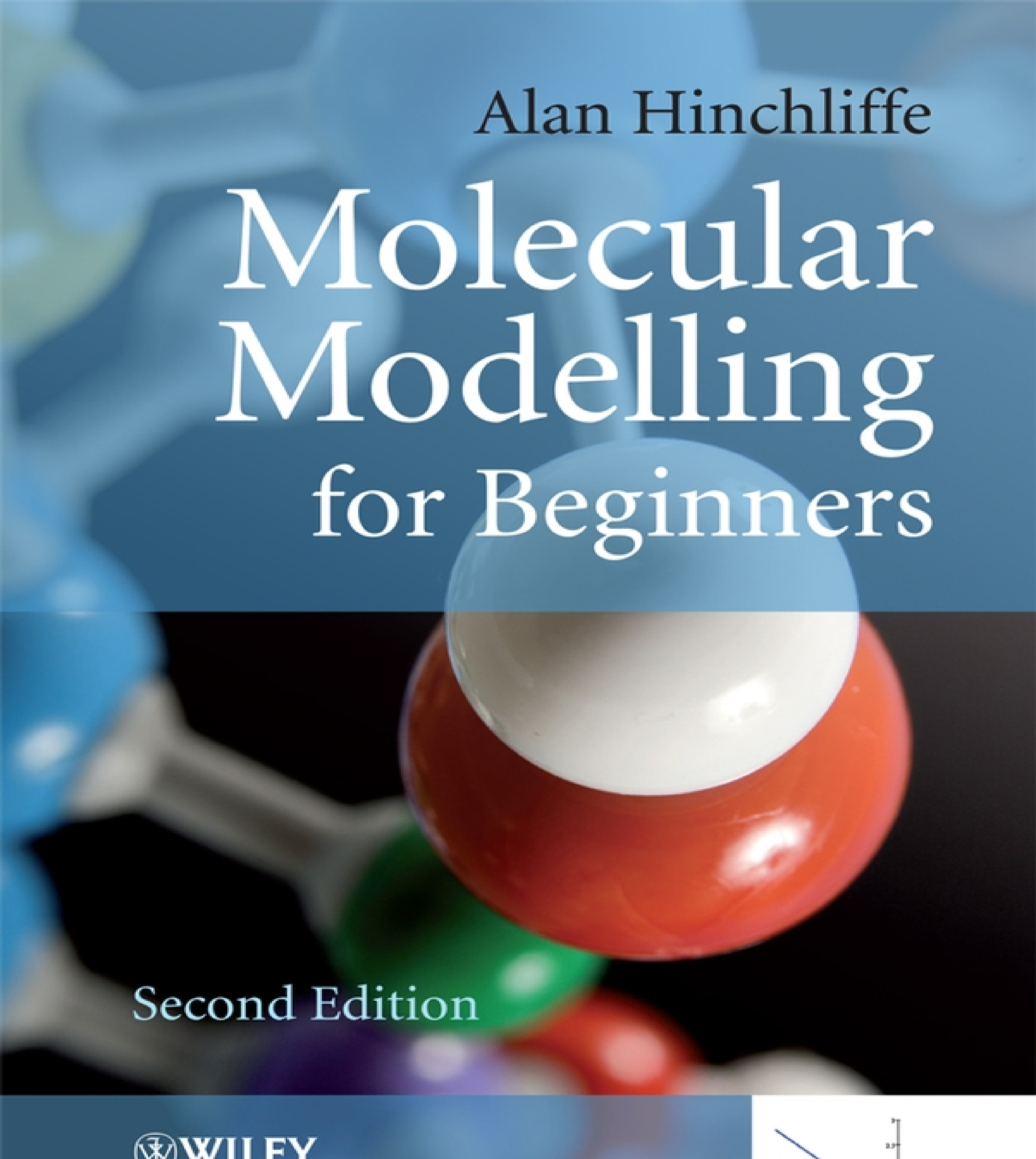
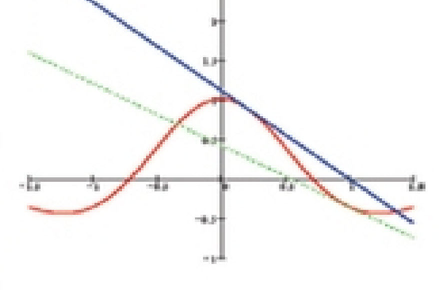


Alan Hinchliffe

Molecular Modelling for Beginners

Second Edition

WILEY



Contents

17 HF-LCAO Examples

18 Semiempirical Models

Molecular Modelling for Beginners

Second Edition

ALAN HINCHLIFFE
The University of Manchester

WILEY
A John Wiley and Sons, Ltd, Publication

This edition first published 2008

Registered office

John Wiley & Sons Ltd, The Atrium, Southern Gate,
Chichester, West Sussex,

PO19 8SQ, United Kingdom

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

other expert assistance is required, the services of a competent professional should be sought.

