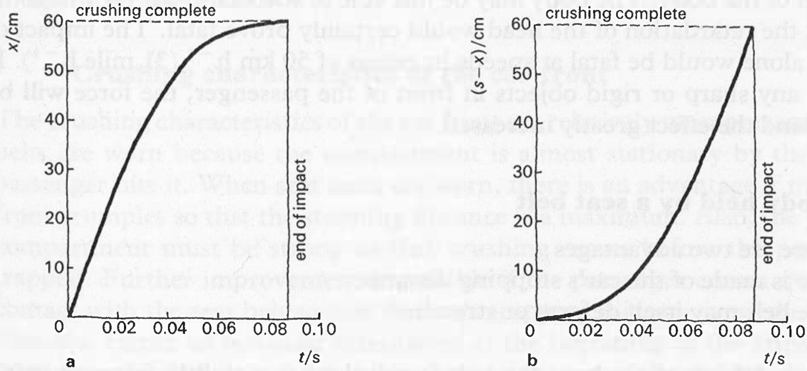


Figure 1.3.5
a Crushing of front of car (or movement of passenger compartment) during a head-on impact.
b Movement of passenger towards dashboard (or relative movement of passenger and car) during a head-on impact.
(Initial velocity of car = 50 km h^{-1}; final deformation of front of car = 60 cm)

Figure 1.3.5*a* shows that most of the crushing takes place during the early part of the impact. In fact, about half of the total crushing occurs in the first quarter of the stopping time. However, in this time, the passenger has only moved about 2 cm (see figure 1.3.5*b*).

Now the basic problem is to disperse the passenger's energy so that the force on him at any instant is minimal. As work done (or energy dissipated), $W = \int F\mathrm{d}x$, it follows that for a given energy dissipation, the force is smallest when the distance is a maximum.

Body free to move within the compartment

Here, the body will move forward with the original velocity of the car even as the bonnet is crumpling. If there is a space of 41 cm in front of the passenger (as in a Rover 2000), figures 1.3.5*a* and *b* show that the car has only 0.5 cm of crushing to complete when the passenger hits the compartment. (Often, the space in front of the head and knees is less than that in front of the body. This complicates the motion of the passenger as pivoting may take place, but for ease of argument, equal distances will be assumed.) When the passenger's body hits the dashboard, the latter will probably crush about 15 cm if it is extremely well designed. The total stopping distance for the body is therefore 15.5 cm and the average retardation $64g$ (from $v^2 = u^2 + 2as$).

The head, however, probably hits padding just above the windscreen and the maximum deformation here will be only about 2.5 cm. The total stopping distance for the head is therefore 3 cm and the retardation of the head about five times that of the body. The body may be just able to withstand $64g$ for this short time, but the retardation of the head would certainly prove fatal. The impact of the body alone would be fatal at speeds in excess of 50 km h^{-1} (31 mile h^{-1}). If there are any sharp or rigid objects in front of the passenger, the force will be localized and the effect greatly increased.

Body held by a seat belt

Here, there are two advantages:

1 use is made of the car's stopping distance;
2 the belt may itself deform or stretch.

If there are 2.5 cm of slack on the belt (equivalent to a tightly fastened belt), figures 1.3.5*a* and *b* show that the slack is taken up in about 0.025 s when the car has crushed by 32 cm. Thus a retarding force is applied to the passenger's body while the car still has 28 cm of crushing to complete. If the belt is designed to stretch 30 cm in these conditions, the total stopping distance is 58 cm. This corresponds to an average retardation of 17*g* with the force applied over a fairly large area of the body. If the seat belt has a lap and shoulder strap, no collision will occur between the passenger and the car compartment and no serious injury will result from the retardation of 17*g* experienced for such a short time.

Collapsible steering column

If the steering wheel is padded and the column is designed to move forward while exerting a force on the driver's body, some advantage may be gained when no seat belt is worn.

Suppose the driver's body is initially 25 cm from the steering wheel. Figures 1.3.5*a* and *b* show that by the time the driver hits the steering wheel, the car has 3 cm of crushing to complete. If the steering column then 'gives' (while exerting a steady force) over a distance of 25 cm, the total stopping distance of the body becomes 28 cm and the retardation 35*g*. This is a marked improvement over the case where the column is rigid, but still no protection is given to the head, the most vulnerable part of the body.

If seat belts are worn, it would be best for the steering column to collapse completely under a relatively low load (about 700 N) so that the extension of the seat belt could continue as before over a maximum distance.

The above calculations show the great importance of close contact between the safety device and the body of the person; in this respect no other device is comparable with a seat belt!

Crushing characteristics of the car front

The crushing characteristics of the car front are relatively unimportant if no seat belts are worn because the compartment is almost stationary by the time the passenger hits it. When seat belts are worn, there is an advantage if most of the front crumples so that the stopping distance is a maximum. Also, the passenger compartment must be strong so that crushing ceases before the passenger is trapped. Further improvement is possible if the passenger comes quickly into contact with the seat belt so that the car's stopping distance is used as much as possible. Hence an *increased* retardation at the beginning of the impact should prove beneficial. Figure 1.3.6 shows three of the many possibilities for retardation of the passenger compartment.

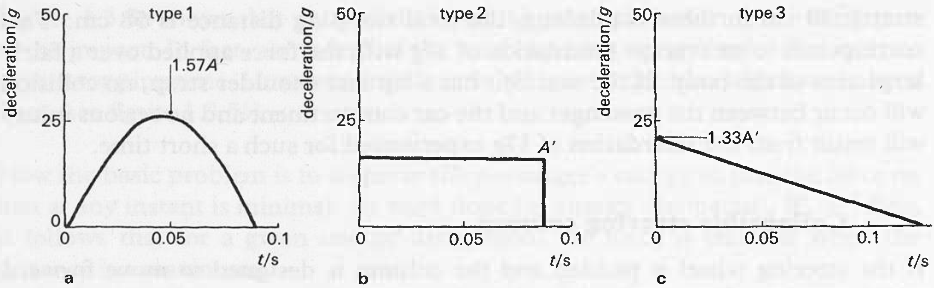


Figure 1.3.6
Different types of deceleration against time curves.
a Normal deceleration curve.
b Constant deceleration.
c Maximum deceleration at start.

The three curves are all adjusted to the same stopping distance of 60 cm from 50 km h^{-1} and have different durations and peak values (the peak values of 1.3.6*a* and 1.3.6*c* are shown compared with the peak value, A', of 1.3.6*b*). The first graph shows the normal deceleration against time curve for the passenger compartment already assumed in the calculations. The second graph shows constant deceleration of the passenger compartment throughout the stopping process. The third graph shows the best arrangement for bringing the passenger into contact with the seat belt in the shortest possible time. It will be noticed that the sooner the peak value occurs, the greater is the total stopping time. This is another advantage for 1.3.6*c*. Whatever the shape of the curve, the area under it must be constant:

$$\text{area under curve} = \int_{v=u}^{v=0} -a\,\mathrm{d}t \qquad \text{where } a \text{ is the deceleration at time } t$$

$$= \int_{v=u}^{v=0} -\frac{\mathrm{d}v}{\mathrm{d}t}\,\mathrm{d}t \qquad \text{where } v \text{ is the velocity at time } t$$

$$= u, \text{ the original velocity of the car}$$

If the initial retardation is increased, the velocity with which the passenger hits the seat belt is slightly increased, but this disadvantage is far outweighed by the additional stopping distance due to the rapid and longer operation of the retarding force.

It is clear that provision of seat belts gives the best protection but alteration of the structure of the car front would help if seat belts are used.

Improvements in safety devices

Although seat belts are already effective in preventing serious injury, there is still room for further improvement.

Three ways in which seat belts might be made more efficient are:

1 improvement in their load against extension characteristics;
2 increase in the restraining force applied to the passenger's body;
3 alterations to the car to enable the belts to work more efficiently.

Load–extension characteristics of seat belts A seat belt should absorb the maximum energy in a given distance while exerting the minimum force on the body. Figure 1.3.7 shows three possible shapes for the curve relating force to extension of the belt.

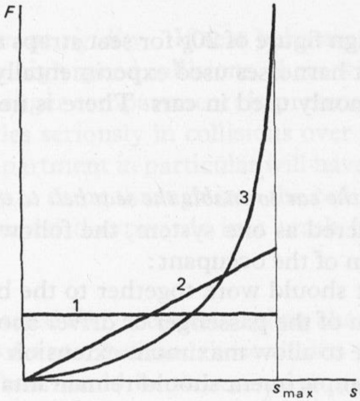


Figure 1.3.7
Force against displacement curves for protection devices.

The area below each curve is given by:

$$\begin{aligned}\text{area} &= \int F\,\mathrm{d}s\\ &= \text{work done by seat belt}\\ &= \text{energy lost by passenger}\end{aligned}$$

The maximum displacements and areas are all equal and it is clear that the maximum force applied to the body is least when that force is constant (curve 1).

The seat belt characteristic is determined by two factors: the properties of the webbing material and the geometry of the harness. The webbing often has the property that extension is proportional to force and this may at first sight appear to give curve 2. However, as the belt stretches, the loops encircling the body tend to straighten out at the sides so that more and more of the belt is stretched along the direction of the applied force. The combined effect of this and the characteristics of the webbing is to produce a curve similar to 3. An obvious improvement would be a force against displacement curve of type 1 and a way of obtaining it will be considered later.

Force applied to passenger's body by the seat belt Present day seat belts are designed so that the maximum force on the passenger's body is about twenty times his weight (a deceleration of $20g$). This figure is an arbitrary one and there is evidence that the human body can withstand larger retardations without injury. Experiments have shown that a retardation of $45g$ could be voluntarily withstood by a person without injury for 0.04 s. However, when it was applied for 1 s, this value dropped to $15g$. For the durations common to car crashes (that is, of the order of 0.1 s), about $40g$ could be tolerated.

Thus the arbitrary design figure of $20g$ for seat straps might well be safely raised to $30g$ although the seat harnesses used experimentally were more elaborate and wider than those commonly used in cars. There is need for further research in this direction.

Modifications to the car to enable the seat belt to work efficiently If the seat belt and car are considered as one system, the following requirements may be stipulated for protection of the occupant:

a the seat and belt should work together to the best advantage;

b the space in front of the passenger or driver should be as large as possible and of the correct shape to allow maximum extension of the seat belt;

c the passenger compartment should remain intact and retain its shape;

d the crushable front of the car should produce the most favourable deceleration conditions for the operation of the seat belt.

a *Seat and harness as one unit* Theoretically, a constant force (curve 1, figure 1.3.7) is most satisfactory for reducing the maximum force on the occupant. A way of obtaining this may be to use an inextensible belt attached to a mechanism which yields when the applied force corresponds to a retardation of $30g$ (say). The mechanism should then exert a constant retarding force on the belt. Alternatively, the inextensible belt may be rigidly attached to the seat and the seat itself move against a constant force once the retardation exceeds $30g$.

b *Space for extension* The space in front of the car occupant should be as large as possible so that the belt may stretch over a large distance. About 43 cm could probably be provided for the passenger's body and head but there is often far less room for the legs. Increasing the height of the space below the instrument panel should help in this respect. Another difficulty is that, in a head-on crash, the limbs tend to move rapidly into a horizontal position and some form of protection is needed here.

The driver has less room in front of his body, but the provision of a collapsible steering column (at low load) may help although it is not entirely satisfactory.

Passengers in the rear seats have a considerable advantage, especially if the front seats slide forward out of the way. A firmly anchored rear seat and an efficient seat belt should be sufficient for protection in this type of collision.

c *Passenger compartment rigidity* If seat belts are improved so that more protection can be given in high speed collisions, there will have to be an increase in the strength of the passenger compartment; in present-day cars, the passenger compartment often buckles seriously in collisions over 56 km h^{-1} (35 m h^{-1}). The front part of the compartment in particular will have to be strengthened and the engine, which is likely to come through into the compartment, should either be mounted low or deflected downwards in a crash by the shape of the redesigned compartment.

d *Front design* As stated above, the crumpling characteristics most favourable to the effectiveness of a seat belt are those shown in figure 1.3.6*c*, but it is difficult to see how the front could be structured to give these characteristics. However, it should be possible to design a front which crumples according to figure 1.3.6*b* and, although inferior to figure 1.3.6*c*, it does bring the passenger into contact with the seat belt more quickly than the usual shape of curve, figure 1.3.6*a*. If the strength of the front bumper and supports were considerably strengthened, there would be a rapid increase in deceleration at the beginning of the impact. If the front then crumpled uniformly until a maximum of one metre of crushing had occurred, the curve should approximate to figure 1.3.6*b*. If the car still had some energy, there would then have to be a rapid increase in strength to save the passenger compartment.

Estimated improvement when all the above devices are used

Suppose:

1 there is a 43 cm gap in front of each part of the passenger's body;
2 there are only 2.5 cm of slack in the seat belt;
3 extension occurs only after the retardation exceeds 30*g*, and
4 the bonnet is capable of crushing by one metre while giving a uniform deceleration to the passenger compartment (as shown in figure 1.3.6*b*).

An analysis of these characteristics in detail shows that, by the time the seat belt slack is taken up, about 63 cm of the car's stopping distance remains. Also, the belt could be designed to stretch 40.5 cm in this limiting condition so that the total stopping distance would become 103.5 cm. If a steady retardation of 30*g* takes place over this distance, the initial velocity of the car is calculated as 88.5 km h^{-1} (55 mile h^{-1}). This represents the maximum velocity for a head-on collision with a rigid obstacle in the above conditions, and is a great improvement over the present maximum of about 56 km h^{-1} (35 mile h^{-1}). However, the above calculation assumes ideal conditions which are hard to realize in practice. For example, the seat belt usually has more than 2.5 cm slack and the body also 'gives' and thereby effectively increases the slackness of the belt. On the other hand, the driver usually brakes sharply before impact and this throws the occupants forward into contact with the seat belt before the collision occurs and therefore reduces the slackness. The car front must also be carefully designed to crumple in the required manner under the enormous forces (of the order of 300 kN) present in these high speed collisions, and much research needs to be undertaken in this direction.

The above demonstrates the importance of inter-relation between the various safety devices, and for the best results the design of the car and the seat belt must be co-ordinated.

Other types of accident

Side impacts Impacts to the sides of vehicles occur in about twelve per cent of accidents which cause injury. The injury is often serious because the strength of the bodywork in that direction is low and because the car occupant is so close to the side of the vehicle. The effects of this type of crash can be seen in figure 1.3.8.

Figure 1.3.8
Result of a side-impact accident.
Reproduced by permission of the Director, Transport and Road Research Laboratory.

There must obviously be a great increase in strength of the bodywork to resist transverse forces on the structure and, in this case, strength must be the main protection because there is no room for crumpling.

Rear impacts These make up about seven per cent of the total and are usually less severe because the relative speeds between colliding vehicles are low in most cases. Also, the back of the seat provides support for the whole body except the head and neck and they can be supported by a head rest, an important safety device in this type of crash.

Passengers in the back seat suffer the same acceleration as the vehicle and therefore the rear end of the car must be suitably designed to produce the lowest accelerating force at any instant. The solution is to give the rear end the same crushing characteristics as the front, that is, those giving constant acceleration. However, the maximum crushing strength must be carefully calculated so that no acceleration greater than $30g$ is produced.

Front seat passengers could be protected to a greater extent by the provision of front seats with some backward movement.

Overturning This occurs in ten to twenty per cent of accidents, mostly in high speed collisions. The chief protections are seat belts and doors which do not open during overturning. Also, the seat should be strongly attached to the floor so that a passenger cannot break free during a crash. A strong passenger compartment is essential.

Protection in head-on collisions when no seat belts are fitted

Suppose there are no sharp projections in front of the occupants, the seats are fixed, and there is some 'give' in the steering column.

As the safety device should be close to the occupant for maximum protection, here it must be the instrument panel that is padded and designed to be as *close* as possible to the occupant. The stopping distance of the car is then utilized to the passenger's benefit during a crash. The front of the compartment should be designed so that the knees and chest hit at the same time. This distributes the forces between the chest and thigh bones, the strongest parts of the body. The space in front of the head should be large so that the head does not hit anything. (In normal car design, the head receives most of the impact during a crash.) As the instrument panel will now extend well into the passenger compartment, the panel should be designed to crumple while exerting a constant retarding force. For the driver, a padded steering wheel and column which buckle under load should provide similar protection. This crumpling, together with the crushing of the car front, make the stopping distance large. Also, an abrupt deceleration at the start of the impact is desirable so that the time duration of the crash is increased.

If a head-on collision is inevitable, two precautions may be taken. First of all, the driver probably brakes hard and this action has the beneficial effect of throwing the occupants forward towards the objects they are likely to hit during the crash. Secondly, it may well be a good plan for the passengers to move forward so that they subsequently come quickly into contact with the objects in front of them, have an increased stopping distance, and a less serious initial collision with the compartment. This movement is in direct opposition to the natural instinct to draw back but the latter action would only increase the chance of serious injury.

Padding of the instrument panel and steering wheel helps because it increases the area of the body over which the total force acts and therefore reduces the pressure on any one part, but padding of the car front would have little effect as the energy absorption would be negligible.

The only type of crash likely to be more serious than a head-on crash with a rigid obstacle is one between a heavy and a light vehicle travelling in opposite directions and colliding head-on. The lighter vehicle suffers far greater decelerations and usually absorbs most of the energy dissipated in the collision because it has a weaker structure. In many cases when two cars collide, one glances off the other and the seriousness of the accident is reduced. Also, stationary objects sometimes buckle when hit by a car and the stopping distance is correspondingly increased. In this respect, further research needs to be undertaken concerning collapsible lamp-posts, etc. However, the seat belt is certainly the best protective mechanism so far devised, and the improvement and integration of seat belts and car design should have a high priority.

Note

If the retardation against time curve for the passenger compartment resembles half a sine curve, then with the symbols defined on page 23:

$$\frac{d^2x}{dt^2} = -A \sin \omega t$$

On integration this gives:

$$\frac{dx}{dt} = \frac{A}{\omega} \cos \omega t + C_1 \text{ (a constant)} \qquad \mathbf{1}$$

Now, $x = 0$, $t = 0$ when $\frac{dx}{dt} = u$

Thus:

$$C_1 = u - \frac{A}{\omega} \qquad \mathbf{2}$$

Also, $\frac{dx}{dt} = 0$ when $t = \frac{\pi}{\omega}$ (half the period of a complete sine wave)

Thus:

$$C_1 = \frac{A}{\omega} \qquad \mathbf{3}$$

From equations 2 and 3:

$$C_1 = \frac{u}{2} = \frac{A}{\omega} \qquad \mathbf{4}$$

Equation 1 becomes:

$$\frac{dx}{dt} = \frac{u}{2} \cos \omega t + \frac{u}{2}$$

On integration this gives:

$$x = \frac{u}{2\omega} \sin \omega t + \frac{u}{2} t + C_2 \text{ (a constant)}$$

When $x = 0$, $t = 0$ and therefore $C_2 = 0$

Thus:

$$x = \frac{u}{2\omega} \sin \omega t + \frac{u}{2} t$$

Therefore the passenger compartment position is:

$$x = \frac{u}{2}\left(t + \frac{\sin \omega t}{\omega}\right)$$

The movement of the passenger is:

$$s = ut$$

Therefore the movement of the passenger relative to the car is:

$$s - x = \frac{u}{2}\left(t - \frac{\sin \omega t}{\omega}\right)$$

1.4 Rocket and jet propulsion

by F. R. McKim

Rocket propulsion

A rocket is a vehicle which is propelled by the thrust obtained when matter contained in the rocket is ejected. The thrust becomes large if the matter is ejected at high velocity, and in the chemical rocket, which is by far the most common, this is achieved by generating high temperature gases by the chemical reaction between suitable fuels and oxidizers. The reaction takes place at high pressure in a combustion chamber, and the gases then expand in a suitably designed nozzle and are ejected from the rocket.

Consider a rocket of instantaneous mass m moving with velocity v in a vacuum (figure 1.4.1*a*). Let its mass change to $(m - \delta m_g)$ in a time interval δt, due to the ejection of gases of mass δm_g with a velocity v_e relative to the rocket (see figure 1.4.1*b*).

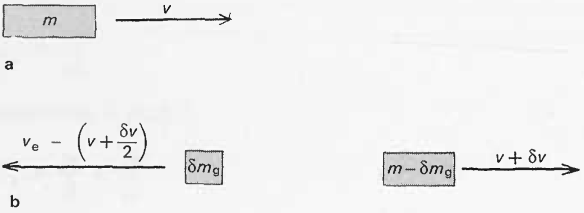


Figure 1.4.1

Then the conservation of momentum equation for the situation at the beginning and end of the time interval δt gives:

$$mv = (m - \delta m_g)(v + \delta v) - \delta m_g \left[v_e - \left(v + \frac{\delta v}{2}\right)\right]$$

Neglecting the second order of small quantities, this reduces to:

$$m\delta v = v_e \delta m_g \qquad 1$$

or in terms of equal and opposite forces (rates of change of momentum):

$$m\frac{dv}{dt} = v_e\frac{dm_g}{dt} = \text{thrust}$$

But δm_g refers to the gas ejected and increases as the mass of the rocket decreases. Hence,

$$\delta m_g = -\delta m_{\text{rocket}}$$

and so we can rewrite equation 1, referring throughout to the rocket mass:

$$m\delta v = -v_e \delta m \qquad \textbf{1a}$$

We suppose that the rocket takes off from the Earth's surface with an initial mass M, and that M_b is its mass and v_b its velocity at burn-out. Then equation 1a may be integrated, giving:

$$\int_0^{v_b} dv = -v_e \int_M^{M_b} \frac{dm}{m}$$

or

$$v_b = v_e \ln \frac{M}{M_b} \qquad \textbf{2}$$

Equation 2 gives the velocity the rocket would achieve in the absence of gravitational effects and air resistance. The practical velocity is always less than this, of course, but the equation does give a first approximation to the behaviour of a rocket. Rocket developments in the past thirty years have been mainly concerned with increasing the values of v_e and the mass ratio M/M_b.

Rocket propellants

For a rocket propellant combination a quantity often quoted is the specific impulse I, defined as the thrust per rate of fuel consumption (by weight). Using the previous notation,

$$I = -\frac{F}{g\,dm/dt}$$

where F is the thrust obtained. But $F = -v_e dm/dt$, so that:

$$I = \frac{v_e}{g} \qquad \textbf{3}$$

So the larger the specific impulse of a particular fuel oxidant combination the higher is the exhaust velocity of the gases and, by equation 2, the higher is the burn-out velocity of the rocket.

The V2 rocket developed by the Germans in the 1940s used ethyl alcohol (ethanol) and liquid oxygen as fuel and oxidant respectively. Liquid oxygen (Lox) rather than gas is often used nowadays. The reason for this is that in order to store large quantities of gas, high pressure storage vessels must be used and these have to have heavy thick walls, whereas the liquid can be stored in thin-walled vacuum storage vessels. A fuel used with Lox in the Atlas rocket (the rocket used for the American Mercury series of manned spaceflights) was a kerosene-type hydrocarbon, RP1. More recently, liquid hydrogen (LH_2) and Lox rockets have been developed. The specific impulses of these fuels, burning under comparable conditions, are shown in the table below.

Fuel and oxidant	**Specific impulse (s) (sea level)**
Ethanol and Lox	256
RP1 and Lox	266
LH_2 and Lox	364

Rocket engines

Liquid propellant engines The heart of the liquid propellant engine is the combustion chamber (figure 1.4.2). The fuel and oxidizer are introduced to this through an injector, designed so as to provide a thorough mixing of the two constituents. Burning takes place in the chamber and the exhaust gases expand through the nozzle.

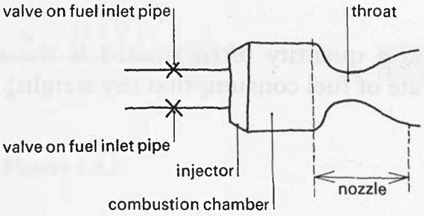


Figure 1.4.2
The main parts of a liquid propellant rocket engine.

The propellants are introduced in one of two ways. They can be pumped in by gas pressure in the tanks, the gas pressure being controlled by valves in the gas inlet pipes, or they can be pumped from the tanks by propellant pumps. The pumps are usually operated by a turbine driven by its own supply of hot gases produced from the burning of rocket fuel. Typically, around two per cent of the propellant flow is to the turbopump (and the other ninety-eight per cent to the main combustion chamber) when the rocket is firing.

The design of the nozzle critically affects the pressure in the combustion chamber, the thrust produced by the gases, and the final velocity produced by the rocket. Its function is to ensure that, with maximum efficiency, the random thermal motion of the hot gases in the combustion chamber is converted to directed motion of the exhaust gases from the end of the nozzle. The first part of the nozzle is a smooth taper to the narrowest part, the throat. From the throat there is a steadily expanding cone-shaped region, the length of this region being sufficient to ensure that the exhaust gas pressure at the end is reasonably low. The cone angle for optimum effect is also important; if it is too small, the cone must be unnecessarily long in order to produce the final low pressure; if it is too large, the loss of momentum due to radial gas flow becomes significant, and the expanding gas separates from the cone walls, producing vortices with a corresponding reduction in thrust. A cone angle of between 30° and 40° is usual, with the inside cone surface slightly concave, as figure 1.4.2 indicates.

The nozzle, particularly in the throat area, and the combustion chamber are subjected to very severe temperatures and must be made of material able to withstand these without melting. The process of cooling is usually helped by the incoming fuel, which circulates past the walls before going to the combustion chamber.

All the fuels mentioned so far require initial ignition before the chemical reaction will take place. Some propellants, for example hydrogen peroxide and hydrazine, are self-igniting (termed *hypergolic*) when brought into contact. The injector for a hypergolic fuel rocket must be different from that for a non-hypergolic one, for the mixing needs to take place in the combustion chamber. Some hypergolic propellants have a high specific impulse, of 300 seconds or better, and unlike the liquefied gases may be relatively easily stored at ordinary temperatures. The lunar landing and take-off rockets for the American Apollo project, which need to be extremely reliable, must be able to switched on and off, and must function days after leaving the Earth, use hypergolic propellants.

Solid propellant engines The solid propellant rocket is simpler than the liquid one, needing no special fuel storage tanks, pumps, and injector. The fuel and oxidizer, intimately mixed together and termed the grain, are packed in the combustion chamber itself. As in the liquid fuel rocket, the chamber ends in a specially shaped nozzle (figure 1.4.3). The chamber is usually cylindrical, with elliptical or spherical ends.

Specific impulses are lower than for liquid propellants, being typically 200 to 250 s, but the higher propellant densities partly compensate for this.

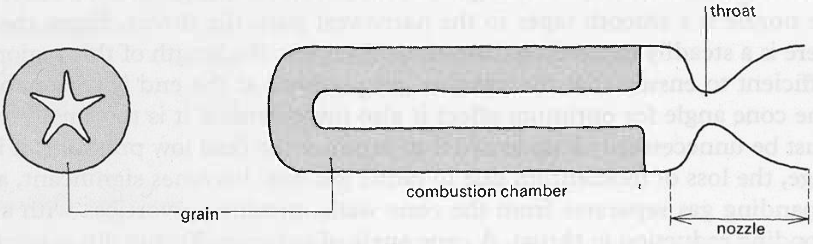


Figure 1.4.3
Cross-section and plan of a simple, solid propellant rocket.

Burning is initiated by some kind of firing charge and then takes place at the exposed surface of the grain. The simplest arrangement would be for the chamber to be packed solid with propellant, burning taking place from the end. This is not used, however, because of the slow burning rate achieved, although it has the advantage that the burning rate remains constant throughout firing. There is a further disadvantage, in that the walls of the chamber would become directly exposed to the high temperature gases. Insulation would therefore have to be used, which as well as adding weight would reduce the volume available for propellant.

For both these reasons, the grain is usually packed into the combustion chamber leaving a specially shaped fluted hole down its axis. The hole shape is chosen to make the exposed area, and hence the burning rate, as constant as possible throughout the burning time.

The nozzle shape is much the same as in the liquid fuel system. The throat of the nozzle, which cannot be cooled by incoming liquid, must be made of a material which can withstand the very high temperatures of up to 3800 K, for example a ceramic or tungsten alloy. In the Polaris missile, which has a solid propellant rocket, a form of transpiration cooling is used. The throat is made of porous tungsten which has about twenty per cent by volume of infiltrated silver. At high temperature the silver melts and vaporizes and the latent heat which is absorbed prevents overheating.

The solid propellant rocket has one great disadvantage in that its thrust cannot be controlled, once burning has begun. The Polaris second stage (solid propellant) does have a thrust termination device; head end ports are opened on command, dropping the internal pressure to a value too low to support combustion. On the other hand, having no pumps or valves, the solid propellant rocket is obviously much simpler and therefore more reliable than a liquid fuel rocket. Furthermore, in contrast to rockets with LH_2 and/or Lox tanks which must be 'topped up' at the last moment, a solid propellant rocket is always fully ready. For obvious reasons, therefore, the American nuclear deterrent missiles, Polaris, Poseidon, and Minuteman, use solid propellant rockets. For the smaller, less spectacular tasks for which rockets are used, for example in giving increased thrust to a plane at take-off from an aircraft carrier or in firing a line during ship rescue work, solid propellant rockets are almost invariably used.

Multi-stage rockets

A method of improving rocket performance is to use the multi-stage method, rather than one single large rocket. With a given fuel, the effective mass ratio is then much increased.

The first stage rocket carries up the second and subsequent stages. When its fuel has been used up it is separated from the rest and the second stage then fires and accelerates further until the second stage fuel has been used up, when separation occurs again, and so on. The advantage of this method over one large single rocket is clear. The only part to be accelerated to the final speed is the final stage plus the payload of spacecraft or warhead. The first stage is only accelerated to the first separation speed, the second to the second speed, and so on.

Illustrative of the advantage of this method are the following figures. They refer to a task of placing a payload on the Moon with a particular chemical fuel, and give the payload, m, that can be delivered for different multi-stage rockets, each system having the same total initial take-off weight. Let the number of stages be denoted by n.

n	1	2	3	4
m/kg	0	45	450	1350

No payload can be landed on the Moon with a single-stage chemical rocket. The rocket would fall back to the Earth's surface, short of its target. This may be seen in the following way. To escape from the Earth a velocity of 7800 m s^{-1} is required. From equations 2 and 3 it follows that to reach this velocity (assuming the rocket starts from rest), the mass ratio M/M_b must be such that

$$\ln\left(\frac{M}{M_b}\right) = \frac{v_b}{v_e} = \frac{v_b}{Ig}$$

Using the known best specific impulse, 364 s, gives

$$\ln\left(\frac{M}{M_b}\right) = \frac{7800}{364 \times 98} = 2.19$$

Therefore:

$$\frac{M}{M_b} = 8.94 = \frac{M}{M - M_f} \quad \text{where } M_f = \text{fuel mass}$$

$$\frac{M_f}{M} = \frac{7.94}{8.94} = 89\%$$

Since for any motor, the best ratio of fuel to total motor mass is seventy to eighty per cent, the required escape velocity rise cannot be realized. Even if the motor were more efficient in terms of mass (up to eighty-nine per cent) there would be no payload of men, instruments, retro motors, etc. Hence to put any payload on the Moon – using present fuels – a multi-stage rocket must be used.

The American Apollo spacecraft, which have now (on several occasions) transported three men to the Moon and back, were launched by giant three-stage Saturn V rockets. Details of the fuel and total thrust, F, produced in each stage are as follows:

	Fuel, oxidant	**F/MN**
1st stage	RP, Lox	33
2nd stage	LH_2, Lox	6.7
3rd stage	LH_2, Lox	0.9

The total loaded mass of the whole rocket and spacecraft is over 3×10^6 kg of which more than 2×10^6 kg is the first stage rocket and its propellant. The spacecraft has a mass of some 4.5×10^4 kg.

Jet propulsion

In a jet engine, air is taken in from the surroundings and compressed before being supplied to the combustion chamber. This requires appreciable power, whereas no power is required for liquid compression in a rocket.

As a means of comparison of the jet and rocket engines, consider a combustion chamber pressure, p_c, and a surrounding pressure, p_0, and imagine a rocket engine with a mass flow m of liquid oxygen per second. If we assume that the liquid is incompressible, no work is done in bringing it to pressure p_0 in the tank, and its density ρ is a constant independent of pressure. Then the work done per second, W_1, in forcing the liquid from the fuel tank through the rocket engine, is given by:

$$W_1 = \frac{(p_c - p_0)m}{\rho} \qquad 4$$

In the jet engine, the gas must first be compressed. If we let the gas flow be m, for the moment, then the work done in compression, W_{comp}, is given by:

$$W_{\text{comp}} = -m\int_{p_0}^{p_c} p\,\mathrm{d}\left(\frac{1}{\rho}\right)$$

On ejection from the compressor through the combustion chamber the work done, W_{ej}, is given by:

$$W_{\text{ej}} = \frac{p_c m}{\rho_c} - \frac{p_0 m}{\rho_0}$$

The total work done per second, W_g, is therefore given by:

$$W_g = W_{\text{comp}} + W_{\text{ej}} = -m\int_{p_0}^{p_c} p\,\mathrm{d}\left(\frac{1}{\rho}\right) + m\left(\frac{p_c}{\rho_c} - \frac{p_0}{\rho_0}\right)$$

$$= m\int_{p_0}^{p_c} \frac{\mathrm{d}p}{\rho} \qquad 5$$

Now the gas compression takes place so quickly as to be adiabatic, that is, without any heat interchange with its surroundings. For such a change it may be shown that $p\rho^{-\gamma}$ is constant, where γ is the ratio of the specific heat of the gas at constant pressure to its specific heat at constant volume. Consequently the variation of ρ with p is known, and equation 5 may be integrated, giving

$$W_g = m\left(\frac{\gamma}{\gamma - 1}\right)\left(\frac{p_c}{\rho_c} - \frac{p_0}{\rho_0}\right) \qquad 6$$

Now for air $\gamma/(\gamma-1) = 3.5$. At $p_0 = 1$ atm, ρ_0 is about $\rho/850$, where ρ is the density of liquid oxygen. At $p_c = 20$ atm, ρ_c is about $\rho/100$. Therefore, for the rocket we have, from equation 4,

$$W_1 = \frac{(20-1)m}{\rho} = \frac{19m}{\rho}$$

whereas for the jet we have, from equation 6,

$$W_g = \frac{m \times 3.5\,(2000-850)}{\rho} \approx \frac{4000m}{\rho}$$

From this it can be seen that air fuel requires about 200 times the pumping power of liquid fuel. Furthermore, air is only about one-fifth oxygen so that a greater mass flow for air is required, making the power consumption even greater. In fact, compression of air is the most critical aspect of the function of a jet engine, and the design of the combustion chamber is comparatively less important. The combustion pressure is decided, for example, by considerations of the efficiency and simplicity with which air compression at this pressure may be achieved. Some sixty per cent of the developed power may be used to drive the air compressor at sea level.

The most usual form of compressor is a turbine compressor, driven by a turbine which takes some energy from the hot gases emerging from the combustion chamber. In early engines the compressor was a centrifugal type, but it is usually nowadays an axial flow turbine compressor because, although more fragile, it has a greater air handling capacity and greater efficiency. This type of engine unit, the turbo-jet, is frequently used to power military aircraft, particularly if they are supersonic.

In the turbo-prop unit, the gases escaping from the combustion chamber drive a propeller, via a turbine, as well as the gas compressor. This unit has a lower fuel consumption than the turbo-jet, higher take-off thrust, and lower engine noise. Its disadvantages are its greater weight and complexity, and the low speed of its peak efficiency. It is best suited to planes with speeds below about 750 km h^{-1} (about 450 m.p.h.) because of the limitations of the propeller.

The turbo-fan unit has an additional by-pass for air, which is pumped through the engine without mixing with fuel. It is ejected at increased speed by a fan which is driven by a turbine in the jet stream. Somewhere between one half and twice as much air goes through the by-pass as goes through the combustion chamber. The turbo-fan is superior in thrust and propulsion efficiency to the turbo-jet at subsonic speeds, because it produces a given thrust by moving more air at a lower speed. In comparison with the turbo-prop, the turbo-fan has a higher speed capability and less complexity. Subsonic jets are therefore tending to become turbo-fan units.

The nozzle of the turbo-jet, like that of the rocket, must convert the random thermal energy of the hot gases to directed kinetic energy. An important difference, however, is that whereas in the rocket combustion chamber the pressure may be more than twenty times atmospheric, in the turbo-jet just behind the turbine it is usually only about twice atmospheric. In consequence, the turbo-jet nozzle does not require the pronounced convergent–divergent shape which characterizes the rocket nozzle. Often, in fact, the jet nozzle converges throughout its length so that the exit has the minimum cross-section.

Simple theory shows that divergence in the nozzle is necessary whenever the ratio of initial to final pressures is sufficiently large to result in supersonic speeds on expansion, that is, in the case of air if the pressure ratio is greater than about 1.9. For subsonic speeds a jet nozzle of fixed dimensions is used. For supersonic flight a variable geometry intake is used to suit different Mach numbers. The throat and exit dimensions of the nozzle are also altered to suit the chamber pressure, which varies also with Mach number, so that a near-constant expansion ratio is obtained.

In comparing rocket propulsion with turbo-jet propulsion, the first difference is that for a given thrust, the gas jet emerging from the rocket is much narrower than that from the turbo-jet, since the pressures are much higher in the rocket. Consequently the rocket engine has a smaller cross-sectional area and fewer mechanical moving parts, and is therefore lighter than the equivalent turbo-jet engine. Roughly, to provide 10 N of thrust at ground level or 2 N at 15 km altitude, the turbo-jet weighs about 0.3 kg. A rocket unit will provide about 10 N of thrust at all altitudes for about 0.01 kg weight.

In contrast to the turbo-jet, however, the rocket must carry its own oxygen as well as its fuel, and this is its chief disadvantage at low altitudes (where oxygen is plentiful) and its overwhelming advantage at very high altitudes (when the air density has fallen almost to zero). The turbo-jet has a practical ceiling of less than 30 km; the rocket has no limit.

1.5 Some biological applications of the centrifuge

by H. Gammage

When a suspension of particles in a fluid is rotated in a container at high angular velocity, the denser particles tend to move towards the periphery of the mixture and, in due course, to meet and pile up against the walls of the container. A device which uses this effect is called a *centrifuge*, and separation techniques dependent upon the effect have been devised and have many applications in the biological sciences. Indeed, centrifuges were first used at the beginning of the century for the fractionation of colloidal suspensions and, like the school laboratory centrifuge of today, these early machines possessed limited angular velocity and fluctuations in speed were so great that they were useless for precise quantitative separations. However, by about 1950, advances in design and performance made it possible to produce centrifuges with rotational speeds up to about 150000 rad s^{-1} which could be held constant and could be estimated with a high degree of precision. In this article we are concerned with two types of centrifuge: the preparative centrifuge used in separation techniques, and the ultra-centrifuge, used for determination of relative molecular masses of large molecules, such as proteins.

Preparative centrifuge

Separation techniques dependent upon using a centrifuge are of great importance in the study of the composition and structure of cells. The general procedure follows a well-defined sequence. First, the outer membranes of the cell need to be disrupted by some mechanical means, such as ultrasonic vibrations or merely through use of a pestle and mortar. To keep all of the organelles of the cell intact, it is usual to suspend the cells in an isotonic medium, such as 0.25 M sucrose. The suspension is then subjected to differential centrifugation at low temperature. In the case of a broken cell preparation of liver, four fractions are usually selected, as shown in the table opposite.

	Fraction	**Maximum centripetal acceleration used**	**Time required for separation**
1	cellular debris nuclei, membranes	1000*g*–6000*g*	600 s
2	mitochondria lysosomes	10 000*g*–15 000*g*	1800 s
3	microsomes ribosomes	100 000*g*	3600 s
4	supernatant fluid		

Note. $g = 9.8\ \mathrm{N\ kg^{-1}}$ or $9.8\ \mathrm{m\ s^{-2}}$. *g* is a convenient unit here because 1*g* is equivalent to the gravitational force on unit mass at rest on the Earth's surface. An acceleration of 1000*g* indicates that the force experienced by the body is increased a thousand-fold compared with normal conditions.

The arbitrary nature of the fractionation will be appreciated from the heterogeneity of the fractions shown above. Thus, it has been found that all the DNA appears in the fraction containing nuclei; the mitochondria fraction contains all of the enzymes required for cell respiration, the oxidation of amino acids (derived from proteins) and of fatty acids; and the microsomal fraction contains enzymes needed in protein synthesis, steroid reductasis, phosphatasis, etc. Under the conditions used, the supernatant fluid retains enzymes needed in the glycolytic system; the hexose monophosphate pathway, glycogen synthesis, fatty acid synthesis, and others. Samples of purified enzymes can be derived from each of these fractions by such techniques as fractional precipitation or chromatography. However, one must not ignore the key characteristic of separations effected by centrifugation, namely that little or no damage is done to sub-cellular particles during the process. Indeed, living cells from the heart of a chick embryo have been subjected to 400 000*g* for as long as 30 minutes without apparent damage.

Interest in DNA and RNA, as molecular species which carry and control the operation of the genetic code, has led to the development of density gradient centrifugation. In essentials, this technique requires the use of a suspension of material in a concentrated solution of either caesium chloride or sucrose. This is centrifuged at high angular speed for a known period of time. During this treatment, an equilibrium is established between sedimentation and diffusion processes and a stable concentration gradient is established along the direction of the centripetal force. In the case of DNA in caesium chloride solution, these large molecules come to rest in a zone in the solution when the density of the solution corresponds to that of the species. The location of these molecules has to be established by a photographic technique. With density gradient centrifugation techniques, density differences of about 1 kg m^{-3} can be detected.

The fact that DNA exists as a double helix can be demonstrated using this same technique. Thus, DNA can be isolated from a culture of *E. coli* in which all division has been allowed to occur in a culture containing nitrogen-15. The DNA separates out at a density of 1717 kg m^{-3}. After isolation and subsequent heating in an appropriate medium for 10 minutes, the sample is again subjected to density gradient centrifugation. Two bands of DNA appear in the caesium chloride solution: one at 1740 kg m^{-3} and one at 1724 kg m^{-3}. Tests demonstrate that only the latter sample contains nitrogen-15. By cooling the preparation rapidly the two bands can be isolated. However, if the preparation is allowed to cool slowly, the two bands combine and appear as a single band again at 1717 kg m^{-3}. Clearly, this sequence of tests suggests that heating the preparation causes the DNA to separate into two strands, one of which is labelled with nitrogen-15. Further, on cooling, the strands must recombine to form DNA.

The separation of RNA can be effected in a sucrose solution in much the same way. Three fractions are obtained: a soluble cytoplasmic RNA and two types of ribosomal RNA. This development led to the construction of so-called zonal centrifuges which possess greater capacity, and can be subjected to greater stress, than other forms of preparation centrifuge. Zonal centrifuges have been used to collect larger quantities of sub-cellular particles, viruses, and macroglobulins.

The ultra-centrifuge

Many substances of biological significance are macromolecules with a relative molecular mass of about 10^4 (for example, albumen has a molecular mass of 44000 and haemoglobin has a molecular mass of 69000). Conventional procedures for determining relative molecular mass (like the depression of freezing point or the elevation of boiling point of a solvent) are inappropriate for such species. (Why?) A variety of experimental methods are feasible for such determinations and depend on properties such as the scattering of light or the measurement of viscosity. The ultra-centrifuge forms the basis of one set of procedures. It was introduced by T. Svedberg in 1925. In essentials, a solution of macromolecules attains an equilibrium condition by achieving a balance between sedimentation and diffusion in a gravitational field. A diagrammatic form of the apparatus is shown in figure 1.5.1. Using such an apparatus, it is possible to determine the solute concentration as a function of distance from the centre of rotation by one of several possible physical methods of analysis (cf. figure 1.5.2), and relative molecular mass is calculated by simply comparing the concentration of solute in solution at two distances from the centre of rotation.