The publisher and the author make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of fitness for a particular purpose. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for every situation. In view of ongoing research, equipment modifications, changes in governmental regulations, and the constant flow of information relating to the use of experimental reagents, equipment, and devices, the reader is urged to review and evaluate the information provided in the package insert or instructions for each chemical, piece of equipment, reagent, or device for, among other things, any changes in the instructions or indication of usage and for added warnings and precautions. The fact that an organization or Website is referred to in this work as a citation and/or a potential source of further information does not mean that the author or the publisher endorses the information the organization or Website may provide or recommendations it may make. Further, readers should be aware that Internet Websites listed in this work may have changed or disappeared between when this work was written and when it is read. No warranty may be created or extended by any promotional statements for this work. Neither the publisher nor the author shall be liable for any damages arising herefrom.

Library of Congress Cataloging in Publication Data

Hinchliffe, Alan.

Molecular modelling for beginners / Alan Hinchliffe. — 2nd ed.

p. cm.

Includes bibliographical references and index.

ISBN 978-0-470-51313-2 (cloth) — ISBN 978-0-470-51314-9
(pbk. : alk. paper)

1. Molecules—Mathematical models. 2. Molecules—
Computer simulation.
I. Title.

QD480.H58 2008

541′.22015118—dc22

2008024099

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British
Library

CLOTH: 978-0-470-51313-2
PAPER: 978-0-470-51314-9

Preface to the Second Edition

It is five years since the first edition was published, and many things have moved on sufficiently to justify this second edition.

Some things never change; I have left the elementary chapters alone and I still believe that Appendix A on relevant mathematical methods is the correct place for you to *start* your studies.

Some topics have matured in the last five years. Density functional theory (and especially the B3LYP choice of functionals) has become the workhorse of modern computational chemistry. I have reworked all the problems and expanded the text as appropriate.

I have also said 'goodbye' to a few of the older topics. For example, everyone can now do chemical drawing, so I do not need to teach it. Thankfully that bane of our lives the Z-matrix has all but disappeared; I still have fond memories of struggling to get cyclic structures symmetrical and so it still gets a page of discussion.

I have completely rewritten the chapters dealing with Monte Carlo and molecular dynamics, the G*n* models, transition states and solvent models. I have also added a completely new chapter called 'Sharing Out the Energy', and I hope you will enjoy reading it.

It is fashionable to have an associated website with any new teaching text, and I have therefore added a website at

http://www.wileyeurope.com/college/hinchliffe

where you will find a number of problem sets and their solutions. Feel free to use them any way you like. I used them in my own teaching. Perhaps you have a corresponding set that you would like to share with the rest of us? Let me know.

I did all the illustrative calculations using either Gaussian 03 or HyperChem; these were done either on a beautiful

Sony Vaio laptop or on the University of Manchester's High Performance Computing parallel computer, a Bull Itanium2 system.

As always, I welcome comments and can be reached at: Alan.Hinchliffe@manchester. ac.uk.

Alan Hinchliffe
Manchester, UK

Preface to the First Edition

There is nothing radically new about the techniques we use in modern molecular modelling. Classical mechanics hasn't changed since the time of Newton, Hamilton and Lagrange, the great ideas of statistical mechanics and thermodynamics were discovered by Ludwig Boltzmann and J. Willard Gibbs amongst others and the basic concepts of quantum mechanics appeared in the 1920s, by which time J.C. Maxwell's famous electromagnetic equations had long since been published.

The chemically inspired idea that molecules can profitably be treated as a collection of balls joined together with springs can be traced back to the work of D.H. Andrews in 1930. The first serious molecular Monte Carlo simulation appeared in 1953, closely followed by B.J. Alder and T.E. Wainwright's classic molecular dynamics study of hard discs in 1957.

The Hartrees' 1927 work on atomic structure is the concrete foundation of our everyday concept of atomic orbitals, whilst C.C.J. Roothaan's 1951 formulation of the HF-LCAO model arguably gave us the basis for much of modern molecular quantum theory.