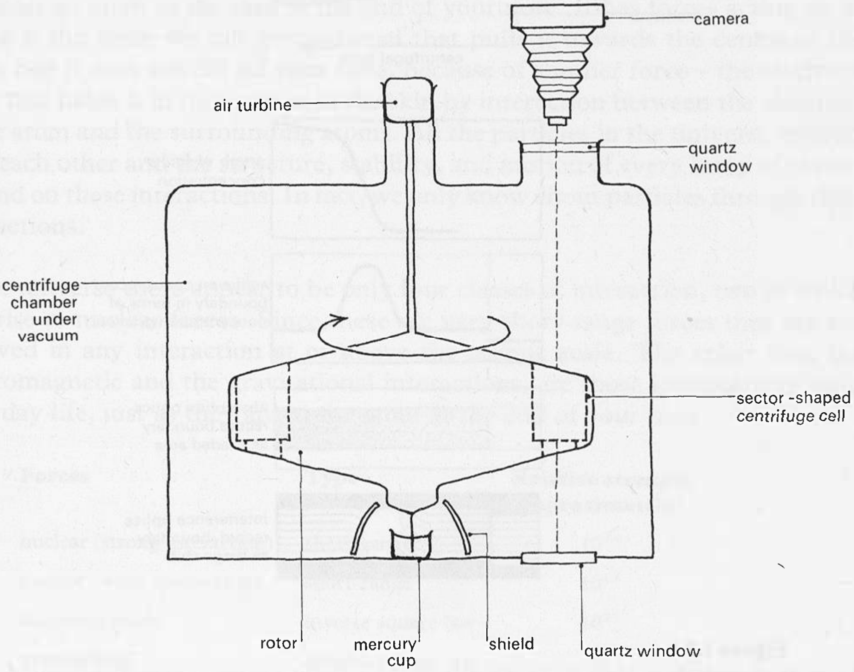


Figure 1.5.1
Vertical cross-sectional view through vacuum chamber of an air-driven ultracentrifuge.
After Svedberg, T. and Pedersen, K. O. (1940) The ultracentrifuge. *The Clarendon Press, Oxford.*

Svedberg's researches demonstrated that haemoglobin and other blood proteins had specific relative molecular masses, an experimental result which indicated that proteins were true molecules and *not* composite particles of varying mass. The technique led to the development of machines capable of asserting a centripetal acceleration of 90 000*g* by 1934. Other workers had meanwhile developed cone-shaped rotors for the centrifuge, and such machines can be operated at extremely low pressure (as 'vacuum centrifuges', in fact). It was through such developments in techniques that the yellow fever virus was shown to possess a relative molecular mass of 2 930 000, and the influenza virus a molecular mass of 650 000.

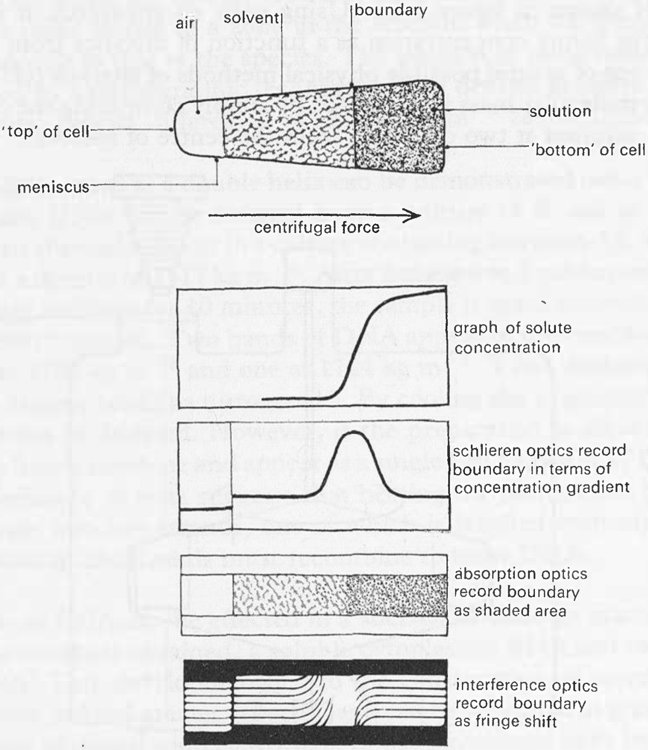


Figure 1.5.2
Summary of methods used for study of high polymer molecules in the analytical ultracentrifuge. *After Moore, W. J. (1962)* Physical chemistry *(3rd edition). Longman. By permission of Beckman Instruments, Inc., Spinco Division.*

1.6 Four kinds of interaction

by C. D. Curling

Consider an atom in the skin at the end of your nose. It has forces acting on it. There is the force we call gravitational that pulls it towards the centre of the earth. But it does not fall off your nose, because of another force – the electrical force that holds it in its position in the skin by interaction between the electrons of the atom and the surrounding atoms. All the particles in the universe interact with each other and the structure, stability, and motion of every lump of matter depend on these interactions. In fact, we only know about particles through their interactions.

In the universe there appear to be only four classes of interaction, two of which give rise to nuclear forces. Since these are very short-range forces they are not involved in any interaction at or above the atomic scale. The other two, the electromagnetic and the gravitational interactions, are those important to us in everyday life, just as they are to the atom at the end of your nose.

Forces	**Type**	**Relative strength (approximately)**
nuclear (strong interaction)	short-range	10^{39}
nuclear (weak interaction)	short-range	10^{25}
electromagnetic	inverse square law	10^{37}
gravitational	inverse square law	1

Not only do the nuclear forces have negligible effect beyond about 10^{-15} m, but there is a general rule that the weaker the interaction the larger the number of particles it embraces. Thus the strong nuclear interaction, acting as a kind of nuclear glue, only operates between nuclear particles, whereas the gravitational force acts on all particles, including photons.

There are certain restrictions on the things that may happen through these interactions. Only those events occur where it is possible for some important quantities to be conserved. Twelve such conservation principles are known (cf. Ford, K. W. (1963) *The world of elementary particles.* Blaisdell) and the most important of these concern the conservation of energy, of momentum, and of charge. It is no exaggeration to say that everything that goes on in the universe does so in accordance with these conservation principles and through these four types of interaction.

And yet these conservation principles are probably all linked to the large-scale properties of the universe. Because the universe is uniform, so that the place and the time at which an experiment is performed do not in themselves affect the result of that experiment, it can be shown that momentum and energy are conserved.

For the strong nuclear interaction all these conservation principles are obeyed, but for other interactions some are violated. Thus, at least three conservation principles do not hold for weak nuclear interactions.

Let us now consider the four interactions in more detail and then seek answers to two basic problems:

1 Why do the long-range forces fall off as the inverse square of the distance?

2 Why is the gravitational force (and it alone) proportional to the mass of the particle on which it is acting?

Nuclear forces

The strong nuclear interaction is responsible for the stability of nuclear matter, and forms the source of the energy of nuclear fission exhibited in nuclear reactors and bombs and in the nuclear reactions inside stars. Heavy nuclei become unstable when the electromagnetic interaction (electrostatic repulsion) of the protons present overcomes the nuclear glue, and the nucleus may either undergo fission or emit alpha-particles. If this electromagnetic interaction were weaker than it is, there would be many more elements in the periodic table heavier than uranium and these would be stable. This is because in the nucleus, electromagnetic forces and strong nuclear interactions act in opposition. In a star, it is the gravitational and the strong nuclear interactions that are in opposition and the cosmic glue of gravitation makes the nucleons come together to form a stellar furnace. Nuclear fusion reactions join nuclear particles together with this nuclear glue – converting hydrogen to helium, helium to carbon, oxygen, etc. in the more massive stars, and eventually producing the whole range of elements up to iron (atomic number 26).

The weak nuclear interaction is displayed in the process of radioactivity and limits the number of stable isotopes of any element.

Both forms of nuclear interaction are at present imperfectly understood but because they are short-range they can often be neglected. The strong interaction is fundamental to an understanding of both stars and nuclei; the weak interaction as displayed in radioactivity is responsible for the instability of free neutrons, provides us with a clock to measure large intervals of time, and energizes the eruption of volcanoes and geysers.

Gravitational and electromagnetic forces

If we were in a freely moving spaceship so that the accelerations due to gravitational forces acted equally on the spaceship and all the matter within it, nearly everything going on in the spaceship would be the result of electromagnetic forces. But on Earth we are always conscious of the force pulling us downwards, and although the gravitational interaction is very weak*, the Earth is massive enough to make its force of attraction a severe limitation to all but the most energetic attempts to leave the Earth's surface.

It is perhaps hard to believe that chemical reactions, viscous forces, elastic forces, and surface tension are all examples of the electromagnetic interaction. But in all these effects, the forces between atoms and between molecules turn out to be forces produced by charge distributions within the atoms. Even barking a shin is really the consequence of the electrical forces between the atoms of your leg and the object you encountered. Electrons are so all-pervasive in matter that it is quite easy to rub some off insulators and produce the varied phenomena of electrostatics.

Magnetic forces, such as those that exist between two parallel wires carrying a current, are lumped with electrical forces as part of the electromagnetic interaction. In fact, it is the relativistic effect of the very small drift-velocity† of the current-carrying electrons that produces the observed repulsion or attraction. (See Osborne, J. M. (1963) 'The origin of the magnetic field', *Sch. Sci. Rev* **XLV** No. 155, 54; and letter by Long, R. J. (1964) 'The origin of the magnetic field', *Sch. Sci. Rev.* **XLVI** No. 158, 239.)

* Compare the gravitational and electrostatic attractions between two grains of sand one metre apart if one in every ten molecules of one grain has lost one electron to the other grain.

† How small is this velocity? Calculate it for a 1 cm^2 cross-section copper rod carrying 10^3 amperes. Assume each copper atom contributes one electron to the job of carrying the current.

An electromagnetic field between any two charged particles can be regarded as a continual exchange of photons, and since this interaction is still present even at very large distances, any charged particle is continually in interaction with the rest of the Universe. If we accelerate a single electron, for example by giving it a twitch, we kink the field-lines associated with it, and this kink travels away from the electron at the velocity of light. This disturbance is called an electromagnetic wave and the effects associated with it enable us to regard the process in terms of photons emitted by the accelerated charge. A single charged particle sends out and reabsorbs photons somewhat like someone hitting away and then catching a ball attached to an elastic string. If two people played this game and one caught the ball on the other's string, the string would produce an attraction. Now there are no strings on photons, and of course, both attraction and repulsion are possible, so our analogy is not a good one. But this idea that forces arise by the exchange of momentum through particle-carriers has been vital to the theory underlying all four of these interactions.

Two problems about forces

There are two features about the long-range interactions that should cause anyone who is scientifically curious to start thinking. Why do both the electromagnetic and gravitational interactions fall off as the inverse square of the distance, and why is it only the gravitational force that is proportional to the mass on which it is acting? It is, of course, this proportionality that enables us to refer to the *acceleration* due to gravity, since all masses are accelerated similarly when subjected to the same gravitational interaction. Following the work of Galileo and Newton, Einstein managed to formulate the close relationship between inertial and gravitational forces which leads to the principle of equivalence of these two forces and hence the equality of 'gravitational' and 'inertial' masses. Indeed, the various inertial forces such as the centrifugal, Coriolis, and the *g*-force experienced in a rocket can all be lumped together with the gravitational interaction, and gravity regarded as the static part of the inertial interaction in the same way as the coulombic force between charges is regarded as the static part of the electromagnetic interaction.

	Electromagnetic interaction	**Inertial interaction**
static	coulomb field	gravitational field
velocity-dependent	magnetic field	?
acceleration-dependent	electromagnetic radiation	inertial effects

There are, however, complicating factors, and Sciama, D. W. (1972) *The physical foundations of general relativity*, Heinemann, deals with these and shows how the inertial interaction – of which gravitation is a part – is a consequence of the structure of the Universe and results from the influence of the very distant parts of the Universe on local phenomena.

We now return to consider our first problem in order to show that long-range forces in a three-dimensional Universe must follow an inverse square relationship.

Consider the force to be produced by the transfer of momentum from source-particle to receiver by a flux of particles (see figure 1.6.1).

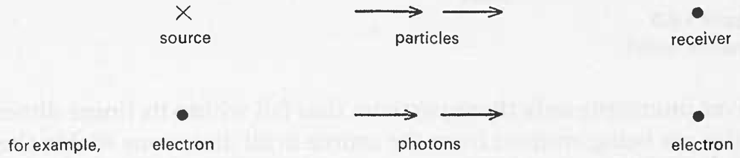


Figure 1.6.1

The number of spatial dimensions in the world in which this event occurs will affect how the force varies with distance. To avoid complication we will assume the flux to occur only from source to receiver.

In a one-dimensional world the entire system is confined to a straight line and particles can only be emitted towards or away from the receiver (figure 1.6.2).

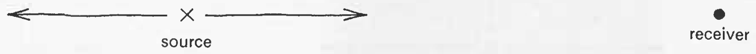


Figure 1.6.2
One-dimensional world.

Hence the receiver intercepts half of the particles emitted, and the force measured as a transfer of momentum is *independent* of distance between source and receiver.

In a two-dimensional world the system is confined to a flat sheet (figure 1.6.3).

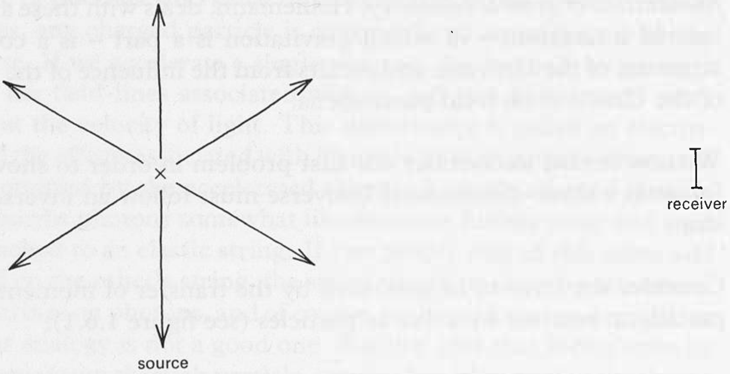


Figure 1.6.3
Two-dimensional world.

The receiver intercepts only those particles that fall within its linear dimensions, and particles are being emitted from the source in all directions within the plane. To collect them all we would need to surround the source by a circle of circumference $2\pi r$. Hence the receiver only intercepts a number proportional to r^{-1}. The force falls off as an inverse first power law.

In a three-dimensional world, the system occupies a volume in space (figure 1.6.4).

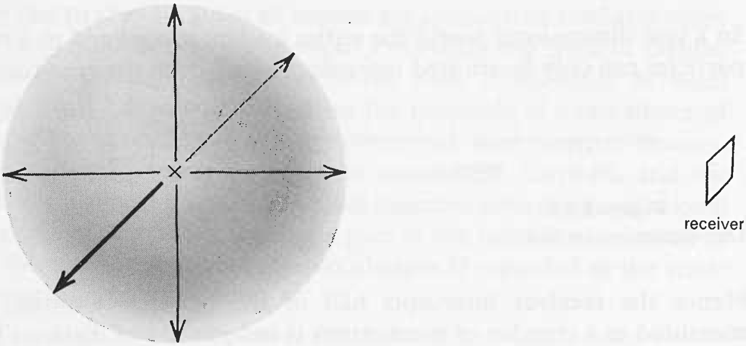


Figure 1.6.4
Three-dimensional world.

The receiver intercepts only those particles that fall within its area, and particles are being emitted from the source in all directions in space. To collect them all, we would need to surround the source by a sphere of surface area $4\pi r^2$. Hence the receiver would only intercept a number proportional to r^{-2}, and the force falls off in a three-dimensional world as the inverse square of the distance.

1-D	force proportional to	r^{0}
2-D	force proportional to	r^{-1}
3-D	force proportional to	r^{-2}

You may like to consider what would happen in a Universe of four spatial dimensions and what the consequences would be. Also, can you now see why the force between two parallel current-carrying wires is proportional to the inverse first-power of the distance, and why this inverse first-power rule holds for acceleration-dependent interactions?

Bibliography

Further reading

Ford, K. W. (1963) *The world of elementary particles.* Blaisdell.
Glasstone, S. (1965) *Sourcebook on the space sciences.* Van Nostrand.
Koelle, H. H. (ed.) (1961) *Handbook of astronautical engineering.* McGraw-Hill.
Neilson, I. D. (1966–67) 'The dynamics of safety belts in motor car head-on impacts.' *Symposium on ergonomics and safety in motor car design.* Proceedings of the Institution of Mechanical Engineers, **181**, Part 3D.
Neilson, I. D. (1969) *Simple representations of car and unrestrained occupant impacts in road accidents.* Road Research Laboratory Report, LR 249. Obtainable from the Transport and Road Research Laboratory, Crowthorne, Berkshire.
Sciama, D. W. (1972) *The physical foundations of general relativity.* Heinemann.

General background reading

Ministry of Transport (1968) *The Highway Code.* H.M.S.O.
Neilson, I. D. (1965) 'Safety belts, choosing and using them.' *Road Safety and Traffic Review* **12**, No. 2. This reprint is available from the Transport and Road Research Laboratory (for address see above).

Abstracts of projects

arranged by D. R. Browning

Motion and forces

1.1 The design and construction of an apparatus for demonstrating Newton's second law

The aim was to design an apparatus for demonstrating: force $= \mathrm{d}(mv)/\mathrm{d}t$ in the case of *variable mass.* A long trolley was made which ran under a hopper from which lead shot was poured. The trolley actuated a device to switch on the flow of shot at the right instant. Usual ticker-tape methods were used to test the law. A limitation was that there was not a continuous record of instantaneous velocity.

SR*: long trolley (trough-shaped) made from Handy-angle; flow wheels; runway; ticker timer; special hopper

P. Ayre, Bedford School, Bedford.

1.2 Nature of inelastic collisions

Masses were dropped onto a pan mounted on springs so that its depression could be recorded on a kymograph. The area under the resulting depression against time trace was compared with the theoretical value. Disappointing results.

I. Bright, Bletchley Grammar School, Bletchley, Bucks.

1.3 Friction-reducing properties of lubricants

This was a joint project. The aim was to push a sled along a horizontal track laying a film of oil as it went. The mass of the sled, the force applied, and the rate of application of oil onto the track were all varied. To compare the friction-reducing properties, the velocity of the sled at known points along the track was measured.

R. P. Jardine, A. C. Fletcher, and J. Knight, Cavendish School, Hemel Hempstead, Herts.

---

* SR – special requirements.

1.4 Pressure and lubrication by fluids

Standard friction experiments were repeated using oil films between the blocks and a Formica surface. The results led to the development of a rotating disc-type friction meter which is of interest.

R. W. Godson, Merchant Taylor's School, Northwood, Middx.

1.5 Rainfall on moving materials

Background – the question of the effect of raindrops incident on supersonic aircraft. The 'whirling arm' technique of study of multiple impact collisions was used. Here the bullet-shaped specimens were attached at the ends of a 1 m wooden arm whirled by a turntable in an area of simulated rain. This latter was produced by a modified garden spray, while the samples were made using a wet-sand mould made from a bullet. The turning rate was assessed stroboscopically. Later the arm was extended to nearly 4 m, using aluminium and suspended with the aid of nylon line – all in an attempt to increase the rate of erosion. The formation of an oxide coating on the metals was also investigated.

SR: motor

C. B. E. Robinson, Cleveland Grammar School, Warwick Road, Redcar, Yorks.

1.6 Rotation of the Sun

A telescope, fitted with a low power eyepiece, was used to project an image of the Sun onto a white card. Attempts to sketch the image failed, largely because of the speed of the moving image. Accordingly, the image on the card was photographed at various times using *smaller* exposures than suggested by a light meter. Analysis of the sun-spot positions (by two methods) gave an apparent rotation rate of 14.2 degrees per day for the Sun at a latitude of 10 degrees on the Sun. This figure checks well with literature values. (**Caution – never** look directly at the Sun, especially through a telescope.)

SR: a ten inch refracting telescope

N. Rowley, Marlborough College, Marlborough, Wilts.

1.7 Terminal velocity of a falling object in a liquid

The velocity of ball-bearings falling through various liquids was measured by a multiple flash photographic technique. Using Stokes's law the viscosity of the liquid was deduced. The variation of viscosity with temperature and the colligative properties of the liquid solution were studied.

C. M. Rowell and E. H. Sawika, City of London Girls' School, Barbican, London EC2.

1.8 Elastic properties of hair

An investigation of elastic properties of human hair, of different types, using a 'Young modulus' method. The hair was then treated with various cosmetic preparations according to manufacturers' instructions, and the elastic properties tested again. Results indicated that shampoos did not affect the dry strength but that wet hair became plastic. Colouring as such did not change the elastic properties. Bleaches rendered the hair brittle.

SR: travelling microscope on a vertical slide

K. Talmadg, Winchester County High School, Winchester, Hants.

1.9 Behaviour of human hair under stress

Load against extension curves were plotted for various samples of human hair and these were interpreted on the basis of structural changes. The effect of chemicals on hair was investigated through changes in the load against extension curves.

SR: travelling microscope

J. G. Malcolm, Aldridge Grammar School, Aldridge, Staffs.

1.10 Relative strengths of plastics

The aim was to construct a gun based on a mouse-trap to fire ball-bearings at a measured velocity at plastic sheets and to photograph the impact and subsequent path of the ball.

SR: fast camera; flashing stroboscope

I. Baxendale, Canon Slade Grammar School, Bolton, Lancs.

Fluid flow

1.11 Fluid flow in two dimensions

A thin film of water was made to flow around test objects, such as an aerofoil and a ship's hull. Streamlines, from potassium permanganate crystals, revealed the varying flow patterns. For a wing section, lift was demonstrated at the leading edge, and for a ship's hull, areas of stagnation at the stern. The streamlines narrowed as the water speed increased around an obstruction. Results were recorded graphically, and with a camera.

SR: a plate-glass flow tank

S. P. Golabek, Richard Taunton College, Southampton, Hants.

1.12 Proving Bernoulli's equation using a disc held up in mid-air equilibrium

When air was blown down onto a light smooth lamina from the middle of a smooth flat platform, the lamina was held in equilibrium just below the platform. The objectives of this project were to demonstrate this experiment, to make measurements to calculate the forces acting on the lamina, and to prove Bernoulli's equation. The power came through an autotransformer (plus voltmeter) to the vacuum cleaner. The air flowed up through the rotameter and down through the platform onto the circular piece of hardboard which was lifted. The lamina stayed under the jet when a nail was pushed through it. Bernoulli's equation was proved correct and the forces acting up and down on the lamina were shown to be equal.

SR: autotransformer; vacuum cleaner; rotameter; syringe needles; large diameter glass tubing; small diameter plastic tube; 50 and 100 cm rulers

J. H. S. Laycock, Merchant Taylors' School, Northwood, Middx.

1.13 Study of vortices

Apparatus was designed and built for the production and investigation of forced vortices in different liquids. A variac controlled record-player turntable was employed giving speeds from 5 to 78 rev min^{-1}, but difficulties with constancy of speeds were found using a spin-drier motor similarly controlled. The shape was investigated, as were the general properties and the decay time.

SR: motor

R. J. Croston, Merchant Taylors' School, Northwood, Middx.

1.14 Vortex tube

Investigation of temperature differences produced in turbulent gas flow in a 'vortex tube'. Several vortex tubes were constructed, involving an appreciation of the suitability of various materials. The validities of several theoretical explanations were tested.

SR: simple workshop facilities; high pressure gas supplies

M. Wingrove, Nottingham High School, Waverley Mount, Nottingham.

1.15 Vortex problems around high buildings

A wind tunnel (approximately 4 × 1 × 1 ft) was built from plywood and glass. Smoke from burning camphor was sucked by a fan around scale models of buildings, made of hardboard. These smoke tests gave flow diagrams which showed the eddies caused very clearly.

SR: fan; supply of camphor

P. Richardson, Merchant Taylors' School, Northwood, Middx.

Aerodynamics

1.16 Car bodies

Models of cars were carved in balsa wood in three sections so that different combinations of front and rear shapes could be investigated. Lift and drag were measured in a wind tunnel which had been constructed by a previous student, using a direct reading chemical balance under the tunnel. Wind speed was estimated using a pitot tube made out of copper tubing.

SR: wind tunnel; balsa wood; copper tubing, etc.; chemical balance

G. W. B. Inchbald, Bryanston School, Blandford, Dorset.

1.17 Wind tunnel experiments

Using a wind tunnel (calibrated by strobe) and aerodynamic balance a series of simple but most useful experiments was performed to investigate the basic relations between thrust, drag, weight, and lift for various shapes in an air-stream.

SR: wind tunnel or equivalent

R. Kiralfy, Merchant Taylors' School, Northwood, Middx.

1.18 Hovercraft or hydrofoil

A comparison of the use of the two for high-speed Channel crossing was made from published data and measurements on small models.

S. Macrae, Bletchley Grammar School, Bletchley, Bucks.

1.19 Hovercraft

Construction of a one-man Hovercraft (approximately 8 × 6 ft) mainly from 'scrap' materials. A wooden frame made airtight with sheet polythene supported a horizontally mounted BSA 4 stroke 420 cm^3 petrol engine driving a rebuilt fan from a diesel lorry forcing air through an aluminium duct into a semi-skirt of neoprenized nylon.

A. N. Kagan, Gordonstoun School, Elgin, Moray, Scotland.

Sea-craft

1.20 Inshore rescue craft

A number of improvements to the standard RNLI 'Zodiac' inflatable boat were made over a long period. This project on hull design arose from that, the aim being to calibrate the speed: thrust ratios observed in models with their full scale counterparts.

SR: boats and water; scaler and photo-timing equipment

L. Celt and I. Bryceson, Atlantic College, Llantwit Major, Glamorgan.

1.21 Engine design: revolution limiter

The ruining of several engine flywheels by running the engines too hard led to this work. The device involved the control of the ignition by the tachometer using a relay and was moderately successful.

SR: this type of engine; relay

Suggested further work: investigate speed: weight and thrust: weight ratio; different model size and the effect of using a small engine on the inshore rescue craft model.

H. Lundberg, Atlantic College, Llantwit Major, Glamorgan.

1.22 Hull design of RNLI inshore rescue boat

A scale model was made of an existing design to test an assumption that a 'step' in the hull would lead to a better performance. This was found to be the case. No conclusions could be drawn about effect of 'step' on seaworthiness.

SR: test tank

M. Henriques, Atlantic College, Llantwit Major, Glamorgan.

Fuels

1.23 Internal combustion engines and fuel consumption

An investigation into how much fuel an internal combustion engine uses, how it varies under different speeds, and how comparative fuels perform. The fuel system had to be modified in order to obtain meaningful results; the electrical system had to be improved and a brake dynamometer constructed. The conclusion reached was that octane value was not the last word in deciding performance; the engine tuning and design were of greater significance. Owing to the time taken in preparing the engine fewer results could be obtained than was hoped at first.

SR: suitable engine; motor-bike petrol tank; railway sleepers; various types of fuels

Suggested further work: the measurement of brake horse power against engine speed for different fuels, determination of the calorific values of the fuels, and an analysis of the exhaust fumes.

T. Lasek and M. Mandzij, Cardinal Hinsley Grammar School, Tong, Bradford, Yorks.

1.24 Rockets: solid fuel

The test-bed consisted of a trolley and ticker-timer system on which was mounted (with particular attention to safety) a rocket fuel chamber, combustion chamber, and nozzle, all this engine being made from a piece of mild steel tubing. This tubing was of total length under 120 mm. The fuel used initially was 'Jetex', and the results were compared with those obtained with a mixture of zinc powder and sulphur. Very many practical difficulties were encountered.

Suggested further work: for ideas see the project on internal combustion engines (1.23).

E. W. Hislop, Bedford School, Bedford.

1.25 Investigation into fuels for self combustion engines

Various fuel mixtures were made using differing quantities of castor oil and ether. A model aircraft diesel engine was then run on these fuels. The output was measured by connecting the engine to a dynamo.

SR: diesel engine, dynamo

T. Gordon, Roan Grammar School for Boys, Blackheath, London, SE3.

1.26 Factors affecting performance of small diesel motor

A small model aeroplane motor was mounted on a platform which swung forward, giving the motor potential energy proportional to the thrust it produced. Thrust was measured by pen and rotating drum, airspeed by manometer. Factors varied were propeller size and pitch, type of fuel, humidity, compression, etc., and graphs were drawn by pen and compared. Performance under varying conditions was thus measured by comparing total impulse, time of running on measured amount of fuel, and overheating problems.

SR: motor; test rig (as described above); slow speed motor to drive drum; and plenty of diesel fuel

P.R.D.H. Greenhouse, Merchant Taylors' School, Northwood, Middx.

Other aspects

1.27 Investigation of reaction times

Voluntary (aural and visual) and involuntary (patellar tendon activation) human response times were measured using a millisecond timer. It was found that in all three cases the taller the subject the longer the reaction time. There was a similar difference in both voluntary response times between hand and foot movement and this allowed a calculation of the speed with which impulses travel in nerves of about 100 $\mathrm{m\ s^{-1}}$. Work was begun on a speed determination in the sciatic nerve of a frog by electrical stimulation.

SR: electronic timer, for example, GM scaler with millisecond timing facility

D. B. Howat, Abingdon School, Abingdon, Berks.

Part 2

Elements and compounds

Contents

Introduction

Our knowledge of the material world depends on experiment and observation. Both approaches are required for a classification of chemical knowledge which employs the concepts of element and compound. The first article considers the development of the modern view of chemical elements. During the nineteenth century, a great deal of attention was given to the notion of 'reacting weights' of elements and compounds and led to the development of precise chemical methods of analysis for the determination of atomic 'weights'. One of these procedures is briefly reviewed in the second article. A separate article deals with the selection of carbon-12 as a standard of mass for relative atomic mass determinations. An extensive review of modern ideas about compounds follows. Three special aspects of the nature of compounds and their reactivity conclude this selection of articles: uses of complexing agents; unusual oxidation states; and acid–base chemistry in the human body.

Related sections of the physical science course

Section 2	The elements of the second short period
Section 4	Some important chemical reactions
Section 5	Electricity and atomic structure
Section 9	Covalent bonds and the compounds of carbon
Section 10	Group relationships in the periodic table
Section 11	Elements of the d-block

Related sections of other Nuffield Advanced Science courses

Chemistry

Topic 1	Amount of substance
Topic 2	Periodicity
Topic 3	The masses of molecules and atoms: the Avogadro constant
Topic 5	The halogens and oxidation numbers
Topic 6	The s-block elements; and the acid–base concept

Topic 13	Carbon chemistry, part 2
Topic 15	Equilibria: redox and acid–base systems
Topic 16	Some d-block elements
Topic 19	Some p-block elements

Physics

Unit 1	Materials and structure
Unit 3	Field and potential
Unit 5	Atomic structure
Unit 10	Waves, particles, and atoms

2.1 Chemical elements

by B. E. Dawson

A study of the history of chemistry reveals that the ideas underlying definitions of such terms as element or compound, although so obvious to us in the twentieth century, were reached only after centuries of effort. Perhaps the first clear expression of the idea of an 'element' occurs in the teachings of the Greek philosophers several centuries before Christ. Many detailed statements exist about such views. Our present concern rests with ideas directly linked to the current interpretation of the term element.

The work of Robert Boyle, as expressed in his book *The Sceptical Chymist* publised in 1661, did much to alter opinion towards our current manner of thinking. Boyle defined the term in the following way:

'And, to prevent mistakes, I must advertize you, that I now mean by elements, as those Chymists that speak plainest do by their principles, certain primitive and simple, or perfectly unmingled bodies; which not being made of any other bodies, or of one another, are the ingredients of which all those called perfectly mixt bodies are immediately compounded, and into which they are ultimately resolved: now whether there be any one such body to be constantly met with in all, and each, of those that are said to be elemented bodies, is the thing I now question.'

This is a clear statement. But, as will be revealed to those who read the book, the weak point of Boyle's argument was his failure to find experimental means to decide whether or not a specified substance is elemental. Even so, Boyle's statement had a profound influence on others.

About a century later, and in what has come to be regarded as the first textbook on chemistry, A-L. Lavoisier put forward the description of an element as 'that which cannot be separated into simpler substances'. The book, *Traité élémentaire de Chimie* published in 1789, contained many novel features. As will be appreciated, his discussion on elements is a natural extension of the idea put forward by Boyle: the view is a tentative one, for, as the text states (in translation*):

'since we have not hitherto discovered the means of separating them, they act with regard to us as simple substances, and we ought never to suppose them compounded until experiment and observation has proved them to be so.'

---

* Translation from Kerr, R. (1790) *Elements of chemistry, in a new systematic order, containing all the modern discoveries* (Preface pp. *xxiv* and *xvi*).

Lavoisier supports this approach by listing elements known to him and this list is reproduced as figure 2.1.1.

192 DES SUBSTANCES SIMPLES.

TABLEAU DES SUBSTANCES SIMPLES.

	Noms nouveaux.	Noms anciens correſpondans.
Subſtances ſimples qui appartiennent aux trois règnes & qu'on peut regarder comme les élémens des corps.	Lumiere	Lumière.
	Calorique	Chaleur. Principe de la chaleur. Fluide igné. Feu. Matière du feu & de la chaleur.
	Oxygène	Air déphlogiſtiqué. Air empiréal. Air vital. Baſe de l'air vital.
	Azote	Gaz phlogiſtiqué. Mofete. Baſe de la mofete.
	Hydrogène	Gaz inflammable. Baſe du gaz inflammable.
Subſtances ſimples non métalliques oxidables & acidifiables.	Soufre	Soufre.
	Phoſphore	Phoſphore.
	Carbone	Charbon pur.
	Radical muriatique.	Inconnu.
	Radical fluorique	Inconnu.
	Radical boracique	Inconnu.
Subſtances ſimples métalliques oxidables & acidifiables.	Antimoine	Antimoine.
	Argent	Argent.
	Arſenic	Arſenic.
	Biſmuth	Biſmuth.
	Cobolt	Cobolt.
	Cuivre	Cuivre.
	Etain	Etain.
	Fer	Fer.
	Manganèſe	Manganèſe.
	Mercure	Mercure.
	Molybdène	Molybdène.
	Nickel	Nickel.
	Or	Or.
	Platine	Platine.
	Plomb	Plomb.
	Tungſtène	Tungſtène.
	Zinc	Zinc.
Subſtances ſimples ſalifiables terreuſes.	Chaux	Terre calcaire, chaux.
	Magnéſie	Magnéſie, baſe du ſel d'Epſom.
	Baryte	Barote, terre peſante.
	Alumine	Argile, terre de l'alun, baſe de l'alun.
	Silice	Terre ſiliceuſe, terre vitrifiable.

Figure 2.1.1
Lavoisier's list of elements.
From Lavoisier, A-L (1789) Traité élémentaire de chimie. *Photographed by courtesy of the Royal Society.*

Experience and knowledge of developments underlines the significance of his counsel requiring a classification of materials based on 'experiment and observation'.

Some years later, John Dalton made a notable advance by endowing the atoms of chemical elements with fixed and definite weights, thereby providing a satisfactory basis for the atomic theory. Also about this same time, Humphry Davy provided a classical illustration of the value of Lavoisier's criteria for elements through his work on chlorine. This work appeared in a paper published in 1810 entitled 'Researches on the oxymuriatic acid, its nature and combinations; and on the elements of the muriatic acid. With some experiments on sulphur and phosphorus, made in the Laboratory of the Royal Institution.' (Davy, H. (1810) *Phil. Trans.* **100**, 231–257).

Much has taken place since these contributions first appeared. Activities in chemical science have been influenced by many factors: personal tastes; personalities of leading investigators; the development of new theoretical concepts; and by general advances in technique. Figure 2.1.2 shows the number of chemical elements known at a given time as a function of time. This record also demonstrates the uneven nature of progress.

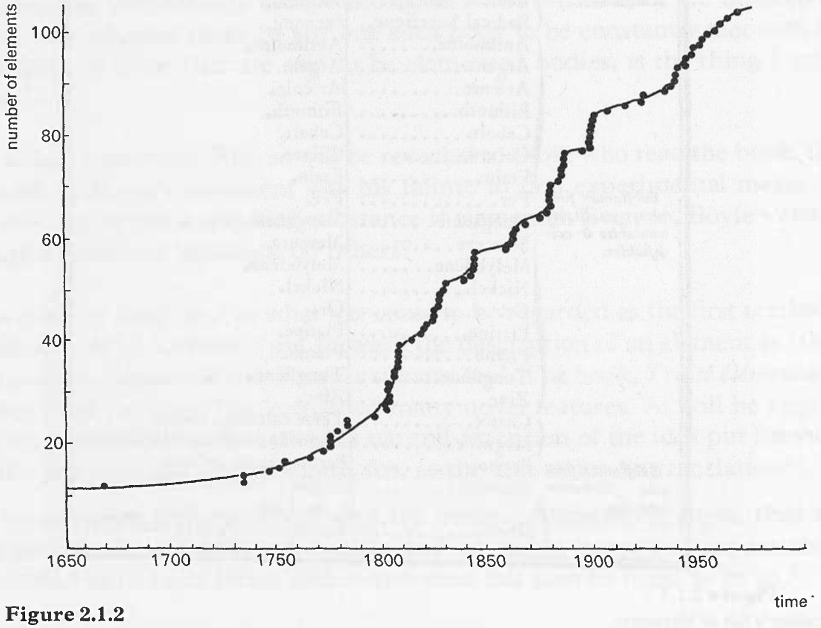


Figure 2.1.2
Number of chemical elements known as a function of time.
After de Solla Price, D. J. (1963) Little science, big science. *Columbia University Press.*

As may be judged from figure 2.1.2, Davy's work on chlorine appeared at a time when there was a steady growth in this aspect of chemical knowledge. A detailed inspection of the literature reveals that a great deal of new knowledge was being acquired at this time (cf. table 1 below). A fresh impetus was acquired once the value of Mendeleev's periodic law had been realized. The isolation of the rare gases and of radioactive elements account for other changes in figure 2.1.2 for the period at about the turn of the century. The most recent escalation shown on the graph arises from the production of trans-uranic elements which are also highly unstable and short-lived. Table 1 contains a complete list of elements, together with details relating to their discovery. Further information is given in Weeks, M. E. (1968) *Discovery of the elements* (7th ed.), Chemical Education Publishing Company.

Our present knowledge of elemental materials permits both the isolation of these substances as well as their production by what we can term 'synthetic methods', such as have been employed to make the trans-uranic elements. Techniques have now been evolved to permit the production of pure elements in large quantities in a number of instances, whereas in others production is only possible in terms of a few atoms. Each chemical element has been identified through its unique properties and it has been the recognition of such properties that has contributed to the evolution of the concept of element.

Table 1

Element	Atomic number	Date of discovery	Origin of name	Discoverer
Actinium	89	1899	Gk *aktis, aktinos,* beam or ray	A. Debierne (France)
Aluminium	13	1827	L *alumen, alum*	F. Wöhler (Germany)
Americium	95	1944	the Americas (analogy with Europium)	G. T. Seaborg, R. James, *et al* (U.S.A.)
Antimony	51	ancient	L *antimonium* (Sb, L *stibium)*	
Argon	18	1894	Gk *argos,* inactive	Lord Rayleigh and W. Ramsay (U.K.)
Arsenic	33	1250	L *arsenicum,* Gk *arsenikon,* yellow orpiment	Albertus Magnus (Germany)
Astatine	85	1940	Gk *astatos,* unstable	E. Segrè, K. McKenzie, D. Corson (U.S.A.)
Barium	56	1808	Gk *barys,* heavy	H. Davy (U.K.)
Berkelium	97	1949	Berkeley, University of California	S. G. Thompson, A. Ghiorso, G. T. Seaborg, (U.S.A.)

Element	Atomic number	Date of discovery	Origin of name	Discoverer
Beryllium	4	1798 1828	Gk *beryllos,* beryl	L. N. Vauquelin (France) – as the oxide F. Wöhler (Germany), A. A. B. Bussy (France)
Bismuth	83	1753	Ger *weisse Masse,* white mass (later, *Wismut*)	C. Geoffroy (France)
Boron	5	1808	Ar *Būraq,* Pers *Būrah*	H. Davy (U.K.), J. L. Gay-Lussac and L. J. Thenard (France)
Bromine	35	1826	Gk *bromos,* stench	A. J. Balard (France)
Cadmium	48	1817	L *cadmia,* Gk *kadmeia,* ancient name for calamine	F. Stromeyer (Germany)
Caesium	55	1860	L *caesius,* sky blue (colour of line in its spectrum)	R. W. Bunsen and G. R. Kirchhoff (Germany)
Calcium	20	1808	L *calx, calcis,* lime	H. Davy (U.K.)
Californium	98	1950	University and State of California	S. G. Thompson, K. Street, *et al* (U.S.A.)
Carbon	6	ancient	L *carbo,* charcoal	
Cerium	58	1803	after the asteroid Ceres, discovered in 1801	J. J. Berzelius and W. Hisinger (Sweden), M. H. Klaproth (Germany) – as the oxide
Chlorine	17	1774	Gk *chloros,* pale green	C. W. Scheele (Sweden)
Chromium	24	1797	Gk *chroma,* colour	L. N. Vauquelin (France)
Cobalt	27	1735	Ger *Kobold,* goblin or evil spirit	G. Brandt (Sweden)
Copper	29	ancient	L *cuprum,* from the island of Cyprus	
Curium	96	1944	in honour of Marie and Pierre Curie	G. T. Seaborg, R. James, A. Ghiorso (U.S.A.)
Dysprosium	66	1886	Gk *dysprositos,* hard to obtain	P. E. L. de Boisbaudran (France)
Einsteinium	99	1952	in honour of Albert Einstein	Laboratories at Berkeley, Los Alamos, and Argonne (U.S.A.)

Element	Atomic number	Date of discovery	Origin of name	Discoverer
Erbium	68	1843	one of four elements obtained from the rock ytterite (now known as gadolinite) which occurs at Ytterby, near Stockholm, Sweden	C. G. Mosander (Sweden)
Europium	63	1896	Europe	E. A. Demarçay (France)
Fermium	100	1953	in honour of Enrico Fermi	Laboratories at Berkeley, Los Alamos, and Argonne (U.S.A.)
Fluorine	9	1886	L *fluere,* to flow	H. Moissan (France)
Francium	87	1939	in honour of France	M. Perey (France)
Gadolinium	64	1880	in honour of J. Gadolin – a Finnish chemist	J. C. G. de Marignac (France)
Gallium	31	1875	L *Gallia,* France	P. E. L. de Boisbaudran (France)
Germanium	32	1886	L *Germania,* Germany	C. A. Winkler (Germany)
Gold	79	ancient	AS *gold* (Au, L *aurum*)	
Hafnium	72	1923	L *Hafnia,* Copenhagen	D. Coster (Holland), G. von Hevesy (Hungary)
*Hahnium**	105	1967		G. N. Flerov, V. A. Druin, *et al* (U.S.S.R.)
		1970	in honour of O. Hahn, co-discoverer of atomic fission	A. Ghiorso *et al* (U.S.A.)
Helium	2	1868	Gk *helios,* Sun	J. N. Lockyer (U.K.) – from spectrum of light from the Sun
		1895		W. Ramsay (U.K.) – isolation
Holmium	67	1879	L *Holmia,* Stockholm	P. T. Cleve (Sweden)
Hydrogen	1	1766	Gk *hudor,* water and *genes,* producing	H. Cavendish (U.K.)
Indium	49	1863	from the brilliant indigo line in its spectrum	F. Reich and H. T. Richter (Germany)

* The name of this element has yet to be approved internationally.

Element	**Atomic number**	**Date of discovery**	**Origin of name**	**Discoverer**
Iodine	53	1811	Gk *iodes,* violet	B. Courtois (France)
Iridium	77	1803	Gk *iris,* rainbow	S. Tennant (U.K.)
Iron	26	ancient	AS *iren* (Fe, L *ferrum*)	
Krypton	36	1898	Gk *kryptos,* hidden	W. Ramsay and M. W. Travers (U.K.)
*Kurchatovium**	104	1964		G. N. Flerov, Yu. Ts. Oganessyan, *et al* (U.S.S.R.)
Lanthanum	57	1839	Gk *lanthanein,* to lie hidden	C. G. Mosander (Sweden) – as the oxide
Lawrencium	103	1961 1965	in honour of E. O. Lawrence, inventor of the cyclotron	A. Ghiorso, T. Sikkeland, *et al* (U.S.A.) E. D. Donets, V. A. Shchegolev, V. A. Ermakov (U.S.S.R.)
Lead	82	ancient	AS *lead* (Pb, L *plumbum*)	
Lithium	3	1818	Gk *lithos,* stone	J. A. Arfvedson (Sweden)
Lutetium	71	1907	L *Lutetia,* ancient name of Paris	G. Urbain (France), A. von Welsbach (Austria)
Magnesium	12	1808	Gk *Magnesia,* district in Thessaly	H. Davy (U.K.)
Manganese	25	1774	F *manganese,* corrupt form of *Magnesia*	J. G. Gahn (Sweden)
Mendelevium	101	1955	in honour of D. Mendeleev	A. Ghiorso, B. Harvey, *et al* (U.S.A.)
Mercury	80	ancient	planet Mercury (Hg, Gk *hydrargyrum,* water silver)	
Molybdenum	42	1778	Gk *molybdos,* lead	C. W. Scheele (Sweden) – as the oxide P. J. Hjelm (Sweden)
Neodymium	60	1841 1885	Gk *neos,* new and *didymos,* twin	C. G. Mosander (Sweden) – as the oxide A. von Welsbach (Austria)
Neon	10	1898	Gk *neos,* new	W. Ramsay and M. W. Travers (U.K.)