If we move on a little, most of my colleagues would agree that the two recent major advances in molecular quantum theory have been density functional theory, and the elegant treatment of solvents using ONIOM. Ancient civilizations believed in the cyclic nature of time and they might have had a point for, as usual, nothing is new. Workers in solid-state physics and biology actually proposed these models many years ago. It took the chemists a while to catch up.

Scientists and engineers first got their hands on computers in the late 1960s. We have passed the point on the computer history curve where every ten years gave us an order of magnitude increase in computer power, but it is no coincidence the that growth in our understanding and application of molecular modelling has run in parallel with growth in computer power. Perhaps the two greatest driving forces in recent years have been the PC and the graphical user interface. I am humbled by the fact that my lowly 1.2 GHz AMD Athlon office PC is far more powerful than the world-beating mainframes that I used as a graduate student all those years ago, and that I can build a molecule on screen and run a B3LYP/6-311++G(3d,2p) calculation before my eyes (of which more in Chapter 20).

We have also reached a stage where tremendously powerful molecular modelling computer packages are commercially available, and the subject is routinely taught as part of undergraduate science degrees. I have made use of several such packages to produce the screenshots; obviously they look better in colour than the greyscale of this text.

There are a number of classic (and hard) texts in the field; if I'm stuck with a basic molecular quantum mechanics problem, I usually reach for Eyring, Walter and Kimball's *Quantum Chemistry* but the going is rarely easy.

Equally there are a number of beautifully produced elementary texts and software reference manuals that can apparently transform you into an expert overnight. It's a two-edged sword, and we are victims of our own success. One often meets self-appointed experts in the field who have picked up much of the jargon with little of the deep understanding. It's no use (in my humble opinion) trying to hold a conversation about gradients, hessians and density functional theory with a colleague who has just run a molecule through one package or another but hasn't the slightest clue what the phrases or the output mean.

It therefore seemed to me (and to the reviewers who read my new book proposal) that the time was right for a middle course. I assume that you are a 'Beginner' in the sense of

Chambers dictionary, *someone who begins; a person who is in the early stages of learning or doing anything*…, and I want to tell you how we go about modern molecular modelling, why we do it, and most important of all, explain much of the basic theory behind the mouse clicks. This involves mathematics and physics, and the book neither pulls punches nor aims at instant enlightenment. Many of the concepts and ideas are difficult ones, and you will have to think long and hard about them; if it’s any consolation, so did the pioneers in our subject. I have given many of the derivations in full, and tried to avoid the dreaded phrase ‘it can be shown that’.

There are various strands to our studies, all of which eventually intertwine. We start off with molecular mechanics, a classical treatment widely used to predict molecular geometries. In Chapter 8, I give a quick guide to statistical thermodynamics (if such a thing is possible), because we need to make use of the concepts when trying to model arrays of particles at nonzero temperatures. Armed with this knowledge, we are ready for an assault on Monte Carlo and molecular dynamics.

Just as we have to bite the bullet of statistical mechanics, so we have to bite the equally difficult one of quantum mechanics, which occupies Chapters 11 and 12. We then turn to the quantum treatment of atoms, where many of the sums can be done on a postcard if armed with knowledge of angular momentum.

The Hartree–Fock and HF-LCAO models dominate much of the next few chapters, as they should. The Hartree–Fock model is great for predicting many molecular properties, but it can’t usually cope with bond breaking and bond making. Chapter 19 treats electron correlation and Chapter 20 deals with the very topical density functional theory (DFT). You won’t be taken seriously if you have not done a DFT calculation on your molecule. Quantum mechanics, statistical mechanics and electromagnetism all have a certain well-deserved reputation amongst science students; they are hard subjects. Unfortunately all three all feature in this new text. In electromagnetism it is mostly a matter of getting to grips with the mathematical notation (although I have spared you Maxwell's beautiful equations), whilst in the other two subjects it is more a question of mastering hard concepts. In the case of quantum mechanics, the concepts are often in direct contradiction to everyday experience and common sense. I expect from you a certain level of mathematical competence; I have made extensive use of vectors and matrices not because I am perverse, but because such mathematical notation brings out the inherent simplicity and beauty of many of the equations. I have tried to help by giving a mathematical appendix, which should also make the text self-contained.