* The name of this element has yet to be approved internationally.

Element	**Atomic number**	**Date of discovery**	**Origin of name**	**Discoverer**
Neptunium	93	1940	Neptune, planet beyond Uranus	E. M. McMillan and P. H. Abelson (U.S.A.)
Nickel	28	1751	Ger *Kupfernickel* (abbrev.), 'Old Nick's copper'	A. F. Cronstedt (Sweden)
Niobium	41	1802	Gk myth, *Niobe,* daughter of Tantalus	C. Hatchett (U.K.)
Nitrogen	7	1772	L *nitrum*; Gk *nitron,* nitre and *genes,* producing	D. Rutherford (U.K.)
Nobelium	102	1957	in honour of Alfred Nobel	Swedish–English–American group
		1957–8		G. N. Flerov, S. M. Polikanov, *et al* (U.S.S.R.)
		1958		A. Ghiorso, T. Sikkeland, *et al* (U.S.A.)
		1963		E. D. Donets, V. A. Shchegolev, V. A. Ermakov (U.S.S.R.)
Osmium	76	1803	Gk *osme,* smell	S. Tennant (U.K.)
Oxygen	8	1774	Gk *oxys,* sharp, acid and *genes,* producing	J. Priestley (U.K.) C. W. Scheele (Sweden) – published 1777
Palladium	46	1803	after the asteroid Pallas, a recent discovery	W. H. Wollaston (U.K.)
Phosphorus	15	1669	Gk *phorphoros,* light-bringing	H. Brand (Germany)
Platinum	78	1735 1741	Sp *platina,* silver	A. de Ulloa (Spain) C. Wood (U.K.)
Plutonium	94	1940	Pluto, planet beyond Neptune	G. T. Seaborg E. McMillan, *et al* (U.S.A.)
Polonium	84	1898	in honour of Poland (L *Polania*)	M. Curie (Poland)
Potassium	19	1807	E *pot-ashes* (K, L *kalium*)	H. Davy (U.K.)
Praseodymium	59	1885	Gk *prasios,* green and *didymos,* twin	A. von Welsbach (Austria)
Promethium	61	1947	Gk myth, *Prometheus,* who stole fire from the gods	J. A. Marinskii, L. E. Glendenin (U.S.A.)
Protactinium	91	1917	Gk *protos,* first	O. Hahn (Germany) and L. Meitner (Austria)

Element	**Atomic number**	**Date of discovery**	**Origin of name**	**Discoverer**
Radium	88	1898	L *radius,* ray	P. Curie (France) and M. Curie (Poland)
Radon	86	1900	from radium	F. E. Dorn (Germany)
Rhenium	75	1925	L *Rhenus,* Rhine	W. Noddack, I. Tacke, O. Berg (Germany)
Rhodium	45	1803	Gk *rhodon,* rose	W. H. Wollaston (U.K.)
Rubidium	37	1861	L *rubidus,* darkest red, the colour of line in its spectrum	R. W. Bunsen and G. R. Kirchhoff (Germany)
Ruthenium	44	1844	L *Ruthenia,* province in Eastern Europe	K. K. Klaus (Russia)
Samarium	62	1879	Samarskite, the mineral from which it was prepared	P. E. L. de Boisbaudran (France)
Scandium	21	1879	L *Scandia,* Scandinavia	L. F. Nilson (Sweden)
Selenium	34	1817	Gk *selene,* Moon – cf. with tellurium	J. J. Berzelius (Sweden)
Silicon	14	1824	L *silex, silicis,* flint	J. J. Berzelius (Sweden)
Silver	47	ancient	AS *seolfor* (Ag, L *argentum*)	
Sodium	11	1807	E *soda* (Na, L *natrium*)	H. Davy (U.K.)
Strontium	38	1808	after Strontian, a town in Argyllshire	H. Davy (U.K.)
Sulphur	16	ancient	S *sulvere,* L *sulphur*	
Tantalum	73	1802	Gk myth *Tantalos,* father of Niobe	A. G. Ekeberg (Sweden)
Technetium	43	1937	Gk *technikos,* artificial	E. Segrè and C. Perrier (U.S.A.)
Tellurium	52	1783	L *tellus, telluris,* earth	F. J. Müller (Austria)
Terbium	65	1843	one of four elements obtained from the rock ytterite (now known as gadolinite) which occurs at Ytterby, near Stockholm, Sweden	C. G. Mosander (Sweden)
Thallium	81	1861	Gk *thallos,* budding twig – because of the green line in its spectrum	W. Crookes (U.K.)
Thorium	90	1828	Thor, Scandinavian god of war	J. J. Berzelius (Sweden)

Element	**Atomic number**	**Date of discovery**	**Origin of name**	**Discoverer**
Thulium	69	1879	Thule, ancient name of Scandinavia	P. T. Cleve (Sweden)
Tin	50	ancient	AS *tin* (Sn, L *stannum*)	
Titanium	22	1791 1825	Gk myth *Titans*, sons of the earth goddess	W. Gregory (U.K.) – as the oxide J. J. Berzelius (Sweden)
Tungsten	74	1783	Sw *tung*, heavy and *sten*, stone. W, from the ore wolframite	J. J. and F. d'Elhuyar (Spain)
Uranium	92	1789 1841	planet Uranus	M. H. Klaproth (Germany) – as the oxide E. M. Peligot (France)
Vanadium	23	1801 1830	Scandinavian goddess Vanadis	A. M. del Rio (Spain) – confused with chromium N. G. Sefström (Sweden)
Xenon	54	1898	Gk *xenos*, stranger	W. Ramsay and M. W. Travers (U.K.)
Ytterbium	70	1878	one of four elements obtained from the rock ytterite (now known as gadolinite) which occurs at Ytterby, near Stockholm, Sweden	J. C. G. de Marignac (France)
Yttrium	39	1843	after Ytterby, a town near Stockholm	C. G. Mosander (Sweden)
Zinc	30	1746	Ger *Zink* – origin obscure	A. S. Marggraf (Germany)
Zirconium	40	1789 1824	Ar *zargun*, gold colour	M. H. Klaproth (Germany) – as the oxide J. J. Berzelius (Sweden)

Abbreviations used: Ar, Arabic; AS, Anglo-saxon; E, English; F, French; Ger, German; Gk, Greek; L, Latin; Pers, Persian; S, Sanskrit; Sp, Spanish; Sw, Swedish.

This table is adapted from Stark, J. G. and Wallace, H. G. (1970) 'Chronology of the discovery of the elements.' *Educ. Chem.* **7**, 152.

2.2 The silver method for the accurate determination of atomic weights

by J. M. Faulkner

Any reader of the history of chemistry in the nineteenth century will be struck by the attention given to the determination of the reacting weights of elements and compounds, and hence to the evaluation of the atomic weights of the elements concerned; increasingly high standards of purity and progressively refined techniques marched hand in hand, and have continued into the earlier decades of the present century as comparisons became desirable between atomic weight values obtained from purely physical as distinct from chemical methods. Of the many workers in this field, the names of J. S. Stas (1813–91), T. W. Richards (1868–1928), and O. Hönigschmid (1878–1945) stand out as those whose work received general acceptance.

Although chemical atomic weights were based (until 1961) upon the scale $O = 16$, the ratios actually determined necessarily involved elements other than oxygen and for many elements the ratios metallic halide: silver or, metallic halide: silver halide were experimentally convenient. Calculation of the atomic weight thus presupposed a knowledge of the atomic weight of silver on the oxygen scale, and to this work Stas brought a degree of patience and skill hitherto unexampled. He determined the ratios $KClO_3:KCl:Ag:3O$ by heating potassium chlorate and determining the loss in weight consequent upon driving off the oxygen, and by a precipitation reaction between potassium chloride and silver nitrate; he determined the value for silver itself by burning silver in a current of chlorine and also by dissolving silver in nitric acid and precipitating silver chloride. His value $Ag = 107.94$ was accepted for many years. By dissolving silver in nitric acid and subsequently weighing the fused nitrate, Stas determined the ratio $Ag:AgNO_3$ and hence calculated the value $N = 14.050$.

Towards the end of the century Richards, profiting by much improved laboratory facilities which had become available, pointed out several possible sources of error in the experiments of Stas and attacked again the problem of the atomic weights of silver, nitrogen, and chlorine. Richards suggested that (1) even the most carefully purified potassium chlorate contains chloride; (2) traces of potassium chloride may be carried down by precipitated silver chloride (aggravated when solid potassium chloride is used); (3) the solubility of silver chloride varies slightly in the presence of excess of one or other of the reagents; (4) Stas's silver may have contained oxygen.

It is to Richards and his co-workers that we owe major advances in technique and values (for example, N = 14.008) fully substantiated by the mass spectrographic method.* Richards made a special study of occlusion, particularly with regard to the contamination of precipitates; it is fortunate that silver chloride is a precipitate of a spongy character from which impurities can be washed with comparative ease. He laid constant stress on the necessity for the utmost purity of both starting and derived materials; for instance, he prepared silver by reduction of silver nitrate which had been fifteen times recrystallized, the silver being fused on lime in an atmosphere of hydrogen. He used platinum or quartz vessels lest precipitates be contaminated by traces of silica from glass; he employed electric heating, and he developed the centrifuge for separating solids from their mother liquors – at that time a technique uncommon among laboratory investigators although in use in industrial processing. He also devised two novelties; a bottling apparatus, and a nephelometer.

The bottling apparatus consisted essentially of a quartz glass tube in which a substance could be dried by heating in any desired atmosphere, then transferred to a weighing bottle and stoppered without exposure to ordinary air. The nephelometer was a device for measuring the amount of silver chloride in the washings by a visual comparison (against a standard) of the opalescence produced when excess silver nitrate was added. Richards obtained for silver the value Ag = 107.881.

Although initiated by Richards, it is mainly to Hönigschmid and his collaborators that we owe a further advance in accuracy by the use of dry reactions. In a determination of new ratios, namely $Ba(ClO_4)_2 : BaCl_2 : 2Ag : 8O$, the perchlorate was reduced to chloride in a stream of dry hydrogen chloride at temperatures between 200 °C and 600 °C. At the same time, Hönigschmid developed the bottling apparatus devised by Richards so that drying, reaction, and weighing in vacuo could all be accomplished without unwanted exposure, the boat being moved and the bottle stoppered by a magnetically operated piston.

---

* See article 9.1 for an account of modern mass spectrometry.

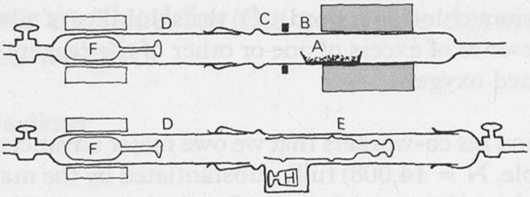


Figure 2.2.1

An example of the operation of the apparatus is as follows. After drying, the boat A (see figure 2.2.1) is pushed into the glass tube D which is then transferred, in a counter current of dry air, to the bottling apparatus, E. After evacuation, the magnetically operated piston, F, is used to push the boat into the weighing tube. On rotation of the apparatus, the piston may again be used to push the stopper, H, into position.

The results of this direct measurement of the ratio Ag: 4O agree completely with the value found by dry reduction of silver nitrate to metallic silver, and confirm the value given by Richards. It would appear from the results so far available that the standard of reproducibility attainable by the use of dry reactions (allied with modern techniques) is highly satisfactory, and the values are in close agreement with those computed from mass spectrographic data.

It would be fitting to complete this paper with a quotation from T. W. Richards*. 'To sum up the matter in a few words, it may be said that the secret of success in a precise chemical or physico-chemical measurement lies in so choosing the particular substance and process and so checking every operation by parallel experiments that both chemical and physical errors may be avoided as effectively as possible: and this choice often involves much study and above all the application of sound common sense. The precautions must be of a consistent order of refinement. Far more depends upon this intelligent choice of conditions than upon mere mechanical execution of the operations, although this too is important.'

---

* Quoted in: Sir Harold Hartley (1933), 'The Theodore William Richards Memorial Lecture.' *Chemical Society Memorial Lectures* **III**.)

2.3 The carbon-12 scale for relative atomic weights

Adapted from Wichers, E. (1964) 'Why the carbon-12 scale?'
Chemistry **37** No. 3, 12–17

Many advances towards understanding natural phenomena have been made possible through the adoption, and subsequent modification, of a simple idea. One of the significant ideas derived from the atomic theory is that the masses of all atomic particles of a given element are identical and differ from the atomic masses of all other elements. This simple idea permeated and formed the key to the structure of chemical science throughout the nineteenth – and even for part of the twentieth – century. The discovery of isotopes (a term coined by F. Soddy, 1913) led to a drastic revision in thinking for all chemists and led ultimately to the adoption in 1961 of a single scale of relative atomic weights based on the standard of the carbon-12 isotope.

All determinations of atomic weights based on chemical procedures yield an average relative atomic mass value dependent upon an arbitrarily assigned mass value for a standard element. Obviously this average value must depend on both the average atomic mass of the naturally occurring mixture of isotopic forms of the element and the isotopic composition of the standard element. When minute quantities of ^{17}O and ^{18}O were discovered in naturally occurring oxygen, physicists used the mass of the ^{16}O isotope as their standard of mass whereas chemists assigned this same value to the average isotopic mass of naturally occurring oxygen. Since natural oxygen contains only traces of ^{17}O and ^{18}O, the two scales were very nearly alike. In fact, the ratio of the 'atomic weight' of the element oxygen (that is, the average atomic mass of the three isotopes) to the atomic mass of ^{16}O is about 1.00028 : 1.

It should be noted that although the differences between these two scales are small, such differences become significant when determinations of fundamental constants such as the Avogadro constant are attempted. Accordingly, it became customary to give the mass scale adopted when quoting such a result.

Other work on oxygen showed that the relative abundance of the three isotopes of oxygen depended on the source used, air, water, or rocks. This variation was extremely small – too small in fact to influence chemically derived values for atomic masses but sufficiently large to influence precise determinations of atomic masses of individual nuclides.

Clearly, the possibility of adopting a uniformly applied standard of atomic mass needed to be considered. The major problem facing chemists was the truly enormous quantity of data published, using the old scale. About one million numerical quantities relevant to physical chemistry were assessed to possess sufficient accuracy such that any adjustment to a new scale band on $^{16}O = 16$ would introduce a significant error. In contrast, the problem for physicists was not of this magnitude.

In principle, any chemical element could be used as a reference standard for a scale of atomic weights provided some simple means were available to allow its atomic mass to be related to one or more other elements. Of the twenty-one elements which did not possess naturally occurring isotopes, only fluorine had an atomic mass on the old chemical scale which was nearly an integral number. This standard might then be feasible for chemists. Unfortunately, fluorine does not lend itself to good operating standards in mass spectroscopy and therefore proved to be unacceptable. Discussion revealed that elements need not be the only feasible materials for use as a standard. A nuclide can be used for this purpose provided means are available for relating its atomic mass accurately to a mixture of nuclides such as may be present in naturally occurring elements. Fortunately the atomic mass of ^{12}C on the old chemical scale differs from the whole number 12 by only about 4 parts in 100 000. This difference is smaller than the estimated uncertainty of any except a very few physico-chemical data, and so conversion from the chemical scale to one based on $^{12}C = 12$ did not require an extensive revision of existing data. The adoption of this new standard required physicists to change their reference from $^{16}O = 16$ to $^{12}C = 12$ and in practice this meant a revision of the masses of only several hundred nuclides.

The table of International Atomic Weights adopted in 1961 contains two types of change not shown in its forerunner, the 1957 table. The first of these changes is due to the change of scale (for example, oxygen is no longer 16 but 15.9994) and amounted to 0.004%, a fraction too small to influence more than a few of the values listed. The second set of changes depended on the recalculation of experimental work done on each of the elements. It should be noted that the official table does not list atomic weights for any radioactive element since the isotopic compositions of such elements (and hence their atomic masses) depend on the source used.

2.4 Chemical compounds

by John Emsley

Atoms come together to form compounds. By so doing they can pair off electrons and the resulting product is energetically more stable. Not all combinations of atoms, however, are known and some elements such as the rare gases are particularly loath to enter into compound formation. To the early chemists the proportion of particular elements in a compound was a mystery in itself and even when simple formulae had been worked out the results were inexplicable. The great Dalton rightly noted that his theory of atoms meant that compounds formed from them must of necessity produce integral ratios of the proportions of the elements. Nevertheless not all ratios were observed and some ratios were found to be the same for widely different compounds; also some compounds stubbornly refused to obey this simple integral ratio requirement. Now we can explain these three effects under the headings of valency, isomerism, and non-stoichiometry respectively. Although on the face of it chemistry appears to be a 'numbers game', this approach is of no real value. A much more fruitful avenue to understanding chemical compounds is via chemical bonding.

The basis of chemical bonding is electron interaction, and in particular electron (spin) pairing. Not all the electrons of an atom are involved in the formation of bonds to other atoms. The electrons in filled inner shells around the nucleus play little or no part. The electrons in the outer, partly filled, shells are the important ones. These are known as the valence or bonding electrons and the shell is called the valence shell. If the atom has a full complement of electrons in its outermost shell, then it shows little inclination to compound formation and for this reason the rare gases, helium, neon, and argon, form no true chemical bonds. Krypton and xenon can form compounds if enough energy is supplied to break into the filled shell and promote some of its electrons to a higher energy level, whereupon bond formation becomes feasible. Even so, only fluorine and oxygen can provide bonds with sufficient energy to do this.

There are three kinds of strong chemical bond – ionic, covalent, and metallic. The part played by the bonding electrons in each of these is quite different. In ionic (or electrovalent) bonding the electrons in the valence shell of a metal atom are transferred to the valence shell of the non-metal atom. In other words, the bonding electrons become associated with one of the atoms and charge separation occurs. The coulombic attractions between the ions so formed are the real basis of the binding energy of ionic compounds.

$$Na + \tfrac{1}{2}Cl_2 \rightarrow Na^+ + Cl^-$$

The structure of ionic compounds is determined mainly by the physical parameters of ion size and charge. The result is a lattice of alternate metal cations and non-metal anions in which there is no particular cation associated with a particular anion. This means that there is no particular direction to the bonding – the ions bond equally to neighbouring ions in all directions.

In covalent bonding the bond is said to be directed. The bonding electrons associated with two atoms occupy the region of space between them. The result of this is that the internuclear repulsion is lowered and the nuclei can approach one another more closely. This is represented in terms of energy in figure 2.4.1. Most covalent bonds are of this type, that is, the bonding electrons are associated with two atoms. There are, in addition, some examples of three-atom covalent bonds in which a pair of bonding electrons serves to hold three nuclei together. These bonds are called three-centre bonds. This ability of electrons to hold several nuclei together reaches its ultimate form in metallic bonding.

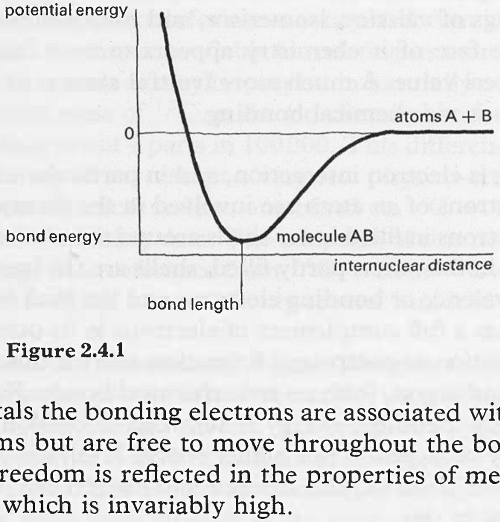


Figure 2.4.1

In metals the bonding electrons are associated with no particular atom or group of atoms but are free to move throughout the body of the metallic compound. This freedom is reflected in the properties of metals such as electrical conductance, which is invariably high.

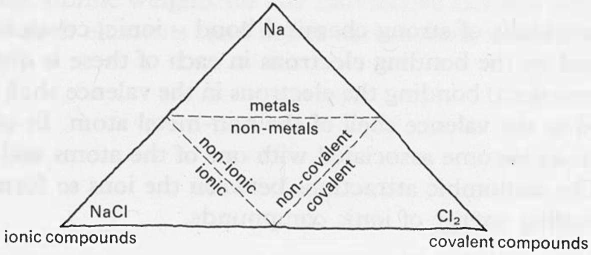


Figure 2.4.2

This method of classifying chemical compounds through their bonding can be illustrated by reference to figure 2.4.2. In this the apices of the triangle represent ideal forms of each kind of compound, but in reality these are not attained. No bond is purely ionic, covalent, or metallic. Because electrons are not fixed they can spend some of their time in a covalent bond with one of the atoms, so that in effect we have an equilibrium between covalent and ionic forms.

$$\underset{\text{covalent form}}{\mathrm{A{-}B}} \leftrightarrow \underbrace{\overset{+}{\mathrm{A}}\,\overset{-}{\mathrm{B}} \leftrightarrow \overset{-}{\mathrm{A}}\,\overset{+}{\mathrm{B}}}_{\text{ionic forms}}$$

Similarly, deformation of the electron clouds of ions can produce a covalent contribution to the bonding in ionic compounds.

Although we do not have ideal types of each compound we do have clearly recognizable examples of each class, and definite physical properties by which they can readily be recognized. Thus, ionic compounds have high melting points, low coefficients of expansion, are strong and hard in the solid phase, and are insulators as solids but good conductors (via ion mobility) in the melt or in solution. Covalent compounds have low melting points, low coefficients of expansion, are weak and soft in the solid phase, and are good insulators in both liquid and solid phases. Some covalent compounds, notably the macromolecules, have high melting points and are strong and hard. Metallic compounds generally have high melting points (although there are some notable exceptions such as mercury), large coefficients of expansion, and are strong and hard in the solid phase which also conducts electricity, in this case via electron mobility.

We shall now look at each class of chemical compound more closely.

Ionic (electrovalent) compounds

The best known ionic compound is sodium chloride (melting point 1081 K, boiling point 1738 K) and it has a cubic structure. In the crystal each sodium cation (and each chloride anion) is surrounded by six ions of the opposite charge. The lattice of this compound, and others, can be exactly determined by the use of X-ray crystallography. This technique enables the positions of the atomic nuclei to be plotted but tells us nothing about the distribution of the electrons. How, then, can we be certain of the existence of ions in these compounds? Indeed, no *direct* evidence of their presence in the solid state is available.

The strongest *indirect* evidence for ions comes from the conductance studies of molten salts and solutions. Their presence in the crystal state is not doubted and electron density studies have shown that virtually complete electron transfer from metal to non-metal atom has occurred. Internuclear distances enable one to deduce the radii of the ions, assuming that these are spherical entities. The potassium cation, K^+, has a radius of 133 pm and the chloride ion, Cl^-, has one of 181 pm. This also illustrates the effect of the ionic charge on the radius (both ions have 18 electrons). The smallest cation is the beryllium ion, Be^{2+}, 30 pm, and the largest is the caesium ion, Cs^+, 167 pm. For anions the smallest is the fluoride ion, F^-, 133 pm, and the largest is the tellurium ion, Te^{2-}, 222 pm. The size and charge of ions are what determines the amount of covalent character of the ionic compound.

Some salts have lower melting points than one would expect of an ionic compound. For example, beryllium chloride, $BeCl_2$, melts at 683 K whereas magnesium chloride, $MgCl_2$, does so at 987 K. The lower melting compounds are the ones with appreciable covalent character. The factors which favour this are (a) large charge upon either cation or anion, and (b) small cation or large anion. Ion deformation or polarization is the key to this behaviour. When ions approach one another the cation will attract the electrons of the anion and repulse its nucleus. The result is to increase electron density between the two atoms as we have in a covalent bond.

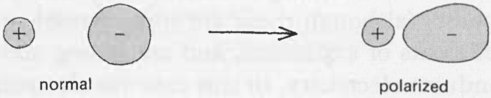


Figure 2.4.3

Since anions are generally much larger than cations it is anion polarization which is the principal factor, but cation polarization must also contribute. Small and highly charged cations have inherently greater polarizing ability, as one would expect. Large and highly charged anions are similarly most easily polarized. Thus compounds containing the ions Li^+, Be^{2+}, Al^{3+}, Sn^{4+}, etc., and Br^-, I^-, S^{2-}, etc., inevitably show physical properties which are out of step with those expected of a typical ionic compound.

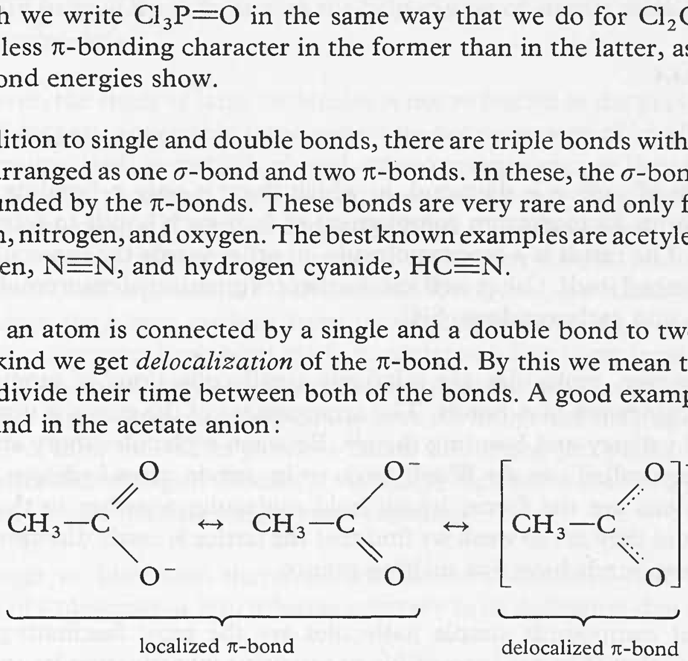
Covalent compounds

There are two kinds of covalent bond: sigma (σ) and pi (π). Both involve the sharing of electron pairs between two atoms. If only one pair of electrons is involved, then it will occupy the space between the two nuclei and this is called the σ-bond. This is the strongest type of covalent bond. In such a bond the distance between the two atoms is shorter than in any other type of bonding. We can illustrate this by comparing the bond length of chlorine, Cl_2, 199 pm, with the internuclear distance in KCl between K^+ and Cl^-, 314 pm, and with the shortest distance between the nuclei in potassium metal, 451 pm. Of all covalent bonds the shortest is that of the H_2 molecule, 75 pm.

In some molecules the formation of the covalent bond involves the sharing of two electron pairs. One of these is the strong σ-bond, lying between the atoms, and the other pair must of necessity occupy the region of space adjacent to the σ-bond. It is thus less effective at reducing internuclear repulsions and its contribution to the strength of the bond correspondingly less. This is the π-bond. Together the σ and one π-bond make up a double bond. Only the atoms carbon, nitrogen, and oxygen form true double bonds. Other elements, such as boron, fluorine, silicon, phosphorus, and sulphur, are also capable of extrabonding (that is, ‘extra’ to the basic σ-bond) but these are of a much less well authenticated type. Thus, although we write $Cl_3P{=}O$ in the same way that we do for $Cl_2C{=}O$, there is much less π-bonding character in the former than in the latter, as bond lengths and bond energies show.

In addition to single and double bonds, there are triple bonds with three electron pairs arranged as one σ-bond and two π-bonds. In these, the σ-bond is effectively surrounded by the π-bonds. These bonds are very rare and only found between carbon, nitrogen, and oxygen. The best known examples are acetylene, $HC{\equiv}CH$, nitrogen, $N{\equiv}N$, and hydrogen cyanide, $HC{\equiv}N$.

When an atom is connected by a single and a double bond to two atoms of the same kind we get *delocalization* of the π-bond. By this we mean that the π electrons divide their time between both of the bonds. A good example of this is to be found in the acetate anion:



Delocalization is not limited to only two bonds and in benzene it embraces all the carbon atoms of the ring, giving to this molecule its 'aromaticity' and concomitant properties. In graphite, which is made up of layers of fused hexagonal carbon rings, the π-bonding is delocalized over the whole layer. Because of this, graphite will conduct electricity but much more readily in the direction of the layers – a remarkable demonstration of the delocalization.

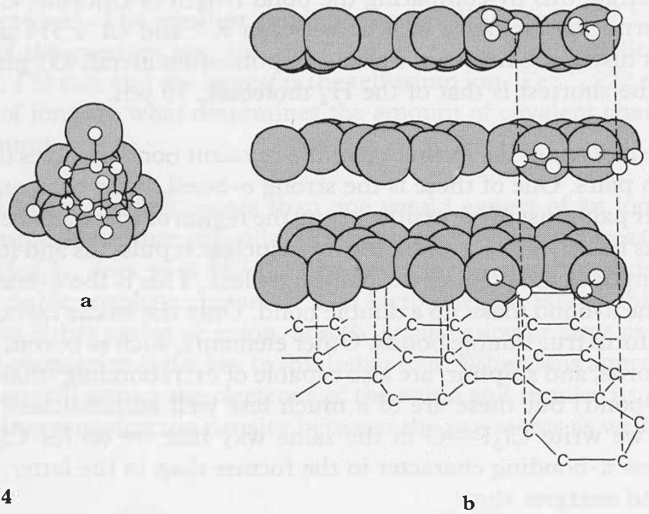


Figure 2.4.4
a Diamond.
b Graphite.

The other form of carbon is diamond, in which there is only σ-bonding. Each carbon atom forms its maximum complement of four such bonds to four other carbon atoms. The result is a macromolecule, in other words the molecule is as large as the diamond itself. Other well known three-dimensional macromolecules are glass, SiO_2 and carborundum, SiC.

Generally, however, molecules are relatively small collections of atoms held together by a framework of σ-bonds. The arrangement of the atoms is governed by the rules of valency and bonding theory. Between molecules there are only weak interactions called *van der Waals forces* or in certain cases *hydrogen bonds*. These interactions are the forces which hold molecules together in the solid state, but because they are so weak we find that the lattice is easily disrupted and hence these compounds have low melting points.

Of all chemical compounds simple molecules are the most fascinating. The reason for this is that they are susceptible to intensive investigation by spectroscopic techniques such as microwave, infra-red, and nuclear magnetic resonance (see Part 9 of this *Sourcebook*). By these methods the bond lengths, bond angles, and the energy of the molecule can be measured exactly. The distribution of electron density within the molecule can be studied, and thereby predictions and explanations about its reactions can be made. The study of matter at the molecular level is the key to all knowledge of the 'materials' world. And as chemists have mastered simple molecules and perfected their techniques and theories on these, so they have moved to larger and larger molecules. As the number of atoms in a molecule increases so does the difficulty involved in its study. This can be illustrated very neatly by reference to one aspect, isomerism.

In principle isomers are possible when a molecule has only three atoms*, such as A—B—A and A—A—B, but examples are as yet unknown. In organic chemistry, where isomers play a very important role, different types of isomerism are found, but the one which best illustrates the complexity is that of the aliphatic hydrocarbons. The number of theoretically possible isomers (without including stereoisomers) rises with the number of carbon atoms in the molecule. Thus CH_4 has one isomer and so has C_2H_6 and C_3H_8. The number of isomers then rises sharply: C_4H_{10} (2 isomers), C_5H_{12} (3), C_6H_{14} (5), C_7H_{16} (9), C_8H_{18} (18), C_9H_{20} (34), $C_{10}H_{22}$ (75), and so on. Concurrent with the rise in the number of isomers is the difficulty of their differentiation and investigation. (You may care to verify some of these facts with the help of a set of atomic models and a copy of the *Book of data*.)

However, the study of large molecules is not so fearful as the previous paragraph would suggest. Invariably large molecules are composed of smaller well-known sub-groups, such as methyl, phenyl, amino groups, etc., or they tend to be built from a basic repeating unit. The latter are known as *polymers*. Polythene, or polyethylene, as it is more correctly called, is built solely from the $—CH_2—$ unit, and most synthetic polymers are likewise very simple. Naturally occurring polymers such as proteins, DNA, animal fibres, etc., are 'naturally' more complex since they have evolved from more complex starting materials. Even so, dramatic advances have been made in understanding these large molecules. To speak of them as covalent chemical compounds, while strictly correct, is to trivialize them, since macromolecules have all-important macroproperties which a study of their individual covalent bonds would neglect. Such polymer properties as tensile strength, elasticity, viscosity, etc., are not directly related to the covalent bonds holding the polymer together.

Although we have seen that a π-bond can be delocalized over more than two atoms of a molecule or ion, it seems contrary to its definition that a σ-bond could

---

* Isotope isomers are even possible with two atom molecules, for example, H—H, H—D, D—D, H—T, D—T, and T—T.

be so delocalized. It thus comes as a surprise to find that in a compound like diborane, B_2H_6, just such a bonding arrangement is present. In this molecule the bridging hydrogen atoms and the boron atoms are held together by only one electron pair. These three-centre bonds, as they are called, are just as strong as conventional two-centre, that is, normal covalent bonds. They tend to occur only in compounds of electron deficient elements, such as boron, beryllium, etc. (See figure 2.4.5.)

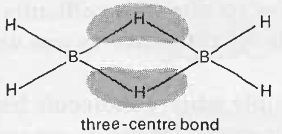


Figure 2.4.5

In a way we can see the three-centre bond as something intermediate between the normal covalent bond and the multi-centre bond which is in essence what we have in metallic bonding.

Metallic bonding

Three-quarters of the elements of the periodic table are metals, and the combinations of these in the formation of alloys is very large. There are two features of metals which account for the nature of their bonding: (1) metals have low ionization energies, in other words the atom has relatively less attraction for its valence electrons than, say, a non-metal; and (2) the number of valence electrons is less than the number of valence orbitals. The result of these two characteristics is that when metal atoms form a bond they are ready to share their valence electrons, and also have empty valence orbitals ready to receive other electrons. The result is essentially complete delocalization of all the valence electrons over all atoms.

On one model of chemical bonding we can see bonding in metals as one end of a spectrum ranging from ionic (that is, one-centre) bonding through covalent (that is, two-centre, three-centre) to multi-centre metallic bonding.

Metallic bonding gives the electrons a larger degree of freedom than found in other chemical compounds, so much so that some authors of textbooks refer to cations of the metal being in a 'sea of electrons'. This explains the exceptional electrical conductance associated with this type of bonding. The electrons possess similar energies and occupy molecular orbitals which are also very similar. Indeed, these energy levels form a *band* and any electron in one of them becomes the property of the crystal as a whole and serves to bind together many nuclei.

Metal atoms in the solid state tend to pack into one of three lattice structures. These are a body-centred cubic (b.c.c.), a hexagonal close-packed (h.c.p.), or a cubic close-packed (c.c.p.) arrangement. In the first of these the coordination number of each atom is eight, in the other two it is twelve. The alkali metals are all b.c.c., the noble metals (copper, silver, gold) are c.c.p. Some metals are polymorphic such as calcium which can be b.c.c. or c.c.p. The packing of any solid spheres (such as billiard balls) leaves spaces or holes in the arrangement. So in the packing of atoms one might also imagine similar holes. There are two types of hole (or interstice) – tetrahedral holes with four surrounding atoms, and octahedral holes with six surrounding atoms. The former have a diameter of only 0.225 that of the surrounding spheres, but the latter have a diameter of 0.414 of the surrounding spheres and can thus accommodate other small atoms. When a metal does this it forms *interstitial compounds*. (Additional features of the metallic state are discussed in Part 10.)

Atoms of hydrogen, nitrogen, boron, oxygen, and carbon can be incorporated fairly easily in metal lattice holes. Not all the holes which can be filled are occupied and so these compounds tend to have non-stoichiometric formulae, such as $ZrH_{1.9}$, although in most cases a definite proportion of the holes are filled and formulae of the type V_2C (half the octahedral holes occupied), and Fe_4N (a quarter of the octahedral holes filled) are found.

With these small atoms only a little distortion of the metal lattice occurs so that interstitial compounds generally have the same bulk properties as the metals themselves; for example, they possess high melting points, hardness, etc. It is worth noting that the exact nature of the bonding is not clear.

In summary, we find that our senses tell us that the world is composed of an infinite variety of compounds. However, closer inspection shows that things are not so complex as they appear and that all compounds can be classified on the basis of their chemical bonding into one of three types. We have briefly reviewed the characteristics of each type and noted some of the subdivisions in each class. Even so, it should be remembered that we classify compounds for our convenience and understanding. The onus is on the chemist to devise classes which correspond to natural divisions and not upon nature to provide us with examples of compounds which fall neatly into our choice of classes. Classification is a disruptive operation being applied to what is essentially a continuous variable. Our present way of classifying compounds, based as it is on the electron and the bond, seems well founded, but we must always bear in mind that other classifications have been used in the past and others might eventually supersede our own. As this article has shown, there are compounds which fit into one of our three classes only very uncomfortably. These may turn out to be the signpost to a future theory.

2.5 Some uses of complexing agents

by H. Derbyshire

A complexing agent is a substance which will combine with metal ions in solution forming either a soluble complex or a precipitate. The soluble complex might be a cation, an anion, or a neutral molecule.

$$Cu^{2+} + 4NH_3 \rightleftharpoons [Cu(NH_3)_4]^{2+}$$
$$Fe^{3+} + 6CN^- \rightleftharpoons [Fe(CN)_6]^{3-}$$
$$Cr^{3+} + 3SCN^- + 3H_2O \rightleftharpoons [Cr(SCN)_3 3H_2O]^0$$

Usually when a complex is formed, the concentration of the free (that is, hydrated) metal ion is considerably reduced; for example,

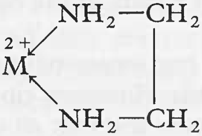
$$Ag^+ + 2CN^- \rightleftharpoons [Ag(CN)_2]^- \quad K = 10^{21}$$

Thus, if equal volumes of 0.1 M Ag^+ and 1 M CN^- are mixed together, the Ag^+ concentration is reduced to below 10^{-20} M. When metal ions are removed from solution in this way, they are sometimes said to be *sequestered.*

The group which attaches itself to the metal ion is called a *ligand.* Examples of ligands are OH^-, NH_3, H_2O, Cl^-, I^-, and many organic substances. Sometimes ligands have more than one point of attachment:

$$\overset{2+}{M}\leftarrow NH_3$$

one point of attachment
(monodentate)



two points of attachment
(bidentate)

and this has led to the terms bidentate ligand, tridentate ligand, polydentate ligand. For example, 1,2-diaminoethane ($NH_2CH_2CH_2NH_2$) is a bidentate ligand; it forms stable complexes with many metal ions.

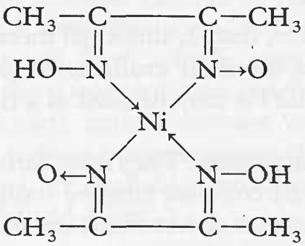
With polydentate ligands, metals form complexes which possess a ring structure and which are very stable because of this. Complexes with ring structures are called *chelates* (because of their claw-like structure) and are very important, especially in biological systems. Examples of biologically important chelates are haemoglobin (containing iron), chlorophyll (containing magnesium), and vitamin B_{12} (containing cobalt).

To illustrate the stability of chelates, sodium hydroxide solution may be added in turn to solutions containing (1) $[Ni(NH_2CH_2CH_2NH_2)_3]^{2+}$ (a chelate), and (2) $[Ni(NH_2CH_3)_6]^{2+}$ (not a chelate). With solution (1), no precipitate is formed but solution (2) gives a precipitate of nickel hydroxide.

Some uses of complexing agents in analytical chemistry

Gravimetric analysis

There are several reasons why it may be useful to precipitate a metal as a complex. Metal complexes often have very low solubilities and so the error due to ions remaining in solution is very low. The metal is also part of a complex of high molecular weight which reduces the effect of errors in weighing and due to loss of precipitate. But one of the most important features of this kind of precipitant is that, by adjusting the pH of the solution, it can often be made specific for only one metal. Dimethylglyoxime, for instance, in dilute ammonia solution, forms complexes with many metals but all are very soluble except for the nickel complex which precipitates.



The nickel–dimethylglyoxime complex

Colorimetric analysis

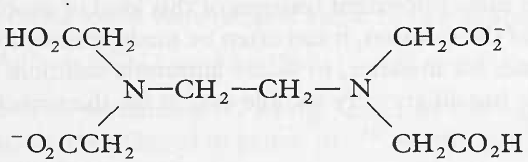
In section 4 of the physical science course, experiment 407 illustrates the use of EDTA and 1,2-diaminoethane ions in determining concentrations of metal ions in solution by means of a colorimeter.

Many metal complexes absorb strongly, and often at characteristic wavelengths, in the visible region of the spectrum. Either by using a filter or by means of a monochromator, that is, a prism or diffraction grating, the concentration of one metal ion in the presence of many others may be found.

The method finds extensive use in industry because it is both rapid and reasonably accurate and is easily adapted to many of the routine determinations carried out in a quality control laboratory. For example, thioglycollic acid may be used to check that the iron content of a commercial sample of sodium carbonate is below 0.001%. The thioglycollic acid–iron complex has a strong red colour and the colour produced by the sample may be compared with that produced by a standard sample of sodium carbonate which has an iron content of exactly 0.001%.

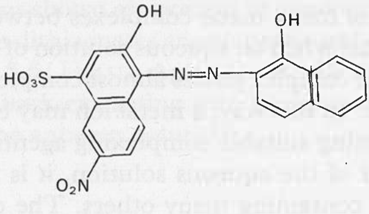
Complexometric analysis

This is the titration of a metal ion with a suitable complexing agent. EDTA (ethylenediamine tetraacetic acid), as its disodium salt, is probably most widely used:

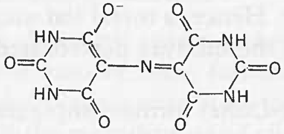


Complexes formed with EDTA are very stable, that is, almost all metal ions are completely removed from solution. If some means is available for indicating when all the metal ions are removed, then EDTA may be used as a titrant.

Other complexing agents are usually used as indicators. They must form strongly coloured complexes with the metal ion and the complex must be stable enough to be formed whenever a very small concentration of free metal ion is available. The metal–indicator complex must, of course, be less stable than the metal–EDTA complex so that when EDTA is in slight excess, all of the other complex will have disappeared. Many dyestuffs are used as indicators, for example, Eriochrome black T for magnesium, calcium, manganese; Pyrocatechol violet for bismuth, thorium; Murexide for calcium, nickel, copper, etc. (See figure 2.5.1.)



Eriochrome Black T



Murexide ion

Figure 2.5.1

Complexing agents for water softening

Hard water contains calcium or magnesium ions which form insoluble compounds with soap (the 'scum' left when the bath water is drained away). Many substances will form soluble complexes with calcium and magnesium ions, removing them from solution and 'softening' the water. Few have been used as water softeners mainly because of their expense, but the polyphosphates, especially sodium pyrophosphate ($Na_4P_2O_7$) and the cyclophosphates ('meta' phosphates), are available commercially (for example, 'Calgon' manufactured by Albright and Wilson).

Solvent extraction

If a solute is added to two immiscible solvents, A and B, in contact with each other, the solute distributes itself between the two and an equilibrium is set up between the solute molecules in solvent A and the solute molecules in solvent B.

$$\text{solute in A} \rightleftharpoons \text{solute in B}$$

The ratio of the concentration in A to that in B is known as the *partition coefficient*. As one would expect from the equilibrium law, it is constant provided the solute, the two solvents, and the temperature are also constant.

The partition coefficients of many metal complexes between water and organic solvents are very high so that when an aqueous solution of the complex is shaken with an organic solvent, the complex passes almost completely from the aqueous phase to the organic phase. In this way, a metal ion may be 'extracted' from an aqueous solution. By choosing suitable complexing agents and organic solvents and by controlling the pH of the aqueous solution, it is possible to separate a metal ion from a solution containing many others. The complexes are usually strongly coloured and this allows the concentration in the organic layer to be determined photochemically. Hence, a metal ion may be extracted from a mixture and its concentration in the mixture determined.

Dithizone (diphenyl thiocarbazone) forms complexes with many metals, which are soluble in organic solvents such as chloroform (see figure 2.5.2).