I have tried to put the text into historical perspective, and in particular I have quoted directly from a number of what I call *keynote papers*. It is interesting to read at first hand how the pioneers put their ideas across, and in any case they do it far better than me. For example, I am not the only author to quote Paul Dirac's famous statement

> The underlying Physical Laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that exact application of these laws leads to equations much too complicated to be soluble.

I hope you have a profitable time in your studies, and at the very least begin to appreciate what all those options mean next time you run a modelling package!

Alan Hinchliffe
Manchester, UK

1

Electric Charges and their Properties

As far as we can tell, there are four fundamental types of interactions between physical objects. There is the *weak nuclear interaction* that governs the decay of beta particles, and the *strong nuclear interaction* that is responsible for binding together the particles in a nucleus. The familiar *gravitational* interaction holds the Earth very firmly in its orbit round the Sun, and finally we know that there is an *electromagnetic* interaction that is responsible for binding atomic electrons to nuclei and for holding atoms together when they combine to form molecules.

Of the four, the gravitational interaction is the only one we would normally come across in our everyday world. This is because gravitational interactions between bodies always add. The gravitational interaction between two atoms is negligible but when large numbers of fundamental particles such as atoms are aggregated together, the gravitational interaction becomes significant.

You may think it bizarre that there are four types of interaction, yet, conversely, you might wonder why there should be just four. Why not one, three or five? Should there not be a unifying theory to explain why there are four, and whether they are related? As I write, there is no such unifying theory despite tremendous research activity.

1.1 Point Charges

In this chapter I am going to concentrate on electric charges and their properties, since electrons and protons are fundamental building blocks for atoms and molecules.

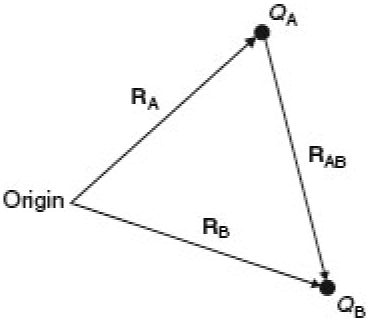
It turns out that there are two types of electric charge in nature, which we might choose to call type X and type Y (or Red and Blue for that matter, but X and Y will do for now). Experimental evidence shows the existence of an electrostatic force between electric charges; the force between two X-type charges is always repulsive, as is the force between two Y-type charges. The force between an X-type and a Y-type is always attractive. For this reason, the early experimenters decided to classify charges as positive or negative, because a positive quantity times a positive quantity gives a positive quantity, a negative quantity times a negative quantity gives a positive quantity whilst a negative quantity times a positive quantity gives a negative quantity. I am sure you know that the best known fundamental particles responsible for these charges are electrons and protons, and you are probably expecting me to tell you that the electrons are the negatively charged particles whilst protons are positively charged. It is actually just a convention that we take: we could just as well have called electrons positive.

Whilst on the subject, it is fascinating to note that the charge on the electron is exactly equal and opposite of that on a proton. Atoms and molecules generally contain exactly the same number of electrons and protons, and so the net charge on a molecule is almost always zero. Ions certainly exist in solutions of electrolytes, but the number of Na^+ ions in a solution of sodium chloride is exactly equal to the number of Cl^- ions and once again we are rarely aware of any imbalance of charge.

A thunderstorm results when nature separates out positive and negative charges on a macroscopic scale. It is thought that friction between moving masses of air and water vapour detaches electrons from some molecules and attaches them to others. This results in parts of clouds being left with an excess of charge, often with spectacular results. It was investigations into such atmospheric phenomena that gave the first clues about the nature of the electrostatic force.

We normally start any study of charges at rest (*electrostatics*) by considering the force between two point charges, as shown in Figure 1.1. The term 'point charge'is a mathematical abstraction; obviously electrons and protons have a finite size. Just bear with me for a few pages, and accept that a point charge is one whose dimensions are small compared to the distance between them. An electron is large if you happen to be a nearby electron, but can normally be treated as a point charge if you happen to be a human being a metre away.

Figure 1.1 *Point charges*



The concept of a point charge may strike you as an odd one, but once we have established the magnitude of the force between two such charges, we can deduce the force between any arbitrary charge distributions on the grounds that they are composed of a large number of point charges.

In Figure 1.1 we have point charge Q_A at position vector $\mathbf{R}_A$ and Q_B at $\mathbf{R}_B$. From the laws of vector analysis, the vector $\mathbf{R}_{AB} = \mathbf{R}_B - \mathbf{R}_A$ joins Q_A to Q_B, and points from Q_A to Q_B as shown. I have indicated the direction of the vectors with arrows.

1.2 Coulomb’s Law

In 1785, Charles Augustin de Coulomb became the first person to give a mathematical form to the force between point charges. He measured the force directly between two very small charged bodies, and was able to show that the force exerted by Q_A on Q_B was

- proportional to the inverse square of the distance between Q_A and Q_B when both charges were fixed;
- proportional to Q_A when Q_B and $\mathbf{R}_{AB}$ were fixed;
- proportional to Q_B when Q_A and $\mathbf{R}_{AB}$ were fixed.

He also noticed that the force acted along the line joining the centres of the two charges, and that the force was either attractive or repulsive depending on whether the charges were different or of the same type. The sign of the product of the charges therefore determines the direction of the force.

A mathematical result of these observations can be written in scalar form as

(1.1)

$$F_{\mathrm{A\ on\ B}} \propto \frac{Q_{\mathrm{A}} Q_{\mathrm{B}}}{R_{\mathrm{AB}}^2}$$

Forces are vector quantities, and Equation (1.1) is better written in vector form as

$$\mathbf{F}_{\mathrm{A\ on\ B}} \propto \frac{Q_{\mathrm{A}} Q_{\mathrm{B}}}{R_{\mathrm{AB}}^3} \mathbf{R}_{\mathrm{AB}}$$

When Coulomb first established his law, he had no means of quantifying charge and so could not identify the proportionality constant. He took it to be unity, and thereby defined charge in terms of the force between charges. Modern practice is to regard charge and force as independent quantities, and because of this a dimensioned proportionality constant is necessary. For a reason that need not concern us, this is taken as $1/4\pi\varepsilon_0$, where the permittivity of free space ε_0 is an experimentally determined quantity with the approximate value $\varepsilon_0 = 8.854 \times 10^{-12}\ \mathrm{C}^2\ \mathrm{N}^{-1}\ \mathrm{m}^{-2}$. Coulomb’s law is therefore

(1.2)

$$\mathbf{F}_{\mathrm{A\ on\ B}} = \frac{1}{4\pi\varepsilon_0}\frac{Q_{\mathrm{A}}Q_{\mathrm{B}}}{R_{\mathrm{AB}}^3}\mathbf{R}_{\mathrm{AB}}$$

and it applies to measurements done in free space. If we repeat Coulomb’s experiments with the charges immersed in different media, then we find that the law still holds but with a different proportionality constant. We modify the proportionality constant using a quantity ε_r called the *relative permittivity*. In older texts, ε_r is called the *dielectric constant*. Our final statement of Coulomb’s law is therefore

(1.3)

$$\mathbf{F}_{\mathrm{A\ on\ B}} = \frac{1}{4\pi\varepsilon_r\varepsilon_0}\frac{Q_{\mathrm{A}}Q_{\mathrm{B}}}{R_{\mathrm{AB}}^3}\mathbf{R}_{\mathrm{AB}}$$

According to Newton’s third law, we know that if Q_A exerts a force $\mathbf{F}_{\mathrm{A\ on\ B}}$ on Q_B, then Q_B should exert an equal and opposite force on Q_A. Coulomb’s law satisfies this requirement, since

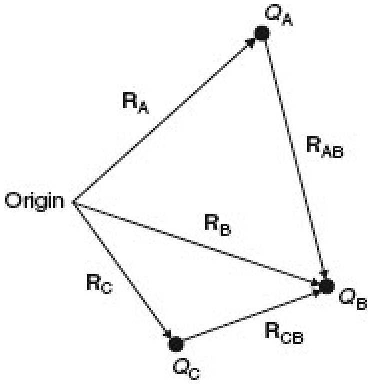
$$\mathbf{F}_{\mathrm{B\ on\ A}} = \frac{1}{4\pi\varepsilon_r\varepsilon_0}\frac{Q_{\mathrm{A}}Q_{\mathrm{B}}}{R_{\mathrm{BA}}^3}\mathbf{R}_{\mathrm{BA}}$$

(the vector $\mathbf{R}_{\mathrm{BA}}$ points in the opposite direction to $\mathbf{R}_{\mathrm{AB}}$ and so one force is exactly the negative of the other, as it should be).

1.3 Pair Wise Additivity

Suppose we now add a third point charge Q_C with position vector $\mathbf{R}_C$ as shown in Figure 1.2. Since Q_A and Q_B are point charges, the addition of Q_C cannot alter the force between Q_A and Q_B.