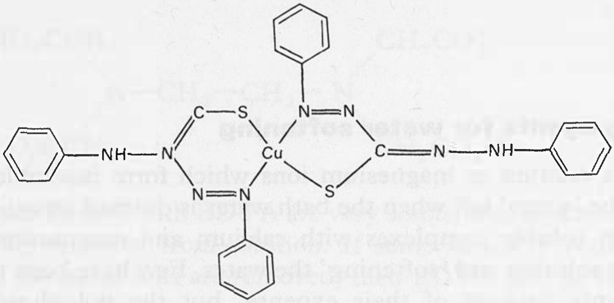


Figure 2.5.2
The dithizone-copper complex.

The graph in figure 2.5.3 shows how the pH may be adjusted to enable certain ions to be separated from the rest.

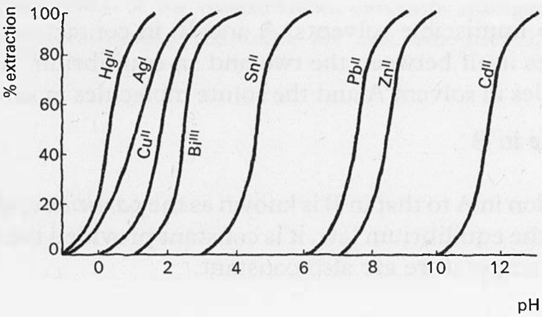


Figure 2.5.3

Suppose that all the ions shown are present in aqueous solution and it is required to separate Sn(II). The dithizonates are prepared and extracted with chloroform from a solution of pH 5.5; Hg(II), Ag(I), Cu(II), Bi(III), and Sn(II) dithizonates will be extracted. On back-extracting with an aqueous solution of pH 3.0, only Sn(II) will pass into the aqueous solution.

The electrodeposition of metals

The fact that metals can be deposited onto a cathode by electrolysis of suitable salt solutions has formed the basis of many industrial processes. Examples are the refining of copper, nickel plating onto steel to prevent corrosion, the preparation of rare metals, and the manufacture of silver plate.

In all these processes, it is important to obtain a metallic deposit which is uniform and does not flake off either during electrolysis or during subsequent use. An electrolytic solution containing a simple salt of the metal does not normally give a suitable deposit; it usually consists of non-adherent coarse particles. The addition of a complexing agent to the electrolytic solution often improves the deposit. An example is the use of cyanide in silver plating.

$$Ag^+ + 2CN^- \rightleftharpoons [Ag(CN)_2]^-$$

The factors which control the nature of the metal deposit are complex and not yet fully understood. Experience shows that high current densities and low non-uniform ion concentrations should be avoided. The presence of colloidal organic matter, for example, gelatine, camphor, or casein, is often helpful, perhaps by preventing high current densities near the surface of the cathode. Metal complexes are often colloidal and may be useful because of this, but probably a more important property is the equilibrium set up between the complex and free metal ions. This will help to avoid local variation in ion concentration near the cathode.

Complexes also have an effect on electrolysis in another way; by reducing the metal ion concentration, the electrode potential changes. Complexes may thus be used to control electrode potentials allowing metals to be selectively deposited. Electrode potentials may also be so adjusted that two metals are deposited simultaneously and such a process finds application in the manufacture of alloys.

For example, the data provided below illustrate how the electrode potentials of zinc in 0.1 M $ZnSO_4$ and copper in 0.1 M $CuSO_4$ may be brought close to each other by adding cyanide.

	With no KCN	**In 0.2 M KCN**	**In 1.0 M KCN**
Zn	−0.816 V	−1.033 V	−1.231 V
Cu	+0.292 V	−0.611 V	−1.169 V

The extraction of metals from ores

Complexing agents have been used in the extraction of metals from their ores. A good example is the use of cyanide in the extraction of gold from gold-bearing quartz. The powdered quartz is agitated with a solution of sodium cyanide and the gold passes slowly into solution:

$$4Au + 8CN^- + 2H_2O + O_2 \rightarrow 4Au(CN)_2^- + 4OH^-$$

The gold is then precipitated by adding zinc dust to the solution and the excess zinc dust is removed through reaction in dilute acid.

With sulphide ores (for example, argentite):

$$Ag_2S + 4CN^- \rightleftharpoons 2Ag(CN)_2^- + S^{2-}$$

With chloride ores (for example, horn silver):

$$AgCl + 2CN^- \rightleftharpoons Ag(CN)_2^- + Cl^-$$

Again, zinc dust is used to precipitate metallic silver and the excess zinc is removed by dilute acid.

Dyes and pigments

Metal complexes are often very strongly coloured, very stable, and sometimes very insoluble in most solvents. These properties make them useful as dyes and pigments in such materials as plastics, paints, printing inks, etc. Among the metals whose complexes have been used in this way are iron, copper, chromium, cobalt, and nickel which, together, give a large range of colours.

The most important group of metal complexes used as pigments are the phthalocyanines. In 1928, Scottish Dyes Ltd (now part of ICI) noticed that during the manufacture of phthalimide from phthalic anhydride and ammonia using iron vessels, an impurity was formed which had a very strong blue colour. This was iron phthalocyanine; it was separated and found to have very good properties as a pigment. Later, the copper derivative was found to be an even better blue pigment, which is now marketed as Monastral Blue (figure 2.5.4).

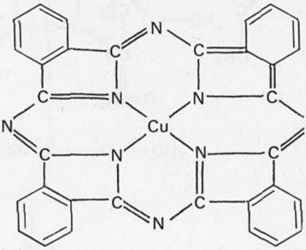


Figure 2.5.4
Monastral Blue.

Complexing agents in biology and medicine

Metal complexes are quite common in biological systems. They are similar to the complexes discussed previously except that the ligands usually have a very high molecular weight and consist mostly of protein. Even the phthalocyanine molecules are quite small in comparison.

Iron, magnesium, and copper are the commonest metals found in these complexes but other metals such as cobalt, zinc, manganese, molybdenum, and vanadium are often used. It is easy to see why trace metals are important in the diet of animals and plants and why the absence of these metals may cause deficiency diseases.

The complexes have a wide variety of uses. They control metal ion concentrations in blood and tissues, act as transporters of metal ions and other substances such as oxygen and carbon dioxide, and they are used to store metal ions in certain organs. Some are important catalysts (enzymes) and play a vital role in metabolism whilst others serve only as pigments.

In spite of these widely differing functions and the complexity of their molecules, it is surprising how similar in molecular structure these substances are. Haemoglobin, found in red blood cells, carries oxygen from the lungs to the tissues and carbon dioxide back to the lungs. Its molecule consists of four protein chains, each enfolding a haem group (see figure 2.5.5).

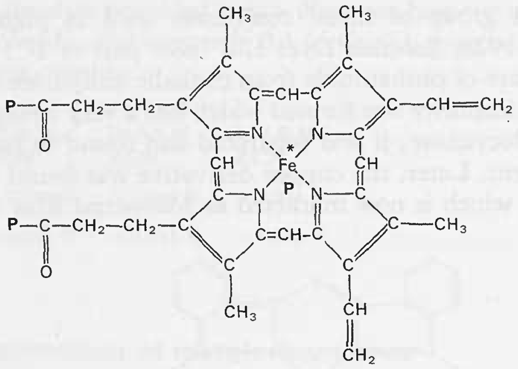


Figure 2.5.5

Haem.

The protein is attached at the points marked **P** and either water or oxygen at *. Iron is in oxidation state II but its behaviour is quite different from that of an Fe^{2+} ion. The effect of this enormous molecule is to modify the properties of the iron so that it easily takes up oxygen (in the lungs) and easily releases it (in the tissues).

Haemoglobin also plays a part in transporting carbon dioxide from the tissues to the lungs. Some of this carbon dioxide (about twenty per cent) becomes attached to certain sites on the protein part of haemoglobin forming 'carbaminohaemoglobin' while the rest is carried as HCO_3^-, distributed evenly through the red blood cells and plasma.

Myoglobin is found in muscle, where its function is to release oxygen to the muscle when required. Structurally it is very similar to haemoglobin but consists of only one haeme group and one protein chain.

Chlorophyll, the green pigment found in plants, is used by them in photosynthesis. It absorbs sunlight and makes this energy available for building up carbohydrates from carbon dioxide and water. Chlorophyll is a mixture of two very similar substances, chlorophyll a and chlorophyll b (see figure 2.5.6).

It is remarkable how this molecule resembles the haem group. Only the side chains are different and the iron atom has been replaced by magnesium.

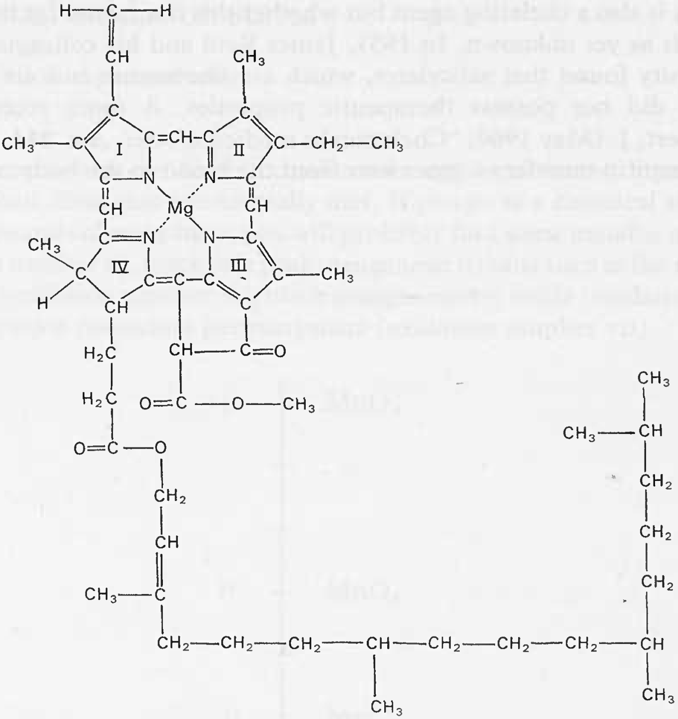


Figure 2.5.6

Chlorophyll a. Note: chlorophyll b differs in having a formyl group in place of a methyl group in ring II.

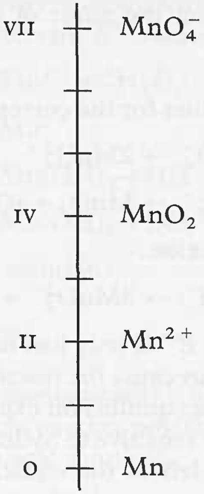
Because so many biological processes make use of metal chelates, it is not surprising that large concentrations of metal ions are very poisonous, both to plants and animals. Metals such as lead, arsenic, antimony, copper, etc. would interfere either by replacing the metal in natural chelates or by forming complexes with biologically active substances. In 1951, the usefulness of EDTA in treating cases of metal poisoning was first demonstrated when it was used to save the life of a three-year-old boy suffering from acute lead poisoning. The EDTA complexed the Pb^{2+} ions and effectively removed them from solution. Since then, chelating agents have also been used in cases of iron poisoning (usually in small children who have eaten large quantities of 'iron' tablets containing iron(II) sulphate [ferrous sulphate]), and poisoning by arsenic, antimony, mercury, nickel, etc. They have also been used to treat skin irritations caused by such metals as copper, chromium, and nickel and for cases of poisoning with radioactive substances such as strontium-90 or plutonium.

Aspirin is also a chelating agent but whether this is a reason for its effectiveness or not is as yet unknown. In 1951, James Reid and his colleagues at Glasgow University found that salicylates, which are like aspirin but are not chelating agents, did not possess therapeutic properties. A more recent suggestion (Schubert, J. (May 1966) 'Chelation in medicine', *Sci. Am.* **214**, No. 5, 40–50) is that aspirin transfers copper ions from the blood to the body cells.

2.6 Unusual oxidation states

by A. W. B. Aylmer-Kelly

If the right conditions can be found, elements can be obtained in oxidation states other than those that are normally met. If you go to a chemical store and look for compounds of *manganese*, you will probably find some metallic manganese (oxidation number O), some pale pink manganese(II) salts such as the sulphate and chloride (oxidation number II), black manganese(IV) oxide (oxidation number IV), and purple potassium permanganate (oxidation number VII).



It is of interest to speculate whether it is feasible to obtain manganese in other oxidation states such as I, III, V, or VI.

Reverse disproportionation is one process that may achieve appropriate conditions. For example, iodate reacts with iodide to give iodine:

$$I(V)+I(-I) \rightarrow I(O)$$

or $$6H^+ + IO_3^- + 5I^- \rightarrow 3I_2 + 3H_2O$$

and one might argue that Mn(VII) may react with Mn(IV) to give Mn(V) or Mn(VI).

The $E^\ominus$ values for such changes in acid solution are:

$$2e^- + 2MnO_4^- \rightarrow 2MnO_4^{2-} \qquad E^\ominus = +0.56\ V$$

or

$$Mn(\text{VII}) \rightarrow Mn(\text{VI})$$

and

$$2e^- + 4H^+ + MnO_4^{2-} \rightarrow MnO_2 + 2H_2O \qquad E^\ominus = +2.26\ V$$

or

$$Mn(\text{VII}) \rightarrow Mn(\text{IV})$$

and indicate that the reaction:

$$2MnO_4^- + MnO_2 + 2H_2O \rightarrow 3MnO_4^{2-} + 4H^+ \qquad E^\ominus = -1.70\ V$$

is not possible.

One might now speculate about the effect of changing the concentration of hydrogen ions in this reaction.

In alkaline solution, the $E^\ominus$ values for the corresponding changes are different:

$$2e^- + 2MnO_4^- \rightarrow 2MnO_4^{2-} \qquad E^\ominus = +0.56\ V$$

$$2e^- + 2H_2O + MnO_4^{2-} \rightarrow MnO_2 + 4OH^- \qquad E^\ominus = +0.6\ V$$

thereby indicating that the reaction:

$$2MnO_4^- + MnO_2 + 4OH^- \rightarrow 3MnO_4^{2-} + 2H_2O \qquad E^\ominus = -0.04\ V$$

is not possible. However, since $E^\ominus$ is only just negative, a change of conditions from the standard conditions *may* cause the reaction to occur. The concentration of hydroxide ions appears in the equilibrium expression to the fourth power, so the equilibrium position is very sensitive to hydroxide ion concentration. As the hydroxide ions appear on the left of the equation above, this means that an increase in the hydroxide ion concentration may make the reaction possible. If a solution of permanganate ions is made alkaline with sodium hydroxide solution, and solid manganese dioxide added and the tube shaken, a green solution of manganate (Mn(VI)) ions is formed.

It is not possible to make manganese(V) by reverse disproportionation as inspection of the $E^\ominus$ values shows:

$$e^- + 2H^+ + MnO_4^{2-} \rightarrow \underset{Mn(\text{V})}{MnO_3^-} + H_2O \qquad E^\ominus = +2.0\ V$$

$$e^- + 2H^+ + MnO_3^- \rightarrow MnO_2 + H_2O \qquad E^\ominus = +2.5\ V$$

$$e^- + H_2O + MnO_4^{2-} \rightarrow MnO_3^- + 2OH^- \qquad E^\ominus = +0.34\ V$$

$$e^- + H_2O + MnO_3^- \rightarrow MnO_2 + 2OH^- \qquad E^\ominus = +0.84\ V$$

It should also be noted that a change of pH will not affect these reactions. (Why?)

Manganese(V) can be made by adding a little solid potassium permanganate to a *very* concentrated solution (12 M) of sodium hydroxide, the solution slowly becoming blue as Mn(V) is formed. Close observation of the solution will show that a gas is also being formed. This gas proves to be oxygen, and hydroxide ions are being oxidized to oxygen molecules:

$$2e^- + H_2O + MnO_4^- \rightarrow MnO_3^- + 2OH^- \qquad E^\ominus = +0.34\ V$$

$$4e^- + 2H_2O + O_2 \rightarrow 4OH^- \qquad E^\ominus = +0.4\ V$$

Many oxidizing agents can theoretically bring this oxidation about. (Why does it not happen in practice?)

Reverse disproportionation between manganese(II) and manganese(IV) may produce manganese(III). The relevant $E^\ominus$ values are:

$$e^- + MnO_2 + 4H^+ \rightarrow Mn^{3+} + 2H_2O \qquad E^\ominus = +0.95\ V$$

$$e^- + Mn^{3+} \rightarrow Mn^{2+} \qquad E^\ominus = +1.51\ V$$

$$e^- + 2H_2O + MnO_2 \rightarrow Mn(OH)_3 + OH^- \qquad E^\ominus = +0.2\ V$$

$$e^- + Mn(OH)_3 \rightarrow Mn(OH)_2 + OH^- \qquad E^\ominus = -0.1\ V$$

Is the reaction possible in acid solution?

Is it possible in alkaline solution?

The reaction between solid manganese(IV) oxide and solid manganese(II) hydroxide is, obviously, slow, but another possibility is to oxidize manganese(II) with manganese(VII). Thus, in acid solution,

$$e^- + Mn^{3+} \rightarrow Mn^{2+} \qquad E^\ominus = +1.51\ V$$

$$5e^- + 8H^+ + MnO_4^- \rightarrow Mn^{2+} + 4H_2O \qquad E^\ominus = +1.51\ V$$

and this reaction occurs when permanganate solution is added to a strongly acid solution of manganese(II) ions, a deep red colour being produced.

What effect will raising the pH of the solution have on this reaction?

All these unusual oxidation states are only stable in particular conditions such as high alkalinity. When these conditions are changed, they decompose into the more stable, more common oxidation states. One way in which oxidation states which are normally unstable in solution can be stabilized, is by forming a complex compound. Manganese(I) is not known as a hydrated Mn^+ ion, but if the hexacyanomanganate(II) ion, $Mn(CN)_6^{4-}$, is reduced with aluminium in an alkaline solution, the hexacyanomanganate(I) ion, $Mn(CN)_6^{5-}$, is formed. Oxidation states which do not form stable ions in solution sometimes occur in the solid state, and copper(I) occurs only in insoluble salts such as copper(I) chloride, $CuCl$, and sulphide, Cu_2S. Copper(I) sulphate, Cu_2SO_4, can be made as a solid, but decomposes in water to give copper(II) sulphate and copper metal. Copper(I) can exist in solution, but only in complexes such as $CuCl_2^-$, $Cu(CN)_2^-$, and $Cu(NH_3)_2^+$. Copper can even exist in the III oxidation state in the complex ion CuF_6^{3-}.

The stability of unusual oxidation states brought about by complexing is particularly well demonstrated by the tetracyanonickelate(0) ion, $Ni(CN)_4^{4-}$, containing Ni(0), the tetracyanocobaltate(0) ion, $Co(CN)_4^{4-}$, containing Co(0), the tricyanonickelate(I) ion, $Ni(CN)_3^{2-}$, containing Ni(I), and by the compound $[FeCl_2 2C_6H_4As(CH_3)_2]FeCl_4$ which contains iron with oxidation numbers of II and IV. The ferrates contain iron in the +VI state and can be made by oxidizing Fe(III) by hypochlorite (chlorate(I) ions) in alkaline solution:

$$2Fe(OH)_3 + 3OCl^- + 4OH^- \rightarrow 2FeO_4^{2-} + 3Cl^- + 5H_2O$$

Some oxidation states which cannot be produced chemically can be made electrochemically: silver(II) and cobalt(III) can be formed on a platinum anode:

$$Ag^+ \rightarrow Ag^{2+} + e^-$$

$$Co^{2+} \rightarrow Co^{3+} + e^-$$

These species are stable for a period of hours in aqueous solution.

2.7 Acid–base chemistry in the human body

by A. W. B. Aylmer-Kelly

The chemical processes which occur in the body are complex, and often present challenging problems to the biochemist. Some of the mechanisms are reasonably well understood, for instance, that by which the body regulates the pH value of the blood.

Although red to the naked eye, blood consists of a pale yellow aqueous fluid, the plasma, in which red and white corpuscles are suspended. The plasma contains proteins, albumen, fibrinogens, and globulins, and also ions such as Na^+, Cl^-, HCO_3^-, and H^+. The red cells contain haemoglobin, an enzyme known as carbonic anhydrase, and ions such as K^+, Cl^-, HCO_3^-, and H^+, also in aqueous solution. The chief functions of the blood are the transport of oxygen and food materials to the body tissues, and of carbon dioxide and other waste products of metabolism from the tissues to the lungs and kidneys, where they are excreted.

Air inhaled into the lungs finds its way into small air sacs, the alveoli, where it is brought into contact with blood across thin membranes. In man, there are about 7×10^8 alveoli, each about 2 μm in diameter. The blood arriving in the lungs is known as venous blood and it contains carbon dioxide collected during its passage round the body. Gaseous exchanges occur across the alveolar membranes, oxygen passing into the blood and carbon dioxide leaving it ready for exhalation. The oxygenated or arterial blood is pumped out of the lungs to circulate through the tissues, giving up oxygen to them and collecting carbon dioxide. The process is summarized in figure 2.7.1.

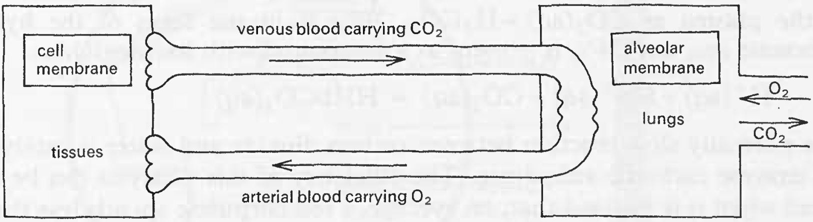


Figure 2.7.1
Transport of oxygen and carbon dioxide in the blood.

Transport of oxygen in the blood

Approximately 0.3 per cent of the oxygen in the blood is in solution, $O_2(aq)$, while the remainder is present as the red compound, oxyhaemoglobin. From Henry's law (a special case of the equilibrium law): $[O_2(aq)] = Kp_{alv}$, where p_{alv} is the partial pressure of oxygen in alveolar air.

Haemoglobin, Hb^-, is a complex organic molecule containing iron which normally carries a negative charge. One molecule combines with one of oxygen:

$$Hb^-(aq) + O_2(aq) \rightleftharpoons HbO_2^-(aq)$$

the molecular weights being such that 1 mol of oxygen combines with 16 700 g of haemoglobin.

The amount of oxygen entering the blood depends upon p_{alv} which in turn depends upon p_{atm}, the partial pressure of oxygen in the atmosphere. In a healthy adult, p_{alv} is about 13 kPa. The figures in the table reveal the necessity for breathing equipment in high altitude flight and mountaineering. (Ht = height above sea-level.)