Figure 1.2 *Third charge added*



The total force on Q_B now comprises two terms, the force due to point charge Q_A and the force due to point charge Q_C. This total force is given by

(1.4)

$$\mathbf{F}_B = \frac{Q_B}{4\pi\varepsilon_0}\left(Q_A\frac{\mathbf{R}_{AB}}{R_{AB}^3} + Q_C\frac{\mathbf{R}_{CB}}{R_{CB}^3}\right)$$

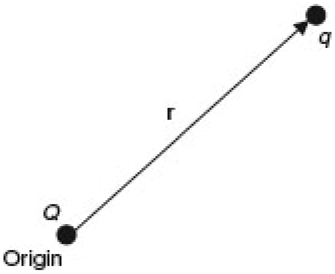
This may seem at first sight to be a trivial statement: surely all forces act this way. Not necessarily, for I have assumed that the addition of Q_C did not have any effect on Q_A and Q_B (and so did not influence the force between them).

The generic term *pair wise additive* describes things like forces that add as above. Forces between point electric charges are certainly pair wise additive, and so you might imagine that forces between atoms and molecules must therefore be pair wise additive, because atoms and molecules consist of (essentially) point charges. I am afraid that nature is not so kind, and we will shortly meet situations where forces between the composites of electrons and protons that go to make up atoms and molecules are far from being pair wise additive.

1.4 Electric Field

Suppose now we have a point charge *Q* at the coordinate origin, and we place another point charge *q* at point P that has position vector **r** (Figure 1.3).

Figure 1.3 *Field concept*



The force exerted by *Q* on *q* is

$$\mathbf{F} = \frac{1}{4\pi\varepsilon_0}\frac{Qq}{r^3}\mathbf{r}$$

which I can rewrite trivially as

$$\mathbf{F} = \left(\frac{1}{4\pi\varepsilon_0}\frac{Q}{r^3}\mathbf{r}\right)q$$

The point is that the term in brackets is to do with *Q* and the vector **r**, and contains no mention of *q*. If we want to find the force on any arbitrary *q* at **r**, we calculate the quantity in brackets once and then multiply by *q*. One way of thinking about this is to imagine that the charge *Q* creates a certain field at point **r**, which determines the force on any other *q* when placed at position **r**.

This property is called the *electric field* **E** at that point. It is a vector quantity, like force, and the relationship is

$$\mathbf{F}\,(\text{on } q \text{ at } \mathbf{r}) = q\mathbf{E}\,(\text{at } \mathbf{r})$$

Comparison with Coulomb’s law, Equation (1.3), shows that the electric field at point **r** due to a point charge *Q* at the coordinate origin is

(1.5)

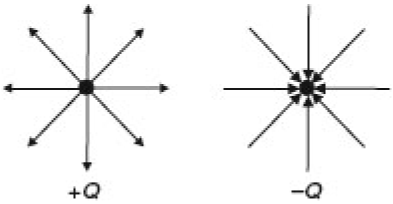
$$\mathbf{E} = \frac{1}{4\pi\varepsilon_0}\frac{Q\mathbf{r}}{r^3}$$

E is sometimes written **E**(**r**) to emphasize that the electric field depends on the position vector **r**.

Electric fields are vector fields and they are often visualized as *field lines*. These are drawn such that their spacing is inversely proportional to the strength of the field, and their tangent is in the direction of the field. They start at positive charges and end at negative charges, and two simple examples are shown in Figure 1.4. Here the choice of eight lines is quite arbitrary.

Electric fields that do not vary with time are called *electrostatic* fields.

Figure 1.4 *Field lines for point charges*



1.5 Work

Look again at Figure 1.3, and suppose we move point charge *q* whilst keeping *Q* fixed in position. When a force acts to make something move, energy is transferred. There is a useful phrase in physical science that is to do with the energy transferred, and it is *work*. Work measures the energy transferred in any change, and can be calculated from the change in energy of a body when it moves through a distance under the influence of a force.

We have to be careful to take account of the energy balance. If a body gains energy, this energy has to come from somewhere, and that somewhere must lose energy. What we do is to divide the universe into two parts: the bits we are interested in called the *system* and the rest of the universe that we call the *surroundings*.

Some texts focus on the work done *by* the system, some concern themselves with the work done *on* the system. According to the law of conservation of energy, one is