Ht/km	0	3	6	15
p_{atm}/kPa	21	15	10	2.4

Transport of carbon dioxide in the blood

When carbon dioxide dissolves in water it reacts slowly to form carbonic acid:

$$CO_2(aq) \rightleftharpoons H_2CO_3 \rightleftharpoons HCO_3^-(aq) + H^+(aq)$$

The blood carries carbon dioxide in three ways; about 6% is present in solution in the plasma as $CO_2(aq) + H_2CO_3$, 70% is in the form of the hydrogen carbonate ion, and 24% is present as a compound with haemoglobin:

$$H^+(aq) + Hb^-(aq) + CO_2(aq) = HHbCO_2(aq)$$

The normally slow reaction between carbon dioxide and water is catalysed by the enzyme carbonic anhydrase. The efficiency of this catalysis can be appreciated when it is realized that, on average, a red corpuscle spends less than half a second in the alveolar membrane during its passage through the lung.

The formation of hydrogen carbonate ions in the blood is accompanied by the formation of H^+ ions; hence if the blood were not buffered, its pH would fall when carbon dioxide passes into it. The blood in fact contains a number of buffer systems which in healthy individuals maintain its pH at about 7.4.

Blood buffers

The plasma proteins and haemoglobin can act as buffers because they contain both acidic ($-CO_2H$) and basic ($-NH_2$) groups in their molecules, which can react with and hence 'absorb' added hydrogen or hydroxide ions.

However, the most important buffer system is:

$$CO_2(aq) \rightleftharpoons H_2CO_3 \rightleftharpoons HCO_3^-(aq) + H^+(aq)$$

The processes discussed are summarized in figure 2.7.2. (Reference is made to 'the chloride shift'. What do you think this is?)

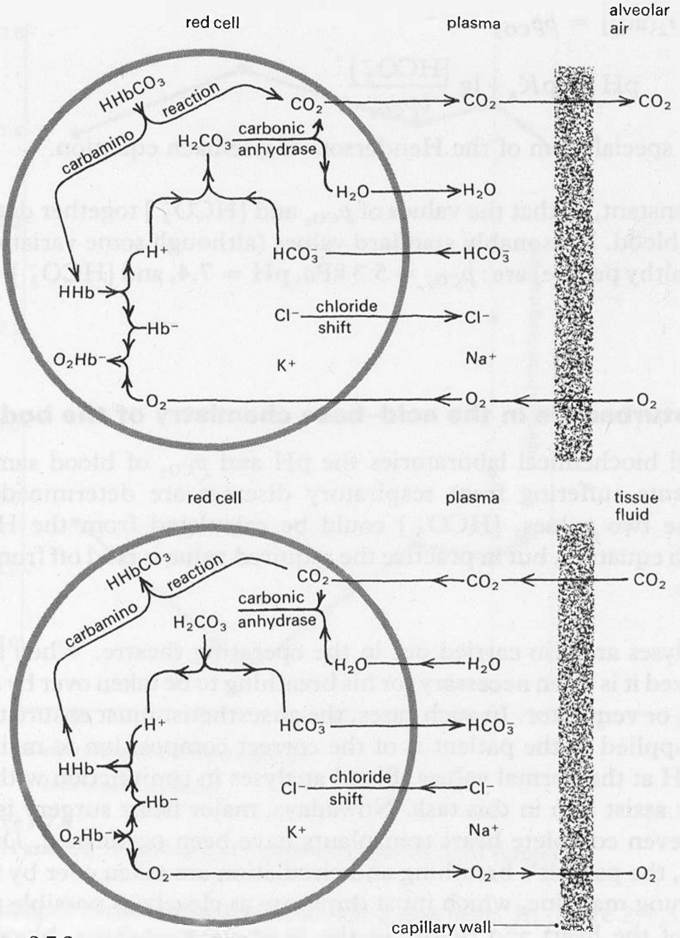


Figure 2.7.2

Transport of oxygen by haemoglobin. *After Roughton, F. J. W. (1935) 'Recent work on carbon dioxide transport by the blood.'* Physiol. Rev. **15,** *241–296.*

The pH of the blood (effectively that of the plasma) is related to the partial pressure of carbon dioxide in the alveolar air, p_{CO_2}, as follows.

For the equilibrium:

$$CO_2(aq) \rightleftharpoons H_2CO_3 \rightleftharpoons H^+(aq) + HCO_3^-(aq)$$

$$K_a = \frac{[H^+(aq)]_{eqm}[HCO_3^-(aq)]_{eqm}}{[CO_2(aq)]}$$

$$K_a = \frac{[H^+][HCO_3^-]}{bp_{CO_2}}$$

since $[CO_2(aq)] = bp_{CO_2}$

that is, $$pH = pK_a + \lg \frac{[HCO_3^-]}{bp_{CO_2}}$$

which is a special form of the Henderson–Hasselbach equation.

pK_a is a constant, so that the values of p_{CO_2} and $[HCO_3^-]$ together determine the pH of the blood. Reasonably standard values (although some variation is found even in healthy people) are: $p_{CO_2} = 5.3$ kPa, pH $= 7.4$, and $[HCO_3^-] = 25$ mmol dm^{-3}.

Disturbances in the acid–base chemistry of the body

In hospital biochemical laboratories the pH and p_{CO_2} of blood samples taken from patients suffering from respiratory diseases are determined regularly. From these two values, $[HCO_3^-]$ could be calculated from the Henderson–Hasselbach equation, but in practice the required value is read off from a suitable nomogram.

Blood analyses are also carried out in the operating theatre. When a patient is anaesthetized it is often necessary for his breathing to be taken over by a mechanical system, or ventilator. In such cases, the anaesthetist must ensure that the gas mixture supplied to the patient is of the correct composition to maintain p_{CO_2}, p_{alv}, and pH at the normal values. Blood analyses in conjunction with a suitable nomogram assist him in this task. Nowadays, major heart surgery is commonplace and even complete heart transplants have been performed. During such operations, the patient's breathing and circulation are taken over by a machine, the heart-lung machine, which must duplicate as closely as possible the normal functions of the heart and lungs. In the heart-lung machine, blood from the patient is oxygenated and carbon dioxide removed from it, so that pH and p_{CO_2} checks must constantly be made if correct conditions are to be maintained.

For example, a man suffering from severe bronchitis was admitted to the intensive care unit of a London hospital. He was unconscious, and blood samples gave the values $p_{CO_2} = 15$ kPa and pH $= 7.31$. The man's disease prevented him adequately ventilating his alveoli, so that p_{CO_2} had risen to the high value found. The buffer reserve of his blood had been exceeded and the plasma pH had fallen; the condition is known as respiratory acidosis. The cure in this case is relatively straightforward; the patient's lungs are ventilated mechanically, after the insertion of a tube in his trachea, with an oxygen-rich gas mixture. Whilst the ventilator maintains his breathing, the lung condition is treated with antibiotics until he is able to breathe normally again and maintain his own acid-base balance.

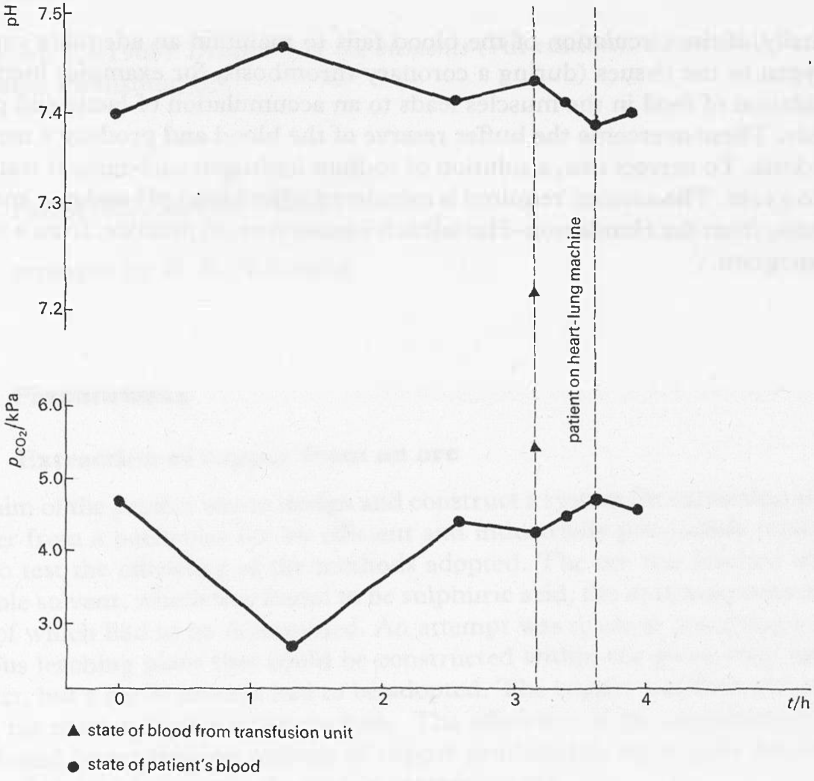


Figure 2.7.3
Variation in pH and p_{CO_2} of the blood of a patient undergoing heart surgery.

Another example is the use of the heart-lung machine. In figure 2.7.3 variations in pH and p_{CO_2} are shown for a patient undergoing heart surgery. From the time the operation commences, the patient is artificially ventilated, and the drop in p_{CO_2} with accompanying rise in pH after $1\frac{1}{4}$ hr is due to overventilation. The anaesthetist corrects this by altering the rate of ventilation and the composition of his gas mixture so that, at $2\frac{1}{2}$ hr, the acid–base balance or status of the patient is restored. The time then approaches when the patient must be attached to the heart-lung machine. The machine is primed with 10 dm^3 of blood from the transfusion unit. Fifteen minutes before the patient is attached, the analysis of the blood in the machine shows pH = 7.25, p_{CO_2} = 42 mmHg (= 5.60 kPa). In order to match the acid–base status of the patient's own blood as closely as possible, two things are done. First, p_{CO_2} in the gases used to oxygenate the blood in the machine will be reduced. As a result carbon dioxide dissolved in the blood passes out of solution and $[HCO_3^-]$ falls. To compensate for this, a calculated amount of sodium hydrogen carbonate is added to the blood.

Finally, if the circulation of the blood fails to maintain an adequate supply of oxygen to the tissues (during a coronary thrombosis, for example) incomplete oxidation of food in the muscles leads to an accumulation of lactic and pyruvic acids. These overcome the buffer reserve of the blood and produce a metabolic acidosis. To correct this, a solution of sodium hydrogen carbonate is transfused into a vein. The amount required is calculated, after blood pH and p_{CO_2} measurements, from the Henderson–Hasselbach equation or, in practice, from a suitable nomogram.

Bibliography

Further reading

Bradley, W. (1958) *Recent progress in the chemistry of dyes and pigments.* Royal Institute of Chemistry Lectures, Monographs, and Reports, No. 5.
Hallas, G. (1965) *Organic stereochemistry.* McGraw-Hill.
Inman, E. R. (1967) *Organic pigments.* Royal Institute of Chemistry Lecture Series, No. 1.
Jaffe, H. H. and Orchin, M. (1965) *Symmetry in chemistry.* Wiley.
Mislow, K. (1965) *Introduction to stereochemistry.* Benjamin.
Perutz, M. F. (1965) 'The anatomy of haemoglobin.' *Chemistry in Britain* **1**, 9–14.
Weeks, M. E. (1968) *Discovery of the elements* (7th edition). Chemical Education Publishing Co.

Abstracts of projects

arranged by D. R. Browning

Preparations

2.1 Extraction of copper from an ore

The aim of the project was to design and construct a system for extracting all the copper from a particular ore by efficient and industrially practicable methods, and to test the efficiency of the methods adopted. The ore was leached with a suitable solvent, which was found to be sulphuric acid, the optimum concentration of which had to be determined. An attempt was made at designing a continuous leaching plant that could be constructed within the given time for the project, but a batch process had to be adopted. The copper was then recovered from the mother liquor by electrolysis. The efficiency of the extraction (94%) was found by relating the amount of copper produced to the copper content of the feed (found by atomic absorption spectrometry).

SR*: copper ore; sintered glass discs; atomic absorption spectrometer

H. Cohen, Merchant Taylors' School, Northwood, Middx.

---

* SR = special requirements.

2.2 Oxide layer formation on metals at high temperatures

The rate of formation and thickness of oxide film formed on iron, aluminium, and copper, heated in air to temperatures of up to 600 °C, was investigated. The thickness of the film was deduced from the gain in weight of the sample. (cf. Abstract 4.15.)

D. Milne, Gordonstoun School, Elgin, Moray, Scotland.

2.3 Halogenated phenols (preparation and antiseptic properties)

Some halogenated phenols were prepared by standard methods. Their antiseptic properties were investigated using a novel method. The products from the metabolic process of bacteria in a gelatin medium are such as to cause an increase in the conductivity of the gelatin solution. By measuring the conductivity of a standard mixture of bacteria culture in gelatin at 40 °C an indication of the rate of metabolism of the bacteria was given. This was found to be inhibited by phenols, and depended on the number of halogen atoms present and on their orientation. Many more phenols remain to be investigated and a comparison with standard methods of testing the activity of bactericidal agents needs to be done.

SR: conductivity cell and meter

J. Turner, Sir Walter St John's School, Battersea, London SW11.

2.4 Preparation of 3-amino-5-butylphenol from aniline

The project concerned the synthesis of a tri-substituted aromatic compound, chosen at random, from a primary amine. The end product was to be tested qualitatively and quantitatively but time did not permit an extensive study.

SR: aniline; Buchner apparatus; reflux apparatus

S. A. Sefton, Merchant Taylors' School, Northwood, Middx.

Analysis

2.5 Analysis scheme for metal ions

The scheme divides up the metal ions encountered at A-level into five groups depending on their reaction, or otherwise, with NaOH and NH_4OH: (1) those which give no precipitate; (2) those which give a white precipitate; (3) those which give a coloured precipitate; (4) those giving a precipitate dissolving in excess NH_4OH; (5) those giving a precipitate dissolving in excess NaOH.

C. J. Crockett, Bedford School, Bedford.

2.6 Analysis of iron (<2%) in ores

Estimation of small quantities of iron in minerals was carried out using (a) oxidation of Fe^{2+} to Fe^{3+} with MnO_4^-, $Cr_2O_7^{2-}$; (b) complex formation with EDTA, salicylic acid, and SCN^-; (c) potentiometric titration with Ce^{4+}; (d) reduction of Fe^{3+} to Fe^{2+} with $HgNO_3$.

S. J. Rattle, Watford Grammar School, Rickmansworth Road, Watford, Herts.

2.7 Complexometric titration of metals

EDTA and 1,2-diaminoethane complexes were used to estimate zinc and copper concentrations in zinc and copper ores.

D. Pijsarczuk, Cardinal Hinsley Grammar School, Tong, Bradford, Yorks.

2.8 Phenol-ferric (iron(III)) complexes

A general picture of reactions between ferric ions (iron(III) ions) and various phenols was first attempted, followed by a more detailed study of the simplest, that of phenol itself, when the results obtained did not fit in with accepted theories. A 'sandwich' structure of the coloured ion was postulated, similar to the structure of the ferrocene type complexes. The reactions were followed using a photoelectric colorimeter. Much work remains to be done on other complexes.

SR: photoelectric colorimeter

D. S. Wethey, Charterhouse, Godalming, Surrey.

2.9 Complex nickel compounds

Some complex reactions of nickel were followed by continuous variation methods to determine the stoichiometry. A 'Vuespec' spectrophotometer was calibrated and then used to obtain absorption spectra of certain nickel complexes.

SR: MLI spectrophotometer (without pen recorder)

J. H. Hackforth-Jones, Queen Elizabeth's Grammar School for Girls, Barnet, Herts.

2.10 Formula of the mercurous ion (mercury(I) ion)

The equilibrium between mercury ions and mercury was investigated by titrimetric methods and by measuring the electrode potential of a platinum electrode dipped into the mixed ions at equilibrium. Reasonable agreement was obtained for the equilibrium constant based on the formula Hg_2^{2+} for the mercurous ion, (mercury(I) ion).

R. G. Cox, Charterhouse, Godalming, Surrey.

2.11 Experiments with a photocolorimeter

Two experiments: (1) Induced redox reactions using potassium dichromate, manganese(II) sulphate, phosphoric acid, and arsenic(III) oxide. Overall induced reaction As(III) → As(V). Induction factor found. (2) Iron(III) sulphosalicylic acid complex formed a red colour in acid solution. Plot of absorbance and hence the moles Fe^{3+}/SSA ratio. Point of maximum absorbance was found (by continuous variation) and hence the formula of the complex was deduced.

SR: photocolorimeter; sodium diphenylamine sulphate

H. Jenkinson, Merchant Taylors' School, Northwood, Middx.

2.12 Construction of an automatic burette

A device was designed and constructed which opened and closed the tap on a burette depending upon the light transmitted through a solution. This version was not sufficiently accurate to perform more than very approximate photocolorimetric titrations, but the principle worked.

P. J. Halden, Bryanston School, Blandford, Dorset.

2.13 Investigation into the alcohol content of wines and a comparison of yeasts

Various fruits and vegetables were used as a basis for making wine using a dried yeast. The alcohol contents of these were measured at stages during fermentation. This was then repeated with a yeast formula.

SR: dried yeast and yeast formula

T. Morris, Canon Slade Grammar School, Bolton, Lancs.

Structure and reactivity

2.14 Investigation into whether solid PCl_5 is ionic

The method chosen for investigating the ionic character of solid PCl_5 was by dielectric constants and dipole moments. Initially a capacitor was built of a very unreactive material to withstand high temperatures. After several attempts a capacitor was made of copper plates in a glass tank. This was then tested with various liquids for accuracy and found to be quite accurate with non-polar liquids but not with polar liquids. The capacitor was modified.

J. C. Robinson, Charterhouse, Godalming, Surrey.

2.15 Investigation into the methods of converting one isomer of tartaric acid into another

(1) Physical study of crystalline structure of the four isomers. (2) Methods of resolving a racemic mixture, (*i*) mechanical separation of crystalline solid, (*ii*) destruction, by bacterial action, of one enantiomorph, at a faster rate than the other, (*iii*) resolution using an optically active base, (*iv*) selective crystallization from a super-saturated solution. (3) (*i*) The manufacture of meso and racemic forms by heating one of the enantiomorphs at 170 °C for six hours, (*ii*) as for (*i*) with water added, (*iii*) as for (*i*) with concentrated sodium hydroxide added.

SR: all four isomers of tartaric acid; d-cinchonine; polarimeter

W. A. G. Reeve, Merchant Taylors' School, Northwood, Middx.

2.16 Phenolic compounds: evidence of hydrogen bonding

The infra-red spectra of a number of substituted phenols were measured and investigated for evidence of hydrogen bonding.

SR: infra-red spectrometer

G. A. Hibbert, Charterhouse, Godalming, Surrey.

2.17 Semiconduction in organic liquids

The intrinsic conduction of samples of benzimidazole was studied. Data for the conductance against temperature was shown to be logarithmic and the 'ohmic' voltage against current relationship demonstrated below 2 volts. The same observations on a sample of the hydrochloride salts supported the thesis that protonation of the nitrogen atom narrowed the energy gap.

SR: d.c. amplifier; potentiometer; special conductivity cell in oven

R. J. Merritt, Royal Grammar School, Guildford, Surrey.

2.18 Para and diamagnetism in the fourth period elements

Published results in this field show many discrepancies and disagreements with theory. A Stenzl electromagnet was calibrated using search coils and ballistic galvanometer. Three methods were tried for measuring the force on specimens: angular deflection on silk thread, vertical deflection on a glass spring, and an adapted torsion balance. Qualitative results were obtained for zinc, copper, nickel, cobalt, magnesium, and chromium but the apparatus was not sensitive enough for quantitative measurements.

SR: electromagnet giving field of order 1–2 kilogauss; pure specimens, supplied by Koch-Light Supply Laboratories

H. Lupton, Dartington Hall School, Totnes, South Devon.

2.19 Investigation of litmus

An investigation of some of the chemical properties of litmus solution by photometric titration, ion exchange, and spectrophotometric experiments. Use of a spectrophotometer indicated molecular structures which are very likely to be present in litmus; thus suggesting the composition of various lichen compounds from which litmus is derived.

SR: litmus powder, spectrophotometer

C. Wronska, St Joan of Arc Convent Grammar School, Rickmansworth, Herts.

2.20 Investigation of phenyl diazonium chloride

Beginning with elementary work on the diazotization reaction, the project was directed towards investigating the azo dyes derived from phenyl diazonium chloride. In particular, the effect of structure on the colour of these dyes was studied. By coupling various simple aromatic compounds with phenyl diazonium chloride solution in the cold and noting the colour of the resultant dye, it was possible to determine the effect of structure on colour. The results showed that colour got darker as molecular weight increased and depended on the presence or absence of various groups. There was also some investigation of the solubility of azo dyes.

N. M. Henderson, King Edward VI School, Southampton, Hants.

2.21 A comparative study of reactions

A comparison was attempted between certain reactions in liquid dinitrogen tetroxide or liquid ammonia and the corresponding reactions in water. The gases were cooled in crushed ice/salt or 'dry ice'/chloroform and liquids stored in test-tubes in the coolants. In liquid dinitrogen tetroxide, silver nitrate and nitrosyl chloride underwent the reaction $NO^+Cl^- + Ag^+NO_3^- \rightarrow AgCl + N_2O_4$ (acid + base → salt + solvent) and metallic sodium underwent the reaction $Na + N_2O_4 \rightarrow NaNO_3 + NO$. Thus acid–base reactions may involve nitrosyl ion transfer $N_2O_4 \rightarrow NO^+ + NO_3^-$. In liquid ammonia, solutions of sodium and calcium metals were made, the former turning colourless with ammonium chloride, an acid in this solvent: $2NH_3 \rightarrow NH_4^+ + NH_2^-$, $Na + NH_4^+ \rightarrow Na^+ + NH_3 + \frac{1}{2}H_2$. The reactions (*i*) $2AgCl + Ba(NO_3)_2 \rightarrow 2AgNO_3 + BaCl_2$ and (*ii*) $K^+NH_2^- + NH_4^+Cl^- \rightarrow KCl + 2NH_3$ were attempted with some success.

SR: ammonia and nitrogen dioxide cylinders; PTFE bungs and tubing; PTFE tape; 'dry ice'

Suggested further work: the study of chemical reactions in non aqueous inorganic and organic solvents is worthy of consideration and could involve an able student in original research.

S. G. Whitaker, King Edward VI School, Southampton, Hants.